TÉCNICO LISBOA

UNIVERSIDADE DE LISBOA

INSTITUTO SUPERIOR TÉCNICO

Biomanufacturing of Non-Viral Protein Nanocages

Jorge Fernandes da Silva João

Supervisor: Doctor Duarte Miguel de França Teixeira dos Prazeres

Thesis approved in public session to obtain the PhD Degree in Biotechnology and Biosciences

Jury Final Classification: Pass with Distinction

2025

UNIVERSIDADE DE LISBOA

INSTITUTO SUPERIOR TÉCNICO

Biomanufacturing of Non-Viral Protein Nanocages

Jorge Fernandes da Silva João

Supervisor: Doctor Duarte Miguel de França Teixeira dos Prazeres

Thesis approved in public session to obtain the PhD Degree in Biotechnology and Biosciences

Jury Final Classification: Pass with Distinction

Jury

Chairperson: Doctor Arsénio do Carmo Sales Mendes Fialho, Instituto Superior Técnico, Universidade de Lisboa

Members of the Committee:

Doctor Ana Cecília Afonso Roque, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa

Doctor Duarte Miguel de França Teixeira dos Prazeres, Instituto Superior Técnico, Universidade de Lisboa

Doctor Gabriel António Amaro Monteiro, Instituto Superior Técnico, Universidade de Lisboa

Doctor Ângela Maria Almeida de Sousa, Centro de Investigação em Ciências da Saúde, Universidade da Beira Interior

Doctor Ana Gabriela Gonçalves Neves Gomes, Escola Superior de Tecnologia do Barreiro, Instituto Politécnico de Setúbal

Funding Institution

FCT: Fundação para a Ciência e a Tecnologia

2025

Abstract

Non-viral protein nanocages (NVPNs) are highly ordered, nanometer scale architectures, which are typically formed by homo- or hetero-self-assembly of multiple monomers into symmetric structures with different dimensions and morphologies. The intrinsic characteristics of protein nanocages make them very attractive as a biological nanomaterial, and include a high surface/volume ratio, multi-functionality, ease to modify or manipulate genetically or chemically, high stability, monodispersibility, and biocompatibility. Several applications have been implemented in a wide range of scientific areas from biotechnology to biomedicine, showing that these nanocages can be a promising and interesting tool. The development of such applications for NVPNs requires large amounts of pure and well-folded assemblies. Consequently, the availability of more effective biomanufacturing processes is needed to transform protein nanocages into valuable bioproducts. Moreover, these manufacturing processes should be supported by analytical techniques adequate for the quality control and characterization of the structure and biological function of nanocages.

The main objective of this work is the development of scalable and efficient processes for the biomanufacturing of NVPNs. Two models of this type of protein nanocages, one natural and the other artificial, were used throughout the work. Firstly, the production of both model nanocages in *Escherichia coli* cells was set up and used as a benchmark manufacturing method. Additionally, the same production was performed in the alternative bacterial host *Vibrio natriegens*, followed by an optimization process of several relevant conditions. Afterwards, a purification strategy was developed using the NVPNs produced both in *E. coli* and *V. natriegens*. The process was based on chromatographic approaches, which were optimized to obtain a highly purified product. The final structure of the protein nanocages synthesized in both bacteria was studied and characterized using a range of analytical techniques, namely microscopy and optical spectroscopy assays.

Key-words: analytical techniques, biomanufacturing, downstream processing, protein nanocages, upstream processing

Resumo

As *nanocages* proteicas não-virais (NVPNs) são arquiteturas altamente ordenadas à escala nanométrica, geralmente formadas pela auto-organização de vários monómeros iguais ou diferentes em estruturas simétricas com dimensões e morfologias variadas. As características intrínsecas das *nanocages* proteicas, que as tornam muito atrativas como nanomaterial biológico, incluem uma razão superfície/volume elevada, multifuncionalidade, facilidade de modificação ou manipulação genética ou química, elevada estabilidade, mono-dispersibilidade e biocompatibilidade. Várias aplicações em diversas áreas científicas têm sido estudadas, desde a biotecnologia à biomedicina, que demonstram que estas *nanocages* podem ser uma ferramenta promissora e interessante. O desenvolvimento deste tipo de aplicações de NVPNs implica a obtenção de elevadas quantidades de nanoestruturas puras e com a montagem correta. Torna-se assim necessário dispor de processos de bio-manufatura eficazes que permitam transformar as *nanocages* proteicas em bioprodutos reais. Para além disso, esses processos devem ser suportados por técnicas analíticas adequadas ao controlo de qualidade e à caracterização da estrutura e função biológica das *nanocages*.

O principal objetivo deste trabalho é desenvolver processos escaláveis e eficientes para a bio-manufatura de NVPNs. Dois modelos de *nanocages* proteicas, um natural e outro artificial, foram utilizados ao longo do trabalho. Primeiramente, foi implementada a produção de ambos os modelos de *nanocages* em células de *Escherichia coli*, tendo depois sido esta considerada como método de referência para a manufatura. Adicionalmente, a produção foi também estabelecida em células de *Vibrio natriegens*, seguida de um processo de otimização de diversas condições relevantes. Posteriormente, foi desenvolvida uma estratégia de purificação utilizando as NVPNs produzidas em *E. coli* e *V. natriegens*. O processo baseou-se em abordagens cromatográficas que foram otimizadas de modo a obter um produto altamente purificado. As estruturas finais das *nanocages* proteicas expressas nas duas bactérias foram analisadas e caracterizadas utilizando várias técnicas analíticas, nomeadamente ensaios de microscopia e de espectroscopia ótica.

Palavras-chave: técnicas analíticas, bio-manufatura, processamento a jusante, *nanocages* proteicas, processamento a montante

Acknowledgements

Firstly, I would like to thank my supervisor Professor Duarte Miguel F. Prazeres for all the support, help, guidance, and availability that he has shown throughout the years of my PhD. I can say that it allowed me to complete this stage of my life successfully.

I would like to thank Professor Isabel Sá-Correia as Coordinator of the BIOTECnico Doctoral Program for giving me the opportunity to take part in this Program and also as Coordinator of my PhD in Biotechnology and Biosciences at IST – ULisboa. I would also like to thank Professor Arsénio M. Fialho as the current person responsible for this PhD.

I would also like to thank Professor Joaquim Sampaio Cabral as Coordinator of the Institute for Bioengineering and Biosciences (iBB), my host institution, without whom it would not have been possible to perform all the experimental work associated with this thesis, as well as of the Institute for Health and Bioeconomy (i4HB). A thank you also to Professor Mário N. Berberan-Santos as the current Head of iBB.

A special thank you to all the collaborators involved in this thesis, who allowed me to complement and enrich my work, and also to learn a lot more on specific topics outside my comfort area. Many thanks to Professor Jonathan Heddle and his team (Karolina Majsterkiewicz and Yu-Chiuan Bau / Jagiellonian University), Professor Ana M. Azevedo (IST – ULisboa) including as member of the Thesis Advisory Committee, Professor Pedro M. Paulo (IST – ULisboa), and Professor Mário S. Rodrigues and his team (Tiago Robalo / Faculdade de Ciências – ULisboa). I would also like to thank Professor Gabriel A. Monteiro for all his help and availability as a member of the Thesis Advisory Committee.

One of my biggest thanks goes to all the current and former iBB laboratory colleagues with whom I have crossed paths throughout these years of my PhD, for their support, help, availability, and friendship. Many thanks to Ana Rita Santos, Cláudia Alves, Cristiana Ulpiano, Diogo Faria, José Santos, Leonor Resina, Miguel Nascimento, Renata Ferreira, Ricardo Pereira, Ricardo Silva, Sara Sousa Rosa, and Sofia Duarte, among many others. Clearly, without all of them it wouldn't have been the same and this long journey would have been even more difficult. I cannot fail to mention Rosa Gonçalves, our "D. Rosa", for everything that she has done for me since I joined iBB, for her availability, attention, and friendship. A very sincere and super special thank you.

A big thank you also to all my friends, some of whom are also still at IST, sharing the love for science. I also thank those I met through a common passion, photography. Here I would like to give special thanks to Marta Bica for her constant support in this PhD adventure that she also decided to embark on, may we continue our friendship for a long time and share our interest in science. A huge thank you also to my friend Cristiana Ulpiano with whom I shared these years at IST in all the good and bad moments, on this long and challenging journey that, although difficult and full of adversities, we will manage to complete successfully.

Finally, I would like to sincerely thank my entire family, also those who are unfortunately no longer with us, like some of my grandparents. A special word to my parents and my sister (and to Fidalgo), for their constant and unconditional understanding, support, patience, and inspiration during all these years. I hope that you are all happy and proud of this journey of mine.

I will therefore end by dedicating this thesis to each of the people that I mentioned here, thank you very much for everything!

List of Figures

List of Tables

List of Abbreviations

3D	Three-dimensional
AC	Affinity chromatography
AEX	Anion exchange
AFM	Atomic force microscopy
APS	Ammonium persulfate
Au-TPPMS	Chloro[diphenyl(3-sulfonatophenyl)phosphine]gold(i) sodium salt hydrate
BCA	Bicinchoninic acid
BHI	Brain heart infusion
BMCs	Bacterial microcompartments
bp	Base pair
BSA	Bovine serum albumin
CEX	Cation exchange
CFU	Colony forming units
CHO	Chinese hamster ovary
CIP	Cleaning-in-place
CV	Column volumes
DEAE	Diethylaminoethyl
DLS	Dynamic light scattering
DMSO	Dimethyl sulfoxide
DOE	Design of experiments
Dps	DNA-binding protein from starved cells
dsDNA	double-stranded DNA
DTT	Dithiothreitol
E2	Dihydrolipoyl acetyltransferase
EDTA	Ethylenediaminetetraacetic acid
EMA	European Medicines Agency
ESI-TOF	Electrospray ionization time-of-flight
EU	Endotoxin unit
FCS	Fluorescence correlation spectroscopy
FDA	Food and Drug Administration
FUV-CD	Far-UV circular dichroism

gDNA	Genomic DNA
GEMMA	Gas-phase electrophoretic mobility molecular analyzer
GVNPs	Gas vesicle protein nanoparticles
HBC	Hydrogen bond chromatography
HIC	Hydrophobic interaction chromatography
HP-SEC	High performance SEC
HPLC	High performance liquid chromatography
HSP	Heat shock protein
IEX	Ion exchange
IFN-γ	Interferon gamma
IMAC	Immobilized metal affinity chromatography
IPTG	Isopropyl β-D-1-thiogalactopyranoside
KDPG	2-keto-3-deoxy-6-phosphogluconate
LALS	Low-angle light scattering
LB	Luria-Bertani
LC/MS	Liquid chromatography/mass spectrometry
LPS	Lipopolysaccharides
MALDI-TOF	Matrix-assisted laser desorption/ionization time-of-flight
MALS	Multi-angle light scattering
MBP	Maltose-binding protein
MES	2-(N-morpholino)ethanesulfonic acid
MJV	Major vault protein
MnP	Manganese peroxidase
mRNA	Messenger RNA
NTA	Nitrilotriacetic acid
NVPNs	Non-viral protein nanocages
OD	Optical density
OIPNC	Oxygen-impermeable protein nanocages
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffered saline
PDB	Protein Data Bank
PDI	Polydispersity index
PDH	Pyruvate dehydrogenase

pDNA	Plasmid DNA
PEG	Polyethylene glycol
PepA	Aminopeptidase
PIPES	1,4-piperazinediethanesulfonic acid
PMBOs	Protein membrane-based organelles
PTMs	Post-translational modifications
qPCR	Quantitative PCR
RALS	Right-angle light scattering
SAXS	Small angle X-ray scattering
SDS-PAGE	Sodium dodecyl-sulfate PAGE
SEC	Size exclusion chromatography
sHSP	Small HSP
siRNA	Small interfering RNA
TCEP	Tris(2-carboxyethyl)phosphine hydrochloride
TEM	Transmission electron microscopy
TEMED	N,N,N′,N′-tetramethylethylenediamine
TRAP	trp RNA-binding attenuation protein
TTTR	Time-tagged time resolved
VLPs	Virus-like particles

Table of Contents

Thesis Outline

This thesis was developed within the scope of the PhD Program in Biotechnology and Biosciences (BIOTECnico) at Instituto Superior Técnico (IST) – Universidade de Lisboa (ULisboa) (Lisbon, Portugal). The scientific project was funded by Fundação para a Ciência e a Tecnologia (FCT) through the PhD fellowship whose reference is PD/BD/150335/2019. The experimental work was mainly performed at the Institute for Bioengineering and Biosciences (iBB) at IST – ULisboa, under the scientific supervision of Professor Duarte Miguel F. Prazeres. Additionally, collaborations with Professor Jonathan Heddle (Jagiellonian University / Krakow, Poland), Professor Ana M. Azevedo (IST – ULisboa), Professor Pedro M. Paulo (IST – ULisboa) and Professor Mário S. Rodrigues (Faculdade de Ciências – Universidade de Lisboa / Lisbon, Portugal) were also established.

The main research topic of this thesis was non-viral protein nanocages (NVPNs) and their biomanufacturing processes. The document is divided into six distinct chapters, which are briefly described below.

In **Chapter I**, a general introduction to protein nanocages is presented, with a particular focus on these non-viral nanostructures. An overview of the progress made since the emergence of protein nanocages as a novel class of biologics is provided. The most important applications of NVPNs in the biotechnological and biomedical areas are also presented. As the main topic, the current processes used for the production and purification of NVPNs are explored and discussed. Simultaneously, conventional characterization techniques and alternative or complementary approaches in development are assessed.

In **Chapter II**, a brief contextualization of the thesis is performed, and the main purpose and associated objectives of the entire work developed are also defined and described. Furthermore, the two NVPNs models used in this research project are presented and their different structural and functional characteristics are briefly explored.

Chapter III is the first dedicated to the research work carried out during this thesis together with **Chapter IV** and **Chapter V**. The initial steps for the preparation of expression vectors appropriate for the *in vivo* production of both nanocages using *Escherichia coli* and *Vibrio natriegens* expression systems are described. A study of the plasmid copy number in both species was also performed. The production of the two model NVPNs in *E. coli* cells was then set up and considered as a benchmark manufacturing method. Additionally, the same production was performed in *V. natriegens* cells, with the existence of some challenges in transforming these cells with the desired plasmids. Then, the production process was optimized by evaluating the impact of five different parameters considered more relevant according to literature.

Chapter IV is focused on the downstream processes studied for the two model NVPNs. Mainly, different strategies for purifying NVPNs from impurities such as host proteins and double stranded DNA were described for both nanocages generated in *E. coli* and *V. natriegens*. The implemented processes were based on chromatographic approaches, which were optimized to obtain a highly purified product.

Due to the alternative characteristics in terms of the assembly process, one of the model structures required an additional step of purification that was addressed.

In **Chapter V**, the conventional and alternative analytical techniques and methodologies used to study and characterize the biological structure of the NVPNs synthesized *in vivo* in both bacteria were described.

To finalize the thesis, **Chapter VI** presents the conclusions obtained from this work as well as future perspectives and challenges.

Chapter I

General Introduction

This chapter was published as:

João J, Prazeres DMF. Manufacturing of non-viral protein nanocages for biotechnological and biomedical applications. Front. Bioeng. Biotechnol. 2023;11:1200729.

Abstract

Protein nanocages are highly ordered nanometer scale architectures, which are typically formed by homo- or hetero-self-assembly of multiple monomers into symmetric structures of different size and shape. The intrinsic characteristics of protein nanocages make them very attractive and promising as a biological nanomaterial. These include, among others, a high surface/volume ratio, multi-functionality, ease to modify or manipulate genetically or chemically, high stability, mono-dispersity, and biocompatibility. Since the beginning of the investigation into protein nanocages, several applications were conceived in a variety of areas such as drug delivery, vaccine development, bioimaging, biomineralization, nanomaterial synthesis and biocatalysis. The ability to generate large amounts of pure and well-folded protein assemblies is one of the keys to transform nanocages into clinically valuable products and move biomedical applications forward. This calls for the development of more efficient biomanufacturing processes and for the setting up of analytical techniques adequate for the quality control and characterization of the biological function and structure of nanocages.

This **Chapter I** concisely covers and overviews the progress made since the emergence of protein nanocages as a new, next-generation class of biologics. A brief outline of non-viral protein nanocages is followed by a presentation of their main applications in the areas of bioengineering, biotechnology, and biomedicine. Afterwards, the current processes used in the manufacturing of protein nanocages are described, with particular emphasis on the most relevant aspects of production and purification. The state-of-the-art on current characterization techniques is then described and future alternative or complementary approaches in development are also discussed. Finally, a critical analysis of the limitations and drawbacks of the current manufacturing strategies is presented, alongside with the identification of the major challenges and bottlenecks.

Keywords: biomanufacturing, bottom-up synthesis, downstream processing, drug delivery, nanostructure engineering, protein nanocages, self-assembly, upstream processing

I.1. Introduction

In recent years, nanoparticles have been explored for applications in several scientific areas from nanobiotechnology and biomedical sciences to materials science and synthetic biology. Examples of nanoparticles studied in the literature include protein-based nanoparticles, metal nanoparticles, polymer micelles, silica nanoparticles and quantum dots (1). Among these, protein-based nanoparticles spurred significant research interest given their enormous potential for biomedical purposes (1–3).

In Nature, the existence of cellular processes essential to life such as metabolic reactions and information exchange is dependent on biological compartmentalization. In addition to lipids, proteins are one of the main components of natural compartmentalization systems such as virus capsids. Nanoparticles based on functional proteins constitute an additional example of such bio-compartments. Several of these protein-based nano-compartments, with different structural and functional characteristics, are described in the literature, including toroid- and donut-shaped proteins, tubes and yoctowells, protein nanocages, bacterial microcompartments (BMCs), protein membrane-based organelles (PMBOs) and gas vesicle protein nanoparticles (GVNPs) (3–5). For example, GVNPs, which have membranes exclusively composed of proteins, are produced in a wide variety of prokaryotic microorganisms, from heterotrophic bacteria to halophilic Archaea (e.g., *Halobacterium* sp. NRC-1). These nanostructures are spindle- or cylinder-shaped with a hydrophobic interior, having dimensions from 30 to 250 nm in width and from 50 to 2 μm in length. The most interesting properties of GVNPs include structural stability, monodispersibility, non-toxicity, self-adjuvanticity and ease of engineering. Some studies in the literature describe applications of GVNPs in drug delivery, in antigen display for vaccines, as contrast agents for ultrasound imaging and as acoustic biosensors (6–11). Nevertheless, among all these protein-based nanoparticles, protein nanocages are one of the most relevant (3–5).

Protein nanocages can be defined as highly ordered, nano-scale architectures. In general, they are produced through the self-assembly of multiple monomers, which may be identical or distinct, into symmetric and homogeneous structures of different shapes and size. These protein-based nanoparticles allow spatial control of biological processes and compartmentalization of toxic, unstable, and sensitive compounds (1,12–14).

The advantages of protein nanocages result from their distinctive intrinsic characteristics, including a high surface/volume ratio, multi-functionality and ease of modification or manipulation through genetic or chemical strategies. Additionally, high stability, monodispersibility, biocompatibility, low toxicity and biodegradability are very attractive properties in the context of biotechnological and biomedical applications. Other applications include biomineralization and nanomaterial synthesis (1,2,14).

Protein nanocages can be classified as virus-like particles (VLPs) or NVPNs. VLPs, which constitute the major group of protein nanocages, present an extensive variability in terms of structures and dimensions (15–21). While structurally similar to viruses, VLPs are not infectious since they lack genetic material (14). NVPNs, on the other hand, are unrelated to viral particles. NVPNs are formed by the self-assembly of protein monomers. This process is critically determined by the nature of the interface between adjoining subunits (2,5). A wide variety of NVPNs with different structural and functional characteristics have been described in the literature.

Key biomedical applications include drug delivery, vaccine development and diagnostic bioimaging (22–31). Non-natural, bioinspired NVPNs can also be designed *de novo* through the assembly of artificial, functional monomers (2,3,14,32).

The development of biomedical applications of NVPNs requires large amounts of pure and well-folded nanoassemblies (33). Consequently, more efficient, flexible, and universal bioprocess technologies are needed to transform NVPNs into clinically valuable products. The development of such biomanufacturing processes should be accompanied by the setting up of adequate quality control strategies to characterize both biological function and structure of the obtained nanocages (13).

This **Chapter I** first overviews general and relevant concepts related to NVPNs, followed by a presentation of their main applications in bioengineering, biotechnology, and biomedicine. The processes currently used to manufacture protein nanocages are described, and the most important aspects of upstream and downstream processing are highlighted. State-of-the-art characterization techniques are then presented and future alternative or complementary approaches in development are also discussed. Finally, a brief critical analysis of the drawbacks of the current manufacturing strategies is presented, alongside the identification of the major challenges.

I.2. Non-Viral Protein Nanocages

I.2.1. General Aspects

Non-viral protein nanocages are formed by multiple protein monomers that self-assemble into precisely defined, symmetric, homogeneous and complex structures (1,5,12,34). These nanometer size (10-100 nm) particles may originate from different prokaryotes and eukaryotes. Their structural characteristics are critical for important cellular functions, which include storage of minerals, regulation of iron homeostasis, chaperone activity for the protection of other proteins in response to high temperature, protection of DNA from oxidative damage, cargo transport of nucleic acids and catalytic support for enzymatic reactions (5,14,34).

I.2.2. Structural and Functional Characteristics

NVPNs are robust, monodisperse and water soluble, and present high biocompatibility and biodegradability (2,5,14,35). Furthermore, they can be chemically or genetically modified to extend functions and applications beyond the natural ones. Such strategies rely on molecular, genetic, and crystal structure information available in the literature (14,30,36,37). This versatility and ability to be used as a multipurpose platform constitutes one of their most interesting features. For example, since most NVPNs have intrinsic catalytic characteristics as well as the ability to carry different molecules in the inner core, it is possible to optimize them as reaction vessels and templates to synthesize metallic nanoparticles (5).

New functionalities can be engineered at interface between monomers, and the external and internal surfaces of protein nanocages (5,14,2,38). The external surface can be conjugated with functional ligands to improve targeting of a therapeutic cargo, cell penetration and biodistribution. These

ligands can be peptides, epitopes, or other small molecules. The possibility for multiple conjugation with different ligands is also attractive and limited only by steric hindrance (5,39–43). The amino acids in the monomers that compose the inner surface of the NVPNs can be replaced by specific amino acids to enable the creation of anchors for the loading of molecules with different dimensions. Depending on the type of nanocage and on the dimensions of the molecule to be encapsulated, this loading process and the subsequent binding to the inner surface can be mediated by chemical interactions (covalent, ionic, hydrophobic) or through protein-protein interactions.

I.2.3. Self-Assembly Mechanism

Self-assembly is the key to nanocage architecture. If the underlying mechanism is known (44–46), self-assembly can be modulated by destabilizing interactions at the subunit interface. Therefore, nanocage disassembly and reassembly can be induced, which allows controlling both the molecular cargo release from the inner core and the encapsulation of payloads (e.g., therapeutic molecules or enzymes). Several reports study the conditions that permit disassembly without irreversibly damaging the protein nanocage and the subsequent reassembly into its original architecture (5,47,48). Some of the most relevant factors are the pH (47,49,50), the ionic strength (51), the presence of reducing agents (52) and the presence of metals (53,54).

I.2.4. Designed NVNPs

Artificial nanocages can be designed and generated *de novo* by mimicking the intrinsic mechanisms of self-assembly of natural NVPNs. Starting with the structural characteristics (e.g., geometry, size) required for the target nanocage, functional monomers are selected and modified accordingly. The amino acid sequences in each monomer can be partially derived from natural nanocages or designed anew to promote self-assembly. Strategies and sub-methodologies used may include directed evolution, use of fusion proteins, redesign of key interfaces and the *de novo* design (46). Ultimately, monomer selection and design must guarantee that protein-protein interactions take place with minimal nonspecific aggregation (2,5,14,32,38,55).

Hybrid protein nanocages, which are conjugated with components such as polymers, nucleotides, carbon hydrates or lipids, are yet another group of protein nanocages that can be considered. These nanocages may be of major importance for diagnosis and therapy applications, for example in the context of targeting or modulation of immune response. However, hybrid NVPNs are not extensively described in the literature (2).

I.2.5. Functionalization

Natural and the artificial NVPNs can be functionalized to create nanoarchitectures more adjusted to the end applications (2,30,37). Two main strategies for the functionalization of NVPNs are described: genetic engineering and bioconjugation (14). Genetic modifications allow a more precise control over the number, position and distribution of the introduced molecules (13). Specific techniques used include

modular assembly (41,56), improvement of payload encapsulation (57–59), interface engineering (60,61), peptide display (62,63), and protein display (64). Bioconjugation on the other hand consists in the attachment of molecules that cannot be introduced through genetic engineering. The conjugation can be performed through covalent (62,65) or non-covalent bonds (2,14). The available literature indicates that genetic modification alone or in combination with bioconjugation is clearly the most efficient approach to modify NVPNs (14).

I.2.6. Examples of Natural and Artificial NVPNs

A broad spectrum of NVPNs, including both natural and artificial variants, are documented in literature. **Figure I.1** presents a schematic representation of the three-dimensional (3D) structure of natural NVPNs that have been more extensively studied. Similarly, **Figure I.2** shows a schematic representation of the 3D structure of some artificial NVPNs.

I.2.6.1. Natural NVPNs

The most relevant structural and functional characteristics of natural NVPNs are summarized in **Table I.1**. Vault nanocages (**Figure I.1A**) have extending caps at their ends, a hinged waist region, 8 small pores (2 nm in diameter) and a significant internal volume. Each vault nanocage is composed of several protein and non-protein elements, with the major vault protein (MJV) representing 70% of the overall mass. An interesting property is the dynamic ability of the vault nanocages, since they can open and close transiently, allowing the incorporation of small molecules and macromolecules within the inner core (66–74).

PepA nanocages are formed by the self-assembly of aminopeptidase (PepA), a zinc-dependent metallopeptidase (**Figure I.1B**). The resulting nanoparticles have an inner cavity with negative charge and 8 pores on the edges and faces with diameters of 1 and 3 nm. The channels are useful for the traffic of molecules. PepA nanocages were used as templates for the size-controlled synthesis of ultrasmall platinum nanoparticles, with the formation of a multifunctional biohybrid catalyst (75,76).

DNA-binding protein from starved cells (Dps) belongs to the ferritin subfamily. Dps nanocages (**Figure I.1C**) have small pores (0.8 nm) through which small molecules can diffuse. Dps nanocages were used as templates for the formation of metallic-protein nanoparticles to improve the endogenous catalytic activity of ferritins (77–79).

Ferritin nanocages (**Figure I.1D**) have an internal cavity that matches the size of the iron core. In this core, iron is stored in an insoluble non-toxic state. The cages have 8 hydrophilic pores (4 nm) that facilitate the movement of iron atoms and other small molecules. Ferritin nanocages are quite stable at high temperature and in a wide range of pH values. Self-assembly, disassembly and reassembly can be controlled by metal ions. The nanocages were applied in biomineralization and as nanocarrier of biological and non-biological molecules (44,60,80–86).

Small heat shock protein (sHSP) nanocages (**Figure I.1E**) have 8 pores (3 nm and 17 nm) that allow molecular trafficking. Advantages of the cages include high stability at high temperatures (up to 70 °C) and in a broad range of pH values (5 to 11). sHSP nanocages were used as nanoreactors, for

biomineralization, for drug delivery and for bioimaging. The functionalization reported includes the introduction of modifications with organic molecules outside and inside of the nanocages (62,87–94).

Dihydrolipoyl acetyltransferase (E2) nanocages from *Bacillus stearothermophilus* (**Figure I.1F**) have 12 pores (5 nm) and high stability at extreme temperatures due to the thermophilic nature of the native organism. E2 proteins can be modified simultaneously at the outer and inner surfaces to allow loading of drugs inside and to display functional epitopes outside (43,95–98).

Encapsulin nanocages (**Figure I.1G**) have a characteristic large central cavity (20 nm) that makes them interesting as a cargo delivery nanoplatform and as a nanoreactor (30,37,57,99–101). Lumazine synthase nanocages (**Figure I.1H**), which have a 7.8 nm inner cavity with negative charge, were used as molecular carriers through encapsulation. Other applications include the biomineralization of iron. A key characteristic of these nanocages is the stability at high temperatures (up to 95 °C) (31,59,102–106).

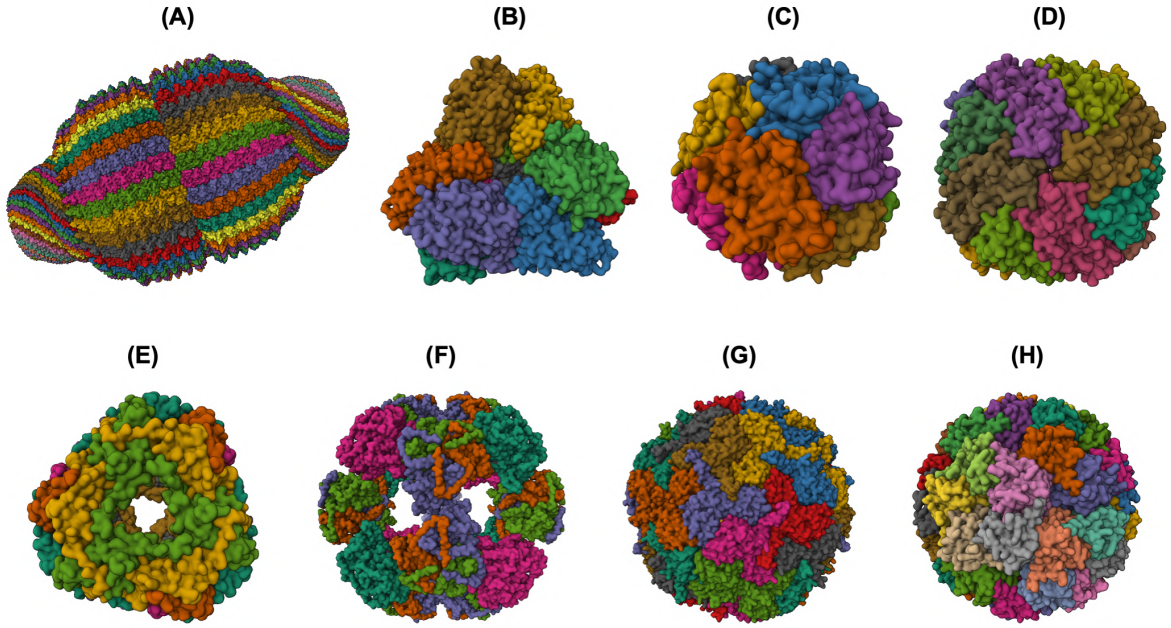


Figure I.1 – 3D structures of the most studied natural NVPNs. **(A)** Vault (PDB ID: 4V60) (66). **(B)** Aminopeptidase (PepA) (PDB ID: 3KL9) (76). **(C)** DNA-binding protein from starved cells (Dps) (PDB ID: 1QGH) (79). **(D)** Ferritin (PDB ID: 2FHA) (80). **(E)** Heat shock protein (HSP) (PDB ID: 1SHS) (87). **(F)** Dihydrolipoyl acetyltransferase (E2) (PDB ID: 1B5S) (95). **(G)** Encapsulin (PDB ID: 3DKT) (99). **(H)** Lumazine synthase (PDB ID: 1RVV) (102). Representations created using the Mol* Viewer tool (107,108).

I.2.6.2. Artificial NVPNs

Table I.2 summarizes structural characteristics of artificial NVPNs reported in the literature. The formation of symmetric and homogenous 16-nm protein nanocages (**Figure I.2A**) results from the self-assembly of monomers obtained by a fusion process. Each monomer is a geometrically controlled fusion of two natural protein oligomers, which are connected by an α-helical linker. One of those oligomers is the trimeric bromoperoxidase and the other is a dimeric M1 virus matrix protein (109).

His6-HuHF and His6-SF nanocages (**Figure I.2B**) with regulatable self-assembly were created based on two recombinant ferritins (rHuHF and rSF). Histidine motifs were incorporated in one of their subunit interfaces. Two different switches (metal- and pH-based) were developed to control the assembly-disassembly of the nanocages. This may lead to more efficient encapsulation of molecules within the nanocages compared to the traditional methods reported (110).

The trp RNA-binding attenuation protein (TRAP) nanocages (**Figure I.2C**) result from the assembly of 24 ring-shaped proteins derived from the natural TRAP from *Bacillus stearothermophilus*. Each constituent TRAP ring is formed by 11 monomers that were engineered to include a cysteine residue. Contrary to the usual situation, the complete nanocage is formed not through a network of protein-protein interactions but through the bridging of opposing thiols of the cysteine residues between TRAP rings via single gold(I) ions. The fully assembled TRAP nanocages present 6 square apertures (4 nm). TRAP nanocages are stable up to 95 °C and at high concentration of denaturing agents (e.g., 7 M urea). Although they are susceptible to reducing agents, this could be a promising characteristic for delivery to targets that contain this type of agents. The fact that the assembly relies on a metal-induced process is very useful since it provides a more rigorous control of assembly-disassembly and, eventually, a programmable mechanism. The nanocages can also be labelled with a dye in each ring-shaped monomer, which could be useful for bioimaging applications (111–116).

I3-01 nanocages (**Figure I.2D**) are hollow architectures that result from the self-assembly of multiple monomers corresponding to a trimeric 2-keto-3-deoxy-6-phosphogluconate (KDPG) aldolase from *T. maritima*. These aldolase monomers were engineered to contain complementary hydrophobic interfaces. The nanocages have several large pores (9 nm). I3-01 nanocages are stable up to 80 °C and in the presence of high concentration of denaturing agents (e.g., 6.7 M guanidine hydrochloride). Applications include synthetic biology, targeted drug delivery, and vaccine development (117–120).

Prototype oxygen-impermeable protein nanocages (OIPNC) (**Figure I.2E**) are derived from the pentameric β-carboxysome from *Thermosynechococcus elongatus* BP-1 (CcmL). Self-assembly of the proteins into nanocages occurs in the presence of quantum dots as templates by protein-quantum dots interfacial engineering. An advantage of these nanocages is the permeability to O_2 in a switchable process controlled by molecular patches. Future interesting applications include their use as nanocarriers and nanoreactors (121).

TIP60 nanocages (**Figure I.2F**) are hollow spheres with 20 triangular pores that were created using a fusion protein design approach. Each monomer is a genetic fusion of two proteins (one pentameric-LSm and one dimeric-MyoX-coil) with a three-residue linker. A reversible assembly and disassembly mechanism based on metal ions and chelators was developed for TIP60 nanocages (122–124).

I.3. Applications of Non-Viral Protein Nanocages

Applications of NVPNs can be found in a wide variety of areas and fields of study, with particular emphasis on those related to bioengineering, biotechnology, and biomedicine. Representative applications in drug delivery, vaccine development, bioimaging and diagnostic imaging,

biomineralization and nanomaterials synthesis, and biocatalysis are briefly discussed in the following sections and presented with more details in the **Supplementary Table I.S1**.

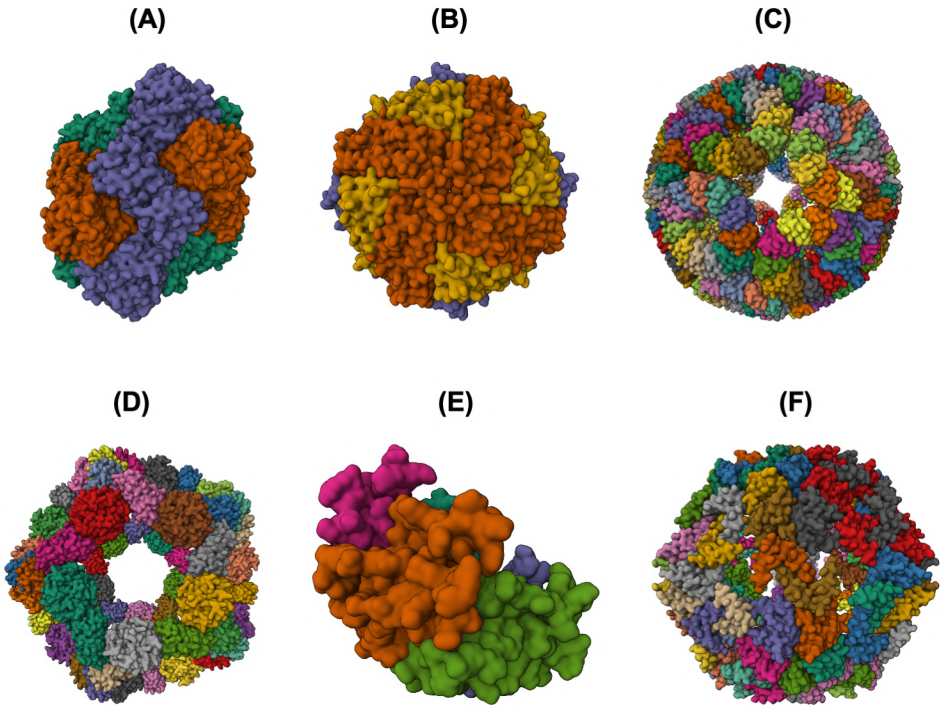


Figure I.2 – 3D structures of some artificial NVPNs. **(A)** 16-nm protein nanocages designed (PDB ID: 3VDX) (109). **(B)** His6-HuHF and His6-SF nanocages (PDB ID: 7CPC) (110). **(C)** trp RNA-binding attenuation protein (TRAP) nanocages (PDB ID: 6RVW) (111). **(D)** I3-01 nanocages (PDB ID: 8ED3) (117). **(E)** Prototype oxygen-impermeable protein nanocages (OIPNC) (PDB ID: 7WKC) (121). **(F)** TIP60 nanocages (PDB ID: 7EQ9) (122). Representations created using the Mol* Viewer tool (107,108).

I.3.1. Drug Delivery

NVPNs constitute an excellent vehicle for the encapsulation, targeted delivery, and controlled release of drugs, which can range from small molecules to larger biomolecules like nucleic acids or proteins. Several NVPNs (e.g., Dps, encapsulin, ferritin, sHSP and TRAP nanocages) were functionalized through genetic or chemical modifications to contain targeting molecules (e.g., biotin, hepatocellular carcinoma cell binding peptides, neuropilin 1-binding peptides and PTD4 cell-penetrating peptides) and to allow the incorporation of different cargos (e.g., SnCe6 photosensitizer, aldoxorubicin, doxorubicin, OSU03012 anticancer drug, small interfering RNA, curcumin). In general, results show that NVPN-mediate delivery can be performed successfully and that the desired effect is achieved efficiently (29,30,39,114,125–127).

Table I.1 – A summary of the structural and functional characteristics of the most studied natural NVPNs. The indicated code corresponds to the respective PDB ID.

Protein nanocages	Structural characteristics				Functional characteristics		References
	Number of monomers	Molecular weight	Geometry	Dimensions*	Native organism	Biological function	
Vault (4V60, **Figure I.1A**)	78-96	13 MDa	Dihedral (39-fold) – Barrel-like structure	OL: 75 nm; OW: 42 nm	Eukarya (Example: *Rattus norvegicus*)	Involved in intracellular transport, cell signalling, cell survival and innate immunity	(66,67,72,74)
Aminopeptidase (PepA) (3KL9, **Figure I.1B**)	12	457 kDa	Tetrahedral	OD: 12 nm; ID: 6 nm	*Streptococcus pneumonia*	Hydrolysis of oligopeptides into free amino acids	(75,76,128)
DNA-binding protein from starved cells (Dps) (1QGH, **Figure I.1C**)	12	216 kDa	Tetrahedral	OD: 9 nm; ID: 5 nm	*Listeria innocua*	Prevention of oxidative damage of DNA	(77–79)
Ferritin (2FHA, **Figure I.1D**)	24	509 kDa	Octahedral	OD: 12 nm; ID: 8 nm	*Homo sapiens*	Regulation of the storage and release of iron	(44,82,128)
Heat shock protein (HSP) (1SHS, **Figure I.1E**)	24	396 kDa	Octahedral	OD: 12 nm; ID: 6.5 nm	*Methanococcus jannaschii*	Chaperone activity in response to cellular stress	(87,88,90–92)
Dihydrolipoyl acetyltransferase (E2) (1B5S, **Figure I.1F**)	60	1.6 MDa	Icosahedral	OD: 24 nm; ID: n.a.	*Bacillus stearothermophilus*	Component of the pyruvate dehydrogenase (PDH) multienzyme complex	(43,97,98,128)
Encapsulin (3DKT, **Figure I.1G**)	60	1.9 MDa	Icosahedral	OD: 24 nm; ID: 20 nm	*Thermotoga maritima*	Involved indirectly in oxidative stress responses through the encapsulation of other related proteins	(37,57,99,101)
Lumazine synthase (1RVV or 1HQK, **Figure I.1H**)	60	1 MDa	Icosahedral	OD: 14.7 nm; ID: 7.8 nm	*Bacillus subtilis, Aquifex aeolicus*	Enzyme complex involved in the synthesis of lumazine (riboflavin precursor)	(102,103,106)

*OL, outer length; OW, outer width; OD, outer diameter; ID, inner diameter; n.a., not available.

Table I.2 – A summary of the structural characteristics for some artificial NVPNs. The indicated code corresponds to the respective PDB ID.

Protein nanocages	Structural characteristics				References
	Number of monomers	Molecular weight	Geometry	Dimensions*	
16-nm protein nanocages designed (3VDX, **Figure I.2A**)	12	600 kDa	Tetrahedral	OD: 16 nm; ID: n.a.	(109)
His6-HuHF and His6-SF nanocages (7CPC, **Figure I.2B**)	24	514 kDa	Octahedral	OD: 12 nm; ID: n.a.	(110)
trp RNA-binding attenuation protein (TRAP) nanocages (6RVW, **Figure I.2C**)	24	2.2 MDa	Octahedral	OD: 22 nm; ID: 16 nm	(111–116)
I3-01 nanocages (8ED3, **Figure I.2D**)	60	1.3 MDa	Icosahedral	OD: 26 nm; ID: n.a.	(117–120,128)
Prototype oxygen-impermeable protein nanocages (OIPNC) (7WKC, **Figure I.2E**)	n.a.	n.a.	Icosahedral	OD: 14 nm; ID: n.a.	(121)
TIP60 nanocages (7EQ9, **Figure I.2F**)	60	1.1 MDa	Icosahedral	OD: 21.7-24.7 nm; ID: 15 nm	(122–124)

*OD, outer diameter; ID, inner diameter; n.a, not available.

I.3.2. Vaccine Development

NVPNs can be used as platforms for antigen display, offering the possibility of co-delivery of adjuvants, targeted delivery, immune modulation, and antigen stabilisation. Engineered protein nanocages including E2, ferritin, I3-01, lumazine synthase, sHSP and vault nanocages were demonstrated as a potential vaccine platform, with the triggering of strong immune responses (namely CD8+ and CD4+ T-cell responses). Different types of antigens and other molecules were displayed on the outer and/or inner surfaces of the nanocages, such as human melanoma-associated antigen gp100, MHC I-restricted SIINFEKL peptide epitopes, SIINFEKL and ISQAVHAAHAEINEAGR peptides, and transmission-blocking and blood-stage malaria antigens (31,119,129–132).

I.3.3. Bioimaging and Diagnostic Imaging

The incorporation of contrast agents into NVPNs offers the possibility of extending their applications to the visualization of biological processes, detection of diseases, and monitoring of therapies. Ferritin and sHSP nanocages were tested in the context of bioimaging and diagnostic imaging applications. For example, the conjugation of NVPNs with targeting peptides (e.g., RGD and DEVD) and a fluorescent molecule (e.g., Cy5.5) allowed the imaging of caspase activity inside tumor cells (40). In another study, labelled protein nanocages loaded with an iron oxide nanoparticle catalyzed the oxidation of peroxidase substrates, which allowed the subsequent visualization of tumor tissue (133). Protein nanocages combined with metallic nanoparticles also showed promise in real-time *in vivo* photoacoustic imaging of tumor cells, and in positron emission tomography imaging when combined with a copper radionuclide (134). Finally, molecular magnetic resonance imaging was possible with a protein nanocage engineered with a targeting molecule (e.g., neuropilin 1-binding peptide) and combined with gadolinium(III)-chelated contrast agents (135).

I.3.4. Biomineralization and Nanomaterials Synthesis

NVPNs can be used to control and direct the formation of nanomaterials with specific properties, for example by serving as templates for the growth of inorganic minerals, by encapsulating metal nanoparticles, quantum dots, or magnetic nanoparticles, or by acting as microreactors for the controlled synthesis of nanomaterials. For example, natural and engineered Dps and sHSP nanocages were used as a nanoscale platform for the synthesis of monodisperse and homogeneous iron oxide nanoparticles (92,136). A synthetic polymer with modifiable groups was successfully incorporated in another type of protein nanocages (93).

I.3.5. Biocatalysis

The use of NVPNs for enzyme encapsulation and immobilization, substrate channelling and modulation/tuning of enzyme properties offers an opportunity to improve catalytic efficiency, alter selectivity and specificity, and enhance stability and recyclability. For example, bioinorganic hybrid

catalysts with interesting characteristics (namely, greater stability and prevention of agglomeration) were created by incorporating an enzyme (e.g., manganese peroxidase) inside a NVPN (such as ferritin, PepA, sHSP and vault nanocages) and using the biomineralization capacities (e.g., iron oxide and platinum) of these nanostructures (94,137).

I.4. Non-Viral Protein Nanocages Manufacturing

I.4.1. Overview

Efficient manufacturing of NVPNs nanostructures, all the way from the laboratory to the industrial scale, is crucial for the development of applications. However, few scientific studies available in the literature have dealt with the biomanufacturing of natural and artificial NVPNs. In addition to basic and applied research, it is essential to focus on bioprocess development, as this will play a pivotal role in bringing protein nanocages closer to the market.

Like in the case of other biological and biopharmaceutical products, the manufacturing of NVPNs involves a sequence of actions that are designed with the objective of producing a certain amount of product with specific quality features (138). These actions can be categorized into the upstream and downstream processing sections (**I.4.2. Upstream Processing** and **I.4.3. Downstream Processing**) (**Figure I.3**). The upstream processing involves the generation of the producer host cells (including host selection and cloning), cell banks implementation, inoculum preparation, cell cultivation and protein nanocage expression (139,140). The downstream processing includes all the unit operations required to purify the NVPNs to a point where final product specifications are met. Lastly, the manufacturing will require a final processing stage, which may comprise formulation, functionalization and sterilization steps to yield the desired final nanocages (14,138).

Since NVPNs are complex biologicals with an intrinsic variability in terms of composition, stability and biological activity, another important aspect to consider during process development is the implementation of analytical and characterization approaches (**I.4.4. Analytical and Characterization Technologies**). Clearly, the set-up of such methodologies is essential to determine and evaluate the structural and functional features of the final nanocages, but also to adequately monitor the performance of the biomanufacturing process throughout all its steps (138).

I.4.2. Upstream Processing

Cultivation of producer host cells and protein expression is at the core of a NVPNs manufacturing process. While the ultimate purpose is to produce large amounts of protein nanocages, scientific studies focused on application development require amounts that can be generated easily with simple lab-based protocols. In general, these protocols are directed at low-volume batches and do not rely on strict control of operating conditions. These results in low host cell densities and in reduced nanocage yields. New and optimized strategies are clearly required that can be used in large-scale settings in order to increase volumetric productivity and consequently decrease production costs (14,38,138,141).

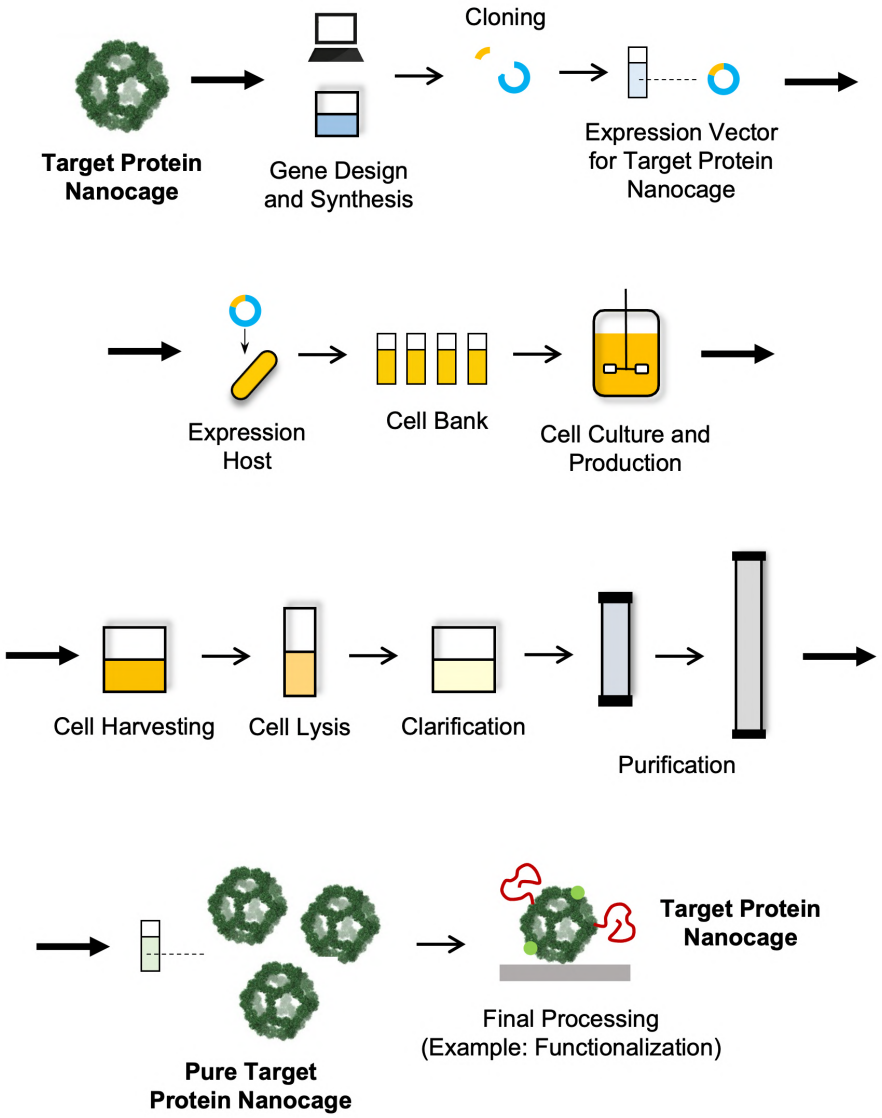


Figure I.3 – Schematic representation of a manufacturing process for NVPNs.

I.4.2.1. Selection of Host Cells

Upstream processing comprises the selection of the host cells, the preparation and optimization of the expression vector and its subsequent transformation/transfection into the host. All these aspects are crucial for obtaining a quality bioproduct and a high productivity (140,141). While the natural host can be selected if the desired NVPN is expressed naturally (99,104), nanocage production uses mainly expression systems/host dedicated to recombinant protein production (14). Here, a variety of expression hosts are available that include bacteria, yeasts and multicellular fungi, insect, mammalian and plant cells (14,140,141). NVPNs are mostly produced recombinantly in bacterial cells, and in particular in *Escherichia coli*, a very well-studied host that grows fast, is easy to cultivate and propagate at low cost, and displays high productivity. However, *E. coli* is also associated with disadvantages such as the lack of proper post-translational modifications (PTMs), the formation of inclusion bodies and the propensity to generate endotoxin contamination due to its Gram-negative nature (139–141). These can be

circumvented using yeasts such as *Saccharomyces cerevisiae* and *Pichia pastoris*, which also grow fast, are easy to manipulate genetically and can perform PTMs. Mammalian cells are more suitable for the production of larger and more complex nanocages that might also require PTMs. The most common and used cell lines are Chinese hamster ovary (CHO), Sp2/0 and NS0. Insect cells like Sf9 can also be used for nanocage production through a baculovirus expression vector system (139–141).

I.4.2.2. Design of Expression Vector and Preparation of Recombinant Host Cells

The selection of host cells is accompanied by the design and construction of an expression vector/system that can drive a highly efficient nanocage expression (140). The expression vectors will comprise the gene that codes for the NVPN but also any additional genetic elements required to improve stability or enable functionalization. These can be point mutations to change or remove standard reactive residues or to introduce unnatural amino acids, or sequences of functional peptides or proteins that are fused to the N- or C-terminus or internal loops of the monomers (14).

Most types of NVPNs reported in the literature (PepA, Dps, E2 nanocages, encapsulin, ferritin, sHSP, lumazine synthase, I3-01 nanocages, Tet8-M nanocages, TIP60 nanocages and TRAP nanocages) were produced in *E. coli* (**Table I.3**) using pET-based expression vectors. **Table I.3** lists the most used key strains of *E. coli* as well as alternatives in terms of producer host organisms. Contrary to most NVPNs, vault nanocages (**Figure I.1A**) cannot be produced in *E. coli* due to their eukaryotic origin, which implies complex post-translational modifications and protein folding (14). In this specific case, yeast, insect, or mammalian cells must be used as producer hosts (**Table I.3**). In yeast, the expression vector comprises a glyceraldehyde 3-phosphate dehydrogenase promoter (142) and in insect cells a Bac-to-Bac method based on infection with a recombinant baculovirus is necessary (58,69,72,130,143–145).

Once recombinant host cells with the expression vector are established, host cell banks are prepared that must be properly characterized and stored. This allows the maintenance of the reproducibility and consistency of the process, since each batch of protein nanocages will be manufactured using the same cell source (138,140).

I.4.2.3. Cultivation of Host Cells and NVPNs Expression

Following bank preparation, host cells are cultivated and the NVPN is expressed. Initially, a screening should be performed using small-scale cultures to identify and evaluate the impact of cultivation and operation parameters on protein expression levels. These parameters include media composition, temperature, agitation, aeration, cell density, pH, inducer concentration and induction time, among others (141,146). After the process is defined at a small scale, scale-up and pilot studies are performed in highly controlled bioreactors as a steppingstone to industrial scale implementation. Bioreactors can be operated in batch, semi-batch, and continuous/perfusion modes, depending on the method used to supply nutrients, circulate the culture medium, or recover the target NVPN. Large scale manufacturing needs to be run using optimal operation conditions that maximize expression and yield of NVPNs. Specific parameters of the bioreactor (aeration, dissolved oxygen, CO_2, and hydrodynamic

shear) also need to be tested and optimized to guarantee high specific and volumetric productivities. A design of experiments (DOE) approach may be used as more efficient strategy to optimize these parameters and understand the interaction effects between them (138,141,146).

NVPNs studied in the literature are produced mostly in shake flask cultures with volumes up to 4 L (147,148), mainly using Luria-Bertani (LB) (125,136,148) or 2× YT growth media (39,126,135,149). When under the control of a T7 promoter, expression is induced with the addition of 0.1 mM (110,123,149) to 1 mM (24,49,97,147,148,150,151) of isopropyl β-D-1-thiogalactopyranoside (IPTG) to the culture during the exponential growth phase. Other parameters that may vary are the induction time, between 2 h (40,87,90,136,150) and 30 h (24,147–149,152), and the temperature at and after induction, which is frequently 20-25°C (24,147,148,152) or 37°C (147,148). Overall, data regarding selection and optimization of process and operating conditions, as well as potential scale-up approaches, is very limited.

A key aspect to consider when using prokaryotic expression systems is whether protein overexpression leads to the accumulation of NVPNs as intracellular inclusion bodies. To minimize this, conditions should be optimized to increase nanocage solubility. For example, Zou and co-workers evaluated the effect of temperature (20 °C, 15 h and 37 °C, 5 h) on the expression of the heavy and light chains of human ferritin and on the co-expression of molecular chaperones. The authors concluded that the amount of soluble protein nanocages increased with the lower temperature and with the presence of chaperones to help in the folding process (147,148). If formation of inclusion bodies cannot be avoided altogether, solubilization and refolding steps must be consider in the downstream processing (141,146). A study by Palombarini and colleagues on the expression of ferritin monomers optimized the concentration of IPTG (0.1, 0.5 and 1 mM), the induction time (4, 8 and 16 h) and the induction temperature (25 and 37 °C). The addition of 0.5 mM IPTG at 25 °C for 16 h was identified as optimal and subsequently implemented at a large-scale production of ferritin nanocages by the biotechnology company GeneScript (152).

Martín and co-workers developed a strategy to produce vault nanocages in mammalian cells as a faster and more efficient alternative to the traditional expression in yeast or insect cells. An engineered vault nanocage (His-tagged major vault protein) was successfully produced in the human embryonic kidney 293F cell line through a transient gene expression (153).

I.4.2.4. Alternative NVPNs Expression Strategy

Cell-free protein synthesis has been widely explored in recent years as an alternative to cell-based expression systems but reports of its use in NVPNs production are scarce. In one example, Mrazek described the obtention of engineered vaults using a cell-free wheat germ expression system and either a DNA vector or an mRNA encoding the major vault protein. Notably, the author was able to simultaneously package passenger molecules in the internal cavity of the formed vaults by adding them to the synthesis mixture (154). However, due to its current limitations and challenges, cell-free protein synthesis does not seem to be the most suitable strategy for consistent production of NVPNs at large-scale (155–160).

Table I.3 – A summary of organisms and respective strains or cell lines used in the production of NVPNs.

Organism	Strain/Cell line	Protein nanocages	References
Bacteria (*E. coli*)	BL21(DE3)	Dps	(29,136)
		E2	(27,49,131)
		Encapsulin	(30)
		Ferritin	(152,161,162)
		His6-HuHF and His6-SF nanocages	(110)
		I3-01 nanocages	(118)
		Lumazine synthase	(31)
		PepA	(76)
		sHSP	(40,62,88–90)
		TIP60 nanocages	(123)
		TRAP nanocages	(111,114)
	BL21(DE3)B	sHSP	(92)
	BL21(DE3)-RIPL	I3-01 nanocages	(119)
	BL21(DE3)C+RIL	E2	(50)
	BL21-Gold(DE3)	sHSP	(126,135,151)
	BL21-CodonPlus(DE3)	sHSP	(39)
	BL21-CodonPlus(DE3)-RIL	Ferritin	(163,164)
	JM109(DE3)	sHSP	(129)
	C43(DE3)	Encapsulin	(101)
Yeast (*P. pastoris*)	SMD1168	Vault	(142)
Insect (*Spodoptera frugiperda*)	Sf9	Vault	(58,69,72,130,143–145)
Mammalian (*Homo sapiens*)	Human embryonic kidney 293F	Ferritin	(165)
		Vault	(153)

I.4.3. Downstream Processing

I.4.3.1. Overview

Downstream processing encompasses the extraction, isolation, and purification of the target NVPNs from the broth culture obtained in the upstream stage. The end-product must conform to a predetermined set of specifications that are established with the final intended use in consideration.

Ideally, the overall downstream process will comprise a small number of high-yield unit operations so as to minimize complexity, residence time and processing costs (138). Typically, unit operations are selected and implemented that explore different physical-chemical properties of the specific protein nanocages and of the associated impurities (genomic DNA, RNA, host proteins, and endotoxins). In general, the downstream processing of NVPNs will include primary isolation and recovery steps such as cell harvesting, cell lysis and clarification, and then a number of purification steps (138,141,146). An overall description of this downstream processing is illustrated in **Figure I.4**.

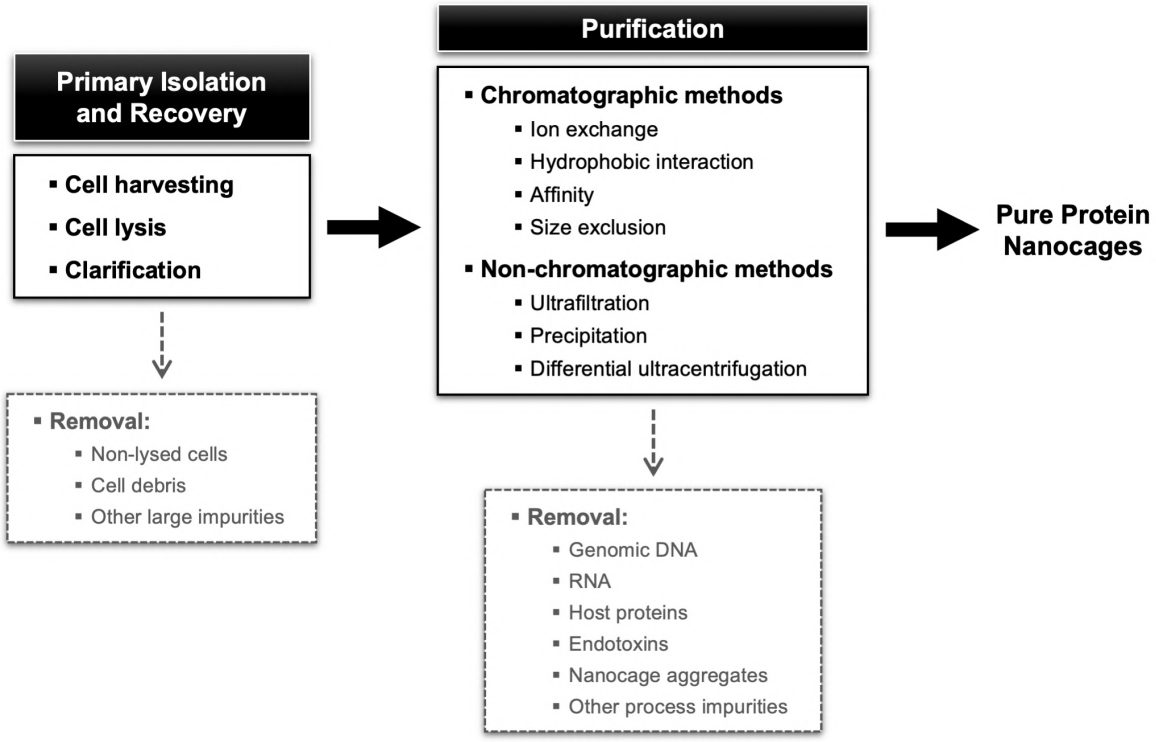


Figure I.4 – Block diagram for a standard NVPNs downstream processing showing different possibilities of unit operations.

I.4.3.2. Primary Isolation and Recovery

In the first step of the downstream processing, NVPN producer cells are harvested from the culture broth, for example by centrifugation at moderate speeds (e.g., 5000×g) (89,147,148,153). Next, cells are resuspended in an appropriate buffer, and the intracellular nanocages are released via a specific cell lysis step. The most common options at lab scale are ultrasonication (23,24,29,40,101,125,136,153), French press (29,50,96) and Dounce homogenization cell lysis (72,145). In some instances, lysozyme is added to the lysis buffer to weaken/break down bacterial cell walls and hence improve lysis (29,125,136). Nucleases (e.g., DNase, RNase) are also added often to lysis buffers to promote degradation of nucleic acids as they are released (29,125,136). Protease inhibitors are also an option in some cases (72,119,145,153). For large-scale operation, high-pressure

homogenizers and bead milling will certainly be more suitable than sonication, French press or Dounce homogenization cell lysis (141).

After cell lysis, a clarification step based on centrifugation or filtration is normally implemented to remove cell debris and particulate matter. If a situation exists where nanocage inclusion bodies are formed, these must be separated from cell debris and recovered, solubilized and then adequately refolded (49,97,141,146). For example, Jeon and co-workers have shown that inclusion bodies of ferritin nanocages can be solubilized with 8 M urea and then refolded on a nickel ion chelate affinity column with a downward gradient of urea from 8 M to 0 M (24).

Clarified lysates containing the NVPNs are sometimes subjected to heat treatment prior to purification. This approach explores the fact that protein nanocages are in general thermostable at high temperatures. Thus, heating clarified *E. coli* cell lysates at 65 to 90 °C will promote the denaturation and precipitation of host proteins, which are subsequently removed by centrifugation (e.g., 12000×g), without affecting the structure of the thermostable nanocages (14). This strategy was applied successfully to lysates containing a range of different nanocages, including Dps (29,136), E2 nanocages (49,97), encapsulin (30,101), ferritin (23,65,82,125,150,163), sHSP (62) and I3-01 nanocages (118).

Several studies also report on the enzymatic treatment of nanocages-containing clarified *E. coli* cell lysates. In the majority of cases, a DNase or an RNase is used directly in the clarified lysates (39,150) or previously added to the lysis buffer (29,62,92,123,125,136) to degrade host cell nucleic acid impurities such as genomic DNA and RNA that are co-released from cells during lysis. The resulting products of degradation (e.g., short oligonucleotides) are easier to remove in subsequent purification steps. Protease inhibitors may also be added at this stage to minimize enzymatic degradation of nanocages (119).

I.4.3.3. Purification

The NVPNs-containing clarified lysates generated in the intermediate recovery stage and entering the final purification stage will often contain host derived impurities that must be removed to obtain a bulk product complying with final-specifications (**Figure I.4**). Apart from genomic DNA, RNA, host proteins and other macromolecules (e.g., lipopolysaccharides in the case of *E. coli*), it is also important to remove misfolded nanocages or aggregates. Given this range of impurities, the purification stage is likely to comprise more than one step. Data on the downstream processing available in the literature primarily comes from laboratory protocols. This limited information makes it difficult to compare and weigh the merits of the different methodologies used.

I.4.3.3.1. Chromatographic Methods

Column chromatography is the preferred method for obtaining highly pure nanocages. The interaction modes explored include ion exchange (IEX), hydrophobic interaction (HIC), affinity chromatography (AC) and size exclusion (SEC). In IEX, HIC and AC, nanocages are retained by interacting with the stationary phase, whereas impurities flowthrough and/or elute under suitable buffer conditions. Nanocages are then eluted using an adequate buffer. In general, the nanocage-containing

fractions obtained are concentrated relatively to the feed (138,139). Despite its known shortcomings, SEC is an attractive purification option because nanocages are in general larger than most host-derived impurities.

The large size of nanocages (10-100 nm) is likely to impact the performance of IEX, HIC and AC. For once, and on account of their size, the diffusion coefficients of protein nanocages are inherently small. Furthermore, since most chromatographic matrices feature pores with diameters that seldom exceed the 30 nm (166), intrapore diffusion will be hindered. This may translate into internal mass transfer limitations, which can result in broad peaks, low recovery, and the need to use small flow rates. This proximity of nanocage and pore size is also likely to translate into poor binding capacities. One possible way to overcome these capacity limitations is to use stationary phases like chromatographic membranes and monoliths that are engineered to accommodate very large pores (> 200 nm). This strategy is in line with what is used for the purification of other very large biologicals such as plasmids, bacteriophages or VLPs (138,167).

I.4.3.3.1.1. Ion Exchange and Hydrophobic Interaction Chromatography

In IEX, anion-exchange (AEX) resins are used almost exclusively due to the anionic nature of the outer surface of most NVPNs. However, negatively charged impurities such as nucleic acids and lipopolysaccharides can also bind to anion exchangers and thus affect binding capacity and performance. The feed to AEX columns is usually a clarified lysate of the producer host cells (microbial or mammalian). Examples of resins used include strong anion-exchangers with quaternary amine functional groups such as Uno-Q, Q Sepharose, HiTrap Q and HiPrep Q (27,39,40,49,61,89,90,97,131,135,136,151,165) and weak anion-exchangers such as DEAE Sepharose (110). In most cases, sodium chloride gradients (up to 1 M) are employed to increase the ionic strength and elute the nanocages. If an adequate combination of stationary phase, operating conditions and elution scheme is used, a substantial amount of host impurities can be removed by AEX. HIC has also been used to purify ferritin nanocages using a Phenyl Sepharose resin (163,164). The key disadvantage of HIC, however, is the need to use large amounts of salts to promote binding.

I.4.3.3.1.2. Size Exclusion Chromatography

The removal of most impurities by AEX is usually followed by a SEC polishing step. Examples of SEC resins used here include Superose 6 (49,97,131,136,165), Superdex 200 (40,61,110), TSKgel G3000SW (39,135,151) and Sephacryl S-200 (89,90). The key goal when using SEC for polishing is to separate nanocages from similarly sized impurities such as nanocage aggregates and misassembled variants. Baseline peak separation of the later impurities will in general be difficult to achieve due to the limitations of resolution inherent to SEC. Removal of traces of host impurities and buffer exchange are also afforded by SEC.

I.4.3.3.1.3. Affinity Chromatography

AC is often used for the purification of engineered variants of natural and artificial NVPNs. The method requires the incorporation of an affinity tag in the primary sequence of monomers, to enable the capture of assembled nanocages by an affinity resin modified with the appropriate ligand. For example, poly-histidine tags in combination with nickel ions chelate affinity resins (e.g., Ni-NTA, Ni Sepharose 6, Ni-NTA-Sefinose, HiTrap Chelating) are widely used to purify a range of protein nanocages (24,76,123,126,147,148,168). Similarly, vault nanocages produced in 293F mammalian cells were purified using commercial magnetic particles in a platform based on immobilized metal affinity chromatography (IMAC) chemistry (153). In another instance, a maltose-binding protein (MBP) domain was fused to the N-terminus of TriEst, an esterase monomer used to assemble artificial protein nanocages Tet8-M. Purification was then performed with a maltose affinity resin (MBP-Trap) (149). While AC provides high selectivity and specificity, leading to high purity levels, in some cases a second purification step is performed, for example by SEC (118,119,149,168) or an AEX (76). Possible downsides of AC are related to the requirement to add affinity tags to monomers. Apart from the extra effort involved, the presence of tags can eventually compromise the self-assemble process or alter the properties of the original nanocages.

I.4.3.3.1.4. Combinations of Chromatographic Steps

Although the AEX-SEC combination is widely used, Santambrogio and co-workers proposed a scheme for the purification of the heavy chain of mouse ferritin nanocages that combines a first SEC step (Sepharose 6B resin) followed by an AEX step (HiTrap Q resin) (150). A similar approach was used by Jung and colleagues when purifying E2 nanocages (SEC with a Superdex 200 resin and AEX with a Mono-Q resin) (96). Finally, Bova and co-workers established a three-step chromatographic purification of sHSP nanocages, which involved AEX (Mono-Q resin), HIC (Phenyl Sepharose resin) and SEC (Superose 6 resin) (88). While combinations of chromatographic steps are the norm when purifying nanocages, some reports describe the purification of NVPNs with either a single AEX (DEAE Sepharose and HiPrep Q) (50,125) or a single SEC step (Superose 6, Superdex 200 and Sephacryl S-400) (23,29,30,92,123,134).

One downstream processing that is particularly interesting and unique among NVPNs involves artificial TRAP nanocages. Unlike other protein nanocages that undergo purification after self-assembly, TRAP nanocages are assembled *in vitro* in the presence of gold ions. This means that the individual sub-units of TRAP nanocages, which are composed of 11 monomeric TRAP proteins, must be purified in advance. Heddle and his group elaborated a downstream strategy that starts with a heat treatment of the clarified *E. coli* lysate containing the TRAP rings followed by AEX (Q Sepharose or HiTrap Q resins) and SEC (Superdex 200 resin) steps (111–114,116).

I.4.3.3.1.5. Pre-Chromatography Processing

In some studies the target protein nanocages are precipitated with ammonium sulphate prior to AEX or SEC chromatography to remove nucleic acid impurities (89,90,96,110,118,150,169). Differential

ultracentrifugation is also described as pre-purification step before chromatography (AEX or SEC) (82). For example, encapsulins were subjected to ultracentrifugation with sucrose gradient (e.g., 10 to 50% (w/v)). However, this approach has clear drawbacks such as the need for high centrifugation speed (e.g., 100000×g) and extensive centrifugation time (e.g., 18 h), and the lack of scalability (30,101).

I.4.3.3.1.6. Chromatographic Process Yields

Despite limited, some process yield data for purified NVPNs can be summarized. For example, yields of 56 mg and 25 mg per liter of cell culture were obtained for ferritin nanocages assembled from heavy chain sub-units and purified with a single AEX step (125) and for ferritin nanocages purified with a single SEC step (23), respectively. On the other hand, yields of 15 mg (purity > 90%) and 10 mg (purity of 96%) per liter of cell culture were obtained when using Ni-affinity chromatography for the purification of the heavy and the light chains of ferritin nanocages, respectively (147,148). For an IMAC-based platform, Martín and co-workers verified a recovery of 90.4% in terms of the vault nanocages in the soluble fraction, obtaining a protein concentration of 20 μg mL^{-1}. These authors concluded that 5 mg of the magnetic nanoparticles can capture 30 μg of vault nanocages, with a purity greater than 85% (153). A two-step chromatography purification (SEC + AEX) yielded 15 mg and 7 mg per liter of cell culture of the heavy and the light chains of ferritin, respectively (150). Other reported values include 20 mg L^{-1} of cell culture for E2 nanocages purified by AEX + SEC (49,97), 400 mg L^{-1} of cell culture for artificial protein nanocages Tet8-M purified with AC (149) and 50 mg L^{-1} of cell culture and 40 mg L^{-1} of cell culture for encapsulin and ferritin nanocages, respectively, when purified by tandem ammonium sulphate precipitation and SEC (101,169). A final yield of purified TRAP rings of 1 to 2 mg L^{-1} of cell culture was obtained with a combination of AEX and SEC steps (111–114,116).

I.4.3.3.2. Non-Chromatographic Methods

Palombarini and co-workers devised and suggested an alternative methodology for large-scale NVPNs purification that eliminates chromatographic steps while maintaining high efficiency and potentially reducing costs (152). Specifically, a ferritin nanocage-containing lysate obtained by sonication was subjected to an initial heat treatment and then clarified by vacuum filtration aided by diatomaceous earth. Next, the clarified lysate was purified by crossflow ultrafiltration using a 100 kDa cut-off membrane module. The system was operated in concentration and diafiltration modes, producing a stream with a final concentration of 20 g L^{-1}. The process was able to eliminate critical impurities, including genomic DNA and non-targeted proteins. The methodology offers the advantage of regenerating and reusing the filtration modules multiple times, depending on the sample source and quality.

The immunogenic and pyrogenic nature of endotoxins, which mainly include lipopolysaccharides (LPS) from Gram-negative bacteria, is a significant concern when purifying biological products. Regulatory agencies such as the Food and Drug Administration (FDA) and the European Medicines Agency (EMA) establish maximum allowed limits for the quantity of endotoxins that must be met when manufacturing of nanocages for biomedical applications (170,171). When exploring methods to remove

endotoxins, it is crucial to consider the substantial size of nanocages, which makes them more prone to interact with endotoxins, as well as the requirement to preserve their architecture (161). Silva and colleagues employed an Endotrap HD resin in conjunction with the detergent Triton X-114 followed by a polishing step with SEC to remove endotoxins from the heavy chain of ferritin nanocages. Recovery of nanocages was 57% and the final concentration of endotoxins was 0.83 EU mL^{-1}, which is lower than the maximum acceptable limit for *in vitro* and *in vivo* biomedical purposes (161). Additionally, Molino and his research group used a method that relied on successive washes of E2 nanocages with Triton X-114 to successfully decrease the concentration of endotoxins to acceptable values (131).

Vault nanocages whose production was performed both in *P. pastoris* SMD1168 and Sf9 insect cells were purified using a discontinuous density gradient ultracentrifugation after cell lysis (72,142,145).

I.4.3.4. Representative Downstream Processes

To illustrate the diversity of options available in terms of producer hosts and unit operations, four different block diagrams of NVPNs manufacturing processes are schematized in **Figure I.5**. These block diagrams were adapted from the literature and are representative of the current panorama in the downstream processing of protein nanocages.

Most of the downstream processing studies available refer to NVPNs produced in *E. coli*, which is clearly the most common microbial production host. Blocks diagrams A (**Figure I.5A**) and B (**Figure I.5B**) exemplify the downstream step respectively of ferritin (150) and sHSP (126) nanocages produced in *E. coli* and purified by chromatography. On the other hand, block diagram C (**Figure I.5C**) illustrates the production of vaults nanocages in an eukaryotic host, the insect (*Spodoptera frugiperda*) cells (145). Finally, block diagram D (**Figure I.5D**) shows a downstream process of ferritin nanocages produced in mammalian (*Homo sapiens*) cells (165).

I.4.4. Analytical and Characterization Technologies

The establishment and validation of analytical and characterization technologies for assessing the structural and functional characteristics is crucial to confirm whether the purified nanocage is within specifications. Furthermore, analytical techniques are also relevant to monitor the performance, robustness and consistency of the different manufacturing steps (138). Some well-established techniques routinely used for the determination and assessment of the biophysical characteristics of different types of NVPNs are summarized in **Table I.4**. Dynamic light scattering (DLS) and transmission electron microscopy (TEM) of negatively stained preparations allow the estimation of the average hydrodynamic diameter of nanocages. Microscopy provides a means to observe the morphology of nanostructures and compare it with the corresponding theoretical 3D structure available in databases. Furthermore, TEM is also useful to visualize the biomineralization of nanocages. The use of analytical SEC (typical resins/columns are Superdex 200, Superose 6 and TSKgel G4000SW) enables the estimation of the molecular weight of the assembled nanostructure by reference to a calibration curve of high and low molecular weight proteins. Further, it can be used to evaluate the oligomeric state of the nanocages at the end of the bioprocess.

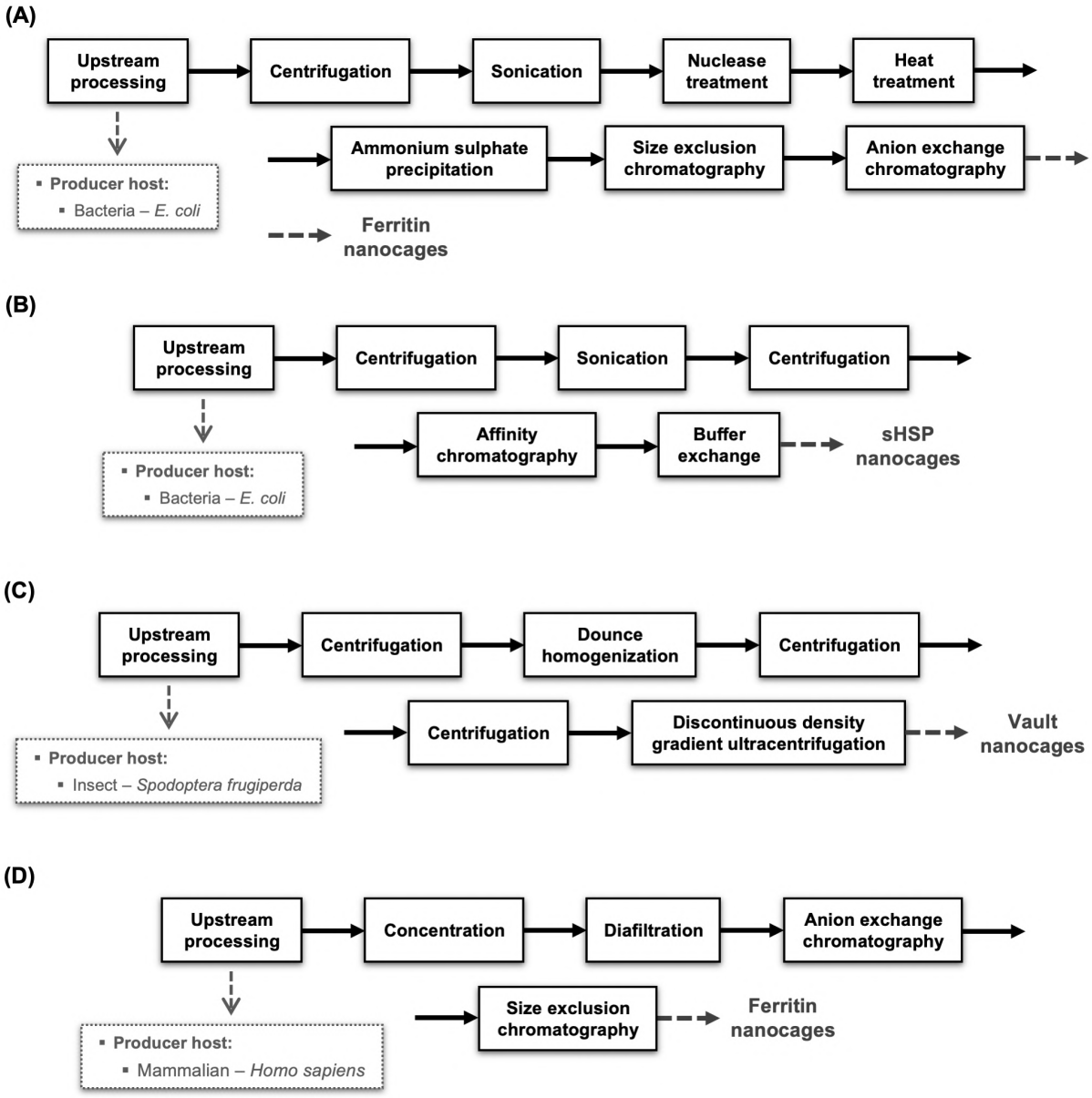


Figure I.5 – Representative examples of downstream processes implemented for distinct NVPNs produced in different organisms. Blocks diagrams **A** and **B** illustrate the downstream step respectively of ferritin and sHSP nanocages produced in *E. coli* and purified by chromatography. Block diagram **C** shows the production of vaults nanocages in insect (*Spodoptera frugiperda*) cells. Block diagram **D** exemplifies a downstream process of ferritin nanocages produced in mammalian (*Homo sapiens*) cells.

Additionally, the literature reports the use of less common analytical techniques (summarized in **Table I.4**). These techniques are used to estimate or determine structural properties of the target protein nanocages including the molecular weight, the particle size distribution, and the oligomeric state of each nanostructure. Furthermore, far-UV circular dichroism (FUV-CD) spectroscopy is used to characterize the secondary structure, the folding, and the thermostability of the protein nanocages.

Atomic force microscopy (AFM) is another promising tool for nanocage characterization. Apart from providing a means to visualize the nanostructures, AFM can be used to determine mechanical properties and study protein-protein interactions. Heddle and his group used AFM and high speed-AFM

to visualize and characterize TRAP nanocages (115,116). AFM was also used with encapsulin nanocages (172), ferritin nanocages (173), lumazine synthase nanocages (174), vault nanocages (175) and O3-33 artificial nanocages (174).

Although not extensively explored, methods based on fluorescence are promising due to the ease with which it is possible to label the exterior or interior of nanocages with fluorophores (e.g., Alexa Fluor 488-maleimide, Alexa Fluor 750-maleimide, fluorescein, Cys 5.5 dye) (39,40,62,65,88,92,151). For example, the detection and study of time-dependent fluctuations in fluorescence intensity afforded by fluorescence correlation spectroscopy (FCS) can be used to determine several physical and chemical parameters including translational and rotational diffusion coefficients (from which hydrodynamic diameters can be inferred), chemical kinetic rate constants and molecular aggregation (176).

Monitoring the concentration, homogeneity, and purity of the target protein nanocages throughout manufacturing is critical. Total protein is estimated by assays such as the bicinchoninic acid (BCA) (50,58,69,111,113,114,119,129,144,153), the Bradford (23,40) and the biuret (136). Methods for the specific quantitation of nanocages in a mixture would be especially useful to monitor the performance of the different steps of manufacturing. High performance liquid chromatography (HPLC) could be a potential approach, in particular based on a size exclusion chromatographic support, as it has been investigated and implemented for other large biological molecules, namely VLPs (177,178). SDS-PAGE is commonly used, but due to its denaturing characteristics it can only detect the presence of the nanocage monomers and other protein impurities (92,153,165). Native PAGE, on the other hand, can be used to check the native quaternary structure of nanocages (101,123,150).

I.4.5. Drawbacks and Challenges

Developing a manufacturing process for NVPNs is a complex task, in part due to the structural and functional nature of protein nanocages. The initial challenge for process developers is to collect available data and information related to the features of the target nanocages (179). Then, improving the efficiency of current processes and incorporating innovative unit operations into an integrated process is vital for enhancing both upstream and downstream processing of NVPNs (179). The existing literature lacks data regarding the upstream processing steps, particularly concerning the optimization of the protein expression conditions and operating parameters that impact quality and concentration of protein nanocages. This optimization process is quite challenging due to the large number of involved factors and their potential interaction. To improve the purity of the final product, it is essential to establish a consistent and effective sequence of unit operations for purification. Furthermore, there should be a greater emphasis on investigating and exploring the conditions and parameters that impact purification, particularly in the chromatography stages. One example is conducting dynamic binding capacity studies. Efforts to improve the manufacturing process of NVPNs should also encompass an investment in the enhancement of current analytics and the exploration of newer techniques (179). This is crucial since the analytical component of the manufacturing process is often challenging due to the complexity, time-consuming nature, or poor reliability of certain quantification methodologies.

Once a NVPNs manufacturing process is established in batch, there is potential for a shift towards continuous processing, which could bring several advantages, including high purity, increased

productivity, and reduced overall process costs (179,180). Improved analytical and monitoring methodologies that are capable of on-line and in-line analysis and can be integrated into the process will be required for continuous monitoring, process control, and product quantification (179). For the upstream processing. this could entail the monitoring of standard physical and physicochemical cell culture parameters (e.g., temperature, pH, dissolved oxygen, optical density, and off-gas composition), as well as the implementation of spectroscopic (e.g., UV-visible, fluorescence, near-infrared, infrared, and Raman) and light scattering techniques (e.g., DLS and multi-angle light scattering (MALS)). In the downstream processing, and apart from the more standard monitoring of UV, conductivity, and refractive index, light scattering sensors such as MALS should be considered (138,141,180–182).

Overall, developing a robust, efficient, and scalable NVPNs manufacturing process with a reduced number of unit operations is crucial to maximizing recovery yield. However, achieving this while maintaining the final quality of nanocages in a cost-effective manner remains a major challenge. As with other biological products, adjusting the manufacturing process for each type of NVPNs will be required (138,179).

I.5. Conclusions

Protein-based nanoparticles, including natural NVPNs, have gained significant attention in bioengineering, biotechnology, and biomedicine due to their intrinsic characteristics. The design of protein-based nanocages based on natural nanostructures has also shown promise for practical applications. However, there is still a need for further research to fully understand the underlying characteristics and assembly mechanisms of nanocages. While implementing biomanufacturing processes suitable for large-scale production is critical to bring nanocages closer to the market, few scientific studies have addressed the upstream and downstream processing of nanocages. *E. coli* is commonly used as the producer host but in the future, alternative host bacteria and other non-bacterial organisms could be explored. Further the standard production processes based on laboratory protocols should be modified and adapted for scale-up. The conceptual design of downstream processing of NVPNs is hampered by limited data and information on alternative steps and process yields. While the combination of AEX and SEC is the most used approach, complementary strategies such as aqueous two-phase extraction and crossflow ultrafiltration could be explored in the future. Finally, it is crucial to invest in more effective and simpler analytical and characterization techniques to determine and assess the structural and functional characteristics of nanocages. Clearly, further research is required to develop cost-effective NVPNs manufacturing processes.

Table I.4 – A summary of well-established and less common analytical techniques used to determine and evaluate the biophysical characteristics of different types of NVPNs.

Technique		Protein nanocages	References
Standard	Dynamic light scattering (DLS)	Dps, E2, encapsulin, ferritin, sHSP and vault	(40,92,101,131,135,136,147,153, 165)
		I3-01 and TRAP nanocages	(111,114,118)
	Transmission electron microscopy (TEM)	Dps, encapsulin, ferritin, sHSP and vault	(40,92,101,126,136,148,153,165)
		I3-01 and TRAP nanocages	(111,114,118)
	Analytical SEC	Dps, encapsulin, ferritin, sHSP and vault	(30,40,89,90,92,136,162,165)
		TIP60 nanocages	(123)
Less Common	High performance SEC (HP-SEC)	Ferritin	(110,152)
	SEC with multi-angle light scattering (SEC-MALS)	TIP60 nanocages	(123)
	SEC with right-angle (RALS)/low-angle (LALS) light scattering	TRAP nanocages	(116)
	Small angle X-ray scattering (SAXS)	Ferritin	(47,183)
		TIP60 nanocages	(123)
	Gas-phase electrophoretic mobility molecular analyzer (GEMMA)	Vault	(72)
	Mass spectrometry	E2	(131)
		Ferritin	(24,41)
		sHSP	(92)
	Native mass spectrometry	Encapsulin	(101)
		TRAP nanocages	(111)
	Electron spray mass spectrometry	Ferritin	(150)
	Liquid chromatography/electrospray mass spectrometry (LC/MS)	Dps	(29)
		sHSP	(62)
		TRAP nanocages	(111)
	Electrospray ionization time-of-flight (ESI-TOF) mass spectrometry	Dps	(77)
		Encapsulin	(30)
	Matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometry	E2	(49,50,61,97)
		sHSP	(39,151)
	Far-UV circular dichroism (FUV-CD) spectroscopy	E2	(27,49,50,61,96,97,184)
		Ferritin	(147)
		sHSP	(88)
	Analytical ultracentrifugation	Ferritin	(65,183)

I.6. Supplementary Material

Supplementary Table I.S1 – Detailed list of representative studies using NVPNs in different research areas in the bioengineering, biotechnology, and biomedicine fields.

Research area	Year	Protein nanocage	Observations	References
Drug delivery	2010	Dps	Engineered Dps nanocages are used to deliver a photosensitizer (a compound that produces reactive oxygen species when excited by light) to biofilm microcolonies of *Aggregatibacter actinomycetemcomitans*. The original Dps from *L. innocua* were modified by functionalization with a targeting moiety (biotin) and covalent incorporation of a SnCe6 photosensitizer. The study concluded that the light-induced activity of the targeted photosensitizer reduced the viability of *A. actinomycetemcomitans* biofilm.	(29)
	2014	Encapsulin	Encapsulins from *T. maritima* were engineered by genetic modification or chemical conjugation. Additionally, hepatocellular carcinoma cell binding peptides were displayed on the surface of the engineered encapsulins to target delivery to HepG2 cells. The anticancer prodrug aldoxorubicin was chemically bound into the nanocages. The drug was effectively delivered and released in the target cancer cells, producing a cytotoxicity effect equivalent to that obtained with free aldoxorubicin.	(30)
	2014	Ferritin	Ferritin nanocages were used to deliver the antitumor drug doxorubicin, which exerts its cytotoxic action by targeting the nuclear DNA of human cancer cells. A genetically engineered apoferritin variant (HFn) was selected. Doxorubicin-loaded HFn nanocages showing selectivity for the target cancer cells were internalized faster and more efficiently compared to free doxorubicin, which increased the drug action efficacy.	(125)
	2015	sHSP	Engineered sHSP nanocages were used as a delivery platform that specifically targets human pancreatic cancer cells. A genetic engineering approach was used to coat the surface of nanocages from *M. jannaschii* with the neuropilin 1-binding peptide (iRGD), which presents affinity for the target cells. An anticancer drug, OSU03012, was successfully incorporated inside these labelled nanocages. The loaded sHSP nanocages induced the death of pancreatic cancer cells by activating the caspase cascade more efficiently when compared with the free drug.	(39)
	2018	sHSP	Genetically modified sHSP nanocages from *M. jannaschii* were used to develop a new small interfering RNA (siRNA) delivery system. The siRNA-loaded sHSP nanocages were introduced efficiently into cancer cells and the delivered siRNA was able to significantly downregulate the desired protein expression.	(126)

Supplementary Table I.S2 – (*continued*).

Research area	Year	Protein nanocage	Observations	References
Drug delivery	2021	TRAP nanocages	Artificial TRAP nanocages were selected to demonstrate that NVPNs can be a valid platform for the delivery of drugs and other active compounds. Original TRAP nanocages were labelled on the surface with a cell-penetrating peptide (PTD4) and a fluorophore (Alexa-647) and the inner core was loaded with a protein cargo (a variant of a green fluorescent protein). The efficiency of the loading of the protein cargo was low, suggesting that additional future improvements are necessary. However, the labelled and loaded TRAP-nanocages successfully entered into MCF-7 and HeLa cells.	(114)
	2022	Ferritin	A pH-sensitive tumor self-targeting drug delivery platform was developed based on ferritin nanocages (HFn). HFn nanocages were successfully loaded with curcumin through a disassembly/reassembly strategy. *In vitro* experiments in breast cancer cell models confirmed that the curcumin-loaded HFn nanocages have higher cytotoxicity, cellular uptake, and targeting performance. Additionally, the conjugated nanocages demonstrated superior *in vivo* therapeutic effect and lower systemic toxicity.	(127)
Vaccine development	2003	sHSP	sHSP 100 nanocages were bound to a human melanoma-associated antigen gp100 during a heat shock with the aim of developing a targeted immunotherapy. Due to its chaperone activity, it was possible to complex the sHSP 110 nanocages with the antigenic protein. These complexes were strongly immunogenic, inducing an antigen-specific IFN-γ production and a cytotoxic T-cell response. Immunization of a mice model with the complexes resulted in an antitumor response.	(129)
	2012	Vault	The immunity produced in response to ovalbumin encapsulated in vault nanocages was characterized. The loaded nanocages induced strong anti-ovalbumin CD8+ and CD4+ T cell memory responses and a reasonable antibody production. Further, the antibody isotypes could be changed *in vivo*, which resulted in modifications of the vault nanocages.	(130)
	2013	E2	E2 nanocages from *B. stearothermophilus* were engineered to develop a cancer vaccine platform. The nanocages with MHC I-restricted SIINFEKL peptide epitopes displayed on the surface were loaded with dendritic cell-activating CpG molecules. The release of CpG from the E2 nanocages by a decrease in pH were able to activate bone marrow-derived dendritic cells at a lower concentration than the one needed with free CpG. The engineered and conjugated nanocages further resulted in an increased and prolonged CD8 T cell activation.	(131)

Supplementary Table I.S3 – (*continued*).

Research area	Year	Protein nanocage	Observations	References
Vaccine development	2014	Ferritin	Ferritin nanocages were tested to create an antigen delivery platform for dendritic cell-based vaccine development. The nanocages were engineered by adding one of two antigen peptides, SIINFEKL or ISQAVHAAHAEINEAGR, to the outer or inner surfaces. Antigen peptide-ferritin nanocages were efficiently delivered to dendritic cells, with both antigen peptides successfully inducing antigen-specific CD8+ or CD4+ T cell proliferations *in vitro* and *in vivo*. The result was a selective destruction of the antigen-specific target cells.	(132)
	2014	Lumazine synthase	An antigen delivery system to dendritic cells was investigated using lumazine synthase protein nanocages. The nanocages were externally labelled with ovalbumin peptides, SIINFEKL and ISQAVHAAHAEINEAGR, through genetic modifications. The peptides were efficient delivered by the labelled protein nanocages, followed by an efficient processing via dendritic cells *in vitro* and *in vivo*. Consequently, proliferation of CD4+ T cells specific for both antigen peptides was induced.	(31)
	2018	I3-01 nanocages	Artificial I3-01 protein nanocages were selected to develop a potential vaccine nanoplatform. The original nanocages were engineered to improve particle uniformity and stability, and conjugated with a SpyCatcher, allowing the binding of the antigen of interest. The modified I3-10 nanocages demonstrated a high stability to temperature, lyophilization, and freeze-thaw, among others. More importantly, they were found to have an efficiency connection of 95% to different transmission-blocking and blood-stage malaria antigens, such as the *Plasmodium falciparum* CyRPA.	(119)
Bioimaging and diagnostic imaging	2011	sHSP	An engineered variant of sHSP nanocages from *M. jannaschii* was chemically and genetically modified by incorporating RGD and DEVD peptides and conjugating Cy5.5 and BHQ3 molecules. The modified nanocages were incorporated in cancer cells via binding and subsequent endocytic internalization. Additionally, imaging of caspase activity of the sHSP nanocages in live cells and *in vivo* experiments showed that the cages accumulate specifically in tumor tissues. The modified sHSP nanocages were used to monitor the therapeutic effect of an anticancer drug treatment by imaging their caspase activity within the tumor cells.	(40)

Supplementary Table I.S4 – (*continued*).

Research area	Year	Protein nanocage	Observations	References
Bioimaging and diagnostic imaging	2012	Ferritin	Magnetoferritin nanoparticles were developed to target and visualize tumor tissues without the use of a contrast agent or a targeting ligand. Recombinant human heavy-chain ferritin (HFn) nanocages were loaded with iron oxide nanoparticles. The nanocages could bind to tumor cells that overexpress transferrin receptor 1. Further, the iron oxide core catalyzed the oxidation of peroxidase substrates in the presence of hydrogen peroxide, producing a color reaction that enabled the visualization of the tumor tissues. *In vivo* experiments showed that the nanocages could distinguish healthy cells from cancer cells with high levels of sensitivity (98%) and specificity (95%).	(133)
	2016	Ferritin	Ferritin nanocages were used as template for the synthesis of ultrasmall copper sulfide nanoparticles inside the inner core. The metal-ferritin nanocages showed a higher photoacoustic tomography improvement for real-time *in vivo* photoacoustic imaging of tumor cells. Coupling of a copper radionuclide resulted in nanoparticles with characteristics of a good positron emission tomography imaging agent, which present higher tumor accumulation in comparison with free copper. Use of the nanocages in photothermal cancer therapy resulted in an increase of the therapeutic efficiency. Further, good biocompatibility was demonstrated both *in vitro* and *in vivo*.	(134)
	2018	sHSP	sHSP nanocages from *M. jannaschii* were conjugated with gadolinium(III)-chelated contrast agents and iRGD peptides, which target neuropilin-1 expressed on pancreatic cancer cells. *In vitro* and *in vivo* studies showed that the molecular magnetic resonance imaging with these protein nanocages allowed detection neuropilin-1-positive cells as well as the creation of a strong signal enhancement of spontaneous pancreatic tumors in engineered mouse models.	(135)
Biomineralization and nanomaterials synthesis	2002	Dps	Dps nanocages from *L. innocua* were used as template for the size constraining of 5 nm, homogeneous nanoparticles of the ferrimagnetic iron oxide maghemite. The electrostatic nature of the protein nanocages interior surface was found to be determinant to spatially direct the mineralization of transition metal oxyhydroxides.	(136)
	2003	sHSP	sHSP nanocages from *M. jannaschii* were used to develop a nanoscale platform for the synthesis of inorganic materials. Two variants of the nanocages (original and engineered) were used as templates to act as spatially controlled reaction vessels for iron oxide mineralization. For both variants, the result was positive, with the formation of monodispersed iron oxide nanoparticles (9 nm).	(92)

Supplementary Table I.S5 – (*continued*).

Research area	Year	Protein nanocage	Observations	References
Biomineralization and nanomaterials synthesis	2009	sHSP	A modified variant of sHSP nanocages from *M. jannnaschii* was filled with a synthetic polymer decorated with modifiable groups. The polymer network allowed a spatial control of the reactive sites and a significant increase in the stability of the nanocages (up to 120 °C). Additionally, the introduced reactive centers were demonstrated to be functional.	(93)
Biocatalysis	2004	Ferritin	The protein-constrained iron oxide mineral core of ferritin (ferrihydrite) was shown to be an efficient catalyst for photoreduction reactions. Specifically, ferrihydrite-mineralized ferritin nanocages catalyzed the photoreduction of Cu(II), yielding a stable and air sensitive colloidal dispersion of Cu(0). The authors show that the properties of a preformed material can be used for the specific synthesis of a second material, adjusting the desired physical properties.	(137)
	2005	sHSP	sHSP nanocages from *M. jannaschii* were used to create an artificial hydrogenase. Firstly, platinum nanoparticles were synthesized within the inner core of the nanocages. The platinum-sHSP nanocages (stable up to 85°C) were able to reduce H^+ to form H_2 at rates comparable to the highly efficient hydrogenase enzymes and superior to those obtained with platinum nanoparticles alone. Maintenance of the integrity in the small clusters, prevention of agglomeration and control of access to active sites were the key advantages of the metallic nanocage complexes.	(94)
	2011	PepA	Aminopeptidase nanocages from *S. pneumonia* were used as templates for the size-controlled synthesis of ultrasmall platinum nanoparticles. The resulting bioinorganic nanohybrid catalysts were found to be active over the Glu-p-nitroanilide substrate.	(75)
	2015	Vault	Manganese peroxidase (MnP) was incorporated into the vaults nanocages via fusion to a packaging domain. MnP-loaded vault nanocages were able to degrade phenol at a rate higher than the one obtained with non-encapsulated MnP. Further, the MnP conjugated to the vault nanocages displayed a significant stability in comparison with the free enzyme.	(144)

Chapter II

Thesis Objectives and Model Non-Viral Protein Nanocages

As explored in **Chapter I**, NVPNs have emerged in recent years as a new class of biologics. Different applications are being explored in several scientific fields, from bioengineering to biomedicine, that showcase the potential of protein nanocages. Although various works advanced the knowledge of the characteristics of protein nanocages and provided proof-of-concept studies on different and innovative functionalization approaches, the focus on the challenges of manufacturing has been practically non-existent. Therefore, process development and manufacturing require special attention from scientists and engineers to bring protein nanocages closer to the market.

In this context, the main objective of this work was the development of scalable and efficient processes for the biomanufacturing of NVPNs. This biomanufacturing process serves three distinct purposes. The goal at the upstream processing stage is to generate the producer host cells (including host selection and cloning), express the protein nanocage and harvest the producer cells. During the downstream processing several unit operations are used to isolate, recover and then purify the nanocages. Finally, suitable analytical techniques are implemented to characterize the structure and biological function of the nanocages.

Given the wide variety of protein nanocages with differing structural and functional characteristics, two were selected as models for this study. One of the main differences is their origin: the small heat shock protein (sHSP) nanocages from *Methanococcus jannaschi* (MjsHSP nanocages) are natural, while the trp RNA-binding attenuation protein (TRAP) nanocages are artificial.

The sHSP (deposited in the PDB database with the PDB ID 1SHS), which originates from the methanogenic hyperthermophilic organism *M. jannaschii*, is able to form a multimeric homogeneous hollow structure with an octahedral geometry (**Figure II.1**). MjsHSP nanocages are composed of 24 monomers, with a total molecular weight of 396 kDa. Each monomer has 147 amino acid residues, presents a molecular weight of 16.5 kDa and is devoid of endogenous cysteine residues. In terms of dimensions, the MjsHSP nanocage has an outer diameter of 12 nm and an inner diameter of 6.5 nm. Additionally, it features 8 trigonal and 6 square pores, which allow trafficking of small molecules. Because their natural biological functions are of chaperone activity in response to cellular stress, these nanocages present elevated stability at high temperatures (up to 70°C) and across a broad range of pH values (5-11), which make them quite advantageous in several bioapplications. The isoelectric point (pI) of the nanocage is 5.02. Interestingly, the self-assembly process occurs naturally *in vivo* through protein-protein interactions. MjsHSP nanocages were already used as a platform for drug delivery (39,126), bioimaging (40,135), biomineralization (92,93) and nanoreactors (94).

The TRAP is a natural ring-shaped protein assembly derived from the organism *Bacillus stearothermophilus*. The assembly of 24 of these TRAP O-rings results in the formation of the TRAP nanocages (deposited in the PDB database with the PDB ID 6RVW) (**Figure II.2**). This assembly is only possible because each ring-shaped TRAP is composed of 11 monomers that were engineered to include a cysteine residue. Each of these monomers (74 amino acid residues) presents a molecular weight of 8.3 kDa, with the TRAP O-rings having a molecular weight of 91.7 kDa. TRAP nanocages present an octahedral geometry, with a total molecular weight of 2.2 MDa. These NVPNs have larger dimensions than MjsHSP nanocages, namely an outer diameter of 22 nm and an inner diameter of 16 nm. When fully assembled, TRAP nanocages present six square openings. These artificial protein

nanocages have a pI of 6.49. In contrast to MjsHSP nanocages, which rely on a network of protein-protein interactions, the assembly of TRAP nanocages involves the bridging of opposing thiol groups from 10 cysteine residues between adjacent TRAP O-rings via single gold (I) ions. After this *in vitro* assembly, TRAP nanocages maintain 24 free cysteine residues, one per O-ring, which can be used for binding to specific molecules or for labeling with fluorescent dyes. As advantages, TRAP nanocages are stable up to 95°C and at high concentration of denaturing agents. The assembly being based on a metal-induced process is interesting, allowing a more rigorous control of assembly-disassembly and an eventual programmable mechanism. Potential applications of TRAP nanocages include mainly bioimaging and diagnostics imaging, with currently available studies focusing on drug delivery (114).

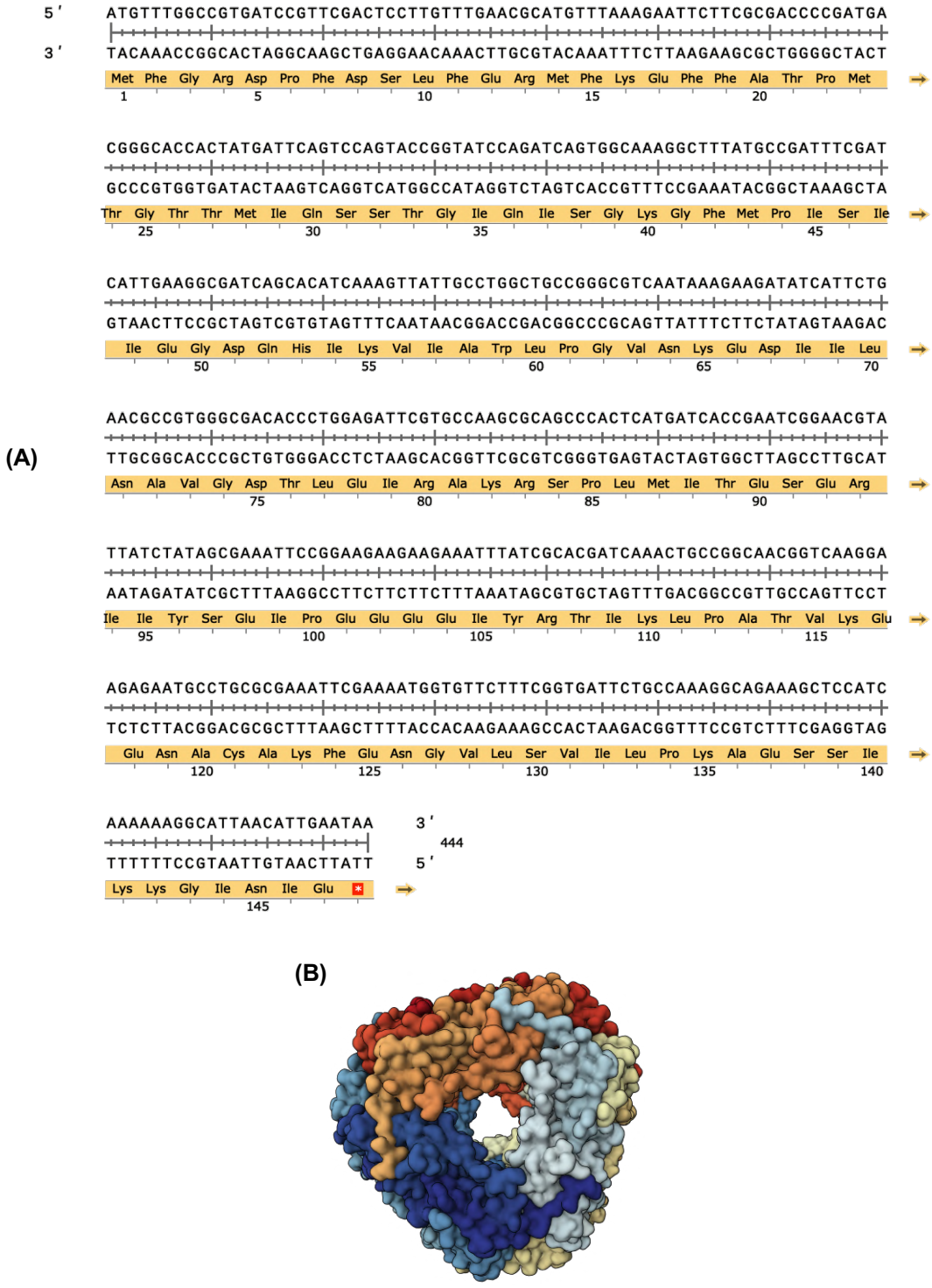


Figure II.1 – MjsHSP nanocages. **(A)** Sequences of nucleotides (441 base pairs) and amino acids (147 residues) of each of the 24 monomers (16.5 kDa) comprising the MjsHSP nanocages (396 kDa). Representation of the nucleotide and amino acid sequences were created using the SnapGene 3.2.1. software. **(B)** Tridimensional structure of the MjsHSP nanocages (PDB ID: 1SHS) (87). 3D representation was created using the Mol* Viewer tool (107,108).

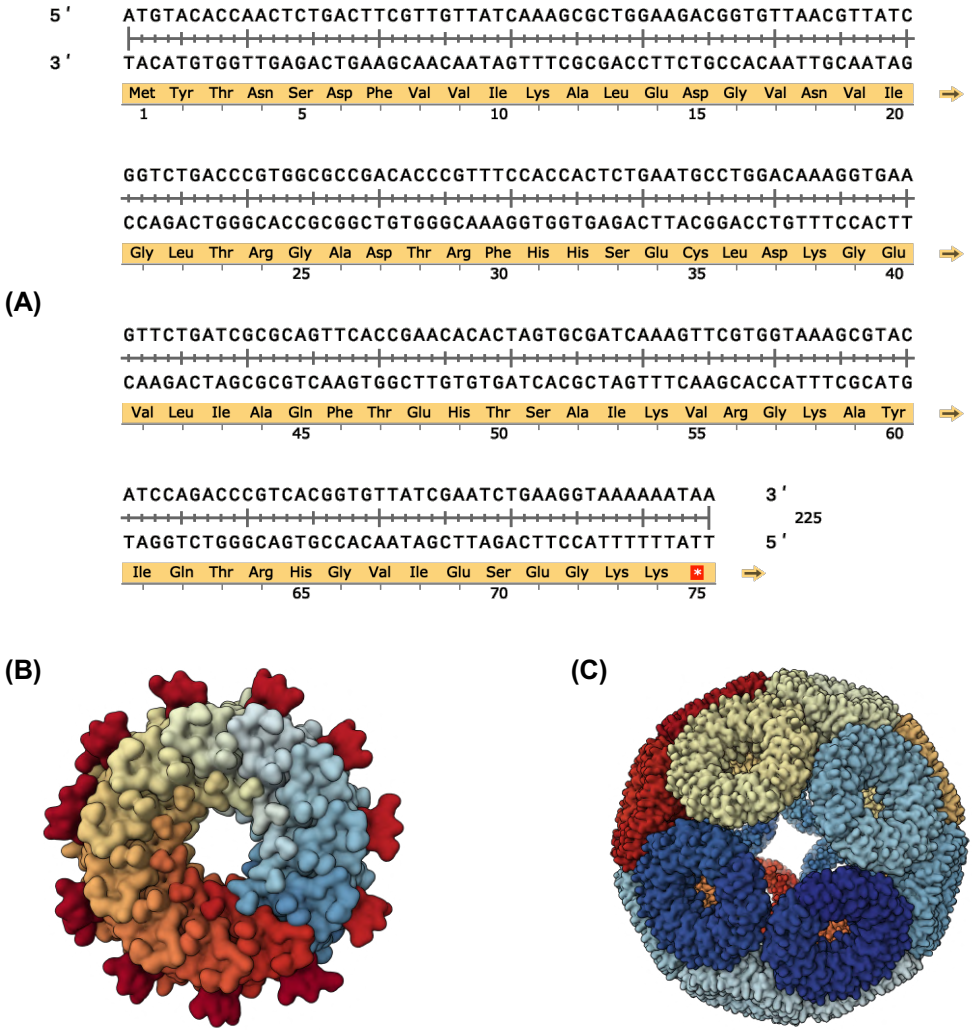


Figure II.2 – TRAP nanocages. **(A)** Sequences of nucleotides (222 base pairs) and amino acids (74 residues) of each of the 11 monomers (8.3 kDa) comprising the TRAP O-ring (91.7 kDa). The *in vitro* assembly of 24 of these O-rings forms the TRAP nanocages (2.2 MDa). Representation of the nucleotide and amino acid sequences were created using the SnapGene 3.2.1. software. **(B)** Tridimensional structure of the TRAP O-rings (PDB ID: 4V4F) (185). **(C)** Tridimensional structure of the TRAP nanocages (PDB ID: 6RVW) (111). Both 3D representations were created using the Mol* Viewer tool (107,108).

Chapter III

Upstream Processing of MjsHSP and TRAP Nanocages with *Escherichia coli* and *Vibrio natriegens* Cells

Part of this chapter is being considered for publication as an original research article:

João J, Prazeres DMF. Optimizing the production of an archaeal small heat shock protein nanocage: A dual-host study using Escherichia coli *and* Vibrio natriegens*. 2025.*

Abstract

NVPNs have recently emerged as versatile tools in several scientific fields, demonstrating significant advantages in applications such as drug delivery, vaccines development, bioimaging, biomineralization, and biocatalysis. Despite the progress in research on the properties of NVPNs for these specific applications, there is a need for more efficient manufacturing processes, particularly in the upstream processing stage. The procedures used to generate the producer host cells and express protein nanocages are based on laboratory protocols, which focus on low-volume batches. Furthermore, data on bioprocess development available in literature is very limited. Therefore, to increase volumetric productivity, it is crucial to develop innovative and highly efficient strategies that maximize host cell densities and protein nanocages yields.

This **Chapter III** is focused on the upstream processing of the MjsHSP and TRAP nanocages, with the exploration of protein expression in *V. natriegens* as one of the novelties of the work. This alternative bacterial host for the expression of proteins has received significant attention in recent years due to its advantages such as extremely fast growth, generation of large amounts of biomass and also protein per volume of cells. The production of the two model NVPNs in *E. coli* cells using standard process conditions was firstly established and used as a benchmark manufacturing method. Next, the production of the same NVPNs was implemented in *V. natriegens* cells. Several challenges were encountered during the transformation of these cells with the TRAP nanocage plasmid. Subsequently, the production process was optimized for both production in *E. coli* and *V. natriegens*. The effect of growth medium composition, concentration of protein expression inducer, temperature of protein expression, time point of induction and aeration conditions on the expression of the nanocages was evaluated.

With the optimized process in *E. coli* cells, a production yield of MjsHSP nanocages of 295.7 $\pm$ 0.0 mg L^{-1} of culture and a volumetric productivity of 53.8 $\pm$ 0.0 mg L^{-1} h^{-1} were obtained after 4 hours of expression, which corresponded, respectively, to an increase of 5.4 and 5.1 times in relation to the non-optimized benchmark production process. In parallel, in *V. natriegens* cells, a production yield of MjsHSP nanocages of 60.9 $\pm$ 13.7 mg L^{-1} of culture and a volumetric productivity of 10.4 $\pm$ 2.3 mg L^{-1} h^{-1} was obtained with the optimized procedure after the same 4 hours of expression. Although these values were lower than those of *E. coli* in direct comparison, an increase was observed in relation to the standard process in *V. natriegens* (2.2 and 1.8 times, respectively, for the production yield and the volumetric productivity). This showed the potential for the production of NVPNs in *V. natriegens*, but highlighted the need for additional studies, particularly in terms of the genetic improvement of this host bacterium. Regarding the optimized TRAP O-rings production process in *E. coli* cells, a production yield of 249.1 $\pm$ 3.7 mg L^{-1} of culture and a volumetric productivity of 46.7 $\pm$ 0.7 mg L^{-1} h^{-1} were obtained after 4 hours of expression, corresponding, respectively, to a slight increase of 1.4 and 1.3 times compared to the reference protein expression process.

Keywords: *Escherichia coli*, production, protein nanocages, upstream processing, *Vibrio natriegens*

III.1. Introduction

In recent years, NVPNs have gained notoriety in numerous fields of study, particularly focusing on those associated with bioengineering, biotechnology, and biomedicine. Applications where these nanocages have demonstrated advantageous characteristics include drug delivery (114,126,127), vaccine development (119,129,132), bioimaging (134,135), biomineralization (93), and biocatalysis (75,144). This topic was explored in more detail in **Chapter I**.

Even with the advancements in research on the features of NVPNs for these specific applications, few scientific studies available in the literature have focused on bioprocess development of natural and artificial protein nanocages (186). Therefore, with the purpose of bringing NVPNs closer to the market, more efficient manufacturing processes are needed.

Similarly to other biological products, the upstream processing of NVPNs comprises as critical steps i) the generation of the producer host cells and ii) the expression of the proteins (139,140). Generically, these steps follow laboratory protocols designed to for low-volume batches, with limited published data available. Thus, more innovative and efficient strategies are clearly required in order to obtain high host cell densities and increased nanocage yields, with the subsequent improvement of volumetric productivities (14,38,138,141).

In terms of host cells for recombinant production of protein nanocages, most publicly accessible studies use the *E. coli* bacterium as the producer organism. This is due to its attractive and well-studied characteristics, namely, ease of genetic modification by tools already established, rapid specific growth rate (a doubling time of approximately 20 minutes in optimized conditions is reported), high cell density, high level of gene expression, facility of cultivation with reduced costs, and possibility of performing a robust scale-up (187–190). Nevertheless, the formation of inclusion bodies, the risk of endotoxin contamination of the final product due to the Gram-negative nature of the bacterium, and the lack of proper post-translational modifications are generally considered as disadvantages of *E. coli* as host for protein expression (13,140,141,187).

Furthermore, the design and construction of an appropriate expression system is essential to obtain a highly efficient production of protein nanocages (140). According to the information available in the literature, the majority of NVPNs expressed in *E. coli* use pET-based expression vectors, typically with the gene encoding the desired nanocage under the control of the strong bacteriophage T7 RNA polymerase promoter system (186,187,191). Additional genetic elements and/or specific point mutations can also be included in order to enhance stability and functionality (14).

The main strain of *E. coli* used for the production of NVPNs is BL21(DE3) (62,90,111,114) or specific strains derived thereof (39,92,135,151). BL21 presents interesting features, such as extensive characterization, high biomass yield, protease deficiency, reduced acetate production, efficient protein folding, and compatibility with the T7 RNA polymerase expression system (186,191–193).

Optimization of the cultivation of host *E. coli* cells and production of NVPNs, calls for screening and assessing the impact of various parameters on protein expression levels in small-scale cultures. Depending on the protein under study, these cultivation and operation parameters can have different impacts, with the most relevant factors being media composition, temperature, agitation, aeration, cell density, pH, inducer concentration, and induction time (141,146,152,186,187). Following small-scale

optimization, the process is then scaled up through pilot experiments in controlled bioreactors. Production of NVPNs typically occurs in shake flasks up to 4 L (147,148), using LB (125,136) or 2× YT media (126,149). T7 promoter-controlled expression is typically induced with 0.1 mM (110,123) to 1 mM (97,150) of IPTG during the exponential growth phase. Induction times ranges from 2 h (40,87) to 30 h (147,152), and temperature at and after induction is around 20-25°C (24,152) or 37 °C (147,148). As identified in **Chapter I**, data related to operating conditions, process optimization and respective evaluation variables such as production yield, productivity, as well as potential scale-up approaches, are very scarce in the literature.

An alternative strategy for expressing NVPNs that has been considered is cell-free protein synthesis, although it presents significant challenges and limitations that make it less robust for large-scale production. A distinctive and promising approach could be the use of an expression system based on an unconventional bacterium, such as *V. natriegens*. This marine, Gram-negative, non-pathogenic bacterium presents an extremely short doubling time at optimal growth conditions (reported to be lower than 10 minutes) due to its strong proteosynthetic activity (194). Studies indicate that the known strains of *V. natriegens* have approximately 115000 ribosomes per cell, a number significantly higher than what is reported for *E. coli* strains (70000 ribosomes per cell) (189). Some intrinsic properties of *V. natriegens* that are potentially advantageous in terms of a cost-effective large-scale production comprise the high biomass-specific substrate consumption rate (twice as high as that of typical industrial hosts), the subsequent high specific growth rate, and the flexible metabolism allowing it to consume a wide variety of carbon substrates (including sucrose and starch) (189,195–199). All these characteristics make it an organism with enormous potential for molecular biology, protein expression, and metabolic engineering (198,200,201). Although it has not yet been applied to the production of NVPNs, there are some studies in the literature that use *V. natriegens* for the production of other proteins, evaluating its efficiency as well as comparing it with the results obtained in *E. coli* (202). A specific *V. natriegens* strain was developed commercially for protein expression (named Vmax™), which integrates a T7 RNA polymerase cassette into the genome, allowing the inducible expression of the protein of interest from compatible plasmids (198,200). Examples of proteins successfully expressed in *V. natriegens* are archaeal catalase-peroxidase (189), enhanced yellow fluorescent protein (203), FK506 binding protein (203), human growth hormone (189), tobacco etch virus protease (203), and yeast alcohol dehydrogenase (189).

In the initial part of this **Chapter III**, the preparation of expression vectors appropriate for the *in vivo* production of both MjsHSP and TRAP nanocages using *E. coli* and *V. natriegens* expression systems is described. The use of the alternative *V. natriegens* bacterial host for protein expression is one of the main novelties of this work. A study of the plasmid copy number in both species was also carried out and the values obtained were compared with those reported in the literature. The production of the two model NVPNs in *E. coli* cells was then established using reference operation parameters and used as a benchmark manufacturing method. Additionally, an equivalent production of MjsHSP nanocages was performed in *V. natriegens* cells. Unfortunately, it was not possible to express TRAP nanocages in *V. natriegens* cells due to difficulties in transforming cells with the corresponding plasmid. Afterwards, the production of MjsHSP nanocages in *E. coli* and *V. natriegens*, and of TRAP nanocages

in *E. coli*, was optimized by evaluating the impact of five parameters considered more relevant according to the literature: i) growth medium composition (from minimal medium with addition of different carbon sources to complex media with distinct components including an eventual supplementation with salts), ii) concentration of protein expression inducer (between 0.1 mM and 1.5 mM), iii) temperature of protein expression (between 15°C and 37°C), iv) timepoint of induction (between $OD_{600\ nm} \approx 0.2$ and $OD_{600\ nm} \approx 1.8$) and v) aeration conditions (non-baffled and baffled shake flasks, maintaining constant the ratio of culture volume to air volume).

III.2. Materials and Methods

III.2.1. Materials

Bacto™ Yeast Extract was acquired from BD Biosciences (Franklin Lakes, NJ, USA). Acrylamide/bis-acrylamide solution (40%), Laemmli sample buffer (4x) and precision plus protein dual color standards were from Bio-Rad (Hercules, CA, USA). Magnesium chloride ($MgCl_2$) was obtained from Fagron Ibérica (Barcelona, Spain). Acetic acid glacial, agarose, ethanol solution (96% v/v), glucose, glycerol, hydrochloric acid (HCl), isopropyl β-D-1-thiogalactopyranoside (IPTG), phosphate buffered saline (PBS) solution (10x), sodium chloride (NaCl), sodium hydroxide (NaOH) and tris(hydroxylmethyl)aminomethane (Tris) were purchased from Fisher Scientific (Hampton, NH, USA). PhastGel™ Blue R (Coomassie R 350 stain) was acquired from GE Healthcare (Chicago, IL, USA). Potassium chloride (KCl) was from Merck (Darmstadt, Germany). NEBuilder® HiFi DNA Assembly Master Mix and restriction enzymes SphI-HF (High-Fidelity), XbaI and XhoI were obtained from New England Biolabs (Ipswich, MA, USA). Dithiothreitol (DTT), dNTPs NZYMix, Luria-Bertani (LB) agar, LB broth, NZYColour protein marker II, NZYDNA ladder III, NZYDNA loading dye (6x), NZYGelpure, NZYProof DNA polymerase, dNTPs NZYMix, NZYSpeedy qPCR Green Master Mix (ROX) (2x) and SDS-PAGE sample loading buffer (5x) were purchased from NZYtech (Lisbon, Portugal). Potassium phosphate dibasic (K_2HPO_4), potassium phosphate dibasic (KH_2PO_4), sodium phosphate dibasic (Na_2HPO_4), sodium phosphate monobasic (NaH_2PO_4) and tryptone were acquired from Panreac AppliChem (Barcelona, Spain). Wizard® Genomic DNA Purification Kit was from Promega (Madison, WI, USA). High Pure Plasmid Isolation Kit was obtained from Roche (Basel, Switzerland). 1,4-piperazinediethanesulfonic acid (PIPES), ammonium chloride (NH_4Cl), ammonium persulfate (APS), ampicillin, brain heart infusion (BHI) broth, calcium chloride ($CaCl_2$), dimethyl sulfoxide (DMSO), ethylenediaminetetraacetic acid (EDTA), glycine, magnesium sulfate ($MgSO_4$), N,N,N′,N′-tetramethylethylenediamine (TEMED), polyethylene glycol 8000 (PEG 8000), sodium azide (NaN_3), sodium dodecyl sulfate (SDS), sodium sulfate (Na_2SO_4), sucrose and urea were purchased from Sigma-Aldrich (St. Louis, MO, USA). Pierce™ BCA Protein Assay Kit, ethidium bromide solution and SYBR™ Safe DNA Gel Stain (10000x) were acquired from Thermo Fisher Scientific (Waltham, MA, USA). All primers used were synthesized and obtained from STAB VIDA (Caparica, Portugal). The pure and ultrapure water used during all experiments was provided by a Milli-Q purification system (Millipore; Bedford, MA, USA).

III.2.2. Bacterial strains

E. coli DH5α [F^-, φ80*lac*ZΔM1, Δ(*lac*ZYA-*arg*F)U169, *rec*A, *end*A1, *hsd*R17(r_K^-, m_K^+), *pho*A, *sup*E44, λ^-*thi*-1, *gyr*A96, *rel*A1] was selected for the cloning and maintenance of the several plasmids used in this work. *E. coli* BL21(DE3) [F^-*omp*T, *hsd*S_B, (r_B^-, m_B^-), *gal, dcm,* (DE3)] and *V. natriegens* Vmax™ X2 [*Vibrio natriegens* 14048, dns::LacI-T7-RNAP] were used for the expression of the two

NVPNs models. Both *E. coli* strains were obtained from Thermo Fisher Scientific (Waltham, MA, USA) and the *V. natriegens* strain was acquired from Codex DNA (San Diego, CA, USA).

Master and working cell banks of these three bacterial strains were prepared according to the procedure described as follows. From the original bacterial stock, an overnight culture (16-18 h) was prepared in a 15 mL conical centrifuge tube with 5 mL of suitable culture medium, specifically LB for both *E. coli* strains and LB supplemented with V2 salts (204 mM NaCl, 4.2 mM KCl and 23.14 mM $MgCl_2$) for the *V. natriegens* strain. The incubation conditions were 37°C and an agitation of 250 rpm. In the next day, an appropriate volume of the overnight culture was used to re-inoculate the same medium at an approximate $OD_{600\ nm}$ of 0.1. Once the culture reached an $OD_{600\ nm}$ of approximately 1.0, cell bank aliquots were prepared by mixing appropriate volumes of the bacterial culture with a stock glycerol solution (50% v/v) to obtain a final concentration of 17.5% v/v glycerol. Cell banks of both *E. coli* and *V. natriegens* were stored at -80°C.

III.2.3. Plasmids

Throughout this work, four different plasmids were used: pUC57-MjsHSP, pDHFR-His, pMjsHSP and pTRAP. Their complete nucleotide sequences are available in the **Supplementary Table III.S1**.

Plasmid pUC57-MjsHSP (3035 bp), which contains the complete gene that codes for MjsHSP (UniProtKB: Q57733), was obtained from NZYtech (Lisbon, Portugal). This plasmid was used to obtain the insert fragment that codes for MjsHSP in subsequent cloning procedures. The MjsHSP gene has a single point mutation at position 121, which was introduced to replace a serine with a cysteine. This unique cysteine constitutes an adequate anchoring point for labeling the nanocages, for example with fluorophores.

Plasmid pDHFR-His (2745 bp), which was obtained from New England Biolabs (Ipswich, MA, USA), was chosen as the starting point to generate a vector backbone for subsequent cloning of nanocage coding fragments.

Plasmid pMjsHSP was constructed as described in the section **III.2.7. Construction of plasmid pMjsHSP**, using the vector backbone obtained from plasmid pDHFR-His and the insert fragment from plasmid pUC57-MjsHSP. The final construction has 2677 bp and comprises the following elements: a hybrid origin of replication (ColE1/pMB1/pBR322/pUC), a kanamycin resistance gene with its promoter, a transcription promoter and a terminator for bacteriophage T7 RNA polymerase, a ribosome binding site from bacteriophage T7, and the gene coding for the MjsHSP nanocage.

Plasmid pTRAP (5592 bp) is a pET-based vector that was kindly provided by the Heddle Lab (Kraków, Poland), and used directly without additional modifications. It comprises all elements of the well-described pET-21b(+) plasmid and the TRAP-K35C gene. A single point mutation in this gene results in the substitution of a lysine with a cysteine at position 35 (185).

III.2.4. Preparation of chemically competent cells

III.2.4.1. *E. coli*

The procedure described next was applied to both strains of *E. coli* (DH5α and BL21(DE3)).

From a frozen cell bank, a pre-inoculum was prepared in a 15 mL conical tube containing 5 mL of LB. Following an overnight incubation (16-18 h) at 37°C and 250 rpm, an appropriate volume of pre-inoculum culture was centrifuged (6000×g, 3 minutes, 4°C) and the resulting cell pellet was resuspended in fresh culture medium. This resuspended cell pellet was inoculated in a 100 mL non-baffled shake flask, containing 20 mL of LB, to obtain an approximate $OD_{600\ nm}$ of 0.1. This culture was then incubated at 37°C and 250 rpm until reaching an $OD_{600\ nm}$ of approximately 1.0. At this point, the culture was centrifuged at 1000×g for 10 minutes at 4°C. The supernatant was discarded, and the cell pellet was resuspended in a freshly prepared, cold transformation and storage solution (20 g L^{-1} LB, 5% v/v DMSO, 0.05 M $MgCl_2$, 10% w/v PEG 8000, pH 6.5; sterilized by filtration) at a ratio 1:10 of the original volume. Chemically competent cells were utilized immediately for the transformation protocol (**III.2.6. Transformation of competent cells**) or stored at -80°C until required.

III.2.4.2. *V. natriegens*

Two distinct procedures were tested based on the work described by Weinstock *et al.* (198) and Tschirhart *et al*. (204).

In accordance with the protocol developed by Weinstock *et al.*, a frozen cell bank vial of *V. natriegens* Vmax™ X2 was used to inoculate 150 mL of BHI medium supplemented with V2 salts in a 500 mL baffled shake flask. Next, the culture was incubated at 30°C and 200 rpm until reaching an $OD_{600\ nm}$ of about 0.4. At room temperature, the culture was divided into three 50 mL conical tubes and centrifuged at 3000×g for 5 minutes. The resulting supernatant was carefully removed, and each cell pellet was gently resuspended in 5 mL of 100 mM $MgCl_2$. The resuspended cells from the three tubes were combined into two 50 mL conical tubes, and the volume in each tube was adjusted to 30 mL with 100 mM $MgCl_2$. The tubes were mixed by gentle inversion. Then, a second centrifugation was performed at 3000×g for 4 minutes. Again, the resulting supernatant was discarded, and each cell pellet was resuspended in 5 mL of 100 mM $CaCl_2$. The resuspended cells were combined into one 50 mL conical tube, and the volume was also adjusted to 30 mL with additional 100 mM $CaCl_2$. The 50 mL conical tube was gently mixed by inversion and incubated at room temperature for 20 minutes. After incubation, the cells were pelleted by centrifugation at 3000×g for 4 minutes. The supernatant was discarded, and the cell pellet was resuspended in approximately 1.5 mL of transformation and storage buffer (55 mM $MnCl_2$, 15 mM $CaCl_2$, 250 mM KCl, 10 mM PIPES, 7% v/v DMSO, pH 6.7; sterilized by filtration). Once resuspended, the cell suspension was divided into aliquots in pre-chilled tubes. These chemically competent cells were either used directly for the transformation protocol (**III.2.6. Transformation of competent cells**) or stored at -80°C until needed.

In accordance with the protocol established by Tschirhart *et al*., a frozen cell bank vial of *V. natriegens* Vmax™ X2 was used to inoculate 30 mL of BHI medium supplemented with V2 salts in a

100 mL non-baffled shake flask. Following an overnight incubation (14-16 h) at 30°C and 275 rpm, an appropriate volume of this pre-inoculum culture was added to a 500 mL baffled shake flask, containing 150 mL of fresh culture medium (BHI supplemented with V2 salts), to obtain an approximate $OD_{600\ nm}$ of 0.1. This re-inoculated culture was incubated at the same operating conditions until reaching an $OD_{600\ nm}$ of approximately 1.0. At this point, the culture was divided into two 250 mL centrifuge tubes and centrifuged at 6500×g for 10 minutes at room temperature. The resulting supernatant was carefully discarded, and each cell pellet was gently resuspended in 3.25 mL of preheated 1 M sorbitol. After a 1-hour incubation, the resuspended chemically competent cells were divided into aliquots, which were used immediately for the transformation protocol (**III.2.6. Transformation of competent cells**) or stored at -80°C until necessary.

III.2.5. Preparation of electrocompetent cells

III.2.5.1. *V. natriegens*

Two different methodologies were tested based on the work described by Weinstock *et al.* (198) and Lee *et al.* (201).

In accordance with the protocol implemented by Weinstock *et al.*, a frozen cell bank vial of *V. natriegens* Vmax™ X2 was used to inoculate 30 mL of BHI medium supplemented with V2 salts in a 100 mL non-baffled shake flask. Following an overnight incubation (14-16 h) at 30°C and 200 rpm, an appropriate volume of this pre-inoculum culture was added to a 1 L baffled shake flask containing 300 mL of fresh culture medium (BHI supplemented with V2 salts), to obtain an approximate $OD_{600\ nm}$ of 0.1. This re-inoculated culture was incubated at 37°C and 200 rpm until reaching an $OD_{600\ nm}$ of approximately 0.5. At this point, the culture was distributed into two pre-cooled 250 mL centrifuge tubes and incubated on ice for 15 minutes. Then, the cells were pelleted by centrifugation at 6500×g for 20 minutes at 4°C. The resulting supernatant was discarded, and each cell pellet was gently resuspended in 10 mL of electroporation buffer (680 mM sucrose, 7 mM K_2HPO_4, pH 7.0; sterilized by filtration). The resuspended cells were combined into one 50 mL conical tube, and the volume was adjusted to 35 mL with additional electroporation buffer. This cell suspension was inverted several times to ensure thorough mixing. The cells were then subjected to a second centrifugation at 7000×g for 15 minutes at 4°C. After this step, the supernatant was removed using a pipette, and the washing process was repeated two more times for a total of three washes. Once the final wash was complete, the cells were gently resuspended in the residual electroporation buffer. The volume was adjusted with additional electroporation buffer to achieve a final $OD_{600\ nm}$ of 16.0. The resuspended electrocompetent cells were divided into pre-cooled aliquots, which were used immediately for the transformation protocol (**III.2.6. Transformation of competent cells**) or stored at -80°C until needed.

In accordance with the protocol developed by Lee *et al.*, a frozen cell bank vial of *V. natriegens* Vmax™ X2 was used to inoculate 5 mL of LB supplemented with 20 g L^{-1} NaCl (LB3) in a 15 mL conical tube. Following an overnight incubation (14-16 h) at 37°C and 225 rpm, an appropriate volume of pre-inoculum culture was centrifuged (20000×g, 1 minute, room temperature) and the resulting cell pellet

was resuspended in fresh culture medium. This resuspended cell pellet was inoculated in a 100 mL non-baffled shake flask, containing 30 mL of LB3, to obtain an approximate $OD_{600\ nm}$ of 0.1. The re-inoculated culture was then incubated at 37°C and 225 rpm until reaching an $OD_{600\ nm}$ of approximately 0.4. At this point, the culture was centrifuged at 3500×g for 5 minutes at 4°C. The supernatant was discarded, and the cell pellet was resuspended in 1 mL of cold 1 M sorbitol. Then, the cells were pelleted by centrifugation at 20000×g for 1 minute at 4°C. This wash procedure was repeated two more times for a total of three washes. The final cell pellet was resuspended in 250 µL of 1 M sorbitol. Aliquots of electrocompetent cells were either utilized directly for the transformation protocol (**III.2.6. Transformation of competent cells**) or stored at -80°C until required.

III.2.6. Transformation of competent cells

III.2.6.1. *E. coli*

For the transformation of chemically competent *E. coli* DH5α and BL21(DE3) cells (prepared as described in section **III.2.4. Preparation of chemically competent cells**), the same heat shock procedure was applied to both strains.

First, 10-50 ng of the desired plasmid DNA (pDNA) was combined with 100 µL of an aliquot of competent cells. This mixture was chilled on ice for 20 minutes, followed by a short heat shock at 42°C for 1 minute, and another 2 minutes on ice. Afterwards, the *E. coli* cells were recovered with the addition of pre-heated LB to achieve a total volume of 1 mL, then incubated at 37°C and 250 rpm for 1 hour. Subsequently, the recovered cells (concentrated to 100 µL) were plated on LB agar supplemented with the appropriate antibiotic and incubated overnight at 37°C.

III.2.6.2. *V. natriegens*

The transformation of *V. natriegens* Vmax™ X2 cells was performed using two distinct procedures, depending on whether the cells were chemically competent or electrocompetent. The preparation of these competent cells was detailed in sections **III.2.4. Preparation of chemically competent cells** and **III.2.5. Preparation of electrocompetent cells**.

For chemically competent cells, 50-100 ng of the selected pDNA were gently mixed with 50 µL of an aliquot of competent cells in a pre-cooled microtube. This mixture was incubated on ice for 30 minutes, followed by a short heat shock at 42°C for 45 seconds, and another 90 seconds on ice. Subsequently, the *V. natriegens* cells were recovered with the addition of pre-heated BHI supplemented with V2 salts (at 30°C) to achieve a total volume of 1 mL in a 15 mL conical tube. The cells were maintained at 30°C for 1 minute in a water bath and then incubated at 30°C and 200 rpm for 2 hours. Afterwards, the recovered cells were diluted and plated on pre-warmed LB agar supplemented with the appropriate antibiotic and incubated overnight at 30°C.

For electrocompetent cells, 50-100 ng of the chosen pDNA were gently combined with 50 µL of an aliquot of competent cells in a pre-cooled microtube. This mixture was transferred to a cooled

electroporation cuvette with a 0.1 cm gap size (Bio-Rad; Hercules, CA, USA). Electroporation was performed in a Bio-Rad electroporation system (Hercules, CA, USA) using the following parameters: 800 V, 25 µF and 200 Ω. Following electroporation, the cells were promptly transferred to a 15 mL conical tube containing 500 µL of recovery medium (BHI supplemented with V2 salts and 680 mM sucrose). To facilitate recovery, the cells were then incubated at 37°C for 2 hours. At the end of this step, the recovered cells were plated on pre-warmed LB agar supplemented with the appropriate antibiotic and incubated overnight at 37°C.

III.2.7. Construction of plasmid pMjsHSP

The construction of the plasmid pMjsHSP was performed as described next, using the Gibson assembly as molecular cloning strategy.

First, the fragment corresponding to the backbone of plasmid pDHFR-His was amplified by PCR using primers pDHFR-His_Backbone_F and pDHFR-His_Backbone_R (**Table III.1**) and the NZYProof DNA polymerase (NZYtech; Lisbon, Portugal). The PCR reaction comprised 10 ng of DNA template, 0.4 mM of each primer (pDHFR-His_Backbone_F and pDHFR-His_Backbone_R), 0.2 mM of dNTPs, 2.5 µL of 10x reaction buffer, 5 µL of 5x stabilizer solution and 0.0625 U of NZYProof DNA polymerase, with the remaining volume of PCR-grade water up to the total reaction volume (25 µL). This PCR reaction was performed according to the following program: initial denaturation at 95°C for 3 minutes; 30 cycles of (i) 95°C for 30 seconds (denaturation), (ii) 52°C for 30 seconds (annealing), (iii) 72°C for 135 seconds (extension); final extension at 72°C for 2 minutes. The resulting fragments of this PCR reaction, prepared in duplicate, were run on a 1% w/v agarose gel, followed by the extraction of the band appearing at 2247 bp. The gel slice was processed next with the NZYGelpure kit (NZYtech; Lisbon, Portugal) following the manufacturer's instructions for agarose gel extraction and purification. Concentration of purified DNA fragments was determined with a NanoDrop One Microvolume UV-Vis Spectrophotometer (Thermo Fisher Scientific; Waltham, MA, USA).

Afterwards, the fragment corresponding to the complete gene encoding MjsHSP was amplified from the plasmid pUC57-MjsHSP by PCR using primers pUC57-MjsHSP_Gene_F and pUC57-MjsHSP_Gene_R (**Table III.1**). This second PCR reaction was set up with the same enzyme and PCR reaction mix as described above for the backbone amplification. The PCR reaction was carried out according to the following program: initial denaturation step at 95°C for 3 minutes; 30 cycles of (i) 95°C for 30 seconds (denaturation), (ii) 52°C for 30 seconds (annealing), (iii) 72°C for 29 seconds (extension); final extension step at 72°C for 2 minutes. The resulting fragments of this second PCR reaction, prepared in duplicate, were purified directly from the reaction mixtures with the NZYGelpure kit (NZYtech; Lisbon, Portugal) following the manufacturer's instructions for PCR clean-up. Concentration of purified DNA fragments was also measured with a NanoDrop One Microvolume UV-Vis Spectrophotometer (Thermo Fisher Scientific; Waltham, MA, USA). At this point, purified DNA fragments were run on a 1% w/v agarose gel, allowing the respective band to be identified at 486 bp.

Reaction of Gibson assembly was performed using the NEBuilder® HiFi DNA Assembly Master Mix (New England Biolabs; Ipswich, MA, USA) according to the manufacturer's protocol. The reaction

was carried out in a total reaction volume of 20 µL using a DNA molar ratio to vector:insert of 1:2 (vector – fragment corresponding to the backbone of plasmid pDHFR-His; insert – fragment corresponding to the complete gene encoding MjsHSP), resulting in a total amount of DNA fragments within the recommended range (0.154 pmol). The reaction was incubated at 50°C for 120 minutes, being withdrawn 2 µL at 15, 30 and 60 minutes of reaction in addition to the sample collected at the end of the process to use in the transformation of chemically competent *E. coli* DH5α cells (**III.2.6. Transformation of competent cells**). The clones obtained were firstly screened by colony PCR with primers pMjsHSP_Colony PCR_F and pMjsHSP_Colony PCR_R (**Table III.1**). The colony PCR reaction was performed with the same enzyme and PCR reaction mix as described above for the backbone and MjsHSP gene amplifications (the source of DNA template was a colony of bacteria sampled). The PCR reaction was set up following the program: initial denaturation step at 95°C for 3 minutes; 30 cycles of (i) 95°C for 30 seconds (denaturation), (ii) 50°C for 30 seconds (annealing), (iii) 72°C for 21 seconds (extension); final extension step at 72°C for 2 minutes. The resulting fragments of this PCR reaction were run on a 1% w/v agarose gel to identify positive clones. These were grown overnight at 37°C and 250 rpm in LB medium supplemented with 100 µg mL^{-1} ampicillin (antibiotic resistance marker). The sequence of the pDNA extracted and purified from these positive clones corresponding to the constructed plasmid pMjsHSP was verified by restriction endonuclease digestion (restriction enzymes XbaI and XhoI – New England Biolabs; Ipswich, MA, USA) and DNA sequencing (STAB VIDA; Caparica, Portugal).

After this verification, the plasmid pMjsHSP was used to transform chemically competent *E. coli* BL21(DE3) and *V. natriegens* Vmax™ X2 cells (**III.2.6. Transformation of competent cells**), with the subsequent preparation of master and working cell banks (**III.2.2. Bacterial strains**).

Table III.1 – Oligonucleotide sequences used for the construction of plasmid pMjsHSP. Note that F and R represent forward and reverse primers, respectively.

Step	Primer	Nucleotide sequence (5' → 3')
Backbone amplification (pDHFR-His)	pDHFR-His_Backbone_F	TGAGGATCCCGGGAATTCT
	pDHFR-His_Backbone_R	ATGTATATCTCCTTCTTAAAGTTAAACAAAATTAT
MjsHSP gene amplification (pUC57-MjsHSP)	pUC57-MjsHSP_Gene_F	GTTTAACTTTAAGAAGGAGATATACATATGTTTGGCCGTGATCC
	pUC57-MjsHSP_Gene_R	TTCCCGGGATCCTCATTATTCAATGTTAATGCCTTTTTTGAT
Colony PCR for screening	pMjsHSP_Colony PCR_F	GGTTATGCTAGTTATTGCTCAGC
	pMjsHSP_Colony PCR_R	GATCACCGAATCGGAACG

III.2.8. Determination of growth curves

The determination of growth curves of *E. coli* BL21(DE3) harboring plasmid pMjsHSP or pTRAP and of *V. natriegens* Vmax™ X2 harboring plasmid pMjsHSP was performed according to the following procedure.

Frozen cell bank vials of *E. coli* BL21(DE3) (pMjsHSP or pTRAP) or *V. natriegens* Vmax™ X2 (pMjsHSP) were used to inoculate 5 mL of appropriate culture medium, specifically LB for the *E. coli* strain and LB with addition of V2 salts for the *V. natriegens* strain, supplemented with 100 μg mL^{-1} ampicillin, in a 15 mL conical centrifuge tube. Cells were incubated overnight (16-18 h) at 37°C and an agitation of 250 rpm. In the next day, a suitable volume of these pre-inoculum cultures was centrifuged (6000×g, 3 minutes, 4°C) and the resulting cell pellet was resuspended in fresh culture medium. This resuspended cell pellet was used to inoculate 100 mL non-baffled shake flasks containing 30 mL of LB (*E. coli*) or LB with addition of V2 salts (*V. natriegens*), and ampicillin supplementation (100 μg mL^{-1}), in such a way as to obtain an approximate $OD_{600\ nm}$ of 0.1. These cultures were then incubated at 37°C and 250 rpm and the $OD_{600\ nm}$ was monitored over time of cell growth. At each sample collection timepoint, additional culture volume may have been collected for use in supplementary analytics.

III.2.9. Determination of plasmid copy number by qPCR

The plasmid copy number in both *E. coli* BL21(DE3) cells harboring plasmid pMjsHSP or pTRAP and in *V. natriegens* Vmax™ X2 cells containing the plasmid pMjsHSP was determined by quantitative PCR (qPCR) according to the procedure described below.

First, two sets of primers were designed and synthetized (**Table III.2**), one specific for the catalytic α subunit of DNA polymerase III (*dnaE*) chromosomal gene, which has only a single copy in the entire genomes of both *E. coli* and *V. natriegens* strains used in this work, and the other specific for the ampicillin resistance (*ampr*) gene, which is present both in pMjsHSP and pTRAP.

Table III.2 – Oligonucleotide sequences used for qPCR amplification of the reference (*dnaE*) and target (*ampr*) genes, and the expected size of the amplified product. Note that F and R represent forward and reverse primers, respectively.

Gene	Primer	Nucleotide sequence (5' → 3')	Amplified product size (bp)
Reference (chromosome): *dnaE*	qPCR_Chromosome_F	CGTCGTGCGATGGGTAAGAA	115
	qPCR_Chromosome_R	GTCGAAGATTTTCATCGCCA	
Target (plasmids): *ampr*	qPCR_Plasmids_F	ACCCAGAAACGCTGGTGAAAG	84
	qPCR_Plasmids_R	GCTGTTGAGATCCAGTTCGA	

Afterwards, the genomic DNA (gDNA) from *E. coli* BL21(DE3) and *V. natriegens* Vmax™ X2 was extracted and purified as described in section **III.2.14.4. gDNA extraction**. Additionally, using the protocol detailed in section **III.2.14.5. pDNA extraction**, plasmid pMjsHSP was obtained and purified from *E. coli* BL21(DE3) and *V. natriegens* Vmax™ X2, and plasmid pTRAP was isolated and purified from *E. coli* BL21(DE3). Then, gDNA and plasmid DNA (pDNA) were serial diluted in PCR-grade water and used as standards in the qPCR reactions (10000, 1000, 100, 10 and 1 pg DNA in each reaction).

Samples of *E. coli* BL21(DE3) with pMjsHSP or pTRAP and of *V. natriegens* Vmax™ X2 with pMjsHSP were collected during a standard bacterial growth as described in section **III.2.8. Determination of growth curves**. The samples were collected at several timepoints throughout the bacterial growth, starting from the initial inoculation ($OD_{600\,nm} \approx 0.1$, 0.3, 0.5, 0.7, 0.9, 1.1 and 1.3). These samples were stored at -20°C until use. At each collection timepoint, an assay was performed to determine the concentration of bacterial cells in colony forming units (CFU) mL^{-1}. This methodology was applied to plasmid-free *E. coli* BL21(DE3) and *V. natriegens* Vmax™ X2 cells. For the qPCR reactions, all samples were diluted in PCR-grade water to add around 1000 cells to each well of a 96-well plate where reactions were prepared. Additionally, it is important to note that to ensure that the samples and the pDNA standards had the same inhibitors present, the qPCR reactions for the pDNA standard curve were spiked with the same number of the corresponding bacterial cells without plasmid (≈ 1000). This procedure was not applied to the gDNA standards.

qPCR reactions were performed in real time using the QuantStudio 5 Real-Time PCR System (Thermo Fisher Scientific; Waltham, MA, USA) and the NZYSpeedy qPCR Green Master Mix (ROX) (2x) (NZYtech; Lisbon, Portugal). Each qPCR reaction (20 µL) comprised 10 µL of 2x NZYSpeedy qPCR Green Master Mix (ROX), 0.8 µL of each primer (qPCR_Chromosome_F and qPCR_Chromosome_R, or qPCR_Plasmids_F and qPCR_Plasmids_R) to a final concentration of 400 nM and 7.4 µL of PCR-grade water, with the remaining volume comprising gDNA standards (1 µL), or pDNA standards (1 µL of pDNA plus 1 µL with 1000 spiking cells of *E. coli* or *V. natriegens*, with the use of 6.4 µL of PCR-grade water instead of 7.4 µL), or samples (1 µL with 1000 *E. coli* or *V. natriegens* cells containing the respective plasmid pMjsHSP or pTRAP). In all cycles of qPCR reactions, negative controls without addition of cells and gDNA or pDNA were included. The qPCR reactions were performed according to the program: initial denaturation at 95°C for 3 minutes; 40 cycles of (i) 95°C for 5 seconds, (ii) 60°C for 20 seconds (polymerization); final extension at 70°C for 30 seconds. At the end of each extension step, the fluorescence signal was automatically detected and quantified by the QuantStudio 5 Real-Time PCR System. Melting curve analyses, through a temperature gradient of 0.05°C s^{-1} from 70°C to 95°C, were performed to confirm the amplification of only the intended specific targets. At the end of the process, the samples were cooled to 40°C for 30 seconds.

The determination of plasmid copy number using the real time qPCR data was based on a relative quantification method previously developed and described in the literature (205). These data were acquired automatically by the QuantStudio 5 Real-Time PCR System software and consisted of the threshold cycle (C_t).

The standard curves for gDNA from *E. coli* BL21(DE3) and *V. natriegens* Vmax™ X2 and for pMjsHSP from *E. coli* BL21(DE3) and *V. natriegens* Vmax™ X2 and pTRAP from *E. coli* BL21(DE3)

were obtained by plotting the logarithm of each concentration against the C_t values (**Supplementary Figure III.S1**, **Supplementary Figure III.S2**, **Supplementary Figure III.S3** and **Supplementary Figure III.S4**). Considering the estimated equation of each linear regression, the slope of the respective standard curve was used to determine the amplification efficiency (E) according to **Equation III.1**.

$$E = 10^{-\frac{1}{slope}} \quad \textbf{III.1}$$

Given the amplification efficiencies of the gDNA (E_{gDNA}) and pDNA (E_{pDNA}) standards, along with the C_t values obtained from each sample amplified separately with the gDNA-specific ($C_{t_{gDNA}}$) and plasmid-specific ($C_{t_{pDNA}}$) primers sets, the plasmid copy number for each sample can be determined using **Equation III.2**.

$$Plasmid\ copy\ number = \frac{{E_{gDNA}}^{C_{t_{gDNA}}}}{{E_{pDNA}}^{C_{t_{pDNA}}}} \quad \textbf{III.2}$$

Regarding the qPCR validation, the specificity of qPCR amplification was confirmed by analyzing the melting temperature (T_m) of each amplified product and by agarose gel electrophoresis. The agarose gel revealed amplified products of expected sizes for each set of primers, validating the specificity of all qPCR reactions. Furthermore, negative controls in all qPCR reactions yielded C_t values greater than 35, significantly higher than samples, indicating negligible background amplification. To validate the experimental determination, in addition to the coefficient of determination (R^2), the percentage of amplification efficiency ($E\ (\%)$) of each standard curve was calculated using **Equation III.3**.

$$E\ (\%) = (10^{-\frac{1}{slope}} - 1) \times 100 \quad \textbf{III.3}$$

III.2.10. Production of MjsHSP nanocages

III.2.10.1. *E. coli*

The experimental methodology defined as reference to produce MjsHSP nanocages using *E. coli* BL21(DE3) cells harboring the plasmid pMjsHSP is described below.

From a frozen cell bank, an overnight culture (16-18 h) was prepared in a 15 mL conical centrifuge tube with 5 mL of LB supplemented with 100 μg mL^{-1} ampicillin. The incubation conditions were 37°C and an agitation of 250 rpm. In the next day, a suitable volume of pre-inoculum culture was centrifuged (6000×g, 3 minutes, 4°C) and the resulting cell pellet was resuspended in fresh culture medium. This resuspended cell pellet was inoculated in a 100 mL non-baffled shake flask, containing 30 mL of LB and ampicillin supplementation (100 μg mL^{-1}), to obtain an approximate $OD_{600\ nm}$ of 0.1. Re-inoculated culture was then incubated at 37°C and 250 rpm and the $OD_{600\ nm}$ was monitored until it reached approximately

a value in the range 0.5-0.6. At this timepoint, the expression of MjsHSP nanocages was induced with 1 mM IPTG. After 4 hours of induction at 37°C and 250 rpm, the bacterial cells were harvested by centrifugation at 6000×g for 20 minutes at 4°C. Following this centrifugation, the cell pellet was washed three times and resuspended in the washing buffer (50 mM Tris-HCl, 100 mM NaCl, 1 mM EDTA, pH 8.0), according to a ratio of 10 mL washing buffer per 1 g wet cell pellet. Then, the resuspended cells were disrupted on ice by sonication using the sonicator Bandelin Sonopuls HD 3200 (Berlin, Germany) and the following operating conditions: MS-72 probe (2 mm diameter), 30% amplitude, 5 cycles comprising 30 seconds ON and 30 seconds OFF. The resulting cell lysate was centrifuged (12000×g, 40 minutes, 4°C) and the supernatant with the produced MjsHSP nanocages was carefully collected. These self-assembled MjsHSP nanocages were stored at 4°C.

III.2.10.2. *V. natriegens*

The benchmark production of MjsHSP nanocages using *V. natriegens* Vmax™ X2 cells containing the plasmid pMjsHSP was performed using the same experimental procedure established for *E. coli* BL21(DE3) cells (**III.2.10. Production of MjsHSP nanocages**), with the exception that LB with addition of V2 salts was used instead of only LB (supplementation with the antibiotic ampicillin at 100 μg mL^{-1} was maintained).

III.2.11. Production of TRAP O-rings

III.2.11.1. *E. coli*

The standardized experimental protocol for producing TRAP O-rings as precursors of TRAP nanocages using *E. coli* BL21(DE3) cells transformed with the plasmid pTRAP was similar to the one implemented for the expression of MjsHSP nanocages in *E. coli* (**III.2.10. Production of MjsHSP nanocages**).

III.2.12. Optimization of production of MjsHSP nanocages

III.2.12.1. *E. coli*

The optimization of the production of MjsHSP nanocages in *E. coli* BL21(DE3) cells used the established standard procedure (**III.2.10. Production of MjsHSP nanocages**) as the starting point.

Considering the information available in the literature regarding the expression of NVPNs and their optimization in different bacterial hosts, five key parameters were selected and analyzed to determine their influence on the final product, enabling the identification of the most favorable settings to produce MjsHSP nanocages. These five parameters were: growth medium composition, concentration of protein expression inducer, temperature of protein expression, timepoint of induction and aeration conditions.

The strategy used consisted of a systematic and step-by-step optimization of protein expression, in which only one parameter was varied, maintaining the others unchanged in relation to the reference protocol. After identifying the best condition in terms of production yield, this methodology was applied to the next parameter. Different operational conditions were tested for each parameter as described below. Similarly, the collection, washing and lysis of the cells, as well as the clarification step was also performed in the same way as in the benchmark procedure.

Regarding the growth medium composition, the following culture media were tested: LB supplemented with V2 salts, LB3 (LB supplemented with 20 g L^{-1} NaCl), 2xYT (10 g L^{-1} yeast extract, 16 g L^{-1} tryptone, 5 g L^{-1} NaCl), enhanced 2xYT (20 g L^{-1} yeast extract, 32 g L^{-1} tryptone, 17 g L^{-1} NaCl, 2 g L^{-1} glucose, 17.6 mM Na_2HPO_4), TB (24 g L^{-1} yeast extract, 12 g L^{-1} tryptone, 4 mL L^{-1} glycerol, 16.9 mM KH_2PO_4, 71.8 mM K_2HPO_4), TB supplemented with V2 salts, and M9 minimal (M9 salts solution: 47.6 mM Na_2HPO_4, 22.0 mM KH_2PO_4, 8.6 mM NaCl and 18.7 mM NH_4Cl, 0.1 mM $CaCl_2$, 2.0 mM $MgSO_4$) supplemented with 4 g L^{-1} glucose. It should be noted that these culture media were chosen after checking for bacterial growth in an overnight growth assay of *E. coli* BL21(DE3) cells harboring the plasmid pMjsHSP (the same procedure as above indicated for an overnight culture). The concentrations of the protein expression inducer evaluated were: 0.1 mM, 0.5 mM, and 1.5 mM of IPTG. In terms of the temperature of protein expression, in order to determine the optimal folding and self-assembly conditions, the following options were evaluated: 15°C, 20°C, and 30°C. To examine the influence of cell growth phase, two time-points of induction were analyzed: $OD_{600\ nm} \approx 0.2$, and $OD_{600\ nm} \approx 1.8$. Finally, *E. coli* BL21(DE3) cells with the plasmid pMjsHSP were grown in 250 mL non-baffled and baffled shake flasks maintaining a constant ratio of culture volume to air volume (75 mL of culture medium) to evaluate oxygen availability and mixing efficiency.

After completion of the optimization study, the production of MjsHSP nanocages under optimized conditions was performed again, with the monitoring and collection of samples of the bacterial culture after 2, 4, 6, 8 and 24 hours of protein expression induction.

III.2.12.2. *V. natriegens*

The optimized production of MjsHSP nanocages using *V. natriegens* Vmax™ X2 cells was performed with the same experimental methodology of optimization that was used for *E. coli* BL21(DE3) cells (**III.2.12. Optimization of production of MjsHSP nanocages**). The only exceptions were the evaluated culture media (LB3, 2xYT, enhanced 2xYT, TB supplemented with V2 salts, M9 minimal supplemented with V2 salts and 4 g L^{-1} glucose, and M9 minimal supplemented with V2 salts and 4 g L^{-1} sucrose) and the two timepoints of induction analyzed ($OD_{600\ nm} \approx 0.2$, and $OD_{600\ nm} \approx 1.4$).

III.2.13. Optimization of production of TRAP O-rings

III.2.13.1. *E. coli*

The optimization of the production of TRAP O-rings as precursors of TRAP nanocages the reference procedure (**III.2.11. Production of TRAP O-rings**) and the most favorable settings verified for the optimization of the production of MjsHSP nanocages (**III.2.12. Optimization of production of MjsHSP nanocages**) as starting point. The strategy consisted of performing the production of TRAP O-rings with the best operating conditions in the five key parameters selected (growth medium composition, concentration of protein expression inducer, temperature of protein expression, timepoint of induction and aeration conditions), evaluating the results and understanding their influence on the production yield. The monitoring and collection of samples of the bacterial culture after 2, 4, 6, 8 and 24 hours from the beginning of the induction of protein expression was also carried out in the final production experiment.

III.2.14. Analytics

III.2.14.1. Total protein quantification

Total protein concentration was determined throughout the work by the bicinchoninic acid (BCA) method, using the commercial Pierce™ BCA Protein Assay Kit (Thermo Fisher Scientific; Waltham, MA, USA). Quantification was performed according to the manufacturer's instructions for the microplate procedure. Transparent 96-well plates were used and, when necessary, the samples were diluted in PBS solution (1x). The absorbance was measured at 562 nm on Spectra MAX (Molecular Devices; San Jose, CA, USA) or Multiskan Go (Thermo Fisher Scientific; Waltham, MA, USA) plate reader. Bovine serum albumin (BSA) was used as a standard curve with a concentration from 2 to 0.025 mg mL^{-1}.

III.2.14.2. SDS-PAGE

Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) was performed to evaluate the composition of the samples both in specific protein nanocage (monomers) and total protein. Samples were diluted with Laemmli sample buffer (4x) (Bio-Rad; Hercules, CA, USA) or SDS-PAGE sample loading buffer (5x) (NZYtech; Lisbon, Portugal), and then denatured in reducing conditions with 100 mM dithiothreitol (DTT) at 95°C for 10 minutes (MjsHSP nanocages) or with 200 mM DTT and 10 M urea at 100°C for 20 minutes (TRAP O-rings). Afterwards, samples were applied in a 12% or 15% acrylamide gel (MjsHSP nanocages or TRAP O-rings, respectively), prepared from an acrylamide/bis-acrylamide solution (40%) (Bio-Rad; Hercules, CA, USA). The molecular weight marker used was precision plus protein dual color standards (Bio-Rad; Hercules, CA, USA) or NZYColour protein marker II (NZYtech; Lisbon, Portugal). The samples were run at 150 V (the first 30 minutes between 80 and 100 V) using a running buffer composed of 192 mM glycine, 25 mM Tris-HCl and 0.1% w/v SDS, pH 8.3. Gels were stained with PhastGel™ Blue R (Coomassie R 350 stain) (GE Healthcare; Chicago, IL,

USA), followed by silver nitrate staining, when necessary. Gels were scanned with GS-800 Calibrated Imaging Densitometer (Bio-Rad; Hercules, CA, USA) and then subject to image processing using Quantity One software (Bio-Rad; Hercules, CA, USA).

III.2.14.3. Protein nanocages quantification by densitometric analysis of SDS-PAGE

For the quantification of MjsHSP nanocages and TRAP O-rings, and specifically of the respective monomers, images of SDS-PAGE were analyzed using the ImageJ software (206). The intensity of the band corresponding to the protein nanocages monomer (I_{PNM}) and the intensity of all protein bands present in the sample under study (I_{TP}) was determined with this analysis. The amount of protein nanocages monomers in relation to the total protein ($Quantity_{PNM}$) in each sample was then determined from the these band intensities using **Equation III.4**.

$$Quantity_{PNM}\ (\%) = \frac{I_{PNM}}{I_{TP}} \times 100 \qquad \textbf{III.4}$$

III.2.14.4. gDNA extraction

Genomic DNA from *E. coli* BL21(DE3) and *V. natriegens* Vmax™ X2 was extracted and purified using the commercial Wizard® Genomic DNA Purification Kit (Promega; Madison, WI, USA), according to the manufacturer's instructions. Rehydration of gDNA was performed with PCR-grade water in overnight incubation (16-18 h) at 4°C. gDNA concentration was determined with a NanoDrop One Microvolume UV-Vis Spectrophotometer (Thermo Fisher Scientific; Waltham, MA, USA). Extracted and purified gDNA was stored at 4°C until use.

III.2.14.5. pDNA extraction

Plasmid DNA from *E. coli* DH5α, *E. coli* BL21(DE3) and *V. natriegens* Vmax™ X2 was extracted and purified using the commercial High Pure Plasmid Isolation Kit (Roche; Basel, Switzerland), according to the manufacturer's instructions. Cell cultures were prepared with an overnight incubation (16-18 h) under conditions appropriate to each species and respective plasmid. Elution of pDNA was performed with pre-warmed PCR-grade water in two successive centrifugations, to increase the recovery efficiency. pDNA concentration was determined with a NanoDrop One Microvolume UV-Vis Spectrophotometer (Thermo Fisher Scientific; Waltham, MA, USA). Extracted and purified pDNA was stored at 4°C (short period until used) or -20°C (long period until needed).

III.2.14.6. Agarose gel electrophoresis

Agarose gel electrophoresis was performed in 1% or 2% w/v agarose gels, which were prepared in TAE buffer (40 mM Tris base, 20 mM acetic acid, 1 mM EDTA, pH 8.0). Samples were mixed with NZYDNA loading dye (6x) (NZYtech; Lisbon, Portugal). Gels were run in TAE buffer under the following

conditions: 100 V for 60 minutes (small gels) or 120 V for 90 minutes (large gels). The molecular weight marker used was NZYDNA ladder III (NZYtech; Lisbon, Portugal). Gels were stained with an ethidium bromide solution (0.4 μg mL^{-1}) or with the SYBR™ Safe DNA Gel Stain (1x) (Thermo Fisher Scientific; Waltham, MA, USA). Gel images were acquired and analyzed with an Eagle Eye II (Stratagene; La Jolla, CA, USA) or an Axygen (Corning, NY, USA) gel documentation system.

III.2.15. Data treatment

III.2.15.1. Protein nanocages production

The growth of bacterial cultures can be characterized by two main parameters, specific growth rate and generation time, whose determination methodology is well described in the literature (207,208). These two parameters were calculated for all the different experimental and/or operational conditions selected and evaluated, considering the respective exponential growth phase.

Additionally, to compare the results obtained in terms of the production of protein nanocages (MjsHSP nanocages and TRAP O-rings) for these experimental tests, three specific indicators were considered: production yield, volumetric productivity and specific productivity.

The production yield (in mg L^{-1}), volumetric productivity (in mg L^{-1} h^{-1}) and specific productivity (in mg $OD_{600\ nm}^{-1}$ h^{-1}) were, respectively, computed from **Equation III.5**, **Equation III.6** and **Equation III.7**.

$$Production\ yield\ (mg\ L^{-1}) = \frac{Mass_{protein\ nanocages}}{Volume_{cell\ culture}} \quad \textbf{III.5}$$

$$Volumetric\ productivity\ (mg\ L^{-1}\ h^{-1}) = \frac{Mass_{protein\ nanocages}}{Volume_{cell\ culture} \times Time_{production\ of\ protein\ nanocages}} \quad \textbf{III.6}$$

$$Specific\ productivity\ (mg\ {OD_{600\ nm}}^{-1}\ h^{-1}) = \frac{Mass_{protein\ nanocages}}{OD_{600\ nm_{cell\ culture}} \times Time_{production\ of\ protein\ nanocages}} \quad \textbf{III.7}$$

Note that the mass of protein nanocages ($Mass_{protein\ nanocages}$) was estimated considering the results obtained in the SDS-PAGE densitometric analysis (**III.2.14.3. Protein nanocages quantification by densitometric analysis of SDS-PAGE**).

The evaluation of the different production conditions tested based on the production yield and specific productivity, allowed to highlight the parameters that have the greatest influence on the upstream processing of both NVPNs.

III.2.16. Statistical analysis

Descriptive statistics, including mean and standard deviation, were calculated for the results obtained across the different variables. The sample sizes for each variable typically ranged between 2 and 3 experiments.

III.3. Results and Discussion

III.3.1. Construction of plasmid pMjsHSP

Plasmid pMjsHSP was constructed by cloning the complete gene encoding MjsHSP into the backbone of plasmid pDHFR-His. First, the backbone of plasmid pDHFR-His was amplified by PCR (**Figure III.1A**). The complete gene that encodes the MjsHSP was amplified from the plasmid pUC57-MjsHSP by PCR using specific primers containing homology to the pDHFR-His backbone (**Figure III.1B**). This insert fragment (complete gene encoding MjsHSP) was cloned into the vector fragment (backbone of plasmid pDHFR-His) through Gibson assembly resulting in the plasmid pMjsHSP (**Figure III.1C**).

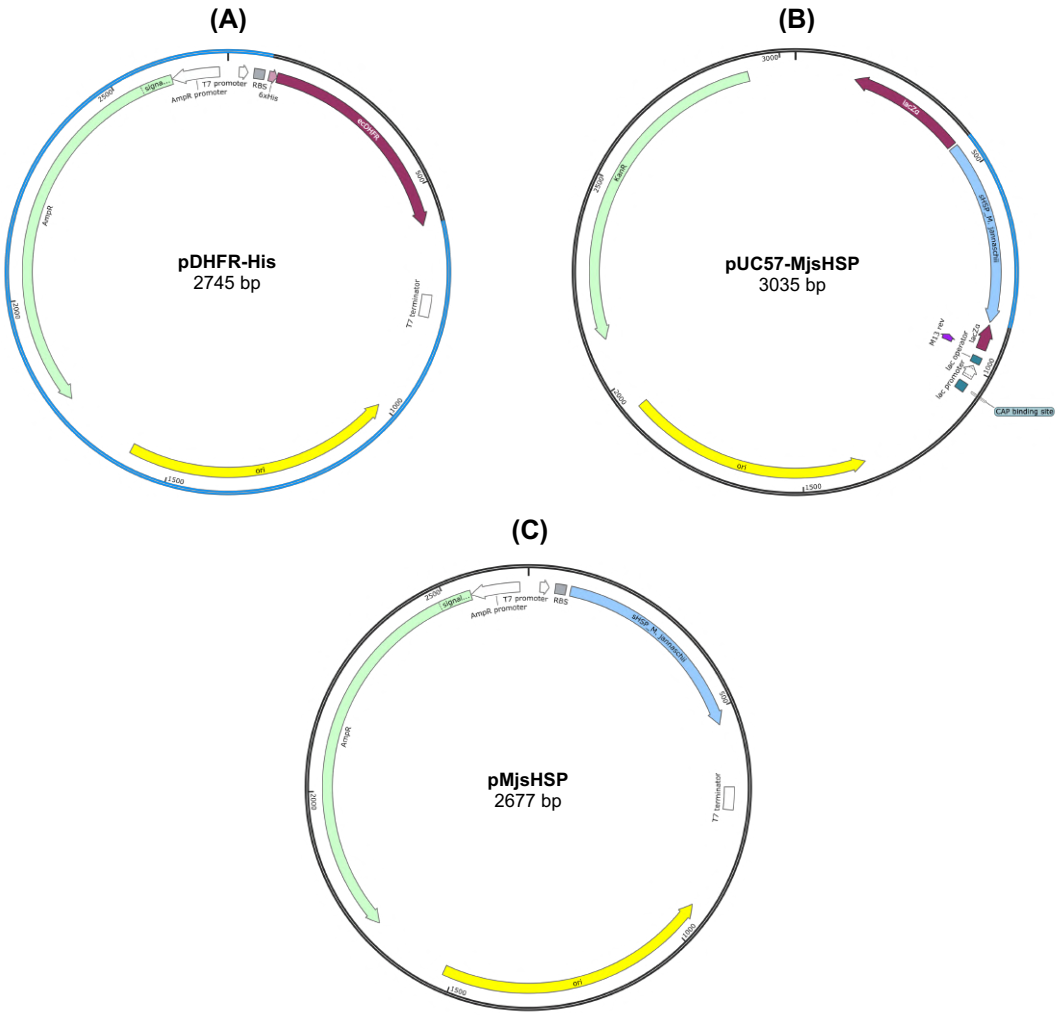


Figure III.1 – Construction of the plasmid pMjsHSP through a molecular cloning strategy based on Gibson assembly. **(A)** Plasmid pDHFR-His with the sequence of the backbone fragment amplified by PCR and used as the vector for cloning colored in blue. **(B)** Plasmid pUC57-MjsHSP, with the sequence of the complete gene encoding MjsHSP amplified by PCR and used as the insert for cloning colored in blue. **(C)** Plasmid pMjsHSP resulting from a Gibson assembly reaction using the fragments (vector and insert) indicated in **(A)** and **(B)**. The plasmid maps presented were created using the SnapGene 3.2.1. software.

The different steps of the cloning procedure were analyzed by agarose gel electrophoresis, with the results obtained shown in **Figure III.2**. The sequence of the final plasmid pMjsHSP was also checked by single and double digestion with restriction enzymes (**Figure III.2C**) and by Sanger sequencing.

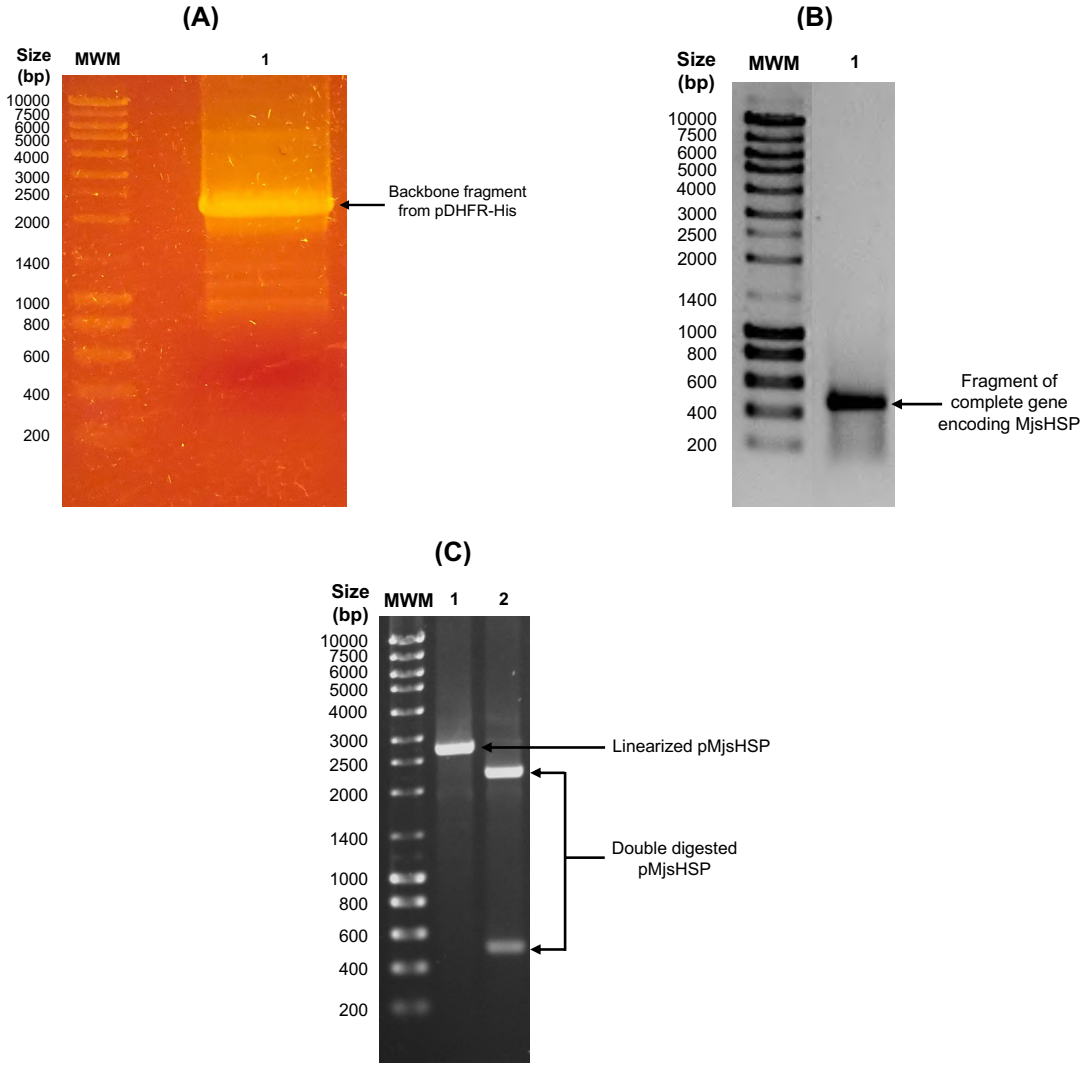


Figure III.2 – Agarose gel analysis obtained during the construction of plasmid pMjsHSP. **(A)** Amplification of the fragment corresponding to the backbone of plasmid pDHFR-His by PCR (lane 1, band size of 2247 bp). **(B)** Amplification of the fragment corresponding to the complete gene encoding MjsHSP by PCR (lane 1, band size of 486 bp). **(C)** Plasmid MjsHSP resulting from the cloning by Gibson assembly of the two fragments mentioned in **(A)** and **(B)**, linearized with the restriction enzyme XhoI (lane 1, band size of 2677 bp) and double digested with the restriction enzymes XbaI and XhoI (lane 2, band sizes of 2188 bp and 489 bp). The abbreviation MWM refers to the molecular weight marker (NZYDNA ladder III; NZYtech Lisbon, Portugal).

Briefly, plasmid pMjsHSP integrates the hybrid origin of replication (ColE1/pMB1/pBR322/pUC), the kanamycin resistance gene, T7 the bacteriophage transcriptional and translational regulatory elements, and the MjsHSP gene.

With the construction of pMjsHSP completed and validated, chemically competent *E. coli* DH5α cells (**III.2.4. Preparation of chemically competent cells**) were transformed by heat shock with this

plasmid as described in section **III.2.6. Transformation of competent cells**. After confirming the transformation, *E. coli* DH5α cells harboring the pMjsHSP were cultured at shake flask scale and master and working cell banks (**III.2.2. Bacterial strains**) were prepared to ensure maintenance, consistency, stability, and reliability during subsequent experiments.

III.3.2. Growth profile of *E. coli* and *V. natriegens* harboring pMjsHSP

Plasmid pMjsHSP was then transformed by heat shock into chemically competent *E. coli* BL21(DE3) and *V. natriegens* Vmax™ X2 cells (**III.2.4. Preparation of chemically competent cells**) by the heat shock strategy (**III.2.6. Transformation of competent cells**). Master and working cell banks of the transformed strains were also prepared (**III.2.2. Bacterial strains**).

With the purpose of obtaining and analyzing the growth profiles of *E. coli* BL21(DE3) and *V. natriegens* Vmax™ X2 cells harboring the plasmid pMjsHSP, in particular the exponential growth phase, bacterial cultures with both strains were performed on a shake flask scale according to the operating conditions described in the section **III.2.8. Determination of growth curves**.

The growth curves obtained for these specific strains of *E. coli* and *V. natriegens* containing the pMjsHSP are shown in **Figure III.3**.

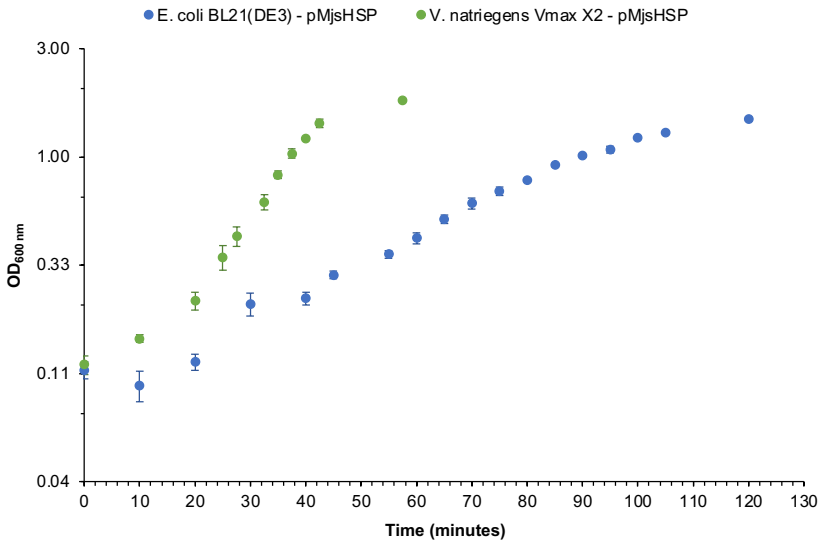


Figure III.3 – Growth curves obtained for *E. coli* BL21(DE3) (●) and *V. natriegens* Vmax™ X2 (●) cells harboring the plasmid pMjsHSP. Bacterial cultures were performed at a shake flask scale using LB (*E. coli*) and LB supplemented with V2 salts (*V. natriegens*), as detailed in section **III.2.8. Determination of growth curves**. Each data point represents the mean of three biological replicates, with error bars indicating the respective standard deviation. The y-axis is on a logarithmic scale.

A lag phase in the initial 20 minutes was observed for *E. coli* BL21(DE3) cells, where cells are in a process of adaptation to the new culture environment after inoculation, with the intracellular resources not being directed towards cell replication. An exponential growth phase was then observed between

approximately 20 and 95 minutes. The specific growth rate was approximately 1.66 ± 0.02 h^{-1} and the generation time around 25.1 ± 0.2 min. Interestingly, the results obtained for these two parameters are slightly higher than those reported in the literature for the BL21(DE3) strain of *E. coli* harboring a plasmid (0.10 to 0.60 h^{-1}) (209–211). Performing these assays under experimental conditions close to the ideal environment for the growth of this bacterial strain, from the culture medium to the temperature and the aeration conditions, among others, could have positively influenced these results. Furthermore, the pMjsHSP plasmid could have contributed with a minimal metabolic burden, which also reduced its impact on the growth of bacterial cells, contrary to what is generally associated with plasmid-harboring bacterial strains. Typically, in the latter there is an increase in the metabolic load for the cells and consequently a delay in their growth profile (212). From this growth curve it was also possible to understand that the deceleration phase and subsequent stationary phase began after 100 minutes of bacterial growth. The maximum cell density was reached after 120 minutes, with a value of 1.47 ± 0.08 being obtained for the $OD_{600\ nm}$.

V. natriegens Vmax™ X2 cells also display lag phase in the first 20 minutes, although this is less prominent than in the case of *E. coli* BL21(DE3). Then, the exponential growth phase occurs between 20 and 42.5 minutes, with specific growth rate and a doubling time of approximately 4.51 ± 0.34 h^{-1} and 9.3 ± 0.7 min, respectively. These results are in agreement with the data available in the literature for the growth of *V. natriegens* Vmax™ X2 cells in optimal conditions (199,209,211,213). The maximum cell density was achieved after 57.5 minutes, with an $OD_{600\ nm}$ of about 1.77 ± 0.02.

A comparison of the growth profile of *E. coli* BL21(DE3) and *V. natriegens* Vmax™ X2 cells harboring the plasmid pMjsHSP, shows that *V. natriegens* presents a much faster growth, which is reflected in a 2.7 times higher specific growth rate, in a 62.8% shorter generation time and in obtaining an equivalent maximum cell density in about half the time. Therefore, this data confirms that *V. natriegens* Vmax™ X2 is an interesting alternative for industrial applications of protein production where fast growth rates and high biomass production are crucial (196,200,203,214,215).

III.3.3. Determination of plasmid copy number for pMjsHSP in *E. coli* and *V. natriegens*

The determination of the pMjsHSP plasmid copy number in *E. coli* BL21(DE3) and *V. natriegens* Vmax™ X2 cells aimed to compare its behavior in both strains, analyze its variation over time and obtain data for correlation with the expression of MjsHSP nanocages.

Standard curves for gDNA from the *E. coli* and *V. natriegens* strains, and for pMjsHSP extracted from cells of the respective strains were determined as described in section **III.2.9. Determination of plasmid copy number by qPCR**. These standard curves are available in **Supplementary Figure III.S1** and **Supplementary Figure III.S2**.

Through the respective linear regressions of these four standard curves and using **Equation III.3**, it was possible to determine their percentage of amplification efficiency. Efficiency values between 86.2% and 100.6% were obtained, which are very close to or even within the range recommended in

the literature (between 90 and 110%) (216). This means that the DNA samples do not contain impurities and contaminants in sufficient quantities to inhibit and/or influence the qPCR reactions. Furthermore, the coefficient of determination was greater than 0.985 in the four standard curves, which also indicates a high precision in the qPCR performance (216).

The plasmid copy number was calculated using **Equation III.2** by considering the C_t values of each sample and the amplification efficiencies determined from the gDNA and pDNA standard curves. The accuracy of the estimated plasmid copy number is critically dependent on the similar amplification efficiency of the reference gene (*dnaE* in genome of *E. coli* and *V. natriegens*) and the target gene (*ampr* in pMjsHSP). When this condition is not met, traditional standard curve methods can lead to significant overestimation of plasmid copy number values, as described in the literature (205,217).

The results obtained for the plasmid copy number of both *E. coli* BL21(DE3) and *V. natriegens* Vmax™ X2 cells harboring the pMjsHSP at different stages of the growth profile (corresponding to an $OD_{600\ nm}$ of 0.1, 0.3, 0.5, 0.7, 0.9, 1.1 and 1.3) are represented in **Figure III.4**.

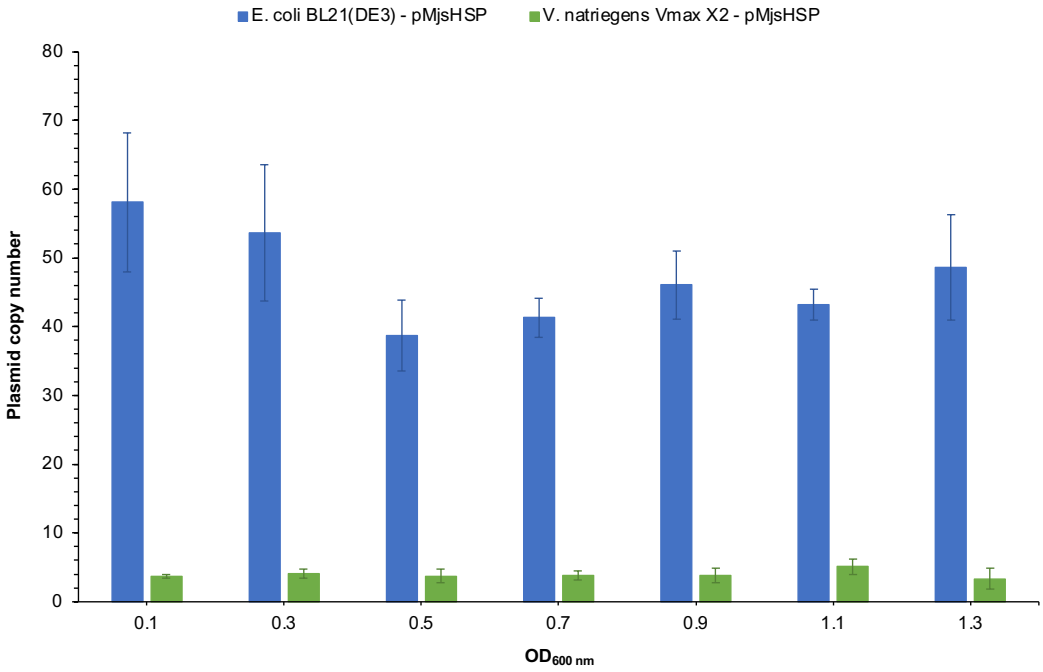


Figure III.4 – Plasmid copy number values obtained for *E. coli* BL21(DE3) (blue bars) and *V. natriegens* Vmax™ X2 (green bars) cells harboring the plasmid pMjsHSP at different growth times (corresponding to an $OD_{600\ nm}$ of 0.1, 0.3, 0.5, 0.7, 0.9, 1.1 and 1.3). These results were determined through a qPCR analysis, following the experimental procedure described in the section **III.2.9. Determination of plasmid copy number by qPCR**. Each data point represents the mean of three biological replicates, with error bars indicating the respective standard deviation.

Regarding *E. coli* BL21(DE3) cells containing the plasmid pMjsHSP, the analysis of **Figure III.4** allowed to verify an oscillation in the plasmid copy number throughout the growth period monitored. At the onset of growth ($OD_{600\ nm}$ of 0.1), a plasmid copy number of about 58 ± 10 was obtained, followed by a slight decrease to about 39 ± 5 in the middle of the exponential growth phase ($OD_{600\ nm}$ of 0.5). From this moment until the beginning of the stationary growth phase ($OD_{600\ nm}$ of 1.1 and 1.3), plasmid

copy numbers increased, reaching 49 ± 8. Considering the entire growth profile, the average plasmid copy number was 47 ± 6. This figure is not consistent with the fact that pMjsHSP has a hybrid ColE1/pMB1/pBR322/pUC origin of replication, which is associated to high plasmid copy number in *E. coli* (≈ 150 to 200 copies). Interestingly, in a comparative study of *E. coli* NEB10β versus *V. natriegens* ATCC 14048, Tschirhart *et al.* obtained a higher plasmid copy number in *E. coli* for a plasmid with this origin of replication (149 copies) (204). A possible explanation may be associated with the host cell physiology, since it is described in the literature that the plasmid copy number is closely related to the bacterial replicative machinery, which is influenced by the host cell's physiological state (218). On the other hand, the results obtained in this sub-section agree with the high growth rate shown by *E. coli* cells harboring pMjsHSP that was discussed above, which suggested that plasmid replication was not significantly impacting the metabolic load of the cells.

Regarding the results obtained with *V. natriegens* Vmax™ X2 cells harboring the plasmid pMjsHSP, the plasmid copy number remained constant throughout the cell growth, counting 4 ± 0 plasmid copies for the initial $OD_{600\ nm}$ of 0.1 and 3 ± 2 plasmid copies for the $OD_{600\ nm}$ of 1.3. The average plasmid copy number was 4 ± 1, which is considerably lower than the value obtained for *E. coli* cells (approximately 12-fold difference). However, the study by Tschirhart *et al.* mentioned above obtained similar results when comparing *E. coli* and *V. natriegens* (6-fold difference) although with a higher plasmid copy number value for *V. natriegens*, ATCC 14048 strain (25 copies). Similarly, host cell physiology may have influenced the results of this work, as the Vmax™ X2 strain was used instead of the wild type ATCC 14048 strain reported by Tschirhart *et al.*. The fact that the plasmid pMjsHSP used to transform *V. natriegens* cells was produced and extracted from *E. coli* may also be of particular relevance since the source organism of the plasmid influences its replication in a different bacterial host, due to modifications in the plasmid (for example, methylations). Generally, and whenever possible, the use of a plasmid produced and extracted from the same host organism is recommended (219).

Additionally, in both *E. coli* and *V. natriegens* cells, as the stationary growth phase approaches and the metabolism shifts towards an increased investment in the plasmid replication, an increase in plasmid copy number at $OD_{600\ nm}$ of 1.1 and 1.3 would be expected; however, this was not observed.

These results can impact protein expression, although there are reported studies in which a reduced plasmid copy number is not necessarily associated with a decrease in protein expression, since it avoids an excessive metabolic overload for the bacterial cells (220).

III.3.4. Production of MjsHSP nanocages in *E. coli* and *V. natriegens*

With the objective of assessing the suitability of *E. coli* BL21(DE3) and *V. natriegens* Vmax™ X2 as production hosts for MjsHSP nanocages, first their growth characteristics under the standard conditions for cultivation and protein expression (non-baffled shake flask with 30% volume of culture medium; LB or LB supplemented with V2 salts, respectively; 37°C and 250 rpm; induction of protein expression at $OD_{600\ nm}$ of 0.5-0.6, with 1 mM IPTG, for 4 hours at 37°C) were analyzed. The growth curves obtained for both strains are shown in **Figure III.5**.

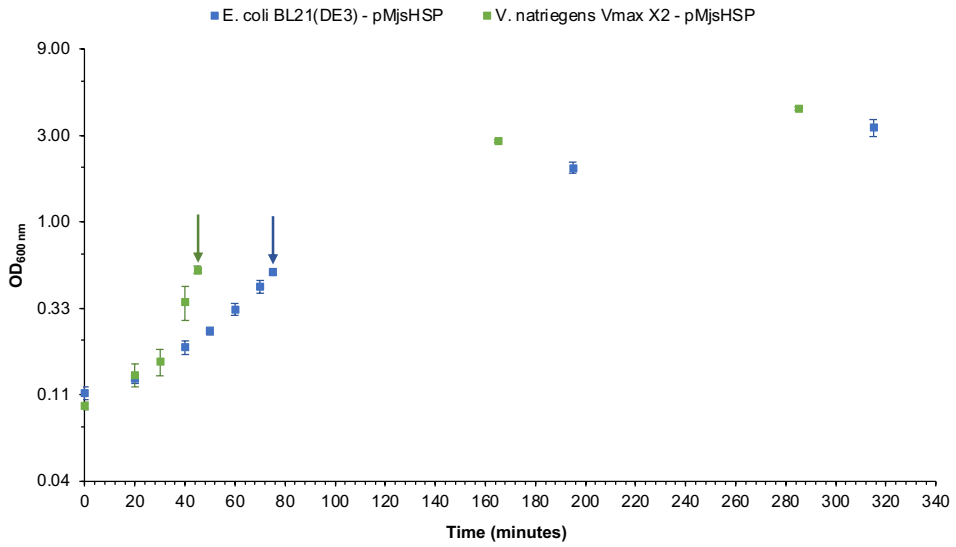


Figure III.5 – Growth curves obtained for the *E. coli* BL21(DE3) (■) and *V. natriegens* Vmax™ X2 (■) cells producing MjsHSP nanocages in standard conditions for cultivation and protein expression (non-baffled shake flask with 30% volume of culture medium; LB or LB supplemented with V2 salts, respectively; 37°C and 250 rpm; induction of protein expression at $OD_{600\ nm}$ of 0.5-0.6, with 1 mM IPTG, for 4 hours at 37°C). Arrows indicate IPTG induction time points. Each data point represents the mean of three biological replicates, with error bars indicating the respective standard deviation. The y-axis is on a logarithmic scale.

As evident from **Figure III.5**, *V. natriegens* Vmax™ X2 cells exhibited a significantly faster growth compared to *E. coli* BL21(DE3). The exponential growth phase for *V. natriegens* began at approximately the same time as *E. coli* (40 minutes from the re-inoculum) but it was significantly more pronounced, indicating a rapid cell division, as expected.

From the growth curves of both strains, it was possible to determine the growth parameters, specific growth rate and generation time, together with maximum $OD_{600\ nm}$ achieved at the end of the protein expression induction period. These data are represented in **Table III.3** for both *E. coli* and *V. natriegens* cells.

V. natriegens demonstrated a remarkably higher specific growth rate of 2.99 ± 0.10 h^{-1}, which was 2 times faster than *E. coli* at 1.50 ± 0.07 h^{-1}. This rapid growth is consistent with previous reports of exceptional generation times obtained for *V. natriegens* (199,213). The higher specific growth rate of this strain of *V. natriegens* indicates its potential for faster biomass accumulation and potentially shorter fermentation times, characteristics well described in the literature (200,214).

The generation time for *V. natriegens* was only 13.9 ± 0.5 minutes, compared to 27.9 ± 1.3 minutes for *E. coli*. This rapid generation time highlights the potential of *V. natriegens* for accelerated bioprocesses. The shorter generation time of *V. natriegens* suggests that it could complete more cell divisions within a given time frame, potentially leading to faster protein production cycles. This rapid growth implies that cells possess highly efficient protein synthesis and cellular machinery.

Interestingly, together with its rapid growth, *V. natriegens* cells reached higher maximum $OD_{600\ nm}$ of 4.25 ± 0.09 compared to *E. coli* (3.33 ± 0.36). Note that these values were obtained at the end of the protein expression induction period. This is an interesting result, since the rapid growth of *V.*

natriegens could mean that the biomass concentration is not so high, with the resources available in the culture medium being consumed more rapidly, becoming limiting.

Table III.3 – Growth parameters of *E. coli* BL21(DE3) and *V. natriegens* Vmax™ X2 cells producing MjsHSP nanocages in standard conditions for cultivation and protein expression (non-baffled shake flask with 30% volume of culture medium; LB or LB supplemented with V2 salts, respectively; 37°C and 250 rpm; induction of protein expression at $OD_{600\ nm}$ of 0.5-0.6, with 1 mM IPTG, for 4 hours at 37°C). Each result represents the mean of three biological replicates, with error bars indicating the respective standard deviation.

Growth parameters	*E. coli* BL21(DE3)	*V. natriegens* Vmax™ X2
Specific growth rate (h^{-1})	1.50 ± 0.07	2.99 ± 0.10
Generation time (min)	27.9 ± 1.3	13.9 ± 0.5
Maximum $OD_{600\ nm}$	3.33 ± 0.36	4.25 ± 0.09

These growth parameters together demonstrate the unique characteristics of *V. natriegens* Vmax™ X2 as a potential host for recombinant protein production. Its rapid growth and short generation time could lead to significant time savings in bioprocesses. However, further optimization of the culture medium composition and the growth conditions may be necessary to fully exploit the advantages of *V. natriegens*.

The rapid cellular metabolism and efficient biomass accumulation of *V. natriegens* Vmax™ X2 observed during the cell growth suggest an underlying molecular machinery capable of supporting high-rate protein synthesis. To investigate this potential, the next step consisted of evaluating the protein production capabilities of *V. natriegens*, focusing on the expression of MjsHSP nanocages.

This production was assessed by obtaining an SDS-PAGE of clarified lysates (soluble fraction) containing the MjsHSP nanocages expressed in both *E. coli* and *V. natriegens* cells using the standard conditions for cultivation and protein expression. **Figure III.6** shows clear bands corresponding to the MjsHSP nanocages (expected molecular weight of 16.5 kDa) in both *E. coli* (**Figure III.6A**) and *V. natriegens* (**Figure III.6B**). Note that since SDS-PAGE is denaturing, the protein bands referring to the MjsHSP nanocages effectively correspond to their monomer. The intensity of the MjsHSP nanocages band appears weaker in *V. natriegens* compared to *E. coli*, suggesting lower expression levels. Comparing the profile of bands obtained for both bacterial strains, there appears to be a larger presence of protein impurities in the case of *V. natriegens*, which may be in some way related to the protein components involved in its fast cellular metabolism (215).

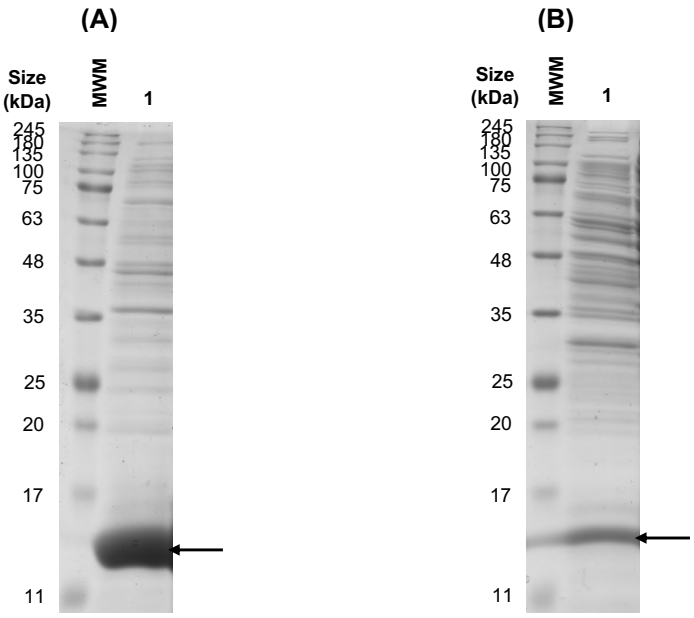


Figure III.6 – SDS-PAGE analysis of cell extracts obtained in the production of MjsHSP nanocages in *E. coli* BL21(DE3) **(B)** cells considering the standard conditions for cultivation and protein expression (non-baffled shake flask with 30% volume of culture medium; LB or LB supplemented with V2 salts, respectively; 37°C and 250 rpm; induction of protein expression at $OD_{600\ nm}$ of 0.5-0.6, with 1 mM IPTG, for 4 hours at 37°C). Arrows indicate the protein bands referring to the monomer of MjsHSP nanocages (expected molecular weight of 16.5 kDa), which is separated and visualized under denaturing conditions of SDS-PAGE. The abbreviation MWM refers to the molecular weight marker (NZYColour protein marker II; NZYtech Lisbon, Portugal).

To quantitatively compare the production efficiency of MjsHSP nanocages in *E. coli* BL21(DE3) and *V. natriegens* Vmax™ X2, three key characterization parameters were determined using **Equation III.5**, **Equation III.6** and **Equation III.7**, respectively: production yield, volumetric productivity, and specific productivity. For this purpose, the densitometric analysis of SDS-PAGE was performed according to the methodology described in the section **III.2.14.3. Protein nanocages quantification by densitometric analysis of SDS-PAGE**, allowing to obtain a ratio of protein nanocages monomers in relation to total protein. With the value for this ratio and considering the total protein mass in the sample under analysis, it was possible to estimate the mass of protein nanocage monomers in the respective sample. Given that each MjsHSP nanocage is composed of 24 identical monomers, this total mass of protein nanocages monomers is identical to the mass of the self-assembled protein nanocages. The results obtained for these three parameters are shown in **Table III.4**.

The analysis of **Table III.4** allowed to compare the parameters of MjsHSP nanocages production between *E. coli* BL21(DE3) and *V. natriegens* Vmax™ X2 under standard cultivation and protein expression conditions. The production yield in *E. coli* was approximately two times higher than that of *V. natriegens*, indicating superior overall production in *E. coli* considering these conditions. Regarding the volumetric productivity, the differences in the values between the two bacterial strains suggested that *E. coli* is more efficient in producing MjsHSP nanocages per unit of volume and time, which could be advantageous for large scale production. In terms of the specific productivity, the results were closer

between *E. coli* and *V. natriegens*, with a slight advantage for the former, which indicated that *E. coli* cells are more efficient in producing MjsHSP nanocages on a per-cell basis.

Table III.4 – Parameters of MjsHSP nanocages production (production yield, volumetric productivity and specific productivity) obtained for *E. coli* BL21(DE3) and *V. natriegens* Vmax™ X2 cells considering the standard conditions for cultivation and protein expression (non-baffled shake flask with 30% volume of culture medium; LB or LB supplemented with V2 salts, respectively; 37°C and 250 rpm; induction of protein expression at $OD_{600\ nm}$ of 0.5-0.6, with 1 mM IPTG, for 4 hours at 37°C). Each result represents the mean of three biological replicates, with error bars indicating the respective standard deviation.

Parameters of MjsHSP nanocages production	*E. coli* BL21(DE3)	*V. natriegens* Vmax™ X2
Production yield (mg L^{-1})	55.0 ± 19.7	27.1 ± 7.0
Volumetric productivity (mg L^{-1} h^{-1})	10.5 ± 3.7	5.7 ± 1.5
Specific productivity (mg $OD_{600\ nm}^{-1}$ h^{-1})	0.09 ± 0.03	0.04 ± 0.01

Overall, *E. coli* BL21(DE3) outperforms *V. natriegens* Vmax™ X2 across all three parameters under the standard conditions used. The standard deviations indicate some variability in the results, particularly for *E. coli*, which shows larger standard deviations compared to *V. natriegens*.

These results suggest that, in these specific conditions defined as benchmark , *E. coli* BL21(DE3) cells remains a more effective host for MjsHSP nanocages production compared to *V. natriegens* Vmax™ X2. Additionally, these outcomes do not validate the eventual correlation between the quick cellular metabolism and effective biomass accumulation with a greater capacity for protein production. As previously discussed, the lack of data in the literature relating to the manufacturing process of NVPNs, particularly regarding the production stage, does not allow comparing and framing the results obtained in this work with those determined in other studies using MjsHSP nanocages as model protein.

Nevertheless, it is important to note that *V. natriegens* Vmax™ X2 was still capable of producing the MjsHSP nanocages. Furthermore, from what is publicly available, this is the first work that used this alternative bacterium for the expression of NVPNs. Thus, additional optimization of expression conditions for *V. natriegens* may be necessary to fully exploit its potential as an alternative production host for MjsHSP nanocages.

III.3.5. Optimized production of MjsHSP nanocages in *E. coli* and *V. natriegens*

The optimization of the production of MjsHSP nanocages in both *E. coli* BL21(DE3) and *V. natriegens* Vmax™ X2 cells was based on a systematic and step-by-step approach to improve upon the standard procedure previously studied (**III.3.4. Production of MjsHSP nanocages in *E. coli* and *V. natriegens***).

Based on the literature regarding NVPNs expression and optimization in various bacterial hosts, five key parameters were identified and analyzed to determine their impact on MjsHSP nanocages production. These parameters, chosen to identify optimal production conditions in both bacterial strains, were growth medium composition, concentration of protein expression inducer, temperature during protein expression, timepoint of induction and aeration conditions.

The optimization strategy involved first varying one parameter at a time while keeping others constant (testing different operational conditions for each parameter), next identifying the best condition for each parameter based on the production yield, volumetric productivity and specific productivity, and then applying the optimized condition to the following parameter. Note that the cell collection, washing, lysis, and clarification steps remained consistent with the benchmark procedure throughout the optimization process.

The first parameter to be studied was the growth medium composition, with the growth and expression of MjsHSP nanocages in *E. coli* BL21(DE3) cells being performed in seven culture media with different compositions: LB supplemented with V2 salts, LB3, 2xYT, enhanced 2xYT, TB, TB supplemented with V2 salts, and M9 minimal supplemented with glucose (**III.2.12. Optimization of production of MjsHSP nanocages**).

The remaining operating parameters were maintained as established in the cultivation and protein expression standard (non-baffled shake flask with 30% volume of culture medium; 37°C and 250 rpm; induction of protein expression at $OD_{600\ nm}$ of 0.5-0.6, with 1 mM of IPTG, for 4 hours at 37°C).

The growth curves obtained for the seven different culture media evaluated are shown in **Figure III.7**. For comparative purposes, standard culture medium (LB) was also included.

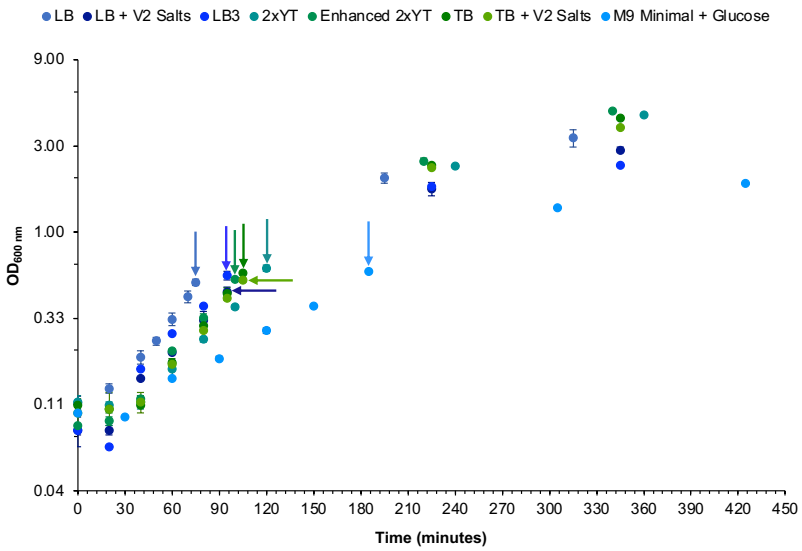


Figure III.7 – Growth curves obtained for the *E. coli* BL21(DE3) cells producing MjsHSP nanocages in seven culture media with different compositions, maintaining the remaining parameters as established in the standard cultivation and protein expression (non-baffled shake flask with 30% volume of culture medium; 37°C and 250 rpm; induction of protein expression at $OD_{600\ nm}$ of 0.5-0.6, with 1 mM of IPTG, for 4 hours at 37°C). Standard culture medium (LB) was also included as a reference point. Arrows indicate IPTG induction time points. Each data point represents the

mean of two biological replicates, with error bars indicating the respective standard deviation. The y-axis is on a logarithmic scale.

A preliminary observation of **Figure III.7** allowed to identify a different cell growth profile mainly between the six enriched culture media and the minimal culture medium, which was expected given the distinct availability of nutrients and other elements essential for the cellular metabolism of *E. coli* BL21(DE3). From these growth curves, it was possible to determine the growth parameters, specific growth rate and generation time, together with the maximum $OD_{600\ nm}$ achieved at the end of the protein expression induction period. These data are graphically represented in **Figure III.8**.

E. coli BL21(DE3) cells harboring the plasmid pMjsHSP presented the highest specific growth rate in standard LB culture medium ($1.50 \pm 0.07\ h^{-1}$), followed by the TB ($1.41 \pm 0.02\ h^{-1}$) and enhanced 2xYT ($1.36 \pm 0.02\ h^{-1}$) culture media (**Figure III.8A**). On the other hand, the lowest growth rate was recorded as expected in the M9 minimal medium supplemented with glucose ($0.63 \pm 0.01\ h^{-1}$).

In terms of the generation time (**Figure III.8B**)., considering that this is a parameter inversely proportional to the specific growth rate, the lowest values were found for these three media (27.9 ± 1.3 minutes, 29.5 ± 0.3 minutes and 30.5 ± 0.4 minutes, respectively, for LB, TB and enhanced 2xYT). For the M9 minimal medium supplemented with glucose, a significantly high value was obtained, exceeding 60 minutes (66.1 ± 0.6 minutes).

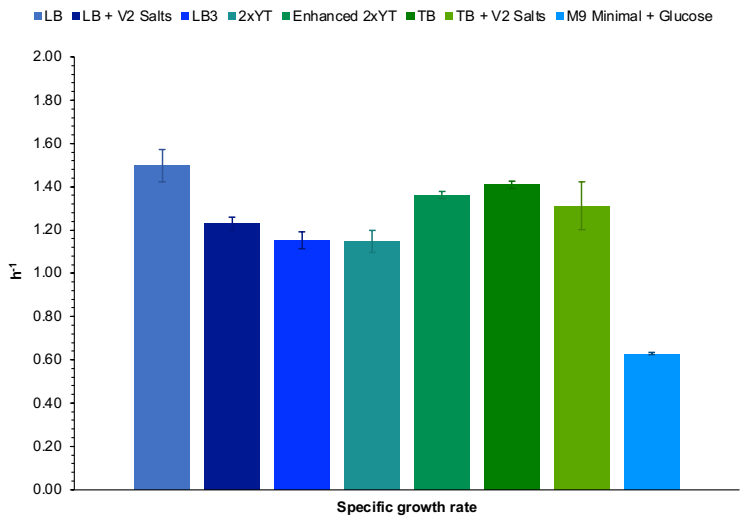
The use of LB supplemented with V2 salts, LB3, 2xYT and TB supplemented with V2 salts media allowed to record a moderate growth, which can be observed by the intermediate values obtained for both the specific growth rate and the generation time.

These results demonstrate that enriched or partially enriched media promote rapid bacterial growth reflected in specific growth rate and generation time due to their high nutrient availability, including amino acids, carbon sources, and essential growth factors. In the case of minimal media, since they only provide basic nutrients, it is necessary for the bacteria to synthesize many essential metabolites, which increases the metabolic burden, reducing the growth rate and consequently increasing the doubling time (221,222).

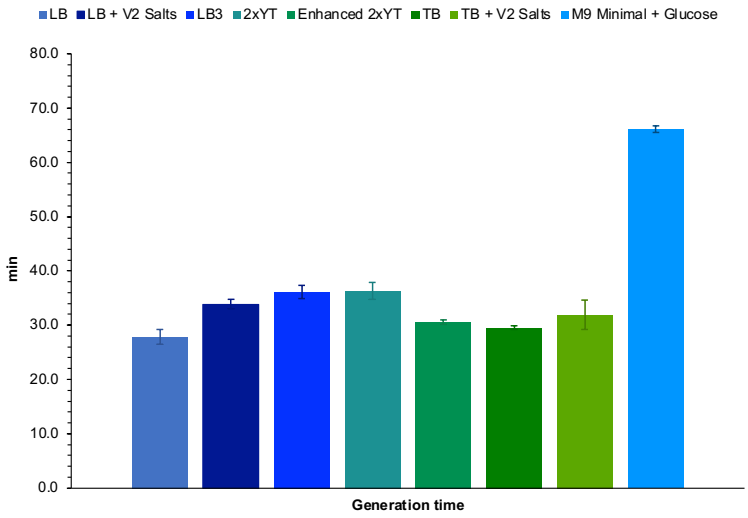
Analyzing **Figure III.8C**, the culture media enhanced 2xYT, 2xYT and TB are those that allowed to maximize the bacterial biomass for the production of MjsHSP nanocages, reaching values of $OD_{600\ nm}$ after the protein expression induction period greater than 4 (4.72 ± 0.02, 4.46 ± 0.04 and 4.30 ± 0.07, respectively, for the culture media enhanced 2xYT, 2xYT and TB). As expected, the M9 minimal medium with glucose allowed to reach only a maximum $OD_{600\ nm}$ of 1.87 ± 0.01.

The enhanced 2xYT seems to be the enriched culture medium with the most potential as it allowed rapid growth while also achieving a high biomass concentration. It is composed of tryptone as a source of nitrogenous compounds (including amino acids and peptides), yeast extract as a supplier of trace elements, vitamins and growth factors, sodium chloride for maintaining the osmotic balance, glucose as a source of carbon and phosphate buffer that is responsible for maintaining the pH stability. Comparing directly the growth parameters obtained with this culture medium with those of the standard LB medium, a similarity was observed for the specific growth rate and generation and a slight increase of 1.4 times in the maximum $OD_{600\ nm}$ reached.

(A)


LB
LB + V2 Salts
LB3
2xYT
Enhanced 2xYT
TB
TB + V2 Salts
M9 Minimal + Glucose
2.00
1.80
1.60
1.40
1.20
1.00
0.80
0.60
0.40
0.20
0.00
h-1
Specific growth rate

(B)


LB
LB + V2 Salts
LB3
2xYT
Enhanced 2xYT
TB
TB + V2 Salts
M9 Minimal + Glucose
80.0
70.0
60.0
50.0
40.0
30.0
20.0
10.0
0.0
min
Generation time

(C)

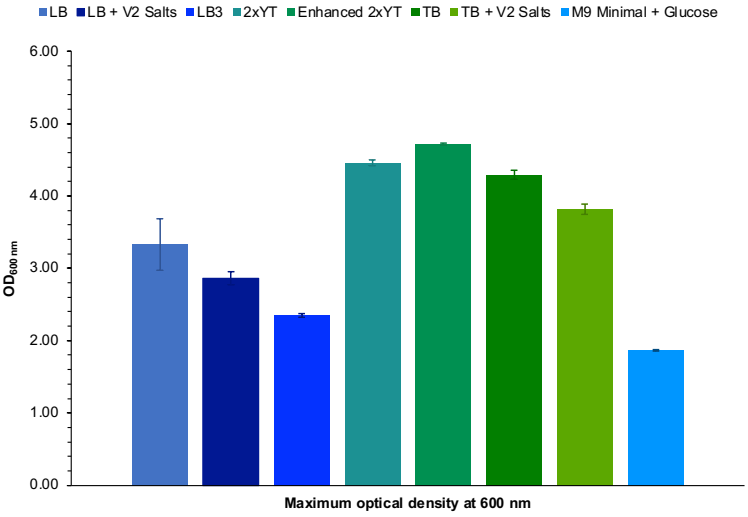

LB
LB + V2 Salts
LB3
2xYT
Enhanced 2xYT
TB
TB + V2 Salts
M9 Minimal + Glucose
6.00
5.00
4.00
3.00
2.00
1.00
0.00
OD600 nm
Maximum optical density at 600 nm

Figure III.8 – Growth parameters of *E. coli* BL21(DE3) producing MjsHSP nanocages in seven culture media with different compositions, maintaining the remaining parameters as established in the standard cultivation and protein expression (non-baffled shake flask with 30% volume of culture medium; 37°C and 250 rpm; induction of protein expression at $OD_{600\ nm}$ of 0.5-0.6, with 1 mM of IPTG, for 4 hours at 37°C). **(A)** Specific growth rate, **(B)** generation time, and **(C)** maximum $OD_{600\ nm}$. Standard culture medium (LB) was also included as a reference point. Each result represents the mean of two biological replicates, with error bars indicating the respective standard deviation.

As previously performed, the next step consisted of analyzing the production of MjsHSP nanocages in *E. coli* BL21(DE3) cells cultivated in the seven alternative culture media.

This production was evaluated by obtaining an SDS-PAGE of the clarified lysates (soluble fraction) containing the MjsHSP nanocages expressed in the culture media with different compositions. **Figure III.9** shows clear bands corresponding to the MjsHSP nanocages (expected molecular weight of 16.5 kDa) in all experiments performed. Note that once SDS-PAGE is denaturing, the protein band referring to the protein nanocages effectively corresponds to their monomer. The intensity of the MjsHSP nanocages band appears to be stronger in the enhanced 2xYT culture medium (**Figure III.9D**) followed by the LB medium supplemented with V2 salts (**Figure III.9A**), which suggests a higher expression level. In the M9 minimal medium with glucose (**Figure III.9G**), it is possible to visualize a band of the protein nanocages monomers much less intense, together with the remaining impurities, which agrees with the results obtained in the analysis of the growth profile of *E. coli* BL21(DE3) under these conditions. In all experiments, a very similar pattern of protein impurities was observed. Interestingly, the band profile obtained for cultivation in standard LB medium allowed to identify a band of MjsHSP nanocages monomers with greater intensity and the remaining impurities with less prominence.

In order to quantitatively assess and compare the production efficiency of MjsHSP nanocages in *E. coli* BL21(DE3) using the seven culture media with different compositions, three key characterization parameters previously presented were determined: production yield (**Equation III.5**), volumetric productivity (**Equation III.6**), and specific productivity (**Equation III.7**). As explained in the section **III.3.4. Production of MjsHSP nanocages in *E. coli* and *V. natriegens***, the determination of these three parameters implied a densitometric analysis of SDS-PAGE, which was performed according to the methodology described above (**III.2.14.3. Protein nanocages quantification by densitometric analysis of SDS-PAGE**). Due to the composition of each MjsHSP nanocage in 24 identical monomers, the total mass of protein nanocages monomers is identical to the mass of the assembled nanocages. The results obtained for these three parameters are shown graphically in **Figure III.10**.

Analyzing **Figure III.10**, it is possible to clearly verify that the cultivation of cells of this strain of *E. coli* in the enhanced 2xYT medium allowed to obtain values for these three parameters that were much higher than those of the remaining six culture media analyzed and also of the standard LB culture medium.

Regarding the production yield (**Figure III.10A**), this parameter is 1.8 times higher in enhanced 2xYT (163.9 ± 6.4 mg L^{-1}) compared to 2xYT (90.6 ± 1.9 mg L^{-1}), which is the second best culture medium, and 3 times higher than standard LB medium (55.0 ± 19.7 mg L^{-1}). For the remaining culture media evaluated there was some fluctuation in the production yield values, with the lowest value being for M9 minimal medium supplemented with glucose (24.0 ± 2.0 mg L^{-1}).

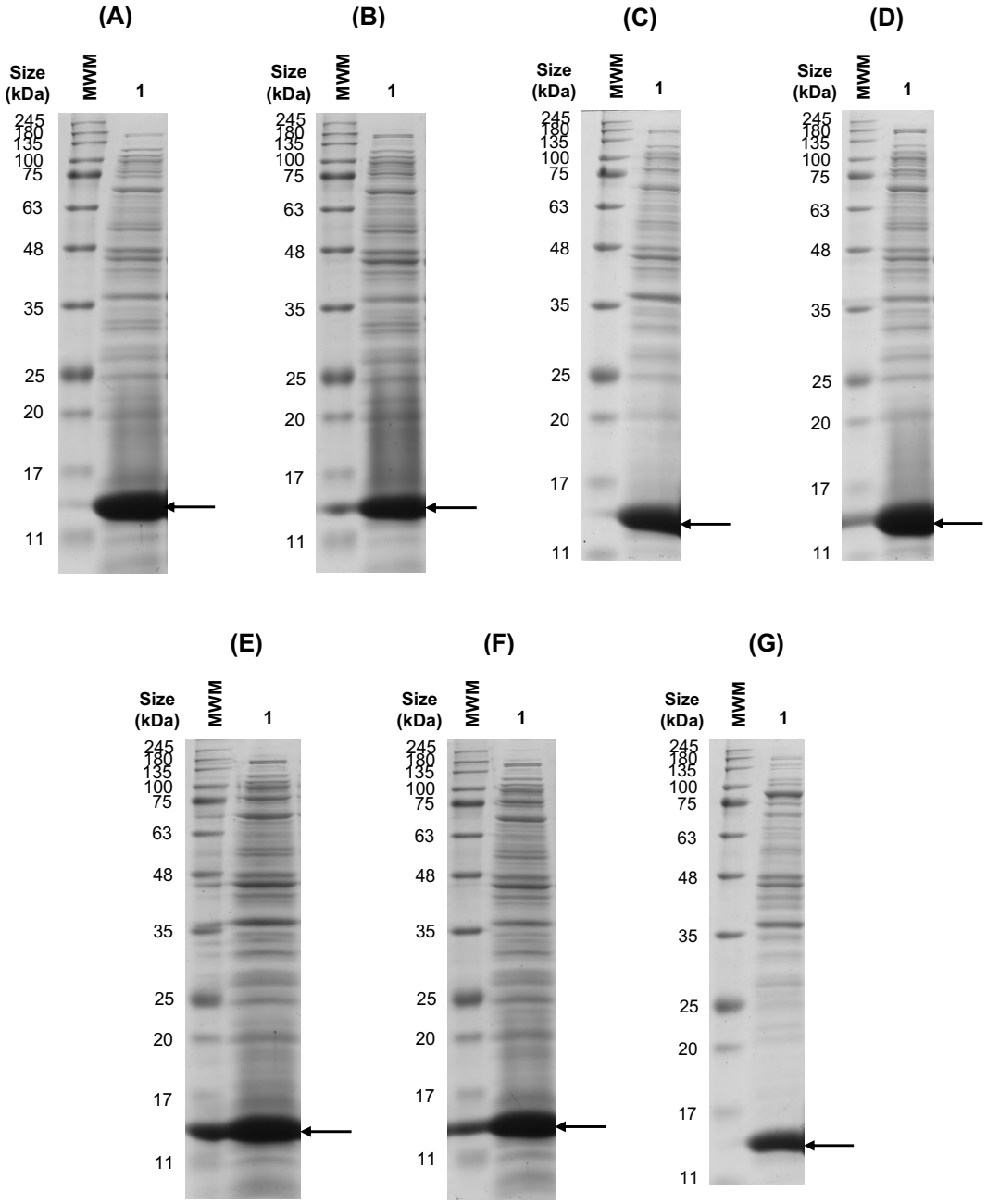
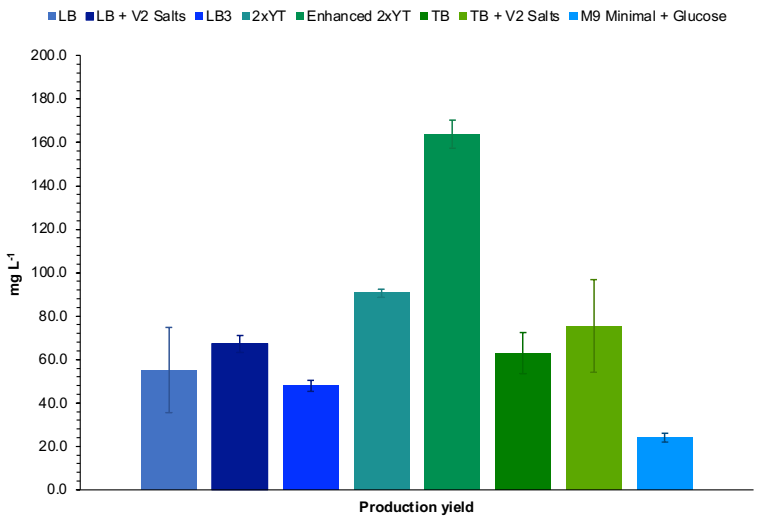
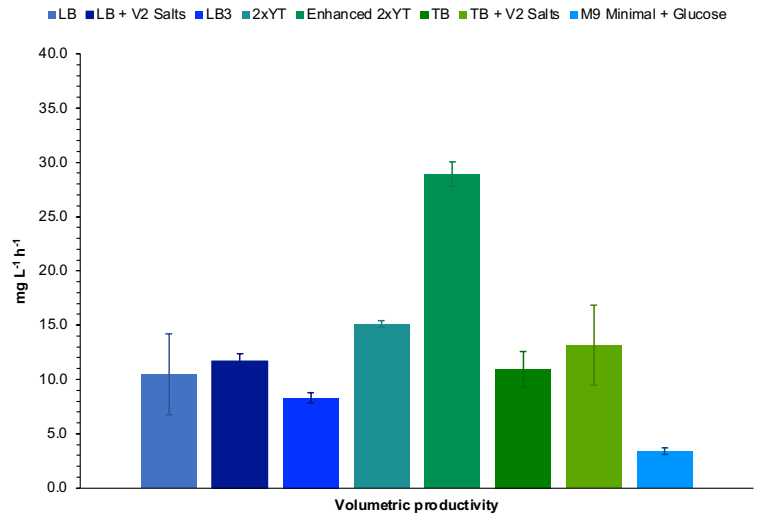


Figure III.9 – SDS-PAGE analysis obtained in the production of MjsHSP nanocages in *E. coli* BL21(DE3) cells using seven culture media with different compositions, maintaining the remaining parameters as established in the standard cultivation and protein expression (non-baffled shake flask with 30% volume of culture medium; 37°C and 250 rpm; induction of protein expression at $OD_{600\ nm}$ of 0.5-0.6, with 1 mM of IPTG, for 4 hours at 37°C). **(A)** LB + V2 salts, **(B)** LB3, **(C)** 2xYT, **(D)** enhanced 2xYT, **(E)** TB, **(F)** TB + V2 salts, and **(G)** M9 minimal + glucose. Arrows indicate the protein bands referring to the monomer of MjsHSP nanocages (expected molecular weight of 16.5 kDa), which is separated and visualized under denaturing conditions of SDS-PAGE. The abbreviation MWM refers to the molecular weight marker (NZYColour protein marker II; NZYtech Lisbon, Portugal).

(A)


LB
LB + V2 Salts
LB3
2xYT
Enhanced 2xYT
TB
TB + V2 Salts
M9 Minimal + Glucose
200.0
180.0
160.0
140.0
120.0
100.0
80.0
60.0
40.0
20.0
0.0
mg L-1
Production yield

(B)


LB
LB + V2 Salts
LB3
2xYT
Enhanced 2xYT
TB
TB + V2 Salts
M9 Minimal + Glucose
40.0
35.0
30.0
25.0
20.0
15.0
10.0
5.0
0.0
mg L-1 h-1
Volumetric productivity

(C)

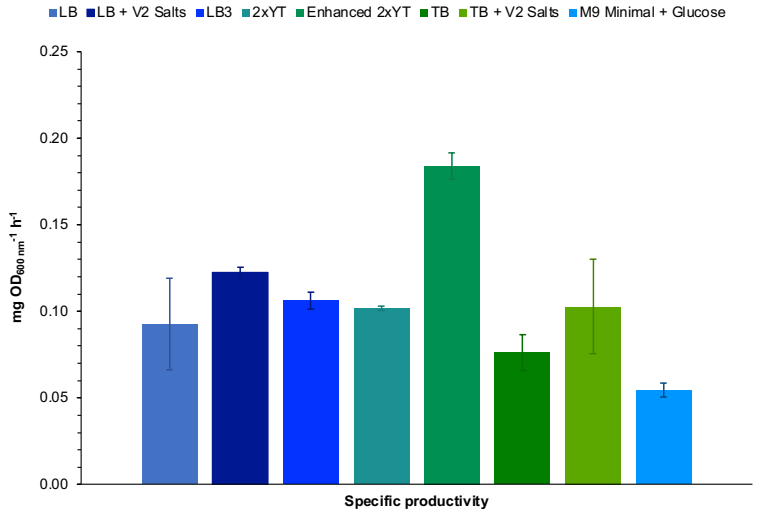

LB
LB + V2 Salts
LB3
2xYT
Enhanced 2xYT
TB
TB + V2 Salts
M9 Minimal + Glucose
0.25
0.20
0.15
0.10
0.05
0.00
mg OD600 nm-1 h-1
Specific productivity

Figure III.10 – Parameters of MjsHSP nanocages production obtained for *E. coli* BL21(DE3) cells in seven culture media with different compositions, maintaining the remaining parameters as established in the standard cultivation and protein expression (non-baffled shake flask with 30% volume of culture medium; 37°C and 250 rpm; induction of protein expression at $OD_{600\ nm}$ of 0.5-0.6, with 1 mM of IPTG, for 4 hours at 37°C). **(A)** Production yield, **(B)** volumetric productivity, and **(C)** specific productivity. Standard culture medium (LB) was also included as a reference point. Each result represents the mean of two biological replicates, with error bars indicating the respective standard deviation.

Similarly, for the volumetric productivity (**Figure III.10B**), in the enhanced 2xYT this value (28.9 ± 1.1 mg L^{-1} h^{-1}) was 1.9 and 2.8 times higher than 2xYT (15.1 ± 0.3 mg L^{-1} h^{-1}) and standard LB (10.5 ± 3.7 mg L^{-1} h^{-1}), respectively. The M9 minimal medium with glucose was the one with the smallest value (3.4 ± 0.3 mg L^{-1} h^{-1}). The enhanced 2xYT is an enriched culture medium that provides additional nutrients and cofactors along with improved pH stability, which subsequently allows for not only higher cell densities but also more expressive production of MjsHSP nanocages.

For the specific productivity that evaluates the production of MjsHSP nanocages per bacterial cell (**Figure III.10C**), the enhanced 2xYT medium presented the highest value (0.18 ± 0.01 mg $OD_{600\ nm}^{-1}$ h^{-1}), followed by LB supplemented with V2 salts (0.12 ± 0.00 mg $OD_{600\ nm}^{-1}$ h^{-1}) and LB3 (0.11 ± 0.01 mg $OD_{600\ nm}^{-1}$ h^{-1}), which represented a slight increase of 1.5 and 1.6 times, respectively. For this parameter, there was a greater proximity between the values obtained for all culture media tested.

It is interesting to note that in general the calculated standard deviations were low, which allowed to conclude that there was proximity in the three independent biological assays performed for each culture medium.

With these results it was concluded that the cultivation of *E. coli* BL21(DE3) cells and the subsequent expression of MjsHSP nanocages in the enhanced 2xYT culture medium was what appears to be most promising, even considering a larger scale production. Thus, this was the culture medium chosen to advance to the optimization of the second selected parameter, the concentration of IPTG as inducer of protein expression.

To test the influence of this second parameter on the production of MjsHSP nanocages in *E. coli* BL21(DE3), the expression of these protein nanocages was induced in distinct experiments with three alternative IPTG concentrations to the standard value of 1 mM (**III.2.12. Optimization of production of MjsHSP nanocages**).

The remaining operating parameters were maintained as established in the cultivation and protein expression standard (non-baffled shake flask with 30% volume of culture medium; 37°C and 250 rpm; induction of protein expression at $OD_{600\ nm}$ of 0.5-0.6, for 4 hours at 37°C).

The growth curves obtained for the three concentrations of IPTG evaluated are shown in **Figure III.11**. For comparative purposes, standard concentration of IPTG (1 mM) was also included.

By analyzing **Figure III.11**, it was found that the growth profile for *E. coli* BL21(DE3) cells in independent experiments for induction with different IPTG concentrations is similar, an expected result since this second parameter have a significant influence on the production of protein nanocages and

not so much on the behavior of the growth curve. The three growth parameters studied (specific growth rate, generation time, and maximum $OD_{600\ nm}$ achieved) are graphically represented in **Figure III.12**.

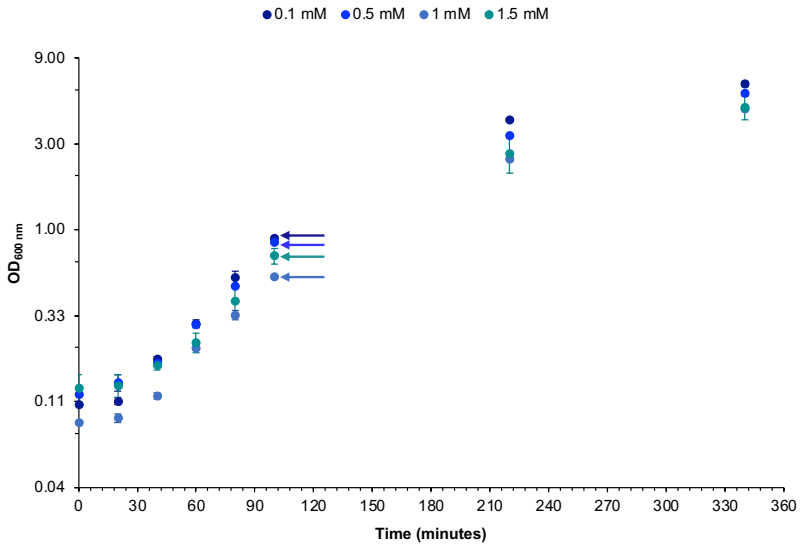
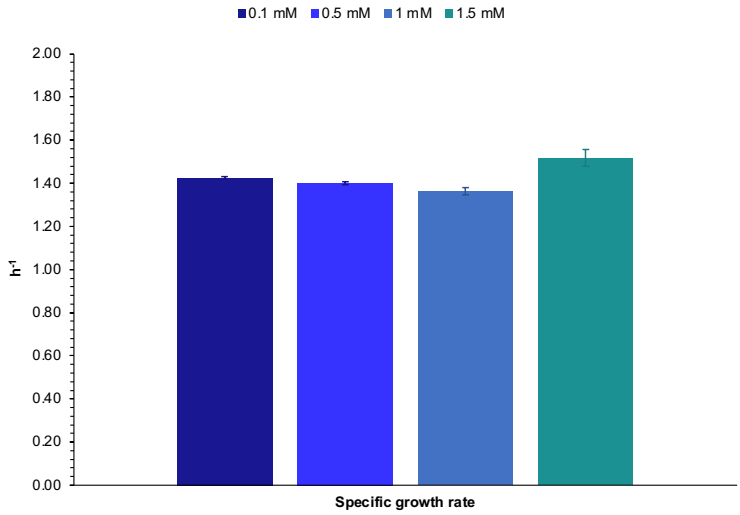


Figure III.11 – Growth curves obtained for the *E. coli* BL21(DE3) cells producing MjsHSP nanocages with the induction of protein expression using three different concentrations of IPTG, maintaining the remaining parameters as previously established (non-baffled shake flask with 30% volume of culture medium; 37°C and 250 rpm; induction of protein expression at $OD_{600\ nm}$ of 0.5-0.6, for 4 hours at 37°C). Standard concentration of IPTG (1 mM) was also included as a reference point. Arrows indicate IPTG induction time points. Each data point represents the mean of two biological replicates, with error bars indicating the respective standard deviation. The y-axis is on a logarithmic scale.

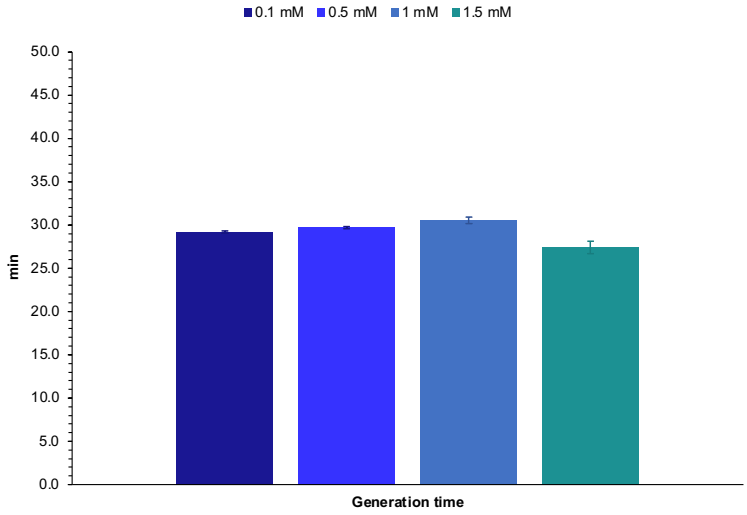
Regarding the specific growth rate (**Figure III.12A**) and the generation time (**Figure III.12B**), the values obtained for the three conditions under evaluation were in agreement with each other as well as with the standard condition of 1 mM of IPTG as inducer (about 1.40 to 1.50 h^{-1} for the specific growth rate and about 27.0 to 30.0 minutes for the generation time). Because these parameters were calculated in the period corresponding to the experiential growth phase and this did not present significant differences in the respective growth curves, it was an expected result and in accordance with what was previously mentioned.

In terms of the maximum $OD_{600\ nm}$ (**Figure III.12C**), it was found that induction with a lower concentration of IPTG (0.1 mM) allowed to obtain a value 1.4 times higher (6.45 ± 0.21) than the standard condition of 1 mM (4.72 ± 0.02) and 1.2 times higher than the second most favorable condition of 0.5 mM (5.75 ± 0.00). This observation may be justified by the fact that a lower concentration of the IPTG inducer means that the protein nanocages were not overexpressed and as such allowed for a better balance between this same expression and the biomass growth. It should be noted that these values correspond to the measurement at the end of the induction period.

(A)


0.1 mM
0.5 mM
1 mM
1.5 mM
2.00
1.80
1.60
1.40
1.20
1.00
0.80
0.60
0.40
0.20
0.00
h-1
Specific growth rate

(B)


0.1 mM
0.5 mM
1 mM
1.5 mM
50.0
45.0
40.0
35.0
30.0
25.0
20.0
15.0
10.0
5.0
0.0
min
Generation time

(C)

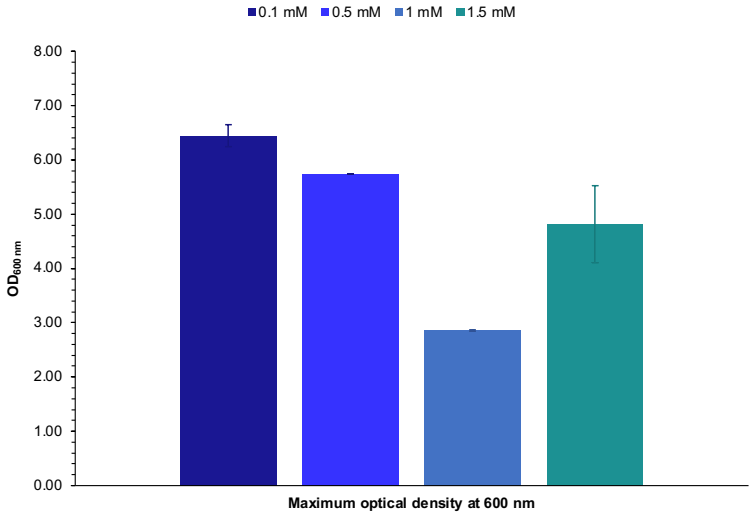

0.1 mM
0.5 mM
1 mM
1.5 mM
8.00
7.00
6.00
5.00
4.00
3.00
2.00
1.00
0.00
OD600 nm
Maximum optical density at 600 nm

Figure III.12 – Growth parameters of *E. coli* BL21(DE3) producing MjsHSP nanocages with the induction of protein expression using three different concentrations of IPTG, maintaining the remaining parameters as previously established (non-baffled shake flask with 30% volume of culture medium; 37°C and 250 rpm; induction of protein expression at $OD_{600\ nm}$ of 0.5-0.6, for 4 hours at 37°C). **(A)** Specific growth rate, **(B)** generation time, and **(C)** maximum $OD_{600\ nm}$. Standard concentration of IPTG (1 mM) was also included as a reference point. Each result represents the mean of two biological replicates, with error bars indicating the respective standard deviation.

Afterwards, the production of MjsHSP nanocages in *E. coli* BL21(DE3) cells using different concentrations of IPTG as inducer was evaluated by obtaining an SDS-PAGE of the clarified lysates (soluble fraction) containing this model of protein nanocages. Clear bands corresponding to the MjsHSP nanocages (expected molecular weight of 16.5 kDa) in all experiments performed can be visualized in **Figure III.13**. As SDS-PAGE is denaturing, the observed protein bands correspond to monomers, regardless of their native oligomeric state. The intensity of the MjsHSP nanocages band appears to be stronger with the use of 0.5 mM of IPTG for the induction of protein production (**Figure III.13B**), which suggests a higher expression level. Again, as expected, a very similar pattern of protein impurities was observed for the tested conditions. The migration profile obtained for the standard induction of protein expression (**Figure III.6A**) allowed the identification of a band of MjsHSP nanocages monomers with lower intensity and the remaining impurities with similar prominence.

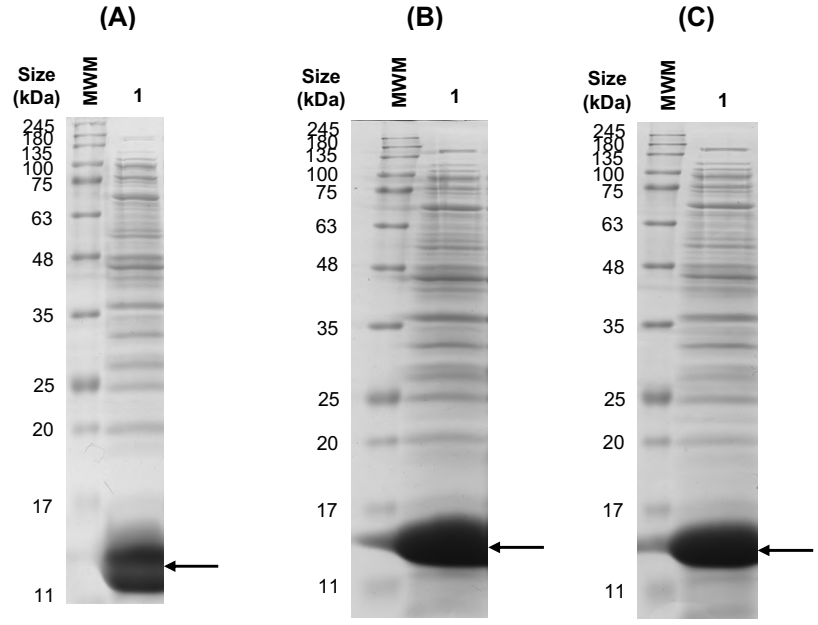


Figure III.13 – SDS-PAGE analysis obtained in the production of MjsHSP nanocages in *E. coli* BL21(DE3) cells using three different concentrations of IPTG for the induction of protein expression, maintaining the remaining parameters as previously established (non-baffled shake flask with 30% volume of culture medium; 37°C and 250 rpm; induction of protein expression at $OD_{600\ nm}$ of 0.5-0.6, for 4 hours at 37°C). **(A)** 0.1 mM, **(B)** 0.5 mM, and **(C)** 1.5 mM of IPTG. Arrows indicate the protein bands referring to the monomer of MjsHSP nanocages (expected molecular weight of 16.5 kDa), which is separated and visualized under denaturing conditions of SDS-PAGE. The abbreviation MWM refers to the molecular weight marker (NZYColour protein marker II; NZYtech Lisbon, Portugal).

In order to quantitatively assess and compare the production efficiency of MjsHSP nanocages in *E. coli* BL21(DE3) using the three concentrations of IPTG inducer, production yield (**Equation III.5**), volumetric productivity (**Equation III.6**) and specific productivity (**Equation III.7**) were determined. The methodology used in this analysis based on densitometric analysis of SDS-PAGE was previously explained (**III.3.4. Production of MjsHSP nanocages in *E. coli* and *V. natriegens***). The results obtained for these three key parameters are shown graphically in **Figure III.14**.

By analyzing **Figure III.14**, it can be seen that the induction of the expression of MjsHSP nanocages in *E. coli* BL21(DE3) with 0.1 mM of IPTG allowed to obtain higher values for production yield, volumetric productivity and specific productivity in comparison with the remaining conditions tested including the standard concentration of IPTG (1 mM).

The production yield (**Figure III.14A**) was about 1.2 times higher with 0.1 mM of IPTG (232.9 ± 18.8 mg L^{-1}) compared to the second best concentration of inducer (0.5 mM; 200.0 ± 30.0 mg L^{-1}), and about 1.4 times higher than the reference 1 mM (10-fold higher concentration; 163.9 ± 6.4 mg L^{-1}).

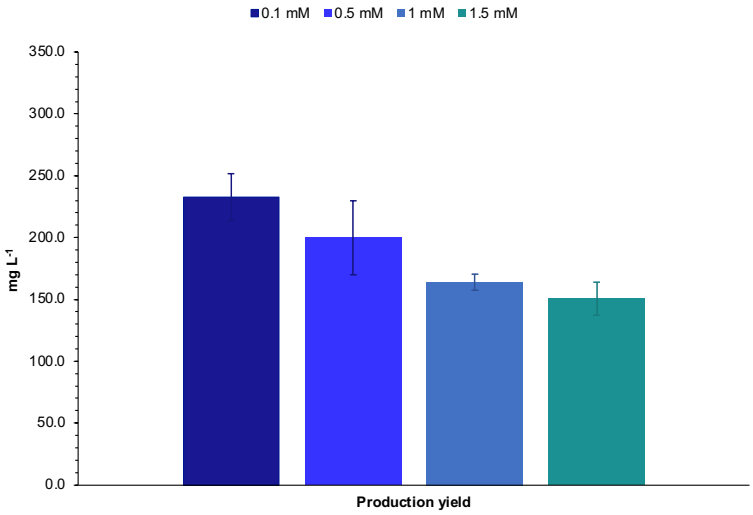
Similarly, for the volumetric productivity (**Figure III.14B**), it was obtained a value 1.4 times higher in 0.1 mM of IPTG (41.1 ± 3.3 mg L^{-1} h^{-1}) than the standard 1 mM (28.9 ± 1.1 mg L^{-1} h^{-1}), also observing a growth profile in the value of this parameter as the concentration of IPTG used as an inducer increased.

For the specific productivity (**Figure III.14C**). the differences in the results for the various conditions analyzed were not so significant (from 0.17 ± 0.01 to 0.19 ± 0.02 mg $OD_{600\ nm}^{-1}$ h^{-1}), although 0.1 mM of IPTG could be highlighted as the most favorable condition (0.19 ± 0.02 mg $OD_{600\ nm}^{-1}$ h^{-1}).

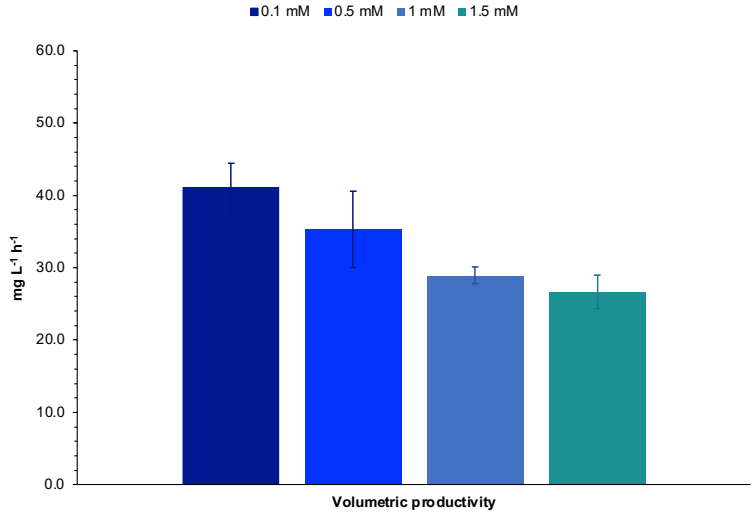
These results are in agreement with what is described in the literature, in which typically higher concentrations of IPTG (above 0.5 mM) can be detrimental to protein expression because this excess of IPTG leads to an overproduction of target protein mRNA, outcompeting endogenous mRNA for ribosomes and reducing cell viability (223). Some studies even recommend the use of IPTG concentrations between 0.05 and 0.2 mM to obtain the optimal amount and quality of proteins (224).

Taking these observations into consideration, it was concluded that the induction of MjsHSP nanocages expression using an IPTG concentration of 0.1 mM was the one that appeared to be most advantageous. Thus, together with the use of the enhanced 2xYT culture medium, this was the inducer concentration selected to advance to the optimization of the second parameter under analysis, the temperature for protein expression.

(A)


0.1 mM
0.5 mM
1 mM
1.5 mM
350.0
300.0
250.0
200.0
150.0
100.0
50.0
0.0
mg L-1
Production yield

(B)


0.1 mM
0.5 mM
1 mM
1.5 mM
60.0
50.0
40.0
30.0
20.0
10.0
0.0
mg L-1 h-1
Volumetric productivity

(C)

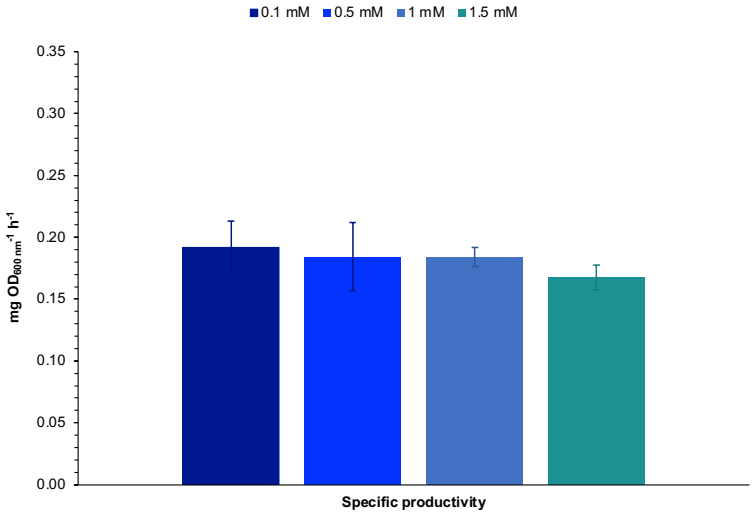

0.1 mM
0.5 mM
1 mM
1.5 mM
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
mg OD600 nm-1 h-1
Specific productivity

Figure III.14 – Parameters of MjsHSP nanocages production obtained for *E. coli* BL21(DE3) cells using three different concentrations of IPTG for the induction of protein expression, maintaining the remaining parameters as previously established (non-baffled shake flask with 30% volume of culture medium; 37°C and 250 rpm; induction of protein expression at $OD_{600\ nm}$ of 0.5-0.6, for 4 hours at 37°C). **(A)** Production yield, **(B)** volumetric productivity, and **(C)** specific productivity. Standard concentration of IPTG (1 mM) was also included as a reference point. Each result represents the mean of two biological replicates, with error bars indicating the respective standard deviation.

With the purpose of assessing the influence of this third parameter on the production of MjsHSP nanocages in *E. coli* BL21(DE3), protein expression was induced in different experiments with three alternative temperatures to the standard value of 37°C (**III.2.12. Optimization of production of MjsHSP nanocages**).

The remaining operating parameters were maintained as previously established (non-baffled shake flask with 30% volume of culture medium; enhanced 2xYT; 37°C and 250 rpm; induction of protein expression at $OD_{600\ nm}$ of 0.5-0.6, with 0.1 mM of IPTG, for 4 hours).

The growth curves obtained for the three temperatures of protein expression evaluated are shown in **Figure III.15**. It should be noted that the temperature was maintained at 37°C in all experiments for bacterial growth until induction with IPTG, and it was only at this point that it was adjusted to the value under analysis. For comparative purposes, standard temperature for protein nanocages expression (37°C) was also included.

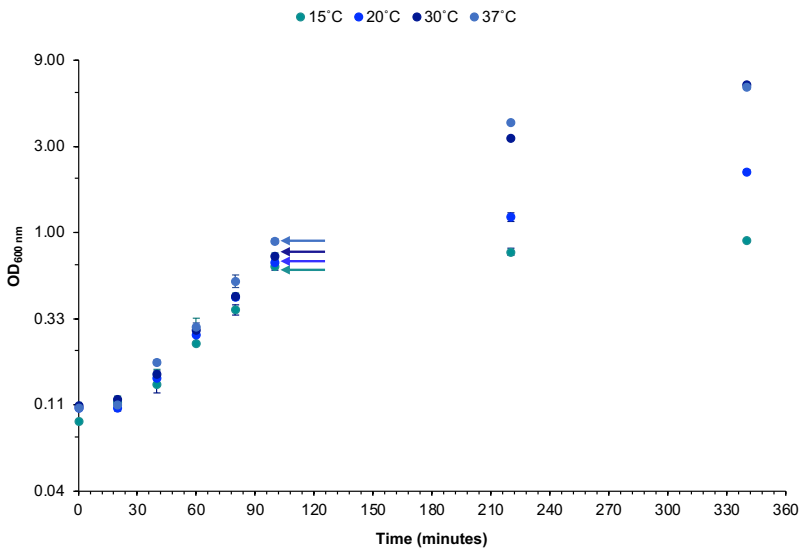


Figure III.15 – Growth curves obtained for the *E. coli* BL21(DE3) cells producing MjsHSP nanocages with the induction of protein expression using three different temperatures, maintaining the remaining parameters as previously established (non-baffled shake flask with 30% volume of culture medium; enhanced 2xYT; 37°C and 250 rpm; induction of protein expression at $OD_{600\ nm}$ of 0.5-0.6, with 0.1 mM of IPTG, for 4 hours). Standard temperature for induction of protein expression (37˚C) was also included as a reference point. Arrows indicate IPTG induction time points. Each data point represents the mean of two biological replicates, with error bars indicating the respective standard deviation. The y-axis is on a logarithmic scale.

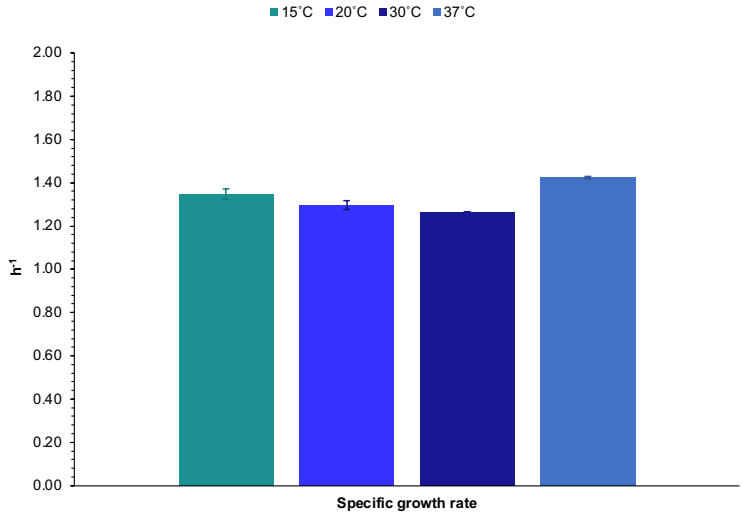
As previously observed in the optimization assays of the IPTG inducer concentration, **Figure III.15** shows that the growth profile for *E. coli* BL21(DE3) cells in the independent induction experiments with different protein induction temperatures was similar to each other until the time point of induction of MjsHSP nanocages expression. Thereafter, the temperature was adjusted for each test, so that the behavior of the bacterial cells also changed. For the two lowest temperatures (15°C and 20°C), it was possible to verify a significant reduction in the growth compared to the 30°C condition and the 37°C reference. Particularly, in the case of the experiment at 15°C, there was a complete interruption of the bacterial growth, probably because the low temperature inhibited all the metabolism of the *E. coli* cells. The three growth parameters studied (specific growth rate, generation time, and maximum $OD_{600\ nm}$ achieved) are graphically represented in **Figure III.16**.

Regarding the specific growth rate (**Figure III.16A**) and the generation time (**Figure III.16B**), the results obtained for the three conditions under study are consistent with each other (about 1.27 to 1.35 h^{-1} for the specific growth rate and about 30.9 to 32.9 minutes for the generation time) and also in relation to the standard temperature condition of 37°C ($1.42 \pm 0.01\ h^{-1}$ and 29.2 ± 0.0 minutes, respectively). This result was expected and is in line with previous observations, considering that these parameters were calculated during the exponential growth phase. Since there were no significant differences in the growth curves corresponding to this phase for the different conditions, it is natural that the values of these parameters were identical.

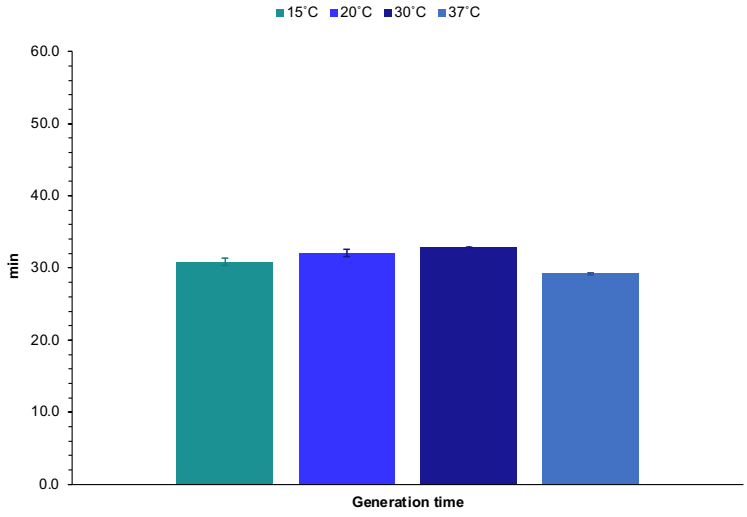
Regarding the maximum $OD_{600\ nm}$ (**Figure III.16C**), the use of induction temperatures of 30°C (6.64 ± 0.07) and 37°C (6.45 ± 0.21) allowed to obtain clearly higher values of final $OD_{600\ nm}$, approximately 7.3 and 7.1 times, respectively, in relation to those of 15°C (0.91 ± 0.02), the worst condition evaluated. The experiment with a temperature of 20°C (2.18 ± 0.06) led to the achievement of a maximum $OD_{600\ nm}$ 3.0 times lower than the most interesting temperature condition (30°C).

Then, the production of MjsHSP nanocages in *E. coli* BL21(DE3) cells using different induction temperatures was evaluated by obtaining an SDS-PAGE of the clarified lysates (soluble fraction) containing the protein nanocages. Observing **Figure III.17**, it can be seen that in the three experiments performed the specific band corresponding to the MjsHSP nanocages (expected molecular weight of 16.5 kDa) is present. Since SDS-PAGE is a denaturing technique, proteins are separated into their individual subunits, and the resulting bands on the gel represent monomeric forms. An increasing pattern of the intensity of the MjsHSP nanocages band was identified from the lowest temperature tested (15°C) to the highest (30°C), even if the result obtained for the standard temperature of 37°C is considered (**Figure III.13A**). This expected result, which corroborates the observations regarding the growth profile of bacterial cells in the various conditions tested, suggests a higher level of expression of protein nanocages in the standard condition. A variable pattern of bands associated with protein impurities was also observed.

(A)


15˚C 20˚C 30˚C 37˚C
2.00
1.80
1.60
1.40
1.20
1.00
0.80
0.60
0.40
0.20
0.00
h-1
Specific growth rate

(B)


15˚C 20˚C 30˚C 37˚C
60.0
50.0
40.0
30.0
20.0
10.0
0.0
min
Generation time

(C)

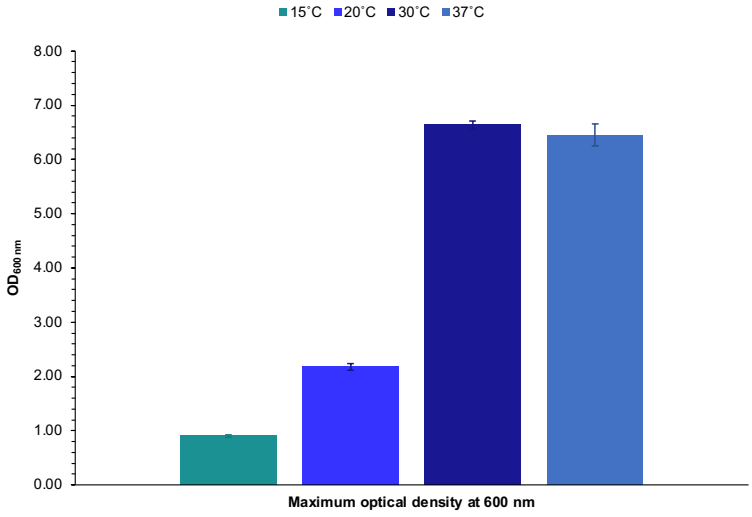

15˚C 20˚C 30˚C 37˚C
8.00
7.00
6.00
5.00
4.00
3.00
2.00
1.00
0.00
OD600 nm
Maximum optical density at 600 nm

Figure III.16 – Growth parameters of *E. coli* BL21(DE3) producing MjsHSP nanocages with the induction of protein expression using three different temperatures, maintaining the remaining parameters as previously established (non-baffled shake flask with 30% volume of culture medium; enhanced 2xYT; 37°C and 250 rpm; induction of protein expression at $OD_{600\ nm}$ of 0.5-0.6, with 0.1 mM of IPTG, for 4 hours). **(A)** Specific growth rate, **(B)** generation time, and **(C)** maximum $OD_{600\ nm}$. Standard temperature for induction of protein expression (37˚C) was also included as a reference point. Each result represents the mean of two biological replicates, with error bars indicating the respective standard deviation.

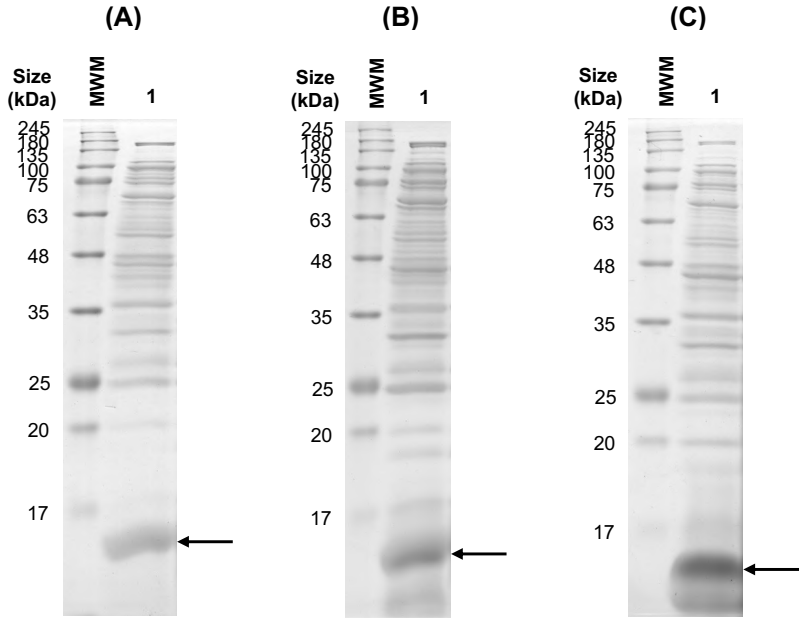
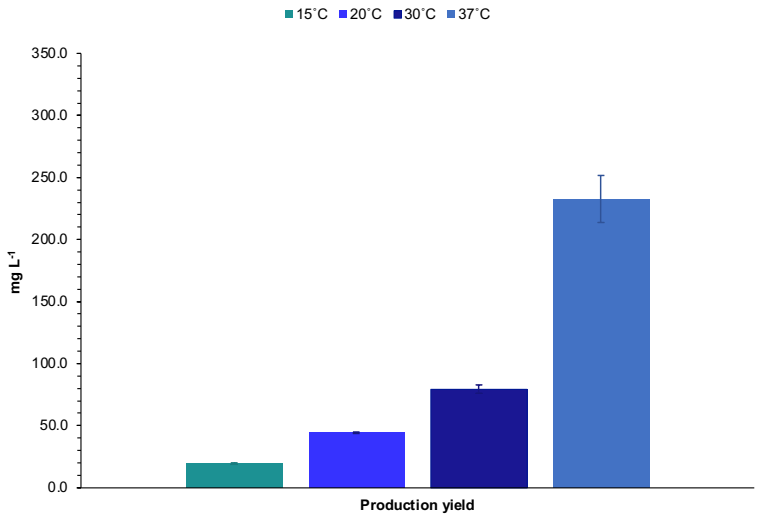


Figure III.17 – SDS-PAGE analysis obtained in the production of MjsHSP nanocages in *E. coli* BL21(DE3) cells using three different temperatures for the induction of protein expression, maintaining the remaining parameters as previously established (non-baffled shake flask with 30% volume of culture medium; enhanced 2xYT; 37°C and 250 rpm; induction of protein expression at $OD_{600\ nm}$ of 0.5-0.6, with 0.1 mM of IPTG, for 4 hours). **(A)** 15˚C, **(B)** 20˚C, and **(C)** 30˚C. Arrows indicate the protein bands referring to the monomer of MjsHSP nanocages (expected molecular weight of 16.5 kDa), which is separated and visualized under denaturing conditions of SDS-PAGE. The abbreviation MWM refers to the molecular weight marker (NZYColour protein marker II; NZYtech Lisbon, Portugal).

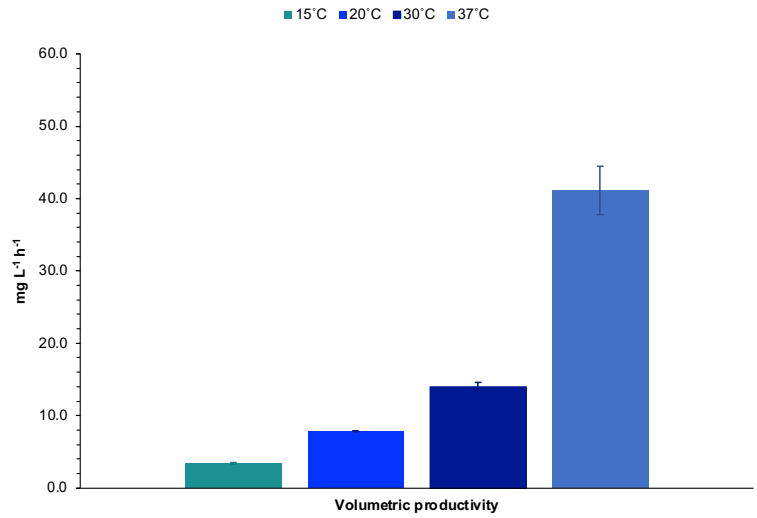
With the purpose of quantitatively evaluating the production efficiency of MjsHSP nanocages in *E. coli* BL21(DE3) under three different temperatures of induction, the production yield (**Equation III.5**), the volumetric productivity (**Equation III.6**) and the specific productivity (**Equation III.7**) were calculated. The approach employed for this analysis, which relies on the densitometric assessment of SDS-PAGE, was previously described in detail. **Figure III.18** graphically presents the results for these three main parameters.

Analyzing **Figure III.18**, it can be seen that the induction of the MjsHSP nanocages expression in *E. coli* BL21(DE3) with a standard temperature of 37°C allowed to obtain higher values for production yield, volumetric productivity and specific productivity in comparison with the three alternative options tested.

(A)


15˚C
20˚C
30˚C
37˚C
350.0
300.0
250.0
200.0
150.0
100.0
50.0
0.0
mg L-1
Production yield

(B)


15˚C
20˚C
30˚C
37˚C
60.0
50.0
40.0
30.0
20.0
10.0
0.0
mg L-1 h-1
Volumetric productivity

(C)

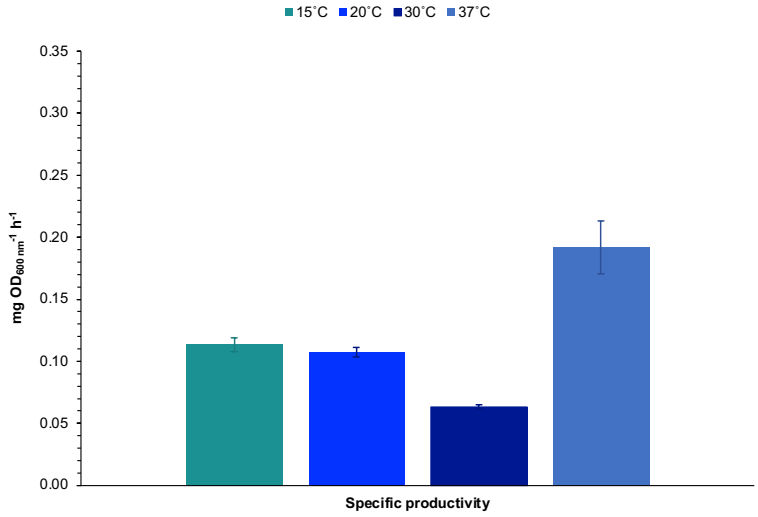

15˚C
20˚C
30˚C
37˚C
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
mg OD600 nm-1 h-1
Specific productivity

Figure III.18 – Parameters of MjsHSP nanocages production obtained for *E. coli* BL21(DE3) cells using three different temperatures for the induction of protein expression, maintaining the remaining parameters as previously established (non-baffled shake flask with 30% volume of culture medium; enhanced 2xYT; 37°C and 250 rpm; induction of protein expression at $OD_{600\ nm}$ of 0.5-0.6, with 0.1 mM of IPTG, for 4 hours). **(A)** Production yield, **(B)** volumetric productivity, and **(C)** specific productivity. Standard temperature for induction of protein expression (37˚C) was also included as a reference point. Each result represents the mean of two biological replicates, with error bars indicating the respective standard deviation.

The production yield (**Figure III.18A**) was approximately 2.9 times higher with the standard temperature in protein nanocages expression of 37°C (232.9 ± 18.8 mg L^{-1}) compared to the second best temperature (30°C; 79.3 ± 3.2 mg L^{-1}), and approximately 12.0 times higher than protein induction at 15°C (19.4 ± 0.6 mg L^{-1}).

Similarly, for the volumetric productivity (**Figure III.18B**), a value 2.9 and 12.1 times higher was obtained for the standard temperature of 37°C (41.1 ± 3.3 mg L^{-1} h^{-1}) in relation to 30°C (14.0 ± 0.6 mg L^{-1} h^{-1}) and 15°C (3.4 ± 0.1 mg L^{-1} h^{-1}), respectively, also observing a profile of growing value of this parameter as the induction temperature increases.

For the specific productivity (**Figure III.18C**), the use of the standard temperature of 37°C was the most favorable condition (0.19 ± 0.02 mg $OD_{600\ nm}^{-1}$ h^{-1}), although curiously it was possible to obtain slightly higher values at 15°C (1.8 times; 0.11 ± 0.01 mg $OD_{600\ nm}^{-1}$ h^{-1}) and 20°C (1.8 times; 0.11 ± 0.00 mg $OD_{600\ nm}^{-1}$ h^{-1}) in relation to 30°C (0.06 ± 0.00 mg $OD_{600\ nm}^{-1}$ h^{-1}).

In the literature, the relationship between protein expression and the temperature of induction in *E. coli* strains is complex and multifaceted. Generally, it is considered that the use of higher temperatures, namely close to 37°C, may lead to a greater accumulation of proteins expressed as inclusion bodies, due to the fact that there is an increased peptide chain elongation rate (225). On the other hand, at low temperatures (around 15°C), there is a pool of non-translating ribosomes, indicating defects in the translation initiation (225). Consequently, inducing protein expression at more moderate temperatures may help to improve protein folding and solubility, reducing the formation of the aforementioned inclusion bodies (187). Nevertheless, there are some reports in the literature that identify the optimal temperature for protein expression in *E. coli* BL21(DE3) as 37°C (226), an observation that, although it refers to the specific protein under study, is completely in line with the results obtained in this work, in which the production of MjsHSP nanocages in this strain of *E. coli* appears to be more advantageous at 37°C.

Thus, it was concluded that the induction of MjsHSP nanocages expression at a temperature of 37°C was the one that appears to be most efficient. Therefore, together with the use of the enhanced 2xYT culture medium and an IPTG concentration of 0.1 mM, this was the temperature of induction selected to advance to the optimization of the fourth parameter under analysis, the time point of induction.

To evaluate the influence of this fourth parameter on the production in *E. coli* BL21(DE3), protein expression was induced in separate experiments using two alternative time points (corresponding to

distinct values of $OD_{600\ nm}$ ≈ 0.2 and $OD_{600\ nm}$ ≈ 1.8) to the standard time point of induction at an $OD_{600\ nm}$ of 0.6 (**III.2.12. Optimization of production of MjsHSP nanocages**).

The remaining operating parameters were maintained as previously established (non-baffled shake flask with 30% volume of culture medium; enhanced 2xYT; 37°C and 250 rpm; induction of protein expression with 0.1 mM of IPTG, for 4 hours at 37°C).

The growth curves obtained for the two time points of protein expression induction evaluated are shown in **Figure III.19**. For comparative purposes, the standard time point ($OD_{600\ nm}$ ≈ 0.6) was also included.

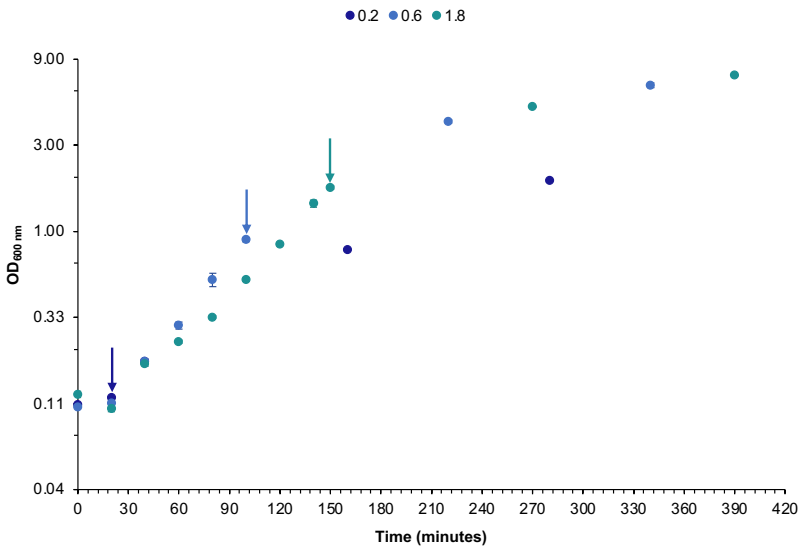


Figure III.19 – Growth curves obtained for the *E. coli* BL21(DE3) cells producing MjsHSP nanocages with the induction of protein expression using two different time points, maintaining the remaining parameters as previously established (non-baffled shake flask with 30% volume of culture medium; enhanced 2xYT; 37°C and 250 rpm; induction of protein expression with 0.1 mM of IPTG, for 4 hours at 37°C). Standard time point for induction of protein expression ($OD_{600\ nm}$ ≈ 0.6) was also included as a reference point. Arrows indicate IPTG induction time points. Each data point represents the mean of two biological replicates, with error bars indicating the respective standard deviation. The y-axis is on a logarithmic scale.

Observing **Figure III.19**, the growth profile for *E. coli* BL21(DE3) cells in independent experiments for induction at two distinct time points (early exponential phase and late exponential/early stationary phase) was equivalent up to this moment. This similarity was also verified in relation to the standard timepoint ($OD_{600\ nm}$ ≈ 0.6). This result can be considered expected because this fourth parameter under study has a more relevant impact on the production of protein nanocages and substantially on the behavior of the growth curve after the time point of induction. The three growth parameters studied (specific growth rate, generation time, and maximum $OD_{600\ nm}$ achieved) are graphically represented in **Figure III.20**.

The specific growth rate (**Figure III.20A**) and the generation time (**Figure III.20B**) values obtained for the two evaluated conditions were consistent with each other and aligned with the standard condition that used the corresponding time point of induction to $OD_{600\ nm}$ ≈ 0.6 (approximately 1.16 to 1.42 h^{-1} for

the specific growth rate and about 29.2 to 33.1 minutes for the generation time). Since these parameters were calculated during the exponential growth phase, which showed no significant differences in the respective growth curves, this outcome was expected and aligns with the observations discussed above.

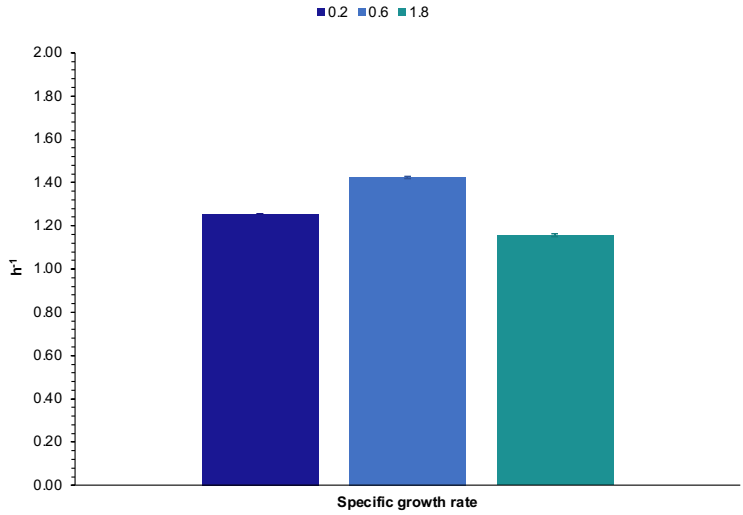
In terms of the maximum $OD_{600\ nm}$ (**Figure III.20C**), it was found that induction at a time point in late exponential/early stationary phase ($OD_{600\ nm} \approx 1.8$) allowed to obtain a growth level 1.1 times higher (7.33 ± 0.02) than the standard condition of induction at middle exponential growth phase (6.45 ± 0.21) and significantly 3.8 times higher than the alternative condition tested of induction at early exponential phase (1.91 ± 0.02). This observation can be explained by the dynamics between cell growth and protein expression at different moments of the exponential growth phase. When induction occurs at the beginning of the exponential phase, there is an early redirection of metabolism and cellular resources to the synthesis of protein nanocages, making biomass production no longer the priority. As a consequence, a reduced $OD_{600\ nm}$ value is observed at the end of the induction process, indicating less cell growth. On the other hand, when the induction of protein expression is performed at a later time point of the exponential growth, close to the beginning of the stationary phase, the cells initially dedicate themselves to the cell growth, resulting in greater biomass production. Only after this point the cells begin to focus their resources on synthesizing protein nanocages.

Then, the production of MjsHSP nanocages in *E. coli* BL21(DE3) cells using two time points of protein expression induction was evaluated. This assessment was performed through SDS-PAGE analysis of the clarified lysates (soluble fraction) containing the MjsHSP nanocages. Distinguishable bands corresponding to the MjsHSP nanocages, with an expected molecular weight of 16.5 kDa, were visible across all the experiments conducted, as illustrated in **Figure III.21**. Given that SDS-PAGE is a denaturing method, these bands reflect the monomeric form of the protein, regardless of its original oligomeric structure in native conditions. The intensity of the MjsHSP nanocages band appeared to be much stronger with the induction of protein expression occurring in the late exponential growth phase, which suggests a relationship with a greater amount of protein nanocages expressed. Compared to the standard time point of induction ($OD_{600\ nm} \approx 0.6$), this band intensity in the SDS-PAGE was much higher (**Figure III.13A**). From **Figure III.21**, it is also possible to observe a greater intensity of the bands corresponding to protein impurities, although their profile was equivalent for all tested conditions.

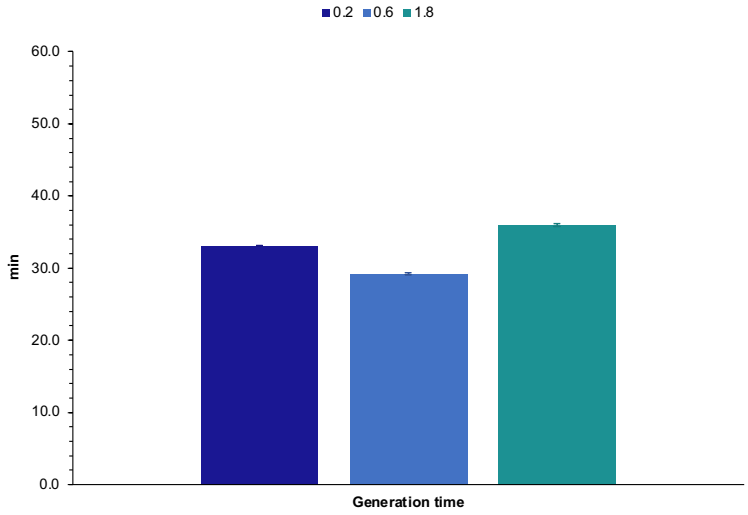
To assess and compare the efficiency of MjsHSP nanocages production in *E. coli* BL21(DE3) under the two alternative time points of protein expression induction, key parameters such as production yield (**Equation III.5**), volumetric productivity (**Equation III.6**) and specific productivity (**Equation III.7**) were estimated. The analysis was performed using densitometric measurements from SDS-PAGE, as previously outlined. The results for these three parameters are presented in **Figure III.22**, providing a clear comparison of production efficiency across the tested conditions.

From **Figure III.22**, the induction of the expression of MjsHSP nanocages at a time point corresponding to the middle of the exponential growth phase ($OD_{600\ nm} \approx 0.6$), the standard condition, allowed to obtain higher values of production yield and volumetric productivity in comparison with the results of the alternative conditions at the early exponential phase ($OD_{600\ nm} \approx 0.2$) and at the late exponential/early stationary phase ($OD_{600\ nm} \approx 1.8$).

(A)


0.2
0.6
1.8
2.00
1.80
1.60
1.40
1.20
1.00
0.80
0.60
0.40
0.20
0.00
h-1
Specific growth rate

(B)


0.2
0.6
1.8
60.0
50.0
40.0
30.0
20.0
10.0
0.0
min
Generation time

(C)

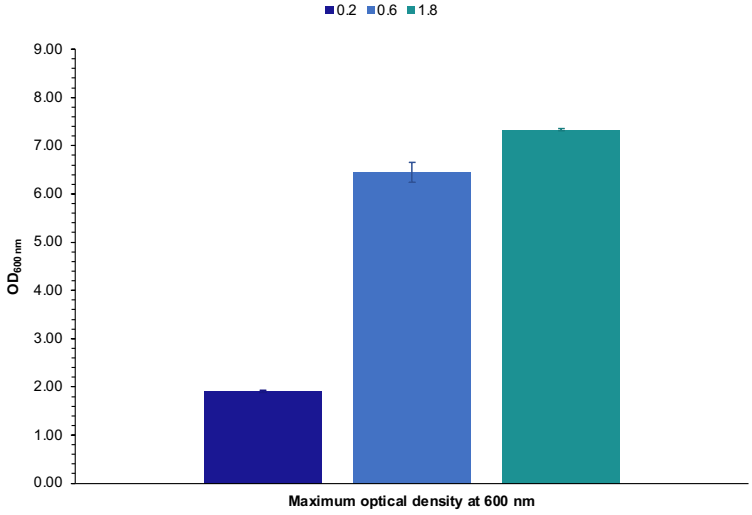

0.2
0.6
1.8
9.00
8.00
7.00
6.00
5.00
4.00
3.00
2.00
1.00
0.00
OD600 nm
Maximum optical density at 600 nm

Figure III.20 – Growth parameters of *E. coli* BL21(DE3) producing MjsHSP nanocages with the induction of protein expression using two different time points, maintaining the remaining parameters as previously established (non-baffled shake flask with 30% volume of culture medium; enhanced 2xYT; 37°C and 250 rpm; induction of protein expression with 0.1 mM of IPTG, for 4 hours at 37°C). **(A)** Specific growth rate, **(B)** generation time, and **(C)** maximum $OD_{600\ nm}$. Standard time point for induction of protein expression ($OD_{600\ nm} \approx 0.6$) was also included as a reference point. Each result represents the mean of two biological replicates, with error bars indicating the respective standard deviation.

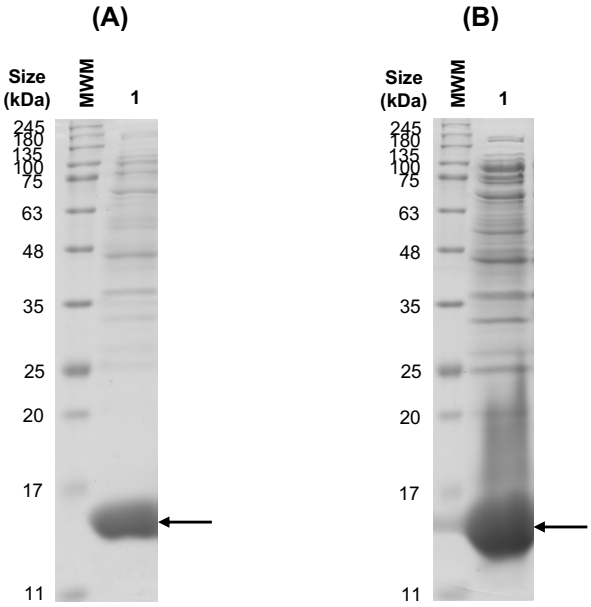
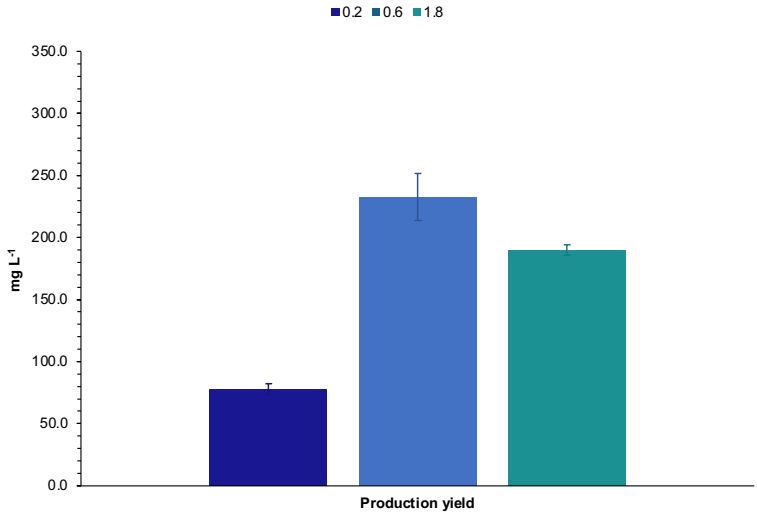


Figure III.21 – SDS-PAGE analysis obtained in the production of MjsHSP nanocages in *E. coli* BL21(DE3) cells using two different time points for the induction of protein expression, maintaining the remaining parameters as previously established (non-baffled shake flask with 30% volume of culture medium; enhanced 2xYT; 37°C and 250 rpm; induction of protein expression with 0.1 mM of IPTG, for 4 hours at 37°C). **(A)** $OD_{600\ nm} \approx 0.2$, and **(B)** $OD_{600\ nm} \approx 1.8$. Arrows indicate the protein bands referring to the monomer of MjsHSP nanocages (expected molecular weight of 16.5 kDa), which is separated and visualized under denaturing conditions of SDS-PAGE. The abbreviation MWM refers to the molecular weight marker (NZYColour protein marker II; NZYtech Lisbon, Portugal).

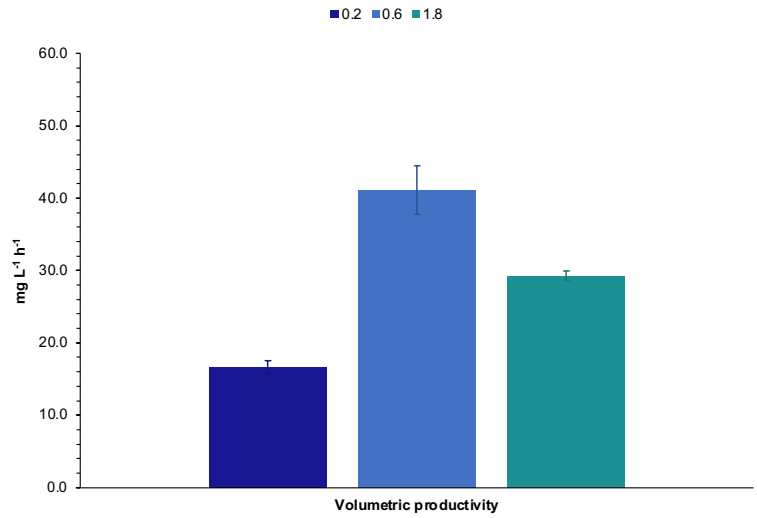
The production yield (**Figure III.22A**) was about 1.2 times higher with time point of induction at $OD_{600\ nm} \approx 0.6$ (232.9 ± 18.8 mg L^{-1}) compared to the second best time point ($OD_{600\ nm} \approx 1.8$; 190.2 ± 4.2 mg L^{-1}), and about 3.0 times higher than the value obtained at the time point of $OD_{600\ nm} \approx 0.2$ (77.5 ± 4.4 mg L^{-1}).

Similarly, for the volumetric productivity (**Figure III.22B**), the induction of MjsHSP nanocages expression at a time point of induction at $OD_{600\ nm} \approx 0.6$ (41.1 ± 3.3 mg L^{-1} h^{-1}) allowed to obtain a result 1.4 and 2.5 times higher than that determined for the time point at $OD_{600\ nm} \approx 1.8$ (29.3 ± 0.7 mg L^{-1} h^{-1}) and at $OD_{600\ nm} \approx 0.2$ (16.6 ± 1.0 mg L^{-1} h^{-1}), respectively. Interestingly, between the two alternatives tested, induction at a very early stage of *E. coli* BL21(DE3) cell growth ends up negatively influencing the expression of MjsHSP nanocages.

(A)


0.2
0.6
1.8
350.0
300.0
250.0
200.0
150.0
100.0
50.0
0.0
mg L-1
Production yield

(B)


0.2
0.6
1.8
60.0
50.0
40.0
30.0
20.0
10.0
0.0
mg L-1 h-1
Volumetric productivity

(C)

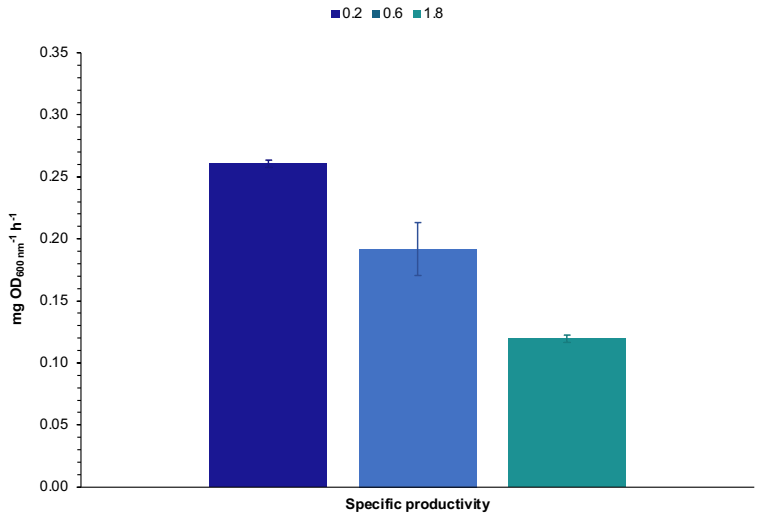

0.2
0.6
1.8
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
mg OD600 nm-1 h-1
Specific productivity

Figure III.22 – Parameters of MjsHSP nanocages production obtained for *E. coli* BL21(DE3) cells using two different time points for the induction of protein expression, maintaining the remaining parameters as previously established (non-baffled shake flask with 30% volume of culture medium; enhanced 2xYT; 37°C and 250 rpm; induction of protein expression with 0.1 mM of IPTG, for 4 hours at 37°C). **(A)** Production yield, **(B)** volumetric productivity, and **(C)** specific productivity. Standard time point for induction of protein expression ($OD_{600\ nm} \approx 0.6$) was also included as a reference point. Each result represents the mean of two biological replicates, with error bars indicating the respective standard deviation.

In terms of the specific productivity (**Figure III.22B**), the results showed that induction at an $OD_{600\ nm} \approx 0.2$ (0.26 ± 0.01 mg $OD_{600\ nm}^{-1}$ h^{-1}) allowed a protein production per cell to be much higher than later induction at $OD_{600\ nm} \approx 1.8$ (0.12 ± 0.00 mg $OD_{600\ nm}^{-1}$ h^{-1}). The standard condition ($OD_{600\ nm} \approx 0.6$; 0.19 ± 0.02 mg $OD_{600\ nm}^{-1}$ h^{-1}) was able to obtain a more favorable equilibrium for the production of protein nanocages considering the several parameters.

Framing these results with the research described in the literature, the induction of protein expression at an early stage of growth typically leads to rapid initial protein production due to increased metabolic activity. However, this situation may result in a significant reduction in protein expression as cells transition into later growth phases, mainly due to resource exhaustion and stress responses (227). This interpretation may be related to the results obtained for the induction of protein expression at $OD_{600\ nm} \approx 0.2$, both in terms of maximum biomass production (lower compared to the alternative conditions) and protein expression (lower values of production yield and volumetric productivity). As for induction at a later phase of exponential growth, closer to the beginning of the stationary phase, protein expression occurs in a less efficient manner due to resource depletion and changes in cellular physiology as cells prepare for stationary period (228). In this work, induction at a later stage of the exponential growth phase led to a very expressive growth followed by a reasonable protein expression. On the other hand, the exponential growth phase is characterized by a high metabolic activity and a robust transcriptional and translational machinery, which provides optimal conditions for protein nanocages production. In this sense, cellular resources are still abundant, and metabolic burden is manageable by the bacterial cells (227,228). This interpretation contextualizes well the results obtained for the induction at the time point of $OD_{600\ nm} \approx 0.6$, being the best condition evaluated.

Thus, it can be concluded that inducing the expression of MjsHSP nanocages at a time point in the middle of the exponential growth phase ($OD_{600\ nm}$ of approximately 0.5 to 0.6) appears to be the most advantageous. Therefore, in combination with the use of enhanced 2xYT culture medium, an IPTG concentration of 0.1 mM and an induction temperature of 37°C, this time point of induction was selected to proceed with the optimization of the fifth and last parameter under analysis, the aeration conditions through the presence or absence of baffles in the culture shake flask.

The influence of this fifth and last parameter on the production of MjsHSP nanocages in *E. coli* BL21(DE3) was evaluated through two separate experiments, one assuming more favorable aeration conditions (shake flask with baffles) and the other considering an aeration condition closer to the standard (shake flasks without baffles). To allow a more reliable comparison, baffled and non-baffled

shake flasks with the same volume of 250 mL were used, maintaining 30% volume of culture medium (**III.2.12. Optimization of production of MjsHSP nanocages**).

The remaining operating parameters were maintained as previously established (enhanced 2xYT; 37°C and 250 rpm; induction of protein expression at $OD_{600\ nm}$ of 0.5-0.6, with 0.1 mM of IPTG, for 4 hours at 37°C).

The growth curves obtained for the two aeration conditions studied are shown in **Figure III.23**. For comparative purposes, the standard aeration condition (non-baffled shake flask with 100 mL and 30% volume of culture medium) was also included.

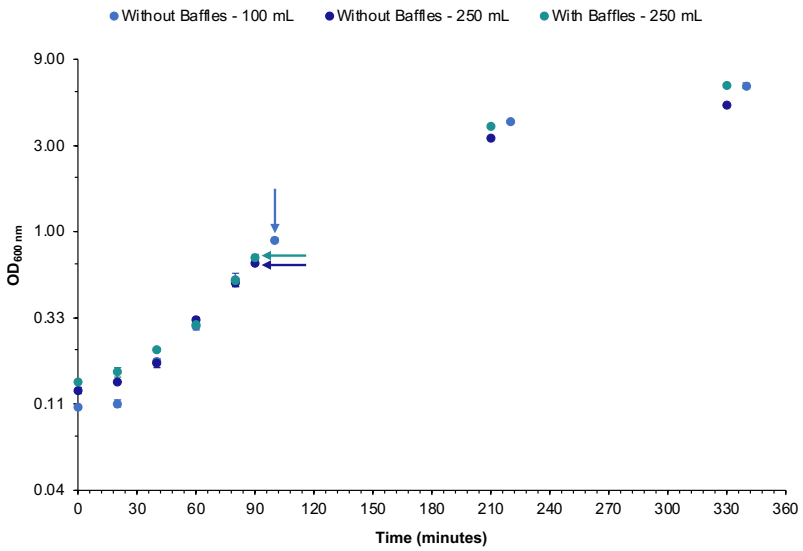


Figure III.23 – Growth curves obtained for the *E. coli* BL21(DE3) cells producing MjsHSP nanocages with the induction of protein expression using two different aeration conditions (250 mL shake flask without and with baffles), maintaining the remaining parameters as previously established (enhanced 2xYT; 37°C and 250 rpm; induction of protein expression at $OD_{600\ nm}$ of 0.5-0.6, with 0.1 mM of IPTG, for 4 hours at 37°C). Standard aeration condition for induction of protein expression (100 mL shake flask without baffles) was also included as a reference point. Arrows indicate IPTG induction time points. Each data point represents the mean of two biological replicates, with error bars indicating the respective standard deviation. The y-axis is on a logarithmic scale.

Analyzing **Figure III.23**, the growth profile for *E. coli* BL21(DE3) cells in both the two alternative aeration conditions and the condition defined as reference was similar in the various stages of the cell growth. The three growth parameters assessed (specific growth rate, generation time, and maximum $OD_{600\ nm}$ achieved) are graphically represented in **Figure III.24**.

The specific growth rate (**Figure III.24A**) and the generation time (**Figure III.24B**) values obtained for the two evaluated conditions were consistent with each other and aligned with the standard condition that used the non-baffled shake flasks with 100 mL (approximately 1.38 to 1.55 h^{-1} for the specific growth rate and about 27.0 to 30.0 minutes for the generation time). A very slight advantage was seen for the experiment with the baffled shake flasks with 250 mL (1.55 ± 0.11 h^{-1} and 27.0 ± 1.9 minutes, respectively). This is a curious result given that potentially baffled shake flasks would allow a greater supply of oxygen, which has the consequence of supporting a faster growth in particular during the

exponential phase and the metabolism of the *E. coli* cells (229). It is also worth noting that in the experiment with the baffled shake flask, greater foam formation was observed, which may potentially interfere with the cell growth.

Regarding the maximum $OD_{600\ nm}$ (**Figure III.24C**), the test using a 250 mL baffled shake flask (6.51 ± 0.14) allowed to obtain a cell density identical to that of the reference aeration condition (6.45 ± 0.21), both values being higher (1.3 times for both parameters) than those of the condition with the 250 mL non-baffled shake flask (5.05 ± 0.02). This is a result that fits with that described in other studies in which baffles increase oxygen transfer rate compared to the non-baffled shake flasks, which can lead to higher cell densities (230,231).

Next, the production of MjsHSP nanocages in *E. coli* BL21(DE3) cells using the two alternative aeration conditions was assessed. This evaluation was performed through SDS-PAGE analysis of the soluble fraction from the clarified lysates containing this model of protein nanocages. Distinct bands corresponding to the MjsHSP nanocages, with an expected molecular weight of 16.5 kDa, are visible across all the experiments conducted, as illustrated in **Figure III.25**. Considering that SDS-PAGE is a denaturing method, these bands reflect the monomeric form of the protein nanocage, regardless of its native oligomeric structure with 24 monomers. The band intensity of the MjsHSP nanocages appeared to be slightly stronger with their production in the 250 mL non-baffled shake flask, although the difference was not very significant. On the other hand, the pattern and intensity of the bands corresponding to protein impurities was quite similar. Compared to the standard condition, in particular the growth and the production of protein nanocages in the 100 mL non-baffled shake flask, the intensity of the bands that composes the entire cell lysate was less pronounced (**Figure III.13A**).

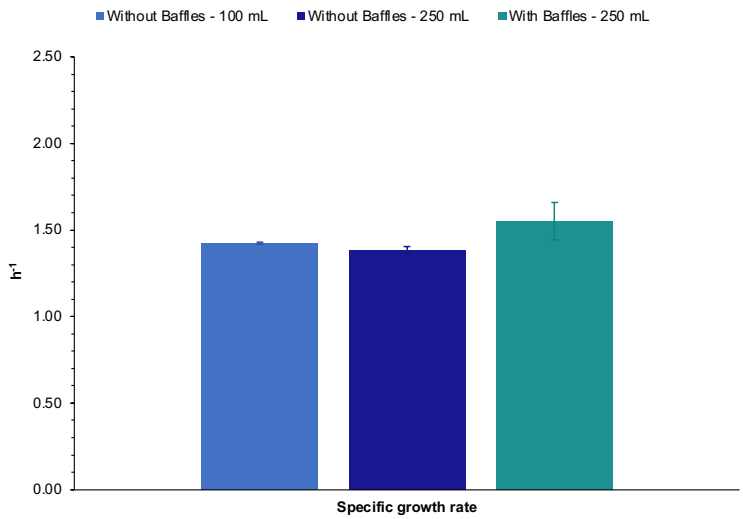
In order to compare and validate the efficiency of MjsHSP nanocages production in *E. coli* BL21(DE3) under the two alternative aeration conditions, key parameters namely production yield (**Equation III.5**), volumetric productivity (**Equation III.6**) and specific productivity (**Equation III.7**) were calculated. The analysis was performed using SDS-PAGE densitometric measurements as previously explained. The results for these parameters are presented in **Figure III.26**.

From **Figure III.26**, the expression of the MjsHSP nanocages in both a 250 mL non-baffled and a 250 mL baffled shake flask allowed to obtain identical values of production yield and volumetric productivity. Compared to the standard condition, interestingly, the values of both parameters were much higher, which may be related to the volume of culture medium being lower (around 30 mL), although the air volume/culture medium volume ratio was always considered.

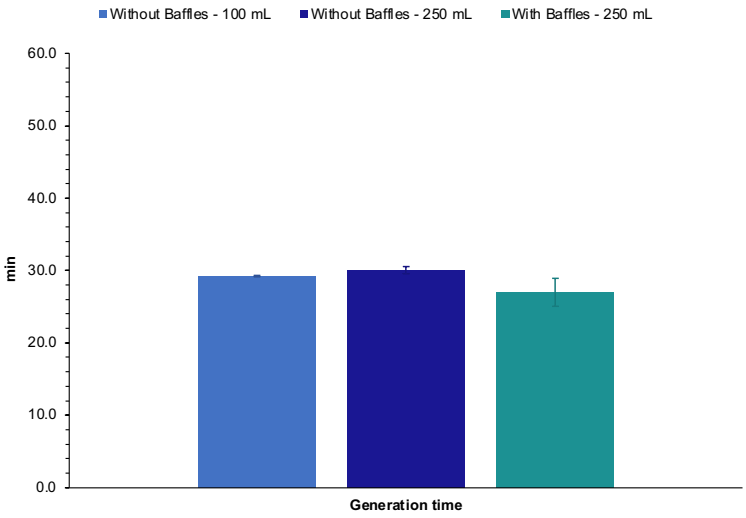
The production yield (**Figure III.26A**) was 1.4 times lower in non-baffled (167.8 ± 1.4 mg L^{-1}) and baffled shake flasks (170.1 ± 9.8 mg L^{-1}) both with 250 mL in relation to the non-baffled shake flask of 100 mL (232.9 ± 18.8 mg L^{-1}).

Similarly, the volumetric productivity (**Figure III.26B**) showed the same behavior, being 1.3 times lower in non-baffled (30.5 ± 0.3 mg L^{-1} h^{-1}) and baffled shake flasks (30.9 ± 1.8 mg L^{-1} h^{-1}) with 250 mL comparing with the non-baffled shake flask of 100 mL (41.1 ± 3.3 mg L^{-1} h^{-1}).

(A)


Without Baffles - 100 mL
Without Baffles - 250 mL
With Baffles - 250 mL
2.50
2.00
1.50
1.00
0.50
0.00
h-1
Specific growth rate

(B)


Without Baffles - 100 mL
Without Baffles - 250 mL
With Baffles - 250 mL
60.0
50.0
40.0
30.0
20.0
10.0
0.0
min
Generation time

(C)

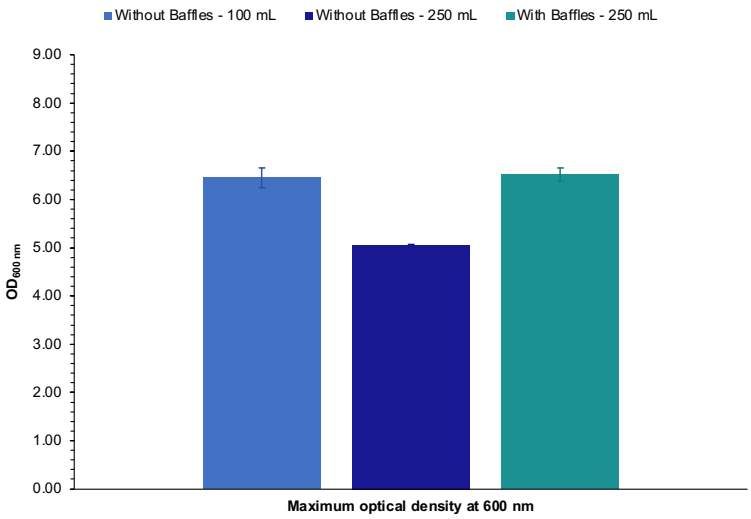

Without Baffles - 100 mL
Without Baffles - 250 mL
With Baffles - 250 mL
9.00
8.00
7.00
6.00
5.00
4.00
3.00
2.00
1.00
0.00
OD600 nm
Maximum optical density at 600 nm

Figure III.24 – Growth parameters of *E. coli* BL21(DE3) producing MjsHSP nanocages with the induction of protein expression using two different aeration conditions (250 mL shake flask without and with baffles), maintaining the remaining parameters as previously established (enhanced 2xYT; 37°C and 250 rpm; induction of protein expression at $OD_{600\ nm}$ of 0.5-0.6, with 0.1 mM of IPTG, for 4 hours at 37°C). **(A)** Specific growth rate, **(B)** generation time, and **(C)** maximum $OD_{600\ nm}$. Standard aeration condition for induction of protein expression (100 mL shake flask without baffles) was also included as a reference point. Each result represents the mean of two biological replicates, with error bars indicating the respective standard deviation.

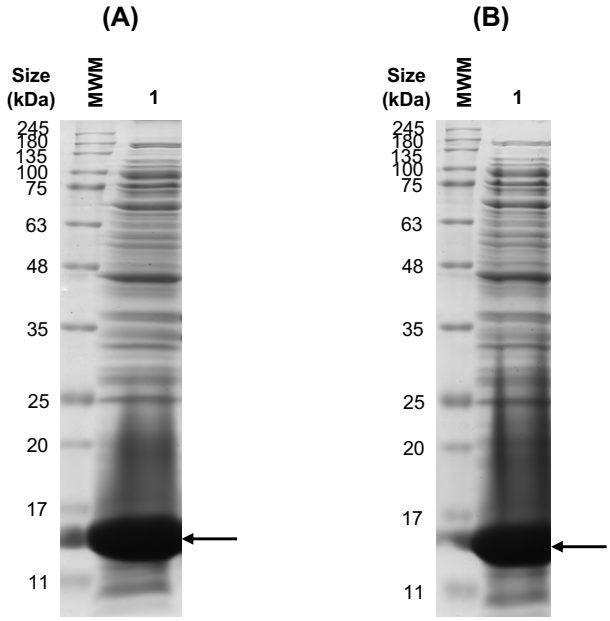
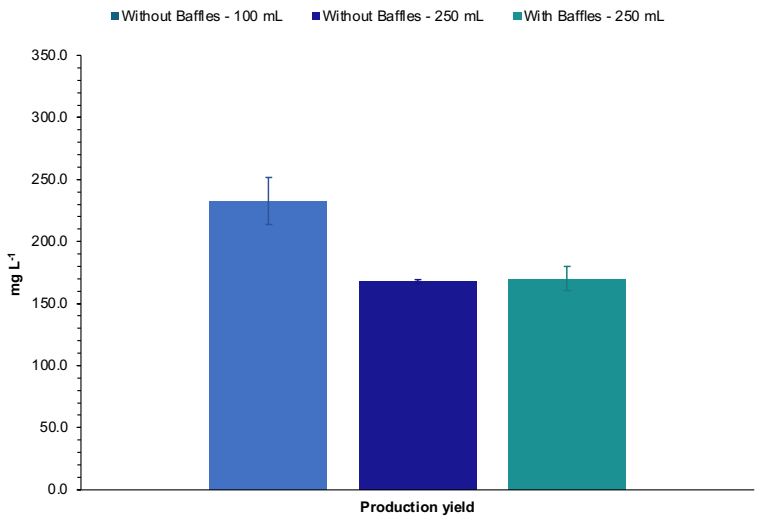


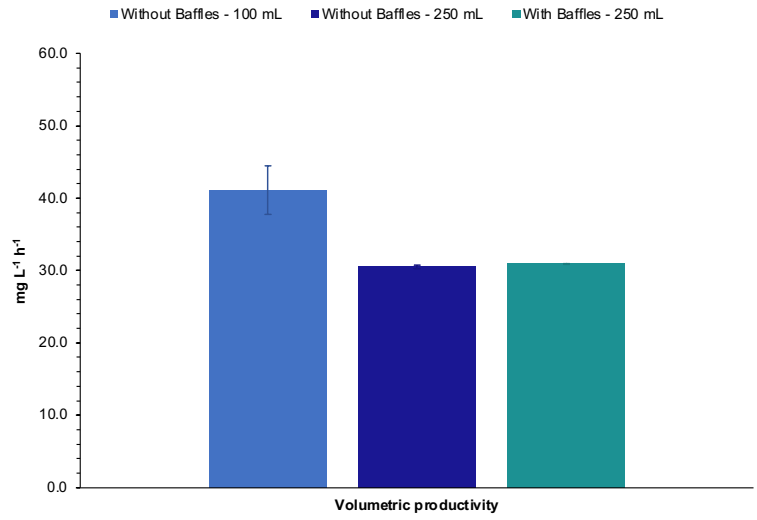
Figure III.25 – SDS-PAGE analysis obtained in the production of MjsHSP nanocages in *E. coli* BL21(DE3) cells using two different aeration conditions (250 mL shake flask without and with baffles) for the induction of protein expression, maintaining the remaining parameters as previously established (enhanced 2xYT; 37°C and 250 rpm; induction of protein expression at $OD_{600\ nm}$ of 0.5-0.6, with 0.1 mM of IPTG, for 4 hours at 37°C). **(A)** 250 mL shake flask without baffles, and **(B)** 250 mL shake flask with baffles. Arrows indicate the protein bands referring to the monomer of MjsHSP nanocages (expected molecular weight of 16.5 kDa), which is separated and visualized under denaturing conditions of SDS-PAGE. The abbreviation MWM refers to the molecular weight marker (NZYColour protein marker II; NZYtech Lisbon, Portugal).

In terms of the specific productivity (**Figure III.26C**), the results show that MjsHSP nanocages production per cell was much higher in the experiment using the 250 mL non-baffled shake flask (0.45 $\pm$ 0.01 mg $OD_{600\ nm}^{-1}$ h^{-1}), approximately 1.3 times higher than the value obtained for the 250 mL baffled shake flask (0.36 $\pm$ 0.01 mg $OD_{600\ nm}^{-1}$ h^{-1}) and 2.4 times higher than the estimation of the reference condition (0.19 $\pm$ 0.02 mg $OD_{600\ nm}^{-1}$ h^{-1}).

(A)


Without Baffles - 100 mL
Without Baffles - 250 mL
With Baffles - 250 mL
350.0
300.0
250.0
200.0
150.0
100.0
50.0
0.0
mg L-1
Production yield

(B)


Without Baffles - 100 mL
Without Baffles - 250 mL
With Baffles - 250 mL
60.0
50.0
40.0
30.0
20.0
10.0
0.0
mg L-1 h-1
Volumetric productivity

(C)

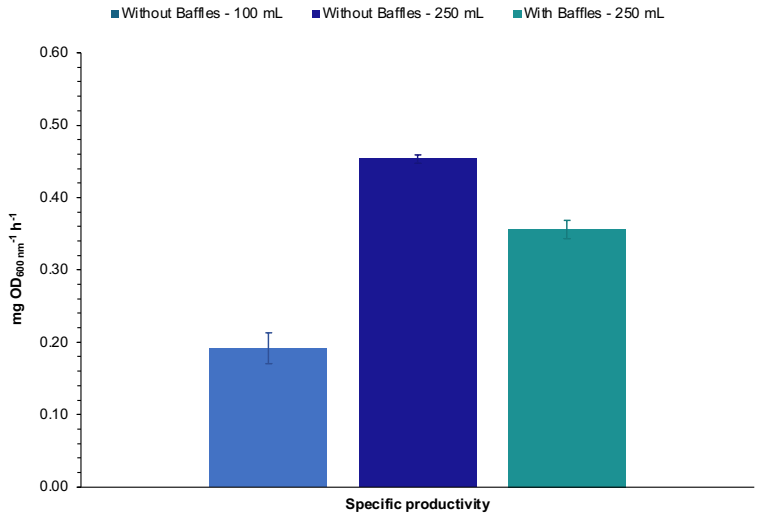

Without Baffles - 100 mL
Without Baffles - 250 mL
With Baffles - 250 mL
0.60
0.50
0.40
0.30
0.20
0.10
0.00
mg OD600 nm-1 h-1
Specific productivity

Figure III.26 – Parameters of MjsHSP nanocages production obtained for *E. coli* BL21(DE3) cells using two different aeration conditions (250 mL shake flask without and with baffles) for the induction of protein expression, maintaining the remaining parameters as previously established (enhanced 2xYT; 37°C and 250 rpm; induction of protein expression at $OD_{600\ nm}$ of 0.5-0.6, with 0.1 mM of IPTG, for 4 hours at 37°C). **(A)** Production yield, **(B)** volumetric productivity, and **(C)** specific productivity. Standard aeration condition for induction of protein expression (100 mL shake flask without baffles) was also included as a reference point. Each result represents the mean of two biological replicates, with error bars indicating the respective standard deviation.

The data described in the literature indicate that the relationship between the use of baffled shake flasks and their impact on protein production depends on several factors that may interact with each other. In general, since the increase in oxygenation with the presence of baffles in the shake flasks helps to obtain greater quantities of biomass, this will consequently lead to an increase in the protein nanocages produced (232). Nevertheless, as this association is not linear, results such as those obtained in this work may be observed, in which the production efficiency of MjsHSP nanocages was equivalent for both aeration conditions.

Then, it can be concluded that inducing the expression of MjsHSP nanocages using a non-baffled shake flask (exemplified in this work by one of 250 mL), maintaining the same ratio of air volume to culture medium volume (about 30% of the latter) appears to have greater potential. Therefore, the final conditions for optimizing the production of this model protein nanocages were the use of enhanced 2xYT culture medium, an IPTG concentration of 0.1 mM, an induction temperature of 37°C, a time point in the middle of the exponential growth phase ($OD_{600\ nm}$ of approximately 0.5 to 0.6) and an aeration setup comprising a non-baffled shake flask.

After identifying these optimized conditions for the production of MjsHSP nanocages in *E. coli* BL21(DE3) cells, an additional experiment was performed with the monitoring and collection of samples of the bacterial culture after 2, 4, 6, 8 and 24 hours from the beginning of the induction of protein expression (**III.2.12. Optimization of production of MjsHSP nanocages**). The objective was to evaluate the parameters of cell growth and production of protein nanocages over a longer period of time, allowing comparison with the results obtained before for the initial conditions defined as standard.

The growth curve obtained for the production of MjsHSP nanocages in *E. coli* BL21(DE3) under the optimized conditions is shown in **Figure III.27**. Note that the growth curve for the production of these protein nanocages under the standard initial conditions is available in **Figure III.5**.

Observation of **Figure III.27** allowed to conclude that although it is necessary to take into account that the temporal periods were different, the growth profile of *E. coli* BL21(DE3) cells under the optimized conditions and under the standard condition was very similar, particularly in the exponential growth phase. As previously described throughout the work, from this growth curve, it was possible to determine the three growth parameters: specific growth rate, generation time, and maximum $OD_{600\ nm}$ achieved at the end of the protein expression induction period.

The values obtained for the specific growth rate ($1.37 \pm 0.01\ h^{-1}$) and the generation time (30.3 ± 0.2 minutes) were very close to the results determined for the standard conditions (**Table III.3**), with a slightly faster growth in the reference experiment (about 1.1 times for both parameters), although the difference was not significant.

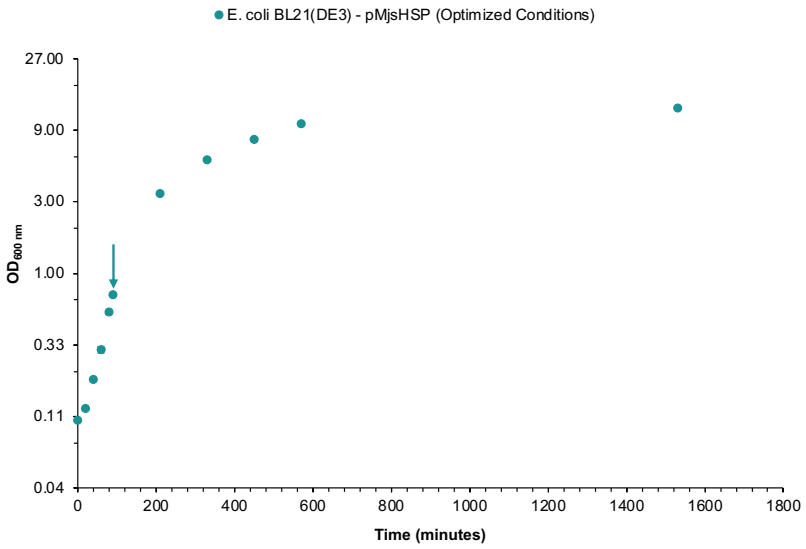


Figure III.27 – Growth curve obtained for the *E. coli* BL21(DE3) cells producing MjsHSP nanocages in the optimized conditions (enhanced 2xYT culture medium, 0.1 mM of IPTG, 37°C during protein expression, time point of induction at $OD_{600\ nm} \approx 0.6$ and shake flask without baffles). Arrow indicates IPTG induction time point. Each data point represents the mean of two biological replicates, with error bars indicating the respective standard deviation. The y-axis is on a logarithmic scale.

Regarding the maximum $OD_{600\ nm}$, the growth of *E. coli* BL21(DE3) cells in the optimized condition led to a much higher value (12.63 ± 0.09) if the period up to 24 hours of protein expression induction is considered, which was expected. In any case, if the values for the same induction period (4 hours) were compared, the experiment under optimized conditions (5.70 ± 0.05) also allowed obtaining a value 1.7 times higher than the standard test (**Table III.3**).

Thus, with regard to the characterization of cell growth, this comparison allowed to conclude that the optimized conditions ensure the maintenance of a promising cell growth profile, with potential for the production of MjsHSP nanocages on a large scale.

Afterwards, the production of this model of non-viral protein nanocages was assessed by obtaining an SDS-PAGE of clarified lysates (soluble fraction) containing the MjsHSP nanocages expressed in *E. coli* cells using these optimized conditions for cultivation and protein expression. **Figure III.28** shows clear bands corresponding to the MjsHSP nanocages (expected molecular weight of 16.5 kDa). As expected, the intensity of the MjsHSP nanocages band increased throughout the induction period from the IPTG addition time point, suggesting that there continued to be significant expression of the protein nanocages as cell growth progressed. It was interesting to note that the basal expression was quite reduced (**Figure III.28**, Lane 1) and that the greatest increase in the band intensity occurred after 2 hours (**Figure III.28**, Lane 2), 4 hours (**Figure III.28**, Lane 3) and 6 hours (**Figure III.28**, Lane 4) of induction, with the differences from this moment on being less considerable. The bands corresponding to protein impurities always had a similar pattern, increasing in intensity as the experiment progressed, although they maintained a similar proportion in relation to the MjsHSP nanocages band.

To understand the differences in the efficiency of MjsHSP nanocages production in *E. coli* BL21(DE3) using these optimized conditions, the aforementioned main parameters, production yield (**Equation III.5**), volumetric productivity (**Equation III.6**) and specific productivity (**Equation III.7**) were determined. Analysis was performed using SDS-PAGE densitometric measurements as previously explained. The results for these parameters are presented in **Figure III.29**.

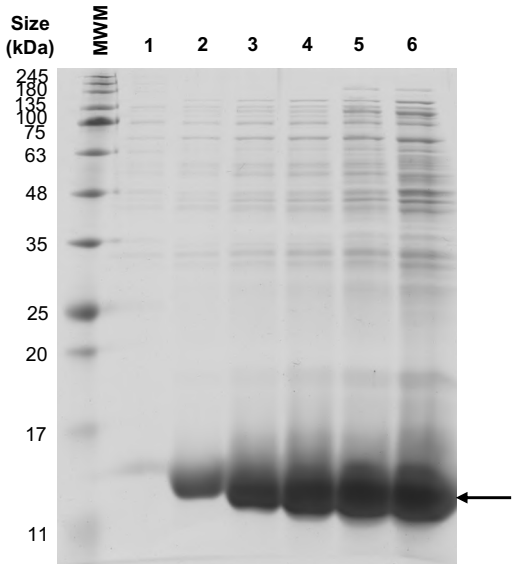
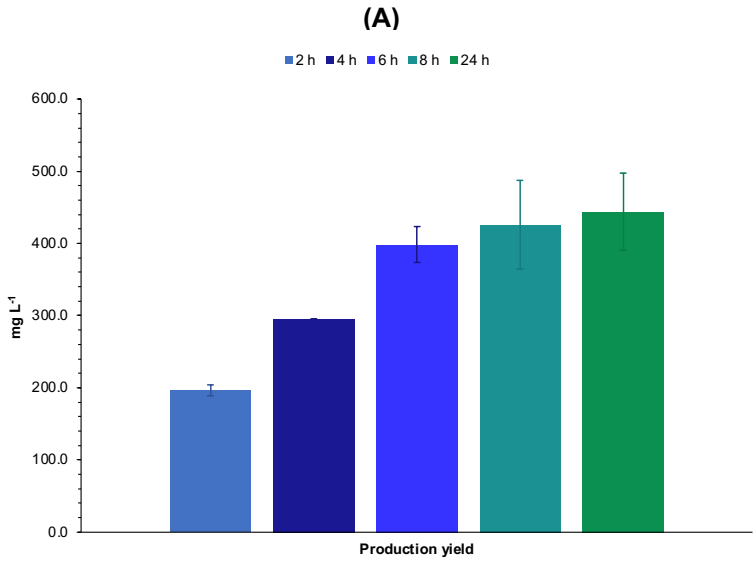
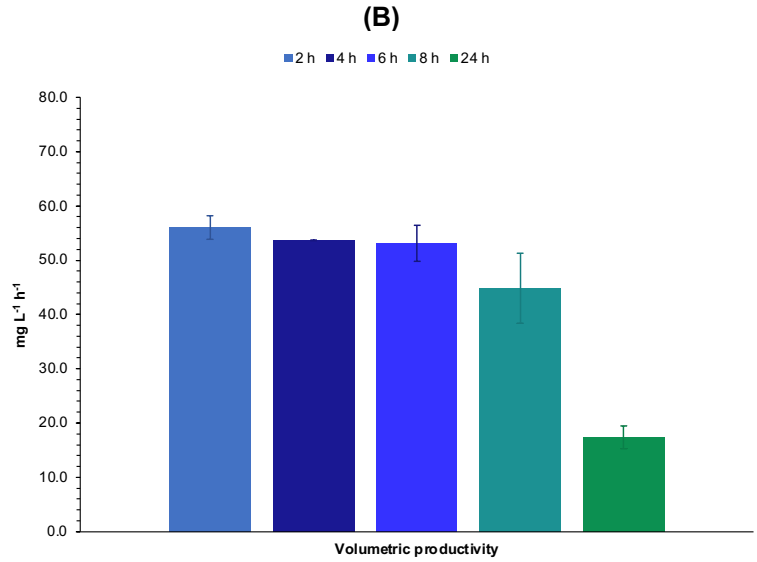


Figure III.28 – SDS-PAGE analysis obtained in the production of MjsHSP nanocages in *E. coli* BL21(DE3) cells using the optimized conditions for the induction of protein expression (enhanced 2xYT culture medium, 0.1 mM of IPTG, 37°C during protein expression, time point of induction at $OD_{600\ nm} \approx 0.6$ and shake flask without baffles). Lane 1 corresponds to the sample collected from the bacterial culture at the IPTG induction time point and lanes 2 to 6 correspond to the samples collected after 2, 4, 6, 8 and 24 hours, respectively, from the beginning of the protein expression induction. Arrow indicates the protein bands referring to the monomer of MjsHSP nanocages (expected molecular weight of 16.5 kDa), which is separated and visualized under denaturing conditions of SDS-PAGE. The abbreviation MWM refers to the molecular weight marker (NZYColour protein marker II; NZYtech Lisbon, Portugal).

According to **Figure III.29A**, for the production yield, a gradual increase was observed throughout the monitored protein expression period. The most considerable increases occurred after 2 hours (18.5 times, for 196.1 ± 7.6 mg L^{-1}), 4 hours (1.5 times, for 295.7 ± 0.0 mg L^{-1}) and 6 hours (1.3 times, for 398.5 ± 24.6 mg L^{-1}), results compatible with the fact that *E. coli* BL21(DE3) cells began to enter in the late exponential growth phase after this initial time, with a slowed metabolism and less directed towards the expression of MjsHSP nanocages, also due to the reduction of available resources in the culture medium.

The volumetric productivity (**Figure III.29B**) recorded values that were in agreement with this last interpretation, since the results after 2 hours (56.0 ± 2.2 mg L^{-1} h^{-1}), 4 hours (53.8 ± 0.0 mg L^{-1} h^{-1}) and 6 hours (53.1 ± 3.3 mg L^{-1} h^{-1}) remained stable and that this parameter verified a considerable decrease after 8 hours (1.2 times, for 44.8 ± 6.5 mg L^{-1} h^{-1}) and mainly after 24 hours (3.1 times, for 17.4 ± 2.1 mg L^{-1} h^{-1}).


(A)
2 h
4 h
6 h
8 h
24 h
600.0
500.0
400.0
300.0
200.0
100.0
0.0
mg L-1
Production yield


(B)
2 h
4 h
6 h
8 h
24 h
80.0
70.0
60.0
50.0
40.0
30.0
20.0
10.0
0.0
mg L-1 h-1
Volumetric productivity

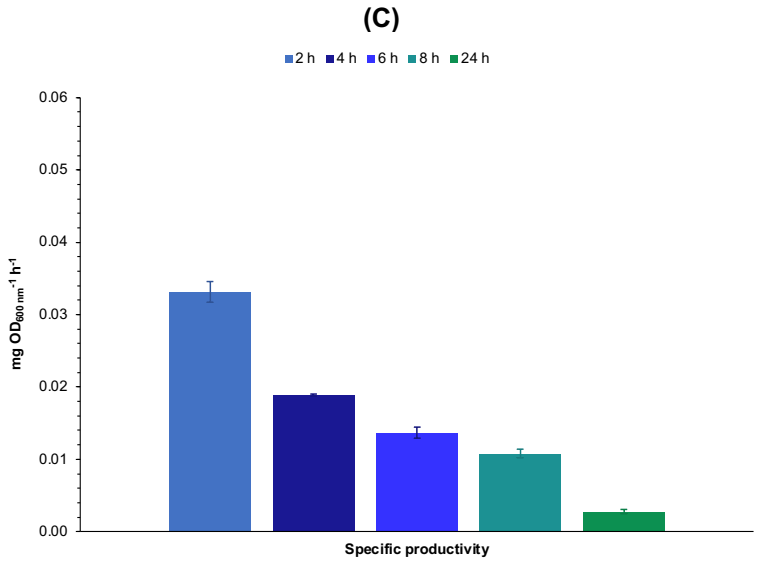

(C)
2 h
4 h
6 h
8 h
24 h
0.06
0.05
0.04
0.03
0.02
0.01
0.00
mg OD600 nm-1 h-1
Specific productivity

Figure III.29 – Parameters of MjsHSP nanocages production obtained for *E. coli* BL21(DE3) cells using the optimized conditions for the induction of protein expression (enhanced 2xYT culture medium, 0.1 mM of IPTG, 37°C during protein expression, time point of induction at $OD_{600\ nm} \approx 0.6$ and shake flask without baffles). **(A)** Production yield, **(B)** volumetric productivity, and **(C)** specific productivity determined considering the samples collected from the bacterial culture after 2, 4, 6, 8 and 24 hours from the beginning of the protein expression induction. Each result represents the mean of two biological replicates, with error bars indicating the respective standard deviation.

Comparing the value obtained after 4 hours of induction with the same time point of the standard condition (**Table III.4**), the process optimization led to a 5.4-fold increase in the production yield and a 5.1-fold increase in the volumetric productivity.

Finally, the specific productivity (**Figure III.29C**) allowed to evaluate the production of MjsHSP nanocages per cell basis, verifying a very high value after the first 2 hours (0.03 ± 0.00 mg $OD_{600\ nm}^{-1}$ h^{-1}), with a subsequent decrease over time until 24 hours of induction of protein expression (10 times, for 0.003 ± 0.000 mg $OD_{600\ nm}^{-1}$ h^{-1}). At an earlier stage of cell growth, the bacterial cells were metabolically more active with greater availability of resources, which may justify a greater production at the individual level. Again, comparing the value obtained after 4 hours of induction with the same time point of the condition defined as standard (**Table III.4**), the process optimization led to an 4.5-fold reduction in the specific productivity (for 0.02 ± 0.00 mg $OD_{600\ nm}^{-1}$ h^{-1}).

Taken together, these results suggest that a promising time point to finish the production of MjsHSP nanocages is close to 6 hours of induction, since up to this moment it is possible to obtain a favorable balance between the parameters analyzed, also combining with the behavior and characteristics of cell growth.

In conclusion, comparing the several results obtained for the standard and optimized experiments, it was possible to successfully optimize this production process of MjsHSP nanocages in *E. coli* BL21(DE3) cells, with the generation of numerous data that can be used as a starting point for future studies, including on a large scale.

The next part of this work consisted of optimizing the production of MjsHSP nanocages in *V. natriegens* Vmax™ X2.

The first parameter to be studied was the growth medium composition, with the growth and expression of MjsHSP nanocages in *V. natriegens* Vmax™ X2 cells being performed in six culture media with different compositions: LB3, 2xYT, enhanced 2xYT, TB supplemented with V2 salts, M9 minimal supplemented with V2 salts and glucose, and M9 minimal supplemented with V2 salts and sucrose (**III.2.12. Optimization of production of MjsHSP nanocages**).

The remaining operating parameters were maintained as established in the cultivation and protein expression standard (non-baffled shake flask with 30% volume of culture medium; 37°C and 250 rpm; induction of protein expression at $OD_{600\ nm}$ of 0.5-0.6, with 1 mM of IPTG, for 4 hours at 37°C).

The growth curves obtained for the six different culture media evaluated are shown in **Figure III.30**. For comparative purposes, standard culture medium (LB + V2 salts) was also included.

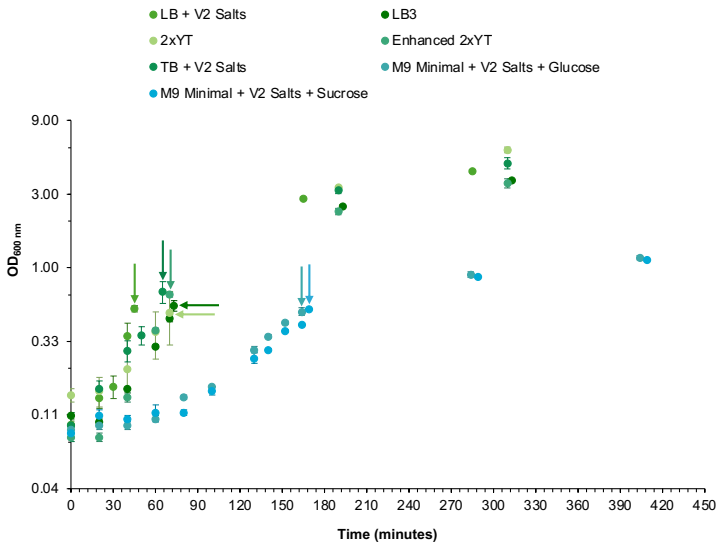


Figure III.30 – Growth curves obtained for the *V. natriegens* Vmax™ X2 cells producing MjsHSP nanocages in six culture media with different compositions, maintaining the remaining parameters as established in the standard cultivation and protein expression (non-baffled shake flask with 30% volume of culture medium; 37°C and 250 rpm; induction of protein expression at $OD_{600\ nm}$ of 0.5-0.6, with 1 mM of IPTG, for 4 hours at 37°C). Standard culture medium (LB + V2 salts) was also included as a reference point. Arrows indicate IPTG induction time points. Each data point represents the mean of two biological replicates, with error bars indicating the respective standard deviation. The y-axis is on a logarithmic scale.

An initial analysis of **Figure III.30** revealed distinct cell growth profiles, primarily contrasting the four enriched culture media (LB3, 2xYT, enhanced 2xYT and TB + V2 salts) with the two minimal culture media (M9 + V2 salts + glucose and M9 + V2 salts + sucrose). This difference was anticipated due to the variable availability of nutrients and other elements essential for the rapid cellular metabolism of *V. natriegens* Vmax™ X2. The analysis of these growth curves enabled the determination of key growth parameters, namely the specific growth rate, the generation time and the maximum $OD_{600\ nm}$ achieved at the end of the protein expression induction period. **Figure III.31** provides a graphical representation of these data.

V. natriegens Vmax™ X2 cells harboring the plasmid pMjsHSP presented the highest specific growth rate in standard LB + V2 salts culture medium ($2.99 \pm 0.10\ h^{-1}$), followed by the enhanced 2xYT ($2.35 \pm 0.13\ h^{-1}$) and LB3 ($1.99 \pm 0.12\ h^{-1}$) culture media (**Figure III.31A**). On the other hand, the lowest growth rate was recorded as expected in the M9 minimal medium supplemented with V2 salts and glucose ($0.85 \pm 0.04\ h^{-1}$), together with the same M9 minimal medium supplemented with V2 salts and sucrose ($0.91 \pm 0.00\ h^{-1}$).

In terms of the generation time (**Figure III.31B**), considering that this is a parameter inversely proportional to the specific growth rate, the lowest values were found for these three media (13.9 ± 0.5 minutes, 17.8 ± 1.0 minutes and 20.9 ± 1.2 minutes, respectively, for LB + V2 salts, enhanced 2xYT and LB3). For the M9 minimal medium supplemented with V2 salts and glucose and with V2 salts and sucrose, significantly high values were obtained, exceeding 45 minutes (49.0 ± 2.5 minutes and $45.7 \pm$

0.2 minutes, respectively), a result justified by the supply of basic nutrients and the need for bacterial cells to synthesize other essential metabolites.

The use of 2xYT and TB supplemented with V2 salts media allowed to record a reasonable growth, which is evident from the moderate values obtained for both the specific growth rate and the generation time.

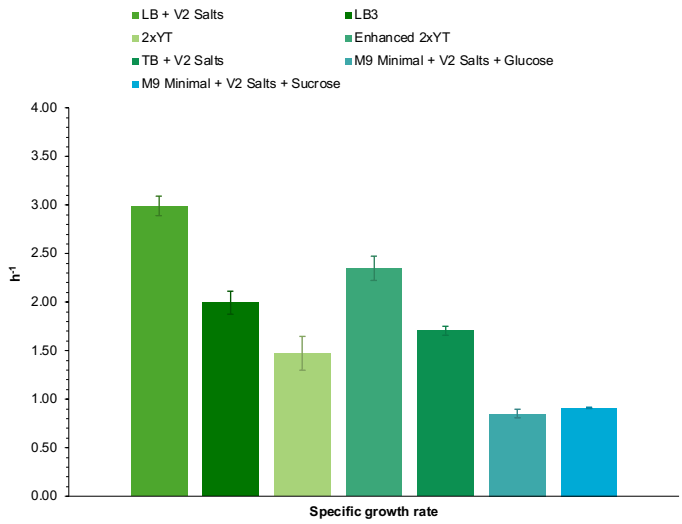
Analyzing **Figure III.31C**, the culture media 2xYT and TB supplemented with V2 salts are those that allowed to maximize the bacterial biomass for the production of MjsHSP nanocages, reaching values of $OD_{600\ nm}$ after the protein expression induction period greater than 4.5 (5.82 ± 0.26 and 4.78 ± 0.40, respectively, for the culture media 2xYT and TB supplemented with V2 salts). As expected, the M9 minimal medium with V2 salts and glucose and the M9 minimal medium with V2 salts and sucrose allowed to reach only a maximum $OD_{600\ nm}$ of 1.17 ± 0.04 and 1.13 ± 0.03, respectively. By relating these results to those of the specific growth rate, it was possible to understand that these two culture media allowed for a more gradual growth of *V. natriegens* Vmax™ X2 cells. Consequently, there was a longer availability of resources over time, which led to a greater biomass concentration at the end of the process. On the contrary, the cells that grew faster due to the presence of a highly enriched medium ended up depleting the resources available in the culture medium more quickly, which influenced the achievement of a lower amount of biomass.

Unlike what was validated for *E. coli* BL21(DE3) producing MjsHSP nanocages, the standard LB + V2 salts appeared to be the enriched culture medium with the most interest, since it allowed a very fast growth and the obtaining of a reasonable biomass concentration. This enriched medium is characterized by high nutrient availability, including amino acids, carbon sources, and essential growth factors, which contributed positively to these results. The additional presence of a high concentration of salts (NaCl, $MgCl_2$ and KCl) also helped to achieve optimal growth due to the halophilic nature of this bacterial species (233).

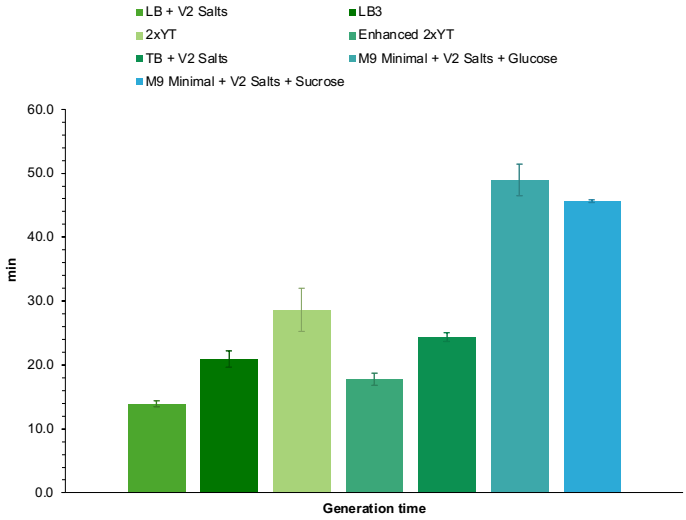
Furthermore, making this same comparison in relation to the results obtained for these three parameters in the optimization of the culture medium in *E. coli* BL21(DE3), the values that characterize *V. natriegens* Vmax™ X2 were significantly higher in all of them (between about 1.2 and 2 times), which proved its potential to be an alternative for the production of biomass and consequently of protein.

Like before, the next step consisted of analyzing the production of MjsHSP nanocages in *V. natriegens* Vmax™ X2 cells cultivated in the six alternative culture media.

(A)


LB + V2 Salts
LB3
2xYT
Enhanced 2xYT
TB + V2 Salts
M9 Minimal + V2 Salts + Glucose
M9 Minimal + V2 Salts + Sucrose
4.00
3.50
3.00
2.50
2.00
1.50
1.00
0.50
0.00
h-1
Specific growth rate

(B)


LB + V2 Salts
LB3
2xYT
Enhanced 2xYT
TB + V2 Salts
M9 Minimal + V2 Salts + Glucose
M9 Minimal + V2 Salts + Sucrose
60.0
50.0
40.0
30.0
20.0
10.0
0.0
min
Generation time

(C)

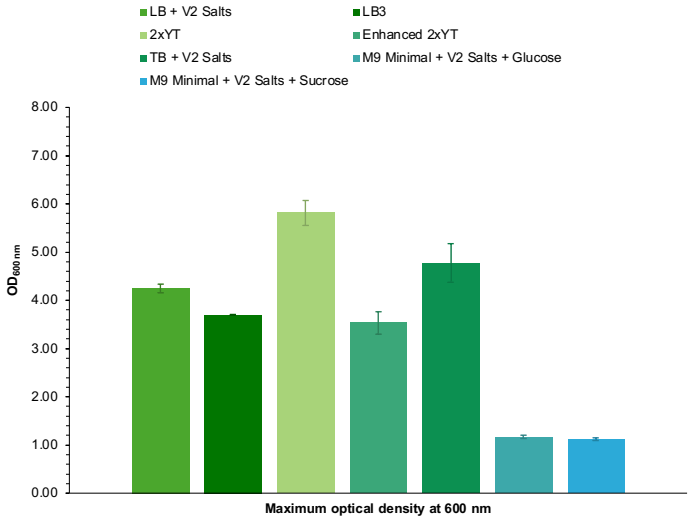

LB + V2 Salts
LB3
2xYT
Enhanced 2xYT
TB + V2 Salts
M9 Minimal + V2 Salts + Glucose
M9 Minimal + V2 Salts + Sucrose
8.00
7.00
6.00
5.00
4.00
3.00
2.00
1.00
0.00
OD600 nm
Maximum optical density at 600 nm

Figure III.31 – Growth parameters of *V. natriegens* Vmax™ X2 producing MjsHSP nanocages in six culture media with different compositions, maintaining the remaining parameters as established in the standard cultivation and protein expression (non-baffled shake flask with 30% volume of culture medium; 37°C and 250 rpm; induction of protein expression at $OD_{600\ nm}$ of 0.5-0.6, with 1 mM of IPTG, for 4 hours at 37°C). **(A)** Specific growth rate, **(B)** generation time, and **(C)** maximum $OD_{600\ nm}$. Standard culture medium (LB + V2 salts) was also included as a reference point. Each result represents the mean of two biological replicates, with error bars indicating the respective standard deviation.

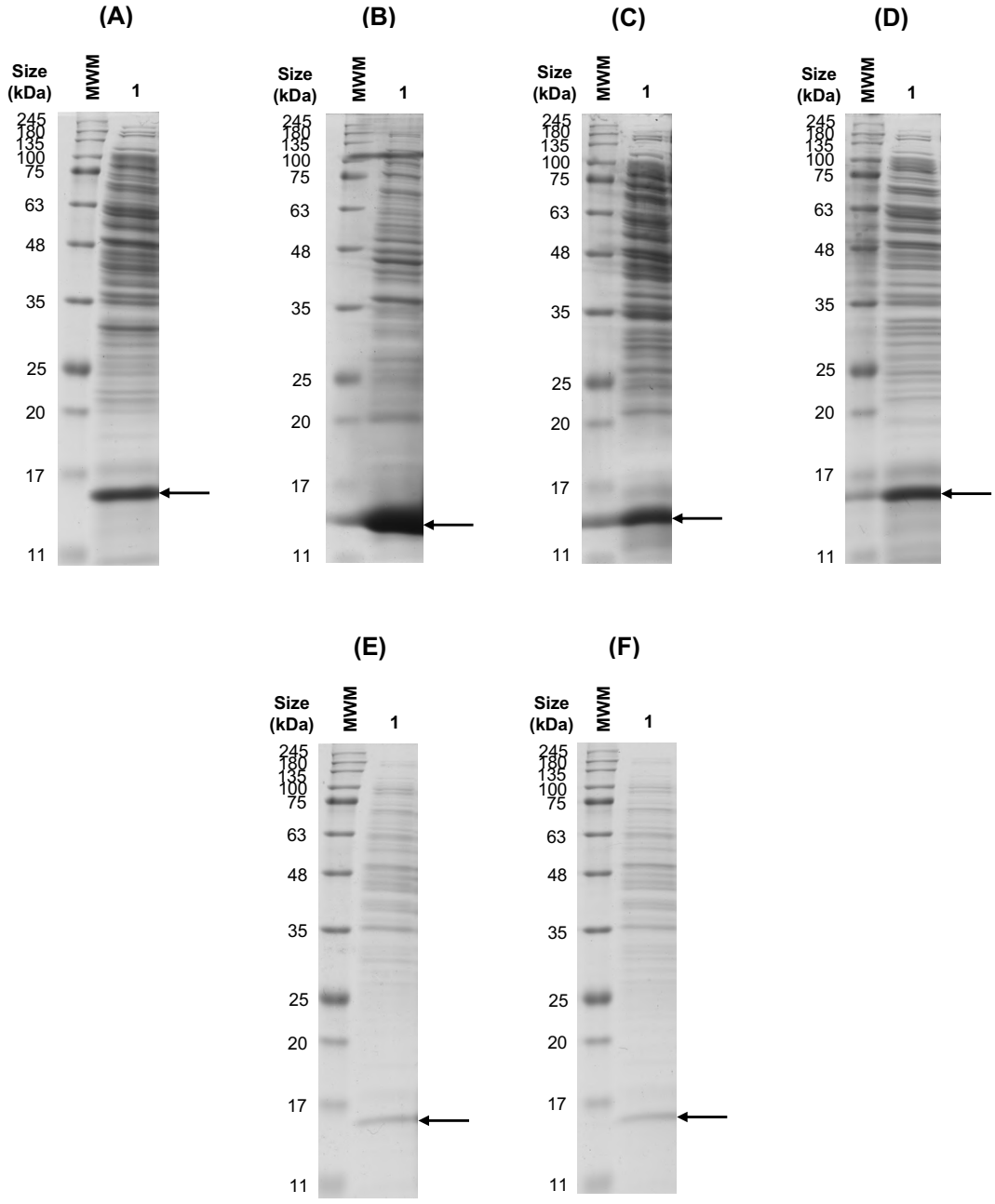


Figure III.32 – SDS-PAGE analysis obtained in the production of MjsHSP nanocages in *V. natriegens* Vmax™ X2 cells using six culture media with different compositions, maintaining the remaining parameters as established in the standard cultivation and protein expression (non-baffled shake flask with 30% volume of culture medium; 37°C and 250 rpm; induction of protein expression at OD600 nm of 0.5-0.6, with 1 mM of IPTG, for 4 hours at 37°C). **(A)**

LB3, **(B)** 2xYT, **(C)** enhanced 2xYT, **(D)** TB + V2 salts, **(E)** M9 minimal + V2 salts + glucose, and **(F)** M9 minimal + V2 salts + sucrose. Arrows indicate the protein bands referring to the monomer of MjsHSP nanocages (expected molecular weight of 16.5 kDa), which is separated and visualized under denaturing conditions of SDS-PAGE. The abbreviation MWM refers to the molecular weight marker (NZYColour protein marker II; NZYtech Lisbon, Portugal).

Comparing these results with those obtained for equivalent experiments in *E. coli* BL21(DE3), it was concluded that the intensity of the MjsHSP nanocages band was substantially reduced, indicating a lower quantity of protein nanocages produced.

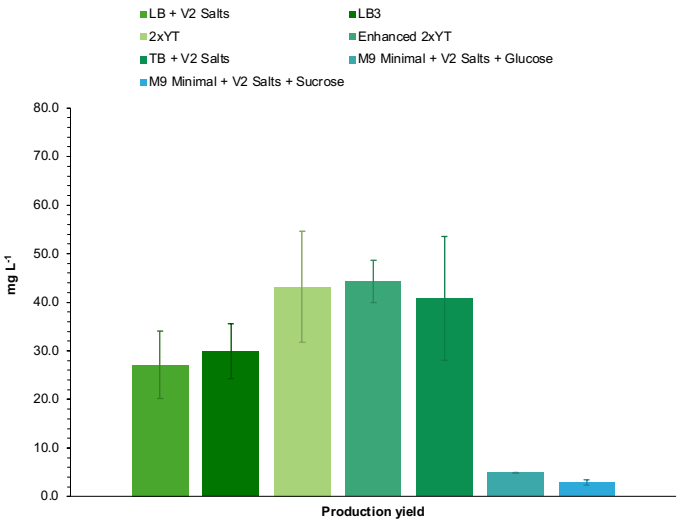
In order to quantitatively assess and compare the production efficiency of MjsHSP nanocages in *V. natriegens* Vmax™ X2 using the six culture media with different compositions, the three key characterization parameters previously presented were determined: production yield (**Equation III.5**), volumetric productivity (**Equation III.6**), and specific productivity (**Equation III.7**). As explained in the section **III.3.4. Production of MjsHSP nanocages in *E. coli* and *V. natriegens***, the determination of these three parameters implied a densitometric analysis of SDS-PAGE, which was performed according to the methodology described above (**III.2.14.3. Protein nanocages quantification by densitometric analysis of SDS-PAGE**). Due to the composition of each MjsHSP nanocage in 24 identical monomers, the total mass of protein nanocages monomers is identical to the mass of the assembled nanocages. The results obtained for these three parameters are shown graphically in **Figure III.33**.

By analyzing **Figure III.33**, it is possible to clearly verify that the cultivation of cells of this strain of *V. natriegens* in the enhanced 2xYT medium allowed to obtain values for these three parameters that were higher than those of the remaining five culture media analyzed and also of the standard LB + V2 salts culture medium. It is important to highlight the achievement of considerable standard deviations for the 2xYT and TB supplemented with V2 salts media, the second and third best conditions tested, contrary to the enhanced 2xYT which showed better reproducibility in the experiments performed.

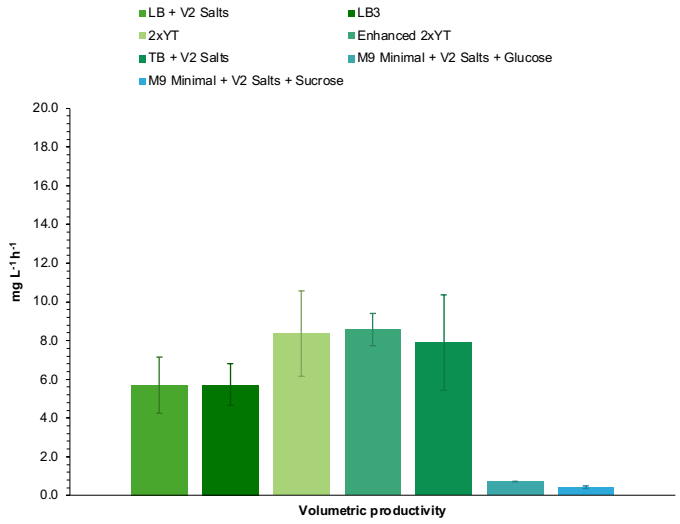
Regarding the production yield (**Figure III.33A**), this parameter was highest in enhanced 2xYT (44.3 ± 4.3 mg L^{-1}), about 1.6 times the value determined for the standard LB + V2 salts medium (27.1 ± 7.0 mg L^{-1}). The 2xYT (43.2 ± 11.4 mg L^{-1}) and TB supplemented with V2 salts (40.8 ± 12.7 mg L^{-1}) presented values close to the enhanced 2xYT, however with more experimental variability. For the remaining culture media evaluated, the lowest value recorded was that of M9 minimal medium supplemented with V2 salts and glucose (4.9 ± 0.0 mg L^{-1}) and of M9 minimal supplemented with V2 slats and sucrose (2.9 ± 0.5 mg L^{-1}).

Similarly, for the volumetric productivity (**Figure III.33B**), in the enhanced 2xYT this value (8.6 ± 0.8 mg L^{-1} h^{-1}) was 1.5 times higher than standard LB + V2 salts (5.7 ± 1.5 mg L^{-1} h^{-1}). The M9 minimal medium with V2 salts and sucrose was the one with the smallest value (0.4 ± 0.1 mg L^{-1} h^{-1}), followed by the M9 medium with V2 salts and glucose (0.7 ± 0.0 mg L^{-1} h^{-1}). The enhanced 2xYT is a nutrient-rich growth medium that improves pH stability, allowing for higher bacterial growth and increased production of non-viral protein nanocages.

(A)


LB + V2 Salts
LB3
2xYT
Enhanced 2xYT
TB + V2 Salts
M9 Minimal + V2 Salts + Glucose
M9 Minimal + V2 Salts + Sucrose
80.0
70.0
60.0
50.0
40.0
30.0
20.0
10.0
0.0
mg L-1
Production yield

(B)


LB + V2 Salts
LB3
2xYT
Enhanced 2xYT
TB + V2 Salts
M9 Minimal + V2 Salts + Glucose
M9 Minimal + V2 Salts + Sucrose
20.0
18.0
16.0
14.0
12.0
10.0
8.0
6.0
4.0
2.0
0.0
mg L-1 h-1
Volumetric productivity

(C)

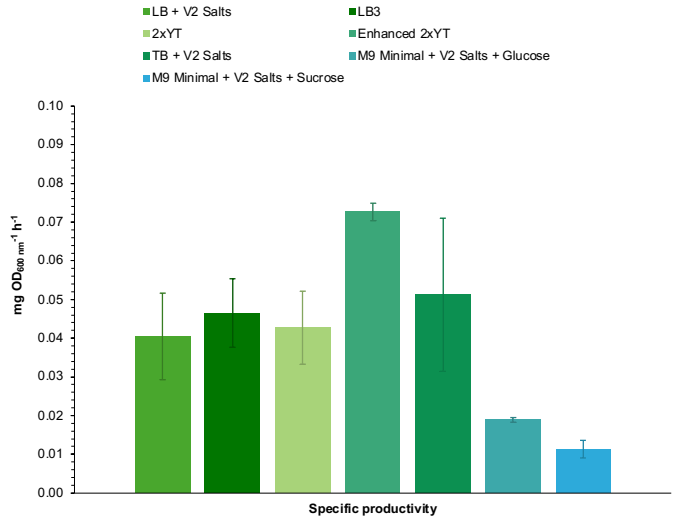

LB + V2 Salts
LB3
2xYT
Enhanced 2xYT
TB + V2 Salts
M9 Minimal + V2 Salts + Glucose
M9 Minimal + V2 Salts + Sucrose
0.10
0.09
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0.00
mg OD600 nm-1 h-1
Specific productivity

Figure III.33 – Parameters of MjsHSP nanocages production obtained for *V. natriegens* Vmax™ X2 cells in six culture media with different compositions, maintaining the remaining parameters as established in the standard cultivation and protein expression (non-baffled shake flask with 30% volume of culture medium; 37°C and 250 rpm; induction of protein expression at $OD_{600\ nm}$ of 0.5-0.6, with 1 mM of IPTG, for 4 hours at 37°C). **(A)** Production yield, **(B)** volumetric productivity, and **(C)** specific productivity. Standard culture medium (LB + V2 salts) was also included as a reference point. Each result represents the mean of two biological replicates, with error bars indicating the respective standard deviation.

For the specific productivity that evaluates the production of MjsHSP nanocages per bacterial cell (**Figure III.33B**), the enhanced 2xYT medium presented the highest value (0.07 ± 0.00 mg $OD_{600\ nm}^{-1}$ h^{-1}), followed by the TB supplemented with V2 salts (0.05 ± 0.02 mg $OD_{600\ nm}^{-1}$ h^{-1}) and LB3 (0.05 ± 0.01 mg $OD_{600\ nm}^{-1}$ h^{-1}), which represented a slight increase of 1.4 times for both of the latter. Note a very high standard deviation for TB supplemented with V2 salts. As expected, M9 minimal medium with V2 salts and glucose and M9 medium with V2 salts and sucrose recorded the lowest values (0.02 ± 0.00 mg $OD_{600\ nm}^{-1}$ h^{-1} and 0.01 ± 0.00 mg $OD_{600\ nm}^{-1}$ h^{-1}, respectively).

Comparing these characterizing parameters of the production of MjsHSP nanocages, the results obtained for *E. coli* BL21(DE3) cells after optimization considering the composition of the culture medium (163.9 ± 6.4 mg L^{-1}, 28.9 ± 1.1 mg L^{-1} h^{-1} and 0.18 ± 0.01 mg $OD_{600\ nm}^{-1}$ h^{-1}) were generally superior to those recorded with the culture medium optimized using *V. natriegens* Vmax™ X2 cells (44.3 ± 4.3 mg L^{-1}, 8.6 ± 0.8 mg L^{-1} h^{-1} and 0.07 ± 0.00 mg $OD_{600\ nm}^{-1}$ h^{-1}). Experimental variability also appears to be greater in the case of *V. natriegens*, something that can be verified by the standard deviation values. However, it is necessary to consider that this *E. coli* strain as undergone many years of study and optimization, in particular for plasmid maintenance and expression of recombinant proteins. On the contrary, this strain of *V. natriegens* results from an improvement in relation to the wild-type strain, more suitable to the protein expression, although it requires more research and genetic optimization. Therefore, the results obtained in this work will be a good starting point for future developments with *V. natriegens* and its application for the production of protein nanocages.

To conclude, the cultivation of *V. natriegens* Vmax™ X2 cells and the subsequent expression of MjsHSP nanocages in the enhanced 2xYT culture medium was what appears to be most promising. Then, this was the culture medium chosen to advance to the optimization of the second selected parameter, the concentration of IPTG as inducer of protein expression.

To test the influence of this second parameter on the production of MjsHSP nanocages in *V. natriegens* Vmax™ X2, the expression of these protein nanocages was induced in different experiments with three alternative IPTG concentrations to the standard value of 1 mM (**III.2.12. Optimization of production of MjsHSP nanocages**).

The remaining operating parameters were maintained as established in the cultivation and protein expression standard (non-baffled shake flask with 30% volume of culture medium; 37°C and 250 rpm; induction of protein expression at $OD_{600\ nm}$ of 0.5-0.6, for 4 hours at 37°C).

The growth curves obtained for the three concentrations of IPTG evaluated are shown in **Figure III.34**. For comparative purposes, standard concentration of IPTG (1 mM) was also included.

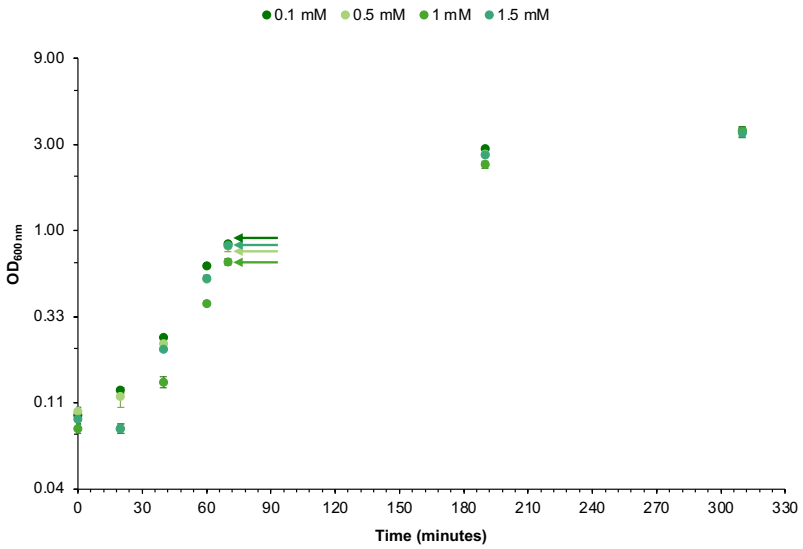


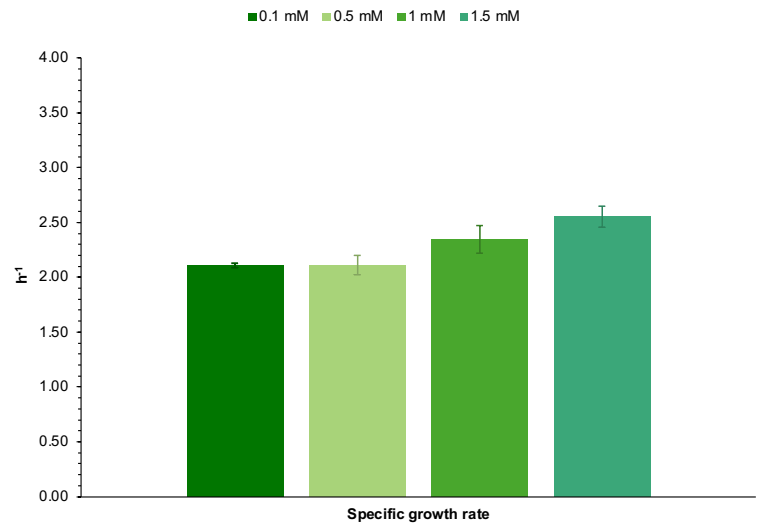
Figure III.34 – Growth curves obtained for the *V. natriegens* Vmax™ X2 cells producing MjsHSP nanocages with the induction of protein expression using three different concentrations of IPTG, maintaining the remaining parameters as previously established (non-baffled shake flask with 30% volume of culture medium; 37°C and 250 rpm; induction of protein expression at $OD_{600\ nm}$ of 0.5-0.6, for 4 hours at 37°C). Standard concentration of IPTG (1 mM) was also included as a reference point. Arrows indicate IPTG induction time points. Each data point represents the mean of two biological replicates, with error bars indicating the respective standard deviation. The y-axis is on a logarithmic scale.

The analysis of **Figure III.34** allowed to verify that the growth profile for *V. natriegens* Vmax™ X2 cells in independent experiments for induction with different IPTG concentrations was similar, an expected result because this second parameter strongly affects proteins nanocage production but has little impact on the growth curve. The three growth parameters studied (specific growth rate, generation time, and maximum $OD_{600\ nm}$ achieved) are graphically represented in **Figure III.35**.

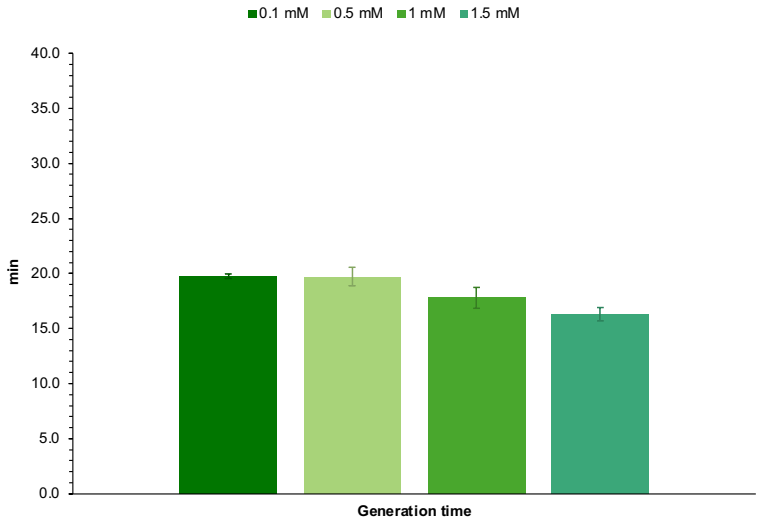
Regarding the specific growth rate (**Figure III.35A**) and the generation time (**Figure III.35B**), the values obtained for the three conditions under evaluation agreed with each other as well as with the standard condition of 1 mM of IPTG as inducer (about 2.11 to 2.55 h^{-1} for the specific growth rate and about 16.3 to 19.7 minutes for the generation time). Since these parameters were calculated during the exponential growth phase, which showed no significant differences in growth curves, this result was expected and consistent with previous observations.

In terms of the maximum $OD_{600\ nm}$ (**Figure III.35C**), it was found that the IPTG concentration also had no influence on this parameter, taking into account that for both the three conditions tested and the standard condition the maximum value obtained was very similar to each other (approximately 3.49 to 3.58). This observation suggests that MjsHSP nanocages expression was not restricting the capacity for cell replication and biomass accumulation throughout the induction period (all values indicated correspond to measurements at the end of this period).

(A)


0.1 mM
0.5 mM
1 mM
1.5 mM
4.00
3.50
3.00
2.50
2.00
1.50
1.00
0.50
0.00
h-1
Specific growth rate

(B)


0.1 mM
0.5 mM
1 mM
1.5 mM
40.0
35.0
30.0
25.0
20.0
15.0
10.0
5.0
0.0
min
Generation time

(C)

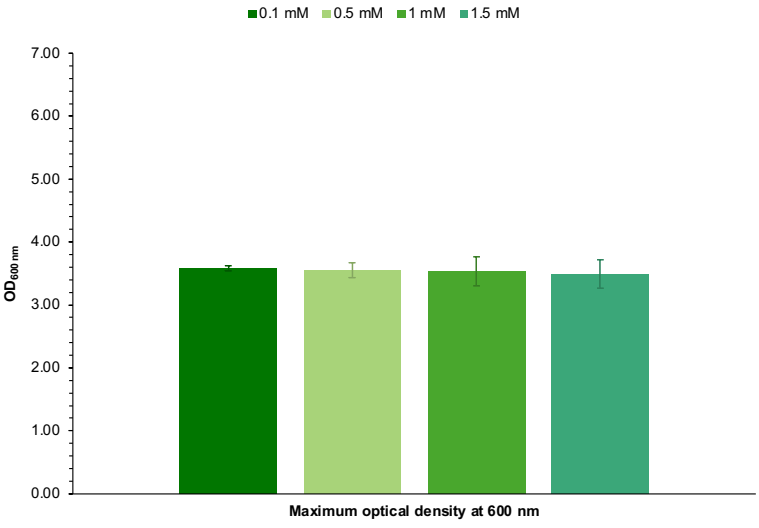

0.1 mM
0.5 mM
1 mM
1.5 mM
7.00
6.00
5.00
4.00
3.00
2.00
1.00
0.00
OD600 nm
Maximum optical density at 600 nm

Figure III.35 – Growth parameters of *V. natriegens* Vmax™ X2 producing MjsHSP nanocages with the induction of protein expression using three different concentrations of IPTG, maintaining the remaining parameters as previously established (non-baffled shake flask with 30% volume of culture medium; 37°C and 250 rpm; induction of protein expression at $OD_{600\ nm}$ of 0.5-0.6, for 4 hours at 37°C). **(A)** Specific growth rate, **(B)** generation time, and **(C)** maximum $OD_{600\ nm}$. Standard concentration of IPTG (1 mM) was also included as a reference point. Each result represents the mean of two biological replicates, with error bars indicating the respective standard deviation.

Furthermore, making this same comparison in relation to the results obtained for these three parameters in the optimization of the concentration of IPTG as inducer for protein expression in *E. coli* BL21(DE3), the values that characterize *V. natriegens* Vmax™ X2 were significantly higher in relation to the specific growth rate and generation time (about 1.8 times) but substantially lower for the maximum $OD_{600\ nm}$ achieved (about 1.9 times).

Afterwards, the production of MjsHSP nanocages in *V. natriegens* Vmax™ X2 cells using different concentrations of IPTG as inducer was evaluated by obtaining an SDS-PAGE of the clarified lysates (soluble fraction) containing this model of protein nanocages. Clear bands corresponding to the MjsHSP nanocages (expected molecular weight of 16.5 kDa) in all experiments performed can be visualized in **Figure III.36** (due to the nature of SDS-PAGE, the observed protein bands correspond to the respective monomers). The intensity of the MjsHSP nanocages band appears to be very similar among the three alternative conditions, with a minor highlight for the experiment in which 0.1 mM of IPTG was used for the induction of protein expression (**Figure III.36A**), which may suggest a slightly higher production. Again, as expected, a very similar pattern of protein impurities was observed for the tested conditions. The migration profile obtained for the standard induction of protein expression (**Figure III.6B**) allowed the identification of a band of MjsHSP nanocages monomers as well as the remaining impurities with lower intensity.

Analysis of these results in comparison to similar experiments performed with *E. coli* BL21(DE3) revealed evident weaker MjsHSP nanocages bands in all conditions. This observation suggests a substantial reduction in the production of protein nanocages.

In order to quantitatively assess and compare the production efficiency of MjsHSP nanocages in *V. natriegens* Vmax™ X2 using the three concentrations of IPTG inducer, production yield (**Equation III.5**), volumetric productivity (**Equation III.6**) and specific productivity (**Equation III.7**) were calculated. The methodology used in this analysis based on densitometric analysis of SDS-PAGE was previously explained (**III.3.4. Production of MjsHSP nanocages in *E. coli* and *V. natriegens***). The results obtained for these three key parameters are shown graphically in **Figure III.37**.

With analysis of **Figure III.37**, it can be seen that the induction of the expression of MjsHSP nanocages in *V. natriegens* Vmax™ X2 with 0.1 mM of IPTG allowed to obtain higher values for production yield, volumetric productivity and specific productivity in comparison with the remaining conditions tested including the standard concentration of IPTG (1 mM).

The production yield (**Figure III.37A**) was about 1.8 times higher with 0.1 mM of IPTG (91.3 ± 9.5 mg L^{-1}) compared to the second best concentration of inducer (1.5 mM; 54.5 ± 4.1 mg L^{-1}), and about 2.1 times higher than the reference 1 mM (10-fold higher concentration; 44.3 ± 4.3 mg L^{-1}).

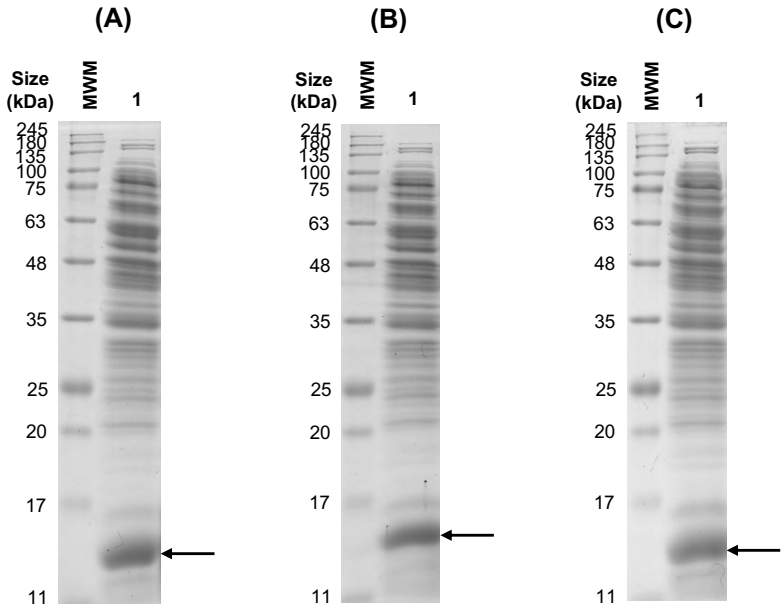
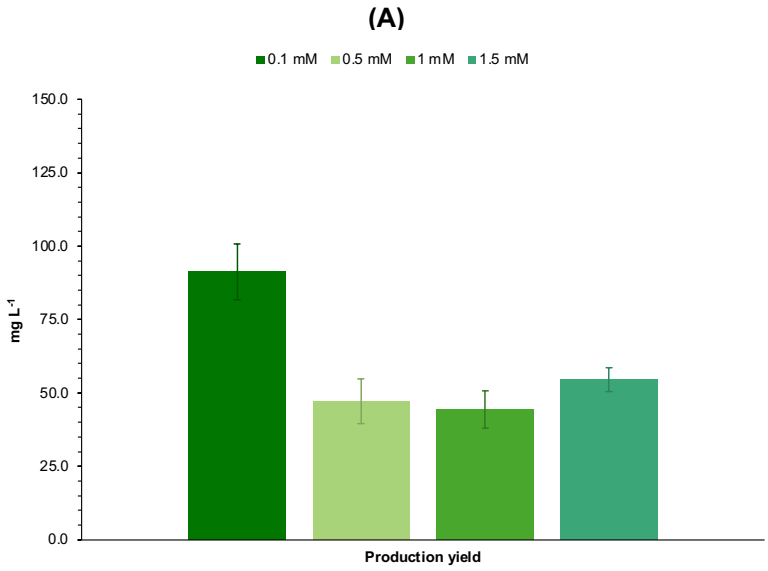


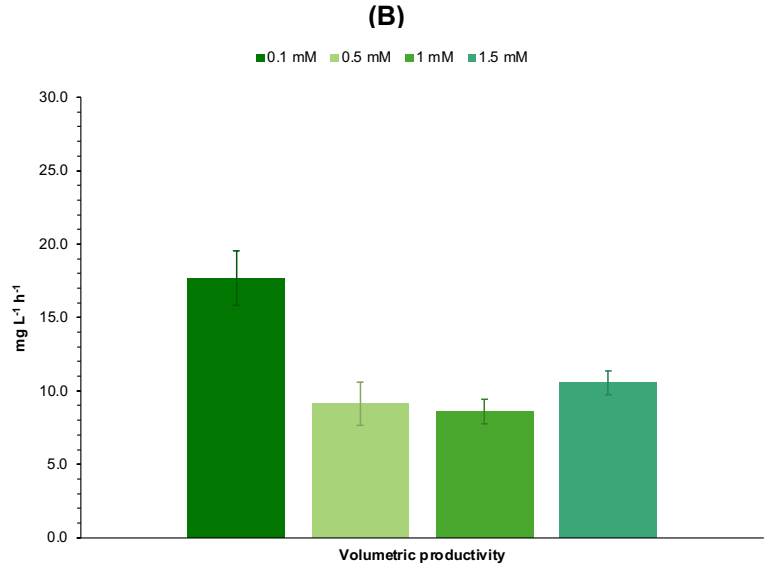
Figure III.36 – SDS-PAGE analysis obtained in the production of MjsHSP nanocages in *V. natriegens* Vmax™ X2 cells using three different concentrations of IPTG for the induction of protein expression, maintaining the remaining parameters as previously established (non-baffled shake flask with 30% volume of culture medium; 37°C and 250 rpm; induction of protein expression at $OD_{600\ nm}$ of 0.5-0.6, for 4 hours at 37°C). **(A)** 0.1 mM, **(B)** 0.5 mM, and **(C)** 1.5 mM of IPTG. Arrows indicate the protein bands referring to the monomer of MjsHSP nanocages (expected molecular weight of 16.5 kDa), which is separated and visualized under denaturing conditions of SDS-PAGE. The abbreviation MWM refers to the molecular weight marker (NZYColour protein marker II; NZYtech Lisbon, Portugal).

Similarly, for the volumetric productivity (**Figure III.37B**), it was obtained a value 2.1 times higher in 0.1 mM of IPTG (17.7 ± 1.8 mg L^{-1} h^{-1}) than the standard 1 mM (8.6 ± 0.8 mg L^{-1} h^{-1}). The use of an IPTG concentration of 1.5 mM allowed to obtain the second best result also in volumetric productivity (10.5 ± 0.8 mg L^{-1} h^{-1}). In comparison with the values estimated for the same study in *E. coli* BL21(DE3), the differences between the optimized and standard conditions were significantly greater in the case of *V. natriegens*, which demonstrated the need for research focused on optimizing the upstream processing.

For the specific productivity (**Figure III.37C**), the differences in the results for the various conditions analyzed followed the same pattern of variation indicated for production yield and volumetric productivity, which can be considered expected. The expression of MjsHSP nanocages with 0.1 mM of IPTG was the most favorable condition (0.15 ± 0.02 mg $OD_{600\ nm}^{-1}$ h^{-1}), by comparison with the reference 1 mM (0.07 ± 0.00 mg $OD_{600\ nm}^{-1}$ h^{-1}).

To date, no published studies have directly examined the relationship between IPTG inducer concentration and its effect on protein expression, specifically in the context of NVPNs. Generally, the most commonly used IPTG concentration values range between 0.5 and 1 mM (215,234). This work ultimately leads to the determination of an optimal concentration value (0.1 mM) that differs from the range reported above. The results align closely with the interpretations previously discussed for *E. coli* BL21(DE3).


(A)
0.1 mM
0.5 mM
1 mM
1.5 mM
150.0
125.0
100.0
75.0
50.0
25.0
0.0
mg L-1
Production yield


(B)
0.1 mM
0.5 mM
1 mM
1.5 mM
30.0
25.0
20.0
15.0
10.0
5.0
0.0
mg L-1 h-1
Volumetric productivity

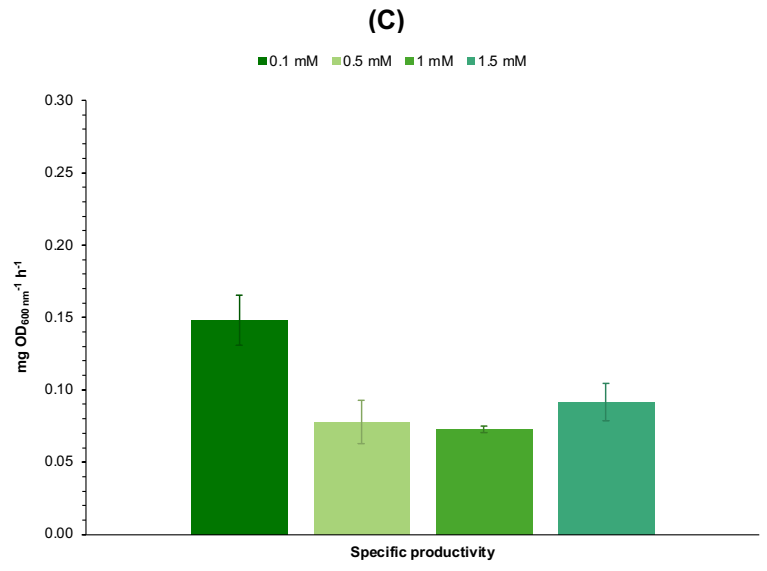

(C)
0.1 mM
0.5 mM
1 mM
1.5 mM
0.30
0.25
0.20
0.15
0.10
0.05
0.00
mg OD600 nm-1 h-1
Specific productivity

Figure III.37 – Parameters of MjsHSP nanocages production obtained for *V. natriegens* Vmax™ X2 cells using three different concentrations of IPTG for the induction of protein expression, maintaining the remaining parameters as previously established (non-baffled shake flask with 30% volume of culture medium; 37°C and 250 rpm; induction of protein expression at $OD_{600\ nm}$ of 0.5-0.6, for 4 hours at 37°C). **(A)** Production yield, **(B)** volumetric productivity, and **(C)** specific productivity. Standard concentration of IPTG (1 mM) was also included as a reference point. Each result represents the mean of two biological replicates, with error bars indicating the respective standard deviation.

Comparing these characterizing parameters of the production of MjsHSP nanocages, the results obtained for *E. coli* BL21(DE3) cells after optimization considering the concentration of IPTG as inducer of protein expression (232.9 ± 18.8 mg L^{-1}, 41.1 ± 3.3 mg L^{-1} h^{-1} and 0.19 ± 0.02 mg $OD_{600\ nm}^{-1}$ h^{-1}) were generally superior to those recorded with the IPTG concentration optimized using *V. natriegens* Vmax™ X2 cells (91.3 ± 9.5 mg L^{-1}, 17.7 ± 1.8 mg L^{-1} h^{-1} and 0.15 ± 0.02 mg $OD_{600\ nm}^{-1}$ h^{-1}).

Considering these observations, it was concluded that the induction of MjsHSP nanocages expression using an IPTG concentration of 0.1 mM was the one that appeared to be most advantageous. Therefore, together with the use of the enhanced 2xYT culture medium, this was the inducer concentration selected to advance to the optimization of the second parameter under analysis, the temperature for protein expression.

With the objective of evaluating the influence of this third parameter on the production of MjsHSP nanocages in *V. natriegens* Vmax™ X2, protein expression was induced in different experiments with three alternative temperatures to the standard value of 37°C (**III.2.12. Optimization of production of MjsHSP nanocages**).

The remaining operating parameters were maintained as previously established (non-baffled shake flask with 30% volume of culture medium; enhanced 2xYT; 37°C and 250 rpm; induction of protein expression at $OD_{600\ nm}$ of 0.5-0.6, with 0.1 mM of IPTG, for 4 hours).

The growth curves obtained for the three temperatures of protein expression evaluated are shown in **Figure III.38**. The temperature was maintained at 37°C for bacterial growth in all experiments until IPTG induction, after which it was adjusted to the specific value under analysis. For comparative purposes, standard temperature for protein nanocages expression (37°C) was also included.

As previously observed in the optimization assays of the IPTG inducer concentration, **Figure III.38** shows that the growth profile for *V. natriegens* Vmax™ X2 cells in the independent induction experiments with different protein induction temperatures was similar to each other until the time point of induction of MjsHSP nanocages expression. Thereafter, the temperature was adjusted for each test, so that the behavior of the bacterial cells also changed. The best growth was observed at a temperature of 30°C, a result that agreed with the studies available in the literature which indicate that this is the optimal temperature for the cultivation of *V. natriegens* (200,235). The use of temperatures of 37°C and 20°C allowed to obtain equivalent values of biomass concentration at the end of the induction period, although this last condition impacted the growth of bacterial cells in the first 2 hours after the addition of the IPTG inducer. For the lowest temperature (15°C), it was possible to verify a complete interruption of the bacterial growth, a quite expected observation since *V. natriegens* is a bacterial species whose survival and subsequent growth is strongly inhibited by low temperatures. This behavior is due to the

reduced catalase activity, which results in a harmful accumulation of reactive oxygen species due to the ongoing metabolic activity (236,237). The three growth parameters studied (specific growth rate, generation time, and maximum $OD_{600\ nm}$ achieved) are graphically represented in **Figure III.39**.

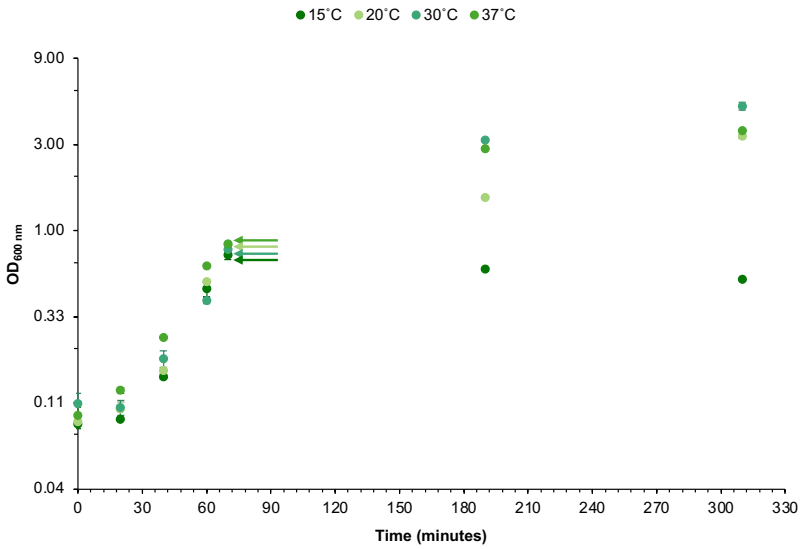
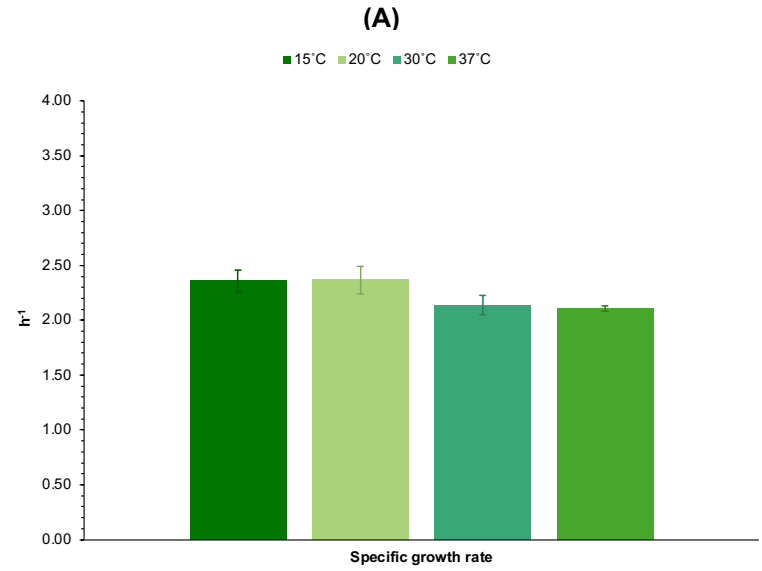


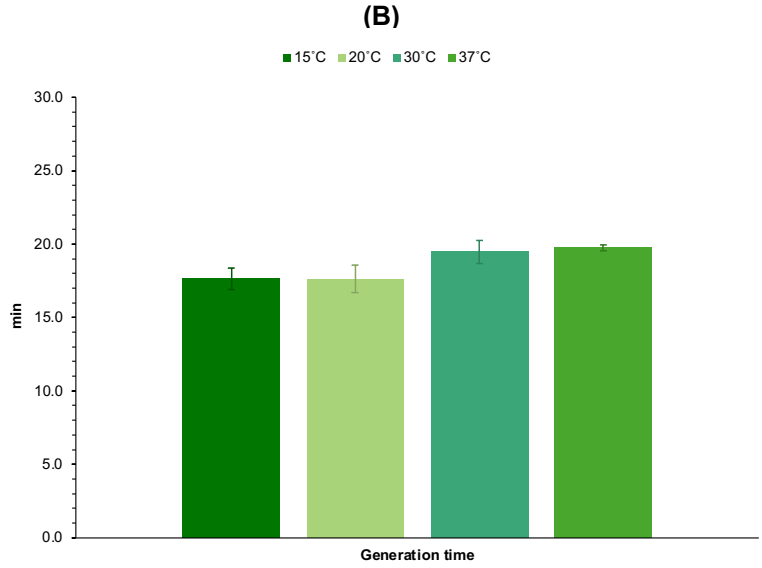
Figure III.38 – Growth curves obtained for the *V. natriegens* Vmax™ X2 cells producing MjsHSP nanocages with the induction of protein expression using three different temperatures, maintaining the remaining parameters as previously established (non-baffled shake flask with 30% volume of culture medium; enhanced 2xYT; 37°C and 250 rpm; induction of protein expression at $OD_{600\ nm}$ of 0.5-0.6, with 0.1 mM of IPTG, for 4 hours). Standard temperature for induction of protein expression (37˚C) was also included as a reference point. Arrows indicate IPTG induction time points. Each data point represents the mean of two biological replicates, with error bars indicating the respective standard deviation. The y-axis is on a logarithmic scale.

Regarding the specific growth rate (**Figure III.39A**) and the generation time (**Figure III.39B**), the results obtained for the three conditions under study are consistent with each other (about 2.14 to 2.37 h^{-1} for the specific growth rate and about 17.6 to 19.5 minutes for the generation time) and also in relation to the standard temperature condition of 37°C (2.11 ± 0.02 h^{-1} and 19.7 ± 0.2 minutes, respectively). The observed result aligns with the expectations and corroborates previous findings, given that these parameters were derived from measurements taken during the exponential growth phase.

Regarding the maximum $OD_{600\ nm}$ (**Figure III.39C**), the use of induction temperature of 30°C (4.88 ± 0.24) allowed to clearly obtain the highest value of final $OD_{600\ nm}$, approximately 9.0 times in relation to those of 15°C (0.54 ± 0.01), the worst condition evaluated. The experiments with a temperature of 37°C (3.58 ± 0.05) and 20°C (3.34 ± 0.09) led to the achievement of a maximum $OD_{600\ nm}$, respectively, 1.4 and 1.5 times lower than the most interesting temperature condition (30°C).

Additionally, making this same comparison in relation to the results obtained for these three parameters in the optimization of temperature of protein expression induction in *E. coli* BL21(DE3), the values that characterize *V. natriegens* Vmax™ X2 were significantly higher in relation to the specific growth rate and generation time (about 1.8 times) but lower for the maximum $OD_{600\ nm}$ achieved (about 1.4 times).


(A)
15˚C 20˚C 30˚C 37˚C
4.00
3.50
3.00
2.50
2.00
1.50
1.00
0.50
0.00
h-1
Specific growth rate


(B)
15˚C 20˚C 30˚C 37˚C
30.0
25.0
20.0
15.0
10.0
5.0
0.0
min
Generation time

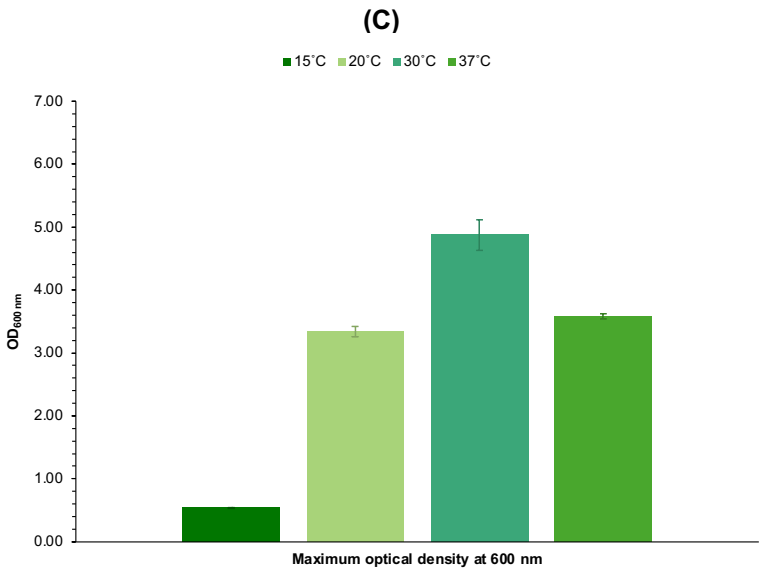

(C)
15˚C 20˚C 30˚C 37˚C
7.00
6.00
5.00
4.00
3.00
2.00
1.00
0.00
OD600 nm
Maximum optical density at 600 nm

Figure III.39 – Growth parameters of *V. natriegens* Vmax™ X2 producing MjsHSP nanocages with the induction of protein expression using three different temperatures, maintaining the remaining parameters as previously established (non-baffled shake flask with 30% volume of culture medium; enhanced 2xYT; 37°C and 250 rpm; induction of protein expression at $OD_{600\ nm}$ of 0.5-0.6, with 0.1 mM of IPTG, for 4 hours). **(A)** Specific growth rate, **(B)** generation time, and **(C)** maximum $OD_{600\ nm}$. Standard temperature for induction of protein expression (37˚C) was also included as a reference point. Each result represents the mean of two biological replicates, with error bars indicating the respective standard deviation.

Then, the production of MjsHSP nanocages in *V. natriegens* Vmax™ X2 cells using different induction temperatures was evaluated by obtaining an SDS-PAGE of the clarified lysates (soluble fraction) containing the protein nanocages. Observing **Figure III.40**, it can be seen that in the three experiments performed the specific band corresponding to the MjsHSP nanocages (expected molecular weight of 16.5 kDa) is present (due to the nature of SDS-PAGE, the observed protein bands correspond to the respective monomers). The band intensity of the MjsHSP nanocages was reducing from the lowest temperature tested (15°C) to the highest (30°C), not considering the standard temperature of 37°C, which was not exactly the expected result. Taking into account the cell growth behavior during the induction period in the three experiments, it would be predicted that as the temperature increased, the expression of MjsHSP nanocages would follow this variation. In the case of the test with the reference temperature, 37°C, it was possible to observe a band with a very high intensity for these protein nanocages. Moreover, the pattern of the bands corresponding to protein impurities had some variations between different experiments, suggesting the expression of host proteins with distinct functions in response to the temperature selected for analysis.

When comparing these findings to analogous experiments conducted using *E. coli* BL21(DE3), weaker bands corresponding to MjsHSP nanocages were observed under all conditions. This indicates a significant decrease in the amount of protein nanocages produced.

With the objective of quantitatively evaluating the production efficiency of MjsHSP nanocages in *V. natriegens* Vmax™ X2 under three different temperatures of induction, the production yield (**Equation III.5**), the volumetric productivity (**Equation III.6**) and the specific productivity (**Equation III.7**) were calculated. The approach employed for this analysis, which relies on the densitometric assessment of SDS-PAGE, was previously described in detail. **Figure III.41** graphically presents the results for these three main parameters.

Analyzing **Figure III.41**, it can be seen that the induction of the MjsHSP nanocages expression in *V. natriegens* Vmax™ X2 with a standard temperature of 37°C allowed to obtain higher values for production yield, volumetric productivity and specific productivity in comparison with the three alternative options tested, although the best results in terms of characterization of bacterial growth were associated with a temperature of 30°C.

The production yield (**Figure III.41A**) was approximately 2.4 times higher with the standard temperature in protein nanocages expression of 37°C (91.3 ± 9.5 mg L^{-1}) compared to the second best temperature (30°C; 38.2 ± 4.0 mg L^{-1}), and approximately 7.7 times higher than protein induction at 15°C (11.8 ± 0.4 mg L^{-1}).

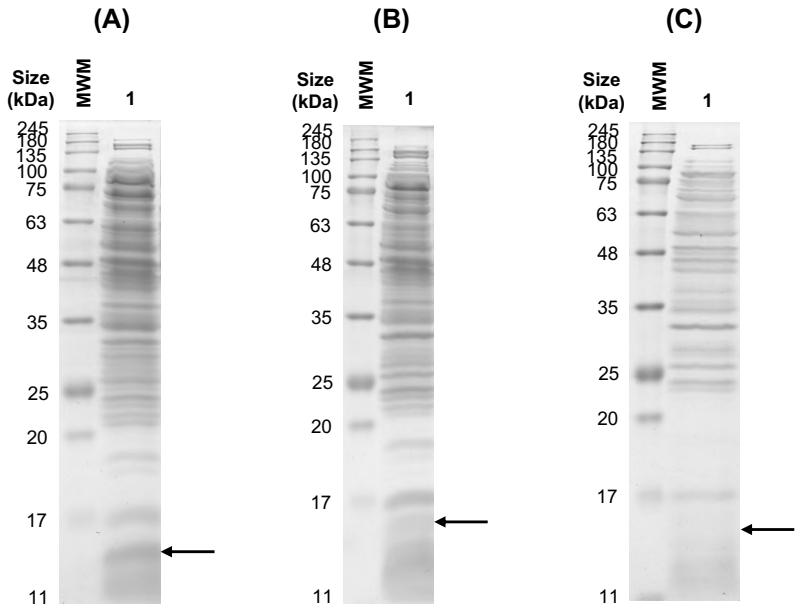
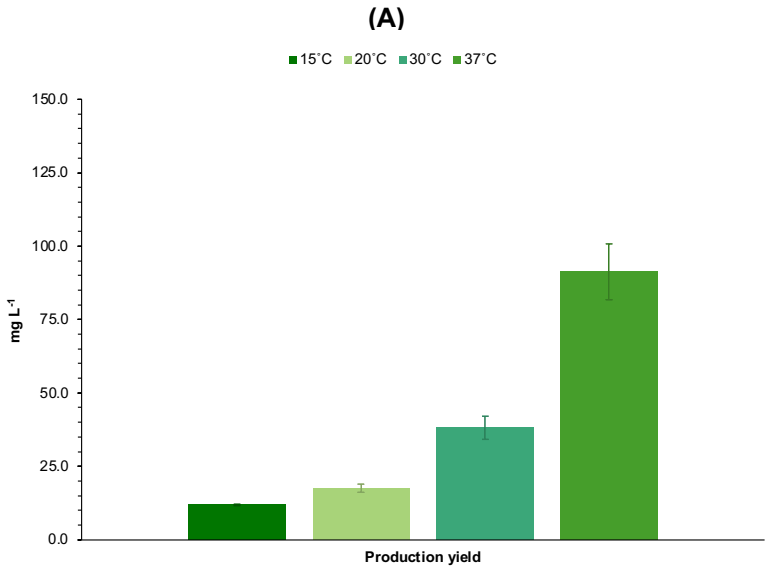


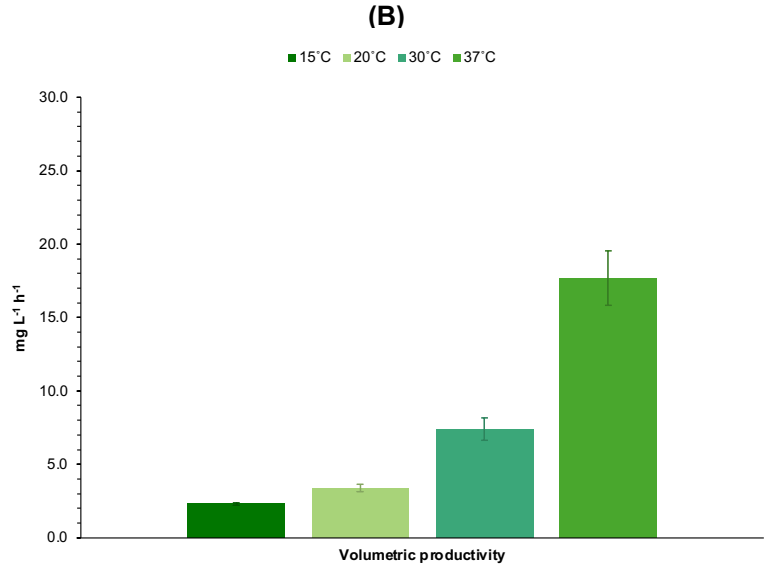
Figure III.40 – SDS-PAGE analysis obtained in the production of MjsHSP nanocages in *V. natriegens* Vmax™ X2 cells using three different temperatures for the induction of protein expression, maintaining the remaining parameters as previously established (non-baffled shake flask with 30% volume of culture medium; enhanced 2xYT; 37°C and 250 rpm; induction of protein expression at $OD_{600\ nm}$ of 0.5-0.6, with 0.1 mM of IPTG, for 4 hours). **(A)** 15˚C, **(B)** 20˚C, and **(C)** 30˚C. Arrows indicate the protein bands referring to the monomer of MjsHSP nanocages (expected molecular weight of 16.5 kDa), which is separated and visualized under denaturing conditions of SDS-PAGE. The abbreviation MWM refers to the molecular weight marker (NZYColour protein marker II; NZYtech Lisbon, Portugal).

Similarly, for the volumetric productivity (**Figure III.41B**), a value 2.4 and 7.7 times higher was obtained for the standard temperature of 37°C (17.7 ± 1.8 mg L^{-1} h^{-1}) in relation to 30°C (7.4 ± 0.8 mg L^{-1} h^{-1}) and 15°C (2.3 ± 0.1 mg L^{-1} h^{-1}), respectively, also observing a profile of growing value of this parameter as the induction temperature increases.

For the specific productivity (**Figure III.41C**), the use of the standard temperature of 37°C was the most favorable condition (0.15 ± 0.02 mg $OD_{600\ nm}^{-1}$ h^{-1}), although curiously it was possible to obtain a high value at 15°C (1.2 times; 0.13 ± 0.01 mg $OD_{600\ nm}^{-1}$ h^{-1}) compared to 30°C (0.05 ± 0.01 mg $OD_{600\ nm}^{-1}$ h^{-1}).

Contextualizing these results obtained with the data available in the literature, several studies have identified that the use of a low induction temperature in the range of 26°C to 30°C led to a greater production of proteins in *V. natriegens* compared to higher temperatures such as 37°C (200,215,238). In this work the most promising temperature for inducing the expression of MjsHSP nanocages in *V. natriegens* Vmax™ X2 was 37°C, which differs from the range indicated above. Factors such as increased ribosome functionality and better conjugation of cellular machinery involved in the transcription/translation processes may justify the observed results (200).


(A)
15˚C
20˚C
30˚C
37˚C
150.0
125.0
100.0
75.0
50.0
25.0
0.0
mg L-1
Production yield


(B)
15˚C
20˚C
30˚C
37˚C
30.0
25.0
20.0
15.0
10.0
5.0
0.0
mg L-1 h-1
Volumetric productivity

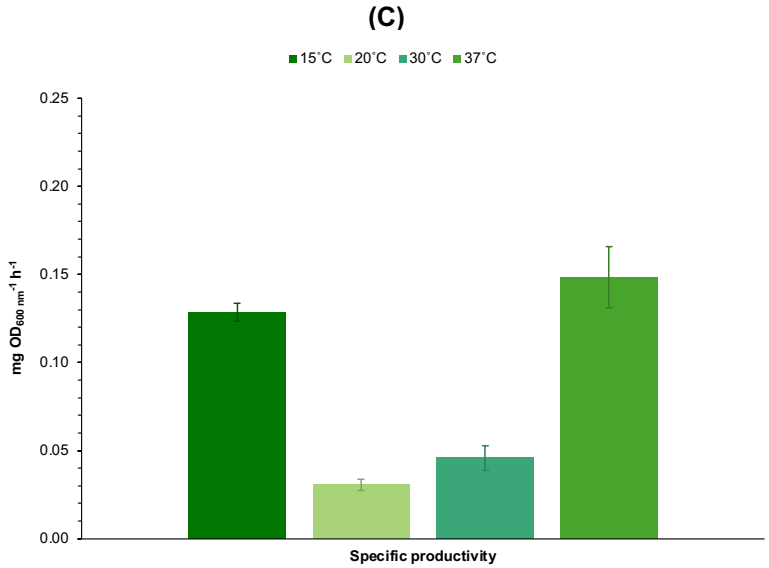

(C)
15˚C
20˚C
30˚C
37˚C
0.25
0.20
0.15
0.10
0.05
0.00
mg OD600 nm-1 h-1
Specific productivity

Figure III.41 – Parameters of MjsHSP nanocages production obtained for *V. natriegens* Vmax™ X2 cells using three different temperatures for the induction of protein expression, maintaining the remaining parameters as previously established (non-baffled shake flask with 30% volume of culture medium; enhanced 2xYT; 37°C and 250 rpm; induction of protein expression at $OD_{600\ nm}$ of 0.5-0.6, with 0.1 mM of IPTG, for 4 hours). **(A)** Production yield, **(B)** volumetric productivity, and **(C)** specific productivity. Standard temperature for induction of protein expression (37˚C) was also included as a reference point. Each result represents the mean of two biological replicates, with error bars indicating the respective standard deviation.

Comparing these characterizing parameters of the production of MjsHSP nanocages, the results obtained for *E. coli* BL21(DE3) cells after optimization considering the temperature of protein expression induction (232.9 ± 18.8 mg L^{-1}, 41.1 ± 3.3 mg L^{-1} h^{-1} and 0.19 ± 0.02 mg $OD_{600\ nm}^{-1}$ h^{-1}) were generally superior to those recorded with the temperature of induction optimized using *V. natriegens* Vmax™ X2 cells (91.3 ± 9.5 mg L^{-1}, 17.7 ± 1.8 mg L^{-1} h^{-1} and 0.15 ± 0.02 mg $OD_{600\ nm}^{-1}$ h^{-1}). It is also worth highlighting that the optimal temperature to induce the expression of these protein nanocages was 37°C, the same condition for both host bacterial species.

In conclusion, the induction of MjsHSP nanocages expression at a temperature of 37°C was the one that appears to be most efficient. Therefore, together with the use of the enhanced 2xYT culture medium and an IPTG concentration of 0.1 mM, this was the temperature of induction selected to advance to the optimization of the fourth parameter under analysis, the time point of induction.

To evaluate the influence of this fourth parameter on the production in *V. natriegens* Vmax™ X2, protein expression was induced in separate experiments using two alternative time points (corresponding to distinct values of $OD_{600\ nm} \approx 0.2$ and $OD_{600\ nm} \approx 1.4$) to the standard time point of induction at an $OD_{600\ nm}$ of 0.6 (**III.2.12. Optimization of production of MjsHSP nanocages**).

The remaining operating parameters were maintained as previously established (non-baffled shake flask with 30% volume of culture medium; enhanced 2xYT; 37°C and 250 rpm; induction of protein expression with 0.1 mM of IPTG, for 4 hours at 37°C).

The growth curves obtained for the two time points of protein expression induction evaluated are shown in **Figure III.42**. For comparative purposes, the standard time point ($OD_{600\ nm} \approx 0.6$) was also included.

Through the analysis of **Figure III.42**, the growth profile for *V. natriegens* Vmax™ X2 cells in independent experiments for induction at two distinct time points (early exponential phase and late exponential/early stationary phase) was equivalent throughout the entire monitored period, including from the time point of induction, which is an interesting observation. This similarity was also verified in relation to the standard timepoint ($OD_{600\ nm} \approx 0.6$). These results suggest that *V. natriegens* cells were able to balance the resources available for both the expression of protein nanocages and for cell growth and subsequent biomass generation. The three growth parameters studied (specific growth rate, generation time, and maximum $OD_{600\ nm}$ achieved) are graphically represented in **Figure III.43**.

Regarding the specific growth rate (**Figure III.43A**) and the generation time (**Figure III.43B**), the induction at a time point in early exponential phase ($OD_{600\ nm} \approx 0.2$; 3.25 ± 0.15 h^{-1} and 12.8 ± 0.6 minutes, respectively) led to obtaining higher values for these two parameters (about 1.5 times for both)

in relation to the induction in the middle of exponential phase ($OD_{600\ nm} \approx 0.6$; $2.11 \pm 0.02\ h^{-1}$ and 19.7 ± 0.2 minutes, respectively) and in the late exponential/early stationary phase ($OD_{600\ nm} \approx 1.4$; $2.16 \pm 0.07\ h^{-1}$ and 19.3 ± 0.6 minutes, respectively). It was a curious result since these parameters were calculated during the exponential growth phase and from the analysis of the respective growth curve (**Figure III.42**) there did not appear to be such a significant difference.

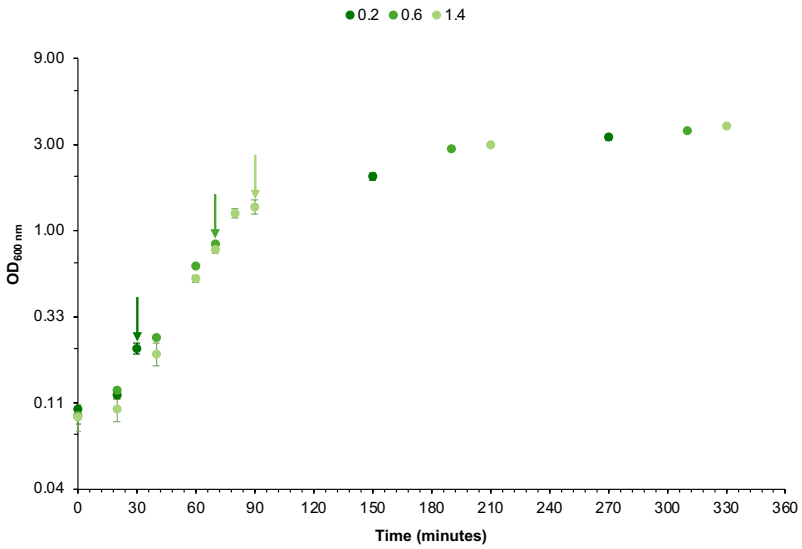
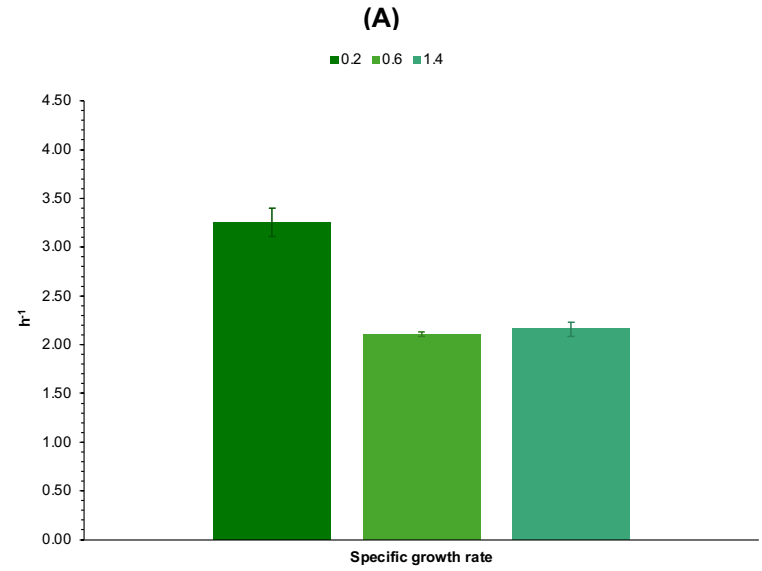
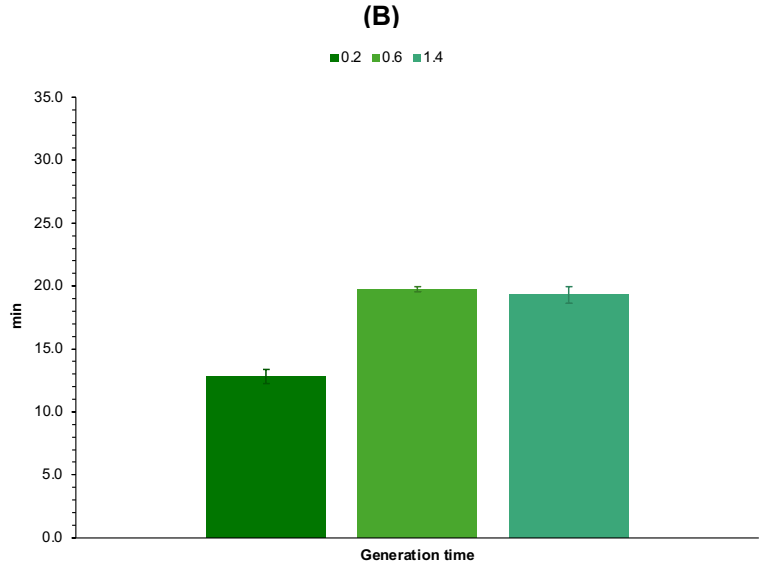


Figure III.42 – Growth curves obtained for the *V. natriegens* Vmax™ X2 cells producing MjsHSP nanocages with the induction of protein expression using two different time points, maintaining the remaining parameters as previously established (non-baffled shake flask with 30% volume of culture medium; enhanced 2xYT; 37°C and 250 rpm; induction of protein expression with 0.1 mM of IPTG, for 4 hours at 37°C). Standard time point for induction of protein expression ($OD_{600\ nm} \approx 0.6$) was also included as a reference point. Arrows indicate IPTG induction time points. Each data point represents the mean of two biological replicates, with error bars indicating the respective standard deviation. The y-axis is on a logarithmic scale.

In terms of the maximum $OD_{600\ nm}$ (**Figure III.43C**), it was found that induction at a time point in late exponential/early stationary phase ($OD_{600\ nm} \approx 1.4$) allowed to obtain a growth level 1.1 times higher (3.81 ± 0.04) than the standard condition of induction at the middle of exponential growth phase (3.58 ± 0.05) and 1.2 times higher than the alternative condition tested of induction at early exponential phase (3.30 ± 0.12). Although the difference between the two alternative conditions tested and the standard experiment was less pronounced than that obtained in the study with *E. coli* BL21(DE3), the slight variation can be explained by the dynamics between cell growth and protein expression at different moments of the exponential growth phase, as discussed in more detail previously.

Furthermore, making this same comparison in relation to the results obtained for these three parameters in the optimization of the time point of protein expression induction in *E. coli* BL21(DE3), the values that characterize *V. natriegens* Vmax™ X2 were significantly higher in relation to the specific growth rate and generation time (about 2.3 times) but lower for the maximum $OD_{600\ nm}$ achieved (about 1.9 times).


(A)
0.2 0.6 1.4
4.50
4.00
3.50
3.00
2.50
2.00
1.50
1.00
0.50
0.00
h-1
Specific growth rate


(B)
0.2 0.6 1.4
35.0
30.0
25.0
20.0
15.0
10.0
5.0
0.0
min
Generation time

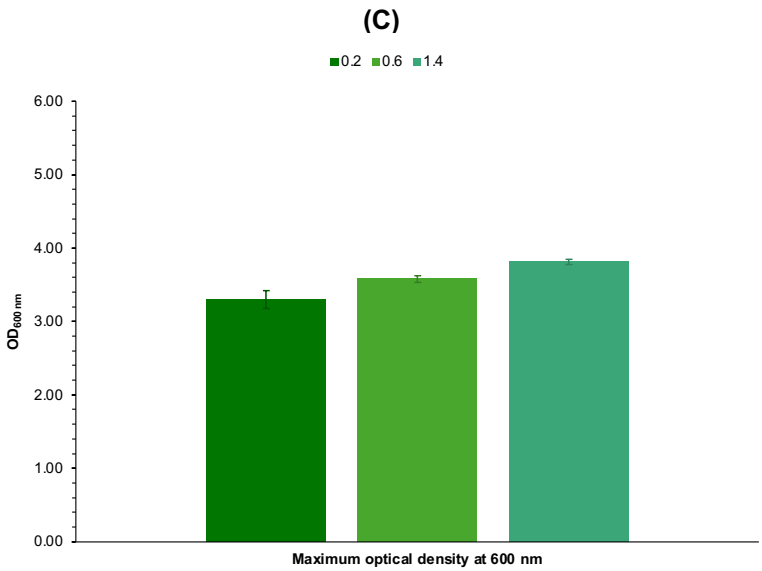

(C)
0.2 0.6 1.4
6.00
5.00
4.00
3.00
2.00
1.00
0.00
OD600 nm
Maximum optical density at 600 nm

Figure III.43 – Growth parameters of *V. natriegens* Vmax™ X2 producing MjsHSP nanocages with the induction of protein expression using two different time points, maintaining the remaining parameters as previously established (non-baffled shake flask with 30% volume of culture medium; enhanced 2xYT; 37°C and 250 rpm; induction of protein expression with 0.1 mM of IPTG, for 4 hours at 37°C). **(A)** Specific growth rate, **(B)** generation time, and **(C)** maximum $OD_{600\ nm}$. Standard time point for induction of protein expression ($OD_{600\ nm} \approx 0.6$) was also included as a reference point. Each result represents the mean of two biological replicates, with error bars indicating the respective standard deviation.

Afterwards, the production of MjsHSP nanocages in *V. natriegens* Vmax™ X2 cells using two time points of protein expression induction was evaluated. This assessment was performed through SDS-PAGE analysis of the clarified lysates (soluble fraction) containing the MjsHSP nanocages. Distinguishable bands corresponding to the MjsHSP nanocages, with an expected molecular weight of 16.5 kDa, were visible across all the experiments conducted, as illustrated in **Figure III.44** (due to the nature of SDS-PAGE, the observed protein bands correspond to the respective monomers). The intensity of the MjsHSP nanocages band appeared to be slightly stronger with the induction of protein expression occurring in the early exponential growth phase, although the difference was not very significant. Compared to the standard time point of induction ($OD_{600\ nm} \approx 0.6$), this band intensity in the SDS-PAGE appears to have been slightly higher (**Figure III.36A**). From **Figure III.44**, it is also possible to observe a greater intensity of the bands corresponding to protein impurities, although their profile presented similarities for all tested conditions.

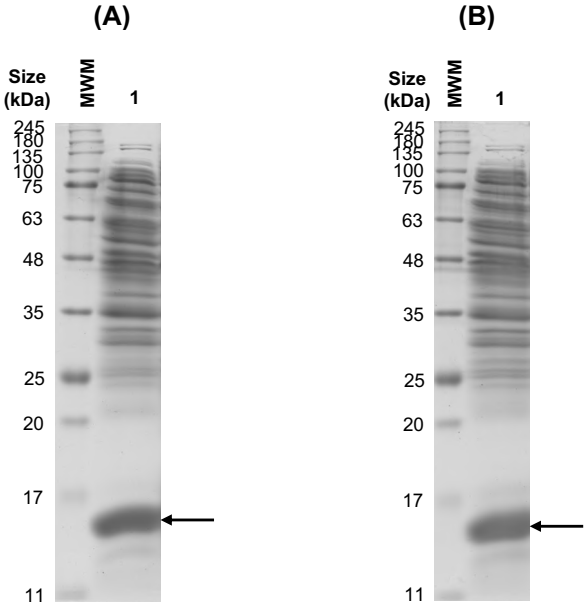


Figure III.44 – SDS-PAGE analysis obtained in the production of MjsHSP nanocages in *V. natriegens* Vmax™ X2 cells using two different time points for the induction of protein expression, maintaining the remaining parameters as previously established (non-baffled shake flask with 30% volume of culture medium; enhanced 2xYT; 37°C and 250 rpm; induction of protein expression with 0.1 mM of IPTG, for 4 hours at 37°C). **(A)** $OD_{600\ nm} \approx 0.2$, and **(B)** $OD_{600\ nm} \approx 1.4$. Arrows indicate the protein bands referring to the monomer of MjsHSP nanocages (expected

molecular weight of 16.5 kDa), which is separated and visualized under denaturing conditions of SDS-PAGE. The abbreviation MWM refers to the molecular weight marker (NZYColour protein marker II; NZYtech Lisbon, Portugal).

In contrast to parallel experiments performed with *E. coli* BL21(DE3), the bands representing MjsHSP nanocages were less intense across all tested conditions. This observation suggests a substantial reduction in the amount of protein nanocages produced. Furthermore, a greater proportion of bands related to protein impurities was observed in relation to the band of protein nanocages in *V. natriegens*, with the latter band being more prominent in the production using *E. coli* BL21(DE3) cells.

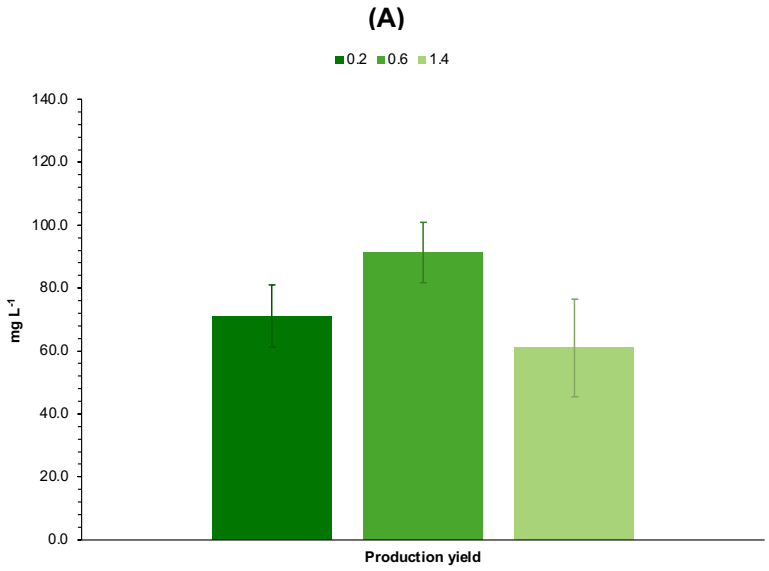
To evaluate and relate the efficiency of MjsHSP nanocages production in *E. coli* BL21(DE3) under the two alternative time points of protein expression induction, key parameters such as production yield (**Equation III.5**), volumetric productivity (**Equation III.6**) and specific productivity (**Equation III.7**) were estimated. The analysis was performed using densitometric measurements from SDS-PAGE, as previously outlined. The results for these three parameters are presented in **Figure III.45**.

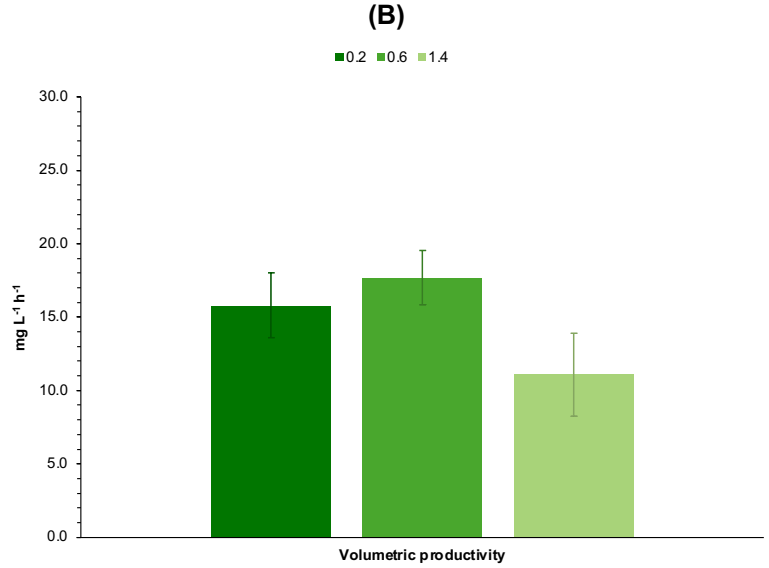
From **Figure III.45**, the induction of the expression of MjsHSP nanocages at a time point corresponding to the middle of the exponential growth phase ($OD_{600\ nm} \approx 0.6$), the standard condition, allowed to obtain higher values of production yield and volumetric productivity in comparison with the results of the alternative conditions at the early exponential phase ($OD_{600\ nm} \approx 0.2$) and at the late exponential/early stationary phase ($OD_{600\ nm} \approx 1.4$).

The production yield (**Figure III.45A**) was about 1.3 times higher with time point of induction at $OD_{600\ nm} \approx 0.6$ (91.3 ± 9.5 mg L^{-1}) compared to the second best time point ($OD_{600\ nm} \approx 0.2$; 71.1 ± 9.8 mg L^{-1}), and about 1.5 times higher than the value obtained at the time point of $OD_{600\ nm} \approx 1.4$ (61.0 ± 15.5 mg L^{-1}).

Similarly, for the volumetric productivity (**Figure III.45B**), the induction of MjsHSP nanocages expression at a time point of induction at $OD_{600\ nm} \approx 0.6$ (17.7 ± 1.8 mg L^{-1} h^{-1}) allowed to obtain a result 1.1 and 1.6 times higher than that determined for the time point at $OD_{600\ nm} \approx 0.2$ (15.8 ± 2.2 mg L^{-1} h^{-1}) and at $OD_{600\ nm} \approx 1.4$ (11.1 ± 2.8 mg L^{-1} h^{-1}), respectively. Interestingly, between the two alternatives tested, induction at an early stage of *V. natriegens* Vmax™ X2 cell growth resulted in a favorable expression of MjsHSP nanocages.

In terms of the specific productivity (**Figure III.45B**), the results showed that induction at an $OD_{600\ nm} \approx 0.2$ (0.15 ± 0.03 mg $OD_{600\ nm}^{-1}$ h^{-1}) and at an $OD_{600\ nm} \approx 0.6$ (0.15 ± 0.02 mg $OD_{600\ nm}^{-1}$ h^{-1}) allowed protein production per cell equivalent to each other. Clearly, induction at a late stage of bacterial growth ($OD_{600\ nm} \approx 1.4$) led to a much lower specific productivity result (about 1.7 times for 0.09 ± 0.02 mg $OD_{600\ nm}^{-1}$ h^{-1}).


(A)
0.2 0.6 1.4
140.0
120.0
100.0
80.0
60.0
40.0
20.0
0.0
mg L-1
Production yield


(B)
0.2 0.6 1.4
30.0
25.0
20.0
15.0
10.0
5.0
0.0
mg L-1 h-1
Volumetric productivity

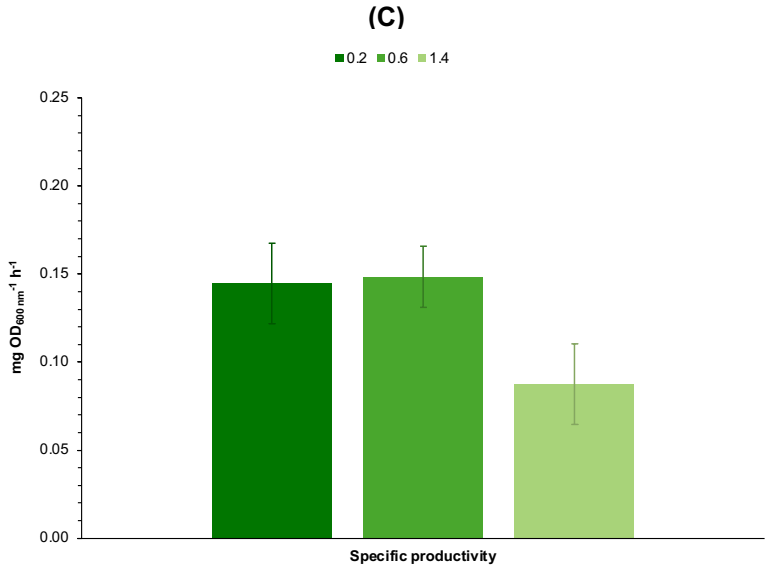

(C)
0.2 0.6 1.4
0.25
0.20
0.15
0.10
0.05
0.00
mg OD600 nm-1 h-1
Specific productivity

Figure III.45 – Parameters of MjsHSP nanocages production obtained for *V. natriegens* Vmax™ X2 cells using two different time points for the induction of protein expression, maintaining the remaining parameters as previously established (non-baffled shake flask with 30% volume of culture medium; enhanced 2xYT; 37°C and 250 rpm; induction of protein expression with 0.1 mM of IPTG, for 4 hours at 37°C). **(A)** Production yield, **(B)** volumetric productivity, and **(C)** specific productivity. Standard time point for induction of protein expression ($OD_{600\ nm} \approx 0.6$) was also included as a reference point. Each result represents the mean of two biological replicates, with error bars indicating the respective standard deviation.

Framing these results with the research described in the literature, most studies have considered induction at an $OD_{600\ nm}$ between 0.5 and 1.0, an interval corresponding to the exponential growth phase, as the most appropriate time point for the existence of a greater expression of proteins in *V. natriegens* (156,198,215,238). The results obtained in this work corroborate these conclusions, since with the induction of the production of MjsHSP nanocages during the exponential phase ($OD_{600\ nm} \approx 0.6$) it was possible to achieve quite promising values for the three parameters characterizing protein expression. High metabolic activity, efficient protein synthesis machinery together with higher number of ribosomes are characteristics of *V. natriegens* Vmax™ X2 that may justify the observed results (197,200,215).

Comparing these characterizing parameters of the production of MjsHSP nanocages, the results obtained for *E. coli* BL21(DE3) cells after optimization considering the time point of protein expression induction (232.9 ± 18.8 mg L^{-1}, 41.1 ± 3.3 mg L^{-1} h^{-1} and 0.19 ± 0.02 mg $OD_{600\ nm}^{-1}$ h^{-1}) were superior to those recorded with the time point of induction optimized using *V. natriegens* Vmax™ X2 cells (91.3 ± 9.5 mg L^{-1}, 17.7 ± 1.8 mg L^{-1} h^{-1} and 0.15 ± 0.02 mg $OD_{600\ nm}^{-1}$ h^{-1}). It is also worth highlighting that the optimal time point to induce the expression of these protein nanocages was at the middle of exponential growth phase, the same condition for both host bacterial species. Furthermore, it is important to note that the experimental variability in *V. natriegens* was greater than in *E. coli*, as demonstrated by the calculated standard deviation values.

To conclude, inducing the expression of MjsHSP nanocages at a time point in the middle of the exponential growth phase ($OD_{600\ nm}$ of approximately 0.5 to 0.6) appears to be the most advantageous. Therefore, in combination with the use of enhanced 2xYT culture medium, an IPTG concentration of 0.1 mM and an induction temperature of 37°C, this time point of induction was selected to proceed with the optimization of the fifth and last parameter under analysis, the aeration conditions through the presence or absence of baffles in the culture shake flask.

The influence of this fifth and last parameter on the production of MjsHSP nanocages in *V. natriegens* Vmax™ X2 was assessed through two separate setups, one with better air flow (using shake flask with baffles) and another with normal air flow (using regular shake flasks without baffles). For a reliable c comparison, 250 mL flasks (with and without baffles) were used, each filled to 30% capacity with culture medium (**III.2.12. Optimization of production of MjsHSP nanocages**).

The remaining operating parameters were maintained as previously established (enhanced 2xYT; 37°C and 250 rpm; induction of protein expression at $OD_{600\ nm}$ of 0.5-0.6, with 0.1 mM of IPTG, for 4 hours at 37°C).

The growth curves obtained for the two aeration conditions studied are shown in **Figure III.46**. For comparative purposes, the standard aeration condition (non-baffled shake flask with 100 mL and 30% volume of culture medium) was also included.

Analyzing **Figure III.46**, the growth profile for *V. natriegens* Vmax™ X2 cells in both the two alternative aeration conditions and the condition defined as reference was similar until the addition of IPTG as protein expression inducer. During the induction period, a slight variation was observed between the three independent experiments with the use of a 250 mL baffled shake flask allowing a more favorable cell growth. The three growth parameters assessed (specific growth rate, generation time, and maximum $OD_{600\ nm}$ achieved) are graphically represented in **Figure III.47**.

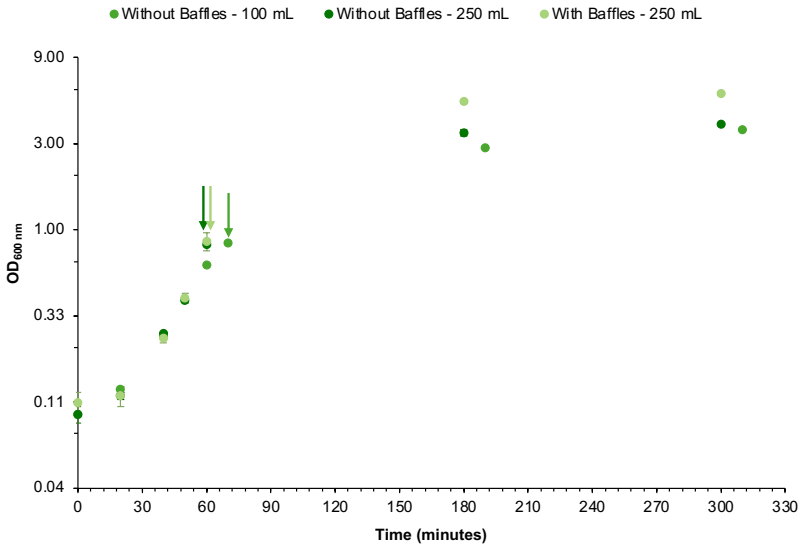
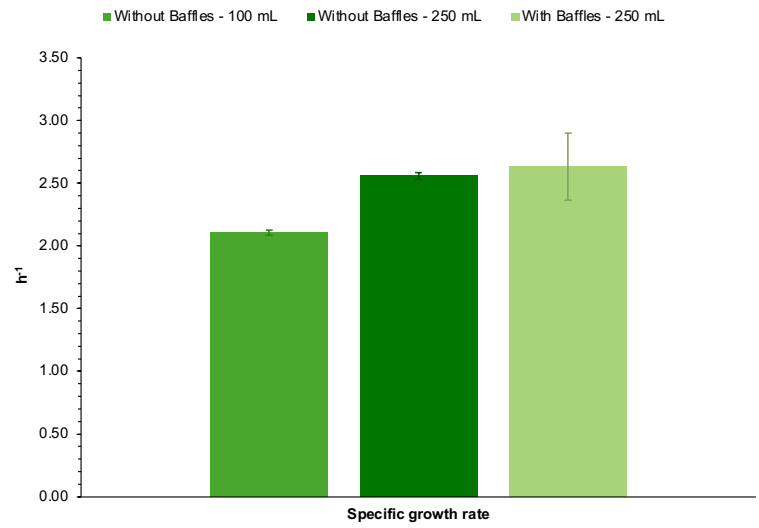


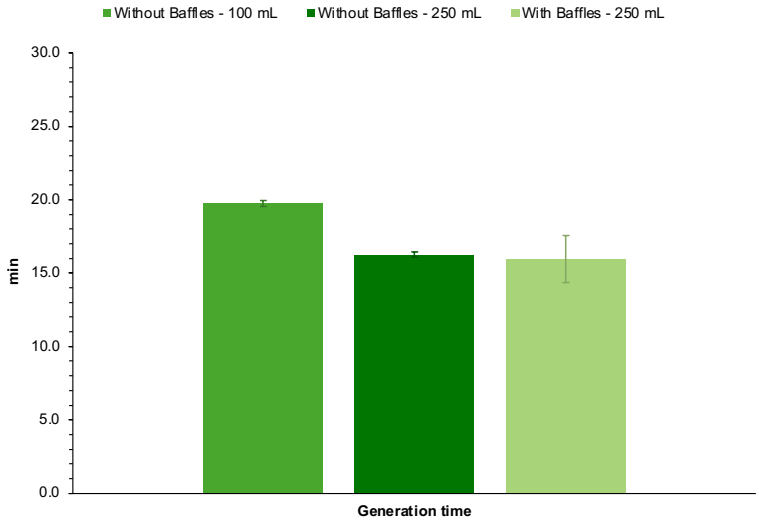
Figure III.46 – Growth curves obtained for the *V. natriegens* Vmax™ X2 cells producing MjsHSP nanocages with the induction of protein expression using two different aeration conditions (250 mL shake flask without and with baffles), maintaining the remaining parameters as previously established (enhanced 2xYT; 37°C and 250 rpm; induction of protein expression at $OD_{600\ nm}$ of 0.5-0.6, with 0.1 mM of IPTG, for 4 hours at 37°C). Standard aeration condition for induction of protein expression (100 mL shake flask without baffles) was also included as a reference point. Arrows indicate IPTG induction time points. Each data point represents the mean of two biological replicates, with error bars indicating the respective standard deviation. The y-axis is on a logarithmic scale.

The specific growth rate (**Figure III.47A**) and the generation time (**Figure III.47B**) values obtained for the two evaluated conditions were consistent with each other (approximately 2.56 to 2.63 h^{-1} for the specific growth rate and about 16.0 to 16.3 minutes for the generation time). It should be noted that the experiment relating to the 250 mL non-baffled shake flask showed greater experimental variability. The standard condition that used the non-baffled shake flask with 100 mL presented a lower specific growth rate (1.2 times; 2.11 $\pm$ 0.02 h^{-1}) and a consequent higher generation time (1.2 times; 19.7 $\pm$ 0.2 minutes). Research studies have shown that baffled shake flasks would allow a greater supply of oxygen, which has the consequence of supporting a faster growth (239), an observation corroborated by this work. As reported for the experiments using *E. coli* BL21(DE3) cells, the cultivation in the baffled shake flask led to greater foam formation, which may potentially interfere with the cell growth.

(A)


Without Baffles - 100 mL
Without Baffles - 250 mL
With Baffles - 250 mL
3.50
3.00
2.50
2.00
1.50
1.00
0.50
0.00
h-1
Specific growth rate

(B)


Without Baffles - 100 mL
Without Baffles - 250 mL
With Baffles - 250 mL
30.0
25.0
20.0
15.0
10.0
5.0
0.0
min
Generation time

(C)

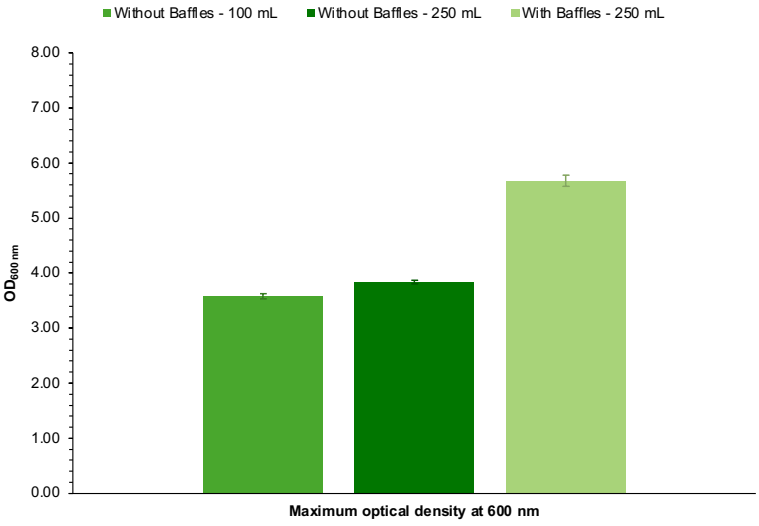

Without Baffles - 100 mL
Without Baffles - 250 mL
With Baffles - 250 mL
8.00
7.00
6.00
5.00
4.00
3.00
2.00
1.00
0.00
OD600 nm
Maximum optical density at 600 nm

Figure III.47 – Growth parameters of *V. natriegens* Vmax™ X2 cells producing MjsHSP nanocages with the induction of protein expression using two different aeration conditions (250 mL shake flask without and with baffles), maintaining the remaining parameters as previously established (enhanced 2xYT; 37°C and 250 rpm; induction of protein expression at $OD_{600\ nm}$ of 0.5-0.6, with 0.1 mM of IPTG, for 4 hours at 37°C). **(A)** Specific growth rate, **(B)** generation time, and **(C)** maximum $OD_{600\ nm}$. Standard aeration condition for induction of protein expression (100 mL shake flask without baffles) was also included as a reference point. Each result represents the mean of two biological replicates, with error bars indicating the respective standard deviation.

Regarding the maximum $OD_{600\ nm}$ (**Figure III.47C**), the test using a 250 mL baffled shake flask (5.68 ± 0.10) allowed to obtain a cell density much higher than the condition with the 250 mL non-baffled shake flask (1.5 times; 3.84 ± 0.04) and the reference aeration condition (1.6 times; 3.58 ± 0.05). This result confirmed that baffles increase oxygen transfer rate compared to the shake flasks without them, which can lead to higher cell densities also in *V. natriegens*.

Moreover, making this same comparison in relation to the results obtained for these three parameters in the optimization of aeration conditions for MjsHSP nanocages expression in *E. coli* BL21(DE3), the values that characterize *V. natriegens* Vmax™ X2 were significantly higher in relation to the specific growth rate and generation time (about 1.7 times) but lower for the maximum $OD_{600\ nm}$ achieved (about 1.1 times).

Following, the production of MjsHSP nanocages in *V. natriegens* Vmax™ X2 cells using the two alternative aeration conditions was assessed. This evaluation was performed through SDS-PAGE analysis of the soluble fraction from the clarified lysates containing this model of protein nanocages. Distinct bands corresponding to the MjsHSP nanocages, with an expected molecular weight of 16.5 kDa, are visible across all the experiments conducted, as illustrated in **Figure III.48** (due to the nature of SDS-PAGE, the observed protein bands correspond to the respective monomers). The band intensity of the MjsHSP nanocages as well as the pattern and intensity of the bands corresponding to protein impurities was similar in both the production in the 250 mL non-baffled shake flask and in the 250 mL baffled shake flask. Compared to the standard condition, in particular the growth and the production of protein nanocages in the 100 mL non-baffled shake flask (**Figure III.36A**), in the latter a much higher intensity was obtained in relation to the two alternative conditions, indicating a greater quantity of protein nanocages produced.

When comparing these findings to analogous experiments conducted using *E. coli* BL21(DE3), quite weaker bands corresponding to MjsHSP nanocages were observed under all conditions. This allowed to infer that the potential capacity for expression of recombinant proteins in *V. natriegens* has not been verified.

In order to compare and validate the efficiency of MjsHSP nanocages production in *E. coli* BL21(DE3) under the two alternative aeration conditions, key parameters namely production yield (**Equation III.5**), volumetric productivity (**Equation III.6**) and specific productivity (**Equation III.7**) were calculated. The analysis was performed using SDS-PAGE densitometric measurements as previously explained. The results for these parameters are presented in **Figure III.49**.

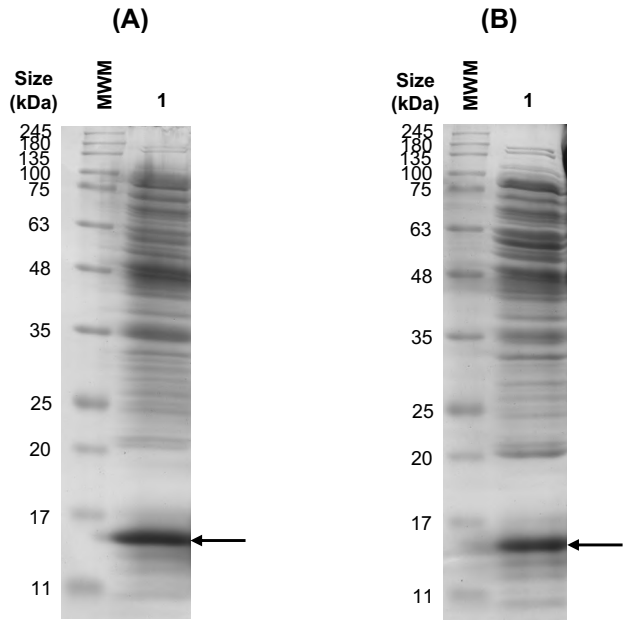
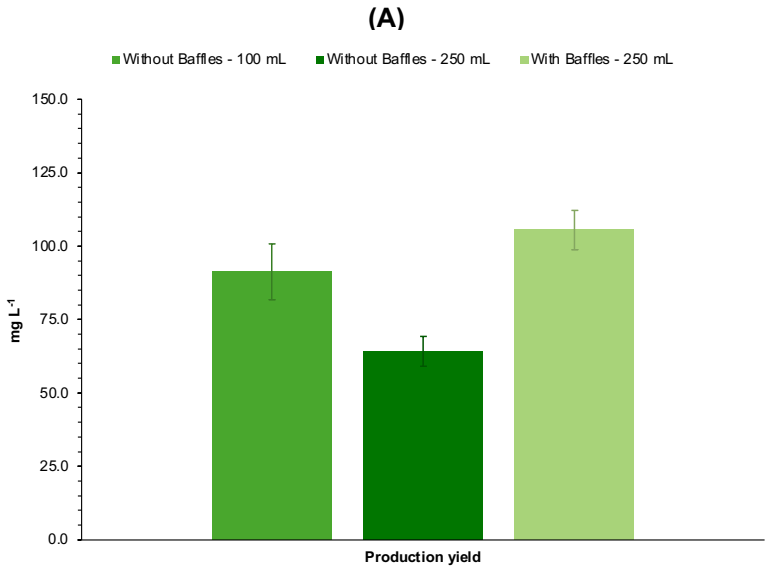


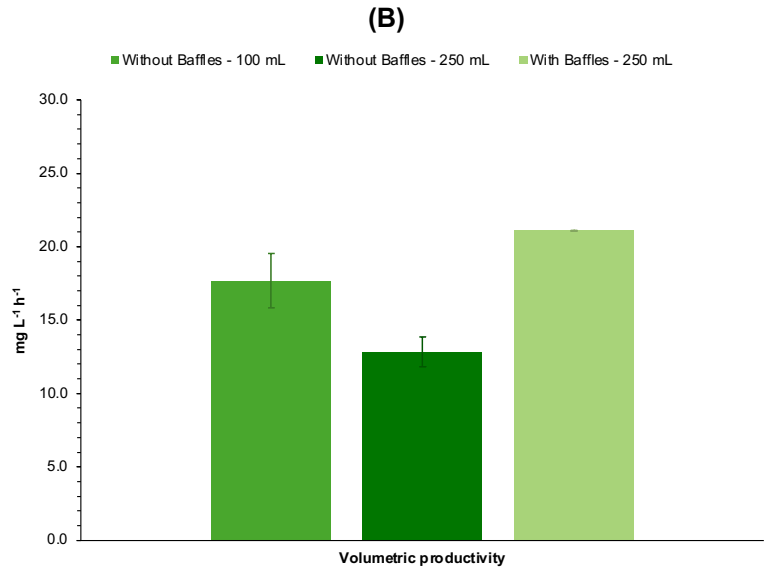
Figure III.48 – SDS-PAGE analysis obtained in the production of MjsHSP nanocages in *V. natriegens* Vmax™ X2 cells using two different aeration conditions (250 mL shake flask without and with baffles) for the induction of protein expression, maintaining the remaining parameters as previously established (enhanced 2xYT; 37°C and 250 rpm; induction of protein expression at $OD_{600\ nm}$ of 0.5-0.6, with 0.1 mM of IPTG, for 4 hours at 37°C). **(A)** 250 mL shake flask without baffles, and **(B)** 250 mL shake flask with baffles. Arrows indicate the protein bands referring to the monomer of MjsHSP nanocages (expected molecular weight of 16.5 kDa), which is separated and visualized under denaturing conditions of SDS-PAGE. The abbreviation MWM refers to the molecular weight marker (NZYColour protein marker II; NZYtech Lisbon, Portugal).

From **Figure III.49**, the expression of the MjsHSP nanocages in a 250 mL baffled shake flask allowed to obtain higher values for production yield and volumetric productivity in relation to the experience with a 250 mL non-baffled shake flask, being the aeration condition more advantageous. Interestingly, the standard procedure (100 mL non-baffled shake flask) was the second best condition, having recorded intermediate values for these two parameters.

The production yield (**Figure III.49A**) was 1.4 times lower in the 250 mL non-baffled shake flask (64.0 ± 5.1 mg L^{-1}) in relation to the baffled shake flask of 100 mL (91.3 ± 9.5 mg L^{-1}). The test with the 250 mL baffled shake flask (105.5 ± 6.8 mg L^{-1}) resulted in a production yield 1.2 times higher compared to the standard condition, standing out as the most interesting aeration setup.

Similarly, the volumetric productivity (**Figure III.49B**) showed the same behavior, being 1.6 times higher in the 250 mL baffled shake flask (21.1 ± 1.4 mg L^{-1} h^{-1}) than the result of the 250 mL non-baffled shake flask (12.8 ± 1.0 mg L^{-1} h^{-1}). With the reference non-baffled shake flask of 100 mL (17.7 ± 1.8 mg L^{-1} h^{-1}), it was possible to obtain a value about 1.2 times lower than the best condition with the 250 mL shake flask with baffles.


(A)
Without Baffles - 100 mL
Without Baffles - 250 mL
With Baffles - 250 mL
150.0
125.0
100.0
75.0
50.0
25.0
0.0
mg L-1
Production yield


(B)
Without Baffles - 100 mL
Without Baffles - 250 mL
With Baffles - 250 mL
30.0
25.0
20.0
15.0
10.0
5.0
0.0
mg L-1 h-1
Volumetric productivity

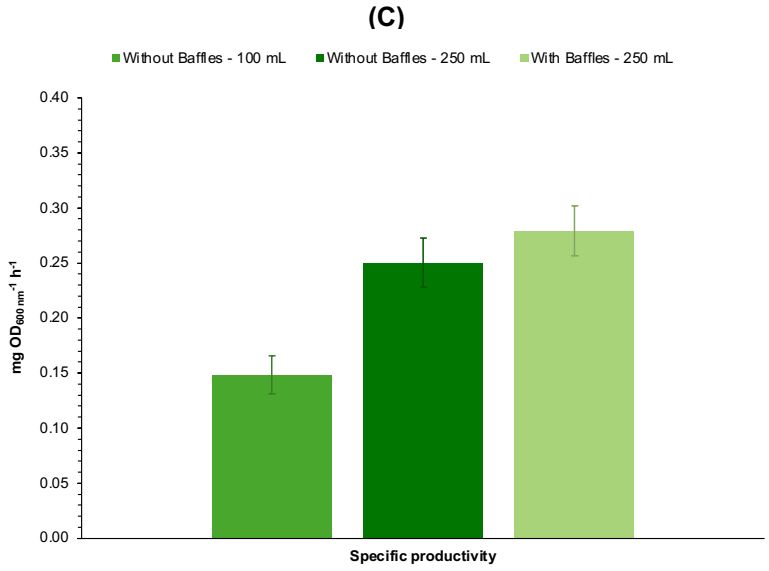

(C)
Without Baffles - 100 mL
Without Baffles - 250 mL
With Baffles - 250 mL
0.40
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
mg OD600 nm-1 h-1
Specific productivity

Figure III.49 – Parameters of MjsHSP nanocages production obtained for *V. natriegens* Vmax™ X2 cells using two different aeration conditions (250 mL shake flask without and with baffles) for the induction of protein expression, maintaining the remaining parameters as previously established (enhanced 2xYT; 37°C and 250 rpm; induction of protein expression at $OD_{600\ nm}$ of 0.5-0.6, with 0.1 mM of IPTG, for 4 hours at 37°C). **(A)** Production yield, **(B)** volumetric productivity, and **(C)** specific productivity. Standard aeration condition for induction of protein expression (100 mL shake flask without baffles) was also included as a reference point. Each result represents the mean of two biological replicates, with error bars indicating the respective standard deviation.

In terms of the specific productivity (**Figure III.49C**), the results indicated that MjsHSP nanocages production per cell was much higher in the experiment using the 250 mL baffled shake flask (0.28 ± 0.02 mg $OD_{600\ nm}^{-1}$ h^{-1}), approximately 1.1 times higher than the value obtained for the 250 mL non-baffled shake flask (0.25 ± 0.02 mg $OD_{600\ nm}^{-1}$ h^{-1}) and 1.9 times higher than the estimation of the reference condition (0.15 ± 0.02 mg $OD_{600\ nm}^{-1}$ h^{-1}).

Currently, there are no studies available in the literature that directly investigate the use of baffled shake flasks with the cell growth behavior and subsequent protein expression, so this is the first work to perform some experiments on this topic. Assuming the same interpretation considered for *E. coli* BL21(DE3), the improved oxygenation in a baffled shake flask leads to higher cell growth. This increased biomass results in greater production of protein nanocages (232). In fact, a large set of articles published using *V. natriegens* choose baffled shake flasks for cell growth and/or protein expression (157,239,240), which may suggest that this variation has a significant influence on the characteristics of this bacterial species. The results of this work corroborate this hypothesis, given that the production efficiency of MjsHSP nanocages was better applying an aeration setup with a baffled shake flask.

Comparing these characterizing parameters of the production of MjsHSP nanocages, the results obtained for *E. coli* BL21(DE3) cells after optimization considering the aeration condition for the protein expression (232.9 ± 18.8 mg L^{-1}, 41.1 ± 3.3 mg L^{-1} h^{-1} and 0.19 ± 0.02 mg $OD_{600\ nm}^{-1}$ h^{-1}) were generally superior to those recorded with the aeration setup optimized using *V. natriegens* Vmax™ X2 cells (105.5 ± 6.8 mg L^{-1}, 21.1 ± 1.4 mg L^{-1} h^{-1} and 0.28 ± 0.02 mg $OD_{600\ nm}^{-1}$ h^{-1}), with the exception of specific productivity, which is an interesting result.

In conclusion, inducing the expression of MjsHSP nanocages using a baffled shake flask (exemplified in this work by one of 250 mL), maintaining the same ratio of air volume to culture medium volume (about 30% of the latter) appears to have greater potential. Therefore, the final conditions for optimizing the production of this model protein nanocages were the use of enhanced 2xYT culture medium, an IPTG concentration of 0.1 mM, an induction temperature of 37°C, a time point in the middle of the exponential growth phase ($OD_{600\ nm}$ of approximately 0.5 to 0.6) and an aeration setup comprising a baffled shake flask.

After identifying these optimized conditions for the production of MjsHSP nanocages in *V. natriegens* Vmax™ X2 cells, similarly to what was implemented for *E. coli*, an additional experiment was performed with the monitoring and collection of samples of the bacterial culture after 2, 4, 6, 8 and 24 hours from the beginning of the induction of protein expression (**III.2.12. Optimization of production of MjsHSP nanocages**). This part of the study aimed to extend the observation period for cell growth

and protein nanocages synthesis, facilitating a comparative analysis with the baseline data from initially established standard conditions.

The growth curve obtained for the production of MjsHSP nanocages in *V. natriegens* Vmax™ X2 under the optimized conditions is shown in **Figure III.50**. Note that the growth curve for the production of these protein nanocages under the standard initial conditions is available in **Figure III.50**.

From **Figure III.50**, it was observed that the growth profile of *V. natriegens* cells under the optimized conditions and the standard conditions presented some variations, particularly in the exponential growth phase. As previously described throughout the work, from this growth curve, it was possible to determine the three growth parameters: specific growth rate, generation time, and maximum $OD_{600\ nm}$ achieved at the end of the protein expression induction period.

The values obtained for the specific growth rate ($1.88 \pm 0.19\ h^{-1}$) and the generation time (22.4 ± 2.3 minutes) were lower than the results determined for the standard conditions (**Table III.3**), with a faster growth in the reference experiment (about 1.6 times for both parameters).

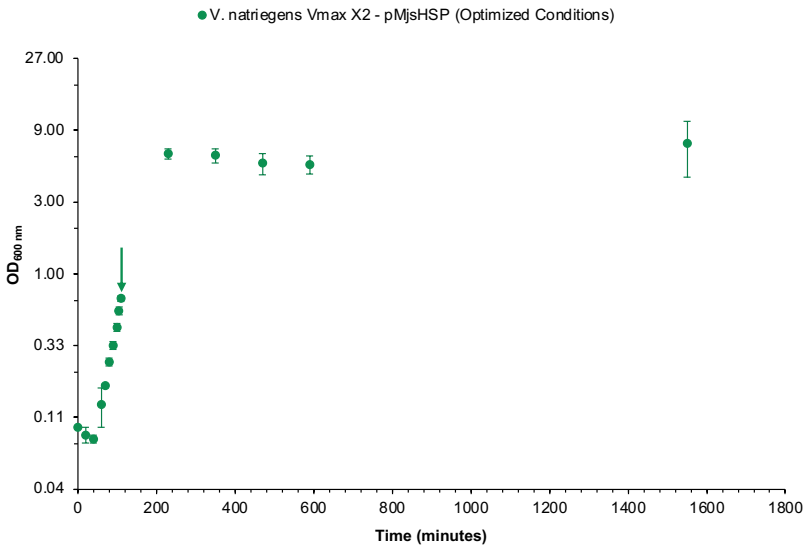


Figure III.50 – Growth curve obtained for the *V. natriegens* Vmax™ X2 cells producing MjsHSP nanocages in the optimized conditions (enhanced 2xYT culture medium, 0.1 mM of IPTG, 37°C during protein expression, time point of induction at $OD_{600\ nm} \approx 0.6$ and shake flask with baffles). Arrow indicates IPTG induction time point. Each data point represents the mean of two biological replicates, with error bars indicating the respective standard deviation. The y-axis is on a logarithmic scale.

Regarding the maximum $OD_{600\ nm}$, the growth of *V. natriegens* Vmax™ X2 cells in the optimized condition led to a much higher value (7.39 ± 2.98) if the period up to 24 hours of protein expression induction is considered, which was expected. In any case, if the values for the same induction period (4 hours) were compared, the experiment under optimized conditions (6.13 ± 0.67) also allowed obtaining a value 1.4 times higher than the standard test (**Table III.3**).

Thus, with regard to the characterization of cell growth, this comparison allowed to conclude that the optimized conditions ensure the maintenance of a promising cell growth profile, with potential for the production of MjsHSP nanocages on a large scale. The difference observed in the specific growth rate

and generation time parameters may be partly related to the volume of culture medium being higher (about 75 mL) in the optimized assay, while maintaining the air volume/culture medium volume ratio.

Complementarily, performing a comparison in relation to the results obtained for these three parameters in the optimized conditions for the production of MjsHSP nanocages in *E. coli* BL21(DE3), the values that characterize *V. natriegens* Vmax™ X2 were significantly higher in relation to the specific growth rate and generation time (about 1.4 times) but lower for the maximum $OD_{600\ nm}$ achieved (about 1.7 times). Note that if only the induction period of 4 hours is considered, the maximum $OD_{600\ nm}$ obtained was slightly higher 1.1 times in *V. natriegens* than in *E. coli*.

Afterwards, the production of these non-viral protein nanocages was assessed by obtaining an SDS-PAGE of clarified lysates (soluble fraction) containing the MjsHSP nanocages expressed in *E. coli* cells using these optimized conditions for cultivation and protein expression. **Figure III.51** shows clear bands corresponding to the MjsHSP nanocages (expected molecular weight of 16.5 kDa). As expected, the intensity of the MjsHSP nanocages band increased throughout the induction period from the IPTG addition time point, suggesting that there continued to be significant expression of the protein nanocages as cell growth progressed. It was interesting to note that the basal expression was quite reduced (**Figure III.51**, Lane 1) and that the greatest increase in the band intensity occurred after 2 hours (**Figure III.51**, Lane 2), 4 hours (**Figure III.51**, Lane 3) and 8 hours (**Figure III.51**, Lane 5) of induction, with the differences from this moment on being less considerable. The bands corresponding to the protein impurities always had a similar pattern, with some variation in their intensity and proportion in relation to the MjsHSP nanocages band throughout the experiment (particularly after 6 and 8 hours).

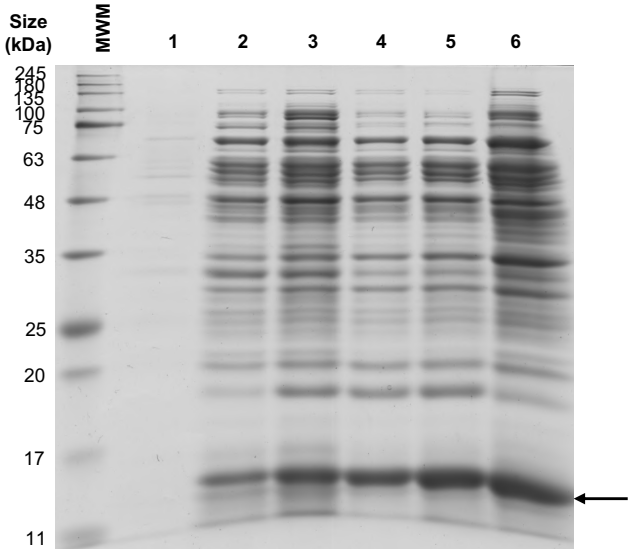


Figure III.51 – SDS-PAGE analysis obtained in the production of MjsHSP nanocages in *V. natriegens* Vmax™ X2 cells using the optimized conditions for the induction of protein expression (enhanced 2xYT culture medium, 0.1 mM of IPTG, 37°C during protein expression, time point of induction at $OD_{600\ nm} \approx 0.6$ and shake flask with baffles). Lane 1 corresponds to the sample collected from the bacterial culture at the IPTG induction time point and lanes 2 to 6 correspond to the samples collected after 2, 4, 6, 8 and 24 hours, respectively, from the beginning of the protein expression induction. Arrow indicates the protein bands referring to the monomer of MjsHSP nanocages (expected

molecular weight of 16.5 kDa), which is separated and visualized under denaturing conditions of SDS-PAGE. The abbreviation MWM refers to the molecular weight marker (NZYColour protein marker II; NZYtech Lisbon, Portugal).

In contrast to the parallel optimized experiment performed with *E. coli* BL21(DE3), the bands representing MjsHSP nanocages were clearly less intense over the various time points of sample collection. This observation suggests a substantial reduction in the amount of protein nanocages produced. Additionally, a greater proportion of bands related to protein impurities was observed in relation to the band of protein nanocages in *V. natriegens*, contrary to what occurred in *E. coli*.

In order to understand the differences in the efficiency of MjsHSP nanocages production in *E. coli* BL21(DE3) using these optimized conditions, the aforementioned main parameters, production yield (**Equation III.5**), volumetric productivity (**Equation III.6**) and specific productivity (**Equation III.7**) were determined. Analysis was performed using SDS-PAGE densitometric measurements as previously explained. The results for these parameters are presented in **Figure III.52**.

According to **Figure III.52A**, for the production yield, a gradual increase was observed throughout the monitored protein expression period. The most considerable increases occurred after 2 hours (5.8 times, for 54.2 $\pm$ 3.4 mg L^{-1}) and 24 hours (1.8 times, for 150.3 $\pm$ 96.9 mg L^{-1}). These results are compatible with a progressive production of MjsHSP nanocages over time in *V. natriegens* Vmax™ X2, in combination with a balance between this and cell growth and biomass generation. It is also important to mention that a very considerable experimental variation was detected, with high standard deviation values being obtained, so more data from additional experiments would be useful.

The volumetric productivity (**Figure III.52B**) recorded values that were in agreement with the previous interpretation, observing an equally gradual decrease in the value of this parameter over time from 14.1 $\pm$ 0.9 mg L^{-1} h^{-1} after 2 hours to 5.8 $\pm$ 3.7 mg L^{-1} h^{-1} after 24 hours. After 4 hours (10.4 $\pm$ 2.3 mg L^{-1} h^{-1}) and after 24 hours (5.8 $\pm$ 3.7 mg L^{-1} h^{-1}) were the time points at which a reduction with the greatest impact was determined, approximately 1.4 and 1.5 times, respectively, in relation to the previous moment.

Comparing the value obtained after 4 hours of induction with the same time point of the standard condition (**Table III.4**), the process optimization led to a 2.2-fold increase in the production yield and a 1.8-fold increase in the volumetric productivity.

Finally, the specific productivity (**Figure III.52C**) allowed to evaluate the production of MjsHSP nanocages per cell basis, verifying a very high value after the first 2 hours (0.008 $\pm$ 0.001 mg $OD_{600\ nm}^{-1}$ h^{-1}), with a subsequent decrease over time until 24 hours of induction of protein expression (8 times, for 0.001 $\pm$ 0.001 mg $OD_{600\ nm}^{-1}$ h^{-1}). During the early phase of cell growth, bacterial cells exhibited enhanced metabolic activity due to abundant resources, potentially explaining increased production at the individual cell level. Once more, comparing the value obtained after 4 hours of induction with the same time point of the condition defined as standard (**Table III.4**), the process optimization led to an 10-fold reduction in the specific productivity (for 0.004 $\pm$ 0.001 mg $OD_{600\ nm}^{-1}$ h^{-1}).

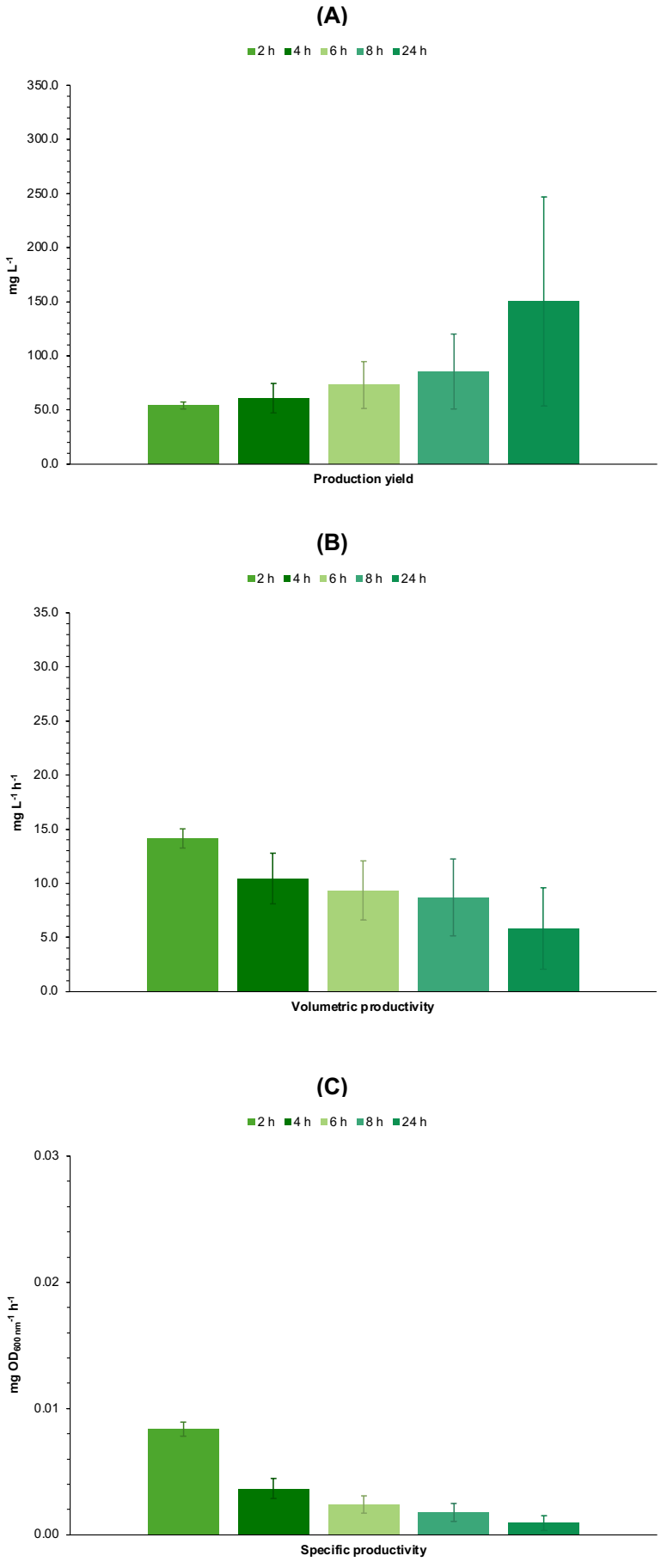

(A)
2 h
4 h
6 h
8 h
24 h
350.0
300.0
250.0
200.0
150.0
100.0
50.0
0.0
mg L-1
Production yield
(B)
2 h
4 h
6 h
8 h
24 h
35.0
30.0
25.0
20.0
15.0
10.0
5.0
0.0
mg L-1 h-1
Volumetric productivity
(C)
2 h
4 h
6 h
8 h
24 h
0.03
0.02
0.01
0.00
mg OD600 nm-1 h-1
Specific productivity

Figure III.52 – Parameters of MjsHSP nanocages production obtained for *V. natriegens* Vmax™ X2 cells using the optimized conditions for the induction of protein expression (enhanced 2xYT culture medium, 0.1 mM of IPTG, 37°C during protein expression, time point of induction at $OD_{600\ nm} \approx 0.6$ and shake flask with baffles). **(A)** Production yield, **(B)** volumetric productivity, and **(C)** specific productivity determined considering the samples collected from the bacterial culture after 2, 4, 6, 8 and 24 hours from the beginning of the protein expression induction. Each result represents the mean of two biological replicates, with error bars indicating the respective standard deviation.

Comparing these characterizing parameters of the production of MjsHSP nanocages, the results obtained for *E. coli* BL21(DE3) cells applying the optimized conditions were always superior throughout the entire monitoring period compared to those recorded with the complete optimized conditions using *V. natriegens* Vmax™ X2 cells. Looking only at the 4 hours interval of protein expression induction, the behavior was the same.

Taken together, these results suggest that a promising time point to finish the production of MjsHSP nanocages is close to 8 hours of induction, since up to this moment it is possible to obtain a favorable balance between the parameters analyzed, also including the behavior and characteristics of cell growth. For the study with *E. coli* BL21(DE3), the apparently most promising condition was near to the 6 hours of induction.

In conclusion, comparing the several results obtained for the standard and optimized experiments, it was possible to successfully optimize this production process of MjsHSP nanocages in *vibrio* cells. Although it has not been possible to obtain an increase in protein expression equal to or greater than that obtained with *E. coli* cells, the numerous data generated can be used as a starting point for future research studies, including in scale-up approaches. Being a complementary host organism to *E. coli* in terms of expression of recombinant proteins, *V. natriegens* has significant potential as a microbial chassis, due to its rapid growth and high metabolic rate.

III.3.6. Validation of plasmid pTRAP

The pTRAP plasmid was kindly provided by the Heddle Lab (Kraków, Poland) and was used without further modifications. Its plasmid map is represented in **Figure III.53** and the respective analysis by agarose gel electrophoresis of the plasmid single and double digested with restriction enzymes is shown in **Figure III.54**. Note that pTRAP was also verified by Sanger sequencing.

Briefly, plasmid pTRAP integrates all elements of the well-described pET-21b(+) plasmid (including the origins of replication ColE1/pMB1/pBR322/pUC and f1) with the TRAP-K35C gene.

With the pTRAP validated, chemically competent *E. coli* DH5α cells (**III.2.4. Preparation of chemically competent cells**) were transformed by the heat shock method with this plasmid as described in section **III.2.6. Transformation of competent cells**. Having confirmed the presence of the plasmid in this bacterial strain, E. coli DH5α cells harboring the pTRAP were cultured at shake flask scale, with the subsequent preparation of master and working cell banks (**III.2.2. Bacterial strains**), in order to ensure maintenance of the plasmid during all experiments.

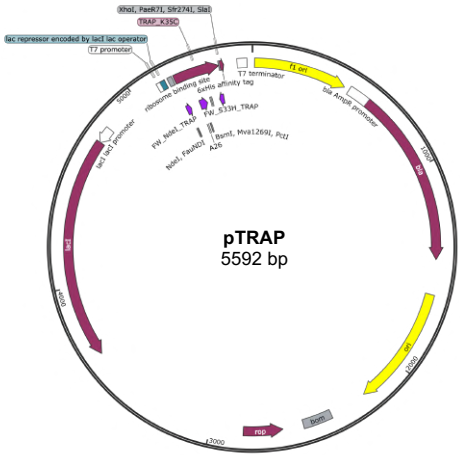


Figure III.53 – Plasmid pTRAP provided by the Heddle Lab (Kraków, Poland). The plasmid map presented was created using the SnapGene 3.2.1. software.

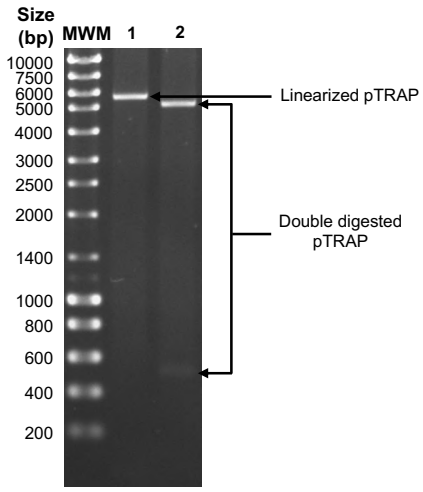


Figure III.54 – Agarose gel analysis of the plasmid pTRAP. This plasmid was linearized with the restriction enzyme XhoI (lane 1, band size of 5592 bp) and double digested with the restriction enzymes SphI-HF and XhoI (lane 2, band sizes of 5070 bp and 522 bp). The abbreviation MWM refers to the molecular weight marker (NZYDNA ladder III; NZYtech Lisbon, Portugal).

III.3.7. Growth profile of *E. coli* harboring pTRAP

Afterwards, plasmid pTRAP was transformed into chemically competent *E. coli* BL21(DE3) cells (**III.2.4. Preparation of chemically competent cells**) by the heat shock strategy (**III.2.6. Transformation of competent cells**), considering that this most appropriate strain was selected for the protein expression in this work. Master and working cell banks were also prepared (**III.2.2. Bacterial strains**).

Regarding *V. natriegens* Vmax™ X2, the initial objective was to transform cells of this bacterial strain with the pTRAP plasmid, allowing the development of a parallel study of *E. coli* vs. *V. natriegens* focused on this plasmid, as it was performed with pMjsHSP. However, different methodologies for the

preparation of chemically competent and electrocompetent cells (described in sections **III.2.4. Preparation of chemically competent cells** and **III.2.5. Preparation of electrocompetent cells**) as well as for their transformation by heat shock and electroporation, respectively (explained in the section **III.2.6. Transformation of competent cells**), were tested, without it being possible to obtain positive results. This difficulty in the transformation of *V. natriegens* can be explained by the source organism of the plasmid being *E. coli*, since it is reported in the literature that transformation efficiencies are higher when a plasmid produced and extracted from the same host organism is used (219). The fact that the pTRAP is a plasmid with significantly larger dimensions than the pMjsHSP (5592 bp vs. 2677 bp) may have contributed to the difficulty of transformation. In terms of the preparation of competent cells, regardless of the protocol used, the cells of this specific strain of *V. natriegens* may not be sufficiently competent to achieve the minimum efficiency required for their transformation. Regarding the transformation itself, the hypothesis that the cell recovery period would be extended and that a more suitable recovery solution would be used, was analyzed, without having the desired impact. All these evaluations were performed based on the information described in the literature as well as the data provided by the company responsible for the manufacturing of *V. natriegens* Vmax™ X2 cells (Codex DNA; San Diego, CA, USA).

With the purpose of obtaining and analyzing the growth profiles of *E. coli* BL21(DE3) cells harboring the plasmid pTRAP, in particular the exponential growth phase, bacterial culture was performed on a shake flask scale according to the operating conditions described in the section **III.2.8. Determination of growth curves**.

The growth curve obtained for *E. coli* BL21(DE3) harboring the pTRAP is shown in **Figure III.55**.

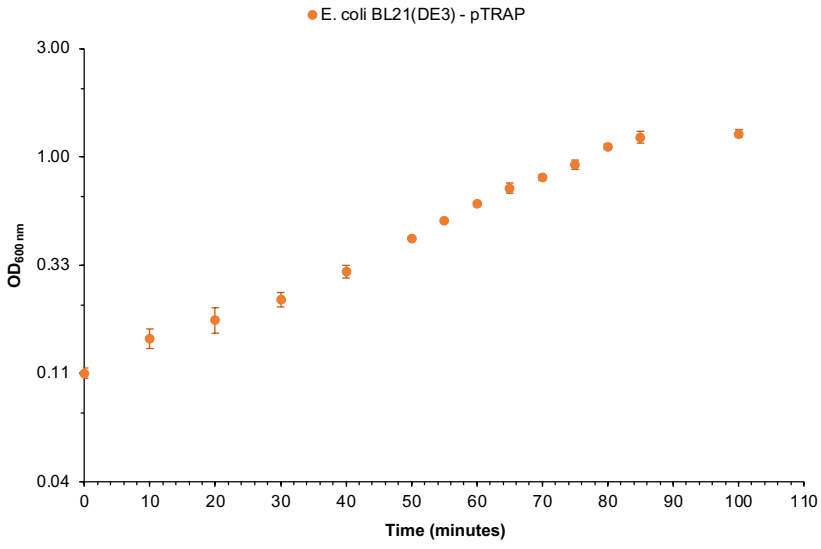


Figure III.55 – Growth curve obtained for the *E. coli* BL21(DE3) (●) cells harboring the plasmid pTRAP. Bacterial culture was performed at a shake flask scale using LB, as detailed in section **III.2.8. Determination of growth curves**. Each data point represents the mean of three biological replicates, with error bars indicating the respective standard deviation. The y-axis is on a logarithmic scale.

The analysis of this growth curve allowed to identify a lag phase in the initial 10 minutes, in which the bacterial cells exhibit metabolic activity as they adapt to the fresh culture conditions following inoculation from the overnight culture. The exponential growth phase was then observed between approximately 10 and 80 minutes, and it was possible to determine a value for the specific growth rate of approximately 1.73 ± 0.06 h^{-1} and for the generation time of about 24.1 ± 0.8 min. These results obtained are slightly superior than those reported in the literature for the BL21(DE3) strain of *E. coli* containing a plasmid (0.10 to 0.60 h^{-1}) (209–211). Conducting these assays under optimal conditions for the growth of this bacterial strain (for example, culture medium composition, temperature, and aeration parameters) may have favorably influenced the results. Additionally, the pTRAP plasmid might have imposed only a minimal metabolic burden on the host cells, which could explain its limited impact on bacterial growth. This observation contrasts with the typical scenario observed in plasmid-harboring bacterial strains, where the presence of plasmids often leads to an increased metabolic load and a consequent delay in growth kinetics (212). From this growth curve it was also possible to understand that the deceleration phase and subsequent stationary phase began after 80 minutes of bacterial growth. The maximum cell density was reached after 100 minutes, with a value of 1.27 ± 0.05 being obtained for the $OD_{600\ nm}$.

III.3.8. Determination of plasmid copy number for pTRAP in *E. coli*

Similar to what was performed for the pMjsHSP, the determination of the plasmid copy number in *E. coli* BL21(DE3) cells aimed to understand the behavior of this parameter for the plasmid pTRAP, analyze its variation throughout different phases of bacterial growth and obtain data to correlate with the expression of the TRAP O-rings.

As described in section **III.2.9. Determination of plasmid copy number by qPCR**, obtaining standard curves for gDNA from the BL21(DE3) strain of *E. coli* used as well as for pDNA (pTRAP) extracted from bacterial cells of the respective strain was the first step. These standard curves are available in **Supplementary Figure III.S3** and **Supplementary Figure III.S4**.

Through the respective linear regressions of these two standard curves and using **Equation III.3**, it was possible to determine their percentage of amplification efficiency. Efficiency values between 99.9 and 105.5% were obtained, which are within the range recommended in the literature (between 90 and 110%) (216). As previously explained, this means that the DNA samples do not contain impurities and contaminants in sufficient quantities to inhibit and/or influence the qPCR reactions. Moreover, the coefficient of determination was greater than 0.992 in the two standard curves, which also indicates a high precision in the qPCR performance (216).

Similar to what was performed for the pMjsHSP, the plasmid copy number was calculated using **Equation III.2** and considered the C_t values of each sample and also the amplification efficiencies determined from the gDNA and pDNA standard curves. The accuracy of the estimated plasmid copy number is critically dependent on the similar amplification efficiency of the reference gene (*dnaE* in genome of *E. coli*) and the target gene (*ampr* in pTRAP). As mentioned before, when this condition is

not met, traditional standard curve methods can lead to significant overestimation of plasmid copy number values, as described in the literature (205,217).

The results obtained for the plasmid copy number of *E. coli* BL21(DE3) cells harboring the plasmid pTRAP at different stages of the growth profile (corresponding to an $OD_{600\ nm}$ of 0.1, 0.3, 0.5, 0.7, 0.9, 1.1 and 1.3) are represented in **Figure III.56**.

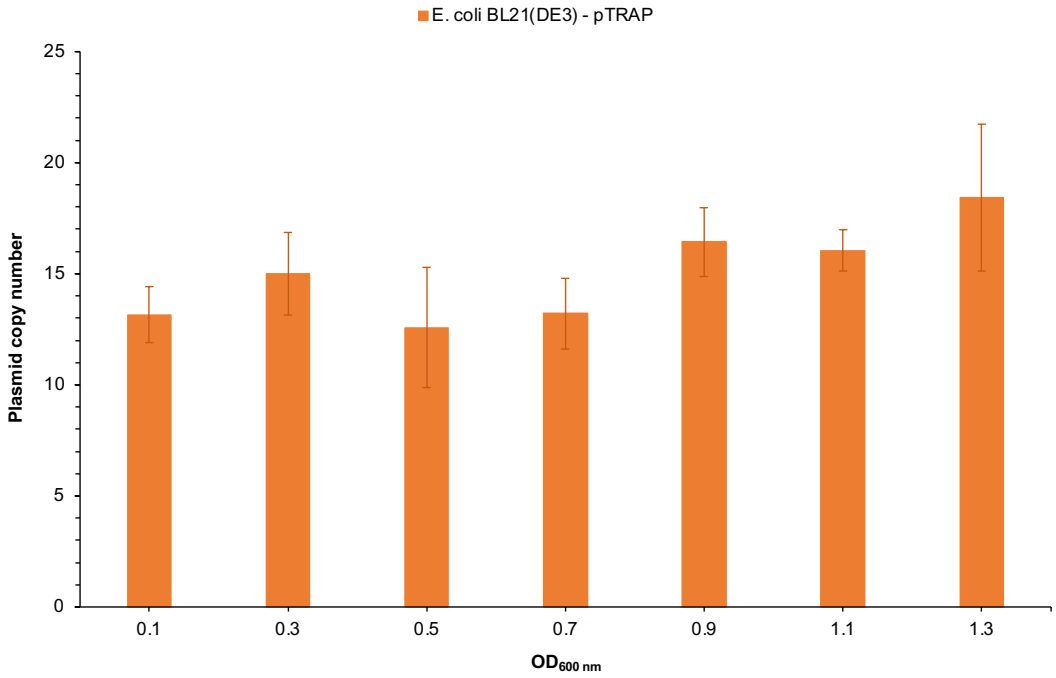


Figure III.56 – Plasmid copy number values obtained for *E. coli* BL21(DE3) cells harboring the plasmid pTRAP at different growth times (corresponding to an $OD_{600\ nm}$ of 0.1, 0.3, 0.5, 0.7, 0.9, 1.1 and 1.3). These results were determined through a qPCR analysis, following the experimental procedure described in the section **III.2.9. Determination of plasmid copy number by qPCR**. Each data point represents the mean of three biological replicates, with error bars indicating the respective standard deviation.

Throughout the monitored growth period, it was possible to observe a variation in the plasmid copy number with an increasing trend. At the beginning of growth with an $OD_{600\ nm}$ of 0.1, a plasmid copy number of about 13 ± 1 was obtained, followed by a stabilization of this value (13 ± 3) in the middle of the exponential growth phase ($OD_{600\ nm}$ equal to 0.5). From this moment until the beginning of the stationary growth phase ($OD_{600\ nm}$ of 1.3), there was a gradual increase in the plasmid copy number reaching 18 ± 3. Considering this entire growth profile, the average plasmid copy number was 15 ± 1, which is a different result than expected for the plasmid pTRAP, given that it has a hybrid ColE1/pMB1/pBR322/pUC origin of replication (in addition to the bacteriophage f1 origin of replication), that is described as having a high plasmid copy number in *E. coli*. These results end up being quite similar to those obtained with the plasmid pMjsHSP, so that the aspects discussed in this section (**III.3.3. Determination of plasmid copy number for pMjsHSP in *E. coli* and *V. natriegens***) can also be applied to the pTRAP. A complete study available in the literature revealed that *E. coli* exhibited a higher plasmid copy number (149 copies) for a plasmid with the same origin of replication (204). Host cell

physiology likely explains the difference observed, as plasmid copy number is closely related to bacterial replication machinery, which depends on the host's physiological state (218). The high plasmid copy number observed aligns with the previously discussed growth profile of *E. coli* cells containing pTRAP. These cells exhibited a rapid growth rate, suggesting that plasmid replication did not significantly burden the cells' metabolism.

Furthermore, an increase in plasmid copy number at $OD_{600\ nm}$ of 1.3 was observed, an expected result, given that it is reported that metabolic resources are redirected to prioritize plasmid replication as the stationary growth phase begins.

These results may impact protein expression, though some studies show that lower plasmid copy number does not always reduce protein production and can actually benefit bacterial host cells by preventing metabolic overload (220).

III.3.9. Production of TRAP-O rings in *E. coli*

With the objective of assessing the suitability of *E. coli* BL21(DE3) as production host for TRAP O-rings, first its growth characteristics under the standard conditions for cultivation and protein expression (non-baffled shake flask with 30% volume of culture medium; LB; 37°C and 250 rpm; induction of protein expression at $OD_{600\ nm}$ of 0.5-0.6, with 1 mM IPTG, for 4 hours at 37°C) were analyzed. The growth curve obtained is shown in **Figure III.57**.

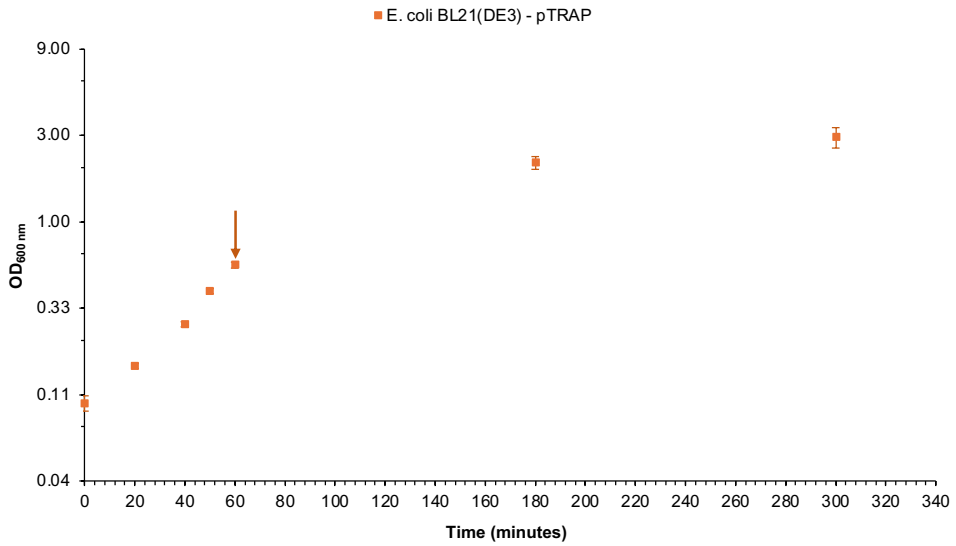


Figure III.57 – Growth curve obtained for the *E. coli* BL21(DE3) (■) cells producing TRAP O-rings in standard conditions for cultivation and protein expression (non-baffled shake flask with 30% volume of culture medium; LB; 37°C and 250 rpm; induction of protein expression at $OD_{600\ nm}$ of 0.5-0.6, with 1 mM IPTG, for 4 hours at 37°C). Arrow indicates IPTG induction time point. Each data point represents the mean of three biological replicates, with error bars indicating the respective standard deviation. The y-axis is on a logarithmic scale.

Through the analysis of the growth curve, it was possible to determine the growth parameters, specific growth rate and generation time, together with maximum $OD_{600\ nm}$ achieved during monitoring. These data are represented in **Table III.5**.

Table III.5 – Growth parameters of *E. coli* BL21(DE3) cells producing TRAP O-rings in standard conditions for cultivation and protein expression (non-baffled shake flask with 30% volume of culture medium; LB; 37°C and 250 rpm; induction of protein expression at $OD_{600\ nm}$ of 0.5-0.6, with 1 mM IPTG, for 4 hours at 37°C). Each result represents the mean of three biological replicates, with error bars indicating the respective standard deviation.

Growth parameters	
Specific growth rate (h^{-1})	1.77 ± 0.08
Generation time (min)	23.5 ± 1.0
Maximum $OD_{600\ nm}$	2.95 ± 0.38

E. coli BL21(DE3) cells harboring the plasmid pTRAP exhibited a significantly fast growth behavior, with very interesting values of specific growth rate (1.77 ± 0.08 h^{-1}) and generation time (23.5 ± 1.0 min). Comparing, for example, the results obtained previously for *E. coli* cells from the same strain containing the plasmid pMjsHSP, the values determined for both parameters are very similar. Furthermore, regarding the maximum $OD_{600\ nm}$ reached (2.95 ± 0.38), this value was obtained at the end of the protein expression induction period. Again, the results analyzed were very similar to those achieved in *E. coli* BL21(DE3) with pMjsHSP.

Afterwards, the next step consisted of evaluating the production of TRAP O-rings by *E. coli* BL21(DE3) cells.

This production was assessed by obtaining an SDS-PAGE of the clarified lysate (soluble fraction) containing the TRAP O-rings expressed using the standard conditions for cultivation and protein expression. **Figure III.58** shows a clear band corresponding to the TRAP O-rings (expected molecular weight of 8.3 kDa). Note that once SDS-PAGE is denaturing, the protein band referring to the TRAP O-rings effectively correspond to their monomer. The intensity of the TRAP O-rings band seems quite strong, suggesting higher expression level. Analyzing the profile of bands obtained, there appears to be a presence of a reasonable amount of protein impurities, which can be an issue for the remaining biomanufacturing process.

To quantitatively evaluate the production efficiency of TRAP O-rings in *E. coli* BL21(DE3), three key characterization parameters previously mentioned were determined using **Equation III.5**, **Equation III.6** and **Equation III.7**, respectively: production yield, volumetric productivity, and specific productivity. As explained in the section **III.3.4. Production of MjsHSP nanocages in *E. coli* and *V. natriegens***, the determination of these three parameters implied a densitometric analysis of SDS-PAGE, which was performed according to the methodology described above (**III.2.14.3. Protein nanocages quantification by densitometric analysis of SDS-PAGE**). Due to the composition of each TRAP O-

ring in 11 identical monomers, the total mass of O-rings monomers is identical to the mass of the assembled O-rings. The results obtained for these three parameters are shown in **Table III.6**.

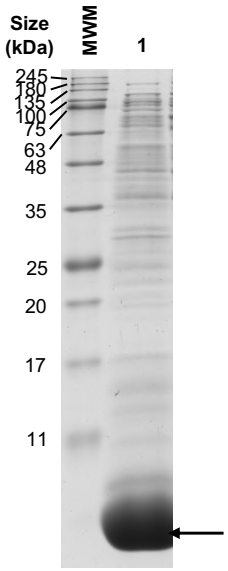


Figure III.58 – SDS-PAGE analysis obtained in the production of TRAP O-rings in *E. coli* BL21(DE3) cells considering the standard conditions for cultivation and protein expression (non-baffled shake flask with 30% volume of culture medium; LB; 37°C and 250 rpm; induction of protein expression at $OD_{600\ nm}$ of 0.5-0.6, with 1 mM IPTG, for 4 hours at 37°C). Arrow indicates the protein band referring to the monomer of TRAP O-rings (expected molecular weight of 8.3 kDa), which is separated and visualized under denaturing conditions of SDS-PAGE. The abbreviation MWM refers to the molecular weight marker (NZYColour protein marker II; NZYtech Lisbon, Portugal).

Table III.6 – Parameters of TRAP O-rings production (production yield, volumetric productivity and specific productivity) obtained for *E. coli* BL21(DE3) cells considering the standard conditions for cultivation and protein expression (non-baffled shake flask with 30% volume of culture medium; LB; 37°C and 250 rpm; induction of protein expression at $OD_{600\ nm}$ of 0.5-0.6, with 1 mM IPTG, for 4 hours at 37°C). Each result represents the mean of three biological replicates, with error bars indicating the respective standard deviation.

Parameters of TRAP O-rings production	
Production yield (mg L^{-1})	182.9 ± 12.1
Volumetric productivity (mg L^{-1} h^{-1})	36.6 ± 2.4
Specific productivity (mg $OD_{600\ nm}^{-1}$ h^{-1})	0.38 ± 0.03

The analysis of **Table III.6** allowed to conclude that *E. coli* BL21(DE3) is a suitable and advantageous host for the production of TRAP O-rings under standard cultivation and protein expression conditions. Obtaining a production yield of 182.9 ± 12.1 mg L^{-1} together with a volumetric productivity and a specific productivity of 36.6 ± 2.4 mg L^{-1} h^{-1} and 0.38 ± 0.03 mg $OD_{600\ nm}^{-1}$ h^{-1},

respectively, are very promising results and a good starting point for further optimization of the protein expression conditions.

As previously discussed, limited published information related to the manufacturing process of NVPNs, particularly regarding to the upstream steps, hinders comparison and contextualization of the results obtained in this work with those determined in other studies using TRAP O-rings as model protein. As an example and comparing with the results obtained for the expression of MjsHSP nanocages also in *E. coli* BL21(DE3) cells, the production of TRAP O-rings in the same strain of *E. coli* presents considerably higher values for these three parameters (about 3 times for production yield and volumetric productivity, and 4 times for specific productivity).

III.3.10. Optimized production of TRAP-O rings in *E. coli*

The optimization of the production of TRAP O-rings in *E. coli* BL21(DE3) cells focused on the improvement of the previously implemented standard procedure (**III.3.9. Production of TRAP-O rings in *E. coli***). Furthermore, the observations and conclusions acquired in the optimization of the production of MjsHSP nanocages were considered (**III.3.5. Optimized production of MjsHSP nanocages in *E. coli* and *V. natriegens***), with the transposition of the most favorable operational and expression conditions for the production of TRAP O-rings.

The optimization strategy involved the production of TRAP O-rings with the best operating conditions in the five crucial parameters selected (growth medium composition, concentration of protein expression inducer, temperature of protein expression, timepoint of induction and aeration conditions). The analysis and evaluation of the results obtained allowed to understand the influence of these improved protein expression conditions on the production yield, volumetric productivity and specific productivity.

After identifying the optimized conditions for the production of MjsHSP nanocages in *E. coli* BL21(DE3) cells, an experiment adapted for the expression of TRAP O-rings was performed with the monitoring and collection of samples of the bacterial culture after 2, 4, 6, 8 and 24 hours from the beginning of the induction of protein expression (**III.2.13. Optimization of production of TRAP O-rings**).

It should be noted that the optimized conditions included the use of enhanced 2xYT culture medium, an IPTG concentration of 0.1 mM, an induction temperature of 37°C, a time point in the middle of the exponential growth phase ($OD_{600\ nm}$ of approximately 0.5 to 0.6) and an aeration setup comprising a non-baffled shake flask.

The growth curve obtained for the production of TRAP O-rings in *E. coli* BL21(DE3) under the optimized conditions is shown in **Figure III.59**. Note that the growth curve for the production of these protein nanocages under the standard initial conditions is available in **Figure III.57**.

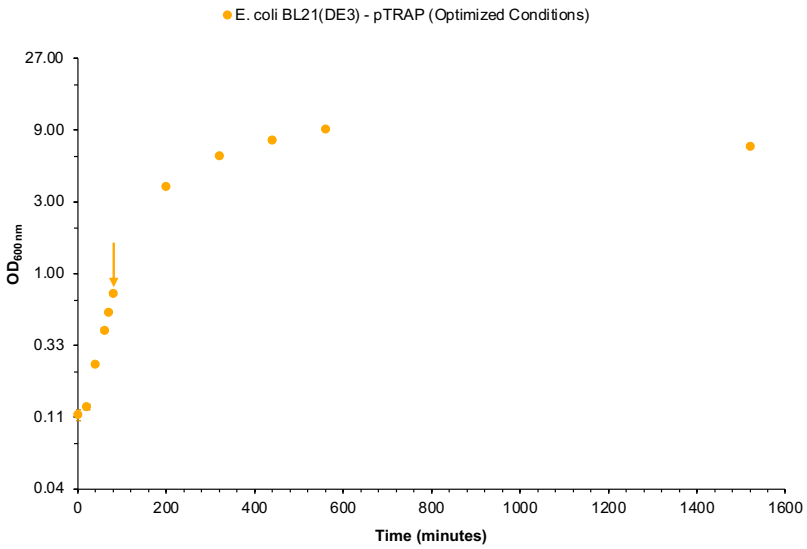


Figure III.59 – Growth curve obtained for the *E. coli* BL21(DE3) cells producing TRAP O-rings in the optimized conditions (enhanced 2xYT culture medium, 0.1 mM of IPTG, 37°C during protein expression, time point of induction at $OD_{600\ nm} \approx 0.6$ and shake flask without baffles). Arrow indicates IPTG induction time point. Each data point represents the mean of two biological replicates, with error bars indicating the respective standard deviation. The y-axis is on a logarithmic scale.

Analysis of **Figure III.59** revealed a similarity between the growth profiles of *E. coli* BL21(DE3) cells under optimized and standard conditions, especially during the exponential growth phase. This similarity was evident despite the differences in the time intervals considered. Using the same approach previously described, from this growth curve, it was possible to determine the growth parameters, specific growth rate, generation time, and maximum $OD_{600\ nm}$ achieved at the end of the protein expression induction period.

The values obtained for the specific growth rate ($1.56 \pm 0.04\ h^{-1}$) and the generation time (26.6 ± 0.6 minutes) were very close to the results determined for the standard conditions (**Table III.5**), with slightly faster growth in the reference experience (approximately 1.1 times for both parameters), although with a non-significant difference.

Regarding the maximum $OD_{600\ nm}$, the growth of *E. coli* BL21(DE3) cells in the optimized condition led to a much higher value (7.00 ± 0.11) for the first 24 hours after initiating protein expression, which was expected. Regardless, if the values for the same induction period (4 hours) were compared, the experiment under optimized conditions (6.09 ± 0.15) also allowed obtaining a value 2.1 times higher than the standard test (**Table III.5**).

Thus, comparative analysis of the cell growth revealed that the optimized conditions provided a highly favorable cell development profile. This profile not only maintained desirable bacterial growth characteristics but also suggested a significant potential for larger scale production of TRAP O-rings.

Next, the effectiveness of the MjsHSP nanocages production was evaluated. This assessment involved analyzing the soluble fraction of clarified cell lysates using SDS-PAGE. The lysates were derived from *E. coli* cells cultivated and induced to express the protein nanocages under the newly optimized conditions. **Figure III.60** shows clear bands corresponding to the TRAP O-rings (expected molecular

weight of 8.3 kDa). Since SDS-PAGE is denaturing, the protein bands referring to the TRAP O-rings effectively correspond to their monomer. The intensity of the TRAP O-rings band increased significantly after 2 hours (**Figure III.60**, Lane 2) and 4 hours (**Figure III.60**, Lane 3) of induction, and from this moment onwards and contrary to expectations, the variation in the intensity of the respective bands was not considerable, suggesting that the expression level ended up stabilizing. It was also found that the basal expression was significantly reduced (**Figure III.60**, Lane 1) and that the bands corresponding to protein impurities increased their intensity as the process of cell growth and protein production progressed over time, presenting a higher proportion in relation to the TRAP O-rings band.

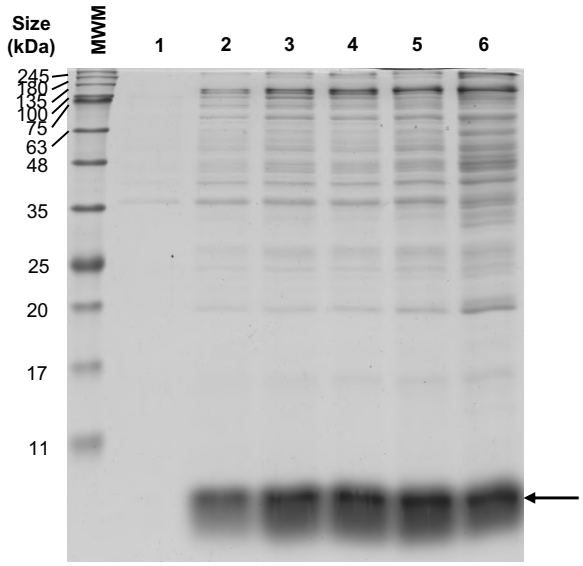
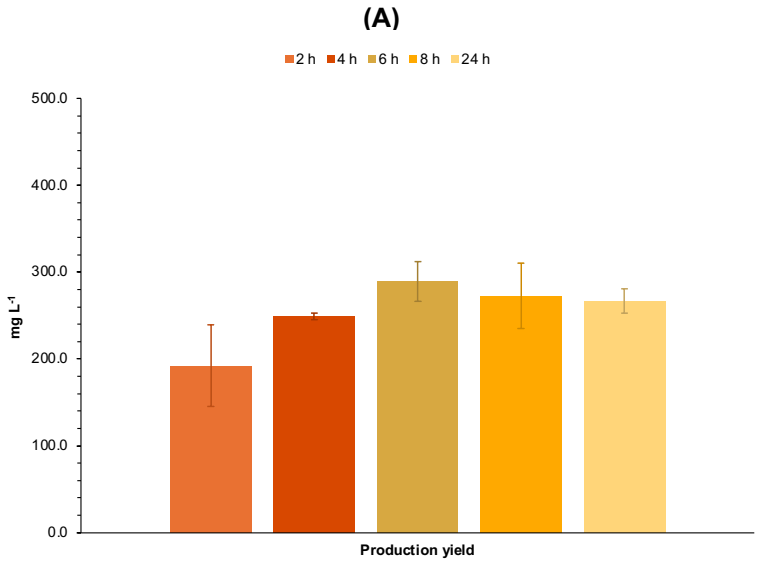
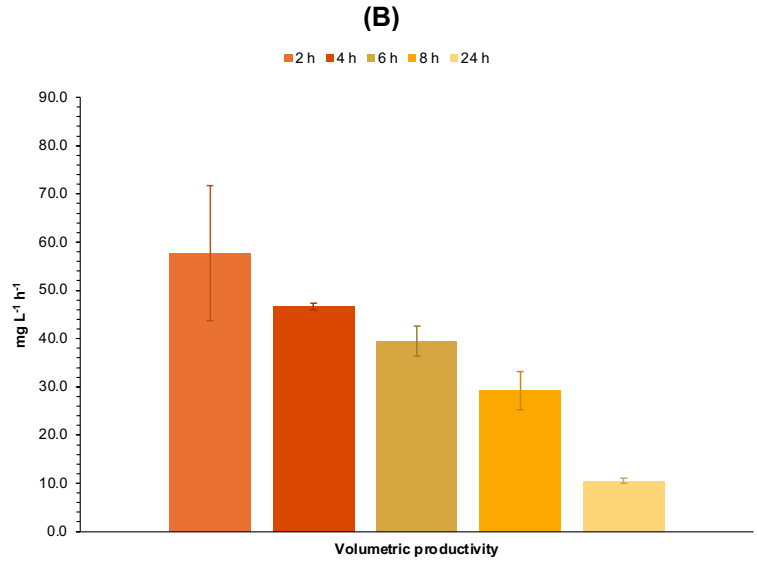


Figure III.60 – SDS-PAGE analysis obtained in the production of TRAP O-rings in *E. coli* BL21(DE3) cells using the optimized conditions for the induction of protein expression (enhanced 2xYT culture medium, 0.1 mM of IPTG, 37°C during protein expression, time point of induction at $OD_{600\ nm} \approx 0.6$ and shake flask without baffles). Lane 1 corresponds to the sample collected from the bacterial culture at the IPTG induction time point and lanes 2 to 6 correspond to the samples collected after 2, 4, 6, 8 and 24 hours, respectively, from the beginning of the protein expression induction. Arrow indicates the protein band referring to the monomer of TRAP O-rings (expected molecular weight of 8.3 kDa), which is separated and visualized under denaturing conditions of SDS-PAGE. The abbreviation MWM refers to the molecular weight marker (NZYColour protein marker II; NZYtech Lisbon, Portugal).

The effectiveness of the optimized conditions on the production of TRAP O-rings in *E. coli* BL21(DE3) was evaluated through a comprehensive analysis, focusing on three crucial parameters: production yield, volumetric productivity and specific productivity, calculated respectively by **Equation III.5**, **Equation III.6** and **Equation III.7**. To quantify these parameters, densitometric measurements of SDS-PAGE were used, a strategy previously validated in this work. The results of these parameters are shown graphically in **Figure III.61**.


(A)
2 h
4 h
6 h
8 h
24 h
500.0
400.0
300.0
200.0
100.0
0.0
mg L-1
Production yield


(B)
2 h
4 h
6 h
8 h
24 h
90.0
80.0
70.0
60.0
50.0
40.0
30.0
20.0
10.0
0.0
mg L-1 h-1
Volumetric productivity

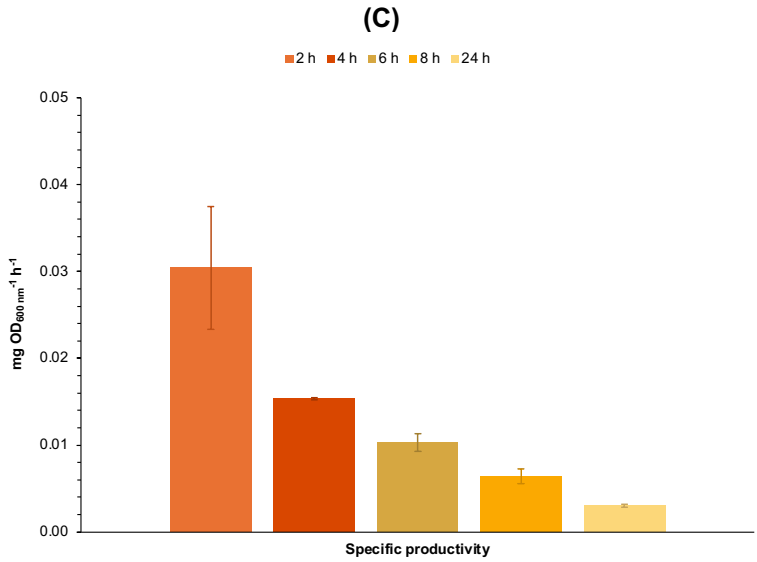

(C)
2 h
4 h
6 h
8 h
24 h
0.05
0.04
0.03
0.02
0.01
0.00
mg OD600 nm-1 h-1
Specific productivity

Figure III.61 – Parameters of TRAP O-rings production obtained for *E. coli* BL21(DE3) cells using the optimized conditions for the induction of protein expression (enhanced 2xYT culture medium, 0.1 mM of IPTG, 37°C during protein expression, time point of induction at $OD_{600\ nm} \approx 0.6$ and shake flask without baffles). **(A)** Production yield, **(B)** volumetric productivity, and **(C)** specific productivity determined considering the samples collected from the bacterial culture after 2, 4, 6, 8 and 24 hours from the beginning of the protein expression induction. Each result represents the mean of two biological replicates, with error bars indicating the respective standard deviation.

Analyzing **Figure III.61A**, a gradual increase in the production yield was observed until approximately 6 hours of induction of protein expression for 289.2 ± 23.0 mg L^{-1}, and from this moment onwards there was a subtle decrease and subsequent stabilization of the value of this parameter in the remaining monitoring period. Between 2 hours (192.3 ± 46.8 mg L^{-1}) and 4 hours (249.1 ± 3.7 mg L^{-1}) of protein expression induction, there was an increase of 1.3 times, and between 4 hours and 6 hours of induction, the increase was 1.2 times. As previously discussed, these experiments proved that for the production of TRAP O-rings, advancing growth to the late exponential phase is not advantageous, a period that is typically associated with a change in the metabolic activity of the bacterial cells and a reduction in the available resources.

For the volumetric productivity (**Figure III.61B**), a clear decrease was recorded throughout the induction of TRAP O-rings expression, with the highest value (57.7 ± 14.0 mg L^{-1} h^{-1}) being reached immediately after 2 hours of induction and decreasing approximately 5.5 times until the lowest value obtained after 24 hours of protein expression induction (10.5 ± 0.6 mg L^{-1} h^{-1}). This result allowed to conclude that in terms of productivity, there is no benefit in extending upstream processing too much, with a better option being a balance between production yield and volumetric productivity parameters.

Comparing the value obtained after 4 hours of induction with the same time point of the standard condition (**Table III.6**), the process optimization led to a 1.4-fold increase in the production yield and a 1.3-fold increase in the volumetric productivity (for 46.7 ± 0.7 mg L^{-1} h^{-1}).

Finally, the specific productivity (**Figure III.61C**) allowed to evaluate the production of TRAP O-rings per cell basis, verifying a very high value after the first 2 hours (0.03 ± 0.01 mg $OD_{600\ nm}^{-1}$ h^{-1}), with a subsequent decrease over time until 24 hours of induction of protein expression (10 times, for 0.003 ± 0.000 mg $OD_{600\ nm}^{-1}$ h^{-1}). In the early stages of growth, bacterial cells have faster metabolism and access to abundant resources. This favorable condition may explain the higher productivity observed at the individual cellular level at this period. Again, comparing the value obtained after 4 hours of induction with the same time point of the condition defined as standard (**Table III.6**), the process optimization led to an 19-fold reduction in the specific productivity (for 0.02 ± 0.00 mg $OD_{600\ nm}^{-1}$ h^{-1}).

All these data together indicated that performing the induced expression of TRAP O-rings for a period of approximately 6 hours appeared to be the approach with the greatest potential, combining the cell growth and the effective production of these protein O-rings.

In summary, assessment between standard and optimized experiments demonstrated a significant improvement in the production of TRAP O-rings in *E. coli* BL21(DE3). This successful optimization generated valuable data that will serve as a basis for future studies, including process scaling. In comparison with the production of MjsHSP nanocages, lower production efficiency values

were obtained for the optimized production process of TRAP-O-rings, consequently additional research could be performed evaluating each parameter more specifically, varying a detailed set of conditions, in order to have an even better improvement.

III.4. Conclusions

NVPNs have emerged as innovative protein nanocages with significant potential across bioengineering, biotechnology, and biomedicine, showcasing remarkable capabilities and versatility in drug delivery (114,126), vaccine development (119,129), bioimaging (134,135), biomineralization (93), and biocatalysis (75,144). Despite significant progress in understanding NVPNs applications, research on bioprocess development for natural and artificial protein nanocages remains limited (186). To advance NVPNs towards commercial viability, there is a pressing need for more efficient manufacturing processes. The critical stages of NVPNs upstream processing, similar to other biological products, involve the development of producer host cells and the expression of the target protein (139,140). Typically, these procedures rely on small-scale laboratory protocols, with limited published data on large-scale production. To develop NVPNs manufacturing, innovative strategies are needed to achieve higher host cell densities and improved nanocage yields, ultimately enhancing volumetric productivity (14,38,138,141).

In this **Chapter III**, the main objective was to explore the upstream processing of the MjsHSP nanocages and the TRAP O-rings as precursors of the TRAP nanocages. The use of the strain Vmax™ X2 of *V. natriegens* as an alternative bacterial host for the expression of proteins was clearly one of the novelties of the work.

Considering the results obtained, it was possible to successfully establish the benchmark manufacturing method for the MjsHSP nanocages both in *E. coli* BL21(DE3) and in *V. natriegens* Vmax™ X2 as well as for the TRAP O-rings in this *E. coli* strain. Several challenges were encountered during the transformation of *V. natriegens* cells with the pTRAP plasmid, which made the expression of TRAP O-rings impossible. Furthermore, the production process was successfully optimized for both production in *E. coli* (MjsHSP nanocages and TRAP O-rings) and *V. natriegens* (MjsHSP nanocages). Growth medium composition, concentration of protein expression inducer, temperature of protein expression, time point of induction and aeration conditions were the five parameters tested and assessed. It was possible to conclude that the concentration of ITPG as inducer was the one that had the greatest influence on the growth profile of bacterial cells and principally on the expression of protein nanocages.

With the optimized process in *E. coli* BL21(DE3) cells (enhanced 2xYT culture medium, 0.1 mM of IPTG, 37°C during protein expression, time point of induction at $OD_{600\ nm} \approx 0.6$ and shake flask without baffles), a production yield of MjsHSP nanocages of 295.7 $\pm$ 0.0 mg L^{-1} of culture and a volumetric productivity of 53.8 $\pm$ 0.0 mg L^{-1} h^{-1} were obtained after 4 hours of protein nanocages expression, which corresponded, respectively, to an increase of 5.4 and 5.1 times in relation to the standard production process. In parallel, in *V. natriegens* cells, a production yield of this first model of NVPNs of 60.9 $\pm$ 13.7 mg L^{-1} of culture and a volumetric productivity of 10.4 $\pm$ 2.3 mg L^{-1} h^{-1} was obtained with the optimized procedure after 4 hours of protein nanocages expression (enhanced 2xYT culture medium, 0.1 mM of IPTG, 37°C during protein expression, time point of induction at $OD_{600\ nm} \approx 0.6$ and shake flask with baffles). Although these values are lower than those of *E. coli* in direct comparison, an increase was observed in relation to the standard process (2.2 and 1.8 times, respectively), which showed the

potential for the production of NVPNs in this specific strain of *V. natriegens*. Regarding the optimized TRAP O-rings production process in *E. coli* BL21(DE3) (enhanced 2xYT culture medium, 0.1 mM of IPTG, 37°C during protein expression, time point of induction at $OD_{600\ nm} \approx 0.6$ and shake flask without baffles), a production yield of 249.1 ± 3.7 mg L^{-1} of culture and a volumetric productivity of 46.7 ± 0.7 mg L^{-1} h^{-1} were obtained after 4 hours of protein expression, corresponding, respectively, at a slight increase of 1.4 and 1.3 times compared to the reference production process.

The work developed in this **Chapter III** allowed to verify that *V. natriegens* cannot be seen as a direct competitor to *E. coli* but rather as a complementary host organism that has the potential to be the starting point for additional studies regarding protein expression. In this sense, in terms of future perspectives, more research is needed for the genetic improvement of *V. natriegens* strains, aiming to optimize this bacterial species for reliable and efficient protein nanocages expression. This optimization is crucial, as *V. natriegens* has significant potential as a microbial chassis for biotechnological applications, due to its rapid growth and high metabolic rate. Furthermore, it is essential to optimize plasmid components, such as promoters, origins of replication and selection markers, to better adapt to the specific physiological and metabolic characteristics of *V. natriegens*, aiming to increase the efficiency and stability of protein expression in this organism. It is essential to perform in-depth studies on transformation methods in *V. natriegens*, regardless of the technique used, as the challenges observed experimentally diverge significantly from the information available in the literature, especially with regard to transformation efficiency, suggesting the need for more specific and optimized protocols for this bacterial species. To optimize the expression of TRAP O-rings, it would be valuable to explore the optimization process in greater detail by implementing a systematic and comprehensive approach as performed for MjsHSP nanocages. This strategy would allow for more precise identification of the key parameters that have the greatest influence on the protein production, thus enabling fine-tuning of expression conditions and, consequently, a significant increase in the production yield of the final product. Considering the advances made in optimizing the expression of NVPNs on a small scale, the next step would be to scale up to pilot scale bioreactors. This transition would allow for more precise control of cultivation and expression conditions, providing a more robust and comprehensive evaluation of the process. This approach would not only validate results obtained in the laboratory but would also provide valuable insights into the challenges and opportunities associated with larger scale production.

III.5. Supplementary Material

Supplementary Table III.S1 – Nucleotide sequences of the four plasmids (pUC57-MjsHSP, pDHFR-His, pMjsHSP and pTRAP) used throughout this work.

Plasmid	Nucleotide sequence (5' → 3')
pUC57-MjsHSP	TCGCGCGTTTCGGTGATGACGGTGAAAACCTCTGACACATGCAGCTCCCGGAGAC GGTCACAGCTTGTCTGTAAGCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCGC GTCAGCGGGTGTTGGCGGGTGTCGGGGCTGGCTTAACTATGCGGCATCAGAGCA GATTGTACTGAGAGTGCACCATATGCGGTGTGAAATACCGCACAGATGCGTAAGG AGAAAATACCGCATCAGGCGCCATTCGCCATTCAGGCTGCGCAACTGTTGGGAAG GGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGGGGGATGTG CTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTGTAA AACGACGGCCAGAGAATTCGAGCTCGGTACCTCGCGAATACATCTAGATCATATGA TGTTTGGCCGTGATCCGTTCGACTCCTTGTTTGAACGCATGTTTAAAGAATTCTTCG CGACCCCGATGACGGGCACCACTATGATTCAGTCCAGTACCGGTATCCAGATCAG TGGCAAAGGCTTTATGCCGATTTCGATCATTGAAGGCGATCAGCACATCAAAGTTA TTGCCTGGCTGCCGGGCGTCAATAAAGAAGATATCATTCTGAACGCCGTGGGCGA CACCCTGGAGATTCGTGCCAAGCGCAGCCCACTCATGATCACCGAATCGGAACGT ATTATCTATAGCGAAATTCCGGAAGAAGAAGAAATTTATCGCACGATCAAACTGCC GGCAACGGTCAAGGAAGAGAATGCCTGCGCGAAATTCGAAAATGGTGTTCTTTCG GTGATTCTGCCAAAGGCAGAAAGCTCCATCAAAAAAGGCATTAACATTGAATAACT CGAGATCGGATCCCGGGCCCGTCGACTGCAGAGGCCTGCATGCAAGCTTGGTGT AATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACA ACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTA ACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGT GCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTG GGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCG GCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGG GATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTA AAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCA CAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATAC CAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGC TTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGC TCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTG TGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCT TGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAAC AGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGC CTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCA GTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTG GTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCT CAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTC ACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTT AAATTAAAAATGAAGTTTTAAATCAAGCCCAATCTGAATAATGTTACAACCAATTAAC CAATTCTGATTAGAAAAACTCATCGAGCATCAAATGAAACTGCAATTTATTCATATC AGGATTATCAATACCATATTTTTGAAAAAGCCGTTTCTGTAATGAAGGAGAAAACTC ACCGAGGCAGTTCCATAGGATGGCAAGATCCTGGTATCGGTCTGCGATTCCGACT CGTCCAACATCAATACAACCTATTAATTTCCCCTCGTCAAAAATAAGGTTATCAAGT GAGAAATCACCATGAGTGACGACTGAATCCGGTGAGAATGGCAAAAGTTTATGCAT TTCTTTCCAGACTTGTTCAACAGGCCAGCCATTACGCTCGTCATCAAAATCACTCG CATCAACCAAACCGTTATTCATTCGTGATTGCGCCTGAGCGAGACGAAATACGCGA TCGCTGTTAAAAGGACAATTACAAACAGGAATCGAATGCAACCGGCGCAGGAACA CTGCCAGCGCATCAACAATATTTTCACCTGAATCAGGATATTCTTCTAATACCTGGA ATGCTGTTTTTCCGGGGATCGCAGTGGTGAGTAACCATGCATCATCAGGAGTACG GATAAAATGCTTGATGGTCGGAAGAGGCATAAATTCCGTCAGCCAGTTTAGTCT…

Supplementary Table III.S1 – (*continued*).

Plasmid	Nucleotide sequence (5' → 3')
pUC57-MjsHSP	…GACCATCTCATCTGTAACATCATTGGCAACGCTACCTTTGCCATGTTTCAGAAAC AACTCTGGCGCATCGGGCTTCCCATACAAGCGATAGATTGTCGCACCTGATTGCC CGACATTATCGCGAGCCCATTTATACCCATATAAATCAGCATCCATGTTGGAATTTA ATCGCGGCCTCGACGTTTCCCGTTGAATATGGCTCATAACACCCCTTGTATTACTG TTTATGTAAGCAGACAGTTTTATTGTTCATGATGATATATTTTTATCTTGTGCAATGT AACATCAGAGATTTTGAGACACGGGCCAGAGCTGCA
pDHFR-His	GCTAGTGGTGCTAGCCCCGCGAAATTAATACGACTCACTATAGGGTCTAGAAATAA TTTTGTTTAACTTTAAGAAGGAGATATACATATGCACCATCACCACCATCACATCAG TCTGATTGCGGCGTTAGCGGTAGATCGCGTTATCGGCATGGAAAACGCCATGCCG TGGAACCTGCCTGCCGATCTCGCCTGGTTTAAACGCAACACCTTAAATAAACCCGT GATTATGGGCCGCCATACCTGGGAATCAATCGGTCGTCCGTTGCCAGGACGCAAA AATATTATCCTCAGCAGTCAACCGGGTACGGACGATCGCGTAACGTGGGTGAAGT CGGTGGATGAAGCCATCGCGGCGTGTGGTGACGTACCAGAAATCATGGTGATTGG CGGCGGTCGCGTTTATGAACAGTTCTTGCCAAAAGCGCAAAAACTGTATCTGACGC ATATCGACGCAGAAGTGGAAGGCGACACCCATTTCCCGGATTACGAGCCGGATGA CTGGGAATCGGTATTCAGCGAATTCCACGATGCTGATGCGCAGAACTCTCACAGC TATTGCTTTGAGATTCTGGAGCGGCGGTAATGAGGATCCCGGGAATTCTCGAGTAA GGTTAACCTGCAGGAGGCCTTTAATTAAGGTGGTGCGGCCGCGCTAGCGGTCCC GGGGGATCGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCTGAGTTGGCTGCTG CCACCGCTGAGCAATAACTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAG GGGTTTTTTGCTGAAAGGAGGAACTATATCCGGAAGCTTGGCACTGGCCGACCGG GGTCGAGCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAG CTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAA GAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTG CTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCT CAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCC TGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTG TCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTA TCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCC GTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGC TAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGC GAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTAC ACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAA AAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTT TTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTG ATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACAGATCCGGGATTT TGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAA GTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTT AATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCT GACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAG TGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAA ACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCT CCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAAT AGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTT TGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCC CCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGT AAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTAC TGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCAT TCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGA TAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTC GGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCA CTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTG…

Supplementary Table III.S1 – (*continued*).

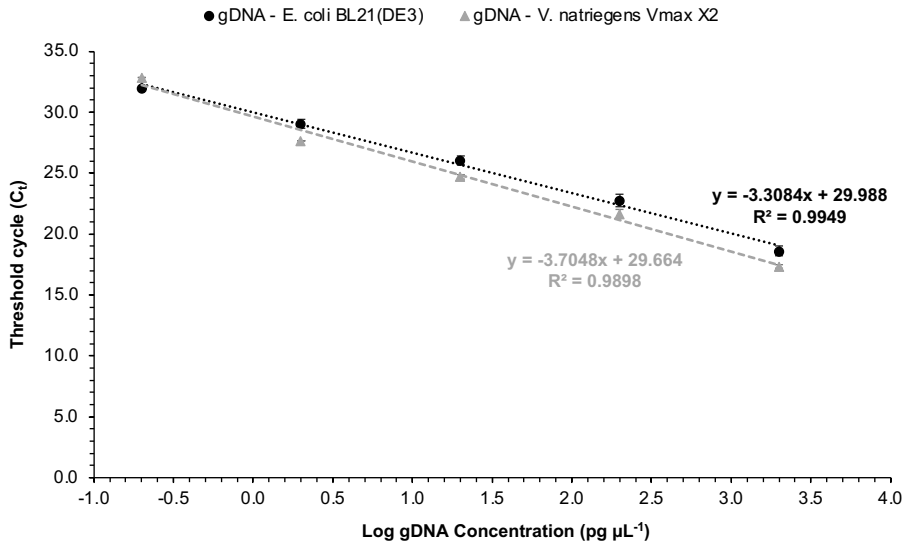
Plasmid	Nucleotide sequence (5' → 3')
pDHFR-His	…GGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACA CGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGG GTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAG GGGTTCCGCGCACATTTCCCCGAAAAGT
pMjsHSP	GCTAGTGGTGCTAGCCCCGCGAAATTAATACGACTCACTATAGGGTCTAGAAATAA TTTTGTTTAACTTTAAGAAGGAGATATACATATGATGTTTGGCCGTGATCCGTTCGA CTCCTTGTTTGAACGCATGTTTAAAGAATTCTTCGCGACCCCGATGACGGGCACCA CTATGATTCAGTCCAGTACCGGTATCCAGATCAGTGGCAAAGGCTTTATGCCGATT TCGATCATTGAAGGCGATCAGCACATCAAAGTTATTGCCTGGCTGCCGGGCGTCA ATAAAGAAGATATCATTCTGAACGCCGTGGGCGACACCCTGGAGATTCGTGCCAA GCGCAGCCCACTCATGATCACCGAATCGGAACGTATTATCTATAGCGAAATTCCGG AAGAAGAAGAAATTTATCGCACGATCAAACTGCCGGCAACGGTCAAGGAAGAGAA TGCCTGCGCGAAATTCGAAAATGGTGTTCTTTCGGTGATTCTGCCAAAGGCAGAAA GCTCCATCAAAAAAGGCATTAACATTGAATAACTCGAGTAAGGTTAACCTGCAGGA GGCCTTTAATTAAGGTGGTGCGGCCGCGCTAGCGGTCCCGGGGGATCGATCCGG CTGCTAACAAAGCCCGAAAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAATA ACTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGCTGAAAG GAGGAACTATATCCGGAAGCTTGGCACTGGCCGACCGGGGTCGAGCACTGACTC GCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGT AATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAA GGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATA GGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGC GAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGT GCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTT CGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTA GGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCG CTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGCTAAGACACGACTTAT CGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCG GTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTA TTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTC TTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGC AGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGG TCTGACGCTCAGTGGAACGAAAACTCACAGATCCGGGATTTTGGTCATGAGATTAT CAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTA AAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCAC CTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGT AGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACC GCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGA AGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTA ATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTT GTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATT CAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAA AAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGT GTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGT AAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTAT GCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACAT AGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTC AAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACT GATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGG CAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACT CTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGAT…

Supplementary Table III.S1 – (*continued*).

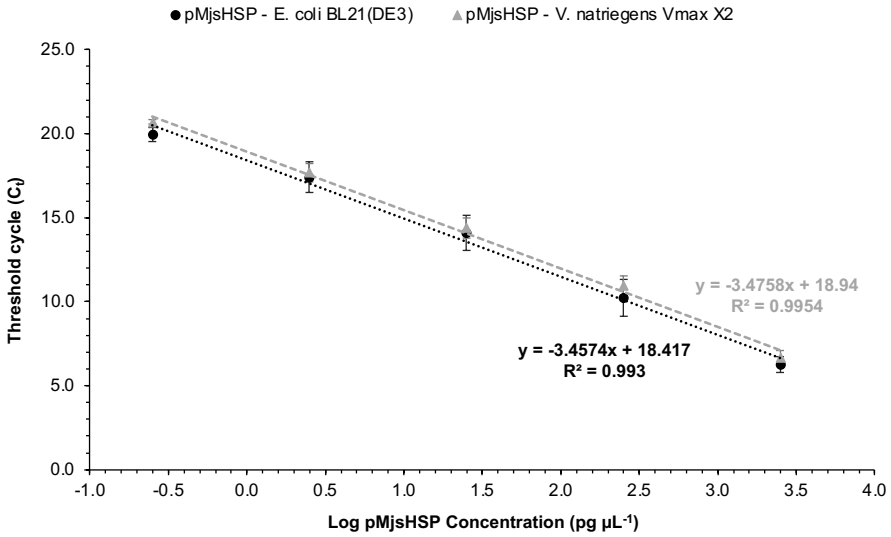
Plasmid	Nucleotide sequence (5' → 3')
pMjsHSP	…ACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCC CCGAAAAGT
pTRAP	TGGCGAATGGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGT TACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCT TTCTTCCCTTCCTTTCTCGCCACGTTCGCCGGCTTTCCCCGTCAAGCTCTAAATCG GGGGCTCCCTTTAGGGTTCCGATTTAGTGCTTTACGGCACCTCGACCCCAAAAAAC TTGATTAGGGTGATGGTTCACGTAGTGGGCCATCGCCCTGATAGACGGTTTTTCGC CCTTTGACGTTGGAGTCCACGTTCTTTAATAGTGGACTCTTGTTCCAAACTGGAAC AACACTCAACCCTATCTCGGTCTATTCTTTTGATTTATAAGGGATTTTGCCGATTTC GGCCTATTGGTTAAAAAATGAGCTGATTTAACAAAAATTTAACGCGAATTTTAACAA AATATTAACGTTTACAATTTCAGGTGGCACTTTTCGGGGAAATGTGCGCGGAACCC CTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAGACAATAACC CTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGAGTATTCAACATTTCCG TGTCGCCCTTATTCCCTTTTTTGCGGCATTTTGCCTTCCTGTTTTTGCTCACCCAGA AACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGTGCACGAGTGGGTTAC ATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAAC GTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTA TTGACGCCGGGCAAGAGCAACTCGGTCGCCGCATACACTATTCTCAGAATGACTT GGTTGAGTACTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGA GAATTATGCAGTGCTGCCATAACCATGAGTGATAACACTGCGGCCAACTTACTTCT GACAACGATCGGAGGACCGAAGGAGCTAACCGCTTTTTTGCACAACATGGGGGAT CATGTAACTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACG ACGAGCGTGACACCACGATGCCTGCAGCAATGGCAACAACGTTGCGCAAACTATT AACTGGCGAACTACTTACTCTAGCTTCCCGGCAACAATTAATAGACTGGATGGAGG CGGATAAAGTTGCAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTAT TGCTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTGCAGCACTG GGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGG CAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAG CATTGGTAACTGTCAGACCAAGTTTACTCATATATACTTTAGATTGATTTAAAACTTC ATTTTTAATTTAAAAGGATCTAGGTGAAGATCCTTTTTGATAATCTCATGACCAAAAT CCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAG GATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAAAC CACCGCTACCAGCGGTGGTTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTTCCG AAGGTAACTGGCTTCAGCAGAGCGCAGATACCAAATACTGTCCTTCTAGTGTAGCC GTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGC TAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTT GGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGGG TTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTA CAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGG TATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGG GGAAACGCCTGGTATCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCG TCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAAC GCGGCCTTTTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTTCCT GCGTTATCCCCTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCTGATAC CGCTCGCCGCAGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGG AAGAGCGCCTGATGCGGTATTTTCTCCTTACGCATCTGTGCGGTATTTCACACCGC ATATATGGTGCACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGCCAGTAT ACACTCCGCTATCGCTACGTGACTGGGTCATGGCTGCGCCCCGACACCCGCCAAC ACCCGCTGACGCGCCCTGACGGGCTTGTCTGCTCCCGGCATCCGCTTACAGACAA GCTGTGACCGTCTCCGGGAGCTGCATGTGTCAGAGGTTTTCACCGTCATCACCGA AACGCGCGAGGCAGCTGCGGTAAAGCTCATCAGCGTGGTCGTGAAGCGATTCACA GATGTCTGCCTGTTCATCCGCGTCCAGCTCGTTGAGTTTCTCCAGAAGCGTTAAT…

Supplementary Table III.S1 – (*continued*).

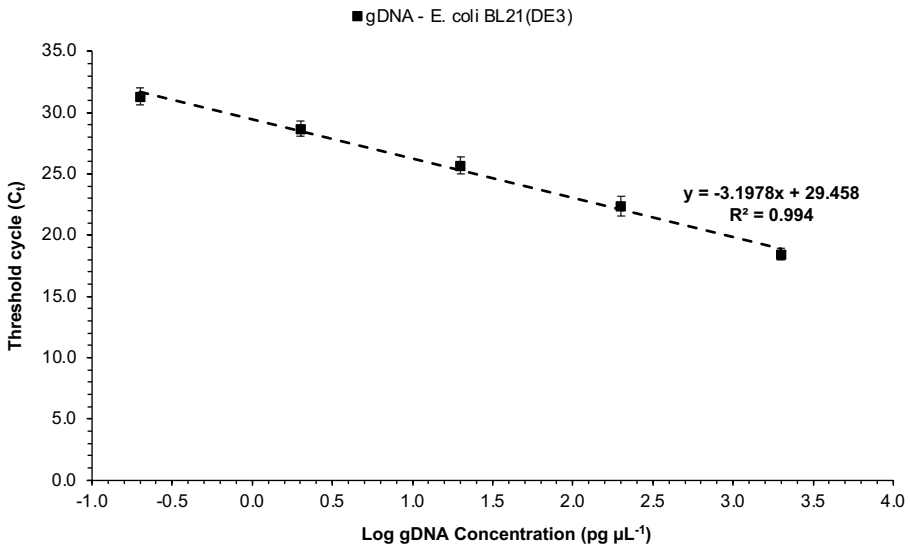
Plasmid	Nucleotide sequence (5' → 3')
pTRAP	...GTCTGGCTTCTGATAAAGCGGGCCATGTTAAGGGCGGTTTTTTCCTGTTTGGTCA CTGATGCCTCCGTGTAAGGGGGATTTCTGTTCATGGGGGTAATGATACCGATGAAA CGAGAGAGGATGCTCACGATACGGGTTACTGATGATGAACATGCCCGGTTACTGG AACGTTGTGAGGGTAAACAACTGGCGGTATGGATGCGGCGGGACCAGAGAAAAAT CACTCAGGGTCAATGCCAGCGCTTCGTTAATACAGATGTAGGTGTTCCACAGGGTA GCCAGCAGCATCCTGCGATGCAGATCCGGAACATAATGGTGCAGGGCGCTGACTT CCGCGTTTCCAGACTTTACGAAACACGGAAACCGAAGACCATTCATGTTGTTGCTC AGGTCGCAGACGTTTTGCAGCAGCAGTCGCTTCACGTTCGCTCGCGTATCGGTGA TTCATTCTGCTAACCAGTAAGGCAACCCCGCCAGCCTAGCCGGGTCCTCAACGAC AGGAGCACGATCATGCGCACCCGTGGGGCCGCCATGCCGGCGATAATGGCCTGC TTCTCGCCGAAACGTTTGGTGGCGGGACCAGTGACGAAGGCTTGAGCGAGGGCG TGCAAGATTCCGAATACCGCAAGCGACAGGCCGATCATCGTCGCGCTCCAGCGAA AGCGGTCCTCGCCGAAAATGACCCAGAGCGCTGCCGGCACCTGTCCTACGAGTT GCATGATAAAGAAGACAGTCATAAGTGCGGCGACGATAGTCATGCCCCGCGCCCA CCGGAAGGAGCTGACTGGGTTGAAGGCTCTCAAGGGCATCGGTCGAGATCCCGG TGCCTAATGAGTGAGCTAACTTACATTAATTGCGTTGCGCTCACTGCCCGCTTTCC AGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGA GAGGCGGTTTGCGTATTGGGCGCCAGGGTGGTTTTTCTTTTCACCAGTGAGACGG GCAACAGCTGATTGCCCTTCACCGCCTGGCCCTGAGAGAGTTGCAGCAAGCGGTC CACGCTGGTTTGCCCCAGCAGGCGAAAATCCTGTTTGATGGTGGTTAACGGCGGG ATATAACATGAGCTGTCTTCGGTATCGTCGTATCCCACTACCGAGATATCCGCACC AACGCGCAGCCCGGACTCGGTAATGGCGCGCATTGCGCCCAGCGCCATCTGATC GTTGGCAACCAGCATCGCAGTGGGAACGATGCCCTCATTCAGCATTTGCATGGTTT GTTGAAAACCGGACATGGCACTCCAGTCGCCTTCCCGTTCCGCTATCGGCTGAAT TTGATTGCGAGTGAGATATTTATGCCAGCCAGCCAGACGCAGACGCGCCGAGACA GAACTTAATGGGCCCGCTAACAGCGCGATTTGCTGGTGACCCAATGCGACCAGAT GCTCCACGCCCAGTCGCGTACCGTCTTCATGGGAGAAAATAATACTGTTGATGGG TGTCTGGTCAGAGACATCAAGAAATAACGCCGGAACATTAGTGCAGGCAGCTTCC ACAGCAATGGCATCCTGGTCATCCAGCGGATAGTTAATGATCAGCCCACTGACGC GTTGCGCGAGAAGATTGTGCACCGCCGCTTTACAGGCTTCGACGCCGCTTCGTTC TACCATCGACACCACCACGCTGGCACCCAGTTGATCGGCGCGAGATTTAATCGCC GCGACAATTTGCGACGGCGCGTGCAGGGCCAGACTGGAGGTGGCAACGCCAATC AGCAACGACTGTTTGCCCGCCAGTTGTTGTGCCACGCGGTTGGGAATGTAATTCA GCTCCGCCATCGCCGCTTCCACTTTTTCCCGCGTTTTCGCAGAAACGTGGCTGGC CTGGTTCACCACGCGGGAAACGGTCTGATAAGAGACACCGGCATACTCTGCGACA TCGTATAACGTTACTGGTTTCACATTCACCACCCTGAATTGACTCTCTTCCGGGCG CTATCATGCCATACCGCGAAAGGTTTTGCGCCATTCGATGGTGTCCGGGATCTCG ACGCTCTCCCTTATGCGACTCCTGCATTAGGAAGCAGCCCAGTAGTAGGTTGAGG CCGTTGAGCACCGCCGCCGCAAGGAATGGTGCATGCAAGGAGATGGCGCCCAAC AGTCCCCCGGCCACGGGGCCTGCCACCATACCCACGCCGAAACAAGCGCTCATG AGCCCGAAGTGGCGAGCCCGATCTTCCCCATCGGTGATGTCGGCGATATAGGCG CCAGCAACCGCACCTGTGGCGCCGGTGATGCCGGCCACGATGCGTCCGGCGTAG AGGATCGAGATCTCGATCCCGCGAAATTAATACGACTCACTATAGGGGAATTGTGA GCGGATAACAATTCCCCTCTAGAAATAATTTTGTTTAACTTTAAGAAGGAGATATAC ATATGTACACCAACTCTGACTTCGTTGTTATCAAAGCGCTGGAAGACGGTGTTAAC GTTATCGGTCTGACCCGTGGCGCCGACACCCGTTTCCACCACTCTGAATGCCTGG ACAAAGGTGAAGTTCTGATCGCGCAGTTCACCGAACACACTAGTGCGATCAAAGTT CGTGGTAAAGCGTACATCCAGACCCGTCACGGTGTTATCGAATCTGAAGGTAAAAA ATAACTCGAGCACCACCACCACCACCACTGAGATCCGGCTGCTAACAAAGCCCGA AAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACCCCTTG GGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGCTGAAAGGAGGAACTATATCCGG AT



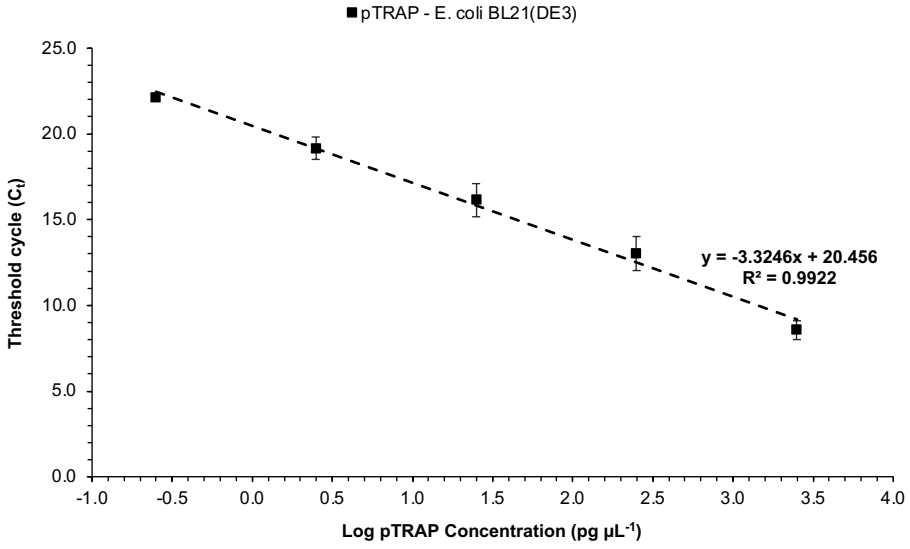
Supplementary Figure III.S1 – Standard curves for gDNA from *E. coli* BL21(DE3) (●) and *V. natriegens* Vmax™ X2 (▲) used in the determination of plasmid copy number by qPCR. Reactions were performed in triplicate for each bacterial strain, with 5 log serial dilutions of the respective gDNA (10000, 1000, 100, 10 and 1 pg). Note that the linear regression trendline is represented, as well as the corresponding equation and the coefficient of determination value (R^2).



Supplementary Figure III.S2 – Standard curves for pDNA (pMjsHSP) from *E. coli* BL21(DE3) (●) and *V. natriegens* Vmax™ X2 (▲) used in the determination of plasmid copy number by qPCR. Reactions were performed in triplicate for each bacterial strain, with 5 log serial dilutions of pMjsHSP (10000, 1000, 100, 10 and 1 pg) spiked with the respective *E. coli* or *V. natriegens* cells. Note that the linear regression trendline is represented, as well as the corresponding equation and the coefficient of determination value (R^2).



Supplementary Figure III.S3 – Standard curve for gDNA from *E. coli* BL21(DE3) (■) used in the determination of plasmid copy number by qPCR. Reactions were performed in triplicate, with 5 log serial dilutions of the respective gDNA (10000, 1000, 100, 10 and 1 pg). Note that the linear regression trendline is represented, as well as the corresponding equation and the coefficient of determination value (R^2).



Supplementary Figure III.S4 – Standard curve for pDNA (pTRAP) from *E. coli* BL21(DE3) (■) used in the determination of plasmid copy number by qPCR. Reactions were performed in triplicate, with 5 log serial dilutions of pTRAP (10000, 1000, 100, 10 and 1 pg) spiked with the respective *E. coli* cells. Note that the linear regression trendline is represented, as well as the corresponding equation and the coefficient of determination value (R^2).

Chapter IV

Downstream Processing of MjsHSP and TRAP Nanocages

Part of this chapter together with part of chapter V are being considered for publication as two original research articles:

João J, Rosa SS, Robalo T, Rodrigues MS, Paulo PMR, Azevedo AM, Prazeres DMF. Downstream processing of an archaeal small heat shock protein nanocage: Development of a chromatography-based purification approach. 2025.

João J, Paulo PMR, Azevedo AM, Prazeres DMF. Exploring a chromatographic strategy for the purification of TRAP O-rings. 2025.

Abstract

In the past few years, NVPNs have gained recognition as versatile tools across various scientific disciplines, showcasing notable benefits in applications like drug delivery, vaccines development, bioimaging, biomineralization, and biocatalysis. While significant advancements have been made in understanding the properties of NVPNs for these purposes, there remains a pressing need to enhance manufacturing efficiency, not only in the upstream stage but also in the downstream processing. The downstream processing is particularly relevant, as it ensures that the final product meets specific quality standards tailored to its intended purpose. This bioprocessing stage involves extracting and isolating NVPNs from the upstream culture broth, with purification standing out as the critical step. Research on both initial isolation and subsequent purification of protein nanocages is fragmented and lacks comprehensive detail, which complicates the process of comparing different methodologies effectively.

This **Chapter IV** is focused on the downstream processing of the MjsHSP and TRAP nanocages, including the TRAP O-rings as precursors of the latter, after their production process in *E. coli* and *V. natriegens* cells. Different strategies for purifying these NVPNs from contaminants such as host proteins and DNA were explored, both in primary isolation and recovery with thermal and enzymatic treatment and in purification based on chromatographic approaches. Various parameters and conditions, such as the type of chromatographic support, the effectiveness of bead and pore size, the composition of binding and elution buffers, in order to obtain a highly purified product. Due to their alternative characteristics, the TRAP nanocages require an additional purification step after the *in vitro* assembly, which was not possible to achieve due to the several challenges encountered in this assembly process of the TRAP O-rings.

The optimized procedure for the primary isolation and purification of MjsHSP nanocages from *E. coli* consisted of enzymatic pre-treatment of the clarified cell lysate with DNase followed by a single-step AEX chromatography, which allowed obtaining a protein nanocages recovery yield of 151.9 ± 1.8 μg mL^{-1}, a recovery efficiency of 63.3 ± 3.3% and a host protein impurities removal of 88.0 ± 1.0%. Regarding the downstream processing benchmark, these results correspond to a 7.2-fold increase in terms of recovery efficiency and a slight 1.1-fold decrease of in terms of impurities removal. In parallel, for the MjsHSP nanocages from *V. natriegens*, the optimized process comprised a similar single-step AEX chromatography, having been obtained protein nanocages recovery yield of 3.4 ± 0.4 μg mL^{-1}, a recovery efficiency of 9.5 ± 1.1% and a host protein impurities removal of 96.8 ± 1.3%. Regarding the optimized TRAP O-rings downstream process (single-step CEX chromatography), a protein O-rings recovery yield of 320.2 ± 5.2 μg mL^{-1}, a recovery efficiency of 69.6 ± 0.8%, a host protein impurities removal of 82.0 ± 2.3% and a host dsDNA impurities removal of 85.0 ± 5.2% were achieved. This optimized strategy improved recovery efficiency by 4-fold compared to the reference process, while impurities removal remained similar.

Keywords: chromatography, downstream processing, primary isolation, protein nanocages, purification

IV.1. Introduction

Recently, NVPNs have become a focal point of research interest, especially in bioengineering, biotechnology, and biomedicine. These protein nanocages have shown promising results in various applications, ranging from targeted drug delivery (114,126,127) and innovative vaccine platforms (119,129,132) to advanced bioimaging techniques (134,135), controlled biomineralization processes (93), and efficient biocatalytic systems (75,144). This topic was explored in more detail in **Chapter I**.

Despite significant progress in proof-of-concept studies focusing on the characteristics of NVPNs, alternative functionalization strategies, and specific applications, there remains a notable gap in research focusing on the bioprocess development of both natural and engineered protein nanocages (186). Numerous challenges have emerged, highlighting the need for increased attention on optimizing the manufacturing processes of these NVPNs.

The downstream processing is particularly relevant, as it ensures that the final product meets specific quality standards tailored to its intended application. This bioprocessing stage involves extracting and isolating NVPNs from the upstream culture broth, with purification standing out as the critical step (138,141,146). To keep costs down and improve efficiency, the ideal downstream process uses just a few high-yield steps. These steps are chosen based on the unique properties of the protein nanocages and the contaminants (such as gDNA, RNA, host proteins, and lipopolysaccharides) that need to be removed. The available research on both initial isolation and subsequent purification of protein nanocages is fragmented and lacks comprehensive detail, making it challenging to establish significant comparisons between different methodologies.

The primary isolation and recovery of NVPNs involve several key steps. Initially, producer cells are harvested from the culture broth by centrifugation at moderate speeds, such as 5000×g (89,147,148,153). Next, intracellular nanocages are released through various cell lysis methods. At the laboratory scale, ultrasonication (23,24,29,40,101,125,136,153), French press (29,50,96), and Dounce homogenization (72,145) are common. For large-scale operations, high-pressure homogenizers and bead milling are more suitable (141). During cell lysis, lysozyme may be added to weaken bacterial cell walls (29,125,136), nucleases to degrade nucleic acids (29,125,136), and protease inhibitors (72,119,145,153). After lysis, a clarification step is performed using centrifugation or filtration to remove cell debris and particulate matter. In some cases, clarified lysates are subjected to a heat treatment, heating them to temperatures between 65°C and 90°C. This exploits the thermostability of the NVPNs, allowing host proteins to denature and precipitate without affecting the protein nanocages, as observed with encapsulin (30,101), ferritin (23,65,82,125,150,163), and sHSP (62). Additionally, enzymatic treatment of clarified lysates is common, using DNase or RNase to degrade host cell nucleic acid impurities (39,150).

The purification stage of NVPNs is crucial for removing host-derived impurities and obtaining a compliant bulk product. Column chromatography is the preferred method for obtaining highly pure nanocages, with two main groups of purification methods: chromatographic and non-chromatographic techniques. The main chromatographic methods used for protein nanocages purification include ion exchange (IEX), hydrophobic interaction (HIC), affinity chromatography (AC) and size exclusion (SEC).

In IEX, HIC, and AC, nanocages are retained by interacting with the stationary phase, while impurities flow through or elute under specific buffer conditions (138,139). SEC is particularly attractive due to the generally larger size of nanocages compared to most host-derived impurities.

IEX primarily uses anion-exchange resins due to the anionic nature of NVPNs, with common resins including strong anion-exchangers like Uno-Q, Q Sepharose, HiTrap Q and HiPrep (27,39,40,49,61,89,90,97,131,135,136,151,165), and weak anion-exchangers like DEAE Sepharose (110). Sodium chloride gradients are typically used for elution. While effective in removing host impurities, negatively charged impurities can interfere with the process. HIC has been used for ferritin nanocages with Phenyl Sepharose resin (163,164), but its main drawback is the requirement for high salt concentrations to promote binding.

SEC is used as a polishing step to separate nanocages from similarly sized impurities like aggregates and misassembled variants. Common resins include Superose 6 (49,97,131,136,165), Superdex 200 (40,61,110), TSKgel G3000SW (39,135,151) and Sephacryl S-200 (89,90). SEC also removes residual impurities and facilitates buffer exchange, though achieving baseline peak separation can be challenging.

AC is commonly used to purify engineered variants of natural and artificial NVPNs. This method requires incorporating an affinity tag into the monomers' primary sequence, allowing assembled nanocages to be captured by an affinity resin with the appropriate ligand. Typical examples include poly-histidine tags with nickel ion chelate affinity resins (e.g., Ni-NTA, Ni Sepharose 6) (24,76,123,126,147,148,168) and maltose-binding protein domain fused to monomers, purified using maltose affinity resin (149). AC offers high selectivity and specificity, resulting in high purity levels. However, it may require a second purification step, such as SEC or AEX. Potential drawbacks include the need to add affinity tags, which may affect self-assembly or alter original protein nanocages properties.

Combinations of chromatographic steps are commonly used for purifying NVPNs. While AEX followed by SEC is widely used, other combinations have been reported. These include SEC-AEX for mouse ferritin nanocages (150) and E2 nanocages (96), as well as a three-step AEX-HIC-SEC purification for sHSP nanocages (88). Some NVPNs have been purified using a single AEX (50,125) or SEC (23,29,30,92,123,134) step. A unique approach is used for artificial TRAP nanocages, where individual subunits are purified before assembly. This process involves heat treatment of clarified E. coli lysate, followed by AEX and SEC steps (111–114,116).

Pre-chromatography processing techniques for protein nanocages include ammonium sulfate precipitation to remove nucleic acid impurities before AEX or SEC (89,90,96,110,118,150,169). Differential ultracentrifugation, sometimes with sucrose gradients, is also used (30,101). However, ultracentrifugation has drawbacks such as high speeds, long processing times, and poor scalability.

Chromatographic process yields for NVPNs vary widely depending on the purification method and the type of nanocage. Single-step methods like AEX or SEC can yield 25 to 56 mg L^{-1} for ferritin nanocages (125), while Ni-affinity chromatography yields 10 to 15 mg L^{-1} with 90-96% purity (147,148). Multi-step methods, such as SEC plus AEX (150) or ammonium sulfate precipitation plus SEC (101,169), typically yield 7 to 50 mg L^{-1} for various nanocages. Special cases include artificial Tet8-M

nanocages purified by AC, yielding 400 mg L^{-1} (149), and IMAC for vault nanocages with 90.4% recovery (153). These diverse yields highlight the variability in purification efficiency across different NVPN types and methods.

Non-chromatographic methods for purifying NVPNs have been developed as alternatives to traditional chromatographic techniques. These include heat treatment, filtration, and ultrafiltration, which efficiently remove impurities (152). For endotoxin removal, techniques like using Endotrap HD resin with Triton X-114 have been effective in reducing endotoxin levels below regulatory limits (161). Additionally, methods such as discontinuous density gradient ultracentrifugation are used for vault nanocages (72,142,145). These approaches offer potential cost reductions and scalability advantages for NVPNs purification.

In this **Chapter IV**, the downstream processing of the two model NVPNs, MjsHSP and TRAP nanocages, was explored. For MjsHSP nanocages produced in *E. coli*, it was evaluated the impact of introducing an enzymatic treatment step using nucleases (DNase and RNase) and/or a thermal treatment with a significant increase in temperature during the primary isolation and recovery phase. The purification process mainly used chromatography, starting from a commonly reported two-step combination: AEX followed by SEC. Several parameters and conditions, such as the type of chromatographic support, the effectiveness of bead and pore size, the composition of binding and elution buffers, were tested in order to obtain a highly purified product. As an alternative to SEC, the potential of multimodal chromatography for the final purification step was analyzed. The optimized purification procedure developed for MjsHSP nanocages from *E. coli* production was then applied and adapted to MjsHSP nanocages produced in *V. natriegens*. The process performance was assessed, and required adjustments were made to accommodate any differences in the production host. The purification process for TRAP nanocages produced in *E. coli* was developed by evaluating additional prior steps and exploring alternative chromatographic methods. A pre-purification thermal treatment of the clarified lysate was assessed. The standard purification approach combined AEX with SEC. As an alternative strategy, CEX was investigated to exploit the different physico-chemical properties of the TRAP O-rings compared to the associated impurities, particularly host proteins and DNA. This approach aimed to enhance recovery yield and purity while potentially simplifying the purification workflow. Due to challenges in the *in vitro* assembly of TRAP O-rings, the purification of TRAP nanocages could not proceed as planned.

IV.2. Materials and Methods

IV.2.1. Materials

Bacto™ Yeast Extract was acquired from BD Biosciences (Franklin Lakes, NJ, USA). Acrylamide/bis-acrylamide solution (40%), Laemmli sample buffer (4x), native sample buffer (2x) and precision plus protein dual color standards were from Bio-Rad (Hercules, CA, USA). Capto™ Core 400, Capto™ Q ImpRes, Capto™ SP ImpRes and Superdex™ 200 resins, Tricorn™ empty chromatography columns (5/50 and 10/300), and HiTrap Capto™ S and HiTrap Q FF pre-packed columns (1 mL) were purchased from Cytiva (Marlborough, MA, USA). Magnesium chloride ($MgCl_2$) was obtained from Fagron Ibérica (Barcelona, Spain). Acetic acid glacial, ethanol solution (96% v/v), glucose, hydrochloric acid (HCl), isopropyl β-D-1-thiogalactopyranoside (IPTG), 2-(N-morpholino)ethanesulfonic acid (MES), phosphate buffered saline (PBS) solution (10x), sodium chloride (NaCl), sodium hydroxide (NaOH), sorbitol, tris(hydroxylmethyl)aminomethane (Tris) and Tween™ 20 were purchased from Fisher Scientific (Hampton, NH, USA). PhastGel™ Blue R (Coomassie R 350 stain) was acquired from GE Healthcare (Chicago, IL, USA). Potassium chloride (KCl) was from Merck (Darmstadt, Germany). Dithiothreitol (DTT), Luria-Bertani (LB) broth, NZYColour protein marker II, NZYSpeedy qPCR Green Master Mix (ROX) (2x), NZY Tris-Glycine precast gel (4-15%, 12 wells) and SDS-PAGE sample loading buffer (5x) were purchased from NZYtech (Lisbon, Portugal). Potassium phosphate dibasic (K_2HPO_4), potassium phosphate dibasic (KH_2PO_4), sodium phosphate dibasic (Na_2HPO_4), sodium phosphate monobasic (NaH_2PO_4) and tryptone were acquired from Panreac AppliChem (Barcelona, Spain). Wizard® Genomic DNA Purification Kit was from Promega (Madison, WI, USA). High Pure Plasmid Isolation Kit and RNase A were obtained from Roche (Basel, Switzerland). CIMac PrimaS® analytical monolith (0.1 mL, 2 µm) and CIMmultus® QA monolith (1 mL, 2 µm) were from Sartorius (Göttingen, Germany). Amicon Ultra-4 centrifugal filters (30 kDa and 100 kDa MWCO), ammonium persulfate (APS), ampicillin, ethylenediaminetetraacetic acid (EDTA), glycine, N,N,N′,N′-tetramethylethylenediamine (TEMED), sodium dodecyl sulfate (SDS) and urea were purchased from Sigma-Aldrich (St. Louis, MO, USA). Chloro[diphenyl(3-sulfonatophenyl)phosphine]gold(i) sodium salt hydrate (Au-TPPMS) was obtained from STREM Chemicals UK (Cambridge, United Kingdom). Invitrogen™ TURBO DNase™ (and respective TURBO™ DNase buffer (10x)), Pierce™ BCA Protein Assay Kit and POROS™ 50 HQ resin were acquired from Thermo Fisher Scientific (Waltham, MA, USA). TSKgel SuperSW3000 column was obtained from Tosoh Bioscience (Tokyo, Japan). All primers used were synthesized and obtained from STAB VIDA (Caparica, Portugal). The pure and ultrapure water used during all experiments was provided by a Milli-Q purification system (Millipore; Bedford, MA, USA).

IV.2.2. Production of MjsHSP nanocages

IV.2.2.1. *E. coli* cells

The expression of MjsHSP nanocages was performed principally using the experimental methodology defined as reference, whose procedure is described in detail in the section **III.2.10. Production of MjsHSP nanocages**. Alternatively, the production of these protein nanocages in *E. coli* BL21(DE3) cells was carried out according to the optimized conditions determined in **Chapter III** of this work (**III.3.5. Optimized production of MjsHSP nanocages in *E. coli* and *V. natriegens***), namely the use of enhanced 2xYT culture medium, an IPTG concentration of 0.1 mM, an induction temperature of 37°C, a time point in the middle of the exponential growth phase ($OD_{600\ nm}$ of approximately 0.5 to 0.6) and an aeration setup comprising a non-baffled shake flask. This respective experimental protocol is available in the section **III.2.12. Optimization of production of MjsHSP nanocages**.

IV.2.2.2. *V. natriegens* cells

Similar to what was applied for *E. coli*, the expression of MjsHSP nanocages in *V. natriegens* Vmax™ X2 cells was performed essentially using the experimental approach defined as benchmark. Note that the procedure is also described completely in the section **III.2.10. Production of MjsHSP nanocages**. Alternatively, the production of these protein nanocages was carried out according to the optimized conditions determined in **Chapter III** of this work (**III.3.5. Optimized production of MjsHSP nanocages in *E. coli* and *V. natriegens***), namely the use of enhanced 2xYT culture medium, an IPTG concentration of 0.1 mM, an induction temperature of 37°C, a time point in the middle of the exponential growth phase ($OD_{600\ nm}$ of approximately 0.5 to 0.6) and an aeration setup comprising a baffled shake flask. Again, this respective experimental protocol is presented in the section **III.2.12. Optimization of production of MjsHSP nanocages**.

IV.2.3. Production of TRAP O-rings

The expression of TRAP O-rings as precursors of TRAP nanocages in *E. coli* BL21(DE3) cells was performed mainly using the experimental method defined as reference, whose procedure is described in detail in the section **III.2.11. Production of TRAP O-rings**. Alternatively, the production of these protein nanocages was performed according to the optimized conditions obtained in **Chapter III** of this work (**III.3.10. Optimized production of TRAP-O rings in *E. coli***), namely the use of enhanced 2xYT culture medium, an IPTG concentration of 0.1 mM, an induction temperature of 37°C, a time point in the middle of the exponential growth phase ($OD_{600\ nm}$ of approximately 0.5 to 0.6) and an aeration setup comprising a non-baffled shake flask. This respective experimental protocol is available in the section **III.2.13. Optimization of production of TRAP O-rings**.

IV.2.4. Assembly of TRAP nanocages

The experimental procedure implemented was based on the work described by Malay *et al.* (111).

The assembly of TRAP nanocages was performed by mixing the previously purified TRAP O-rings (**IV.2.6. Primary isolation and purification of TRAP O-rings**) and Au-TPPMS in an aqueous solution. The specific conditions for this reaction involved combining equal molar quantities between 0.5 mM and 1 mM of the TRAP O-rings as protein nanocages monomers and the Au-TPPMS as a source of gold ions. It should be noted that both the samples containing the TRAP O-rings and the Au-TPPMS stock solution were prepared in a defined buffer (50 mM Tris-HCl, 150 mM NaCl, pH 7.9). This reaction mixture was subsequently incubated for a minimum of 5 days and a maximum of 12 days at room temperature with slight shaking. These assembled TRAP nanocages were stored at room temperature or at 4°C.

IV.2.5. Primary isolation and purification of MjsHSP nanocages

IV.2.5.1. Production in *E. coli* cells

IV.2.5.1.1. Treatment of cell lysate with nucleases (DNase and RNase)

After the clarification step following the production of MjsHSP nanocages in *E. coli* BL21(DE3) cells, the resulting cell lysate was subjected to an enzymatic treatment in order to reduce the amount of host nucleic acid impurities (gDNA, pDNA and RNA) before proceeding to the main purification steps of the protein nanocages.

With this aim, 1 mL of the clarified cell lysate under study containing the MjsHSP nanocages (concentration of approximately 2 mg of total protein per mL) was combined with 4 µL of TURBO™ DNase (2 U μL^{-1}; corresponding to 8 U) and 100 µL of the respective TURBO™ DNase buffer (10x) (Thermo Fisher Scientific; Waltham, MA, USA) or with 80 µL of RNase A (0.05 U μL^{-1}, corresponding to 4 U) (Roche; Basel, Switzerland). Each enzymatic reaction was incubated for 1 hour at 37°C. When applicable, inactivation of both nucleases was performed according to the manufacturer's instructions. In addition to this sequential DNase and RNase reaction, experiments were also performed with only each of these enzymes, in order to evaluate their influence on the reduction of nucleic acids impurities and consequently on the following steps based on chromatography.

The enzymatically treated samples were directly used in the next step of purification of the MjsHSP nanocages based on chromatography.

IV.2.5.1.2. Anion-exchange chromatography

AEX chromatography was performed using a ÄKTA Start system (Cytiva; Marlborough, MA, USA). equipped with a 2 mm path length UV cell with a fixed wavelength of 280 nm. Chromatography parameters were monitored and controlled using UNICORN Start software (Cytiva; Marlborough, MA, USA). The system was operated at a flow rate of 1 mL min^{-1}.

As standard condition, it was used the HiTrap Q FF pre-packed column (1 mL, particle size of 90 µm and pore size of about 50 to 60 nm) (Cytiva; Marlborough, MA, USA), a mobile phase composed of the low salt buffer A_1 (50 mM Tris-HCl, 100 mM NaCl, pH 8.0) and the high salt buffer B_1 (50 mM Tris-HCl, 1 M NaCl, pH 8.0), and a bind/elute mode strategy, with the elution being performed through a linear gradient from 0% to 100% of buffer B_1 corresponding to 20 column volumes (CV).

During the optimization process, two strong AEX resins with quaternary ammonium groups Capto™ Q ImpRes (Cytiva; Marlborough, MA, USA) and POROS™ 50 HQ (Thermo Fisher Scientific; Waltham, MA, USA) were tested, both packed in Tricorn 5/50 columns for a target volume of 1 mL (following the manufacturer's guidelines), and a CIMmultus® QA monolith (1 mL, channel size of 2 µm) (Sartorius; Göttingen, Germany). The particle size of Capto™ Q ImpRes and POROS™ 50 HQ is of 40 µm and 50 µm, respectively, while the pore size is approximately of 70 nm and 360 nm, respectively. The use of an alternative buffer solution with the main component being 50 mM sodium phosphate was also evaluated (also varying the pH between 8.0 and 6.0), resulting in a low salt buffer A_2 composed of 50 mM sodium phosphate, 250 mM NaCl, pH 6.0 and a high salt buffer B_2 consists of 50 mM sodium phosphate, 1 M NaCl, pH 6.0. In all test experiments, a linear gradient elution was initially considered (20 CV), allowing for an analysis of the behavior and results obtained, which were later adjusted to an improved step gradient.

For the optimized protocol, the column or monolith were equilibrated with 5 CV of buffer A_2 ($\kappa \approx 26.2$ mS cm^{-1}). Then, a 5 mL sample (comprising 1 mL of the enzymatically treated or non-enzymatically treated clarified lysate containing the MjsHSP nanocages together with 4 mL of buffer A_2) was injected into the column or monolith and all unbound material was washed with 8 CV of buffer A_2. Note that the approximate mass of total protein in the sample was 2 mg. It was implemented a two-step gradient for the elution of bound species, the first step with 8 CV of 23% buffer B_2 ($\kappa \approx 40.5$ mS cm^{-1}) and the second step with 8 CV of 100% buffer B_2 ($\kappa \approx 81.5$ mS cm^{-1}). The eluate was collected in fractions of 1 mL during the chromatographic run with a fraction collector. All fractions were stored at 4°C until further utilization and/or analysis.

After each chromatography run, both columns and monolith were washed with Milli-Q water until the conductivity was below 0.1 mS cm^{-1} and then stored in 20% v/v ethanol. When necessary, a cleaning-in-place (CIP) procedure was performed (30 CV of 0.5 to 1 M NaOH and/or 2 M NaCl at a flow rate of 0.5 to 1 mL min^{-1}). This maintenance protocol was considered in all experiments regardless of the type of chromatography under study.

IV.2.5.1.3. Concentration and diafiltration by centrifugal filtration

When required, the previous fractions containing the MjsHSP nanocages were pooled and concentrated using 100 kDa Amicon Ultra-4 mL centrifugal filters (Sigma-Aldrich; St. Louis, MO, USA), following the parameters recommended by the manufacturer (5000×g, for 5 to 15 minutes in a fixed-angle rotor at room temperature). Note that these centrifugal filters were passivated overnight at room temperature with 5% v/v Tween 20 and immediately prior to use they were washed with Milli-Q water and equilibrated with 50 mM sodium phosphate, 250 mM NaCl, pH 6.0. If applicable, diafiltration was

also performed with the passage of 3 to 4 times of PBS solution (1x) (137 mM NaCl, 2.7 mM KCl, 8 mM Na_2HPO_4, 2 mM KH_2PO_4, pH 7.4). All samples were stored at 4°C until further utilization and/or analysis.

IV.2.5.1.4. Size exclusion chromatography

SEC was performed using the ÄKTA Start system and the respective UNICORN Start software (Cytiva; Marlborough, MA, USA). The system was operated at a flow rate of 0.5 or 1 mL min^{-1}.

The selected size exclusion resin Superdex™ 200 (particle size of 13 µm) (Cytiva; Marlborough, MA, USA) was packed in a Tricorn 10/300 column for a target volume of 24 mL. The mobile phase consisted of PBS solution (1x) (137 mM NaCl, 2.7 mM KCl, 8 mM Na_2HPO_4, 2 mM KH_2PO_4, pH 7.4). The column was equilibrated with 2 CV of PBS solution (1x) ($\kappa \approx 14.2$ mS cm^{-1}). Then, a 200 µL sample containing the concentrated MjsHSP nanocages was injected into the column and this was washed with 2 CV of PBS solution (1x). The eluate was collected in fractions of 0.5 mL during the chromatographic run with a fraction collector. At the end, all the fractions were stored at 4°C until further utilization and/or analysis.

As previously mentioned, the column maintenance followed the manufacturer specifications when required.

IV.2.5.1.5. Multimodal chromatography

Multimodal chromatography was performed using the ÄKTA Start system and the respective UNICORN Start software (Cytiva; Marlborough, MA, USA). The system was operated at a flow rate of 1 mL min^{-1}.

The selected multimodal resin Capto™ Core 400 (size exclusion and binding chromatography, particle size of 90 µm and pore size of about 19 nm, average molecular weight cut-off of 400 kDa) (Cytiva; Marlborough, MA, USA) was packed in a Tricorn 5/50 column for a target volume of 1 mL. The mobile phase consisted of PBS solution (1x) (137 mM NaCl, 2.7 mM KCl, 8 mM Na_2HPO_4, 2 mM KH_2PO_4, pH 7.4). It was used a flowthrough mode strategy due to the intrinsic characteristics of this resin. The column was equilibrated with 5 CV of PBS solution (1x) ($\kappa \approx 14.2$ mS cm^{-1}). Next, a 2 to 3 mL sample containing the MjsHSP nanocages was injected into the column and this was washed with 6 CV of PBS solution (1x). The flowthrough was collected in fractions of 1 mL during the chromatographic run with a fraction collector. Then, the column was washed with a 1 M NaCl solution to remove material that had bound to the resin. The resulting 1 mL fractions were also collected. At the end, all the fractions were stored at 4°C until further utilization and/or analysis.

As previously mentioned, the column maintenance followed the manufacturer specifications when required.

IV.2.5.2. Production in *V. natriegens* cells

IV.2.5.2.1. Anion-exchange chromatography

The AEX experimental methodology for purification of MjsHSP nanocages produced in *V. natriegens* Vmax™ X2 was the same as that used for *E. coli* BL21(DE3) cells (sub-section **IV.2.5.1.2. Anion-exchange chromatography** in **IV.2.5.1. Production in *E. coli* cells**, section **IV.2.5. Primary isolation and purification of MjsHSP nanocages**), considering the optimized conditions with the greatest potential and implementing slight adjustments for a better fit to the samples of the clarified lysates from *V. natriegens*.

IV.2.5.2.2. Concentration and diafiltration by centrifugal filtration

The concentration and diafiltration by centrifugal filtration of samples containing MjsHSP nanocages previously produced in *V. natriegens* Vmax™ X2 was performed using the same experimental methodology that was established for samples containing this type of protein nanocages originating from *E. coli* BL21(DE3) (sub-section **IV.2.5.1.3. Concentration and diafiltration by centrifugal filtration** in **IV.2.5.1. Production in *E. coli* cells**, section **IV.2.5. Primary isolation and purification of MjsHSP nanocages**).

IV.2.5.2.3. Size exclusion chromatography

Similar to the above mentioned for the AEX stage, the SEC approach was also explored for MjsHSP nanocages from *V. natriegens* Vmax™ X2 based on an experimental strategy similar to that used with *E. coli* BL21(DE3) cells (sub-section **IV.2.5.1.4. Size exclusion chromatography** in **IV.2.5.1. Production in *E. coli* cells**, section **IV.2.5. Primary isolation and purification of MjsHSP nanocages**).

IV.2.6. Primary isolation and purification of TRAP O-rings

IV.2.6.1. Treatment of cell lysate with heating

After the clarification step following the production of TRAP O-rings in *E. coli* BL21(DE3) cells, the resulting cell lysate was subjected to a thermal treatment in order to reduce the amount of host protein impurities before proceeding to the main purification steps of the protein nanocages.

With this purpose, 750 µL to 1 mL of the clarified cell lysate under study containing the TRAP O-rings (concentration of approximately 2 mg of total protein per mL) was exposed to a significant increase in temperature in relation to the room temperature. Two alternative strategies were evaluated to understand their influence on the reduction of protein impurities and consequently on the following steps based on chromatography. The first approach involved a first incubation step at 37°C for 1 hour followed by a second incubation step at 85°C for 30 minutes. The second strategy involved a single incubation step at 85°C for 30 minutes. At the end of the incubation period, the samples were centrifuged at

12000×g for 20 minutes at room temperature. The resulting supernatant containing the soluble TRAP O-rings was collected and the respective pellet containing the protein impurities precipitated by their sensitivity to high temperatures was resuspended in 500 µL of 50 mM Tris-HCl, 100 mM NaCl, 1 mM EDTA, pH 8.0.

The thermally treated samples were directly used in the next step of purification of the TRAP O-rings based on cation-exchange chromatography.

IV.2.6.2. Anion-exchange chromatography

AEX chromatography was performed using the ÄKTA Start system and the respective UNICORN Start software (Cytiva; Marlborough, MA, USA). The system was operated at a flow rate of 1 mL min^{-1}.

In the experiments performed, the standard conditions were implemented, which consisted of the use of the HiTrap Q FF pre-packed column (1 mL) (Cytiva; Marlborough, MA, USA), a mobile phase composed of the low salt buffer A (50 mM Tris-HCl, 50 mM NaCl, pH 7.9) and the high salt buffer B (50 mM Tris-HCl, 1 M NaCl, pH 7.9), and a bind/elute mode strategy. The column was equilibrated with 5 CV of buffer A ($\kappa \approx 7.2$ mS cm^{-1}). Then, a 5 mL sample (comprising 1 mL of the clarified lysate containing the TRAP O-rings together with 4 mL of buffer A) was injected into the column and all unbound material was washed with 8 CV of buffer A. Note that the approximate mass of total protein in the sample was 2 mg. The elution was performed through a linear gradient from 0% to 100% of buffer B corresponding to 20 CV ($\kappa \approx 80.2$ mS cm^{-1}). The eluate was collected in fractions of 1 mL during the chromatographic run with a fraction collector. All fractions were stored at 4°C until further utilization and/or analysis.

As previously indicated, the column maintenance followed the manufacturer specifications when required.

IV.2.6.3. Concentration and diafiltration by centrifugal filtration

The concentration and diafiltration by centrifugal filtration of samples containing TRAP O-rings previously produced in *E. coli* BL21(DE3) was performed using a similar experimental methodology that was implemented for samples containing the MjsHSP nanocages originating from this strain of *E. coli* (sub-section **IV.2.5.1.3. Concentration and diafiltration by centrifugal filtration** in **IV.2.5.1. Production in *E. coli* cells**, section **IV.2.5. Primary isolation and purification of MjsHSP nanocages**). The differences in relation to this protocol described above were the use of 30 kDa Amicon Ultra-4 mL centrifugal filters (Sigma-Aldrich; St. Louis, MO, USA) and its equilibration with the buffer 50 mM Tris-HCl, 150 mM NaCl, pH 7.9. When applicable, the diafiltration was also performed with this last buffer.

IV.2.6.4. Size exclusion chromatography

The SEC approach was also explored for TRAP O-rings from *E. coli* BL21(DE3) based on an experimental strategy similar to that used with MjsHSP nanocages (sub-section **IV.2.5.1.4. Size exclusion chromatography** in **IV.2.5.1. Production in *E. coli* cells**, section **IV.2.5. Primary isolation**

and purification of MjsHSP nanocages). The differences in relation to this protocol described above consisted of operating at a flow rate of 1 mL min^{-1}, using a suitable mobile phase (50 mM Tris-HCl, 150 mM NaCl, pH 7.9 buffer; $\kappa \approx 16.4$ mS cm^{-1}), injecting 250 µL of sample containing the concentrated TRAP O-rings into the column and collecting the eluate in fractions of 1 mL during the chromatographic run.

As previously indicated, the column maintenance followed the manufacturer specifications when required.

IV.2.6.5. Cation-exchange chromatography

CEX was performed using the ÄKTA Start system and the respective UNICORN Start software (Cytiva; Marlborough, MA, USA). The system was operated at a flow rate of 1 mL min^{-1}.

During the establishment and optimization process, two strong CEX chromatographic supports with sulfonate groups were tested, the HiTrap Capto™ S pre-packed column (1 mL) (Cytiva; Marlborough, MA, USA) and the Capto™ SP ImpRes resin packed in a Tricorn 5/50 column for a target volume of 1 mL (following the manufacturer's guidelines) (Cytiva; Marlborough, MA, USA). The particle size of HiTrap Capto™ S and Capto™ SP ImpRes is of 90 µm and about 36 to 44 µm, respectively. The use of different combinations of mobile phases (binding/washing – A and elution – B buffers) was also analyzed, considering the variation in the composition of the elution buffer in terms of NaCl concentration and/or pH (50 mM MES, 2 M NaCl, pH 5.5; 50 mM Tris-HCl, pH 8.0; and 50 mM Tris-HCl, 2 M NaCl, pH 8.0). This resulted in a buffer without salt (A) composed of 50 mM MES, pH 5.5 and a high salt buffer (B) consisting of 50 mM MES, 2 M NaCl, pH 5.5. In all test experiments, a linear gradient elution was initially assessed (20 CV), allowing for an analysis of the behavior and results obtained, which were later adjusted to an improved step gradient.

For the optimized protocol, the column was equilibrated with 5 CV of 8% buffer B ($\kappa \approx 20.7$ mS cm^{-1}). Then, a 5 mL sample (comprising approximately 1 mL of thermally treated or non-thermally treated clarified lysate containing the TRAP O-rings together with 4 mL of 8% buffer B) was injected into the column and all unbound material was washed with 8 CV of 8% buffer B. Note that the approximate mass of total protein in the sample was 2 mg. It was implemented a two-step gradient for the elution of bound species, the first step with 8 CV of 35% buffer B ($\kappa \approx 61.9$ mS cm^{-1}) and the second step with 8 CV of 100% buffer B ($\kappa \approx 141.3$ mS cm^{-1}). The eluate was collected in fractions of 1 mL during the chromatographic run with a fraction collector. All fractions were stored at 4°C until further utilization and/or analysis.

As previously indicated, the column maintenance followed the manufacturer specifications when required.

IV.2.7. Analytics

IV.2.7.1. Total protein quantification

Total protein concentration was determined by the bicinchoninic acid (BCA) method, according to the experimental procedure described in detail in section **III.2.14.1. Total protein quantification** of **Chapter III**.

IV.2.7.2. SDS-PAGE

Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) was implemented to evaluate the composition of the samples both in specific protein nanocage (monomers) and total protein. The experimental procedure used is described in detail in section **III.2.14.2. SDS-PAGE** of **Chapter III**.

IV.2.7.3. Native PAGE

Native polyacrylamide gel electrophoresis (Native PAGE) was performed with purified samples for evaluating the conformational state and assembly of protein nanocages. Samples were diluted with native sample buffer (2x) (Bio-Rad; Hercules, CA, USA), without any denaturation step. Afterwards, samples were applied in a 4-15% NZY Tris-Glycine precast gel (NZYtech; Lisbon, Portugal). The molecular weight marker used was NZYColour protein marker II (NZYtech; Lisbon, Portugal). The samples were run at 150 V for 4 hours using a running buffer composed of 192 mM glycine and 25 mM Tris-HCl, pH 8.3. Similarly to SDS-PAGE, gels were stained with PhastGel™ Blue R (Coomassie R 350 stain) (GE Healthcare; Chicago, IL, USA). Gels were scanned with GS-800 Calibrated Imaging Densitometer (Bio-Rad; Hercules, CA, USA) and then subject to image processing using Quantity One software (Bio-Rad; Hercules, CA, USA).

IV.2.7.4. Protein nanocages quantification by densitometric analysis of SDS-PAGE

The quantification of MjsHSP nanocages and TRAP O-rings, specifically the respective monomers, was performed through densitometric analysis of SDS-PAGE images analyzed by the ImageJ software (206). The strategy implemented was the same as that previously described in section **III.2.14.3. Protein nanocages quantification by densitometric analysis of SDS-PAGE** of **Chapter III**.

IV.2.7.5. Analytical hydrogen bond chromatography

Clarified lysates and several samples from the recovery and purification process of MjsHSP nanocages were analyzed by analytical hydrogen bond chromatography (HBC) using a CIMac PrimaS® analytical monolith (1 mL, channel size of 2 µm) (Sartorius; Göttingen, Germany). Analytical HBC was performed using a ÄKTA Purifier 10 system (Cytiva; Marlborough, MA, USA). equipped with a 2 mm

path length UV cell with a fixed wavelength of 280 and 260 nm. Chromatography parameters were monitored and controlled using UNICORN 5 software (Cytiva; Marlborough, MA, USA). The system was operated at a flow rate of 1 mL min^{-1}. The column was equilibrated with 10 CV of 50 mM sodium phosphate, 200 mM NaCl, pH 6.0 buffer (A). Then, 25 µL of each sample was loaded into the column using a suitable injection loop and then all unbound material was washed with 8 CV of buffer A. The elution was performed through a linear gradient from 0% to 100% of 50 mM sodium phosphate, 2.5 M NaCl, pH 6.0 buffer (B) corresponding to 20 CV. A set of alternative binding/washing and elution buffers were tested with the aim of improving the elution process, specifically buffer A was equally composed of 50 mM sodium phosphate, 200 mM NaCl, pH 6.0 and buffer B was composed of 50 mM sodium phosphate, 200 mM NaCl, 200 mM sorbitol, pH 6.0. When applicable, eluate fractions (0.5 mL) were collected during the chromatographic run with a fraction collector. All these fractions were stored at 4°C until further analysis.

It should be noted that the monolith maintenance procedure was performed in accordance with the manufacturer's instructions, when necessary, as previously explained.

IV.2.7.6. Analytical size exclusion chromatography

Clarified lysates and several samples from the recovery and purification process of both protein nanocages were analyzed by analytical SEC using a TSKgel SuperSW3000 column (Tosoh Bioscience; Tokyo, Japan). Analytical SEC was performed using the ÄKTA Purifier 10 system previously mentioned and the respective UNICORN 5 software (Cytiva; Marlborough, MA, USA). The system was operated at a flow rate of 0.35 mL min^{-1}. The column was equilibrated with 3 CV of 100 mM sodium phosphate, 100 mM Na_2SO_4, 0.05% w/v NaN_3, pH 6.7 buffer. Then, 150 µL of each sample was loaded into the column using the autosampler for automatic sample injection. The column was washed with 2 CV of the same buffer mentioned above. When relevant, eluate fractions (0.5 mL) were collected during the chromatographic run with a fraction collector. All these fractions were stored at 4°C until further analysis.

As previously mentioned, the column maintenance followed the manufacturer specifications when required.

IV.2.7.7. gDNA extraction

The experimental approach used for the extraction and purification of genomic DNA from *E. coli* BL21(DE3) and *V. natriegens* Vmax™ X2 is explained in detail in section **III.2.14.4. gDNA extraction** of **Chapter III**.

IV.2.7.8. pDNA extraction

Similarly to gDNA extraction, plasmid DNA from *E. coli* BL21(DE3) and *V. natriegens* Vmax™ X2 was performed according to the experimental method indicated in section **III.2.14.5. pDNA extraction** of **Chapter III**.

IV.2.7.9. Total dsDNA quantification by qPCR

The quantification of double-stranded DNA (dsDNA) in both the clarified lysates and several samples from the recovery and purification process relative to TRAP O-rings in *E. coli* BL21(DE3) was performed through a strategy of quantitative PCR (qPCR) according to the procedure described below. It should be noted that the dsDNA comprised the gDNA originating from the *E. coli* strain cells and the pDNA corresponding to the plasmid for expression of the TRAP O-rings.

The two sets of primers designed and synthesized for the determination of plasmid copy number in **Chapter III** (**III.2.9. Determination of plasmid copy number by qPCR**) were used for this dsDNA quantification since they are equally suitable (**Table III.2**). Briefly, one set of primers was specific for the catalytic α subunit of DNA polymerase III (*dnaE*) chromosomal gene (present in the genome of the *E. coli* strain) and the other specific for the ampicillin resistance (*ampr*) gene present in the plasmid pTRAP.

Subsequently, the gDNA from *E. coli* BL21(DE3) was extracted and purified as described in section **IV.2.7.7. gDNA extraction**. Moreover, using the protocol detailed in section **IV.2.7.8. pDNA extraction**, and plasmid pTRAP was isolated and purified from *E. coli* BL21(DE3) cells. Then, gDNA and pDNA were serial diluted in PCR-grade water to be used as standards in the qPCR reactions (420, 42, 4.2, 0.42 and 0.042 ng in each reaction).

The samples to be analyzed were collected directly from the clarified lysates and throughout the downstream processing steps. For the qPCR reactions all these samples were diluted in PCR-grade water considering a dilution factor of 1:10. These dilutions were prior to placing them in each well of the 96-well plate where reactions were prepared.

qPCR reactions were performed in real time using the QuantStudio 5 Real-Time PCR System (Thermo Fisher Scientific; Waltham, MA, USA) and the NZYSpeedy qPCR Green Master Mix (ROX) (2x) (NZYtech; Lisbon, Portugal). Each qPCR reaction (20 µL) comprised 10 µL of 2x NZYSpeedy qPCR Green Master Mix (ROX), 0.8 µL of each primer (qPCR_Chromosome_F and qPCR_Chromosome_R, or qPCR_Plasmids_F and qPCR_Plasmids_R) to a final concentration of 400 nM and 8.4 µL of gDNA standards, pDNA standards or samples. In all cycles of qPCR reactions, negative controls without addition of gDNA, pDNA or sample were included. The qPCR reactions were performed according to the program: initial denaturation at 95°C for 3 minutes; 40 cycles of (i) 95°C for 5 seconds, (ii) 60°C for 20 seconds (polymerization); final extension at 70°C for 30 seconds. At the end of each extension step, the fluorescence signal was automatically detected and quantified by the QuantStudio 5 Real-Time PCR System. Melting curve analyses, through a temperature gradient of 0.05°C s^{-1} from 70°C to 95°C, were performed to confirm the amplification of only the intended specific targets. At the end of the process, the samples were cooled to 40°C for 30 seconds.

Real time qPCR data were acquired automatically by the QuantStudio 5 Real-Time PCR System software and consisted of the threshold cycle (C_t).

The standard curves for gDNA from *E. coli* BL21(DE3) and for pTRAP from *E. coli* BL21(DE3) were obtained by plotting the logarithm of each concentration against the C_t values (**Supplementary Figure IV.S1** and **Supplementary Figure IV.S2**). Considering the estimated equation of each linear regression, the slope of the respective standard curve was used to determine the amplification efficiency (E) according to **Equation III.1** (**Chapter III**).

Furthermore, based on the equations obtained from both linear regressions and knowing the C_t values determined for each sample, it was possible to calculate the concentration of gDNA and pDNA (pTRAP) in ng μL^{-1} in the original samples, considering the previously performed 1:10 dilution factor.

Regarding the qPCR validation, the specificity of qPCR amplification was confirmed by analyzing the melting temperature (T_m) of each amplified product. Negative controls in all qPCR reactions yielded C_t values greater than 20 to 25, significantly higher than samples, indicating negligible background amplification. To validate the experimental determination, in addition to the coefficient of determination (R^2), the percentage of amplification efficiency (E (%)) of each standard curve was calculated using **Equation III.3** (**Chapter III**).

IV.2.8. Data treatment

IV.2.8.1. Protein nanocages recovery

In order to compare the results obtained in terms of the downstream processing of MjsHSP nanocages, TRAP O-rings and TRAP nanocages for these experimental tests, the recovery yield of protein nanocages in each unit operation was determined. This parameter (in µg mL^{-1}) was computed using **Equation IV.1**.

$$Recovery\ yield\ (\mu g\ mL^{-1}) = \frac{Mass_{protein\ nanocages_{after\ each\ unit\ operation}}}{Volume_{after\ each\ unit\ operation}} \quad \textbf{IV.1}$$

Following this first parameter, it was also possible to determine a value in percentage relative to the efficiency of protein nanocages recovery in each unit operation through **Equation IV.2**. Considering the mass of protein nanocages obtained before and after each unit operation, its quotient multiplied by 100, allowed the determination of this percentage. Note that clarified lysate was considered as the starting point for the calculation.

$$Recovery\ efficiency\ (\%) = \frac{Mass_{protein\ nanocages_{after\ each\ unit\ operation}}}{Mass_{protein\ nanocages_{before\ each\ unit\ operation}}} \times 100 \quad \textbf{IV.2}$$

The global efficiency of the downstream processing of NVPNs ($Global\ recovery\ efficiency$ (%)) was calculated by multiplying the recovery efficiency of each unit operation.

IV.2.8.2. Impurities removal

In the downstream processing, it is crucial to implement a parameter that evaluates the reduction of different types of impurities throughout the various stages of purification.

The global efficiency of impurities removal in the downstream processing of NVPNs was determined from **Equation IV.3**. Considering the previously mentioned global recovery efficiency together with the mass of protein nanocages and each of the impurities (host proteins, gDNA and pDNA) at the beginning and at the end of the complete downstream processing, it was possible to estimate the value for this parameter.

$$Impurities\ removal\ efficiency\ (\%) = \frac{Mass_{impurities_{beginning}} - Mass_{impurities_{end}}}{Mass_{impurities_{beginning}}} \times 100 \qquad \textbf{IV.3}$$

IV.2.9. Statistical analysis

Descriptive statistics, including mean and standard deviation, were calculated for the results obtained across the different variables. The sample sizes for each variable typically ranged between 2 and 3 experiments.

IV.3. Results and Discussion

IV.3.1. Primary isolation and purification of MjsHSP nanocages produced in *E. coli* cells

IV.3.1.1. Standard purification by two-step chromatography

The purification process of NVPNs and specifically of MjsHSP nanocages is typically based on a two-step combination of chromatography (AEX followed by SEC), with an intermediate step of concentration of these protein nanocages, according to what is reported in the literature. This approach was chosen as a starting point for the establishment of a more robust and consistent optimized process for the purification of MjsHSP nanocages.

Thus, the first step was the AEX chromatography, in which a traditional agarose resin was used in a 1 mL pre-packed column (HiTrap Q FF). This strong AEX resin can mediate electrostatic interactions with the solutes in the feed stream through the charged nitrogen of the corresponding quaternary amine ligands. Due to its more conventional composition, it has a large particle size of 90 µm and a small pore size of about 50 to 60 nm.

After some preliminary evaluation tests, the mobile phases used were a buffer A composed of 50 mM sodium phosphate, pH 6.0 and a buffer B composed of 50 mM sodium phosphate, 1 M NaCl, pH 6.0. The AEX column was equilibrated with buffer A ($\kappa \approx 5.8$ mS cm^{-1}). The sample injected into the HiTrap Q FF column consisted of about 1 mL of clarified lysate from *E. coli* BL21(DE3) cells containing MjsHSP nanocages (approximate concentration of 2 mg of total protein per mL). For this standard strategy, unbound material was washed with buffer A, and elution was performed through a linear gradient (20 CV) with increasing salt concentration up to 100% of buffer B ($\kappa \approx 79.6$ mS cm^{-1}).

The chromatogram obtained is shown in **Figure IV.1A**, characterized by a major flowthrough peak (*1*) and three partially separated peaks during the elution stage (*2*, *3*, and *4*). The corresponding fractions were analyzed by SDS-PAGE alongside the feed sample (procedure available in sub-section **IV.2.7.2. SDS-PAGE**). The resulting gel is shown in **Figure IV.1B**. Note that since SDS-PAGE is a denaturing technique, the protein band corresponding to the MjsHSP nanocages effectively represents their monomeric form. All these fractions were also analyzed using the BCA method for total protein quantification (methodology described in sub-section **IV.2.7.1. Total protein quantification**).

The feed sample exhibited a significant amount of protein impurities, as evidenced by the identification of bands with different molecular weights in addition to the MjsHSP nanocages. The nanocages were represented by the band corresponding to their respective monomers (with a molecular weight of approximately 16.5 kDa). An estimated mass for the NVPNs in this feed sample was 837.2 ± 36.8 µg. In the flowthrough (**Figure IV.1A**, *1*), the elution of impurities with low charge density occurred, including host proteins and nucleic acids, as evidenced by the recorded peak. The SDS-PAGE analysis of this fraction showed weak bands corresponding to protein impurities and a small band relating to the MjsHSP nanocages monomers. This indicates that while most of these protein nanocages were retained by the HiTrap Q FF, some loss occurred in the flowthrough due to the NVPNs that did not bind to the resin. The analysis of the other collected fractions confirmed that the peak eluting at approximately

45.0% of buffer B ($\kappa \approx 33.8$ mS cm^{-1}) with a retention volume of about 34 mL contained the MjsHSP nanocages (**Figure IV.1A**, *3*). In the first peak of elution (**Figure IV.1A**, *2*), the respective fractions likely contained some protein impurities. In the last peak of elution (**Figure IV.1A**, *4*), a small quantity of MjsHSP nanocages was detected, as indicated by the SDS-PAGE result. This final peak probably also contained some dsDNA with a higher charge density compared to other impurities and the MjsHSP nanocages themselves. Considering the pI of the MjsHSP nanocages (5.02), the use of a pH of 6.0 ensured that the MjsHSP nanocages were negatively charged. Simultaneously, a large portion of the protein impurities exhibited a lower charge density, allowing their elution from the resin while the MjsHSP nanocages remained bound. The nanocages were retained until their elution was achieved when a higher NaCl concentration was reached. Nevertheless, dsDNA is a challenging impurity due to its highly negatively charged nature at this pH by the phosphate backbone, which caused it to compete with the MjsHSP nanocages for binding to the functional groups of the resin.

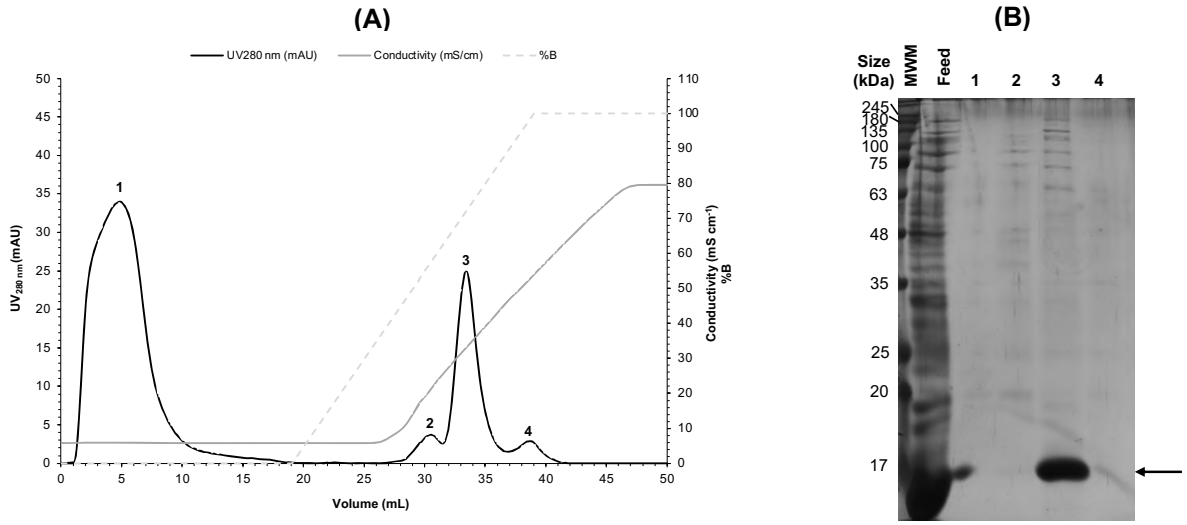


Figure IV.1 – AEX chromatography as the first step for the standard purification of MjsHSP nanocages produced in *E. coli* BL21(DE3). **(A)** Chromatogram obtained using a 1 mL pre-packed column with the traditional agarose resin HiTrap Q FF (Cytiva; Marlborough, MA, USA). A feed stream corresponding to approximately 2 mg of total protein was injected into this column previously equilibrated with buffer A composed of 50 mM sodium phosphate, pH 6.0 ($\kappa \approx 5.8$ mS cm^{-1}). Unbound material was washed with this buffer A and elution was performed using a linear gradient (20 CV) with increasing salt concentration up to 100% of buffer B composed of 50 mM sodium phosphate, 1 M NaCl, pH 6.0 ($\kappa \approx 79.6$ mS cm^{-1}). **(B)** SDS-PAGE analysis of fractions collected during the chromatographic run (peaks *1*, *2*, *3*, and *4*) together with the clarified lysate feed sample. Peak *1* corresponded to the flowthrough and peaks *2* to *3* to the eluted material. Arrow indicates the protein bands referring to the monomer of MjsHSP nanocages (expected molecular weight of 16.5 kDa), which is separated and visualized under denaturing conditions of SDS-PAGE. The abbreviation MWM refers to the molecular weight marker (NZYColour protein marker II; NZYtech Lisbon, Portugal).

The fractions of peak *3* containing the MjsHSP nanocages (**Figure IV.1A** and **B**) yielded a mass of 161.0 ± 5.4 µg of these protein nanocages, corresponding to a recovery yield of 35.8 ± 1.2 µg mL^{-1} and subsequent recovery efficiency of 19.2 ± 0.2% relative to the feed with the clarified lysate. These

last two parameters (recovery yield and recovery efficiency) were estimated using **Equation IV.1** and **Equation IV.2**, respectively. As in **Chapter III**, the determination of these three parameters implied a densitometric analysis of SDS-PAGE, which was performed according to the method described above (**IV.2.7.4. Protein nanocages quantification by densitometric analysis of SDS-PAGE**). Due to the composition of each MjsHSP nanocage in 24 identical monomers, the total mass of protein nanocages monomers is identical to the mass of the assembled nanocages.

Afterwards, the fractions containing MjsHSP nanocages were concentrated by centrifugal filtration, utilizing commercial centrifugal filters with a 100 kDa molecular weight cut-off. The need for this methodology aimed to overcome one of the limitations of SEC, the second step of this chromatography-based standard approach, which consists of the necessity to have a small sample injection volume relative to the total resin volume (about 1%). It was possible to obtain a mass of MjsHSP nanocages of 96.3 ± 2.0 µg, corresponding to 385.2 ± 7.9 µg mL^{-1} of recovery yield at the end of this unit operation. The recovery efficiency was 59.9 ± 3.2%, a somewhat unsatisfactory result, although several measures were implemented to reduce the loss of protein nanocages during the procedure, namely the passivation of the centrifugal filters membrane with 5% v/v Tween 20 and its equilibration with a suitable buffer to minimize non-specific interactions.

Next, this sample containing the concentrated MjsHSP nanocages was subjected to the second step of chromatography-based purification with SEC. A size exclusion resin (Superdex™ 200) packed in a chromatographic column with a target volume of 24 mL was selected. In this resin, separation results exclusively from the hydrodynamic size, with the stationary phase having no elements that promote the occurrence of secondary chemical interactions. It presents particles with a size of 13 µm, allowing the separation of molecules with molecular weights between 10 and 600 kDa. The mobile phase used was a PBS solution (1x), with the resin equilibrated using this buffer at a flow rate of 1 mL min^{-1}. After injecting the referred sample into the column, this was washed with 2 CV of PBS solution (1x), maintaining a constant conductivity value ($\kappa \approx 14.2$ mS cm^{-1}).

The chromatogram obtained is shown in **Figure IV.2A**, being characterized by two partially separated peaks (*1* and *2*). The corresponding fractions were analyzed by SDS-PAGE in comparison with the feed sample. The resulting gel is shown in **Figure IV.2B**. All these fractions were similarly analyzed using the BCA method for total protein quantification.

The feed presented an intense band corresponding to the monomers of the MjsHSP nanocages and many subtle bands with higher molecular weight. Both peaks *1* and *2* were characterized by unique bands corresponding to these NVPNs, which aligns with the low resolution of these two peaks, suggesting some overlap in their content (**Figure IV.2A**). However, this may indicate the presence of two populations of MjsHSP nanocages, for example, one composed of individual nanocages (peak *2*, retention volume of 8.4 mL) and another containing some aggregates (peak *1*, retention volume of 9.6 mL). Additionally, the size exclusion resin used could not provide sufficient accuracy to achieve complete separation of these two populations. These potential aggregates, being larger in size, eluted earlier from the column as they do not pass through its pores. A drawback of SEC is that it dilutes the samples, which means that some host protein impurities present in low concentrations may not have been detected by SDS-PAGE, even when using the more sensitive silver staining technique.

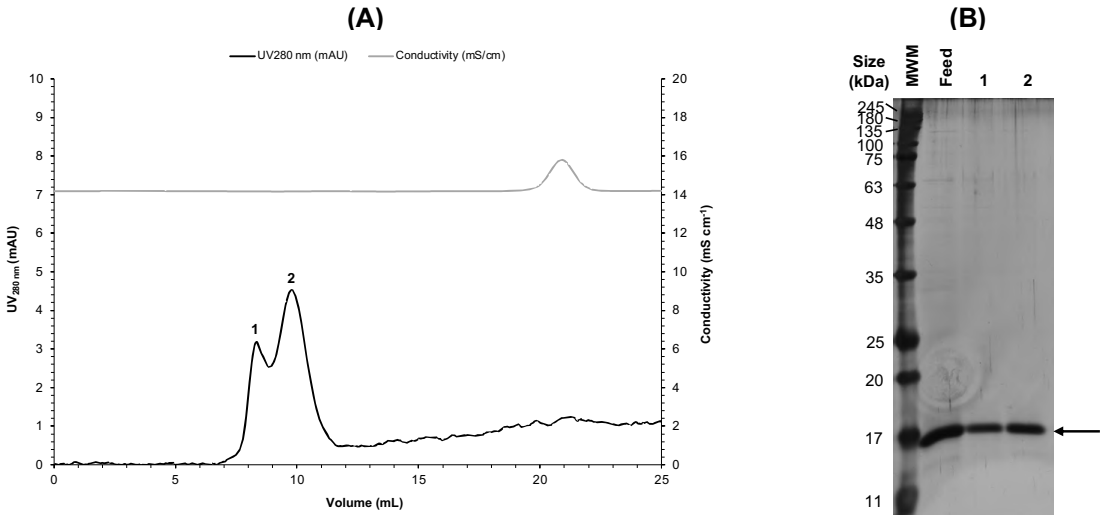


Figure IV.2 – SEC as the second step for the standard purification of MjsHSP nanocages produced in *E. coli* BL21(DE3). **(A)** Chromatogram obtained using a 24 mL size exclusion resin Superdex™ 200 (Cytiva; Marlborough, MA, USA) packed in an appropriate chromatographic column. The feed stream was comprised by the previously concentrated MjsHSP nanocages. The column was equilibrated with a PBS solution (1x) ($\kappa \approx 14.2$ mS cm^{-1}) at a flow rate of 1 mL min^{-1}. **(B)** SDS-PAGE analysis of fractions collected during the chromatographic run (peaks *1* and *2*) together with the concentrated protein nanocages feed sample. Arrow indicates the protein bands referring to the monomer of MjsHSP nanocages (expected molecular weight of 16.5 kDa), which is separated and visualized under denaturing conditions of SDS-PAGE. The abbreviation MWM refers to the molecular weight marker (NZYColour protein marker II; NZYtech Lisbon, Portugal).

The fractions related to peak *1* yielded a mass of 11.2 ± 1.3 µg of MjsHSP nanocages, corresponding to a recovery yield of 7.5 ± 0.9 µg mL^{-1} and a recovery efficiency of 15.1 ± 1.5% relative to the feed, while the fractions related to peak *2* yielded a mass of 44.6 ± 1.3 µg of protein nanocages, corresponding to a recovery yield of 17.8 ± 0.5 µg mL^{-1} and a recovery efficiency of 60.1 ± 0.6% (**Figure IV.2A** and **B**). Overall, if the content of these MjsHSP nanocages were considered from both peaks, a mass of 55.8 ± 2.6 µg of MjsHSP nanocages, a recovery yield of 25.3 ± 1.4 µg mL^{-1} and a recovery efficiency of 75.2 ± 2.1% were obtained.

Finally, considering this standard purification process composed of three steps, the global recovery efficiency of MjsHSP nanocages was 8.7 ± 0.1% with the removal efficiency of the host protein impurities having been 97.7 ± 1.2%, which provides a pathway to explore and optimize this downstream processing.

IV.3.1.2. Optimized purification by chromatography

Following the establishment of the standard methodology for purifying MjsHSP nanocages, a more detailed study was initiated to optimize the process, with the ultimate goal of increasing the recovery of protein nanocages while removing the maximum possible quantity of contaminants. AEX

chromatography was maintained as the core purification strategy due to its advantages in exploiting the properties of MjsHSP nanocages in relation to the host impurities.

Three strong AEX chromatographic supports were selected as alternatives to the traditional HiTrap Q FF: Capto™ Q ImpRes, POROS™ 50 HQ and CIMmultus® QA. Capto™ Q ImpRes and POROS™ 50 HQ are classified as gigaporous resins due to their reduced particle size and large pore size. CIMmultus® QA is a monolith with a channel size of 2 µm. Due to their characteristics, these two resins and the monolith could offer superior separation performance for large biomolecules such as the MjsHSP nanocages.

For these exploratory assays, the same conditions described previously for the standard approach were utilized in terms of mobile phases (buffers A and B), the mass of sample injected into the columns and monolith, and the binding/washing and elution procedures (sub-section **IV.3.1.1. Standard purification by two-step chromatography**).

The chromatograms obtained are presented in **Figure IV.3A**, **C** and **E**, along with the corresponding SDS-PAGE images (**Figure IV.3B**, **D** and **F**). All fractions were analyzed by SDS-PAGE and quantified for total protein content using the BCA method (sub-sections **IV.2.7.2. SDS-PAGE** and **IV.2.7.1. Total protein quantification**).

Generally, the chromatogram profiles were quite similar among the three evaluated alternatives and also in comparison to the standard HiTrap Q FF column. They all present a major flowthrough peak and three to four peaks during the elution phase. Note that the feed sample was equivalent for all experiments, showing bands of host protein impurities with different molecular weights in addition to the MjsHSP nanocages. The mass of these NVPNs loaded ranged from 510.2 ± 12.3 to 843.3 ± 2.4 µg. In all flowthrough fractions (**Figure IV.3A**, **C** and **E**, *1*), a loss of protein nanocages that did not bind to the resin and monolith was observed. This may indicate that the binding capacity of the chromatographic support was reached, not only due to the target protein nanocages but also because some impurities that were interacting with the functional groups of these separation media. Regarding the peaks eluting with the linear NaCl gradient, the most prominent peak corresponded to the MjsHSP nanocages (**Figure IV.3A**, **C** and **E**, *4*) and it was confirmed by SDS-PAGE results. For all the three alternatives, the retention volume was approximately 33 mL when conductivity ranged between 30 and 35 mS cm^{-1}. These results were consistent among themselves and also with the standard HiTrap Q FF method.

For Capto™ Q ImpRes, the fractions of peak *4* containing the MjsHSP nanocages (**Figure IV.3A** and **B**) yielded a mass of 172.20 ± 1.6 µg of these protein nanocages, corresponding to a recovery yield of 43.0 ± 0.4 µg mL^{-1} and subsequent recovery efficiency of 33.8 ± 1.1% relative to the feed with the clarified lysate. These values represented an increase of approximately 1.76 times compared to the standard methodology. For POROS™ 50 HQ, considering the corresponding peak (**Figure IV.3C** and **D**), the mass of MjsHSP nanocages was 105.3 ± 0.6 µg, the recovery yield was 30.1 ± 0.2 µg mL^{-1}, and the recovery efficiency was 13.2 ± 0.1%. These values represented a decrease of approximately 1.46 times compared to the benchmark method. In terms of the CIMmultus® QA monolith (**Figure IV.3E** and **F**), the mass of MjsHSP nanocages of 53.1 ± 3.2 µg, a recovery yield of 17.7 ± 1.1 µg mL^{-1} and a recovery efficiency of 6.3 ± 0.4% were estimated. Similarly to the POROS™ 50 HQ results, the data for these parameters revealed a decrease of 3.05 times compared to the standard strategy. As previously

explained, the parameters recovery yield and recovery efficiency were calculated using **Equation IV.1** and **Equation IV.2**, respectively. These calculations were based on the quantification obtained through densitometric analysis of SDS-PAGE (**IV.2.7.4. Protein nanocages quantification by densitometric analysis of SDS-PAGE**).

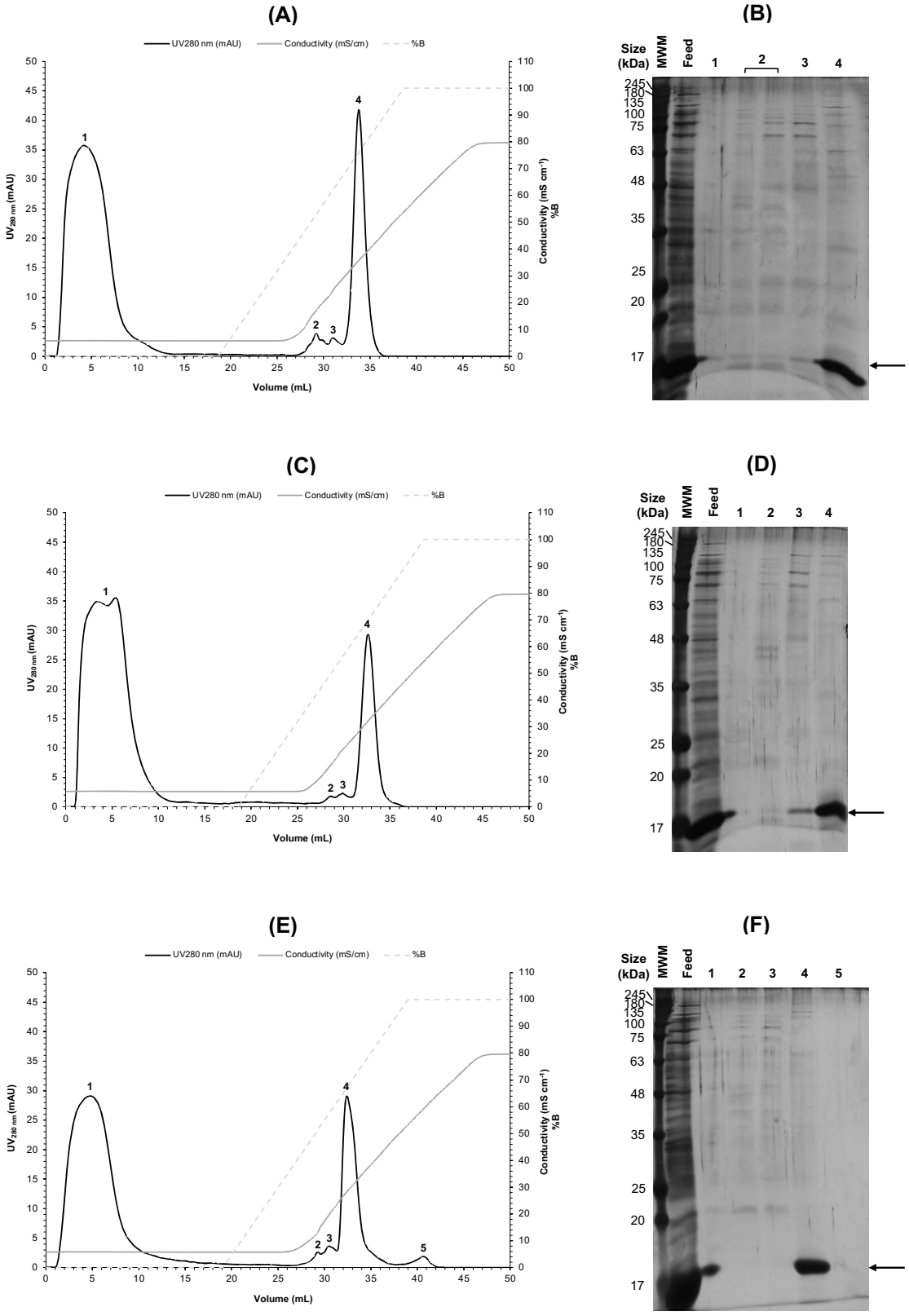


Figure IV.3 – Optimization of the AEX chromatography for the purification of MjsHSP nanocages produced in *E. coli* BL21(DE3). Chromatograms obtained using a 1 mL chromatographic column with the resin Capto™ Q ImpRes (Cytiva; Marlborough, MA, USA) **(A)** and the resin POROS™ 50 HQ (Thermo Fisher Scientific; Waltham, MA, USA) **(C)**, and a 1 mL monolith CIMmultus® QA (Sartorius; Göttingen, Germany) **(E)**. A similar feed stream corresponding to approximately 2 mg of total protein was injected into the columns and the monolith previously equilibrated with buffer A composed of 50 mM sodium phosphate, pH 6.0 ($\kappa \approx 5.8$ mS cm^{-1}). Unbound material was washed with this buffer A and elution was performed using a linear gradient (20 CV) with increasing salt concentration up to 100% of buffer B composed of 50 mM sodium phosphate, 1 M NaCl, pH 6.0 ($\kappa \approx 79.6$ mS cm^{-1}). SDS-PAGE analysis of fractions collected during the chromatographic run together with the clarified lysate feed sample for the three strong AEX chromatographic supports tested: Capto™ Q ImpRes **(B)**, POROS™ 50 HQ **(D)** and CIMmultus® QA **(E)**. Peak *1* corresponded to the flowthrough and peaks *2* to *4* or *5* to the eluted material. Arrow indicates the protein bands referring to the monomer of MjsHSP nanocages (expected molecular weight of 16.5 kDa), which is separated and visualized under denaturing conditions of SDS-PAGE. The abbreviation MWM refers to the molecular weight marker (NZYColour protein marker II; NZYtech Lisbon, Portugal).

Therefore, the results demonstrated that Capto™ Q ImpRes was the strong AEX chromatographic support that exhibited the highest efficiency for recovering MjsHSP nanocages from the clarified lysate among the four tested options. Furthermore, Capto™ Q ImpRes showed superior peak resolution characteristics, with a sharper elution peak corresponding to the MjsHSP nanocages and minimal tailing effects. These observations can be justified by the smaller bead size of Capto™ Q ImpRes (40 µm) compared to the other two resins (HiTrap Q FF with 90 µm and POROS™ 50 HQ with 50 µm). Capto™ Q ImpRes has a pore size of 70 nm, which is smaller than that of CIMmultus® QA monolith (2 µm).

Considering these promising results for Capto™ Q ImpRes, it was decided to proceed with a more optimized procedure for the AEX-based purification of MjsHSP nanocages. This optimization involved defining a stepwise elution gradient, with the ultimate goal of establishing a defined and consistent strategy.

Regarding the experimental conditions, the mobile phases used were a buffer A composed of 50 mM sodium phosphate, 250 mM NaCl, pH 6.0 and a buffer B consists of 50 mM sodium phosphate, 1 M NaCl, pH 6.0. Capto™ Q ImpRes column was previously equilibrated with buffer A ($\kappa \approx 26.12$mS cm^{-1}). The sample injected into the column comprised approximately 2 mg of total protein from the clarified lysate containing the MjsHSP nanocages. After washing the unbound material with buffer A, a two-step gradient for the elution of bound species was applied, the first step with 23% of buffer B ($\kappa \approx 40.5$ mS cm^{-1}) and the second step with 100% of this buffer B ($\kappa \approx 81.5$ mS cm^{-1}).

The chromatogram obtained is presented in **Figure IV.4A**, along with the resulting SDS-PAGE image (**Figure IV.4B**). As before, the fractions from the chromatographic run were analyzed by SDS-PAGE and quantified for total protein content using the BCA method (sub-sections **IV.2.7.2. SDS-PAGE** and **IV.2.7.1. Total protein quantification**).

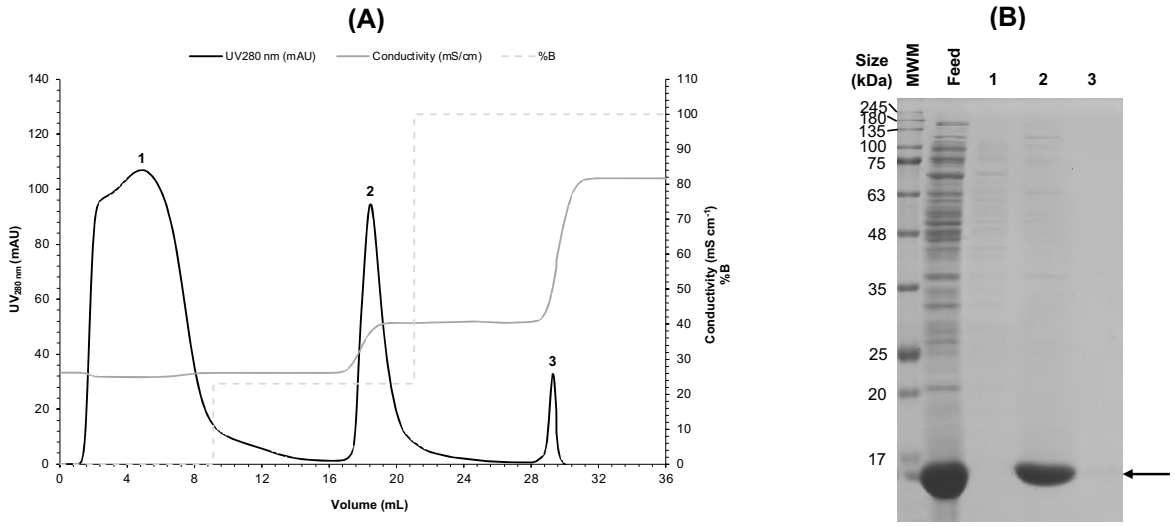


Figure IV.4 – Optimized AEX chromatography strategy for the purification of MjsHSP nanocages produced in *E. coli* BL21(DE3). **(A)** Chromatogram obtained using a 1 mL chromatographic column with the gigaporous resin Capto™ Q ImpRes (Cytiva; Marlborough, MA, USA). A feed stream corresponding to approximately 2 mg of total protein was injected into this column previously equilibrated with buffer A composed of 50 mM sodium phosphate, 250 mM NaCl, pH 6.0 ($\kappa \approx 26.1$ mS cm^{-1}). Unbound material was washed with this buffer A and elution was performed using a two-step gradient, the first step with 23% of buffer B ($\kappa \approx 40.5$ mS cm^{-1}) and the second step with 100% of this buffer B ($\kappa \approx 81.5$ mS cm^{-1}). Buffer B was composed of 50 mM sodium phosphate, 1 M NaCl, pH 6.0. **(B)** SDS-PAGE analysis of fractions collected during the chromatographic run (peaks *1*, *2*, and *3*) together with the clarified lysate feed sample. Peak *1* corresponded to the flowthrough and peaks *2* and *3* to the eluted material. Arrow indicates the protein bands referring to the monomer of MjsHSP nanocages (expected molecular weight of 16.5 kDa), which is separated and visualized under denaturing conditions of SDS-PAGE. The abbreviation MWM refers to the molecular weight marker (NZYColour protein marker II; NZYtech Lisbon, Portugal).

Considering these results, a highly interesting profile was observed with a major flowthrough peak, a peak in the first elution step corresponding to the fractions containing the MjsHSP nanocages, and a second elution step related to the host impurities, probably nucleic acids as previously hypothesized. The feed sample contained approximately 955.0 ± 18.4 µg of MjsHSP nanocages. In the flowthrough (**Figure IV.4A**, *1*), no protein nanocages were observed to be washed, which was an improvement compared to the previous experiments with the linear elution gradient. SDS-PAGE analysis revealed the removal of protein impurities, with several bands of different molecular weights appearing. Peak *2* of the NVPNs exhibited a sharper profile with reduced tailing, as expected, presenting a retention volume of approximately 18 mL at a conductivity of 36.2 mS cm^{-1} (**Figure IV.4A**, *2*). The SDS-PAGE of peak *2* showed a well-defined band of strong intensity corresponding to the monomers of the MjsHSP nanocages, along with very weak bands that may represent residual host protein impurities, although in virtually negligible quantities (**Figure IV.4B**, *2*). For the peak *3*, no protein bands were observed in the SDS-PAGE, which supported the hypothesis that some nucleic acids, particularly ds DNA, were being retained in the resin and only eluted when the NaCl concentration increased to the corresponding conductivity of 49.7 mS cm^{-1} (**Figure IV.4A** and **B**, *3*).

The fractions of peak *2* containing the MjsHSP nanocages (**Figure IV.4A** and **B**) yielded a mass of 701.6 ± 8.0 µg of these NVPNs, corresponding to a recovery yield of 140.3 ± 1.6 µg mL^{-1} and subsequent recovery efficiency of 73.5 ± 0.6% relative to the feed with the clarified lysate. These values represent an increase of approximately 2.18 times compared to the standard methodology, which is a highly significant improvement.

To evaluate the complete process with all three purification steps, the MjsHSP nanocages resulting from AEX chromatography were concentrated by centrifugal filtration. This was accomplished using commercial centrifugal filters with a 100 kDa molecular weight cut-off. The experimental methodology employed was the same as described in the results section above (**IV.3.1.1. Standard purification by two-step chromatography**). It was possible to obtain a mass of MjsHSP nanocages of 357.1 ± 4.8 µg, corresponding to 1428.4 ± 19.0 µg mL^{-1} and a recovery efficiency of 46.8 ± 0.4% at the end of this unit operation.

Subsequently, these concentrated MjsHSP nanocages were subjected to the second step of chromatography-based purification using SEC. Again, the experimental methodology implemented was the same as described in section **IV.3.1.1. Standard purification by two-step chromatography**. The results obtained, both in terms of the chromatogram and the corresponding SDS-PAGE, are shown in **Figure IV.5A** and **B**, respectively.

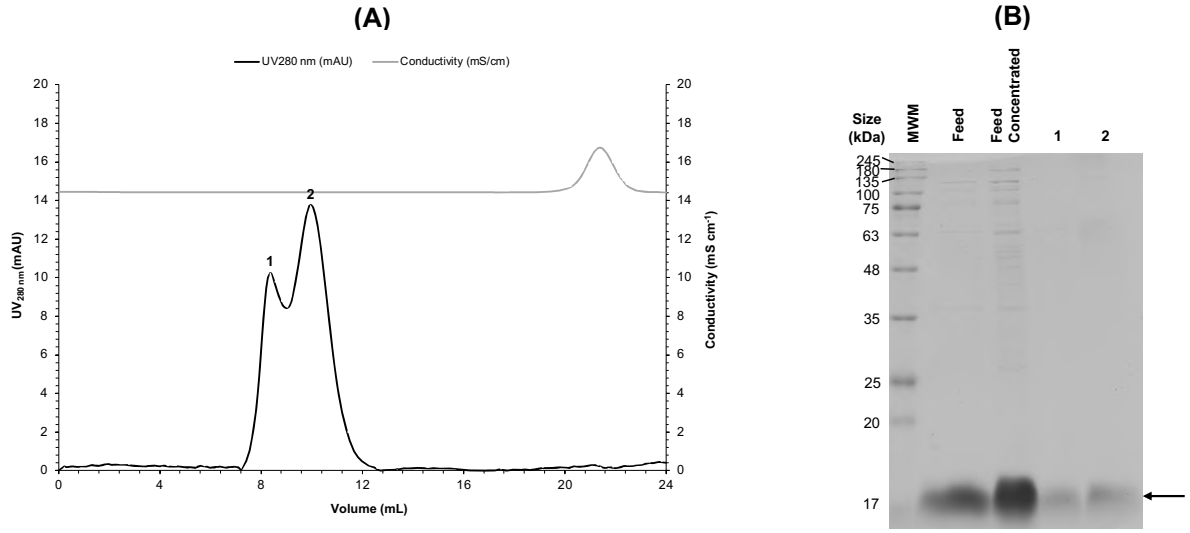


Figure IV.5 – SEC as the second step for the optimized purification of MjsHSP nanocages produced in *E. coli* BL21(DE3). **(A)** Chromatogram obtained using a 24 mL size exclusion resin Superdex™ 200 (Cytiva; Marlborough, MA, USA) packed in an appropriate chromatographic column. The feed stream was comprised by the previously concentrated MjsHSP nanocages (from optimized AEX chromatography). The column was equilibrated with a PBS solution (1x) ($\kappa \approx 14.2$ mS cm^{-1}) at a flow rate of 1 mL min^{-1}. **(B)** SDS-PAGE analysis of fractions collected during the chromatographic run (peaks *1* and *2*) together with the original and concentrated protein nanocages feed samples. Arrow indicates the protein bands referring to the monomer of MjsHSP nanocages (expected molecular weight of 16.5 kDa), which is separated and visualized under denaturing conditions of SDS-PAGE. The abbreviation MWM refers to the molecular weight marker (NZYColour protein marker II; NZYtech Lisbon, Portugal).

The results were quite similar to those previously described, which confirms that the optimization of the AEX chromatography did not interfere with the content of the MjsHSP nanocages samples and, consequently, with the separation behavior in the SEC. This consistency demonstrated the robustness of the purification process and validated the effectiveness of the optimized AEX step in the overall purification strategy.

The fractions related to peak *1* yielded a mass of 19.4 ± 0.7 µg of MjsHSP nanocages, corresponding to a recovery yield of 9.7 ± 0.3 µg mL^{-1} and a recovery efficiency of 7.0 ± 0.1% relative to the feed, while the fractions related to peak *2* yielded a mass of 53.4 ± 1.4 µg of protein nanocages, corresponding to a recovery yield of 15.2 ± 0.54 µg mL^{-1} and a recovery efficiency of 19.4 ± 1.4% (**Figure IV.5A** and **B**). Overall, if the content of these MjsHSP nanocages were considered from both peaks, a mass of 72.8 ± 0.7 µg of MjsHSP nanocages, a recovery yield of 24.9 ± 0.1 µg mL^{-1} and a recovery efficiency of 26.4 ± 1.5% were obtained.

Considering this optimized purification process composed of three steps, the global recovery efficiency was 9.5 ± 1.8%, with a host protein impurities removal efficiency of 97.5 ± 1.6%. This corresponded to a 1.10-fold increase in the recovery efficiency and a similar impurities removal efficiency compared to the standard protocol with the non-optimized AEX chromatography.

However, particularly in terms of the protein nanocages recovery, about 9.5% as a relatively low value, which led to the hypothesis that the three-step process could be further improved. When considering only the AEX chromatography step, it was obtained a recovery efficiency of 73.5 ± 0.7%, and an impurities removal efficiency of 73.5 ± 1.2%. This represents a better value for the first parameter, although slightly lower for the second. If the AEX chromatography step followed by centrifugal filtration concentration is considered, the protein nanocages recovery decreased to 36.1 ± 1.8%, but the removal of host protein impurities increased to 91.1 ± 1.4%.

Therefore, these calculations suggest that the procedure using the AEX chromatography step alone may be the best option to achieve a good balance between the amount of protein nanocages recovered and the quantity of host protein impurities removed. Even in comparison to the standard method, it was observed an increase of 8.4 times in terms of the protein nanocages recovery while the protein impurities removal decreased 1.3 times. Additional purification steps may increase the level of purification, but they significantly reduce (up to approximately 7.8 times) the mass of MjsHSP nanocages recovered, which is not desirable for a scalable strategy.

In this regard, SEC is also a challenging approach for scale-up due to its intrinsic characteristics, so implementing a downstream processing without its presence can be favorable.

It was then decided to evaluate a multimodal resin Capto™ Core 400, which has an inactive shell that excludes large molecules (average molecular weight cut-off of 400 kDa) while allowing smaller impurities to enter and bind to the ligand-activated core (combination of AEX and SEC). Since MjsHSP nanocages have a molecular weight of approximately 396 kDa, it was intended to test its effectiveness. Capto™ Core 400 presents advantages such as the possibility of a higher sample load, which would avoid the need to use a concentration step before. Additionally, it would allow the use of a faster flow rate, reducing the process time.

In terms of the experimental conditions, a flowthrough mode strategy was used due to the intrinsic characteristics of this resin, with a pre-equilibration of the column with PBS solution (1x). The feed sample was a pool of fractions from AEX chromatography (2 to 3 mL). At the end, the column was washed with 1 M NaCl to elute any components that might have bound to the interior of the Capto™ Core 400 resin.

The chromatogram obtained is presented in **Figure IV.6A**, together with the resulting SDS-PAGE image (**Figure IV.6B**). Once more, the fractions from the chromatographic run were analyzed by SDS-PAGE and quantified for total protein using the BCA method (sub-sections **IV.2.7.2. SDS-PAGE** and **IV.2.7.1. Total protein quantification**).

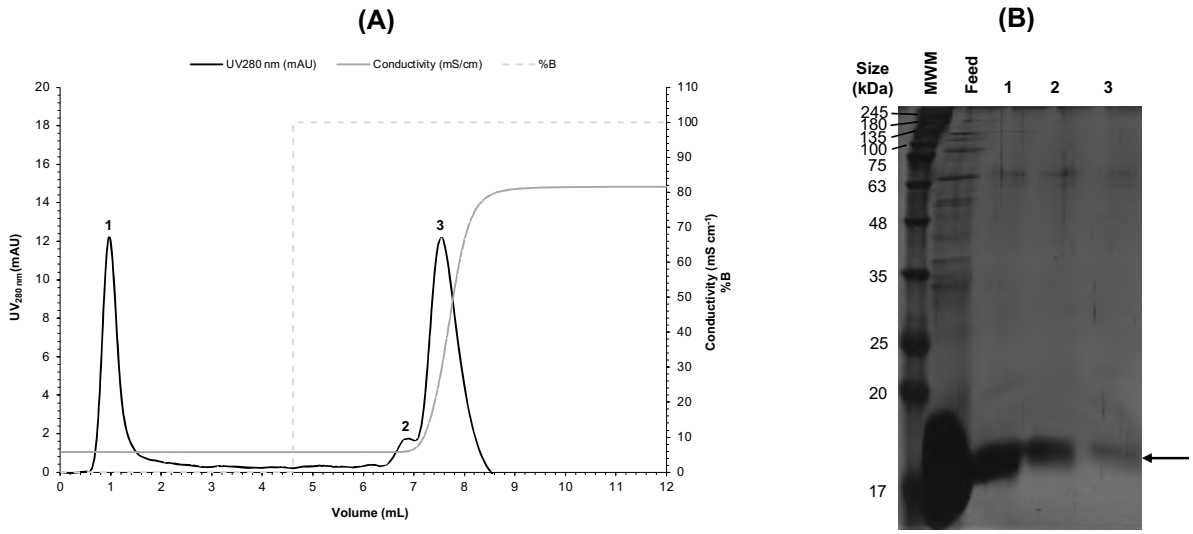


Figure IV.6 – Multimodal chromatography as an alternative to SEC as a second step for the optimized purification of MjsHSP nanocages produced in *E. coli* BL21(DE3). **(A)** Chromatogram obtained using a 1 mL resin Capto™ Core 400 (Cytiva; Marlborough, MA, USA) packed in an appropriate chromatographic column. The feed stream was comprised by the MjsHSP nanocages previously subject to the optimized AEX chromatography. The column was equilibrated with a PBS solution (1x) ($\kappa \approx 14.2$ mS cm^{-1}). At the end, the column was washed with a 1 M NaCl solution to remove material that had bound to the resin. **(B)** SDS-PAGE analysis of fractions collected during the chromatographic run (peaks *1*, *2*, and *3*) together with the protein nanocages feed sample. Arrow indicates the protein bands referring to the monomer of MjsHSP nanocages (expected molecular weight of 16.5 kDa), which is separated and visualized under denaturing conditions of SDS-PAGE. The abbreviation MWM refers to the molecular weight marker (NZYColour protein marker II; NZYtech Lisbon, Portugal).

It was possible to observe a flowthrough peak, which contained a large quantity of MjsHSP nanocages as illustrated by SDS-PAGE (**Figure IV.6A** and **B**, *1*). However, these NVPNs were also retained by the resin and eluted after washing with high salt concentration (**Figure IV.6A** and **B**, *2* and *3*). This resulted in a mass recovery of 118.1 ± 2.7 µg, a recovery yield of 78.8 ± 1.8 µg mL^{-1} and a recovery efficiency of 11.1 ± 0.4%. These observations can be justified, on one hand, by the fact that the average molecular weight cut-off of the resin is very close to the size of fully assembled MjsHSP nanocages, which may lead to some nanocages entering the resin core. On the other hand, these

protein nanocages may not be completely assembled, meaning that some of them will enter due to their significantly smaller size.

To attempt to evaluate whether this chromatographic approach could help distinguish between the various oligomeric forms of the MjsHSP nanocages, the collected samples were analyzed by native SDS-PAGE according to the procedure described in the methodological sub-section **IV.2.7.3. Native PAGE**. However, this is a challenging technique with difficulties in identifying optimal running conditions. As a result, it was not possible to establish with certainty that the relationship between the separation on Capto™ Core 400 and the assembly state of the protein nanocages can be considered clear and consistent.

A limitation identified throughout this work was in the quantification of the MjsHSP nanocages at different stages of the downstream processing by densitometric analysis of SDS-PAGE. In this sense and trying to explore the intrinsic properties of protein nanocages in terms of large size, it was decided to try to implement chromatography-based quantification, which would make the process more consistent, robust, and efficient.

A CIMac PrimaS® analytical monolith was used, which is a multimodal media that combines primarily hydrogen bond chromatography with a secondary AEX chromatography component. The objective was not only to distinguish the MjsHSP nanocages from other impurities but also within the protein nanocages themselves to try to distinguish different oligomeric compositions. Particularly, hydrogen bond chromatography could be useful for this objective since it is designed for large biomolecules, based on hydrogen bond interactions which are more prominent in this type of biological molecules. However, it was not possible to obtain positive and consistent results between the various experiments, with difficulties identified in having complete elution of the protein-based nanostructures from the monolith as well as difficulty in separating peaks with good resolution particularly in less purified samples. Further research in this topic is necessary in order to achieve a clearly robust strategy for this analytical quantification.

IV.3.1.3. Evaluation of pre-treatment of cell lysate with DNAse and RNAse

Considering the optimized purification process of MjsHSP nanocages produced in *E. coli* BL21(DE3) based on chromatography, one of the most challenging impurities with the greatest impact was host nucleic acids. Therefore, based on the available data in the literature, the hypothesis of adding a preliminary step of treating the clarified lysate with a nuclease, specifically DNase and/or RNase, was proposed in order to reduce the amount of this type of impurities, helping in the chromatography step and eventually increasing its effectiveness.

To evaluate this hypothesis, two main approaches were tested. The first involved subjecting a clarified cell lysate containing the MjsHSP nanocages to a TURBO™ DNase treatment (8 U), while the second approach treated a similar sample of the clarified lysate using only RNase A (4 U) (sub-section **IV.2.5.1.1. Treatment of cell lysate with nucleases (DNase and RNase)**). Additionally, the sequential combination of both enzymes was assessed. The results obtained from this combined approach were identical to those observed when using TURBO™ DNase individually.

After undergoing the corresponding enzymatic treatment, each sample was injected into the 1 mL column containing the strong AEX resin Capto™ Q ImpRes to validate the impact of its use on the purification profile of MjsHSP nanocages. Regarding experimental conditions, the same parameters described earlier for the optimized AEX chromatography approach were applied, including the mobile phases (buffers A and B), the mass of enzymatically treated sample injected, and the binding/washing and elution procedures (sub-section **IV.3.1.2. Optimized purification by chromatography**).

The chromatograms obtained are presented in **Figure IV.7A** and **Figure IV.8A**, together with the corresponding SDS-PAGE images (**Figure IV.7B** and **Figure IV.8B**). All fractions were analyzed by SDS-PAGE and quantified for total protein content using the BCA method (sub-sections **IV.2.7.2. SDS-PAGE** and **IV.2.7.1. Total protein quantification**).

Generally, the chromatogram profile was largely similar between the two evaluated approaches and also comparable to the AEX chromatography of the non-treated sample (**Figure IV.4A**). A predominant flowthrough peak was observed, where no removal of MjsHSP nanocages was detected in either situation (**Figure IV.7A** and **B**, and **Figure IV.8A** and **B**, *1*). Peak *2* from the first elution step corresponded to the MjsHSP nanocages, as illustrated by SDS-PAGE (**Figure IV.7A** and **B**, and **Figure IV.8A** and **B**, *2*).

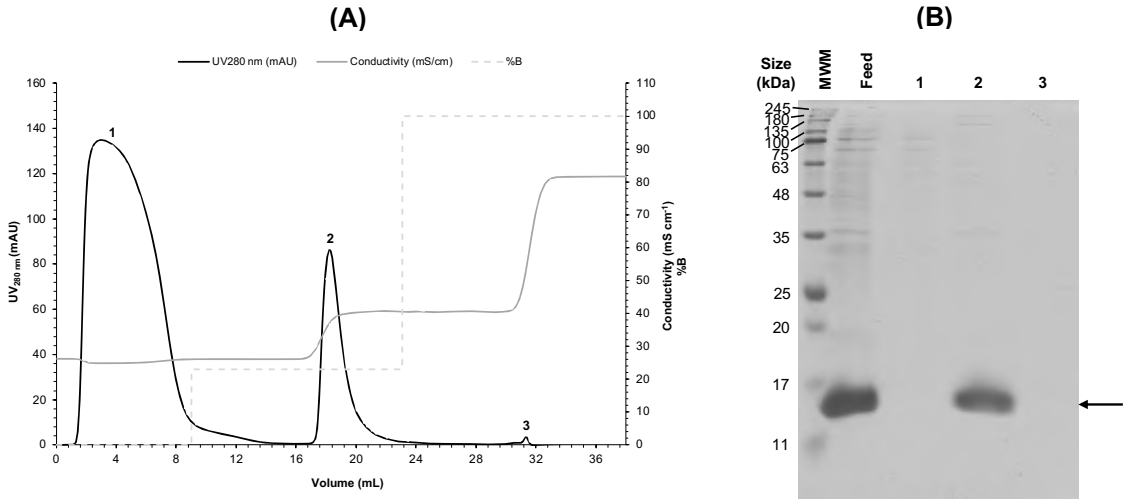


Figure IV.7 – Two-step approach for the purification of MjsHSP nanocages produced in *E. coli* BL21(DE3): pre-treatment with DNase followed by optimized AEX chromatography. **(A)** Chromatogram obtained using a 1 mL chromatographic column with the gigaporous resin Capto™ Q ImpRes (Cytiva; Marlborough, MA, USA). A feed stream (initial total protein mass of about 2 mg) previously subjected to an additional treatment with TURBO™ DNase (8 U) (Thermo Fisher Scientific; Waltham, MA, USA) was injected into this column earlier equilibrated with buffer A composed of 50 mM sodium phosphate, 250 mM NaCl, pH 6.0 ($\kappa \approx 26.1$ mS cm^{-1}). Unbound material was washed with this buffer A and elution was performed using a two-step gradient, the first step with 23% of buffer B ($\kappa \approx 40.5$ mS cm^{-1}) and the second step with 100% of this buffer B ($\kappa \approx 81.5$ mS cm^{-1}). Buffer B was composed of 50 mM sodium phosphate, 1 M NaCl, pH 6.0. **(B)** SDS-PAGE analysis of fractions collected during the chromatographic run (peaks *1*, *2*, and *3*) together with the clarified lysate feed sample. Peak *1* corresponded to the flowthrough and peaks *2* and *3* to the eluted material. Arrow indicates the protein bands referring to the monomer of MjsHSP nanocages (expected molecular weight of 16.5 kDa), which is separated and visualized under denaturing

conditions of SDS-PAGE. The abbreviation MWM refers to the molecular weight marker (NZYColour protein marker II; NZYtech Lisbon, Portugal).

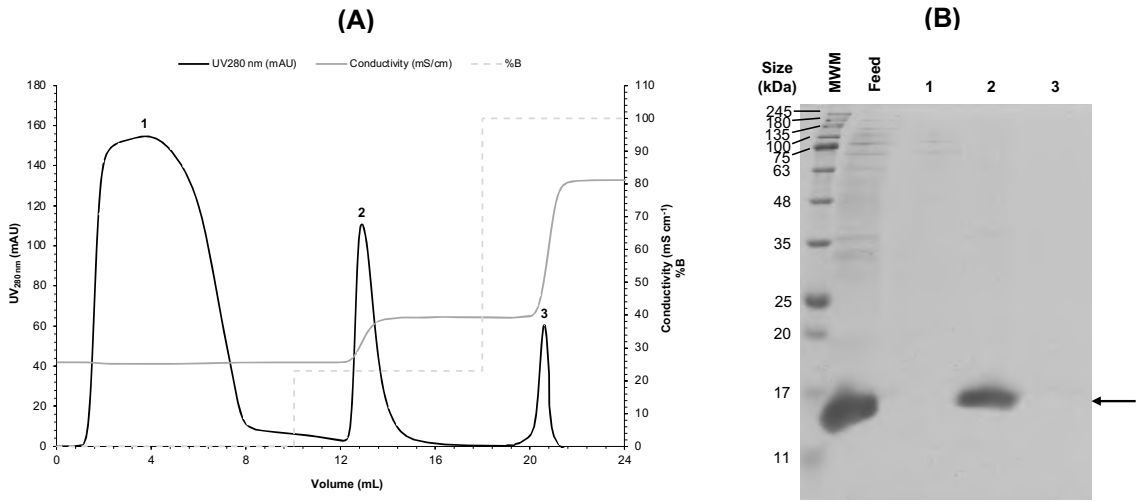


Figure IV.8 – Two-step approach for the purification of MjsHSP nanocages produced in *E. coli* BL21(DE3): pre-treatment with RNase followed by optimized AEX chromatography. **(A)** Chromatogram obtained using a 1 mL chromatographic column with the gigaporous resin Capto™ Q ImpRes (Cytiva; Marlborough, MA, USA). A feed stream (initial total protein mass of about 2 mg) previously subjected to an additional treatment with RNase A (4 U) (Roche; Basel, Switzerland) was injected into this column earlier equilibrated with buffer A composed of 50 mM sodium phosphate, 250 mM NaCl, pH 6.0 ($\kappa \approx 26.1$ mS cm^{-1}). Unbound material was washed with this buffer A and elution was performed using a two-step gradient, the first step with 23% of buffer B ($\kappa \approx 40.5$ mS cm^{-1}) and the second step with 100% of this buffer B ($\kappa \approx 81.5$ mS cm^{-1}). Buffer B was composed of 50 mM sodium phosphate, 1 M NaCl, pH 6.0. **(B)** SDS-PAGE analysis of fractions collected during the chromatographic run (peaks *1*, *2*, and *3*) together with the clarified lysate feed sample. Peak *1* corresponded to the flowthrough and peaks *2* and *3* to the eluted material. Arrow indicates the protein bands referring to the monomer of MjsHSP nanocages (expected molecular weight of 16.5 kDa), which is separated and visualized under denaturing conditions of SDS-PAGE. The abbreviation MWM refers to the molecular weight marker (NZYColour protein marker II; NZYtech Lisbon, Portugal).

In the DNase-treated experiment, a protein nanocages mass of 835.6 ± 10.0 µg was obtained, with a recovery yield of 151.9 ± 1.8 µg mL^{-1} and a recovery efficiency of 63.3 ± 3.3%, representing a 1.16-fold decrease in recovery efficiency compared to the non-treated sample. Host protein impurities removal was approximately 88.0 ± 1.0%. The chromatogram profile for the RNase-treated experiment showed results similar to the DNase test. The MjsHSP nanocages mass was 701.4 ± 65.8 µg, with a recovery yield of 140.3 ± 13.2 µg mL^{-1} and a recovery efficiency of 62.9 ± 0.9%, representing a 0.86-fold decrease in recovery efficiency compared to the non-treated sample. Host protein impurities removal was 84.0 ± 0.9%.

Peak *3* from the second elution step likely corresponded to nucleic acid impurities, consistent with the absence of protein bands in SDS-PAGE (**Figure IV.7A** and **B**, and **Figure IV.8A** and **B**, *3*). This final peak exhibited significant differences across the experimental tests conducted. In the case of the DNase-treated experiment, the peak decreased significantly, almost disappearing, which suggests that

this peak was clearly composed of dsDNA from the clarified lysate. On the other hand, in the case of RNase treatment, no considerable difference was observed between this assay and the non-treated sample, indicating that probably the amount of RNA in the clarified lysate was not significant.

Therefore, a prior enzymatic treatment with DNase could be an interesting approach to apply in the optimized downstream processing of MjsHSP nanocages (single-step AEX chromatography). Additional studies to identify the optimal conditions for enzymatic digestion with DNase could be conducted to achieve a consistent and highly efficient removal of host-derived dsDNA from the clarified lysates containing the MjsHSP nanocages.

This part of the work highlights the importance of conducting further research on this topic, with a particular focus on quantifying host-derived dsDNA throughout the purification process. Such quantification would allow for a more accurate assessment of its variation and the effectiveness and efficiency of each downstream processing operation unit. Implementing methods like qPCR for nucleic acid quantification could provide valuable insights.

IV.3.2. Primary isolation and purification of MjsHSP nanocages produced in *V. natriegens* cells

IV.3.2.1. Optimized purification by chromatography

Since the production of MjsHSP nanocages in *V. natriegens* Vmax™ X2 represents a novel aspect of this work, there is no existing data in the literature that could serve as a starting point for implementing a downstream processing strategy. Thus, the optimized approach previously developed for the purification of this type of NVPNs from clarified lysates of *E. coli* BL21(DE3) was adapted for the equivalent MjsHSP nanocages expressed in *V. natriegens*.

First, this optimized strategy involved the use of the strong AEX chromatography resin Capto™ Q ImpRes and the previously defined stepwise elution gradient.

Regarding the experimental conditions, the mobile phases used were a buffer A composed of 50 mM sodium phosphate, 250 mM NaCl, pH 6.0 and a buffer B consists of 50 mM sodium phosphate, 1 M NaCl, pH 6.0. The column was previously equilibrated with buffer A ($\kappa \approx 26.12$mS cm^{-1}). The sample injected into the column comprised approximately 2 mg of total protein from the clarified lysate containing the MjsHSP nanocages. After washing the unbound material with buffer A, a two-step gradient for the elution of bound species was applied, the first step with 23% of buffer B ($\kappa \approx 40.5$ mS cm^{-1}) and the second step with 100% of this buffer B ($\kappa \approx 81.5$ mS cm^{-1}).

The chromatogram obtained is presented in **Figure IV.9A**, along with the resulting SDS-PAGE image (**Figure IV.9B**). As before, the fractions from the chromatographic run were analyzed by SDS-PAGE and quantified for total protein content using the BCA method (sub-sections **IV.2.7.2. SDS-PAGE** and **IV.2.7.1. Total protein quantification**).

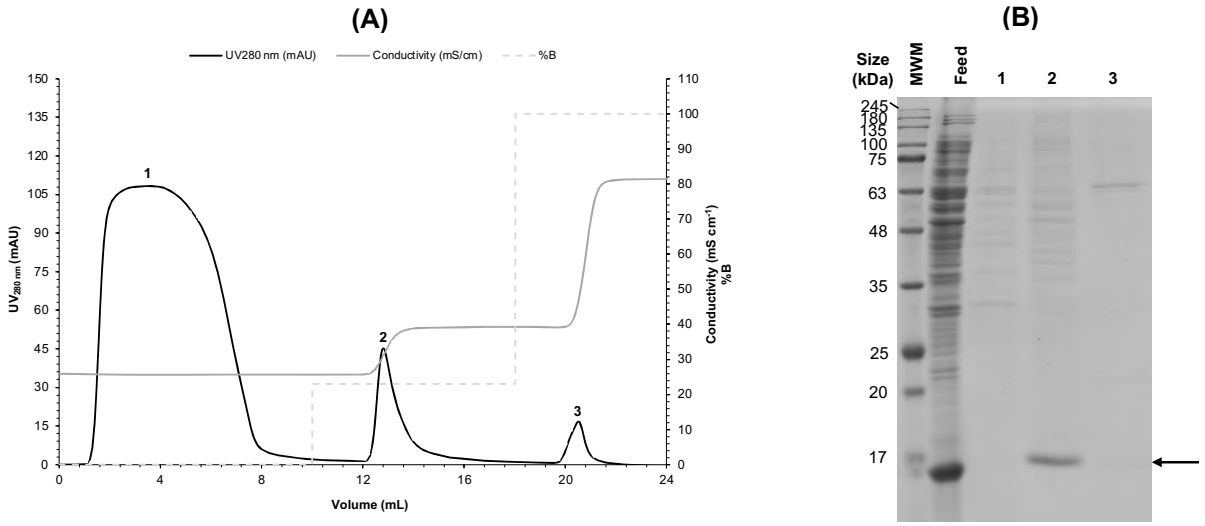


Figure IV.9 – Optimized AEX chromatography strategy for the purification of MjsHSP nanocages produced in *V. natriegens* Vmax™ X2. **(A)** Chromatogram obtained using a 1 mL chromatographic column with the gigaporous resin Capto™ Q ImpRes (Cytiva; Marlborough, MA, USA). A feed stream corresponding to approximately 2 mg of total protein was injected into this column previously equilibrated with buffer A composed of 50 mM sodium phosphate, 250 mM NaCl, pH 6.0 ($\kappa \approx 26.1$ mS cm^{-1}). Unbound material was washed with this buffer A and elution was performed using a two-step gradient, the first step with 23% of buffer B ($\kappa \approx 40.5$ mS cm^{-1}) and the second step with 100% of this buffer B ($\kappa \approx 81.5$ mS cm^{-1}). Buffer B was composed of 50 mM sodium phosphate, 1 M NaCl, pH 6.0. **(B)** SDS-PAGE analysis of fractions collected during the chromatographic run (peaks *1*, *2*, and *3*) together with the clarified lysate feed sample. Peak *1* corresponded to the flowthrough and peaks *2* and *3* to the eluted material. Arrow indicates the protein bands referring to the monomer of MjsHSP nanocages (expected molecular weight of 16.5 kDa), which is separated and visualized under denaturing conditions of SDS-PAGE. The abbreviation MWM refers to the molecular weight marker (NZYColour protein marker II; NZYtech Lisbon, Portugal).

Based on the results obtained, a chromatographic profile very similar to that achieved with the clarified lysate from *E. coli* was observed, which was an interesting outcome. This profile consisted of a major flowthrough peak, a peak in the first elution step corresponding to the fractions containing the MjsHSP nanocages, and a second elution step associated with host impurities, namely nucleic acids (dsDNA). The feed sample contained approximately 325.9 ± 4.4 µg of MjsHSP nanocages. In the flowthrough (**Figure IV.9A**, *1*), no protein nanocages were observed to be washed. On the contrary, SDS-PAGE analysis revealed the removal of protein impurities, with several bands of varying molecular weights being detected. Peak *2* of the NVPNs exhibited a sharper profile with some tailing, presenting a retention volume of approximately 13 mL at a conductivity of 31.8 mS cm^{-1} (**Figure IV.9A**, *2*). The SDS-PAGE of peak *2* showed a well-defined band of higher intensity corresponding to the monomers of the MjsHSP nanocages, along with very weak bands that may represent residual host protein impurities (**Figure IV.9B**, *2*). In direct comparison with the results from *E. coli*, more host impurities appeared to co-elute with the MjsHSP nanocages. Additionally, the intensity of the band corresponding to the protein nanocages was significantly lower than in the samples from *E. coli* BL21(DE3). For the peak *3*, very weak protein bands were observed in the SDS-PAGE, which was a difference compared to *E. coli* (**Figure IV.9A** and **B**, *3*). However, the possibility that some nucleic acids, particularly dsDNA,

were being retained on the resin and only eluted when the NaCl concentration increased to a conductivity of 45.7 mS cm^{-1} can still be considered.

The fractions of peak *2* containing the MjsHSP nanocages (**Figure IV.9A** and **B**) yielded a mass of 30.8 ± 3.7 µg of these NVPNs, corresponding to a recovery yield of 7.7 ± 0.9 µg mL^{-1} and subsequent recovery efficiency of 9.5 ± 1.3% relative to the feed with the clarified lysate. These values represent a very significant reduction of approximately 7.77-fold compared to the results obtained from the purification of MjsHSP nanocages produced in *E. coli* BL21(DE3).

To evaluate the complete process with all three purification steps, the MjsHSP nanocages resulting from AEX chromatography were concentrated by centrifugal filtration. This was performed using commercial centrifugal filters with a 100 kDa molecular weight cut-off (experimental procedure available in sub-section **IV.3.1.1. Standard purification by two-step chromatography** of **IV.3.1. Primary isolation and purification of MjsHSP nanocages produced in *E. coli* cells**). It was possible to obtain a mass of MjsHSP nanocages of 20.7 ± 1.0 µg, corresponding to 82.9 ± 4.0 µg mL^{-1} and a recovery efficiency of 68.8 ± 4.1% at the end of this unit operation. As expected, the values for these parameters were significantly lower than those recorded for the *E. coli* experiments.

Subsequently, these concentrated MjsHSP nanocages were subjected to the second step of chromatography-based purification using SEC (experimental methodology described in sub-section **IV.3.1.1. Standard purification by two-step chromatography** of **IV.3.1. Primary isolation and purification of MjsHSP nanocages produced in *E. coli* cells**). The results obtained, both in terms of the chromatogram and the corresponding SDS-PAGE, are shown in **Figure IV.10A** and **B**, respectively.

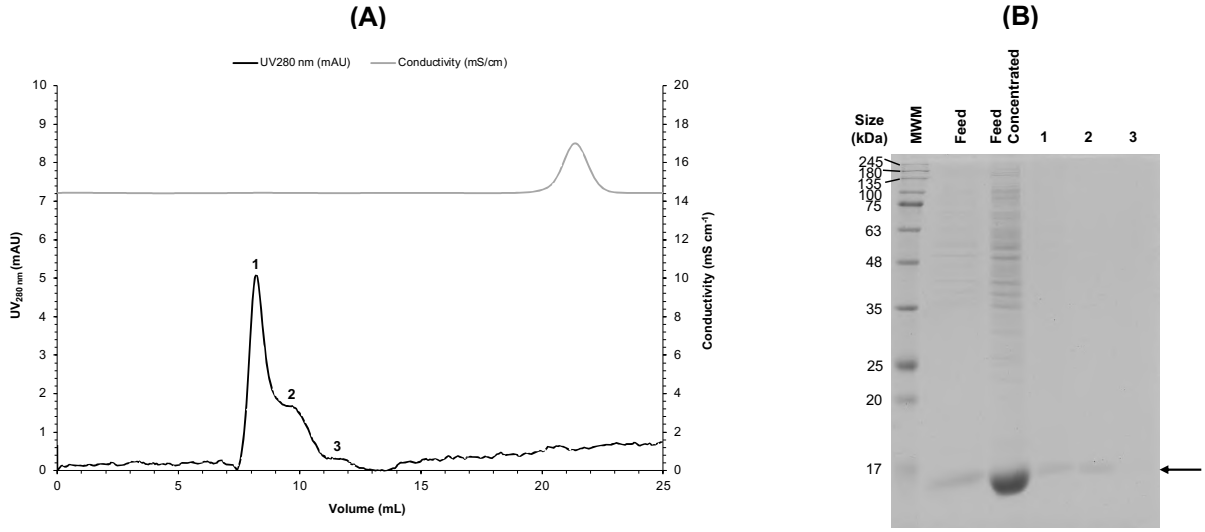


Figure IV.10 – SEC as the second step for the optimized purification of MjsHSP nanocages produced in *V. natriegens* Vmax™ X2. **(A)** Chromatogram obtained using a 24 mL size exclusion resin Superdex™ 200 (Cytiva; Marlborough, MA, USA) packed in an appropriate chromatographic column. The feed stream was comprised by the previously concentrated MjsHSP nanocages (from optimized AEX chromatography). The column was equilibrated with a PBS solution (1x) ($\kappa \approx 14.2$ mS cm^{-1}) at a flow rate of 1 mL min^{-1}. **(B)** SDS-PAGE analysis of fractions collected during the chromatographic run (peaks *1*, *2*, and *3*) together with the original and concentrated protein nanocages feed samples. Arrow indicates the protein bands referring to the monomer of MjsHSP nanocages (expected molecular weight of 16.5 kDa), which is separated and visualized under denaturing conditions of SDS-

PAGE. The abbreviation MWM refers to the molecular weight marker (NZYColour protein marker II; NZYtech Lisbon, Portugal).

The results obtained in terms of the chromatographic profile for SEC were somewhat different from those observed in the equivalent experiments with the MjsHSP nanocages produced in *E. coli*, with the appearance of three peaks, the most prominent being peak *1* (**Figure IV.10A**). As previously discussed, the possible presence of aggregates of MjsHSP nanocages and/or incomplete or unexpected assembled nanostructures may be contributing to this behavior. In terms of SDS-PAGE analysis (**Figure IV.10B**), bands corresponding to the monomers of MjsHSP nanocages were also observed in peaks *1* and 2. In both cases, these bands displayed significantly lower intensity compared to those seen in *E. coli*, suggesting that the purification yield was also lower. The absence of bands associated with protein impurities in the eluted fractions was also observed, which may be related to the fact that SEC contributes to the dilution of the injected sample, something particularly critical when the components of that sample are present in low quantities.

Overall, if the reduced content of MjsHSP nanocages from peaks *1* and *2* was considered (**Figure IV.10A** and **B**), a mass of 5.1 ± 0.6 µg of MjsHSP nanocages, a recovery yield of 3.4 ± 0.4 µg mL^{-1} and a recovery efficiency of 30.5 ± 5.7% were obtained.

Considering this optimized purification process composed of three steps, the global recovery efficiency was 2.0 ± 1.9%, with a host protein impurities removal efficiency of 98.7 ± 0.7%. This corresponded to a 4.75-fold decrease in the recovery efficiency and a similar impurities removal efficiency compared to the optimized protocol for purification of MjsHSP nanocages from *E. coli*.

Particularly in terms of the protein nanocages recovery, about 2.0% was an extremely low value and incompatible with potential downstream processing on a large scale. When considering only the AEX chromatography step, it was obtained a recovery efficiency of 9.5 ± 1.1%, and an impurities removal efficiency of 96.8 ± 1.3%. This represents a better value for the first parameter, although slightly lower for the second. If the AEX chromatography step followed by centrifugal filtration concentration is considered, the protein nanocages recovery decreased to 6.5 ± 1.0%, but the removal of host protein impurities increased to 98.6 ± 1.9%.

Therefore, these calculations suggest that the procedure using the AEX chromatography step alone may be the best option to achieve a good balance between the amount of protein nanocages recovered and the quantity of host protein impurities removed. Even considering the 9.5% of MjsHSP nanocages recovery efficiency, this value remained extremely low, suggesting that additional research could be conducted to identify possible improvements. These may include optimizing binding/elution conditions, pretreating samples with DNase and/or RNase as explored for *E. coli* lysates and/or implementing thermal pretreatment of the samples. Such adjustments could aim to enhance protein nanocages recovery while ensuring efficient removal of impurities from the clarified lysate, making the process more suitable for scalable applications.

IV.3.3. Primary isolation and purification of TRAP O-rings

IV.3.3.1. Standard purification by chromatography

Similarly to the MjsHSP nanocages, TRAP O-rings, precursors of TRAP nanocages and considered monomeric nanostructures, have been purified using a two-step chromatographic process combining AEX chromatography followed by SEC, with an intermediate step for protein nanocages concentration. Considering its description in the literature, this approach was chosen as the starting point for developing a more robust, reliable, and optimized purification process for TRAP O-rings.

The first step involved AEX chromatography, where the traditional agarose resin HiTrap Q FF (1 mL column) already used in the standard purification experiments for MjsHSP nanocages, was selected. As previously mentioned, this strong AEX resin facilitates electrostatic interactions with solutes through the charged nitrogen of its quaternary amine ligands. It features a conventional composition with a large particle size (90 µm) and small pore size (50 to 60 nm).

Regarding the experimental conditions, the mobile phases used consisted of buffer A (50 mM Tris-HCl, 50 mM NaCl, pH 7.9) and buffer B (50 mM Tris-HCl, 1 M NaCl, pH 7.9). The AEX column was equilibrated with buffer A ($\kappa \approx 7.2$ mS cm^{-1}). The sample injected into the HiTrap Q FF column consisted of approximately 1 mL of clarified lysate from *E. coli* BL21(DE3) cells containing TRAP O-rings (with an approximate total protein concentration of 2 mg mL^{-1}), mixed with 4 mL of buffer A. For this standard protocol, unbound material was washed with buffer A, and elution was performed using a linear gradient (20 CV) with increasing salt concentration up to 100% of buffer B ($\kappa \approx 80.2$ mS cm^{-1}).

The chromatogram obtained is presented in **Figure IV.11A**, along with the corresponding SDS-PAGE images (**Figure IV.11B**). All fractions were analyzed by SDS-PAGE as well as quantified for total protein content using the BCA method and for total dsDNA using qPCR (sub-sections **IV.2.7.2. SDS-PAGE**, **IV.2.7.1. Total protein quantification** and **IV.2.7.9. Total dsDNA quantification by qPCR**).

The chromatographic profile obtained was characterized by a significant flowthrough peak and four main peaks reasonably well-separated during the elution stage (**Figure IV.11A**). The feed sample showed a substantial amount of protein impurities, as evidenced by the identification of bands with different molecular weights in addition to the TRAP O-rings (**Figure IV.11B**). In SDS-PAGE, these O-rings were represented by the band corresponding to their respective monomers, with a molecular weight of approximately 8.3 kDa. The estimated mass of TRAP O-rings in this feed sample was 1204.3 ± 60.3 µg. In the flowthrough (**Figure IV.11A**, *1*), the elution of impurities with low charge density occurred, including host proteins and nucleic acids, as evidenced by the recorded peak. SDS-PAGE analysis of this fraction revealed a weak band corresponding to the TRAP O-rings monomers and virtually no other protein impurities (**Figure IV.11B**, *1*). This latter observation is somewhat unexpected, given the signal detected for the flowthrough at $UV_{280\ nm}$. A possible explanation could be that the flowthrough consisted of multiple 1 mL samples combined into a total volume close to 8 mL, which may have caused dilution of the components, rendering them undetectable in SDS-PAGE. Additionally, other types of impurities, such as nucleic acids and endotoxins, may have been washed from the column but they were not detectable by this analytical technique. Regarding the TRAP O-rings that did not bind to the resin, their estimated mass was 39.1 ± 3.5 µg. The analysis of the other collected fractions confirmed

that the peak eluting at approximately 25.5% buffer B ($\kappa \approx 17.9$ mS cm^{-1}) with a retention volume of about 17.6 mL contained the TRAP O-rings (**Figure IV.11A**, *2*). However, in this peak, other host protein impurities with different molecular weights were co-eluted, as observed in the SDS-PAGE analysis (**Figure IV.11B**, *2*). In the remaining elution peaks, neither TRAP O-rings nor other host protein impurities were detected. On the other hand, qPCR analysis of the fractions corresponding to peaks *4* and *5* revealed the presence of approximately 966.691 ± 22.378 µg of dsDNA eluted at higher NaCl concentrations ($\kappa \approx 57$ to 67 mS cm^{-1}) (**Figure IV.11A**, *4* and *5*). This result was expected, as previously mentioned in studies with MjsHSP nanocages, where dsDNA exhibited higher charge density compared to other impurities and the target proteins themselves. Considering the pI of the TRAP O-rings (6.49), using a pH of 7.9 ensured that these protein O-rings were negatively charged. Protein impurities with lower charge density were eluted from the resin, while TRAP O-rings remained bound until elution at higher NaCl concentrations. However, dsDNA presented a challenge due to its strong negative charge at this pH, which competed with the protein O-rings for binding to the functional groups of the resin.

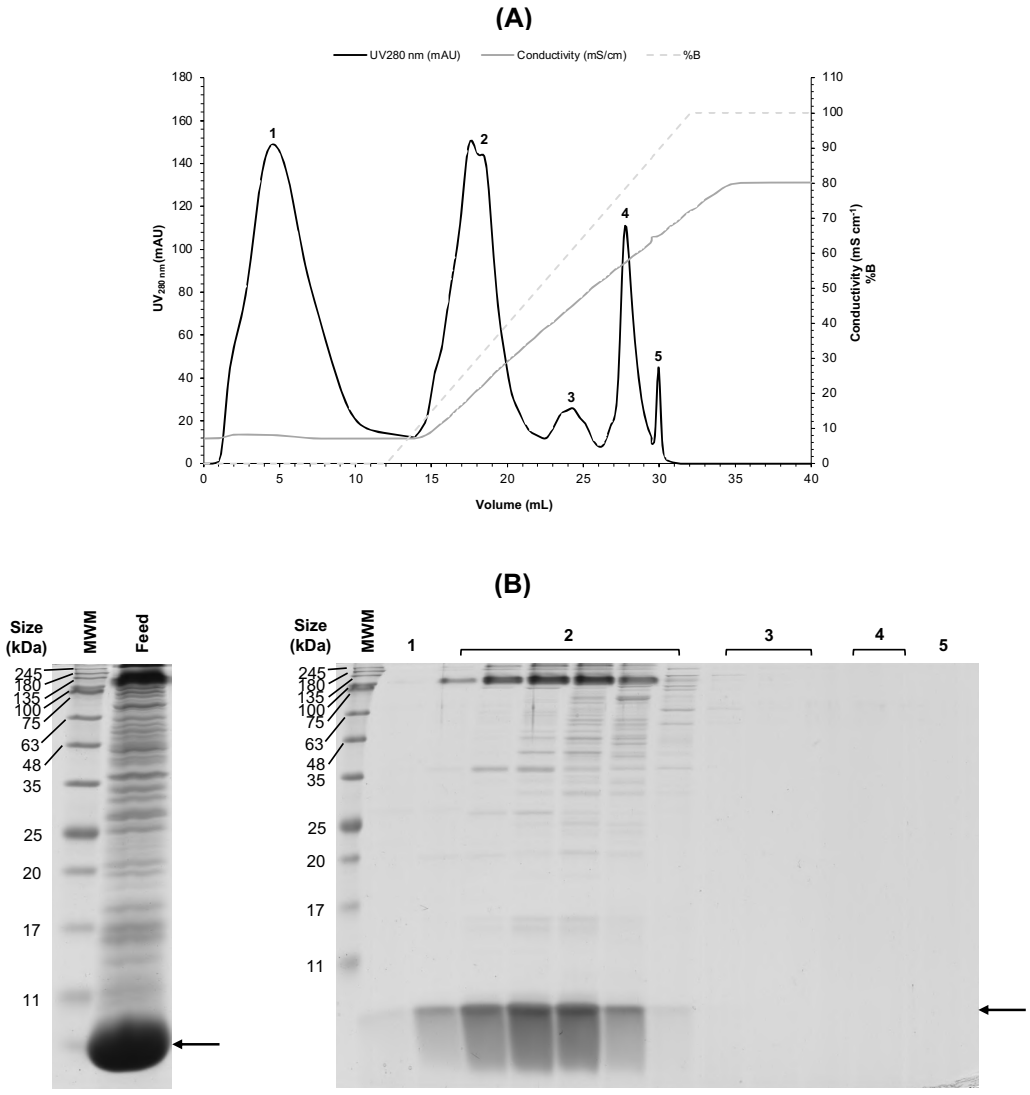


Figure IV.11 – AEX chromatography as the first step for the standard purification of TRAP O-rings produced in *E. coli* BL21(DE3). **(A)** Chromatogram obtained using a 1 mL pre-packed column with the traditional agarose resin

HiTrap Q FF (Cytiva; Marlborough, MA, USA). A feed stream comprising 1 mL of the clarified lysate containing the TRAP O-rings together with 4 mL of buffer A was injected into this column previously equilibrated with this buffer A composed of 50 mM Tris-HCl, 50 mM NaCl, pH 7.9 ($\kappa \approx 7.2$ mS cm^{-1}). Unbound material was washed with this buffer A and elution was performed using a linear gradient (20 CV) with increasing salt concentration up to 100% of buffer B composed of 50 mM Tris-HCl, 1 M NaCl, pH 7.9 ($\kappa \approx 80.2$ mS cm^{-1}). **(B)** SDS-PAGE analysis of fractions collected during the chromatographic run (peaks *1*, *2*, *3*, *4*, and *5*) together with the clarified lysate feed sample. Peak *1* corresponded to the flowthrough and peaks *2* to *5* to the eluted material. Arrow indicates the protein bands referring to the monomer of TRAP O-rings (expected molecular weight of 8.3 kDa), which is separated and visualized under denaturing conditions of SDS-PAGE. The abbreviation MWM refers to the molecular weight marker (NZYColour protein marker II; NZYtech Lisbon, Portugal).

The fractions from peak *2* containing the TRAP O-rings (**Figure IV.11A** and **B**) yielded a total mass of 998.6 ± 4.1 µg of these target proteins, corresponding to a recovery yield of 124.8 ± 0.5 µg mL^{-1} and a recovery efficiency of 83.1 ± 3.8% relative to the feed from the clarified lysate.

Afterwards, the fractions containing TRAP O-rings were concentrated using centrifugal filtration with commercial centrifugal filters featuring a 30 kDa molecular weight cut-off. As discussed earlier in this work, this methodology was designed to address a limitation of SEC, specifically the requirement for a small sample injection volume (around 1% of the total resin volume). The selection of a 30 kDa molecular weight cut-off also served a purification purpose, as it allowed the removal of TRAP O-rings monomers (estimated molecular weight of approximately 8.3 kDa). Using this approach, a mass of TRAP O-rings amounting to 360.1 ± 6.3 µg was obtained, corresponding to a recovery yield of 1200.3 ± 21.2 µg mL^{-1} at the end of this unit operation. The recovery efficiency was 36.1 ± 0.5%, despite implementing measures such as membrane passivation with 5% v/v Tween 20 and buffer equilibration to minimize protein nanocage loss by reducing non-specific interactions.

Next, the sample containing the concentrated TRAP O-rings was subjected to the second step of chromatography-based purification using SEC. Similar to the MjsHSP nanocages, the Superdex™ 200 resin (24 mL column) was selected for this process. The mobile phase used was a buffer composed of 50 mM Tris-HCl, 150 mM NaCl, pH 7.9. The resin was equilibrated with this buffer at a flow rate of 1 mL min^{-1}. After injecting the sample into the column, it was washed with the same buffer solution while maintaining a constant conductivity value ($\kappa \approx 16.4$ mS cm^{-1}).

The chromatogram obtained is shown in **Figure IV.12A**, being characterized by two distinct peaks (*2* and *3*), along with a pronounced shoulder associated with peak *2* (*1*). The corresponding fractions were analyzed by SDS-PAGE in comparison with the feed sample. The resulting gel is shown in **Figure IV.12B**. All these fractions were similarly analyzed using the BCA method for total protein quantification and using qPCR for total dsDNA quantification.

The feed sample displayed an intense band corresponding to the monomers of TRAP O-rings, along with other protein bands of higher molecular weight with varying intensities (**Figure IV.12B**). In peak *2*, a significantly stronger signal was observed, corresponding to these nanocages and some host protein impurities, which also extended to shoulder *1* (**Figure IV.12A** and **B**, *2*). This result seems reasonable due to the overlap in content between peak *2* and shoulder *1*. The presence of this shoulder may indicate that part of the nanostructures had dimensions larger than those washed from the column

in peak *2* (retention volume of about 13 mL), such as aggregates or defective structures resulting from incorrect self-assembly. In peak *3* (**Figure IV.12A** and **B**, *3*), no bands corresponding to TRAP O-rings or protein impurities were detected. It was likely that the SDS-PAGE technique as not entirely reliable in this context, as SEC diluted the samples, and therefore some host protein impurities presented at low concentrations may not be detected by this method, even with the use of the more sensitive silver staining technique.

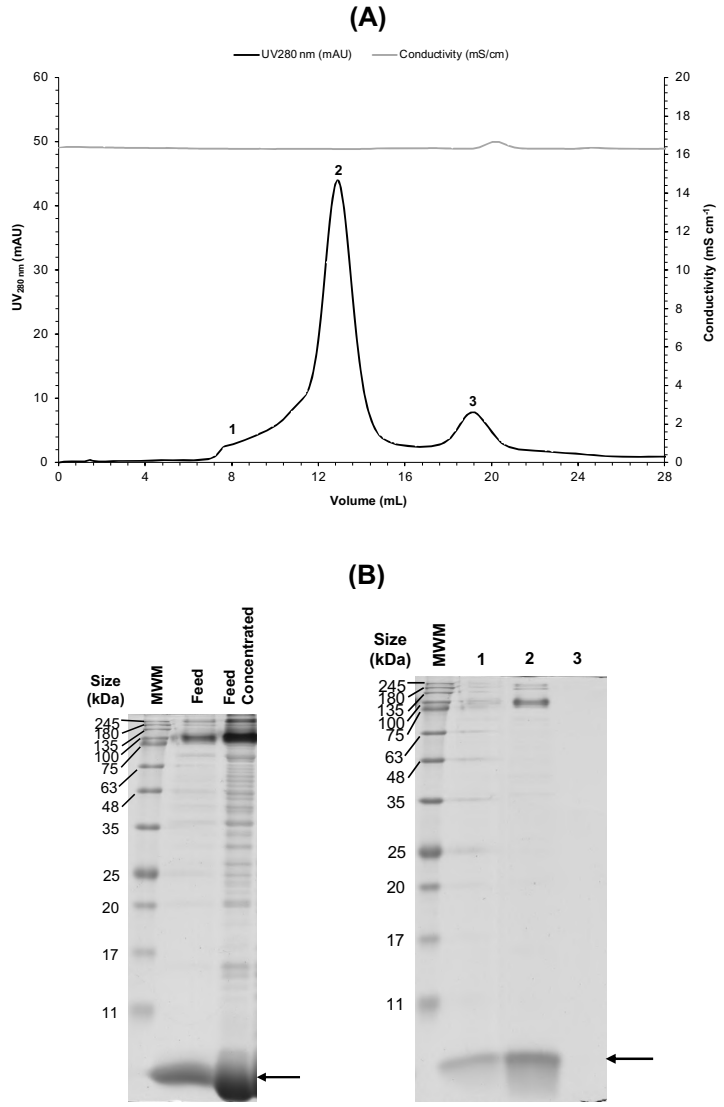


Figure IV.12 – SEC as the second step for the standard purification of TRAP O-rings produced in *E. coli* BL21(DE3). **(A)** Chromatogram obtained using a 24 mL size exclusion resin Superdex™ 200 (Cytiva; Marlborough, MA, USA) packed in an appropriate chromatographic column. The feed stream was comprised by the previously concentrated MjsHSP nanocages. The column was equilibrated with 50 mM Tris-HCl, 150 mM NaCl, pH 7.9 buffer ($\kappa \approx 16.4$ mS cm^{-1}) at a flow rate of 1 mL min^{-1}. **(B)** SDS-PAGE analysis of fractions collected during the chromatographic run (peaks *1*, *2*, and *3*) together with the original and concentrated protein nanocages feed samples. Arrow indicates the protein bands referring to the monomer of TRAP O-rings (expected molecular weight of 8.3 kDa), which is separated and visualized under denaturing conditions of SDS-PAGE. The abbreviation MWM refers to the molecular weight marker (NZYColour protein marker II; NZYtech Lisbon, Portugal).

Overall, when considering the content of TRAP O-rings from peak *2* and shoulder *1*, a mass of 172.4 ± 4.3 µg of protein O-rings was obtained, corresponding to a recovery yield of 48.6 ± 1.2 µg mL^{-1} and a recovery efficiency of 57.9 ± 2.1% (**Figure IV.12A** and **B**).

Finally, considering this standard purification process composed of three steps, the global recovery efficiency of TRAP O-rings was 17.3 ± 0.6%, while the removal efficiency of host proteins and dsDNA impurities was 82.4 ± 1.9% and 83.9 ± 0.4%, respectively. These results highlight an opportunity to further explore and optimize the downstream processing.

IV.3.3.2. Alternative and optimized purification by chromatography

Following the establishment of the standard methodology for purifying TRAP O-rings, a more detailed study was initiated to optimize the process. The primary objective was to enhance the recovery of protein O-rings while maximizing the removal of contaminants.

Preliminarily, an alternative strong AEX chromatographic support to the conventional HiTrap Q FF resin (Capto™ Q ImpRes) was evaluated. However, after obtaining equivalent results, it was decided to explore a different chromatographic approach that could potentially deliver better performance by exploiting distinct chemical interactions between the resin and the sample components. This represents one of the novelties of this work in the context of downstream processing for TRAP O-rings.

TRAP O-rings are stable at low pH values (until around 5.0), which led to the selection of a buffer with a pH of 5.5. This ensured that the TRAP O-rings carried a positive charge due to their pI of 6.49, while at this same pH, most protein impurities, particularly dsDNA, were negatively or highly negatively charged. Thus, the choice of a CEX chromatography resin was made to ensure that these negatively charged impurities would not bind to the resin, as they share the same negative charge as the functional groups of the resin at this pH. In contrast, the positively charged TRAP O-rings would be retained on the resin and subsequently eluted by varying either the NaCl concentration or the pH.

For this research study, two strong CEX chromatographic supports (HiTrap Capto™ S and Capto™ SP ImpRes) were selected, and the impact of different compositions of three elution buffers (B) on the elution step was evaluated: 50 mM MES, 2 M NaCl, pH 5.5; 50 mM Tris-HCl, pH 8.0; and 50 mM Tris-HCl, 2 M NaCl, pH 8.0. The selection of these buffers was justified by the objective of assessing the effect of increasing NaCl concentration and pH on the chromatographic run.

Regarding the experimental conditions, the mobile phases used included buffer A (50 mM MES, pH 5.5) and one of the three buffer B solutions mentioned earlier. The CEX column was equilibrated with buffer A ($\kappa \approx 1.0$ mS cm^{-1}). The sample injected into the column consisted of approximately 1 mL of clarified lysate from *E. coli* BL21(DE3) cells containing TRAP O-rings (with an approximate total protein concentration of 2 mg mL^{-1}), mixed with 4 mL of buffer A. Unbound material was washed with buffer A, and elution was performed using a linear gradient (20 CV) up to 100% of one of the buffer B solutions ($\kappa \approx 2.9$ to 140.9 mS cm^{-1}).

Starting with the experiments using the HiTrap Capto™ S, the chromatograms obtained are presented in **Figure IV.13A**, **C** and **E**, along with the corresponding SDS-PAGE images (**Figure IV.13B**, **D** and **F**). SDS-PAGE analysis of the clarified lysate feed sample used in all these experiments is shown

in **Figure IV.11B**. All fractions were analyzed by SDS-PAGE as well as quantified for total protein content using the BCA method and for total dsDNA using qPCR (sub-sections **IV.2.7.2. SDS-PAGE**, **IV.2.7.1. Total protein quantification** and **IV.2.7.9. Total dsDNA quantification by qPCR**).

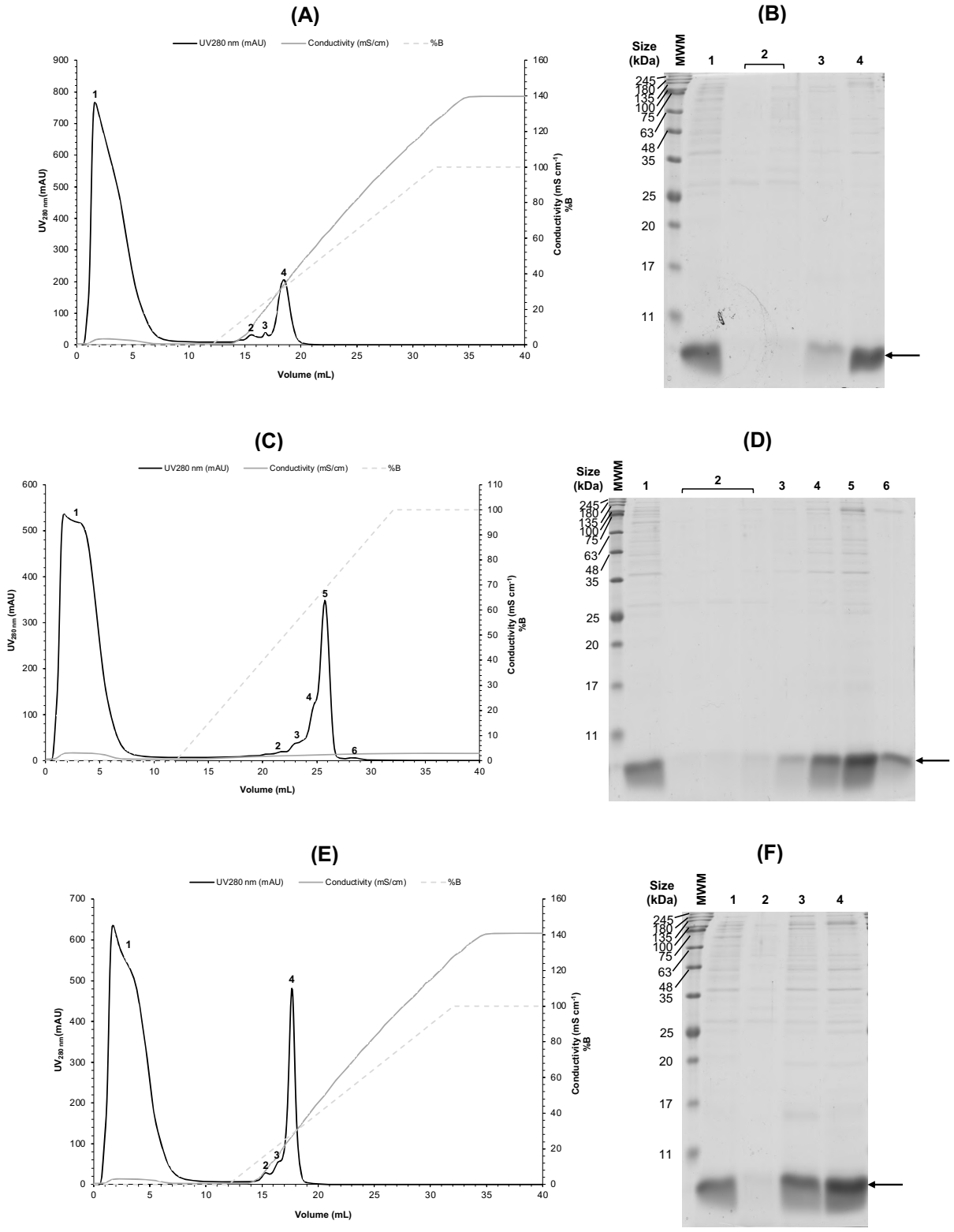


Figure IV.13 – CEX chromatography as an alternative for the purification of TRAP O-rings produced in *E. coli* BL21(DE3). Chromatograms obtained using the 1 mL HiTrap Capto™ S pre-packed column (Cytiva; Marlborough,

MA, USA). For all experiments, a feed stream corresponding to approximately 2 mg of total protein was injected into the column previously equilibrated with buffer A composed of 50 mM MES, pH 5.5 ($\kappa \approx 1.0$ mS cm^{-1}). Unbound material was washed with this buffer A and elution was performed using a linear gradient (20 CV) with increasing salt concentration up to 100% of buffer B composed of 50 mM MES, 2 M NaCl, pH 5.5 ($\kappa \approx 139.8$ mS cm^{-1}) **(A)**, 50 mM Tris-HCl, pH 8.0 ($\kappa \approx 2.9$ mS cm^{-1}) **(C)** and 50 mM Tris-HCl, 2 M NaCl, pH 8.0 ($\kappa \approx 140.9$ mS cm^{-1}) **(E)**. SDS-PAGE analysis of fractions collected during the chromatographic run for the three distinct compositions of the buffer B in terms of NaCl concentration and/or pH tested: 50 mM MES, 2 M NaCl, pH 5.5 **(B)**, 50 mM Tris-HCl, pH 8.0 **(D)** and 50 mM Tris-HCl, 2 M NaCl, pH 8.0 **(F)**. Peak *1* corresponded to the flowthrough and peaks *2* to *4* or *6* to the eluted material. SDS-PAGE analysis of the clarified lysate feed sample used in all these experiments is shown in **Figure IV.11B**. Arrow indicates the protein bands referring to the monomer of TRAP O-rings (expected molecular weight of 8.3 kDa), which is separated and visualized under denaturing conditions of SDS-PAGE. The abbreviation MWM refers to the molecular weight marker (NZYColour protein marker II; NZYtech Lisbon, Portugal).

Generally, the chromatogram profiles showed some differences between the three alternative elution buffer compositions, particularly between elution based on a significant increase in NaCl concentration and elution focused on increasing the system's pH. All chromatograms displayed a major flowthrough peak and three to five peaks during the elution phase. It is important to note that the feed sample was consistent across all experiments, showing bands of host protein impurities with varying molecular weights, in addition to the TRAP O-rings. In all flowthrough fractions (**Figure IV.13A**, **C** and **E**, *1*), a loss of TRAP O-rings that did not bind to the resin was observed. This may suggest that the binding capacity of the chromatographic support had been reached, not only due to the target protein O-rings but also because some impurities were interacting with the functional groups of the separation media. On the other hand, bands associated with protein impurities were also washed out, which was a desired outcome. Regarding the peaks eluting with the linear gradient, the most prominent peak corresponded to the TRAP O-rings (**Figure IV.13A** and **E**, *4*; **Figure IV.13C**, *5*), as confirmed by SDS-PAGE results (**Figure IV.13B**, **D** and **F**). For the two alternative elution buffer compositions 50 mM MES, 2 M NaCl, pH 5.5 and 50 mM Tris-HCl, 2 M NaCl, pH 8.0, the retention volume was approximately 18 mL when conductivity ranged between 30 and 35 mS cm^{-1}. However, when using 50 mM Tris-HCl, pH 8.0, the retention volume shifted to 26 mL with a conductivity of 2.3 mS cm^{-1}.

For the elution buffer composed of 50 mM MES, 2 M NaCl, pH 5.5, the fractions from peak *4* containing the TRAP O-rings (**Figure IV.13A** and **B**) yielded a mass of 576.1 ± 5.6 µg of these protein nanocages, corresponding to a recovery yield of 192.0 ± 1.9 µg mL^{-1} and a recovery efficiency of 49.9 ± 0.9% relative to the feed from the clarified lysate. Additionally, as visualized in the SDS-PAGE analysis, peak *1* resulted in the washing out of 882.8 ± 0.7 µg of TRAP O-rings, while peak *3* led to the elution of approximately 11.7 ± 0.7 µg of these target protein O-rings, which contributed to the reduction in recovery efficiency. In terms of the host nucleic acid impurities, peak *1* showed the removal of 678.3 ± 32.2 µg of dsDNA, along with co-elution with TRAP O-rings in peak *4*, where approximately 408.6 ± 20.1 µg of dsDNA was eluted. Nevertheless, in these experiments, a total removal of approximately 64.5 ± 0.8% of host dsDNA from the original feed sample was achieved.

For the elution buffer composed of 50 mM Tris-HCl, pH 8.0, the fractions from peaks *4* and *5* containing the TRAP O-rings (**Figure IV.13C** and **D**) yielded a mass of 566.8 ± 13.5 µg of these protein

nanocages, corresponding to a recovery yield of 161.9 ± 3.9 μg mL^{-1} and a recovery efficiency of 48.9 ± 0.5% relative to the feed from the clarified lysate. Additionally, as visualized in the SDS-PAGE analysis, peak *1* resulted in the washing out of 821.1 ± 12.3 μg of TRAP O-rings, while the remaining elution peaks (excluding peaks *4* and *5*) showed the elution of approximately 90.7 ± 3.7 μg of these target proteins, contributing to the reduction of the recovery efficiency. In peak *5*, the elution also included host protein impurities with higher molecular weights. In terms of host dsDNA impurities, the greatest removal was observed in peak *1* (approximately 607.6 ± 7.3 μg of dsDNA), followed by co-elution with TRAP O-rings in peaks *4* and *5* (approximately 480.4 ± 9.8 μg of dsDNA). In the remaining elution peaks, an additional 62.4 ± 4.6 μg of dsDNA was eluted. Nevertheless, in these experiments, an overall removal of approximately 58.3 ± 1.3% of dsDNA from the original feed sample was achieved, despite significant co-elution of TRAP O-rings and dsDNA during the process.

For the elution buffer composed of 50 mM Tris-HCl, 2 M NaCl, pH 8.0, the fractions from peak *4* containing the TRAP O-rings (**Figure IV.13E** and **F**) yielded a mass of 471.8 ± 28.0 μg of these protein nanocages, corresponding to a recovery yield of 235.9 ± 14.0 μg mL^{-1} and a recovery efficiency of 40.7 ± 1.9% relative to the feed from the clarified lysate. Moreover, as observed in the SDS-PAGE analysis, peak *1* resulted in the washing out of 769.7 ± 13.8 μg of TRAP O-rings, while the remaining elution peaks, particularly peak *3*, showed the elution of approximately 111.5 ± 6.9 μg of these target protein O-rings, contributing to the reduction in their recovery efficiency. Due to the overlap between peaks *3* and *4*, with peak *3* appearing as a small shoulder, this could explain the higher amount of TRAP O-rings being eluted. In this set of peaks (*3* and *4*), a reasonable amount of protein impurities with medium and high molecular weights was also eluted. In terms of host dsDNA impurities, the greatest removal was observed in peak *1* (approximately 578.0 ± 8.2 μg of dsDNA), followed by co-elution with TRAP O-rings in peak *4* (approximately 339.4 ± 11.9 μg of dsDNA). In the remaining elution fractions, around 12.4 ± 2.2 μg of dsDNA was removed. However, in these experiments, an overall removal of approximately 70.5 ± 1.8% of dsDNA from the original feed sample was achieved, despite substantial co-elution of both TRAP O-rings and dsDNA during the chromatographic run.

Therefore, the results demonstrated that the purification of TRAP O-rings using CEX chromatography with the HiTrap Capto™ S column and the elution buffer composed of 50 mM MES, 2 M NaCl, pH 5.5 allowed for a higher recovery efficiency of the protein O-rings at 49.9 ± 0.9%, compared to the two tested alternatives (approximately 1.0 and 1.2 times lower). This outcome was balanced with good removal efficiencies for host protein and dsDNA impurities, achieving 83.0 ± 0.9% and 64.5 ± 0.8%, respectively. In terms of elution buffer composition, increasing the NaCl concentration appeared to be the best strategy for the purification of TRAP O-rings.

Regarding the experiments with the use of Capto™ SP ImpRes, the chromatograms obtained are presented in **Figure IV.14A**, **C** and **E**, along with the corresponding SDS-PAGE images (**Figure IV.14B**, **D** and **F**). SDS-PAGE analysis of the clarified lysate feed sample used in all these experiments is shown in **Figure IV.11B**. All fractions were analyzed by SDS-PAGE as well as quantified for total protein content using the BCA method and for total dsDNA using qPCR (sub-sections **IV.2.7.2. SDS-PAGE**, **IV.2.7.1. Total protein quantification** and**IV.2.7.9. Total dsDNA quantification by qPCR**).

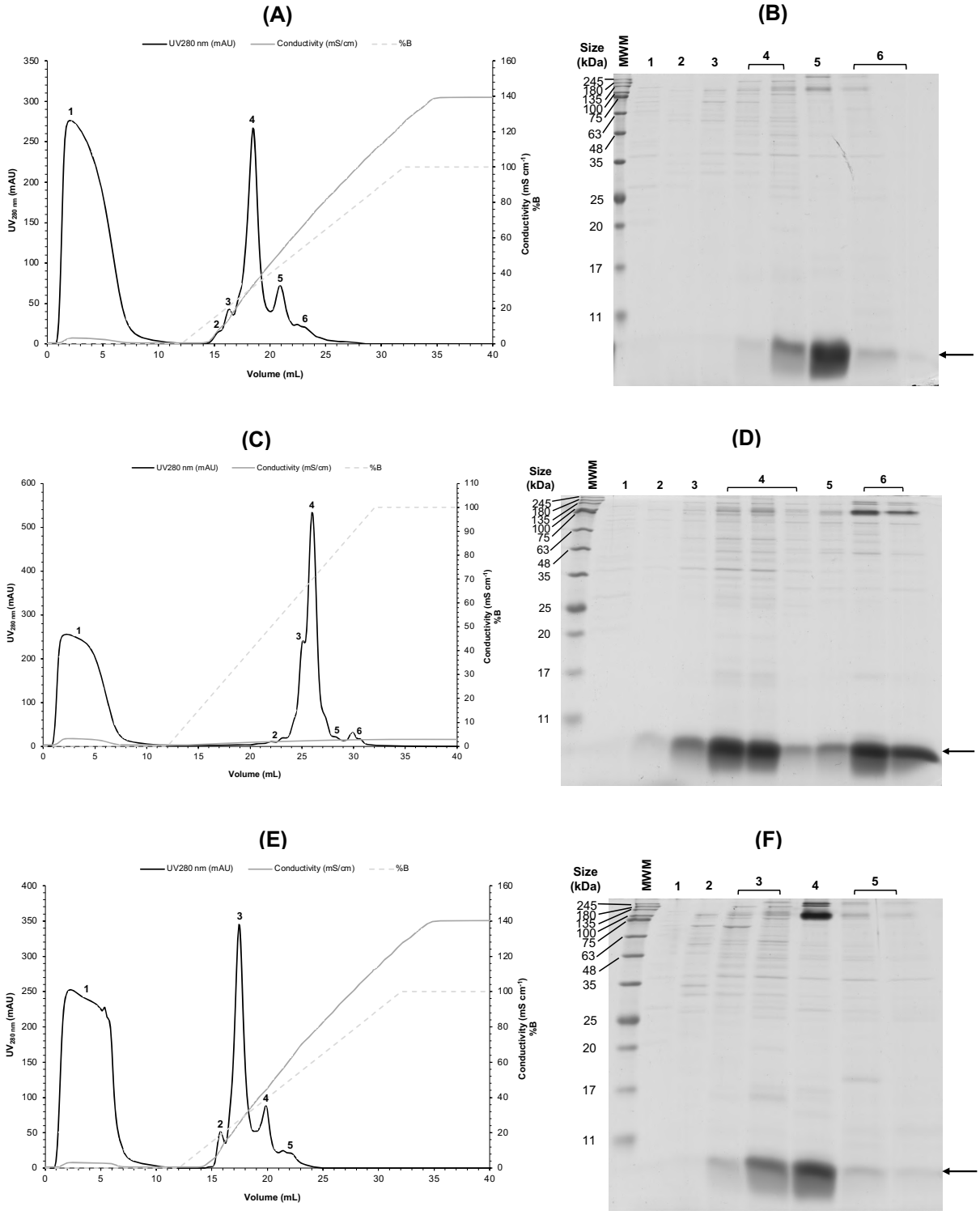


Figure IV.14 – CEX chromatography as an alternative for the purification of TRAP O-rings produced in *E. coli* BL21(DE3). Chromatograms obtained using a 1 mL chromatographic column with the resin Capto™ SP ImpRes (Cytiva; Marlborough, MA, USA). For all experiments, a feed stream corresponding to approximately 2 mg of total protein was injected into the column previously equilibrated with buffer A composed of 50 mM MES, pH 5.5 ($\kappa \approx 1.0$ mS cm^{-1}). Unbound material was washed with this buffer A and elution was performed using a linear gradient (20 CV) with increasing salt concentration up to 100% of buffer B composed of 50 mM MES, 2 M NaCl, pH 5.5 ($\kappa \approx$ 139.8 mS cm^{-1}) **(A)**, 50 mM Tris-HCl, pH 8.0 ($\kappa \approx 2.9$ mS cm^{-1}) **(C)** and 50 mM Tris-HCl, 2 M NaCl, pH 8.0 ($\kappa \approx$

140.9 mS cm^{-1}) **(E)**. SDS-PAGE analysis of fractions collected during the chromatographic run for the three distinct compositions of the buffer B in terms of NaCl concentration and/or pH tested: 50 mM MES, 2 M NaCl, pH 5.5 **(B)**, 50 mM Tris-HCl, pH 8.0 **(D)** and 50 mM Tris-HCl, 2 M NaCl, pH 8.0 **(F)**. Peak *1* corresponded to the flowthrough and peaks *2* to *4* or *6* to the eluted material. SDS-PAGE analysis of the clarified lysate feed sample used in all these experiments is shown in **Figure IV.11B**. Arrow indicates the protein bands referring to the monomer of TRAP O-rings (expected molecular weight of 8.3 kDa), which is separated and visualized under denaturing conditions of SDS-PAGE. The abbreviation MWM refers to the molecular weight marker (NZYColour protein marker II; NZYtech Lisbon, Portugal).

As with the previously tested CEX chromatography resin, the chromatogram profiles showed some differences between the three alternative elution buffer compositions, particularly between elution based on a significant increase in NaCl concentration and elution focused on increasing the system's pH. All chromatograms presented a major flowthrough peak and four to five peaks during the elution phase. Note that the feed sample was equivalent for all experiments, showing bands of host protein impurities with different molecular weights in addition to the TRAP O-rings. In all flowthrough fractions (**Figure IV.14A**, **C** and **E**, *1*), no significant loss of TRAP O-rings was observed, indicating that the binding capacity of the chromatographic support was not reached and that the majority of the protein O-rings remained retained in the resin. On the other hand, weak bands associated with protein impurities were also washed out, which was a desired result. Regarding the peaks eluting with the linear gradient, interestingly, the most prominent peak did not correspond in all experiments to the TRAP O-rings (**Figure IV.14A** and **C**, *4*; **Figure IV.14E**, *3*), as confirmed by SDS-PAGE results (**Figure IV.14B**, **D** and **F**). For the two alternative elution buffer compositions 50 mM MES, 2 M NaCl, pH 5.5 and 50 mM Tris-HCl, 2 M NaCl, pH 8.0, the retention volume was approximately 20.5 mL when conductivity ranged between 50 and 55 mS cm^{-1} (**Figure IV.14A** and **E**, *5*). This was the second most prominent peak, and there was a loss of TRAP O-rings before this point of elution. When using the elution buffer composed of 50 mM Tris-HCl, pH 8.0, the retention volume shifted to 26 mL with a conductivity of 2.3 mS cm^{-1} (**Figure IV.14C**, *4*). In this case, there was a loss during elution that occurred mainly afterwards, likely due to variations in the oligomeric structure of the TRAP O-rings, which may have altered their chemical interactions with the functional groups of the resin.

For the elution buffer composed of 50 mM MES, 2 M NaCl, pH 5.5, the fractions from peak *5* containing the TRAP O-rings (**Figure IV.14A** and **B**) yielded a mass of 847.0 ± 5.8 µg of these protein O-rings, corresponding to a recovery yield of 423.5 ± 2.9 µg mL^{-1} and a recovery efficiency of 73.7 ± 0.2% relative to the feed from the clarified lysate. Furthermore, as visualized in the SDS-PAGE analysis, peak *1* resulted in the washing out of 78.6 ± 3.6 µg of TRAP O-rings, while peak *4* primarily accounted for the elution of 294.8 ± 13.0 µg of these protein O-rings, and peak *6* showed a loss of 4.8 ± 2.4 µg. All these undesired losses of the target TRAP O-rings contributed to the decrease in recovery efficiency. In terms of host dsDNA impurities, peak *4* mainly showed the removal of 307.3 ± 3.7 µg of dsDNA, in addition to co-elution with TRAP O-rings in peak *5*, where approximately 73.9 ± 1.8 µg of dsDNA was eluted. Also, during elution, an additional loss of 178.4 ± 5.3 µg of dsDNA was observed. Nevertheless, in this experiment, an overall removal of approximately 93.6 ± 2.3% of dsDNA from the original feed sample was achieved.

For the elution buffer composed of 50 mM Tris-HCl, pH 8.0, the fractions from peak *4* containing the TRAP O-rings (**Figure IV.14C** and **D**) yielded a mass of 779.1 ± 22.2 µg of these protein O-rings, corresponding to a recovery yield of 222.6 ± 5.3 µg mL^{-1} and a recovery efficiency of 67.8 ± 1.6% relative to the feed from the clarified lysate. Meanwhile, the fractions related to peak *6* yielded a mass of 368.9 ± 3.1 µg of TRAP O-rings, corresponding to a recovery yield of 105.4 ± 0.9 µg mL^{-1} and a recovery efficiency of 32.1 ± 0.4%. Furthermore, as visualized in the SDS-PAGE analysis, peaks *3* and *5* showed the elution of 144.2 ± 16.7 µg of protein O-rings. However, what contributed most to the reduction in the recovery efficiency for peak *4*, which was indicated as the preferred elution point for TRAP O-rings under these conditions, was the loss of these target proteins in the last elution peak (*6*). In general, an undesirable behavior was observed, with the distribution of TRAP O-rings elution occurring at multiple points throughout the chromatography run, making it challenging to implement a defined and consistent process for purifying TRAP O-rings. In terms of host dsDNA impurities, the highest removal was observed in peak *4* (approximately 694.6 ± 5.8 µg of dsDNA), which co-eluted with TRAP O-ring. This was an unfavorable result as it demonstrated an inability to efficiently separate protein O-rings from the main contaminants such as dsDNA. In peak *6*, approximately 26.2 ± 4.8 µg of dsDNA was eluted, while other peaks accounted for an additional 159.8 ± 13.6 µg. In this experiment, an overall removal of approximately 39.7 ± 2.2% of dsDNA from the original feed sample was achieved, despite significant co-elution of both TRAP O-rings and dsDNA during the process.

For the elution buffer composed of 50 mM Tris-HCl, 2 M NaCl, pH 8.0, the fractions from peak *4* containing the TRAP O-rings (**Figure IV.14E** and **F**) yielded a mass of 838.0 ± 2.7 µg of these protein nanocages, corresponding to a recovery yield of 419.0 ± 1.4 µg mL^{-1} and a recovery efficiency of 72.9 ± 0.1% relative to the feed from the clarified lysate. Meanwhile, the fractions related to peak *3* yielded a mass of 207.6 ± 2.7 µg of protein nanocages, corresponding to a recovery yield of 103.8 ± 1.4 µg mL^{-1} and a recovery efficiency of 18.1 ± 0.1%. Clearly, there was a separate elution of the TRAP O-rings by peaks *3* and *4*, although peak *4* appeared to be the most suitable location given its higher yield. This ultimately led to a reduction in the recovery efficiency of the protein O-rings. The elution loss of TRAP O-rings outside this peak *3* was approximately 1.4 ± 1.2 µg, which was negligible. In this set of peaks (*3* and *4*), a reasonable amount of protein impurities with medium and high molecular weights was also eluted. In terms of host dsDNA impurities, the highest removal occurred in peak *3* (approximately 491.2 ± 7.6 µg of dsDNA), with co-elution alongside TRAP O-rings, and a similar behavior was observed in peak *4* with an elution of approximately 57.2 ± 3.6 µg of dsDNA. In the remaining elution fractions, approximately 22.5 ± 3.9 µg of dsDNA was removed. Nonetheless, in this experiment, an overall removal of approximately 95.0 ± 1.2% of dsDNA from the original feed sample was achieved, despite significant co-elution of both TRAP O-rings and dsDNA during the chromatographic purification.

Therefore, the results demonstrated that the purification of TRAP O-rings by CEX chromatography using the Capto™ SP ImpRes resin and the elution buffer composed of 50 mM MES, 2 M NaCl, pH 5.5 allowed for a higher recovery efficiency of the protein O-rings at 73.7 ± 0.8%, compared to the two tested alternatives (approximately 1.1 times lower). This outcome was balanced with good removal efficiencies for host protein and dsDNA impurities, achieving 93.0 ± 1.6% and 93.6 ± 0.8%, respectively. Furthermore, this method clearly achieved better separation of the TRAP O-rings from

other impurities, particularly dsDNA. In terms of the elution buffer composition, increasing the NaCl concentration appeared to be the best strategy for the purification of TRAP O-rings, as previously identified. These values were approximately 1.1 to 1.5 times higher than those obtained under the best conditions with the HiTrap Capto™ S column and the elution buffer with the same composition (50 mM MES, 2 M NaCl, pH 5.5). A possible explanation for this difference may be related to the particle size of the two resins. HiTrap Capto™ S has a particle size of 90 µm, while Capto™ SP ImpRes has a particle size of 36 to 44 µm, which likely contributes to its superior resolution and performance.

In this way, an overall assessment of this initial phase of TRAP O-rings purification suggested that the use of the Capto™ SP ImpRes resin combined with the mobile phases comprising buffers A and B with the mentioned compositions (50 mM MES, pH 5.5 and 50 mM MES, 2 M NaCl, pH 5.5, respectively) appeared to be the most promising settings, with the greatest potential for large scale downstream processing of TRAP O-rings.

Given the promising results obtained with Capto™ SP ImpRes, it was decided to advance to a more optimized procedure for the CEX-based purification of TRAP O-rings. This optimization focused on implementing a stepwise elution gradient, aiming to establish a well-defined and consistent purification strategy.

Regarding the experimental conditions, the mobile phases used consisted of buffer A, composed of 50 mM MES, pH 5.5, and buffer B, composed of 50 mM MES, 2 M NaCl, pH 5.5. The Capto™ Q ImpRes column was pre-equilibrated with 8% of buffer B ($\kappa \approx 20.7$ mS cm^{-1}). The sample injected into the column contained approximately 2 mg of total protein from the clarified lysate, which included the TRAP O-rings. After washing out unbound material using 8% of buffer B, a two-step gradient was applied for the elution of bound species. The first step involved 35% of buffer B ($\kappa \approx 61.9$ mS cm^{-1}), followed by the second step with 100% of this buffer B ($\kappa \approx 141.3$ mS cm^{-1}).

The chromatogram obtained is presented in **Figure IV.15A**, along with the resulting SDS-PAGE image (**Figure IV.15B**). SDS-PAGE analysis of the clarified lysate feed sample used in all these experiments is shown in **Figure IV.11B**. As before, the fractions from the chromatographic run were analyzed by SDS-PAGE as well as quantified for total protein content using the BCA method and for total dsDNA using qPCR (sub-sections **IV.2.7.2. SDS-PAGE**, **IV.2.7.1. Total protein quantification** and**IV.2.7.9. Total dsDNA quantification by qPCR**).

Considering these results, a highly interesting profile was observed, characterized by a major flowthrough peak, a peak during the first elution step corresponding to the fractions containing the TRAP O-rings, and a second elution step associated with host impurities. In the flowthrough (**Figure IV.15A**, *1*), TRAP O-rings were visualized being eluted, which was not an expected result. It would be necessary to determine the binding capacity of the column for this type of sample and also to adjust the conductivity during the binding/washing stages to ensure that the loss of TRAP O-rings in the flowthrough is negligible. SDS-PAGE analysis revealed the removal of protein impurities, with several weak bands of different molecular weights appearing. The washing loss of TRAP O-rings in this first peak was approximately 699.0 ± 25.6 µg. Peak *2* of the protein O-rings exhibited a reasonably sharper profile, with a retention volume of approximately 17.5 mL at a conductivity of 35 mS cm^{-1} (**Figure IV.15A**, *2*). The SDS-PAGE analysis of peak *2* showed a well-defined band with strong intensity corresponding to

the monomers of TRAP O-rings, along with very faint bands that may represent residual host protein impurities, although in virtually negligible amounts (**Figure IV.15B**, *2*). For the peak *3*, no protein bands were observed in the SDS-PAGE analysis, supporting the hypothesis that it may contain some host impurities, such as proteins and dsDNA, but in quantities so low that they were undetectable by SDS-PAGE. This was further supported by the low $UV_{280\ nm}$ signal observed for this peak (**Figure IV.15A** and **B**, *3*).

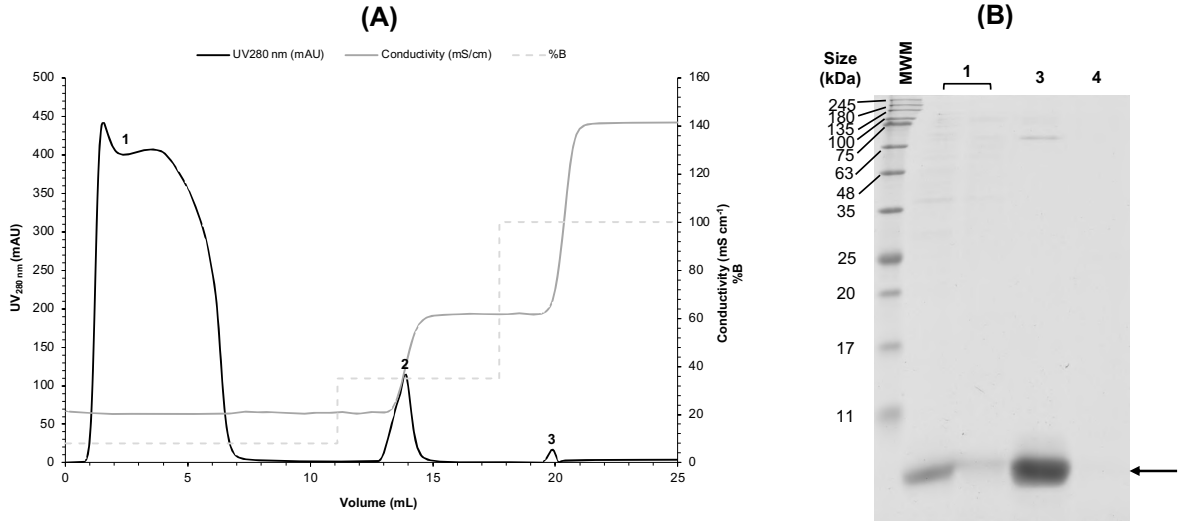


Figure IV.15 – Optimized CEX chromatography strategy for the purification of TRAP O-rings produced in *E. coli* BL21(DE3). **(A)** Chromatogram obtained using a 1 mL chromatographic column with the resin Capto™ SP ImpRes (Cytiva; Marlborough, MA, USA). A feed stream corresponding to approximately 2 mg of total protein was injected into this column previously equilibrated with 8% of buffer B ($\kappa \approx 20.7$ mS cm^{-1}). Unbound material was washed with 8% of buffer B and elution was performed using a two-step gradient, the first step with 35% of buffer B ($\kappa \approx 61.9$ mS cm^{-1}) and the second step with 100% of buffer B ($\kappa \approx 141.3$ mS cm^{-1}). Buffers A and B were composed of 50 mM MES, pH 5.5 and 50 mM MES, 2 M NaCl, pH 5.5, respectively. **(B)** SDS-PAGE analysis of fractions collected during the chromatographic run (peaks *1*, *2*, and *3*). Peak *1* corresponded to the flowthrough and peaks *2* and *3* to the eluted material. SDS-PAGE analysis of the clarified lysate feed sample used is shown in **Figure IV.11B**. Arrow indicates the protein bands referring to the monomer of TRAP O-rings (expected molecular weight of 8.3 kDa), which is separated and visualized under denaturing conditions of SDS-PAGE. The abbreviation MWM refers to the molecular weight marker (NZYColour protein marker II; NZYtech Lisbon, Portugal).

The fractions from peak *2* containing the TRAP O-rings (**Figure IV.15A** and **B**) yielded a mass of 800.5 ± 13.0 µg of these protein O-rings, corresponding to a recovery yield of 320.2 ± 5.2 µg mL^{-1} and a recovery efficiency of 69.6 ± 0.8% relative to the feed from the clarified lysate. In terms of the host impurities removal, approximately 82.0 ± 2.3% of proteins and about 85.0 ± 5.2% of dsDNA were removed. The recovery efficiency values represented an increase of approximately 4.0 times compared to the standard methodology, which was a highly significant improvement. The efficiency of impurities removal was similar between this optimized strategy and the standard approach.

The hypothesis of using centrifugal filtration to remove potential impurities and TRAP O-rings with incorrect assembly was preliminarily evaluated (for this purpose, centrifugal filters with a cut-off of 50 kDa were chosen). However, a preliminary test estimated a significant loss of TRAP O-rings due to adsorption onto the centrifugal filter membrane, even after applying a passivation step to reduce non-specific interactions, decreasing the recovery efficiency to a value lower than 25%.

In conclusion, a single CEX chromatography step may be the best option to achieve a suitable balance between the amount of recovered protein O-rings and the removal of host protein impurities.

As previously discussed for MjsHSP nanocages, one of the major challenges in evaluating downstream processing was the quantification of TRAP O-rings at various stages of the process. Densitometric analysis of SDS-PAGE was the strategy considered. However, it is not the most appropriate method to obtain highly precise and robust estimates. To address this objective as well as to evaluate the oligomeric state of these TRAP O-rings, it was decided to investigate the use of analytical SEC. The idea was to explore the considerable size differences between TRAP O-rings and other impurities, using the fundamental principles of SEC. Initial studies were conducted using an analytical SEC column, specifically the TSKgel SuperSW3000. The preliminary results suggested potential for this approach when applied to samples containing nearly or fully purified TRAP O-rings. Nevertheless, analysis of samples directly from clarified lysates did not yield reliable results. Furthermore, no positive outcomes were observed in terms of distinguishing oligomeric states, possibly because the size differences were insufficient for precise resolution. Additional experiments with this strategy or alternative methods, such as hydrogen bond chromatography, are necessary to establish a clearly robust strategy for analytical quantification. A chromatography-based quantification method would undoubtedly be a very suitable approach for this type of application.

IV.3.3.3. Evaluation of pre-treatment of cell lysate with heating

In order to facilitate and improve the purification process of TRAP O-rings produced in *E. coli* BL21(DE3) through chromatography, it was decided to explore a strategy frequently described in the literature, which involved adding a preliminary step of treating the clarified lysate with heat. This approach explored the stability of these TRAP O-rings at elevated temperatures, where other host impurities, such as proteins, typically lose stability and precipitate. At the same time, it was also intended to evaluate whether the presence of nucleases in the clarified lysate could assist in the removal of dsDNA by creating favorable conditions for the enzymatic reaction.

To this end, two main approaches were tested. The first involved subjecting a clarified cell lysate containing the TRAP O-rings to an initial incubation step at 37°C for 1 hour, followed by a second incubation step at 85°C for 30 minutes. The purpose of using a longer first step at 37°C was to determine whether the nucleases present in the clarified cell lysate were still active and capable of degrading dsDNA present in this lysate. This temperature was chosen because it is within the typical optimal range for nuclease activity. The second approach treated a similar sample of the clarified lysate with a single incubation step at 85°C for 30 minutes. At this higher temperature, most host protein impurities were

expected to precipitate, leaving the TRAP O-rings soluble, as they are stable at this temperature. Experimental details are available in sub-section **IV.2.6.1. Treatment of cell lysate with heating**.

In **Figure IV.16**, the SDS-PAGE results obtained from the analysis of clarified cell lysates subjected to a pre-treatment with heating using each of the previously mentioned approaches are presented, considering both the soluble and insoluble fractions.

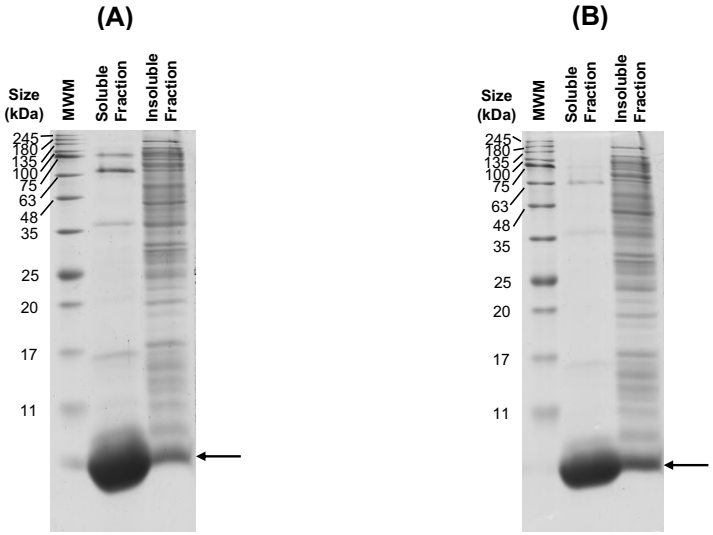


Figure IV.16 – SDS-PAGE analysis of cell lysates containing the TRAP O-rings produced in *E. coli* BL21(DE3) that were subjected to a pre-treatment with heating during the development of a primary isolation and purification strategy. Two alternative methodologies were evaluated, each using 750 µL to 1 mL of the clarified cell lysate (with a concentration of about 2 mg mL^{-1} in terms of total protein). The first method involved a first incubation step at 37°C for 1 hour followed by a second incubation step at 85°C for 30 minutes **(A)**. The second strategy involved a single incubation step at 85°C for 30 minutes **(B)**. In both experiments, at the end of the incubation period, the samples were centrifuged. The resulting supernatant containing the TRAP O-rings (soluble fraction) was collected and the respective pellet containing the protein impurities precipitated (insoluble fraction) was resuspended in 50 mM Tris-HCl, 100 mM NaCl, 1 mM EDTA, pH 8.0 buffer. Arrow indicates the protein bands referring to the monomer of TRAP O-rings (expected molecular weight of 8.3 kDa), which is separated and visualized under denaturing conditions of SDS-PAGE. The abbreviation MWM refers to the molecular weight marker (NZYColour protein marker II; NZYtech Lisbon, Portugal).

For both approaches, it was clearly observed that, as expected, the majority of the TRAP O-rings remained in the soluble fraction, while a significant amount of host protein impurities precipitated and were retained in the insoluble fraction. This very interesting result was confirmed by a more detailed quantification. Regarding the approach with two heating steps (**Figure IV.16A**), it was possible to recover approximately 1268.0 ± 18.1 µg of TRAP O-rings in the soluble fraction, with only about 97.4 ± 6.5 µg remaining in the insoluble fraction. These values corresponded to a recovery efficiency of 88.4 ± 2.3% for TRAP O-rings and a host protein impurities removal efficiency of 62.5 ± 3.2%, which was a very interesting and promising result. For the approach with a single heating step (**Figure IV.16B**), approximately 668.5 ± 6.0 µg of TRAP O-rings were recovered in the soluble fraction, with only about

2.5 ± 1.8 µg remaining in the insoluble fraction. These values corresponded to a recovery efficiency of 46.5 ± 4.8% for TRAP O-rings and a host protein impurities removal efficiency of 75.5 ± 2.5%.

Comparing the results of both strategies, the two-step heating strategy appeared to have better overall performance, achieving a better balance between recovery efficiency and impurities removal efficiency. In the case of the first parameter there was an increase of 1.9 times, and for the second parameter a decrease of 1.2 times. Inopportunely, due to technical issues, it was not possible to quantify dsDNA, which prevented testing the impact of the additional 1 hour incubation at 37°C on nucleic acids degradation. Additional experiments on this topic would be necessary to confirm this hypothesis and consequently improve the procedure.

After undergoing the corresponding thermal treatment by heating, each sample was injected into a 1 mL column containing the strong CEX resin Capto™ SP ImpRes to validate the impact of its use on the purification profile of TRAP O-rings. Regarding experimental conditions, the same parameters described earlier for the CEX chromatography approach using this resin were applied, including elution buffer B (composed of 50 mM MES, 2 M NaCl, pH 5.5), a linear elution gradient incorporating buffer A, the mass of the thermally treated sample injected, and the binding/washing and elution procedures (sub-section **IV.3.3.2. Alternative and optimized purification by chromatography**).

The chromatograms obtained are presented in **Figure IV.17A** and **Figure IV.18A**, together with the corresponding SDS-PAGE images (**Figure IV.17B** and **Figure IV.18B**). All fractions were analyzed by SDS-PAGE and quantified for total protein content using the BCA method (sub-sections **IV.2.7.2. SDS-PAGE** and **IV.2.7.1. Total protein quantification**).

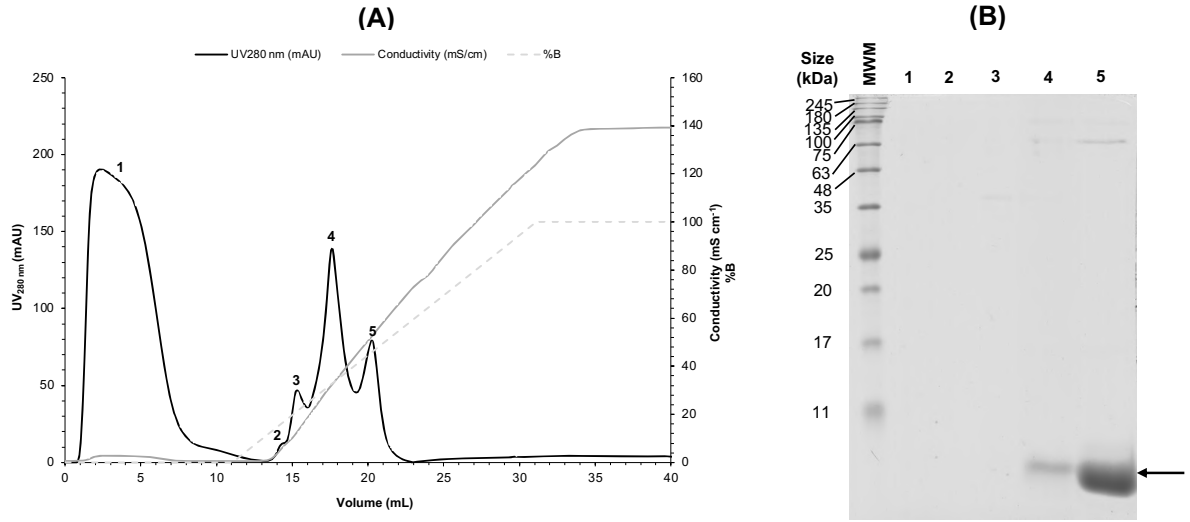


Figure IV.17 – Two-step approach for the purification of TRAP O-rings produced in *E. coli* BL21(DE3): pre-treatment with two sequential heating steps followed by CEX chromatography. **(A)** Chromatogram obtained using a 1 mL chromatographic column with the resin Capto™ SP ImpRes (Cytiva; Marlborough, MA, USA). A feed stream (initial total protein mass of about 1.5 to 2 mg) previously subjected to an additional treatment with heating at 37°C for 1 hour followed by heating at 85°C for 30 minutes was injected into this column earlier equilibrated with buffer A composed of 50 mM MES, pH 5.5 ($\kappa \approx 0.8$ mS cm^{-1}). Unbound material was washed with this buffer A and elution was performed using a linear gradient (20 CV) with increasing salt concentration up to 100% of buffer B composed

of 50 mM MES, 2 M NaCl, pH 5.5 ($\kappa \approx 138.9$ mS cm^{-1}). **(B)** SDS-PAGE analysis of fractions collected during the chromatographic run (peaks *1*, *2*, *3*, *4*, and *5*). Peak *1* corresponded to the flowthrough and peaks *2* to *5* to the eluted material. SDS-PAGE analysis of the pre-treated clarified lysate feed used is shown in **Figure IV.16A** (soluble fraction). Arrow indicates the protein bands referring to the monomer of TRAP O-rings (expected molecular weight of 8.3 kDa), which is separated and visualized under denaturing conditions of SDS-PAGE. The abbreviation MWM refers to the molecular weight marker (NZYColour protein marker II; NZYtech Lisbon, Portugal).

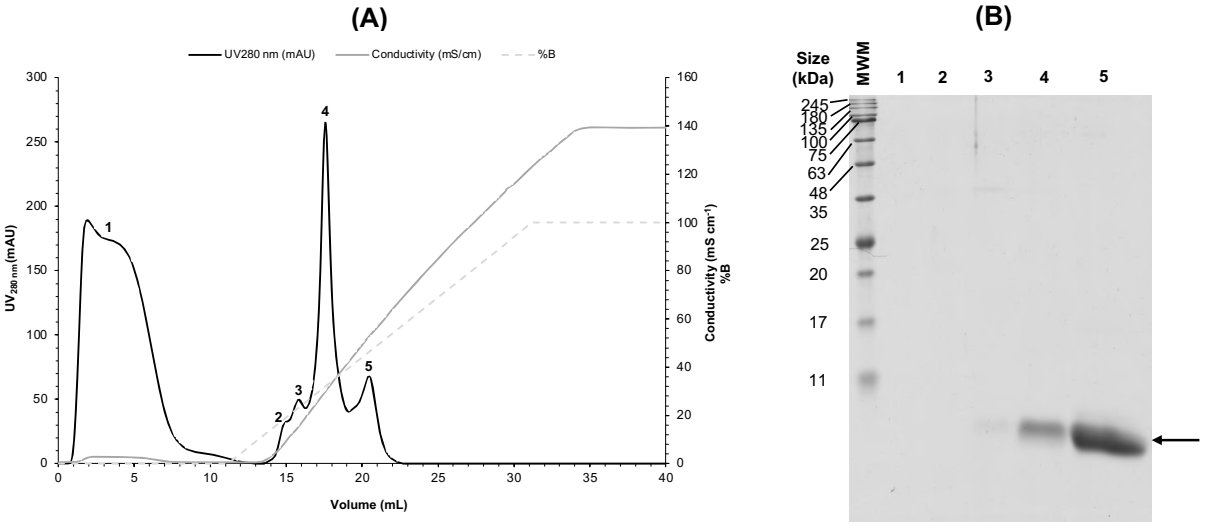


Figure IV.18 – Two-step approach for the purification of TRAP O-rings produced in *E. coli* BL21(DE3): pre-treatment with a single heating step followed by CEX chromatography. **(A)** Chromatogram obtained using a 1 mL chromatographic column with the resin Capto™ SP ImpRes (Cytiva; Marlborough, MA, USA). A feed stream (initial total protein mass of about 1.5 to 2 mg) previously subjected to an additional treatment with heating at 85°C for 30 minutes was injected into this column earlier equilibrated with buffer A composed of 50 mM MES, pH 5.5 ($\kappa \approx 0.8$ mS cm^{-1}). Unbound material was washed with this buffer A and elution was performed using a linear gradient (20 CV) with increasing salt concentration up to 100% of buffer B composed of 50 mM MES, 2 M NaCl, pH 5.5 ($\kappa \approx 139.3$ mS cm^{-1}). **(B)** SDS-PAGE analysis of fractions collected during the chromatographic run (peaks *1*, *2*, *3*, *4*, and *5*). Peak *1* corresponded to the flowthrough and peaks *2* to *5* to the eluted material. SDS-PAGE analysis of the pre-treated clarified lysate feed used is shown in **Figure IV.16B** (soluble fraction). Arrow indicates the protein bands referring to the monomer of TRAP O-rings (expected molecular weight of 8.3 kDa), which is separated and visualized under denaturing conditions of SDS-PAGE. The abbreviation MWM refers to the molecular weight marker (NZYColour protein marker II; NZYtech Lisbon, Portugal).

Generally, the chromatogram profile was largely similar between the two evaluated approaches and also comparable to the CEX chromatography of the non-treated sample (**Figure IV.14A**). A predominant flowthrough peak was observed, where no removal of TRAP O-rings was detected in either situation (**Figure IV.17A** and **B**, and **Figure IV.18A** and **B**, *1*). Peak *5* from the elution step corresponded to the TRAP O-rings, as illustrated by SDS-PAGE (**Figure IV.17A** and **B**, and **Figure IV.18A** and **B**, *5*). However, as previously observed for the non-treated sample, in peak *4*, which was the elution point of dsDNA, occurred its co-elution with some TRAP O-rings, resulting in the loss of these target proteins from the resin (**Figure IV.17A** and **B**, and **Figure IV.18A** and **B**, *4*). It was possible to observe that the

SDS-PAGE band was more prominent in peak *4* in the case of the single heating step approach, which may suggest that a larger quantity of TRAP O-rings did not bind to the column until reaching the desired retention volume. Interestingly, peak *4* as much more pronounced in this single heating step experiment, which could also be related to the enzymatic degradation of dsDNA by nucleases during the initial 37°C step. This enzymatic activity likely reduced the signal significantly in terms of $UV_{280\ nm}$ absorbance.

In the experiment with the two-step heating approach, a TRAP O-rings mass of 893.7 ± 27.0 µg was obtained, with a recovery yield of 297.9 ± 9.0 µg mL^{-1} and a recovery efficiency of 70.9 ± 2.5%, representing a similar recovery efficiency compared to the non-treated sample. The loss of TRAP O-rings during the elution was only about 77.8 ± 2.8 µg. The removal of protein impurities in this step was limited to just 7.1 ± 1.1%. Considering both steps, the initial thermal pre-treatment and the subsequent CEX chromatography, it was possible to achieve a total efficiency of 62.7 ± 0.5% and a removal of host protein impurities of 69.2 ± 0.2%.

In the experiment with a single heating step, a TRAP O-rings mass of 791.6 ± 10.1 µg was obtained, with a recovery yield of 263.9 ± 3.4 µg mL^{-1} and a recovery efficiency of 66.1 ± 2.5%, representing a 1.11-fold decrease in recovery efficiency compared to the non-treated sample. The loss of TRAP O-rings during the elution stage was only about 83.6 ± 0.8 µg. The removal of protein impurities in this step was 32.6 ± 1.1%. Considering both steps, the initial thermal pre-treatment with heating and the subsequent CEX chromatography, it was possible to achieve a total efficiency of 36.7 ± 0.0% and a removal of host protein impurities of 68.6 ± 0.2%. In direct comparison with the first approach, this represents a decrease of 1.7 times in recovery efficiency while achieving a similar value for the removal of protein impurities.

Therefore, a prior thermal treatment with heating could be an interesting approach to apply in the optimized downstream processing of TRAP O-rings (single-step CEX chromatography). Additional studies to identify the optimal conditions for heating could be conducted to achieve consistent and highly efficient removal of host-derived proteins and dsDNA from clarified cell lysates from *E. coli* BL21(DE3) containing TRAP O-rings. Based on the results and data collected in this part of the work, in terms of recovery efficiency and impurities removal, it can only be suggested that single-step CEX chromatography without additional thermal pre-treatment steps appeared to be more promising. The quantification of dsDNA by qPCR could clearly be a useful tool to evaluate the indirect impact of heating through nuclease activity on this challenging type of impurity.

After establishing an efficient and robust optimized process for the purification of TRAP O-rings from clarified lysates of *E. coli* BL21(DE3), it would be necessary to advance to the *in vitro* assembly of these O-rings for the subsequent formation of TRAP nanocages. However, several challenges emerged during this assembly process, with multiple reactions failing to allow the formation of fully assembled nanostructures or resulting in their presence in quantities that were too low, making it impossible to obtain a consistent and sufficient sample for their purification. In this regard, further experiments would be necessary to optimize the *in vitro* assembly procedure. Once this is achieved, it would then be possible to proceed with different optimization experiments for the purification of TRAP nanocages using chromatography-based techniques.

If the *in vitro* assembly process of TRAP O-rings was successful, three chromatographic approaches would be evaluated for the purification of TRAP nanocages: SEC, multimodal chromatography, and hydrogen bond chromatography. The main justification for each of these approaches lies in the considerable dimensions of these protein nanostructures, which could be more easily exploited compared to the significantly smaller sizes of eventual impurities. For SEC, a conventional resin such as Superdex™ 200 would be considered. In the case of multimodal chromatography, the Capto™ Core 400 resin would be evaluated, while for the hydrogen bond chromatography, the CIMmultus® H-Bond monolith would be explored. As was performed for the TRAP O-rings, the objective would be to investigate not only the chromatographic supports but also the mobile phases, the amount of sample injected, the binding/washing/elution conditions, and the capacity to distinguish between different oligomeric states, among others.

IV.4. Conclusions

NVPNs have become a focal point of research in bioengineering, biotechnology, and biomedicine, showing promising results in targeted drug delivery (114,126,127), innovative vaccine platforms (119,129,132), advanced bioimaging techniques (134,135), controlled biomineralization processes (93), and efficient biocatalytic systems (75,144). Research on NVPNs has made substantial progress in conceptual studies and applications, but there is a significant lack of attention on developing bioprocesses for both natural and engineered protein nanocages (186). Downstream processing is essential for NVPNs, as it ensures the final product meets specific quality standards. This process involves extracting and isolating NVPNs from the culture broth, with purification being a critical step (138,141,146). The current research on isolating and purifying protein nanocages is scattered and lacks detailed information, impeding significant comparisons between different methods. Further development is necessary to improve recovery yields of protein nanocages, making them viable for large scale production.

In this **Chapter IV**, the main objective was to explore the downstream processing of the MjsHSP nanocages as well as of the TRAP nanocages and their precursors, the TRAP O-rings. The use of protein nanocages produced by two different bacterial hosts (*E. coli* and *V. natriegens*) was one of the novelties of this work, as well as the investigation of alternative purification strategies that can take advantage of the physico-chemical properties of these protein-based nanostructures.

Considering the results obtained, a standard purification approach for MjsHSP nanocages from *E. coli* BL21(DE3) was successfully established based on the literature information. Using this as a starting point, various optimization strategies were implemented to enhance recovery efficiency and impurities removal, testing parameters and conditions such as the type of chromatographic support, the effectiveness of bead and pore size, and the composition of binding and elution buffers. An optimized approach for MjsHSP nanocages was defined for *E. coli*, consisting of an enzymatic pre-treatment of the clarified cell lysate followed by a single-step AEX chromatography. The equivalent methodology for MjsHSP nanocages produced in *V. natriegens* consisted only of a single step of AEX chromatography. For TRAP O-rings, a conventional standard approach was initially implemented, followed by alternative methods to exploit their distinct physico-chemical properties compared to impurities like host proteins and dsDNA. Different resins and mobile phase compositions were tested. The fully optimized approach should involve thermal heating pre-treatment of the clarified lysate followed by a single-step CEX chromatography. However, due to challenges in the *in vitro* assembly of TRAP O-rings, the purification of TRAP nanocages could not proceed as planned.

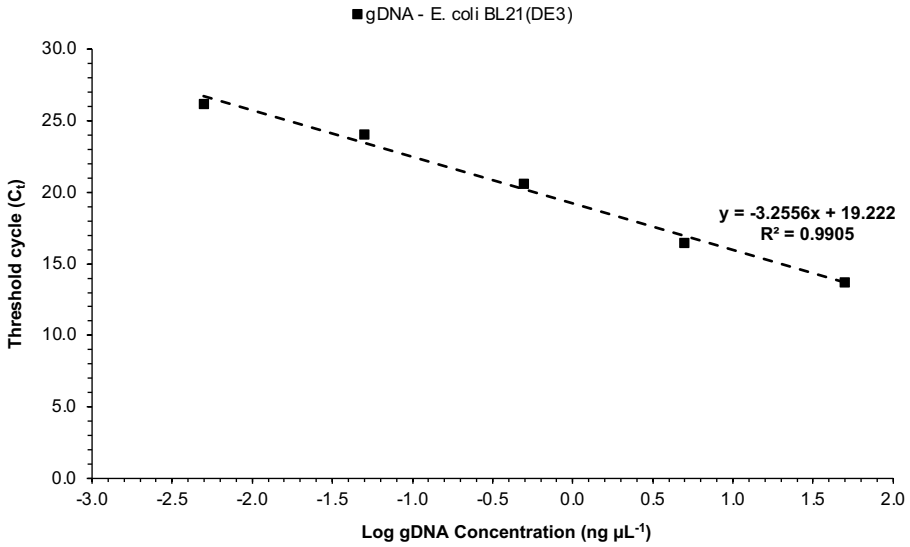
With the optimized procedure for the primary isolation and purification of MjsHSP nanocages from *E. coli* (enzymatic pre-treatment of the clarified cell lysate with TURBO™ DNase followed by a single-step AEX chromatography with the strong Capto™ Q ImpRes resin in a bind/elute mode with a stepwise elution gradient involving an increase in NaCl concentration), a protein nanocages recovery yield of 151.9 ± 1.8 μg mL^{-1}, a recovery efficiency of 63.3 ± 3.3% and a host protein impurities removal of 88.0 ± 1.0% were obtained. Regarding the downstream processing benchmark, these results correspond to a 7.2-fold increase in terms of recovery efficiency and a slight 1.1-fold decrease of in terms of impurities removal. In parallel, for the MjsHSP nanocages from *V. natriegens*, the optimized process comprised a

similar single-step AEX chromatography with the strong Capto™ Q ImpRes resin in a bind/elute mode with a stepwise elution gradient involving an increase in NaCl concentration. A protein nanocages recovery yield of 3.4 ± 0.4 μg mL^{-1}, a recovery efficiency of 9.5 ± 1.1% and a host protein impurities removal of 96.8 ± 1.3% were obtained. Although these values were significantly lower than those of *E. coli* in direct comparison, these experiments confirmed the possibility of purifying MjsHSP nanocages derived from an alternative host bacterial strain for protein expression. Regarding the optimized TRAP O-rings downstream process (single-step CEX chromatography with the strong Capto™ SP ImpRes resin in a bind/elute mode with a stepwise elution gradient involving an increase in NaCl concentration), a protein O-rings recovery yield of 320.2 ± 5.2 μg mL^{-1}, a recovery efficiency of 69.6 ± 0.8%, a host protein impurities removal of 82.0 ± 2.3% and a host dsDNA impurities removal of 85.0 ± 5.2% were achieved. This optimized strategy improved recovery efficiency by 4-fold compared to the reference process, while impurities removal remained similar.

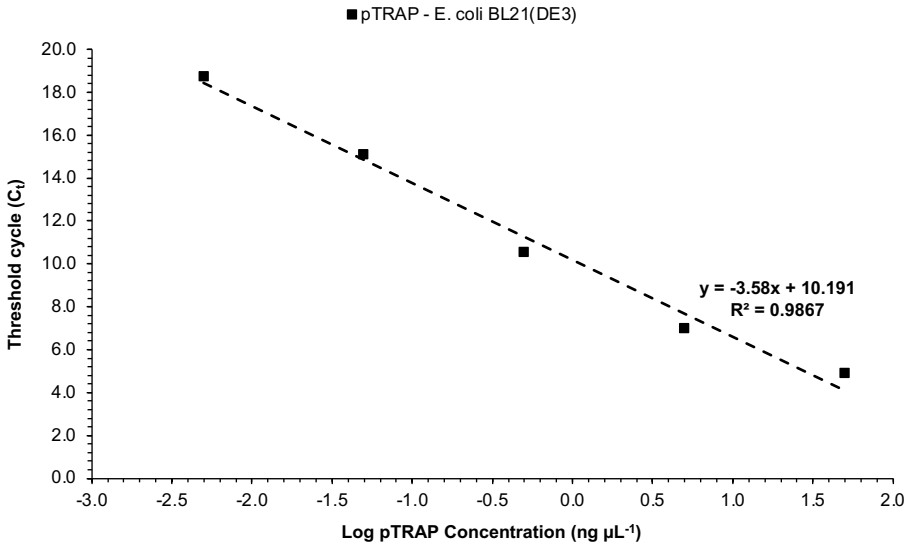
The work developed in this **Chapter IV** can be considered an excellent starting point for the area of process development dedicated to the extraction, primary isolation and purification of protein nanocages. The data generated can be used in future research studies and even help to establish specific protocols for different NVPNs. From the perspective of future work, the first priority will be to repeat certain key experiments to obtain additional replicates, ensuring the reliability and reproducibility of the results. This will help refine the experimental protocols and enable the identification of optimal conditions for the efficient purification of MjsHSP nanocages and TRAP O-rings. By improving consistency in the data, it will be possible to better evaluate critical parameters such as target proteins recovery efficiency and impurities removal efficiency of the purification process. Complementarily, the enzymatic treatment of clarified cell lysates (using DNase and RNase) could be evaluated for the purification of both MjsHSP nanocages produced in *V. natriegens* Vmax™ X2 and TRAP O-rings derived from *E. coli* BL21(DE3) similar to what was already done for MjsHSP nanocages from *E. coli* production. Future studies could focus on fine-tuning enzyme concentrations, incubation times, and reaction conditions to maximize impurity removal while preserving the integrity of the MjsHSP nanocages and TRAP O-rings. One of the major challenges in evaluating downstream processing was the quantification of MjsHSP nanocages and TRAP O-rings at various stages of the process. Densitometric analysis of SDS-PAGE was the strategy considered. Nevertheless, it is not the most appropriate method for obtaining highly precise and robust estimates. In this study, two distinct approaches were explored for MjsHSP nanocages and TRAP O-rings quantification: multimodal chromatography with AEX and hydrogen bond chromatography components, as well as SEC. However, it was a challenging task, and it was not possible to achieve the best results. Clearly, further exploration of this topic in research studies will be necessary, with the ultimate purpose of defining an analytical technique that enables quick, reproducible, and easily applicable quantification of protein nanocages throughout upstream and downstream processing. Implementing chromatography-based quantification would make the process more consistent, robust, and efficient. An optimized process for purifying TRAP O-rings from *E. coli* BL21(DE3) lysates was developed, but challenges in the *in vitro* assembly of these O-rings into their nanocages arose, including incomplete assembly or insufficient quantities, preventing consistent purification. Further experiments will be needed to optimize the assembly procedure before advancing

to purification experiments. If the assembly process succeeds, three chromatography methods will be evaluated for purifying TRAP nanocages: SEC, multimodal chromatography, and hydrogen bond chromatography. These approaches explore the large dimensions of these protein-based nanostructures relative to the impurities. SEC would use Superdex™ 200 resin, multimodal chromatography would use Capto™ Core 400 resin, and hydrogen bond chromatography would explore CIMmultus® H-Bond monolith. The objective will be to optimize chromatographic supports, mobile phases, sample injection amounts, chromatographic run conditions, and the ability to distinguish oligomeric states.

IV.5. Supplementary Material



Supplementary Figure IV.S1 – Example of a standard curve for gDNA from *E. coli* BL21(DE3) (■) used in the quantification of total dsDNA (gDNA and pDNA) by qPCR. Reactions were performed with 5 log serial dilutions of the respective gDNA (420, 42, 4.2, 0.42 and 0.042 ng). Note that the linear regression trendline is represented, as well as the corresponding equation and the coefficient of determination value (R^2).



Supplementary Figure IV.S2 – Example of a standard curve for pDNA (pTRAP) from *E. coli* BL21(DE3) (■) used in the quantification of total dsDNA (gDNA and pDNA) by qPCR. Reactions were performed with 5 log serial dilutions of pTRAP (420, 42, 4.2, 0.42 and 0.042 ng). Note that the linear regression trendline is represented, as well as the corresponding equation and the coefficient of determination value (R^2).

Chapter V

Structural Characterization of MjsHSP and TRAP Nanocages

Abstract

Over the last years, NVPNs have emerged as highly versatile tools in a wide range of scientific fields, offering significant advantages in areas such as drug delivery, vaccine development, bioimaging, biomineralization, and biocatalysis. While substantial progress has been achieved in understanding the properties of NVPNs for various applications, critical gaps persist not only in the manufacturing efficiency (both upstream and downstream processes) but also in the analytical methodologies for structural characterization. The development of robust analytical and characterization methodologies is crucial for accurately assessing the structural features of the final nanocages and ensuring comprehensive monitoring of the biomanufacturing process at every stage.

This **Chapter V** presents a comparative structural analysis of two NVPNs, MjsHSP nanocages produced in *E. coli* and *V. natriegens* and TRAP nanocages produced in *E. coli* (including precursor TRAP O-rings) using biophysical techniques. For *E. coli*-derived MjsHSP nanocages, dynamic light scattering (DLS) and transmission electron microscopy (TEM) provided complementary insights into the hydrodynamic diameter, the protein nanocages aggregate detection and the morphological validation. Atomic force microscopy (AFM) revealed protein nanocages topography but highlighted challenges in optimizing sample preparation, while fluorescence correlation spectroscopy (FCS) enabled analysis of the diffusion coefficient and nanocages aggregation assessment via fluorophore labeling. MjsHSP nanocages produced by *V. natriegens* were characterized exclusively using DLS and TEM. TRAP O-rings and nanocages were similarly characterized using DLS, TEM and FCS, demonstrating methodology transferability.

The structural analysis revealed distinct characteristics for each type of NVPNs studied. MjsHSP nanocages from *E. coli* demonstrated consistent assembly, with DLS indicating an average hydrodynamic diameter of 12.9 ± 2.4 nm, corroborated by TEM imaging with negative staining, which showed uniform octahedral structures (12.8 ± 1.1 nm). AFM provided a reasonable resolution surface topology. FCS measurements suggested a partially monodisperse population of MjsHSP nanocages with a diffusion coefficient of 41 $\mu m^2\ s^{-1}$, indicating reduced aggregation. Interestingly, MjsHSP nanocages produced in *V. natriegens* cells showed minimal structural differences, with a slightly larger hydrodynamic diameter (15.3 ± 3.8 nm). TEM analysis revealed properly assembled cages, with considerable similarity in terms of the observable morphology in relation to the results obtained in *E. coli* (18.4 ± 1.9 nm). TRAP O-rings and nanocages exhibit a more complex dynamics of assembly and disassembly. By DLS, TRAP O-rings showed a hydrodynamic diameter of 8.3 ± 0.5 nm, which was confirmed through TEM images that displayed a clear morphology of the O-rings (18.7 ± 2.0 nm). TRAP nanocages were not detected by these two techniques. FCS data suggested the presence of partially assembled TRAP nanocages, with a diffusion coefficient of 18 ± 4.7 $\mu m^2\ s^{-1}$, alongside TRAP O-rings with a diffusion coefficient of 59 ± 11 $\mu m^2\ s^{-1}$ (estimated hydrodynamic diameter of 8.3 ± 1.6 nm).

Keywords: atomic force microscopy, dynamic light scattering, fluorescence correlation spectroscopy, protein nanocages, transmission electron microscopy

V.1. Introduction

NVPNs have recently emerged as a significant focus in bioengineering, biotechnology, and biomedical research. These nanoscale protein assemblies have demonstrated remarkable versatility across various applications. Studies have shown promising results in utilizing NVPNs for targeted drug delivery (114,126,127), novel vaccine platforms (119,129,132), advanced bioimaging techniques (134,135), controlled biomineralization processes (93), and efficient biocatalytic systems (75,144). The diverse applications of NVPNs highlight their potential to advance multiple areas of scientific and medical research. This topic was explored in more detail in **Chapter I**.

Although proof-of-concept studies have advanced the knowledge and understanding of this type of protein-based nanostructures, including their structural properties, functionalization approaches, and application-specific implementations, limited attention has been directed toward systematic bioprocess development for both natural and engineered variants (186). The rise of various challenges emphasizes the importance of enhancing NVPNs manufacturing processes, including in terms of the analytical and characterization techniques.

Developing and validating robust analytical and characterization methods is essential for assessing the structural and functional properties of protein nanocages, ensuring that the purified product meets specified criteria. Moreover, these analytical techniques play a crucial role in monitoring the performance, consistency, and reliability of each step in the manufacturing process (138).

Based on available research studies in the literature, there are three well-established techniques routinely used for the determination and assessment of the biophysical characteristics of protein nanocages: dynamic light scattering (DLS), transmission electron microscopy (TEM), and analytical size exclusion chromatography (SEC). DLS enables the estimation of the average hydrodynamic diameter of NVPNs (40,92,101,131,135,136,147,153,165). TEM, using negatively stained preparations, not only serves as a tool to observe the morphology of this protein-based nanostructures and compare it with the corresponding theoretical three-dimensional structures but also acts as a complementary strategy for estimating the dimensions of these NVPNs (40,92,101,126,136,148,153,165). Analytical SEC is used to determine the molecular weight of the assembled nanostructures by comparison with a calibration curve, as well as to assess the oligomeric state of the protein nanocages post-production (30,40,89,90,92,136,162,165).

Furthermore, other less common analytical techniques with potential to complement or serve as alternatives in the characterization of NVPNs have been explored. Examples include SEC with multi-angle light scattering (SEC-MALS) (123), SEC with right-angle (RALS)/low-angle (LALS) light scattering (116), small angle X-ray scattering (SAXS) (47,183), mass spectrometry (24,41), native mass spectrometry (101), electron spray mass spectrometry (150), liquid chromatography/electrospray mass spectrometry (LC/MS) (62), electrospray ionization time-of-flight (ESI-TOF) mass spectrometry (77), matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometry (39,151), far-UV circular dichroism (FUV-CD) spectroscopy (49,97,184) and analytical ultracentrifugation (65,183). These techniques are employed to evaluate the structural properties of target protein nanocages, such as the molecular weight, the particle size distribution, and the oligomeric state, as well as to characterize their secondary structure, folding, and thermostability.

Additionally, atomic force microscopy (AFM) is a promising technique for characterizing protein nanocages. In addition to enabling visualization of nanostructures, AFM allows for the determination of mechanical properties and the investigation of protein-protein interactions, as demonstrated in the literature for various types of NVPNs (115,172–175).

Fluorescence-based methods for detecting and studying protein nanocages show great promise, despite being relatively underexplored. These techniques are particularly attractive due to the ease of labeling NVPNs with fluorescent molecules, or fluorophores (39,40,62,65,88,92,151). These fluorescent markers can be readily attached to either the exterior surface or interior cavity of protein nanocages. Common fluorophores used in this process include Alexa Fluor 488-maleimide, Alexa Fluor 750-maleimide, fluorescein, and Cys 5.5 dye. Fluorescence correlation spectroscopy (FCS) utilizes time-dependent fluctuations in fluorescence intensity to determine several key physical and chemical parameters (176). This technique allows to quantify translational and rotational diffusion coefficients, from which hydrodynamic diameters can be inferred. Additionally, FCS can measure chemical kinetic rate constants and assess molecular aggregation dynamics. By analyzing these fluctuations, FCS provides non-invasive, real-time insights into nanoscale interactions and motion, making it a relevant tool for studying molecular behavior in complex biological systems, such as NVPNs.

From a functional characterization perspective in order to verify the biological activity of NVPNs after the biomanufacturing process, *in vitro* studies employing various targeting strategies have been used to evaluate interactions between targeted protein nanocages and mammalian cells, as well as the subsequent internalization process. Exemplifying, a research work of Suci *et al.* demonstrated that Dps nanocages functionalized with biotin could successfully deliver a photosensitizer (SnCe6) to targeted cells (29). Similarly, Bellini *et al.* engineered ferritin nanocages with hepatocellular carcinoma cell-binding peptides, improving their specificity for liver cancer cells (125). Moon *et al.* modified sHSP nanocages with neuropilin 1-binding peptides to enhance their targeting capabilities (30). To assess the biological activity of NVPNs after internalization, various strategies have been employed, including gene silencing experiments using nanoparticles loaded with siRNA or antisense oligonucleotides (241), drug delivery efficacy evaluations (126,127), and antigen presentation studies for vaccine development (129,130). For example, scientific works of Guan *et al.* and Ji *et al.* assessed the delivery of anticancer drugs like doxorubicin and OSU03012 using functionalized NVPNs (126,127). Moreover, Wang *et al.* and Kar *et al.* demonstrated that protein nanocages displaying specific antigens could trigger strong CD8+ and CD4+ T-cell responses (129,130).

In this **Chapter V**, the structural characterization of two model NVPNs, MjsHSP and TRAP nanocages, including the precursors of the latter (TRAP O-rings), was explored. For MjsHSP nanocages produced in *E. coli*, DLS and TEM were employed as two well-established techniques to analyze their biophysical properties. DLS provided estimates of hydrodynamic diameter, aiding in assessing assembly correctness and detecting potential aggregates. These findings were corroborated by TEM analyses, which offered visual confirmation of nanostructure morphology and assembly states. Complementarily, AFM and FCS were implemented as less conventional techniques. AFM enabled visualization of MjsHSP nanocages surface topology, though challenges arose in identifying optimal sample preparation conditions to maximize analysis quality with this microscopic method. Given limited

prior exploration of fluorescence-based approaches in the literature, FCS was investigated after labeling NVPNs with an appropriate fluorophore. This allowed estimation of the diffusion coefficient and subsequently inferring the hydrodynamic diameter of the protein-based nanostructures, in addition to evaluating the formation of possible protein nanocages aggregates. MjsHSP nanocages produced by *V. natriegens* were characterized exclusively using DLS and TEM. Regarding the TRAP O-rings and nanocages, these were analyzed and explored using DLS, TEM and FCS, employing experimental strategies similar to those applied to *E. coli*-derived MjsHSP nanocages. This comprehensive approach allowed for a thorough structural characterization of both the precursor O-rings and the fully assembled TRAP nanocages.

V.2. Materials and Methods

V.2.1. Materials

Bacto™ Yeast Extract was acquired from BD Biosciences (Franklin Lakes, NJ, USA). Acrylamide/bis-acrylamide solution (40%), Laemmli sample buffer (4x) and precision plus protein dual color standards were from Bio-Rad (Hercules, CA, USA). Capto™ Q ImpRes resin, Capto™ SP ImpRes resin, Sephadex G-25 columns, Superdex™ 200 resin, and Tricorn™ empty chromatography columns (5/50 and 10/300) were purchased from Cytiva (Marlborough, MA, USA). Magnesium chloride ($MgCl_2$) was obtained from Fagron Ibérica (Barcelona, Spain). Acetic acid glacial, ethanol solution (96% v/v), glucose, hydrochloric acid (HCl), isopropyl β-D-1-thiogalactopyranoside (IPTG), 2-(N-morpholino)ethanesulfonic acid (MES), phosphate buffered saline (PBS) solution (10x), sodium chloride (NaCl), sodium hydroxide (NaOH), tris(hydroxylmethyl)aminomethane (Tris) and Tween™ 20 were purchased from Fisher Scientific (Hampton, NH, USA). PD SpinTrap G-25 columns and PhastGel™ Blue R (Coomassie R 350 stain) were acquired from GE Healthcare (Chicago, IL, USA). Disposable polystyrene cuvettes (DTS0012) and folded capillary zeta cells (DTS1070) were obtained from Malvern Panalytical (Malvern, United Kingdom). Potassium chloride (KCl) was from Merck (Darmstadt, Germany). Dithiothreitol (DTT), Luria-Bertani (LB) broth, NZYColour protein marker II and SDS-PAGE sample loading buffer (5x) were purchased from NZYtech (Lisbon, Portugal). Sodium phosphate dibasic (Na_2HPO_4), sodium phosphate monobasic (NaH_2PO_4) and tryptone were acquired from Panreac AppliChem (Barcelona, Spain). RNase A was obtained from Roche (Basel, Switzerland). Amicon Ultra-4 centrifugal filters (3 kDa, 10 kDa and 100 kDa MWCO), ammonium persulfate (APS), ampicillin, ethylenediaminetetraacetic acid (EDTA), glycine, Millex™ PVDF syringe filters (0.1 µm), nickel(II) chloride ($NiCl_2$), N,N,N′,N′-tetramethylethylenediamine (TEMED), sodium dodecyl sulfate (SDS), syringes PP/PE without needle (1 mL), tris(2-carboxyethyl)phosphine hydrochloride (TCEP) and urea were purchased from Sigma-Aldrich (St. Louis, MO, USA). Chloro[diphenyl(3-sulfonatophenyl)phosphine]gold(i) sodium salt hydrate (Au-TPPMS) was obtained from STREM Chemicals UK (Cambridge, United Kingdom). Invitrogen™ Alexa Fluor™ 647 Maleimide, Invitrogen™ Alexa Fluor™ 647 NHS Ester (Succinimidyl Ester), Invitrogen™ Alexa Fluor™ 647 Protein Labelling Kit, Invitrogen™ TURBO DNase™ (and respective TURBO™ DNase buffer (10x)) and Pierce™ BCA Protein Assay Kit were acquired from Thermo Fisher Scientific (Waltham, MA, USA). All the specific materials necessary for negative staining TEM were provided by the HEMS core facility at i3S, Universidade do Porto, Portugal. Specific materials required for AFM were provided by the AFMaRT Laboratory, Bio-PhysNano, BioISI at Faculdade de Ciências, Universidade de Lisboa, Portugal. The pure and ultrapure water used during all experiments was provided by a Milli-Q purification system (Millipore; Bedford, MA, USA).

V.2.2. Production and purification of MjsHSP nanocages

V.2.2.1. *E. coli* cells

The expression of MjsHSP nanocages was performed principally using the experimental methodology defined as reference, whose procedure is described in detail in the section **III.2.10. Production of MjsHSP nanocages**. Alternatively, the production of these protein nanocages in *E. coli* BL21(DE3) cells was carried out according to the optimized conditions determined in **Chapter III** of this work (**III.3.5. Optimized production of MjsHSP nanocages in *E. coli* and *V. natriegens***), namely the use of enhanced 2xYT culture medium, an IPTG concentration of 0.1 mM, an induction temperature of 37°C, a time point in the middle of the exponential growth phase ($OD_{600\ nm}$ of approximately 0.5 to 0.6) and an aeration setup comprising a non-baffled shake flask. This respective experimental protocol is available in the section **III.2.12. Optimization of production of MjsHSP nanocages**.

Subsequently, the resulting samples containing the produced MjsHSP nanocages were isolated and purified using the standard and/or optimized chromatography-based approaches (with and without previous enzymatic treatment with nucleases), which were explored in **Chapter IV** of this work (**IV.3.1. Primary isolation and purification of MjsHSP nanocages produced in *E. coli* cells**). AEX chromatography combined with SEC after concentration of samples by centrifugal filtration or only AEX followed by the same step of concentration were the methodologies considered (sub-section **IV.2.5.1. Production in *E. coli* cells** of section **IV.2.5. Primary isolation and purification of MjsHSP nanocages**).

V.2.2.2. *V. natriegens* cells

Similar to what was applied for *E. coli*, the expression of MjsHSP nanocages in *V. natriegens* Vmax™ X2 cells was performed essentially using the experimental approach defined as benchmark. Note that the procedure is also described completely in the section **III.2.10. Production of MjsHSP nanocages**. Alternatively, the production of these protein nanocages was carried out according to the optimized conditions determined in **Chapter III** of this work (**III.3.5. Optimized production of MjsHSP nanocages in *E. coli* and *V. natriegens***), namely the use of enhanced 2xYT culture medium, an IPTG concentration of 0.1 mM, an induction temperature of 37°C, a time point in the middle of the exponential growth phase ($OD_{600\ nm}$ of approximately 0.5 to 0.6) and an aeration setup comprising a baffled shake flask. Again, this respective experimental protocol is presented in the section **III.2.12. Optimization of production of MjsHSP nanocages**.

Afterwards, the resulting samples containing the produced MjsHSP nanocages were isolated and purified using the optimized chromatography-based approach, which was explored in **Chapter IV** of this work (**IV.3.2. Primary isolation and purification of MjsHSP nanocages produced in *V. natriegens* cells**). AEX chromatography combined with SEC after concentration of samples by centrifugal filtration or only AEX followed by the same step of concentration were also the methods considered (sub-section **IV.2.5.2. Production in *V. natriegens* cells** of section **IV.2.5. Primary isolation and purification of MjsHSP nanocages**).

V.2.3. Production and purification of TRAP O-rings

The expression of TRAP O-rings as precursors of TRAP nanocages in *E. coli* BL21(DE3) cells was performed mainly using the experimental method defined as reference, whose procedure is described in detail in the section **III.2.11. Production of TRAP O-rings**. Alternatively, the production of these protein nanocages was performed according to the optimized conditions obtained in **Chapter III** of this work (**III.3.10. Optimized production of TRAP-O rings in *E. coli***), namely the use of enhanced 2xYT culture medium, an IPTG concentration of 0.1 mM, an induction temperature of 37°C, a time point in the middle of the exponential growth phase ($OD_{600\ nm}$ of approximately 0.5 to 0.6) and an aeration setup comprising a non-baffled shake flask. This respective experimental protocol is available in the section **III.2.13. Optimization of production of TRAP O-rings**.

Then, the resulting samples containing the produced TRAP O-rings were isolated and purified using the optimized chromatography-based approach (without thermal treatment with high temperature), which was explored in **Chapter IV** of this work (**IV.3.3. Primary isolation and purification of TRAP O-rings**). CEX chromatography followed by a step of sample concentration through centrifugal filtration was the strategy considered (**IV.2.6. Primary isolation and purification of TRAP O-rings**).

V.2.4. Assembly of TRAP nanocages

The assembly of TRAP nanocages was performed according to the procedure described in detail in the section **IV.2.4. Assembly of TRAP nanocages** from **Chapter IV**. Note that the methodology used was based on the work described by Malay *et al.* (111).

V.2.5. Characterization of MjsHSP nanocages

V.2.5.1. Production in *E. coli* cells

V.2.5.1.1. Dynamic light scattering

DLS experiments were performed using the equipment Zetasizer Nano ZS (Malvern Panalytical; Malvern, United Kingdom) and the respective Zetasizer Nano software (Malvern Panalytical; Malvern, United Kingdom).

For particle size measurements, the preparation of samples containing MjsHSP nanocages produced in *E. coli* BL21(DE3) and purified by different approaches included an initial step to remove large particles and dust. This was achieved through either filtration using syringe filters with a 0.1 µm pore size, or centrifugation at 12000×g for 2 minutes at room temperature, or both methods in combination. When necessary, an additional dilution step using PBS solution (1x) was included to ensure that the protein concentration of the sample was consistent with the detection limit of the technique (from 0.1 to 1 mg mL^{-1}). The analyzed samples were always originally in PBS solution (1x) or

diluted with this buffer solution. Each sample (1 mL) was analyzed in a DTS0012 disposable polystyrene cuvette (Malvern Panalytical; Malvern, United Kingdom) after a pre-incubation for 2 minutes at 25°C. The measurements were performed using both the standard default parameters and settings suitable for a sample in phosphate buffer. Results were obtained as intensity and number distributions. All samples were measured 3 times, with each measurement corresponding to approximately 12 readings. Blank measurements were performed using the PBS solution (1x).

For zeta potential measurements, the same initial step to remove larger particles and dust was applied, as well as the subsequent dilution step. The latter was performed not only to ensure that the protein concentration of the sample was within the detection range of the technique, but also to guarantee that the sample conductivity was below the maximum allowed limit (200 mS cm^{-1}). In these cases, the analyzed samples were diluted with Milli-Q water before measurement, since these samples were generally in PBS solution (1x). Each sample (1 mL) was analyzed in a DTS1070 folded capillary zeta cell (Malvern Panalytical; Malvern, United Kingdom) after a pre-incubation for 2 minutes at 25°C. The measurements were performed using both the standard default parameters and settings suitable for a sample in water. Results were obtained as zeta potential distributions. All samples were measured 3 times, with each measurement corresponding to approximately 12 readings. Blank measurements were performed using the PBS solution (1x) or Milli-Q water.

V.2.5.1.2. Transmission electron microscopy

For negative staining TEM, the preparation of samples containing MjsHSP nanocages produced in *E. coli* BL21(DE3) and purified by different approaches included, when necessary, a dilution step using PBS solution (1x) or Milli-Q water according to a dilution factor of 1:10 and/or 1:100. Subsequently, 10 µL of samples were mounted on formvar/carbon film-coated mesh nickel grids (Electron Microscopy Sciences; Hatfield, PA, USA) and left standing for 2 minutes. The liquid in excess was removed with filter paper, and 10 µL of 1% uranyl acetate were added on to the grids and left standing for 10 seconds, after which, liquid in excess was removed with filter paper. Visualization was carried out on a JEM 1400 TEM at 80 kV (JEOL; Tokyo, Japan). Images were digitally recorded using a CCD digital camera PHURONA (EMSIS; Münster, Germany).

Note that the negative staining TEM was performed at the HEMS core facility at i3S, Universidade do Porto, Portugal with the assistance of Sofia Pacheco and Rui Fernandes. These responsible technicians acknowledge the support of the i3S Scientific Platform HEMS, member of the national infrastructure PPBI – Portuguese Platform of Bioimaging (PPBI-POCI-01-0145-FEDER-022122).

V.2.5.1.3. Atomic force microscopy

For AFM analysis, the preparation of samples containing purified MjsHSP nanocages produced in *E. coli* BL21(DE3) included an initial step to remove large particles and dust, and then a dilution step to a concentration suitable for AFM experiments. The first step was achieved through filtration using syringe filters with a 0.1 µm pore size. The second step was performed with PBS solution (1x), buffer I-AFM (50 mM Tris-HCl, 50 mM NaCl, 5 mM $NiCl_2$, pH 7.5) or buffer II-AFM (50 mM Tris-HCl, 50 mM

NaCl, 10 mM $MgCl_2$, pH 7.5) for a final concentration of protein nanocages ranging from 10 to 1 nM. Subsequently, each sample (10 µL) was deposited onto a freshly cleaved mica surface. After 15 minutes to allow for adsorption onto this supporting substrate, the sample was washed with Milli-Q water to remove non-immobilized protein nanocages and then imaged using AFM. The passivation of mica with $MgCl_2$ was also tested before sample placement, using a stock solution with a concentration of 1 M, to promote the adsorption of the MjsHSP nanocages. The topography/morphology of the samples was characterized by AFM using a Pico SPM LE system (Molecular Imaging; Tempe, AZ, USA). AFM experiments were conducted in both air and aqueous environments, with the latter proving to be more appropriate. Thus, the images were obtained at room temperature with HiRes-C14/Cr–Au (force constant 5 N m^{-1}, tip radius < 1 nm) probes which are optimized for imaging relatively soft samples and obtaining a better phase contrast-reducing surface deformations caused by the AFM tip.

V.2.5.1.4. Fluorescence correlation spectroscopy

For FCS measurements, two types of samples containing purified MjsHSP nanocages previously produced in *E. coli* BL21(DE3) were labelled with the Alexa 647 NHS ester fluorophore (absorption and fluorescence emission maxima of approximately 650 nm and 668 nm, respectively).

Labelling and subsequent purification were performed using the Invitrogen™ Alexa Fluor™ 647 Protein Labelling Kit (Thermo Fisher Scientific; Waltham, MA, USA), following the manufacturer's instructions. The conjugation reaction was performed in the dark at room temperature for 1 hour with gentle stirring. In this reaction, both MjsHSP nanocages samples with a total protein concentration ranging from 1.1 to 1.3 mg mL^{-1} were used. Purification of the protein conjugates was carried out by injecting each reaction mixture into a size exclusion chromatographic resin, which allowed the separation of the labelled MjsHSP nanocages from the unbound fluorophore. Both fractions were collected for the two samples. When not used immediately, the samples of MjsHSP nanocages labelled with Alexa Fluor 647 were stored protected from the light at 4°C.

At the end, the absorption spectrum between 250 and 750 nm of the protein conjugate for both samples was acquired against a reference PBS solution (1x) by spectrometry ultraviolet/visible (UV-Vis) using quartz cuvettes with an optical path of 1 cm. The protein samples used were diluted with PBS solution (1x) considering a dilution factor of 1:20. With the absorbance values at 280 nm and 650 nm, it was possible to determine the degree of labelling.

In the FCS measurements, the fractions containing the labelled MjsHSP nanocages and the unbound Alexa 647 fluorophore for both samples were diluted with PBS solution (1x) considering a dilution factor of 1:2000. Measurements were also performed on samples as negative control, the PBS solution (1x) and the unlabeled MjsHSP nanocages samples.

The FCS measurements were performed with a MicroTime 200 time-resolved confocal fluorescence microscope (PicoQuant; Berlin, Germany), which has single molecule sensitivity. An acquisition time of 180 seconds at room temperature was considered for all the analysis. Considering the characteristics of the Alexa Fluor 647 used in the conjugation with the MjsHSP nanocages, the FCS analysis was performed with a pulsed diode laser at 635 nm (LDH-635; PicoQuant; Berlin, Germany), operating at a repetition rate of 20 MHz by a PDL 800-B single channel picosecond diode laser driver.

The intensity of the laser light was defined to 30000 arbitrary units (which corresponded to an optical power at the objective exit of approximately 15 kW cm^{-2} for this objective and laser). Fluorescence emitted was spectrally selected through the dichroic mirror and a bandpass filter with transmission between 668 and 723 nm. For the confocal detection, it was used a pinhole with a diameter of 50 μm, which is placed in the path to a beam splitter. The signal detected was digitized using a TimeHarp 200 TCSPC PC board (PicoQuant; Berlin, Germany) in a time-tagged time resolved (TTTR) mode. The FCS experiments for all the samples including the calibration measurements were performed with the laser beam focused approximately 10 μm deep into the sample solution of 3 μL deposited on a microscope coverslip (thickness of 0.15 mm). Finally, the SymPhoTime software (PicoQuant; Berlin, Germany) was used for the data acquisition and subsequent analysis.

V.2.5.2. Production in *V. natriegens* cells

V.2.5.2.1. Dynamic light scattering

DLS experiments for particle size measurement applied to the samples containing MjsHSP nanocages produced in *V. natriegens* Vmax™ X2 and purified by different approaches were performed using the same experimental procedure established for these protein nanocages originating from *E. coli* BL21(DE3) cells (sub-section **V.2.5.1.1. Dynamic light scattering** in **V.2.5.1. Production in *E. coli* cells**, section **V.2.5. Characterization of MjsHSP nanocages**).

V.2.5.2.2. Transmission electron microscopy

Negative staining TEM was performed on samples containing MjsHSP nanocages produced in *V. natriegens* Vmax™ X2 and purified using different approaches. The experimental procedure applied was identical to that used for the same protein nanocages originally produced in *E. coli* BL21(DE3) cells (sub-section **V.2.5.1.2. Transmission electron microscopy** in **V.2.5.1. Production in *E. coli* cells**, section **V.2.5. Characterization of MjsHSP nanocages**).

As previously mentioned, the negative staining TEM was performed at the HEMS core facility at i3S, Universidade do Porto, Portugal.

V.2.6. Characterization of TRAP O-rings and nanocages

V.2.6.1. Dynamic light scattering

DLS experiments for particle size measurement were performed on two types of samples: TRAP O-rings produced in *E. coli* BL21(DE3) and purified by different approaches, and *in vitro* assembled TRAP nanocages. The experimental methodology applied was similar to that used for the MjsHSP nanocages also produced in *E. coli* BL21(DE3) cells (sub-section **V.2.5.1.1. Dynamic light scattering** in **V.2.5.1. Production in *E. coli* cells**, section **V.2.5. Characterization of MjsHSP nanocages**). The

only exception was the buffer solution used, 50 mM Tris-HCl, 150 mM NaCl, pH 7.9 instead of the PBS solution (1x).

V.2.6.2. Transmission electron microscopy

Negative staining TEM was implemented to two types of samples: TRAP O-rings produced in *E. coli* BL21(DE3) and purified by different approaches, and *in vitro* assembled TRAP nanocages. It was used the same experimental procedure established for the MjsHSP nanocages also produced in *E. coli* BL21(DE3) (sub-section **V.2.5.1.2. Transmission electron microscopy** in **V.2.5.1. Production in *E. coli* cells**, section **V.2.5. Characterization of MjsHSP nanocages**). Again, the only exception was the buffer solution used, 50 mM Tris-HCl, 150 mM NaCl, pH 7.9 instead of the PBS solution (1x).

As previously mentioned, the negative staining TEM was performed at the HEMS core facility at i3S, Universidade do Porto, Portugal.

V.2.6.3. Fluorescence correlation spectroscopy

For FCS measurements, two samples were labelled with the Alexa 647 maleimide fluorophore, which has absorption and fluorescence emission maxima of approximately 650 nm and 668 nm, respectively. The first sample consisted of TRAP O-rings purified after the CEX chromatography step, while the second sample contained TRAP nanocages formed through the *in vitro* assembly reaction of TRAP O-rings in the presence of gold ions.

The Alexa 647 fluorophore solution was prepared according to the manufacturer's instructions (Thermo Fisher Scientific; Waltham, MA, USA), with concentrations ranging between 100 nM and 100 µM, depending on the tests performed. The samples containing TRAP O-rings and TRAP nanocages were prepared at concentrations ranging from 10 nM to 10 µM. The conjugation reaction was performed under gentle agitation in the dark at room temperature for 6 hours or overnight at 4°C. For the preparation of the TRAP O-rings reaction, an excess molar amount of TCEP (stock solution at 10x concentration) relative to the molar concentration of protein O-rings was added as a reducing agent. The fluorophore was added to each protein sample in a molar ratio of 10x excess fluorophore relative to the protein monomers concentration.

Several purification methods for the protein conjugates were tested, including Sephadex G-25 columns (Cytiva; Marlborough, MA, USA), PD SpinTrap G-25 columns (GE Healthcare; Chicago, IL, USA) and Amicon Ultra centrifugal filters (Sigma-Aldrich; St. Louis, MO, USA), all following the respective manufacturers' instructions. For the optimized procedure, Amicon Ultra centrifugal filters were selected as they proved to be more effective. Each labeled sample was passed through centrifugal filters with a molecular weight cut-off of 3 kDa eight times, followed by two passes through 100 kDa centrifugal filters for the samples containing TRAP nanocages. Prior to use, the centrifugal filters were passivated overnight at room temperature with 5% v/v Tween 20 and subsequently washed with Milli-Q water and equilibrated with 50 mM Tris-HCl, 150 mM NaCl, pH 7.9. All labeling procedures utilized this last buffer solution.

The actual protein concentrations in TRAP O-rings and nanocages samples were determined by measuring absorbance values at 280 nm and 650 nm using UV-Vis spectrometry with a quartz cuvette (optical path: 1 cm). During sample preparation and quantification, as well as FCS measurements, when necessary, samples were diluted using either the buffer solution mentioned above or Milli-Q water.

Measurements were also performed on negative control samples, including Milli-Q water, buffer solution, and unlabeled TRAP O-rings and nanocages samples.

Additionally, experiments were conducted to evaluate the mechanisms of assembly and disassembly. In one experiment, labeled TRAP O-rings were used for an assembly reaction (following procedure described in sub-section **V.2.4. Assembly of TRAP nanocages**), which was monitored via FCS measurements during 5 days of incubation. In another experiment, TRAP nanocages were treated with TCEP (approximately 10x molar concentration relative to protein molar concentration) as a reducing agent to break bonds between O-rings within the nanocages, leading to partial or complete disassembly. This second sample was similarly monitored by FCS over 5 days of incubation.

The FCS measurements were conducted using a MicroTime 200 time-resolved confocal fluorescence microscope (PicoQuant, Berlin, Germany), using an experimental procedure similar to that applied in the analysis of MjsHSP nanocages produced in *E. coli* BL21(DE3) (**V.2.5.1.4. Fluorescence correlation spectroscopy** in **V.2.5.1. Production in *E. coli* cells** of **V.2.5. Characterization of MjsHSP nanocages**). All analyses were performed at room temperature with an acquisition time of 180 or 360 seconds. Given the properties of Alexa Fluor 647 used for conjugating TRAP O-rings and nanocages, the FCS analysis utilized a pulsed diode laser at 635 nm (LDH-635; PicoQuant), operating at a repetition rate of 20 MHz via a PDL 800-B picosecond diode laser driver. The laser intensity was set to 30000 arbitrary units, corresponding to an optical power of approximately 15 kW cm^{-2} at the objective exit. Fluorescence emission was filtered spectrally using a dichroic mirror and a bandpass filter transmitting between 668 and 723 nm. A pinhole with a diameter of 50 µm was employed for confocal detection, positioned along the path to a beam splitter. The detected signal was digitized in TTTR mode using a TimeHarp 200 TCSPC PC board (PicoQuant; Berlin, Germany). The FCS experiments, including calibration measurements, were conducted by focusing the laser beam approximately 10 µm into the sample solution (3 µL) deposited on a microscope coverslip with a thickness of 0.15 mm. Data acquisition and subsequent analysis were performed using SymPhoTime software (PicoQuant; Berlin, Germany).

V.2.7. Analytics

V.2.7.1. Total protein quantification

Total protein concentration was determined by the bicinchoninic acid (BCA) method, according to the experimental procedure described in detail in section **III.2.14.1. Total protein quantification** of **Chapter III**.

V.2.7.2. SDS-PAGE

Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) was implemented to evaluate the composition of the samples both in specific protein nanocage (monomers) and total protein. The experimental procedure used is described in detail in section **III.2.14.2. SDS-PAGE** of **Chapter III**.

V.2.8. Statistical analysis

Descriptive statistics, including mean and standard deviation, were calculated for the results obtained across the different variables. The sample sizes for each variable typically ranged between 2 and 3 experiments.

V.3. Results and Discussion

V.3.1. Characterization of MjsHSP nanocages produced in *E. coli* cells

V.3.1.1. Dynamic light scattering

To analyze the biophysical properties of NVPNs, one of the well-established techniques is DLS. This methodology is able to measure the size distribution of particles in suspension or solution by analyzing the fluctuations in scattered light caused by the Brownian motion of these particles (242). Its operating principle consists of shining a laser onto the sample and detecting the scattered light intensity fluctuations over time. These fluctuations are faster for smaller particles due to their faster diffusion, allowing DLS to determine particle sizes based on the rate of these oscillations.

For the MjsHSP nanocages produced in *E. coli* BL21(DE3), four distinct samples were analyzed by DLS, selected at various stages during the downstream processing of these protein-based nanoparticles, which was explored and developed in **Chapter IV**. The selected samples included fractions obtained after the AEX chromatography step, specifically those where the MjsHSP nanocages eluted; after concentration using centrifugal filtration of this sample derived from AEX chromatography; and after the SEC purification step, specifically fractions corresponding to the two main peaks where MjsHSP nanocages were eluted.

In terms of experimental conditions for particle size measurements, details regarding sample preparation are available in sub-section **V.2.5.1.1. Dynamic light scattering** in **V.2.5.1. Production in *E. coli* cells** of **V.2.5. Characterization of MjsHSP nanocages**. It is important to highlight that results were reported as intensity and number distributions. The presence of peaks corresponding to particles with dimensions exceeding 1000 nm likely represented waste particles, such as dust present in the samples, even after procedures aimed at avoiding contamination with such environmental particles. Data related to these large particles were excluded from the analysis of the results.

Regarding the sample from the fractions post-AEX chromatography containing MjsHSP nanocages (**Chapter IV**, **Figure IV.4A**, peak *2*), a polydispersity index (PDI) of 0.74 was obtained, indicating a heterogeneous population (242). This heterogeneity was reflected in the intensity distribution, which revealed two main peaks corresponding to particle dimensions of 76.3 ± 23.3 nm and 21.5 ± 6.7 nm. The analysis of the number distribution allowed the identification of a single peak with particle sizes of 14.1 ± 1.0 nm.

For the sample from the fractions post-centrifugal filtration containing concentrated MjsHSP nanocages (**Chapter IV**) a PDI of 0.67 was obtained, suggesting a slightly heterogeneous population (242). The intensity distribution showed two principal peaks with particle sizes of 127.9 ± 90.8 nm and 21.5 ± 7.3 nm. The corresponding number distribution analysis identified a single peak with particle dimensions of 11.9 ± 1.0 nm.

Regarding the sample from the first fractions post-SEC purification containing theses protein nanocages (**Chapter IV**, **Figure IV.5A**, peak *1*), a PDI of 0.41 was obtained, indicating a population closer to homogeneity (242). The intensity distribution revealed two main peaks with particle dimensions

of 155.5 ± 37.3 nm and 21.01 ± 3.3 nm, while the number distribution analysis identified a single peak with particle dimensions of 16.2 ± 0.5 nm.

Finally, for the sample from the second fractions post-SEC purification containing purified MjsHSP nanocages (**Chapter IV**, **Figure IV.5A**, peak *2*), a PDI of 0.27 was obtained, indicating a highly homogeneous population (242). The intensity distribution showed a single principal peak with particle dimensions of 14.5 ± 4.1 nm, and the corresponding number distribution analysis identified a single peak with particle dimensions of 10.0 ± 0.1 nm.

In all samples, a similar behavior was observed in terms of detecting two populations in the intensity distribution and only one in the number distribution. For the intensity distribution, there was a trend of shifting the major contribution toward the peak corresponding to smaller dimensions, increasingly closer to the expected size for MjsHSP nanocages as purification steps progressed. The exception was the sample corresponding to peak *2* from SEC, where only a single population was detected in both the intensity and number distributions. It is also worth noting that, in the intensity distribution, the peak related to larger particles exhibited a significantly higher standard deviation, suggesting that this result was less consistent and likely associated with contaminants or aggregates.

These results suggested that the samples were predominantly composed of a population of particles with an average hydrodynamic diameter 12.9 ± 2.4 nm, which was compatible with the MjsHSP nanocages described as having a dimension of 12 nm. However, the sample was not homogeneous and did not consist solely of this population, as it also contained particles or aggregates with larger dimensions than 40 to 50 nm, which significantly contributed to the total scattered light intensity. On the other hand, the distribution based on number was influenced by smaller particles, which aligned with this observation and the expected results. As the sample was purified, particles related to impurities were removed, and consequently, they no longer contributed to the light scattering in the analyzed sample.

Focusing only on the samples from post-SEC, there was clearly a difference between peak *1* and peak *2*, with the second peak showing a population closer to homogeneity (**Chapter IV**, **Figure IV.5A**). In contrast, peak *1* consisted of a mix of a population of MjsHSP nanocages with the expected size and other larger particles that contribute to an additional signal in terms of intensity distribution. These results may support the hypothesis of the presence of some type of protein nanocages aggregates or nanocages combined with other impurity particles in peak *1*, while peak *2* predominantly contains a single purified population of MjsHSP nanocages.

A decrease in the PDI was also observed as the purification process advanced, suggesting that the samples became increasingly homogeneous. This supported the above hypothesis that the MjsHSP nanocages were the dominant population. As impurities were removed, other populations ceased to be measured by DLS, leaving only this type of NVPNs. It is also worth noting that these determinations assumed that the MjsHSP nanocages are spherical structures with a filled interior. Nevertheless, this assumption represents only an approximation and does not fully reflect the true structure of these protein nanocages.

Still related to DLS, some measurements of the zeta potential were performed. For these measurements, only samples from post-AEX chromatography, post-concentration by centrifugal

filtration, and post-SEC purification (specifically fractions from peak 2 of MjsHSP nanocages elution) were considered.

For the sample corresponding to the fractions post-AEX chromatography containing these protein nanocages (**Chapter IV**, **Figure IV.4A**, peak *2*), a zeta potential of -29.2 ± 10.2 mV was obtained, measured in a buffer solution with a pH of 6.0 and an estimated conductivity of 3.48 mS cm^{-1}. For the sample corresponding to the fractions post-centrifugal filtration containing concentrated MjsHSP nanocages, a zeta potential of -21.2 ± 16.8 mV was obtained, measured in a buffer solution with a pH of 6.0 and an estimated conductivity of 3.34 mS cm^{-1}. For the sample corresponding to the second fractions from the post-SEC purification containing MjsHSP nanocages (**Chapter IV**, **Figure IV.5A**, peak *2*), a zeta potential of -3.2 ± 7.4 mV was obtained, measured in a buffer solution with a pH of 7.4 and an estimated conductivity of 3.90 mS cm^{-1}.

These preliminary zeta potential results should be further explored in future studies, where evaluations of pH dependence could be conducted to understand how the zeta potential varies with pH and to experimentally identify the pI of MjsHSP nanocages, comparing it with the theoretical value of 5.02. Additionally, it may be interesting to perform stability assessments over time to determine whether prolonged periods under certain pH and conductivity conditions result in variations in zeta potential and to assess the magnitude of such variations.

V.3.1.2. Transmission electron microscopy

The second well-established technique for analyzing and structurally characterizing NVPNs is TEM using negative staining. Visualizing the structural details of protein nanocages is achieved by enhancing contrast with an electron-dense stain, which creates a dark background around the sample.

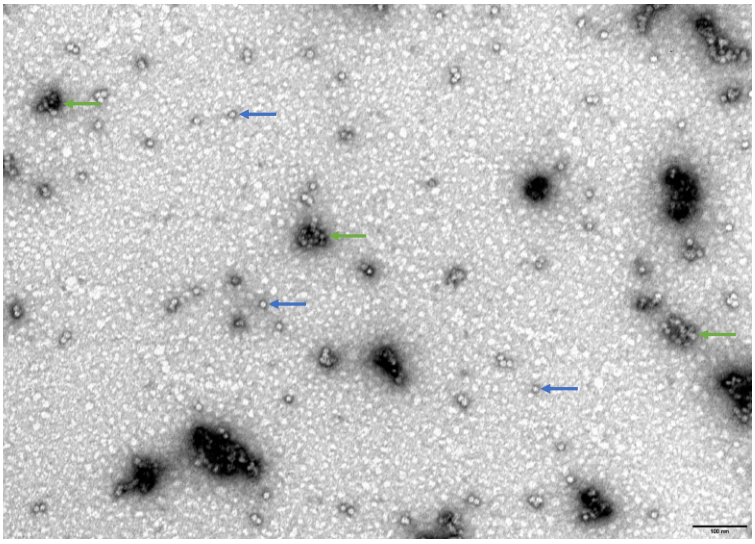
Similarly to what was done in the DLS measurements of MjsHSP nanocages, three distinct samples were analyzed by negative staining TEM, selected at various stages during the downstream processing of these nanocages, which was studied in **Chapter IV**. The selected samples comprised fractions obtained after concentration using centrifugal filtration of this sample derived from AEX chromatography; and after the SEC purification step, specifically fractions corresponding to the two main peaks where MjsHSP nanocages were eluted.

In terms of experimental conditions for particle size measurements, details regarding sample preparation and subsequent procedure are available in sub-section **V.2.5.1.2. Transmission electron microscopy** in **V.2.5.1. Production in *E. coli* cells** of **V.2.5. Characterization of MjsHSP nanocages**.

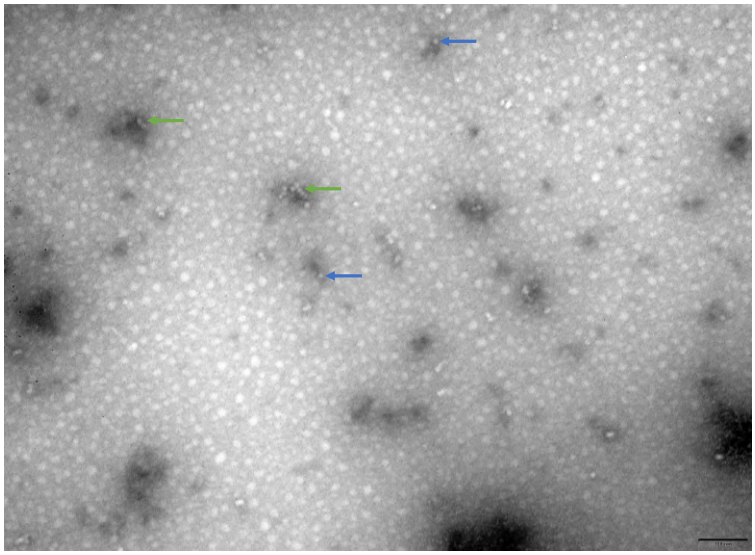
The corresponding negative staining TEM images obtained are shown in **Figure V.1**, clearly identified according to the stage of the downstream processing at which they were collected.

The analysis of these results showed that, across the three stages of downstream processing, it was possible to visualize structures with morphology and dimensions compatible with the MjsHSP nanocages. Protein nanocages aggregates were observed in all samples, which was expected. However, individual structures were also visible, confirming not only the presence of protein nanocages but also their biophysical characteristics in terms of morphology and dimensions.

(A)



(B)



(C)

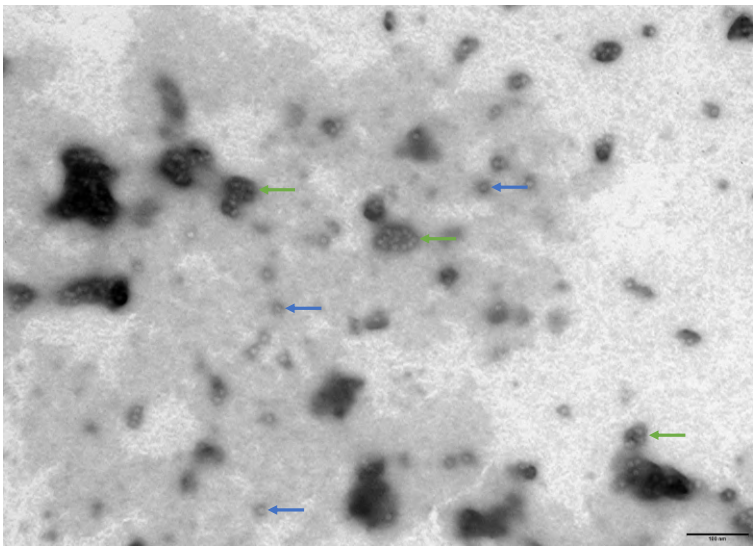


Figure V.1 – Negative staining TEM images of the MjsHSP nanocages produced in *E. coli* BL21(DE3) at distinct stages of the downstream processing. The three samples were collected from the fractions post-centrifugal filtration containing concentrated protein nanocages **(A)**, from the peak *1* fractions post-SEC purification **(B)**, and from the peak *2* fractions post-SEC purification **(C)**. Blue arrows identify individual MjsHSP nanocages, where structural details can even be observed. Green arrows indicate potential aggregates of MjsHSP nanocages. Scale bars: 100 nm.

It is worth noting the presence of some elements in the images that may either be naturally present in the sample or artifacts resulting from its preparation, particularly related to the use of negative staining. These artifacts could include uneven staining or background noise that might interfere with the clarity of the observed structures. Despite these limitations, the analysis provided valuable insights into the structural integrity and distribution of MjsHSP nanocages.

Regarding particle dimensions, for the sample obtained after centrifugal filtration, particles with a hydrodynamic diameter of 14.2 ± 0.9 nm were detected. For the sample corresponding to peak *1* post-SEC, several particles exhibited a hydrodynamic diameter of 12.8 ± 0.8 nm, while for the sample from peak *2* post-SEC, the average hydrodynamic diameter was 11.3 ± 1.5 nm. All these values were in complete agreement with the theoretical dimensions calculated from the predicted model for MjsHSP nanocages, which estimate an external diameter of approximately 12 nm.

V.3.1.3. Atomic force microscopy

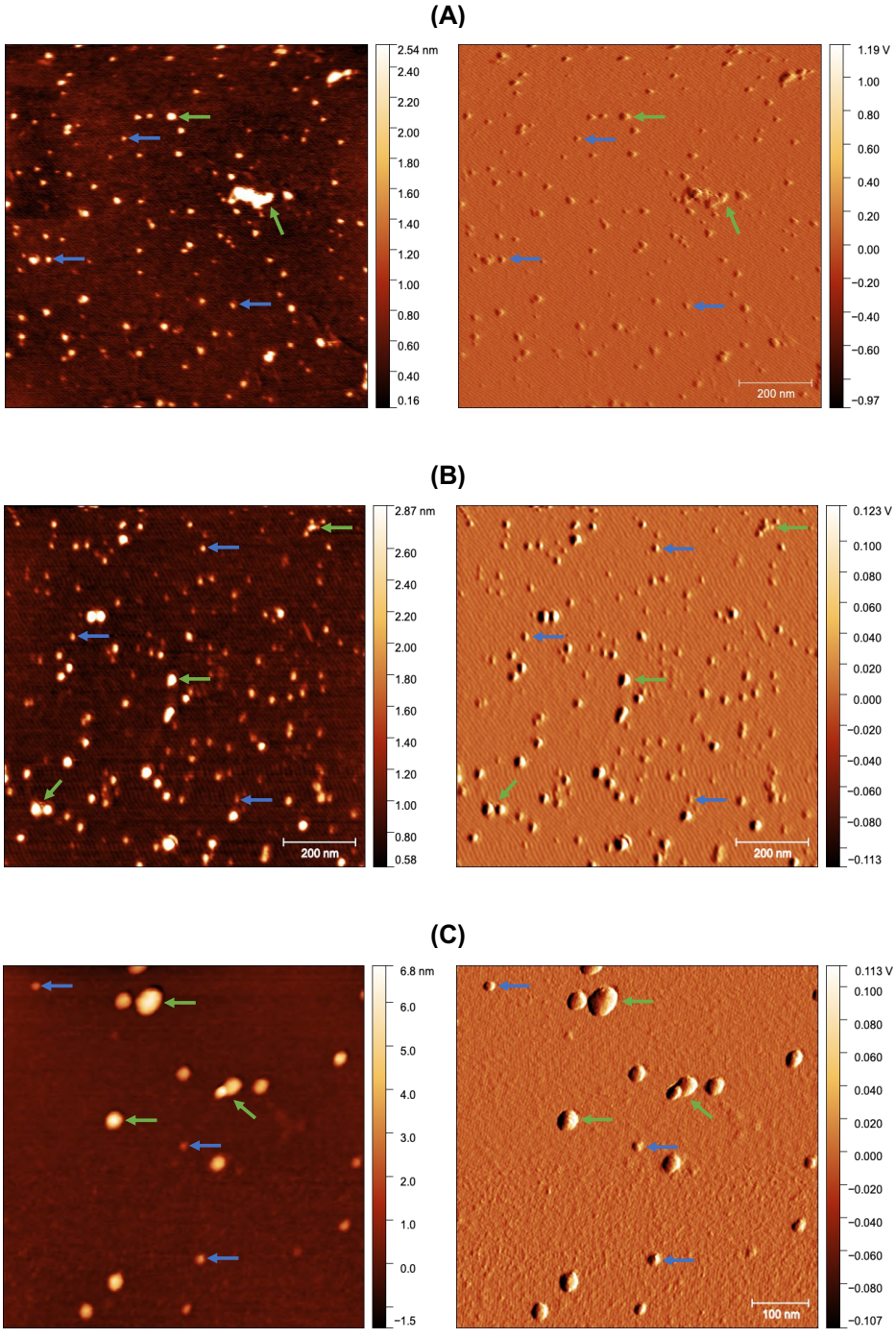
A less common technique used to structurally characterize NVPNs is AFM. AFM is a high-resolution imaging method that uses a sharp probe attached to a cantilever to scan surfaces at the nanoscale, providing detailed topographical maps and enabling the characterization of the physical and mechanical properties of the protein nanocages sample. In this particular case, the main objective was topographical analysis, specifically to explore the surface morphology and structural details of MjsHSP nanocages, including their size, shape, and assembly states.

All analyzed samples were obtained after the SEC purification step, specifically from the fractions containing MjsHSP nanocages from peak *2* (**Chapter IV**, **Figure IV.5A**).

One of the main challenges of AFM was sample preparation (details available in sub-section **V.2.5.1.3. Atomic force microscopy** in **V.2.5.1. Production in *E. coli* cells** of **V.2.5. Characterization of MjsHSP nanocages**). This is because AFM analysis required a sample that was as clean as possible, containing only the structures to be visualized. Any contaminant particles can interfere with the process and make it even more time-consuming. For this reason, a highly purified sample containing MjsHSP nanocages was used. Another critical aspect was the concentration of the sample, which must fall within the nanomolar range. Several attempts were made to dilute the samples to achieve concentrations within this range. Additionally, another challenge was related to the adsorption of MjsHSP nanocages onto the supporting substrate. In this case, mica was used as the substrate, and it was freshly cleaved immediately before analysis to ensure that it was in optimal condition. Difficulties arose regarding the adsorption of MjsHSP nanocages, either they were not adsorbing at all, resulting in no visualization, or movements of the nanostructures were detected during measurement, rendering the analysis

unfeasible. To address this issue, various buffers were tested, including those containing $MgCl_2$ and $NiCl_2$, two salts described in the literature as effective aids for this adsorption process. Additionally, attempts were made to passivate the mica with $MgCl_2$ before applying the sample. The salt concentration in the solutions also had to be carefully controlled since excessive salt could remain in the sample after water evaporation and interfere with image acquisition. The use of alternative substrates, such as silicon, was also evaluated as part of these efforts to optimize sample preparation and improve adsorption for successful AFM imaging.

In **Figure V.2**, the main results obtained under the different tested conditions are presented. For each condition, both a topography image and the corresponding amplitude image are provided.



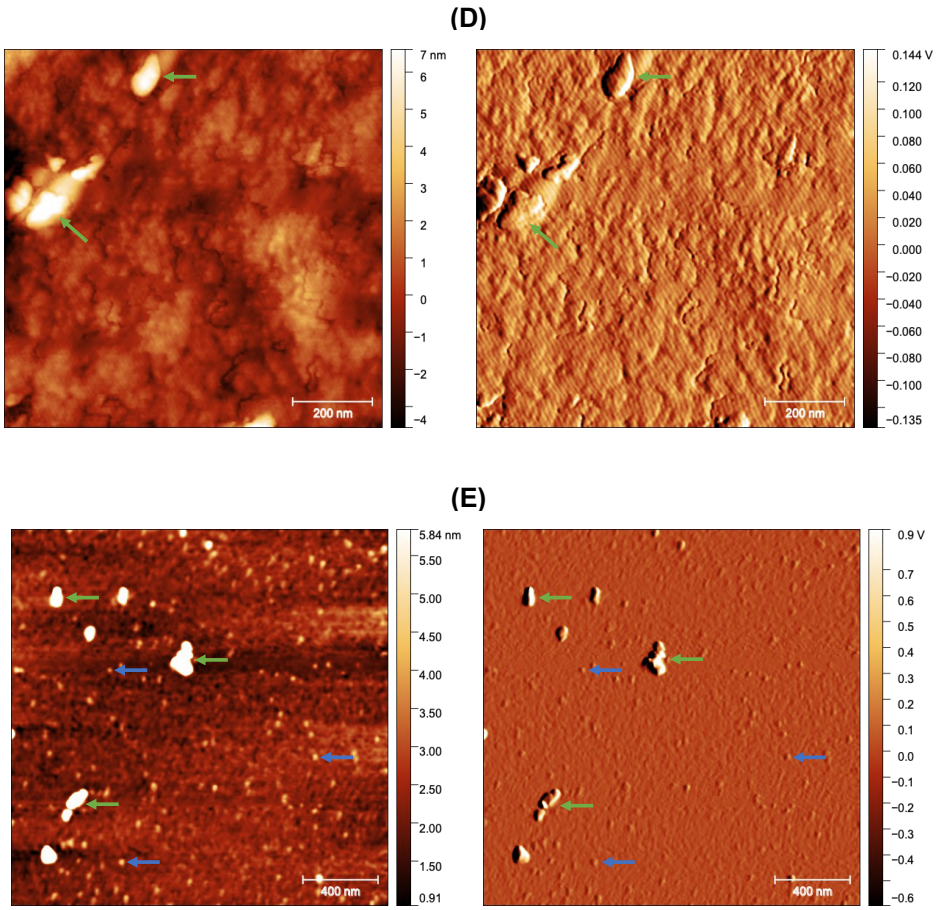


Figure V.2 – AFM analysis of MjsHSP nanocages produced in *E. coli* BL21(DE3) at the final stage of the downstream processing. The five samples analyzed were MjsHSP nanocages in an air environment at a molar concentration of 1 nM in a PBS solution (1x) **(A)**, MjsHSP nanocages in an aqueous environment at a molar concentration of 1 nM in a PBS solution (1x) **(B)**, MjsHSP nanocages at a molar concentration of 10 nM in PBS solution (1x), with mica as supporting substrate passivated with $MgCl_2$ prior to sample application, in an aqueous environment **(C)**, MjsHSP nanocages at a molar concentration of 10 nM in a Tris-based buffer containing $MgCl_2$, in an aqueous environment **(D)**, and MjsHSP nanocages at a molar concentration of 10 nM in a Tris-based buffer containing $NiCl_2$, in an aqueous environment **(E)**. For each sample, both a topography image and the corresponding amplitude image are presented. Blue arrows identify individual MjsHSP nanocages. Green arrows indicate potential aggregates of MjsHSP nanocages. Scale bars: 100 nm, 200 nm and 400 nm.

Initially, analyses were performed on protein nanocages samples in air at a concentration of 1 nM in a PBS solution (1x). As shown in **Figure V.2A**, several structures with dimensions compatible with MjsHSP nanocages were observed, recording diameters of 16.3 ± 1.9 nm. Although the images did not provide sufficient resolution to visualize detailed structural features, the general topography was visible. However, visualizing these samples in an air environment leads to water droplets surrounding the pores and protein structures, hindering clear nanocage characterization.

Subsequently, the same samples at 1 nM concentration in PBS solution (1x) were analyzed in an aqueous environment within the same buffer (**Figure V.2B**). The results were quite similar, with the identification of several structures with compatible dimensions (estimated diameter of 19.7 ± 1.8 nm)

and a morphology consistent with that of MjsHSP nanocages. On the other hand, a rather heterogeneous population was observed in terms of dimensions, including some structures with sizes larger or smaller than expected.

The next sample consisted of MjsHSP nanocages at a concentration of 10 nM in PBS solution (1x), with mica passivated with $MgCl_2$ prior to sample application, in an aqueous environment (**Figure V.2C**). This sample exhibited structures with dimensions significantly larger than typical nanocages (greater than 30 nm), as well as structures with sizes that are highly compatible (estimated dimension of 19.0 ± 2.3 nm). This observation suggests that the larger structures may be aggregates formed by two or three nanocages. Due to their close proximity, the AFM may lack the resolution to distinguish them as individual structures.

Next, a sample of MjsHSP nanocages at a concentration of 10 nM in a Tris-based buffer containing $MgCl_2$, in an aqueous environment, was analyzed to assess the impact of this salt on the measurement. The results were not very favorable, as shown in **Figure V.2D**, where a background layer was visible across the entire sample analysis. Only a few distinguishable structures were identified, but they exhibited dimensions significantly larger than expected for individual nanocages (greater than 50 nm). These larger structures may potentially be aggregates of the NVPNs, and as previously discussed, AFM may lack the resolution to present them as individual structures due to their close proximity.

The final sample consisted of MjsHSP nanocages at a concentration of 10 nM in a Tris-based buffer containing $NiCl_2$, in an aqueous environment (**Figure V.2E**). Some structures with sizes significantly larger than expected were observed, which could once again correspond to aggregates of nanocages, as also noted in the TEM images (**V.3.1.2. Transmission electron microscopy** of **V.3.1. Characterization of MjsHSP nanocages produced in *E. coli* cells**). These larger structures exhibited dimensions exceeding 60 nm. On the other hand, smaller structures were also observed, which presented dimensions closer to the expected size for the target MjsHSP nanocages (estimated sizes around 20 nm), showing greater potential for being individual nanocages.

In addition to further developing and replicating these experiments, efforts should focus on improving not only the sample preparation conditions, one of the critical aspects of this technique for achieving an optimal process, but also on enhancing image quality to better visualize and distinguish individual nanocages. It would also be interesting to use AFM for complementary studies that are typically applied to other types of biological components. These could include investigations into the mechanical properties of NVPNs, such as stiffness, elasticity, and deformation behavior, which could provide valuable insights into their structural integrity. Furthermore, AFM could be utilized to study protein-protein interactions, which would be particularly relevant for these protein nanocages since their assembly is based on such interactions. Measuring intermolecular forces and studying interaction dynamics could offer a deeper understanding of the mechanisms underlying their assembly and stability.

V.3.1.4. Fluorescence correlation spectroscopy

Unlike the three techniques more commonly described in the literature, fluorescence-based approaches have more limited published data. For this reason, FCS was investigated in this work for

MjsHSP nanocages after their labeling with an appropriate fluorophore. FCS is a methodology that uses time-dependent fluctuations in fluorescence intensity to determine several key physical and chemical parameters. This technique enables the quantification of translational and rotational diffusion coefficients, which can be used to infer hydrodynamic diameters.

Two distinct samples were analyzed by FCS, both obtained after the SEC purification step. These samples corresponded to fractions containing MjsHSP nanocages from peak *1* (referred to as S2) and peak *2* (referred to as S1) (**Chapter IV**, **Figure IV.5A**).

The fluorescent molecule used for labelling these protein nanocages was the Alexa 647 fluorophore, conjugated via an NHS ester, which covalently binds to primary amines on the protein structure. Details regarding the experimental procedure used are available in sub-section **V.2.5.1.4. Fluorescence correlation spectroscopy** in **V.2.5.1. Production in *E. coli* cells** of **V.2.5. Characterization of MjsHSP nanocages**.

With the purpose of determining the labeling efficiency of the MjsHSP nanocages with Alexa 647 fluorophore, the absorption of the labelled fraction for each of the samples S1 and S2 was measured between 250 and 750 nm (**Figure V.3**).

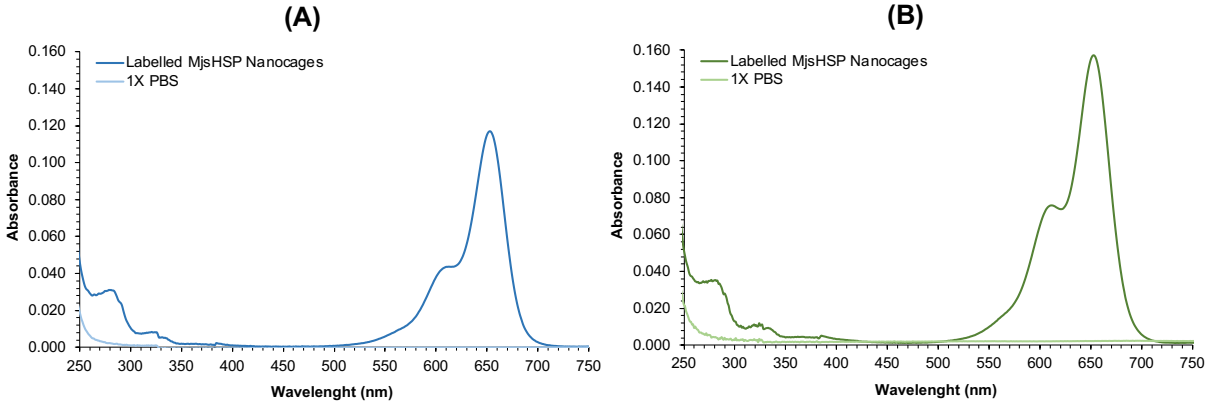


Figure V.3 – Absorption spectra between 250 and 750 nm of the fraction containing the labelled MjsHSP nanocages from samples S1 **(A)** and S2 **(B)**. The baseline spectra (PBS solution at 1x concentration) are also shown.

It was possible to observe that the maximum absorption wavelength for Alexa 647 was at 650 nm as expected, thus confirming the suitability of this fluorophore to be excited with selected laser (LDH-635). On the other hand, the absorption of the MjsHSP nanocages corresponded to the spectral region around 280 nm.

For both samples S1 and S2, with the absorbance values of the protein conjugates at 280 nm ($Abs_{280\,nm}$) and at 650 nm ($Abs_{650\,nm}$), it was possible to first determine the protein concentration ($c_{protein}$) in each sample using **Equation V.1** (243):

$$c_{protein} = \frac{[Abs_{280\,nm} - Abs_{650\,nm} \times CF] \times DF}{\varepsilon_{protein}} \qquad \textbf{V.1}$$

where CF is a correction factor to considerer the absorption of Alexa 647 fluorophore at 280 nm (corresponds to the value of 0.03), DF is the dilution factor and $\varepsilon_{protein}$ is the molar absorptivity coefficient of the protein (MjsHSP nanocages) at 280 nm (≈ 203,520 M^{-1} cm^{-1}). This last value was determined using the ProtParam tool (128).

Next, the degree of labelling (moles fluorophore per mole protein) was calculated through the **Equation V.2** (243):

$$Degree\ of\ labelling = \frac{Abs_{650\,nm} \times DF}{\varepsilon_{fluorophore} \times c_{protein}} \qquad \textbf{V.2}$$

where $\varepsilon_{fluorophore}$ is the approximate molar absorptivity coefficient of the Alexa 647 fluorophore at 650 nm (≈ 239,000 M^{-1} cm^{-1}).

Table V.1 summarizes the values considered for each of the variables requested in equations before that were not previously indicated as well as the final results of the MjsHSP nanocages concentration and the respective degree of labelling obtained for the two samples S1 and S2.

The degree of labelling for both samples S1 and S2 corresponded to approximately 4 moles of Alexa 647 fluorophore per mole of MjsHSP nanocages, which is in accordance with the recommended range (3 to 7 moles fluorescent dye per mole protein) (243).

Table V.1 – Summary of the values considered for each of the variables required for the determination of the MjsHSP nanocages concentration and the respective degree of labelling as well as the final results for these two parameters.

Sample	$Abs_{280\,nm}$	$Abs_{650\,nm}$	Dilution factor	MjsHSP nanocages concentration (M)	Degree of labelling
S1	0.0310	0.1152	20	2.7×10^{-6}	≈ 4
S2	0.0353	0.1548		3.0×10^{-6}	≈ 4

The use of FCS to physically characterize the labeled MjsHSP nanocages allowed the determination of the corresponding diffusion coefficient and subsequent inference of the hydrodynamic radius in the two types of samples. This should allow the verification of the correct assembly of the MjsHSP nanocages and also the evaluation of the presence of aggregates of these protein nanoparticles.

In all the FCS measurements, the time trace of fluorescence intensity (MCS-trace) acquired over the duration of the experiment was recorded and used to determine the respective autocorrelation function. The MCS-trace curves obtained for the fractions containing the labelled MjsHSP nanocages and the unbound Alexa 647 fluorophore for both samples S1 and S2 are shown in **Figure V.3** during a window of only 5 seconds from a total of 180 seconds used for signal acquisition.

Moreover, with the MCS-trace data, it was possible to build a fluorescence intensity histogram, which allowed observing the frequency in number of events of each intensity. The histograms obtained for the two fractions from samples S1 and S2 are shown in **Figure V.4**.

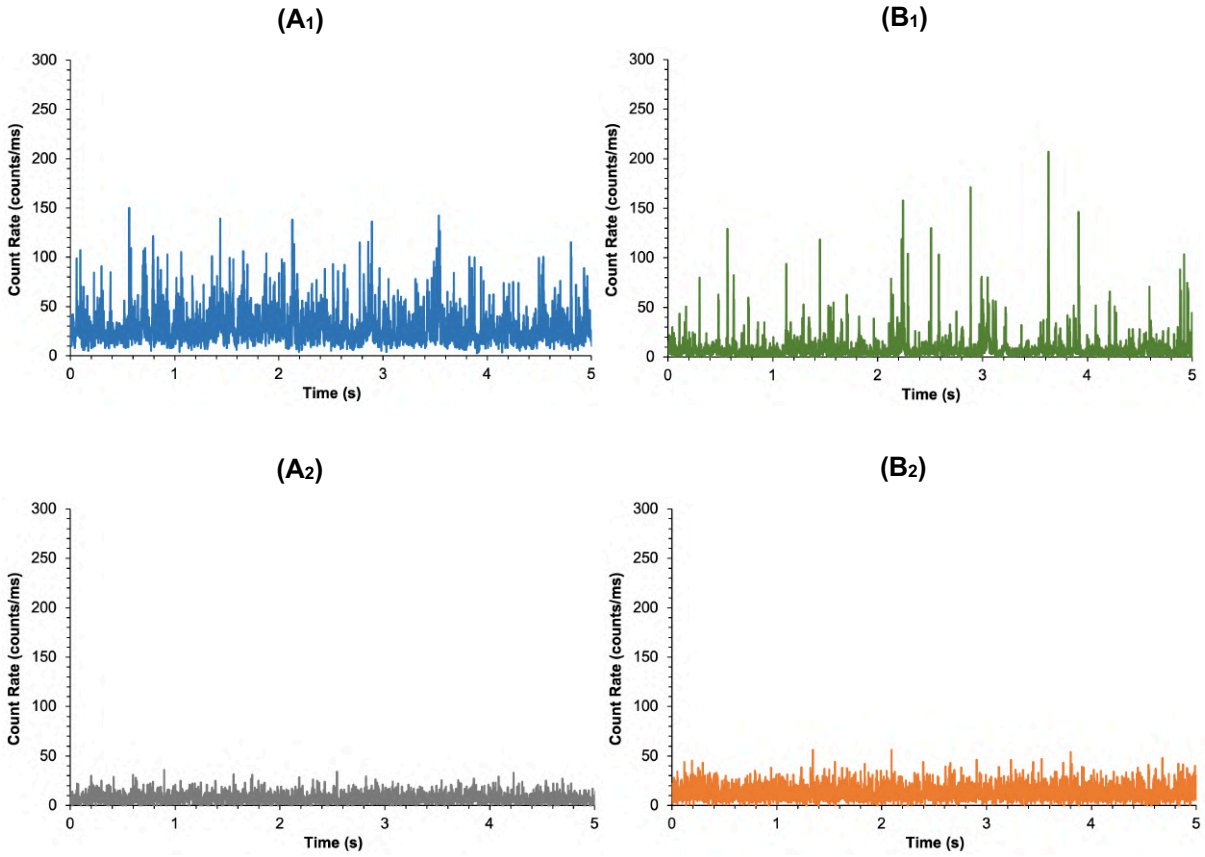


Figure V.4 – MCS-trace curves obtained for the fractions containing the labelled MjsHSP nanocages (**A_1** and **B_1**) and the unbound Alexa 647 fluorophore (**A_2** and **B_2**) from the samples S1 **(A)** and S2 **(B)**.

Analyzing the MCS-trace curves relative to the fraction containing the MjsHSP nanocages of samples S1 and S2, it was possible to observe a more regular and homogeneous fluorescent intensity pattern in the case of the sample S1, while in the sample S2 this pattern was not only more irregular but also displayed occasional peaks of considerable fluorescent intensity. This observation is compatible with the suggested hypothesis that some aggregates of MjsHSP nanocages were present in the sample S2, which, because of their larger size, should also contain a larger number of dye labels than individual MjsHSP nanocages, therefore, causing the observed large bursts of emission. The data from the fluorescence intensity histogram also evidences these events, since the highest detected intensities were present mainly in the case of the sample S2. In any case, these results also suggest that the population of MjsHSP nanocages aggregates was not the majority in sample S2, which could also be induced by the fact that the FCS measurements were performed at low concentrations, thereby, promoting sample disaggregation.

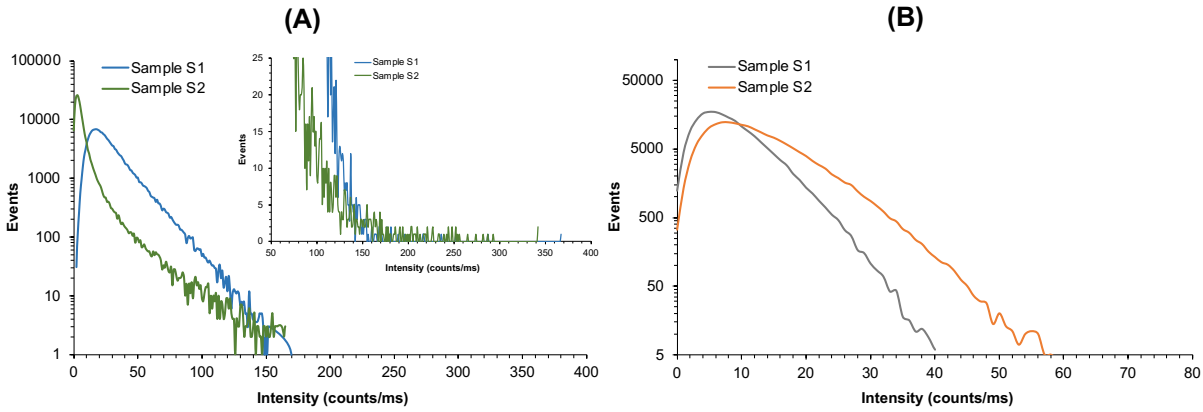


Figure V.5 – Fluorescence intensity histograms obtained for the fractions containing the labelled MjsHSP nanocages **(A)** and the unbound Alexa 647 fluorophore **(B)** from the samples S1 and S2. Inset in **(A)** shows an amplification of the respective histogram in the interval of higher intensity values (both axes with linear scale).

Regarding the MCS-traces curves and the fluorescence intensity histograms of the fraction constituted by the unbound Alexa 647 fluorophore from the samples S1 and S2, very similar results were obtained for both experiments with a regular intensity fluorescence pattern, which was expected.

With the MCS-trace data, the autocorrelation curves were determined for both fractions from the two samples S1 and S2 assuming a pure 3D Brownian diffusion model. The normalized results (experimental date and fitted model) are represented in **Figure V.6**.

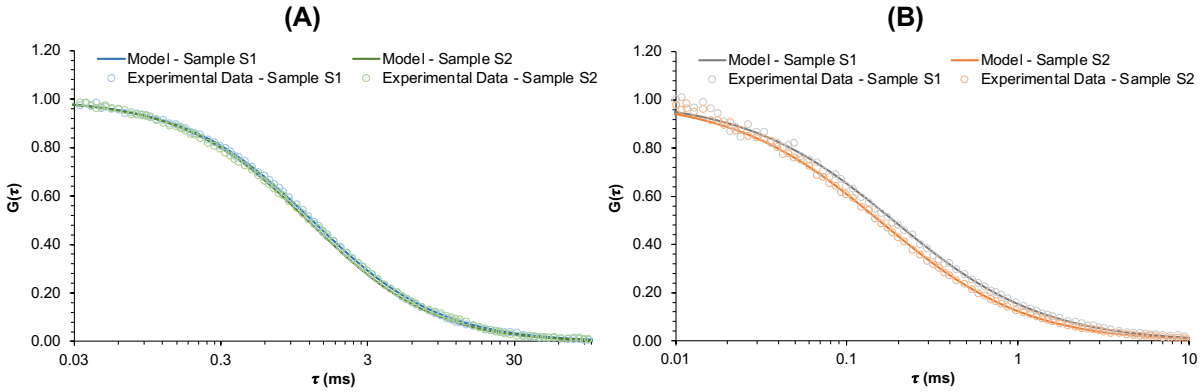


Figure V.6 – Autocorrelation curves obtained in the FCS measurements of the fractions containing the labelled MjsHSP nanocages **(A)** and the unbound Alexa 647 fluorophore **(B)** from the samples S1 and S2. Light colors correspond to the experimental data and dark colors correspond to the fitted model.

The analysis of the autocorrelation curves for the fraction comprising the MjsHSP nanocages of the samples S1 and S2 showed that there was a better fit between the experimental data and the model function in the case of the sample S1. This observation corroborates the conclusions mentioned above, i.e., that sample S1 is constituted by individual MjsHSP nanocages, whereas sample S2 contains aggregates, which lead to changes in the diffusion time due to their larger dimensions.

Concerning the autocorrelation curves of the fraction composed by the unbound Alexa 647 fluorophore of both samples S1 and S2, consistent results were also obtained for the two analyses, with a good fit between the experimental points and the model function.

The FCS measurements allowed to determine the diffusion time, the translational diffusion coefficient and, indirectly, the corresponding hydrodynamic radius of structures present in both fractions from the two samples S1 and S2. These results are summarized in **Table V.2**.

Knowing that there is a relation between the dimensions of a molecule and its respective diffusion rate, with the value of the translational diffusion coefficient, it is possible to determine the hydrodynamic radius (R_h) of that fluorescent species (an isolated fluorophore or a fluorophore-labelled molecule) through the Stokes-Einstein relation, which is shown in **Equation V.3**, where k is the Boltzmann constant, T is the temperature and η is the solvent viscosity. It is important to note that is assumed that the fluorescent molecule is a rigid sphere (244,245).

$$D_t = \frac{kT}{6\pi\eta R_h} \Leftrightarrow R_h = \frac{kT}{6\pi\eta D_t} \quad \textbf{V.3}$$

The value obtained for the hydrodynamic radius and the subsequent hydrodynamic diameter (approximately 12.0 nm) of MjsHSP nanocages in sample S1 agreed with the typical dimensions of the individual MjsHSP nanocages available in the literature (12 nm). On the other hand, the value obtained for the same parameter in the sample S2 was slightly higher, which, together with the other aspects already mentioned before in the analysis of the FCS data, reinforced the hypothesis that the sample S2 was constituted not only by the individualized MjsHSP nanocages but also by eventual protein aggregates.

Table V.2 – Summary of the values of the diffusion time, the translational diffusion coefficient and the hydrodynamic radius obtained for the fractions containing the labelled MjsHSP nanocages and the unbound Alexa 647 fluorophore from the samples S1 and S2.

Sample		τ_D (ms)	D_t (μm² s⁻¹)	R_h (nm)
S1	Fraction containing the labelled MjsHSP nanocages	1.3 ± 0.3	41	6.0
S2		1.2 ± 0.7	33	7.4
S1	Fraction containing the unbound Alexa 647 fluorophore	0.2 ± 0.1	274	0.90
S2		0.16 ± 0.09	278	0.88

Furthermore, it was also possible to analyze the value of the diffusion coefficient obtained for the fraction containing the unbound Alexa 647 fluorophore from the samples S1 (274 $\mu m^2\ s^{-1}$) and S2 (278 $\mu m^2\ s^{-1}$), which were close to the value available in the literature of 330 ± 10 $\mu m^2\ s^{-1}$ in water at 25°C (246). However, the difference verified could be explained by an eventual modification of the fluorophore

molecule due to the chain where the NHS ester group is located. One approach to evaluating this hypothesis would involve testing by FCS the Alexa 647 fluorophore completely pure and before any application.

Finally, with the experimental data obtained from the FCS measurements, the fluorescence decay curves for the fractions from the samples S1 and S2 were determined (**Figure V.7**).

For the fraction containing the MjsHSP nanocages from the samples S1 and S2, the excited-state decay kinetics was described by a bi-exponential function since the protein nanocages can create a microenvironment that affects the photodynamics of the fluorophore. Based on the respective fitting analysis, a value was obtained for the fluorescent intensity lifetime of 1.6 ns (χ^2 = 1.1) and 1.2 ns (χ^2 = 1.0) for the samples S1 and S2, respectively.

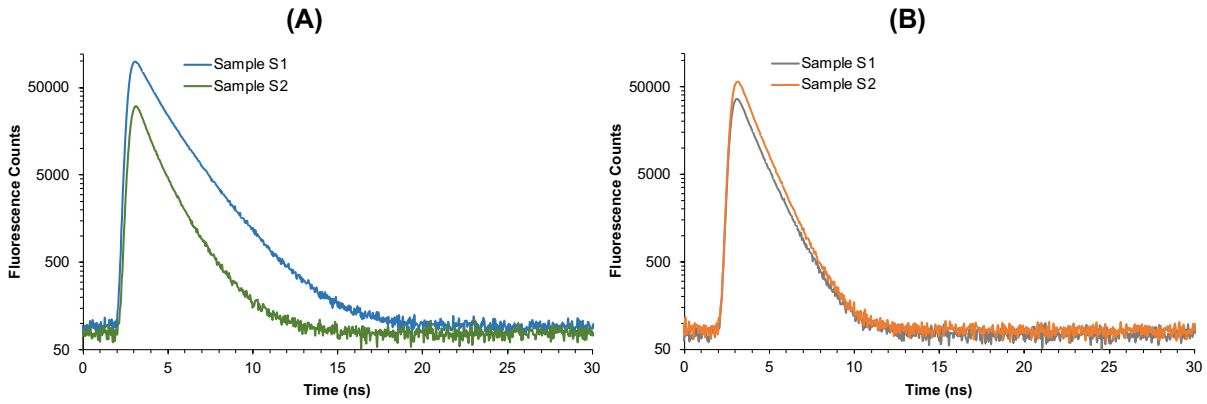


Figure V.7 – Fluorescent decay curves obtained for the fractions containing the labelled MjsHSP nanocages **(A)** and the unbound Alexa 647 fluorophore **(B)** from the samples S1 and S2.

On the other hand, in the fraction composed of the unbound Alexa 647 fluorophore from the sample S1, the fluorescence decay can be described by a mono-exponential function, since the free fluorophore was found in a more homogeneous medium. Considering the respective fitting analysis, a value for the fluorescent lifetime of 1.0 ns (χ^2 = 1.5) was obtained. In the case of the equivalent fraction in the sample S2, this decay was described by a bi-exponential function, although with a fluorescent lifetime of 1.0 ns (χ^2 = 1.1) obtained. These values are in full agreement with what is established in the literature for the fluorescence lifetime of this Alexa 647 fluorophore (in the order of 1 ns) (243).

V.3.2. Characterization of MjsHSP nanocages produced in *V. natriegens* cells

V.3.2.1. Dynamic light scattering

Similarly to what was performed for the MjsHSP nanocages derived from *E. coli* BL21(DE3), these protein nanocages produced in *V. natriegens* Vmax™ X2 and subsequently purified were also analyzed by DLS. Four different samples equivalent to those from *E. coli* were selected at various stages during

the downstream processing (**Chapter IV**). The selected samples consisted of fractions collected after the AEX chromatography step, specifically those containing the eluted MjsHSP nanocages; fractions obtained after concentrating the AEX-derived sample using centrifugal filtration; and fractions collected following the SEC purification step, corresponding to the two primary peaks where MjsHSP nanocages were eluted.

In terms of experimental conditions for particle size measurements, details regarding sample preparation are available in sub-section **V.2.5.2.1. Dynamic light scattering** in **V.2.5.2. Production in *V. natriegens* cells** of **V.2.5. Characterization of MjsHSP nanocages**. It is important to note that results were reported as intensity and number distributions. Large particles (greater than 1000 nm), likely contaminants such as environmental dust, were detected despite anti-contamination measures. These data were excluded from the final analysis to ensure accuracy.

Regarding the sample from the fractions post-AEX chromatography containing MjsHSP nanocages (**Chapter IV**, **Figure IV.9A**, peak *2*), a PDI of 0.48 was obtained, indicating a population closer to homogeneity (242). The intensity distribution revealed two main peaks with particle dimensions of 138.5 ± 38.3 nm and 16.6 ± 3.8 nm, while the number distribution analysis identified a single peak with particle dimensions of 19.2 ± 4.6nm.

For the sample from the fractions post-centrifugal filtration containing concentrated MjsHSP nanocages (**Chapter IV**) a PDI of 0.44 was obtained, indicating a population that was also partially homogeneous (242). The intensity distribution revealed two main peaks with particle dimensions of 77.3 ± 42.6 nm and 17.3 ± 4.4 nm. The analysis of the number distribution allowed the identification of a single peak with particle sizes of 11.5 ± 0.9 nm.

Regarding the sample from the first fractions post-SEC purification containing theses protein nanocages (**Chapter IV**, **Figure IV.10A**, peak *1*), a PDI of 0.30 was obtained, indicating a highly homogeneous population (242). The intensity distribution showed a single principal peak with particle dimensions of 71.17 ± 18.4 nm, and the corresponding number distribution analysis identified a single peak with particle dimensions of 53.8 ± 1.4 nm.

For the sample from the second fractions post-SEC purification containing purified MjsHSP nanocages (**Chapter IV**, **Figure IV.10A**, peak *2*), a PDI of 0.66 was obtained, indicating a heterogeneous population (242). This theoretical heterogeneity was not prominently reflected in the intensity distribution, which showed a single main peak corresponding to particle dimensions of 274.0 ± 52.4 nm. The analysis of the number distribution allowed the identification of a single peak with particle sizes of 282.9 ± 24.2 nm.

In the samples obtained from the AEX chromatography step and subsequently concentrated by centrifugal filtration, a similar behavior was observed among themselves and compared to the equivalent *E. coli* samples. Two populations were detected in the intensity distribution, but only one population was observed in the number distribution. For the intensity distribution, there was a noticeable trend of shifting the major contribution toward the peak corresponding to smaller particle dimensions, increasingly closer to the expected size for MjsHSP nanocages. In contrast, the peak related to larger particles in the intensity distribution exhibited a significantly higher standard deviation, indicating less consistency and suggesting that these results were possible associated with contaminant particles or aggregates. In

terms of the number distribution, the results also suggested the presence of a significant population of structures with a measured hydrodynamic diameter of 15.3 ± 3.8 nm. This value was consistent with the expected dimensions of MjsHSP nanocages, described as having an external diameter of approximately 12 nm.

However, in the post-SEC samples, a clear difference was observed between peaks *1* and *2*, as well as compared to the post-AEX chromatography and post-centrifugal filtration samples. The post-SEC samples showed a single peak in both intensity and number distributions, but with values significantly higher than expected for MjsHSP nanocages. This discrepancy may be related to the presence of contaminants with dimensions exceeding 70 nm. The fact that there was agreement between the number and intensity distributions did not strongly support the hypothesis that protein nanocages aggregates were contributing to these results. Additionally, the variation in PDI between the two post-SEC samples (peaks *1* and *2*) was an unexpected result, further reinforcing the possibility of multiple populations within these samples that are contributing differently to the light scattering signal.

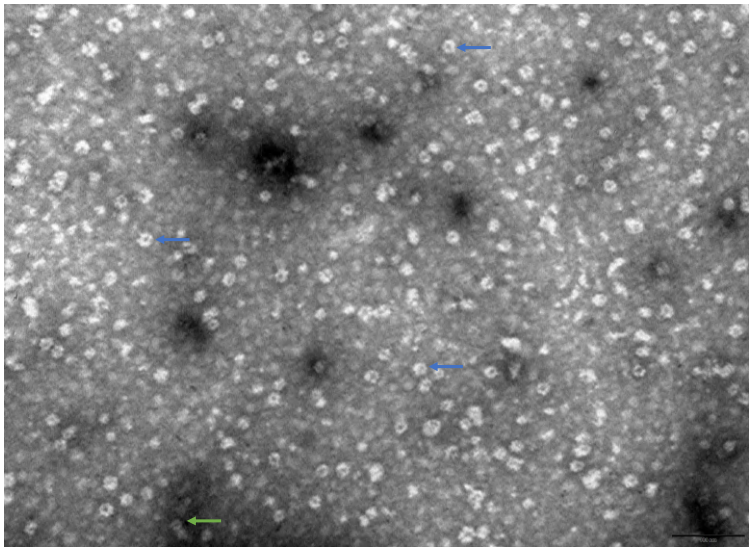
Overall, the results obtained for the two primary samples of MjsHSP nanocages from *V. natriegens* Vmax™ X2 (post-AEX chromatography and post-centrifugal filtration) allowed for estimates of hydrodynamic diameters that can be considered equivalent to those obtained for the corresponding samples from *E. coli*, as expected. Slight variations may even be influenced by variability in the experimental technique itself. The presence of some protein nanocages aggregates was also supported by these results, although they represented a minority population within the two samples. The post-SEC samples appeared to consist predominantly of other contaminant particles with larger dimensions. This part of the study suggested that purifying MjsHSP nanocages derived from a standard protein expression host and an alternative host did not interfere with the structural characteristics of the NVPNs, which is a very interesting and promising result.

Similar to what was done for the MjsHSP nanocages derived from *E. coli* BL21(DE3), it would be interesting to perform zeta potential measurements on these samples to better understand variations in their charge under different experimental conditions. Additionally, experimental verification of their pI and stability capacity would provide valuable insights

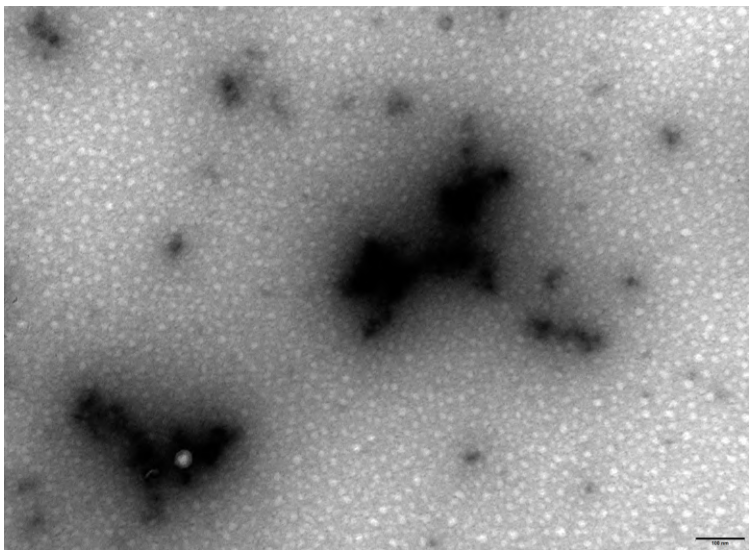
V.3.2.2. Transmission electron microscopy

The MjsHSP nanocages produced in *V. natriegens* Vmax™ X2 were also visualized and analyzed using negative staining TEM. For this purpose, three distinct samples were selected at various stages of downstream processing (**Chapter IV**), following a strategy similar to that applied in the DLS experiments. The selected samples included fractions obtained after concentrating the AEX chromatography-derived sample using centrifugal filtration, as well as fractions collected following the SEC purification step. Specifically, the SEC fractions corresponded to the two main peaks where MjsHSP nanocages were eluted.

(A)



(B)



(C)

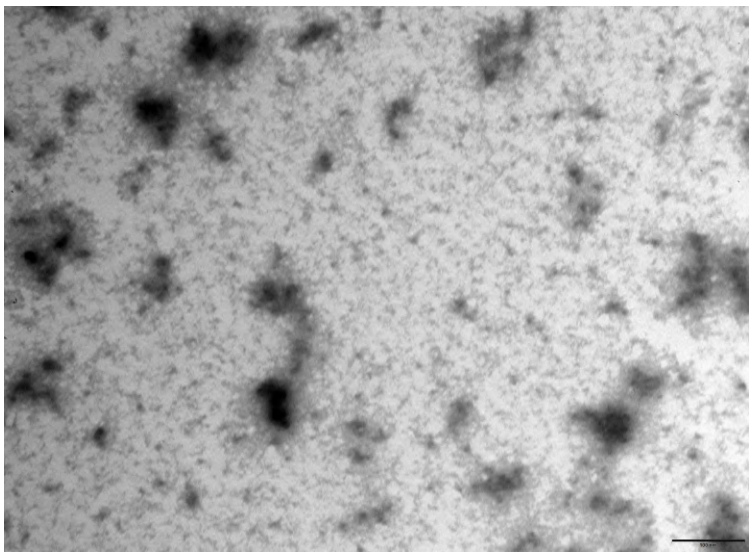


Figure V.8 – Negative staining TEM images of the MjsHSP nanocages produced in *V. natriegens* Vmax™ X2 at distinct stages of the downstream processing. The three samples were collected from the fractions post-centrifugal filtration containing concentrated protein nanocages **(A)**, from the peak *1* fractions post-SEC purification **(B)**, and from the peak *2* fractions post-SEC purification **(C)**. Blue arrows identify individual MjsHSP nanocages, where structural details can even be observed. Green arrow indicates potential aggregate of MjsHSP nanocages. Scale bars: 100 nm.

In terms of experimental conditions for particle size measurements, details regarding sample preparation and subsequent procedure are available in sub-section **V.2.5.2.2. Transmission electron microscopy** in **V.2.5.2. Production in *V. natriegens* cells** of **V.2.5. Characterization of MjsHSP nanocages**.

The negative staining TEM images obtained are presented in **Figure V.8**, with each image clearly labeled to indicate the corresponding stage of downstream processing at which the samples were collected.

The analysis of the results identified structures with morphology and dimensions consistent with MjsHSP nanocages, but only in the sample obtained after centrifugal filtration. In this sample, aggregates of protein nanocages were clearly visualized, which aligned with previous DLS results. Additionally, structures resembling the morphology and dimensions of MjsHSP nanocages were observed, including some displaying darker spots that appeared to correspond to the external pores theoretically described for MjsHSP nanocages.

In contrast, in the post-SEC samples from both peaks *1* and *2*, no nanostructures with dimensions or morphology similar to MjsHSP nanocages were identified, suggesting their presence was minimal. Although this result was unexpected, it aligned with DLS observations. It was likely that their concentration in these samples was so low that they could not be detected in the areas analyzed by TEM, and only other contaminants (larger than 20 nm), either residual from the sample or introduced during the negative staining procedure, were observed.

Similar to the images of MjsHSP nanocages derived from *E. coli* BL21(DE3), some artifacts were present, which may primarily result from the original sample or its preparation process.

Regarding particle dimensions, in the sample obtained after centrifugal filtration, particles with a hydrodynamic diameter of 18.4 ± 1.9 nm were detected. This value was consistent with the theoretical dimensions predicted by the model for MjsHSP nanocages, which estimated an external diameter of approximately 12 nm.

V.3.3. Characterization of TRAP O-rings and nanocages

V.3.3.1. Dynamic light scattering

To analyze the structural characteristics of TRAP O-rings, as well as their assembly state and the potential TRAP nanocages formed after *in vitro* assembly in the presence of gold ions, DLS was utilized. Two distinct samples were selected: purified TRAP O-rings obtained after the CEX chromatography

step (**Chapter IV**) and a reaction mixture for *in vitro* assembly aimed at forming TRAP nanocages. The latter sample was subjected to DLS measurements at the beginning of the reaction, after 5 days, and after 10 days of incubation.

In terms of experimental conditions for particle size measurements, details regarding sample preparation are available in sub-section **V.3.3.1. Dynamic light scattering** of **V.3.3. Characterization of TRAP O-rings and nanocages**. The results were reported as intensity and number distributions. Peaks corresponding to particles larger than 1000 nm were attributed to contaminants like environmental dust, despite anti-contamination efforts, and were excluded from analysis.

Regarding the sample obtained from the fractions post-CEX chromatography containing TRAP O-rings (**Chapter IV**, **Figure IV.15A**, peak *2*), a PDI of 1.00 was recorded, indicating a highly heterogeneous population (242). This heterogeneity was reflected in the intensity distribution, which revealed two main peaks corresponding to particle dimensions of 593.3 ± 3.0 nm and 10.4 ± 3.0 nm. In contrast, the analysis of the number distribution identified a single peak with particle sizes of 7.7 ± 1.6 nm.

For the sample from the *in vitro* assembly reaction aimed at forming TRAP nanocages, a PDI of 0.69 was observed at the beginning of incubation, which increased to 1.00 after 5 days and remained at this value until 10 days post-incubation, indicating an increase in population heterogeneity over time (242). At the beginning of incubation, the intensity distribution revealed two main peaks with particle dimensions of 196.1 ± 62.4 nm and 13.4 ± 4.1 nm, while the number distribution analysis identified a single peak with particle dimensions of 7.9 ± 2.6 nm. After 5 days of incubation, the intensity distribution showed peaks corresponding to hydrodynamic diameters of 141.0 ± 57.1 nm and 11.3 ± 2.1 nm, whereas the number distribution revealed a single peak with particle dimensions of 8.6 ± 3.5 nm. Finally, after 10 days of incubation, the intensity distribution again recorded peaks at hydrodynamic diameters of 160.5 ± 74.5 nm and 12.5 ± 4.5 nm, while the number distribution estimated particle sizes at 8.5 ± 2.3 nm.

In all samples, a similar behavior was observed, with two populations detected in the intensity distribution and only one in the number distribution. No significant differences were found between the samples or during the 10-day monitoring of the *in vitro* assembly reaction of TRAP O-rings into nanocages. In terms of intensity distribution, one of the peaks displayed average dimensions (12.4 ± 1.0 nm) that could be consistent with the hydrodynamic diameters of TRAP O-rings, estimated to be around 9 to 10 nm. On the other hand, the peak corresponding to larger structures (exceeding 140 nm) surpassed even the estimated dimensions of TRAP nanocages (approximately 22 nm of external diameter, as reported in the literature). This suggested that these larger structures could represent aggregates of TRAP nanocages, clusters of TRAP O-rings, or another type of contaminant. The number distribution remained consistent across all samples, and the average value (8.3 ± 0.5 nm) indicated that a majority population in the samples consisted solely of TRAP O-rings. However, even in the initial sample containing protein O-rings, a homogeneous population of particles based on size was not observed, as also reflected by the PDI value.

As future work, it could also be interesting to perform zeta potential measurements for both the TRAP O-rings and the TRAP nanocages. This would help to understand the variation in their surface

charge under different experimental conditions and to identify the differences between the monomeric form and the assembled form.

V.3.3.2. Transmission electron microscopy

Negative staining TEM was used to visualize and analyze the structural morphology of TRAP O-rings, their assembly state, and the potential TRAP nanocages formed through *in vitro* assembly in the presence of gold ions. Three distinct samples were examined: purified TRAP O-rings obtained after the CEX chromatography step (**Chapter IV**), and reaction mixtures for *in vitro* assembly of TRAP O-rings into nanocages, analyzed after 5 days and 1 month of incubation, respectively.

In terms of experimental conditions for particle size measurements, details regarding sample preparation and subsequent procedure are available in sub-section **V.3.3.2. Transmission electron microscopy** of **V.3.3. Characterization of TRAP O-rings and nanocages**.

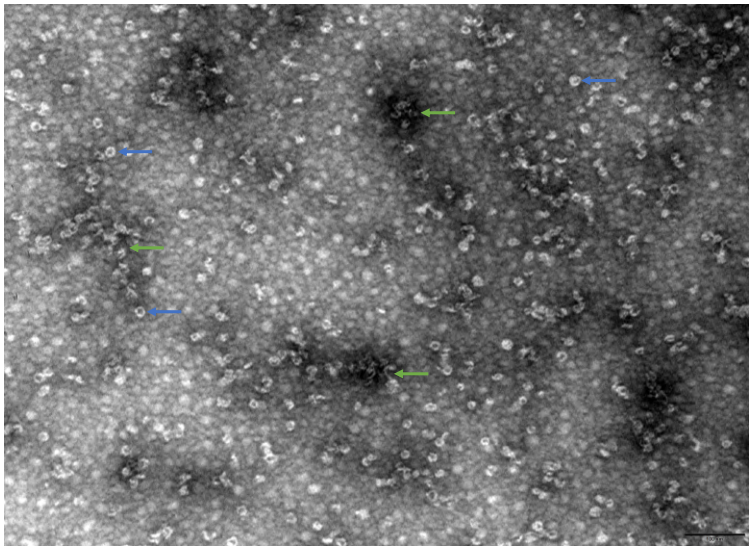
The negative staining TEM images obtained are presented in **Figure V.9**, with each image clearly labeled to indicate the corresponding conditions of the analyzed samples.

The analysis of the results identified structures with morphology and dimensions consistent with TRAP O-rings in all samples, which was expected since, even in the samples from the assembly reactions, it was likely that the assembly process was incomplete, leaving free protein O-rings in solution. However, the number of TRAP O-rings observed in the *in vitro* reaction samples after 5 days and 1 month of incubation was quite high. Combined with the absence of structures with morphological characteristics matching TRAP nanocages, this suggested that the assembly process was unsuccessful. This observation aligned with previous DLS results. Regarding the TRAP O-rings, several images clearly showed their characteristic ring shape with a central opening that appeared darker. Additionally, incomplete structures were observed that might represent TRAP O-rings containing fewer than the standard 11 monomers. This observation highlighted the challenge of distinguishing between different forms of these nanostructures, ranging from the monomeric form (8.3 kDa) to fully assembled nanocages (2.2 MDa). In various samples, aggregates of TRAP O-rings were also visible, which may contribute, at least partially, to the light scattering signal detected in DLS measurements. It is possible that TRAP nanocages were formed but existed in very low number, making them undetectable in the areas analyzed by TEM.

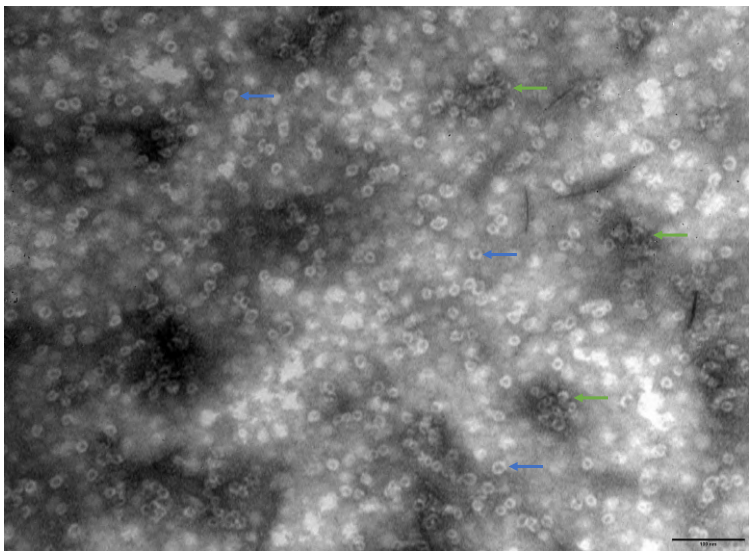
Similar to the images of MjsHSP nanocages derived from both *E. coli* BL21(DE3) and *V. natriegens* Vmax™ X2, some artifacts were observed, probable resulting from the original sample or the preparation process.

Regarding particle dimensions, for the sample of purified TRAP O-rings obtained after CEX chromatography, particles with a hydrodynamic diameter of 17.6 ± 2.8 nm were detected. In the sample from the *in vitro* assembly reaction of TRAP O-rings into nanocages after 5 days of incubation, particles exhibited a hydrodynamic diameter of 18.9 ± 2.4 nm. Meanwhile, for the equivalent reaction after 1 month of incubation, the average hydrodynamic diameter was 18.7 ± 2.0 nm. All these values were consistent with the theoretical dimensions predicted by the model for TRAP O-rings, which estimate an external diameter of about 9 to 10 nm.

(A)



(B)



(C)

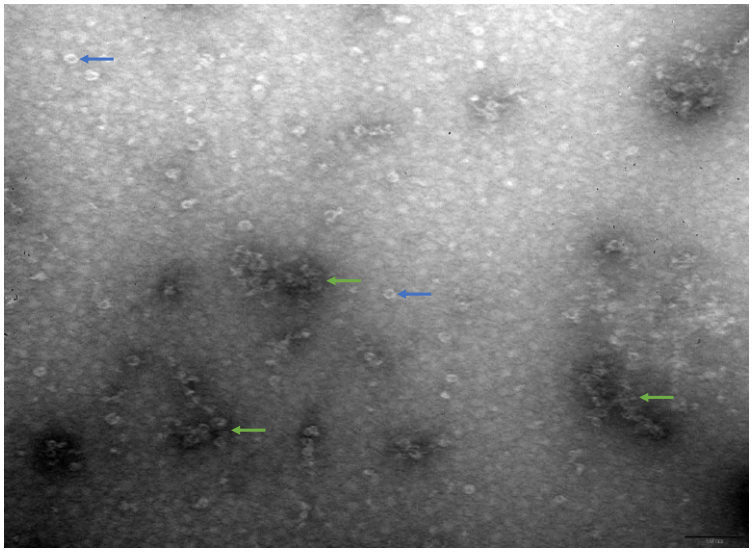


Figure V.9 – Negative staining TEM images of the TRAP O-rings produced in *E. coli* BL21(DE3) and subsequent TRAP nanocages formed through an *in vitro* assembly reaction. Three samples were analyzed: purified TRAP O-rings obtained after the CEX chromatography step **(A)**, reaction mixture for *in vitro* assembly of TRAP O-rings into nanocages after 5 days of incubation **(B)**, and reaction mixture for in vitro assembly of TRAP O-rings into nanocages after 1 month of incubation **(C)**. Blue arrows identify individual MjsHSP nanocages, where structural details can even be observed. Green arrows indicate potential aggregates of MjsHSP nanocages. Scale bars: 100 nm.

Clearly, it would be essential in future research works to improve the *in vitro* assembly of TRAP O-rings into nanocages, so that it becomes possible to obtain an adequate quantity of TRAP nanocages for further analysis and application in functional assays and characterization studies.

V.3.3.3. Fluorescence correlation spectroscopy

The determination of the biophysical properties of TRAP O-rings and nanocages included evaluating potential aggregation or the formation of larger structures, as well as analyzing the processes of assembly and disassembly of these protein-based nanostructures to investigate the dynamics involved. These were the main objectives for utilizing FCS in these NVPNs.

Two samples were initially analyzed, one containing TRAP O-rings purified after the CEX chromatography purification step (**Chapter IV**), and another potentially containing TRAP nanocages formed as a result of the *in vitro* assembly reaction with gold ions (**Chapter V**). Following this initial approach, the aim was to investigate the mechanisms of assembly and disassembly. For the first case, involving the purified TRAP O-rings mentioned above, the focus was on inducing the formation of fully assembled structures. For the second case, based on the sample with TRAP nanocages as described earlier, a chemical compound was added to induce disassembly.

The fluorescent molecule used for labelling the TRAP O-rings and the TRAP nanocages was Alexa Fluor 647, conjugated via a maleimide group, which covalently bond to thiol groups on the protein structure (cysteine residues). Details regarding the experimental procedure used are available in sub-section **V.3.3.3. Fluorescence correlation spectroscopy** of **V.3.3. Characterization of TRAP O-rings and nanocages**.

Through the experiments conducted, it was possible to explore the dynamics that characterize TRAP O-rings and TRAP nanocages in their various assembly and disassembly reactions. These processes involve four distinct populations: the monomers constituting the TRAP O-rings, the O-rings themselves, the formation of TRAP nanocages from these O-rings, and finally, the dissociation of these nanocages back into partially or completely incomplete structures. **Figure V.10** illustrates this dynamic equilibrium of association and dissociation involving TRAP O-rings and nanocages. The first step in the formation of TRAP O-rings from their respective monomers is a reversible chemical equilibrium, as the interaction between these monomers occurs through non-covalent protein interactions (**Figure V.10A**). The second step involves the formation of TRAP nanocages in an *in vitro* reaction induced by a chemical stimulus, specifically the presence of gold ions (Au^{3+}), which is practically irreversible (**Figure V.10B**). The third step consists of the dissociation of TRAP nanocages triggered by reducing conditions, such

as the use of the compound TCEP. This reaction is also irreversible and results in partially or completely dissociated structures (**Figure V.10C**).

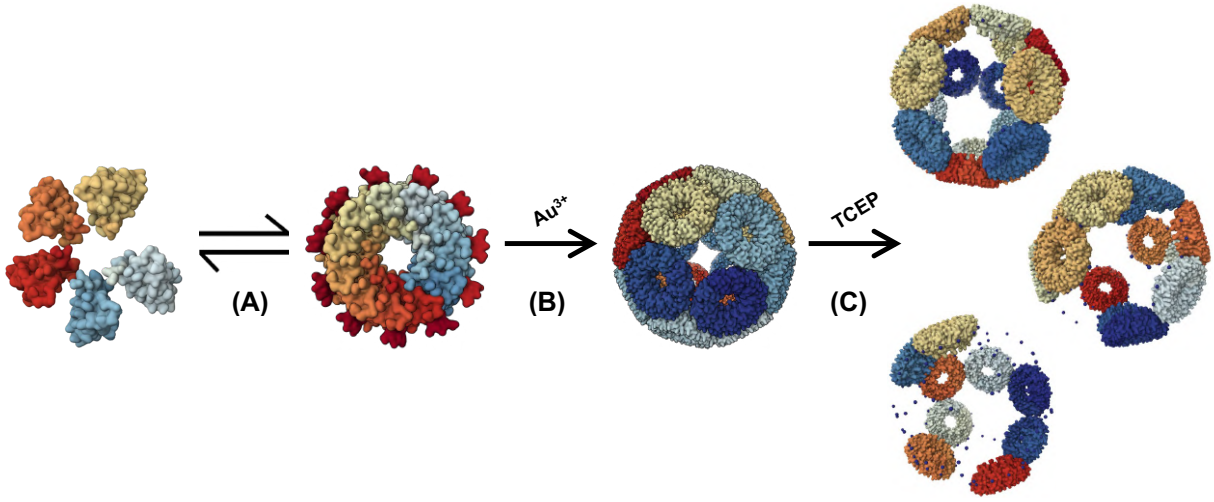


Figure V.10 – Schematic representation of the dynamic equilibrium involving the association and dissociation processes that characterize TRAP O-rings and TRAP nanocages. **(A)** Irreversible reaction between TRAP O-rings and their constituent monomers. **(B)** Irreversible reaction induced by a chemical stimulus (gold in the form of Au^{3+} ions) leading to the assembly of TRAP nanocages from TRAP O-rings. **(C)** Irreversible reaction induced by a chemical stimulus (addition of a reducing agent such as TCEP) causing the disassembly of TRAP nanocages into partially or completely incomplete structures. 3D representations were created using the Mol* Viewer tool (107,108).

The initial experiments aimed to assess the equilibrium between TRAP O-rings and their respective monomers in a sample containing TRAP O-rings labeled with the Alexa 647 fluorophore. For this purpose, the original labeled sample was concentrated and diluted to obtain three different concentrations: high concentration (concentrated 4 times), original concentration, and low concentration (diluted 10 times). FCS measurements were performed on these three conjugate samples, and fluorescence intensity time traces were recorded. From these MCS-trace data, autocorrelation curves were determined for each sample, assuming two components, as this process of association and dissociation represents a dynamic equilibrium between the associated form (TRAP O-rings) and the free form (monomers of TRAP O-rings).

The normalized results (experimental data and fitted model) for the sample with high concentration are shown in **Figure V.11A**. This sample was selected as representative of the three samples with different concentrations, as all exhibited similar profiles. Additionally, the autocorrelation curve for unbound Alexa Fluor 647 was included for comparative purposes.

The FCS measurements enabled the determination of the translational diffusion coefficients of TRAP O-rings and/or their monomers labeled with Alexa Fluor 647. The preparation of these three samples, resulting from the dilution of the TRAP O-rings-Alexa 647 conjugate, allowed for establishing a ratio between the parameter ρ of one population (with a higher diffusion coefficient) and that of another population (with a lower diffusion coefficient) present in each sample. Assuming that the fluorophore brightness did not vary, this parameter ρ was proportional to each population's relative abundance.

Thus, **Figure V.11B** presents the obtained values for the diffusion coefficients of the labeled components in each of the three samples, as well as the corresponding ρI/ρII ratios.

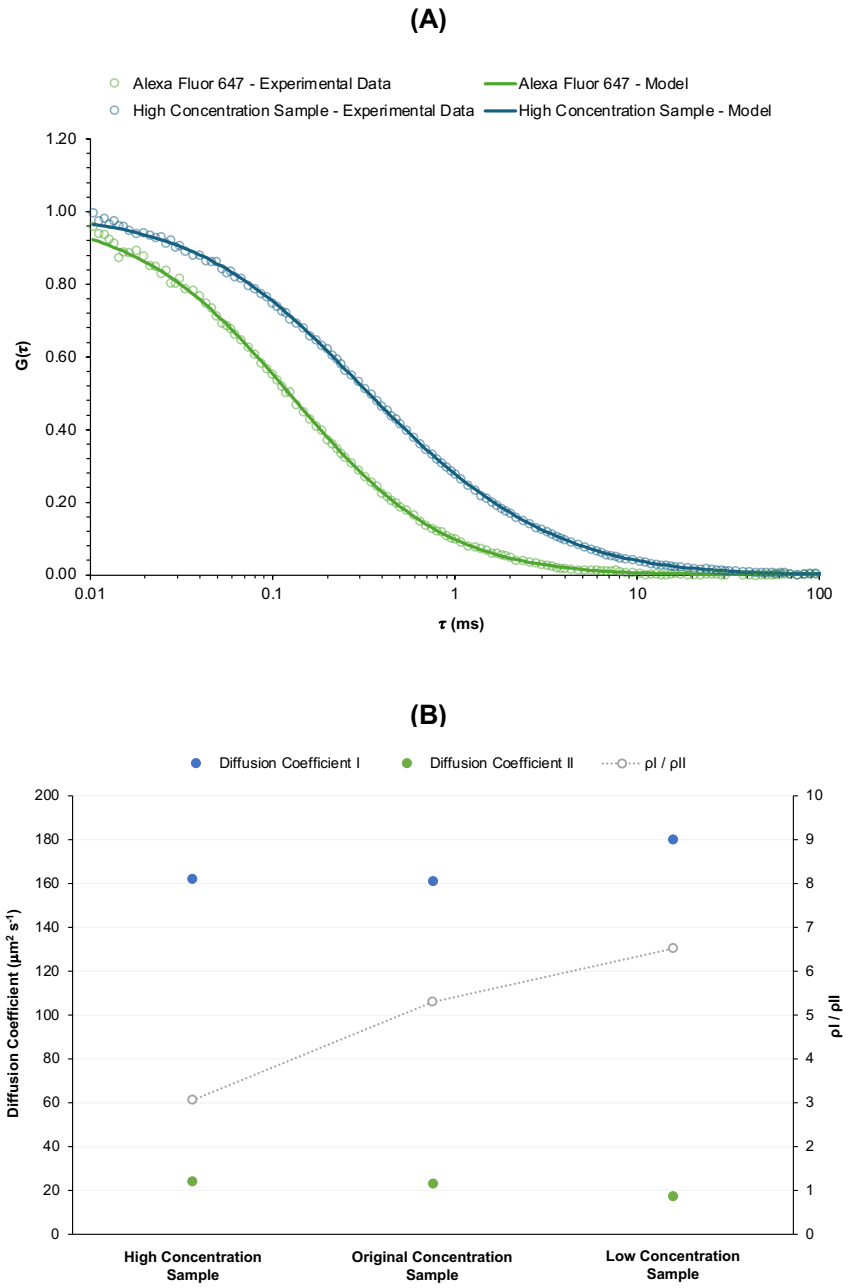


Figure V.11 – FCS measurements to evaluate the dynamic equilibrium between TRAP O-rings and their constituent monomers. **(A)** Autocorrelation curves obtained for the unbound Alexa 647 fluorophore and for a concentrated sample containing TRAP O-rings labeled with this fluorophore. Light colors correspond to the experimental data and dark colors correspond to the fitted model. **(B)** Graphical representation of the diffusion coefficients for the two populations (I and II) present in three samples of TRAP O-rings labeled with different concentrations (high, original, and low). Additionally, the respective ratios (ρI/ρII) are shown to assess the dynamics of association and dissociation.

Analyzing **Figure V.11A**, it was evident that the autocorrelation curve profiles differ between the unbound Alexa 647 fluorophore and the high concentration sample containing TRAP O-rings. These results aligned with expectations, as the unbound fluorophore had smaller dimensions and, consequently, a higher diffusion coefficient (290 ± 33 $\mu m^2 s^{-1}$). In contrast, TRAP O-rings labeled with the same fluorophore had a larger hydrodynamic radius, which reduced their diffusion coefficient. This was observed as a shift of the autocorrelation curve to the right in the graphical representation. For the high concentration sample, which exhibited a behavior consistent with two components, two diffusion coefficients were identified: 162 ± 7 $\mu m^2 s^{-1}$ for one population and 24 ± 17 $\mu m^2 s^{-1}$ for the other.

Regarding the unbound fluorophore, the measured diffusion coefficient aligned with values reported in the literature, such as 330 ± 10 $\mu m^2 s^{-1}$in water at 25°C (246), confirming the presence of a single population of unbound fluorescent molecules.

For the TRAP O-rings sample, the higher diffusion coefficient likely corresponded to individual monomers of TRAP O-rings. Using **Equation V.3**, a hydrodynamic diameter of approximately 3.0 nm was estimated, which was consistent with the dimensions of TRAP monomers. For the population with a lower diffusion coefficient, a hydrodynamic diameter of approximately 20.5 nm was estimated using the same methodology. This value was slightly higher than expected for TRAP O-rings (theoretically about 9 to 10 nm). However, it must be considered that these calculations assume spherical objects, whereas TRAP O-rings have a ring-like structure with an open center. In this context, the observed diffusion coefficient possible resulted from two modes of diffusion since TRAP O-rings can adopt two distinct orientations during movement, leading to different diffusion coefficients. Consequently, the global diffusion coefficient was influenced by this phenomenon. It is important to note that this type of FCS measurements cannot distinguish between these two modes of diffusion.

Regarding **Figure V.11B**, the objective was to evaluate how the dynamic association reaction between TRAP O-ring monomers and O-rings, and vice versa, was progressing. Since the interactions between proteins are non-covalent, there is a possibility of dissociation of O-rings into their monomers. This is particularly relevant as the experiments were conducted at concentration ranges close to the affinity constant, meaning the reaction does not exclusively shift toward association. In future studies, it may be possible to add unlabeled TRAP O-rings to this type of sample to force the equilibrium to shift toward O-ring formation, enabling a more homogeneous population containing only protein O-rings. Currently, two populations were observed: one consisting of TRAP O-rings and another of free monomers resulting from dissociation. The diffusion coefficient values for both populations remained relatively constant across the three distinct concentrations tested. Population I, consisting of free monomers, exhibits diffusion coefficients ranging from 161 ± 10 $\mu m^2 s^{-1}$ to 180 ± 11 $\mu m^2 s^{-1}$. Population II, consisting of O-rings, shows diffusion coefficients between 17 ± 3 $\mu m^2 s^{-1}$ and 24 ± 2 $\mu m^2 s^{-1}$. When considering the ratio ρI/ρII, it was evident that as the sample concentration increased, the contribution of free monomers (with higher diffusion coefficients) decreased, leading to a reduction in this ratio, as observed in the graph. Conversely, as the sample was diluted, the contribution of free species increased, causing the ratio to rise. This behavior was clearly demonstrated in **Figure V.11B**, with the ratio ρI/ρII serving as a reliable indicator of this phenomenon.

Afterwards, using samples of TRAP O-rings and TRAP nanocages, both labeled, the aim was to compare their diffusion behavior by analyzing values for this characteristic parameter and, subsequently, their hydrodynamic diameters. In **Figure V.12A**, the normalized results (experimental data and fitted model) for the autocorrelation curves of both samples are presented, along with the curve for Alexa Fluor 647 for comparative purposes.

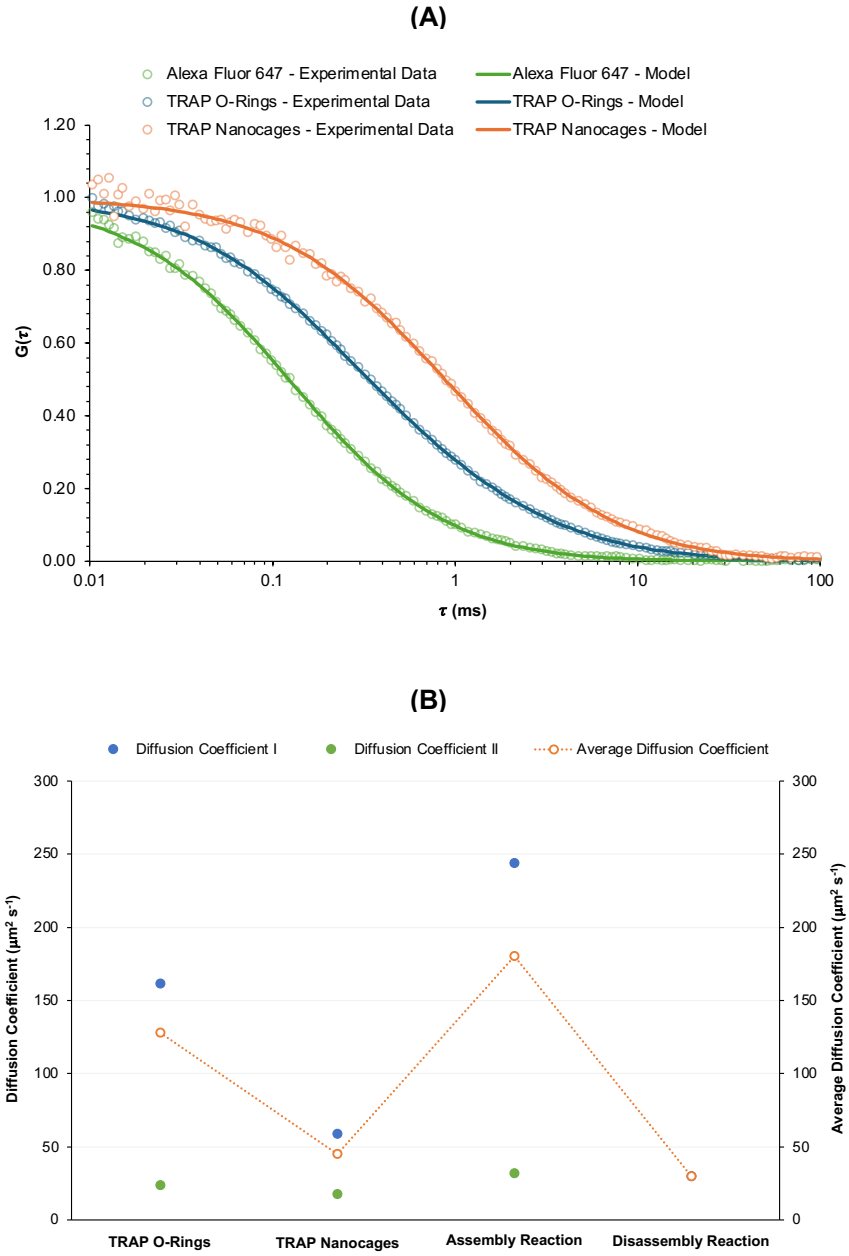


Figure V.12 – FCS measurements to evaluate the dynamics of association and dissociation between TRAP O-rings and TRAP nanocages. **(A)** Autocorrelation curves obtained for the unbound Alexa 647 fluorophore, for a concentrated sample containing TRAP O-rings labeled with this fluorophore, and for a sample containing TRAP nanocages also labeled with the same fluorophore. Light colors correspond to the experimental data and dark colors correspond to the fitted model. **(B)** Graphical representation of the diffusion coefficients of one or two populations (I or I and II) present in these two samples (TRAP O-rings and TRAP nanocages). The data reflect the assembly reaction of TRAP nanocages from pre-labeled TRAP O-rings (triggered by a chemical stimulus via the addition of

gold ions in the form of Au^{3+}) and the disassembly reaction of pre-labeled TRAP nanocages (induced by a chemical stimulus via the addition of the reducing agent TCEP). Calculated average diffusion coefficients for each of the four analyzed samples are also indicated.

A clear shift in the autocorrelation curve was observed from the unbound Alexa 647 fluorophore with smaller dimensions, to TRAP O-rings with intermediate dimensions, and finally to the TRAP nanocages sample with larger dimensions. This trend reflected the relationship between size and diffusion coefficient, where larger structures exhibited lower diffusion coefficients. For the TRAP O-rings sample, two populations were identified with previously mentioned diffusion coefficients. For the TRAP nanocages sample, the diffusion coefficients for its two populations were 59 ± 11 $\mu m^2\ s^{-1}$ and 18 ± 4.7 $\mu m^2\ s^{-1}$, corresponding to hydrodynamic diameters of approximately 8.3 nm and 27.3 nm.

Subsequently, two experiments were conducted to understand the effects of assembly and disassembly mechanisms on diffusion coefficients. For the TRAP O-rings sample, the experimental procedure was followed to induce the assembly reaction, while for the TRAP nanocages sample, a reducing agent was added to induce the disassembly reaction. In the assembly reaction, the data revealed two components in the autocorrelation curve, whereas in the disassembly reaction, only one component was observed. Using these data, it was possible to estimate an average diffusion coefficient for each of the four samples subjected to FCS measurements, utilizing the respective parameters ρI and ρII as indicators of each component's contribution. Thus, **Figure V.12B** presents the diffusion coefficients for each population in the four samples, as well as their respective average diffusion coefficients. This analysis provided insights into how assembly and disassembly processes influence molecular diffusion behavior.

Transitioning from the TRAP O-rings sample to the TRAP nanocages sample revealed a decrease in the average diffusion coefficient, which made sense, as some degree of assembly naturally led to a reduction in diffusion. However, the sample from the assembly reaction exhibited a higher average diffusion coefficient. If partial or complete assembly had occurred, a shift in the population toward structures with lower diffusion coefficients would be expected, resulting in a decrease in the measured average value. Since this was not observed, it can be explained by the fact that cysteine residues on the TRAP O-rings were used as binding sites for the Alexa 647 fluorophore. This interfered with the assembly process, as it relied precisely on thiol groups from these amino acids. Thus, incomplete objects may have formed due to covalent bonding at positions not blocked by Alexa Fluor 647 upon gold ion addition. In this sample, diffusion coefficients of 244 ± 75 $\mu m^2\ s^{-1}$ for one population and 32 ± 7.1 $\mu m^2\ s^{-1}$ for another population were recorded. The presence of partially assembled TRAP nanocages was plausible, as fully assembled nanocages would exhibit a lower diffusion coefficient. Regarding disassembly, an increase in diffusion coefficient was expected since dissociation of the TRAP nanocages would produce smaller objects compared to the intact nanocages, leading to higher diffusion coefficients. Nevertheless, this was not observed, with a diffusion coefficient of approximately 30 ± 2.4 $\mu m^2\ s^{-1}$ being recorded. Overall, these results indirectly suggested that some degree of TRAP nanocages assembly was achieved during the experiments.

In terms of future work, it would be essential to ensure and improve the labelling of TRAP O-rings and subsequent TRAP nanocages. For this, it is necessary to have a sufficient quantity of protein O-rings and nanocages to be labeled and fluorophore within the micromolar range, ideally with a molar excess of the fluorophore. These conditions will promote the labelling reaction. Another important aspect to consider is the purification of the protein O-rings/nanocages and fluorophore conjugate from the remaining free fluorophore, which is often present in significant amounts due to the excess added during the labelling reaction. It is crucial to use an efficient methodology for this purification, which is not always straightforward, as observed in this study. Based on the data obtained, the use of centrifugal filters in several sequential washes could be an effective alternative.

Furthermore, it was verified that FCS is a promising technique for evaluating the assembly and disassembly processes of TRAP O-rings into nanocages and vice-versa. However, several challenges were identified that could be explored in future studies, such as labelling the protein O-rings and nanocages with a different fluorophore, for example, using NHS ester chemistry. This approach would allow labelling at other sites besides cysteines, leaving them free to react with gold ions. Additionally, for monitoring the disassembly reaction, it would be useful to add a mixture of labeled and unlabeled TRP O-rings (ideally with a higher proportion of unlabeled ones) to promote the reaction, as the excess of unlabeled protein O-rings would help ensure the formation of fully or more completely assembled TRAP nanocages. This strategy would likely result in a greater quantity of formed and simultaneously labeled nanocages, enabling FCS measurements to yield better results compatible with a more homogeneous population. Regarding disassembly, it would also be interesting to develop further studies using different compounds for this process beyond the reducing agent TCEP and testing other experimental conditions. Moreover, once optimal conditions for assembly and its detection via FCS are established, it would be valuable to monitor this process in real time over an extended period to gain insights into the kinetics of both assembly and disassembly reactions. This approach could provide additional data on the dynamics of this model of NVPNs.

V.4. Conclusions

NVPNs have emerged as a key area of interest across bioengineering, biotechnology, and biomedicine fields, demonstrating significant potential in various applications. These include enhancing targeted drug delivery systems (114,126,127), developing novel vaccine platforms (119,129,132), advancing bioimaging techniques (134,135), controlling biomineralization processes (93), and creating more efficient biocatalytic systems (75,144). Research on NVPNs has advanced significantly in both conceptual studies and practical applications. However, there remains a notable gap in efforts to develop bioprocesses and analytical and characterization methodologies tailored for the production and optimization of natural and engineered protein nanocages (186). Development and validation of robust analytical and characterization methods is crucial for evaluating structural and functional characteristics of protein nanocages, ensuring product quality and monitoring biomanufacturing process performance and consistency (138).

In this **Chapter V**, the main objective was to explore the structural characterization of two model NVPNs, MjsHSP and TRAP nanocages, including the precursors of the latter (TRAP O-rings). Protein nanocages produced by two different bacterial hosts (*E. coli* and *V. natriegens*) were used as well as a combination of well-established techniques with less common ones including a fluorescence-based strategy not yet applied in NVPNs according to the current publications available in the literature.

Considering the results obtained, the structural characterization of both MjsHSP and TRAP nanocages, along with the precursors of the latter (TRAP O-rings), was achieved. For MjsHSP nanocages produced in *E. coli* BL21(DE3), DLS and TEM were employed as primary techniques to estimate hydrodynamic diameters, assess assembly quality, and detect aggregates. TEM provided visual confirmation of nanostructure morphology and assembly states. Complementary techniques included AFM, which offered insights into surface topology despite sample preparation challenges, and FCS, which estimated diffusion coefficients, inferred hydrodynamic diameters, and evaluated aggregation dynamics as well as monomer-nanocage association/dissociation. MjsHSP nanocages produced by *V. natriegens* were characterized using DLS and TEM exclusively. For TRAP O-rings and nanocages, DLS, TEM, and FCS were applied using similar strategies to those used for MjsHSP nanocages. TRAP nanocages were formed via gold-mediated assembly of TRAP O-rings, leveraging thiol groups for precise control over assembly-disassembly processes.

The structural analysis revealed distinct characteristics for each type of NVPNs studied. MjsHSP nanocages produced in *E. coli* displayed consistent assembly, as shown by DLS, which measured an average hydrodynamic diameter of 12.9 ± 2.4 nm. This was corroborated by TEM imaging with negative staining, which revealed uniform octahedral structures measuring 12.8 ± 1.1 nm. AFM provided detailed surface topology, and FCS data indicated a partially monodisperse population of MjsHSP nanocages with a diffusion coefficient of 41 $\mu m^2\ s^{-1}$, suggesting reduced aggregation. In comparison, MjsHSP nanocages produced in *V. natriegens* showed only minimal structural differences, with a slightly larger hydrodynamic diameter of 15.3 ± 3.8 nm according to DLS. TEM analysis confirmed proper assembly and revealed similar morphology to the results observed in *E. coli*, with a diameter of 18.4 ± 1.9 nm. TRAP O-rings and nanocages exhibited more complex assembly and disassembly dynamics. DLS measurements indicated a hydrodynamic diameter of 8.3 ± 0.5 nm for TRAP O-rings, which was

supported by TEM images showing clear protein ring morphology with a diameter of 18.7 ± 2.0 nm. Nevertheless, TRAP nanocages were not detected using these techniques. FCS data suggested the presence of partially assembled TRAP nanocages with a diffusion coefficient of 18 ± 4.7 $\mu m^2 s^{-1}$, alongside TRAP O-rings with a diffusion coefficient of 59 ± 11 $\mu m^2 s^{-1}$, corresponding to an expected hydrodynamic diameter of 8.3 ± 1.6 nm.

The work established in this **Chapter V** allowed the validation of a combined set of conventional and innovative methodologies that can be used for the structural characterization of protein nanocages, even at the end of the manufacturing process. The data generated can inform future research works and contribute to the development of specific protocols tailored to various protein nanocages. From the perspective of future studies, the structural characterization and functional analysis of NVPNs, including MjsHSP and TRAP nanocages, can be significantly advanced by addressing key areas. Zeta potential studies should investigate pH dependence to determine the pI of MjsHSP nanocages and compare it to theoretical values. Stability assessments under varying pH and conductivity conditions could reveal changes in zeta potential over time. Expanding these studies to include MjsHSP nanocages from *V. natriegens* and TRAP O-rings and nanocages would enable comparative analyses of surface properties and assembly mechanisms. AFM can play a key role in future studies by improving sample preparation to enhance image quality and better visualize individual nanocages. It can also be used to investigate mechanical properties such as stiffness, elasticity, and deformation behavior, providing insights into NVPNs structural integrity. Additionally, AFM could study protein-protein interactions within protein nanocages, measuring intermolecular forces and interaction dynamics to deepen understanding of their assembly mechanisms. Future works should focus on improving the labelling of TRAP O-rings and nanocages by ensuring sufficient protein quantities and fluorophore concentrations, ideally with a molar excess of fluorophore. Efficient purification methods, such as sequential washes with centrifugal filters, are essential to remove excess unbound fluorophore. Alternative labelling strategies, like NHS ester chemistry, could target different sites, leaving cysteines available for gold-mediated assembly. For the disassembly studies, testing new reducing agents and experimental conditions could provide further insights. Real-time monitoring of assembly and disassembly using FCS would help elucidate reaction kinetics and dynamics. Additionally, exploring alternative fluorescence-based strategies could enhance the characterization of these protein-based nanostructures and improve understanding of their biophysical properties. Further studies could involve cultured cells to evaluate the biological activity of protein nanocages. Confocal microscopy might be used to investigate the internalization of empty, labeled nanocages in live cells, such as CHO or HeLa, to assess their ability to permeate cell membranes and evaluate their cellular uptake and intracellular fate. Alternatively, labeled protein nanocages could be loaded with specific molecules or drugs, such as doxorubicin, using either a disassembly/reassembly method for MjsHSP nanocages or the gold-mediated assembly system of TRAP nanocages. These loaded nanocages might then be used in assays to examine cell binding, internalization, and intracellular behavior. Such experiments could provide valuable insights into the mechanisms of cellular uptake and the potential applications of NVPNs in drug delivery and therapeutic interventions.

Chapter VI

Conclusions and Future Perspectives

Protein nanocages, among the most relevant protein-based nanoparticles, are highly ordered nano-scale structures self-assembled from multiple monomers, offering advantages such as spatial control of biological processes, compartmentalization capabilities, high surface-to-volume ratio, multi-functionality, ease of modification, stability, monodispersibility, biocompatibility, and biodegradability, making them attractive for various biotechnological and biomedical applications, including biomineralization and nanomaterial synthesis (1,2,14). Protein nanocages are classified into two main categories: VLPs and NVPNs. VLPs, the predominant group, exhibit diverse structures and sizes but lack genetic material, making them non-infectious (15–21). NVPNs, unrelated to viral particles, form through protein monomer self-assembly, with their structure determined by subunit interfaces (2,5). As explored in **Chapter I**, several applications of NVPNs have been implemented in an extensive range of scientific areas from biotechnology to biomedical sciences. Particular emphasis on various biomedical applications (22–31), including drug delivery, vaccine development, and diagnostic bioimaging, with the potential for de novo design of bioinspired NVPNs using artificial monomers (2,3,14,32). Nevertheless, the development of all these possible applications for NVPNs requires large amounts of pure and well-folded assemblies (33). In the literature, several studies have advanced the knowledge about the characteristics of protein nanocages and provided proof-of-concept experiments on different and innovative functionalization approaches but works focused on the challenges of biomanufacturing have been practically non-existent. Therefore, in this thesis, biomanufacturing processes of NVPNs were addressed from the upstream to the downstream steps and including analytical techniques suitable to characterize the structure of protein nanocages. MjsHSP and TRAP nanocages were the two NVPNs selected as models for all the work developed (**Chapter II**).

Regarding the upstream processing (**Chapter III**), the benchmark manufacturing method for MjsHSP nanocages was successfully established in both *E. coli* BL21(DE3) and *V. natriegens* Vmax™ X2, as well as for TRAP O-rings in *E. coli*. The production process was optimized by testing five parameters: growth medium composition, concentration of protein expression inducer, temperature of protein expression, time point of induction and aeration conditions. Among these, IPTG concentration had the greatest impact on bacterial growth and protein nanocages expression. The optimized production process in *E. coli* BL21(DE3) (enhanced 2xYT culture medium, 0.1 mM of IPTG, 37°C during protein expression, time point of induction at $OD_{600\ nm} \approx 0.6$ and shake flask without baffles) yielded 295.7 ± 0.0 mg L^{-1} of MjsHSP nanocages with a volumetric productivity of 53.8 ± 0.0 mg L^{-1} h^{-1} after 4 hours, representing 5.4- and 5.1-fold increases over the standard process. In *V. natriegens*, the same optimization strategy (with a baffled shake flask) produced 60.9 ± 13.7 mg L^{-1} of MjsHSP nanocages and a volumetric productivity of 10.4 ± 2.3 mg L^{-1} h^{-1}, achieving 2.2- and 1.8-fold improvements compared to its standard process, despite lower yields than *E. coli*. For TRAP O-rings in *E. coli*, the optimized conditions resulted in a yield of 249.1 ± 3.7 mg L^{-1} and a volumetric productivity of 46.7 ± 0.7 mg L^{-1} h^{-1} after 4 hours, reflecting modest increases of 1.4- and 1.3-fold over the reference process. These results highlight significant improvements in protein nanocage production efficiency across both strains and processes.

Future work should focus on several key areas to further advance the production of protein nanocages It is evident that while *V. natriegens* shows promise as a complementary host organism to

E. coli, additional research is necessary to fully realize its potential. The following aspects should be prioritized: genetic improvement of *V. natriegens* strains for optimized protein nanocages expression; optimization of plasmid components tailored to physiological and metabolic characteristics of *V. natriegens*; in-depth studies on transformation methods for *V. natriegens* to address current challenges; systematic optimization of TRAP O-rings expression, similar to the approach used for MjsHSP nanocages; and scale-up of the optimized processes to pilot scale bioreactors for more comprehensive evaluation and validation. These future directions aim to enhance the efficiency and reliability of NVPNs production, particularly in *V. natriegens*, and to connect the gap between laboratory scale experiments and larger scale biotechnological applications.

Concerning the downstream processing (**Chapter IV**), a standard purification approach for MjsHSP nanocages from *E. coli* BL21(DE3) was successfully established and optimized, involving enzymatic pre-treatment with TURBO™ DNase followed by single-step AEX chromatography using Capto™ Q ImpRes resin in bind/elute mode with a stepwise NaCl gradient. This process achieved a recovery yield of 151.9 ± 1.8 µg mL^{-1}, a recovery efficiency of 63.3 ± 3.3%, and host protein impurities removal of 88.0 ± 1.0%, representing a 7.2-fold increase in recovery efficiency compared to the benchmark process. For MjsHSP nanocages from *V. natriegens* Vmax™ X2, a similar single-step AEX chromatography approach achieved lower values, with a recovery yield of 3.4 ± 0.4 µg mL^{-1}, recovery efficiency of 9.5 ± 1.1%, and host protein impurities removal of 96.8 ± 1.3%, confirming the feasibility of purifying nanocages from an alternative host strain. For TRAP O-rings, an optimized downstream process involving thermal pre-treatment followed by single-step CEX chromatography with Capto™ SP ImpRes resin achieved a recovery yield of 320.2 ± 5.2 µg mL^{-1}, recovery efficiency of 69.6 ± 0.8%, host protein impurities removal of 82.0 ± 2.3%, and dsDNA impurities removal of 85.0 ± 5.2%. This strategy improved recovery efficiency by 4-fold compared to the reference process while maintaining similar impurity removal levels, though challenges in the *in vitro* assembly prevented further purification of TRAP nanocages.

Further developments should focus on refining the processes established in this **Chapter IV**, which provide a strong starting point for the extraction, isolation, and purification of protein nanocages. Key priorities include repeating critical experiments to ensure reproducibility and optimize conditions for the efficient purification of MjsHSP nanocages and TRAP O-rings, while improving recovery and impurities removal efficiency. Enzymatic treatments with DNase and RNase could be further explored for both *V. natriegens* and *E. coli* systems to enhance impurity removal while preserving protein nanocages integrity. A major challenge remains the quantification of MjsHSP nanocages and TRAP O-rings throughout the process, as current methods like SDS-PAGE densitometry lack precision. Future research should aim to develop a quick, reproducible, and chromatography-based quantification technique to improve process consistency and efficiency. Additionally, challenges in the *in vitro* assembly of TRAP O-rings into their nanocages must be addressed through further optimization before advancing to purification experiments. If successful, three chromatography methods, SEC (Superdex™ 200 resin), multimodal chromatography (Capto™ Core 400 resin), and hydrogen bond chromatography (CIMmultus® H-Bond monolith) will be evaluated to exploit the structural properties of TRAP nanocages

while optimizing chromatographic parameters such as mobile phases, chromatographic run conditions, and oligomeric state separation.

Finally, in relation to the structural characterization of both protein nanocages (**Chapter V**), distinct characteristics for each NVPNs type were revealed. MjsHSP nanocages produced in *E. coli* showed consistent assembly, with DLS measuring a hydrodynamic diameter of 12.9 ± 2.4 nm, confirmed by TEM imaging (12.8 ± 1.1 nm) and AFM, which provided detailed surface topology. FCS indicated a partially monodisperse population with a diffusion coefficient of 41 $\mu m^2\ s^{-1}$, suggesting reduced aggregation. MjsHSP nanocages from *V. natriegens* exhibited minimal structural differences, showing a slightly larger hydrodynamic diameter (15.3 ± 3.8 nm) by DLS and similar morphology to *E. coli* protein nanocages in TEM (18.4 ± 1.9 nm). TRAP O-rings and nanocages displayed more complex assembly and disassembly dynamics. DLS measured a hydrodynamic diameter of 8.3 ± 0.5 nm for TRAP O-rings, supported by TEM imaging (18.7 ± 2.0 nm). TRAP nanocages were not detected by these methods, but FCS suggested partially assembled nanocages with a diffusion coefficient of 18 ± 4.7 $\mu m^2\ s^{-1}$, alongside TRAP O-rings with a coefficient of 59 ± 11 $\mu m^2\ s^{-1}$, corresponding to an expected hydrodynamic diameter of 8.3 ± 1.6 nm.

Future directions could advance the structural and functional analysis of NVPNs, including MjsHSP and TRAP nanocages, by focusing on key areas. Zeta potential studies should investigate pH dependence to determine the pI of MjsHSP nanocages and assess stability under varying pH and conductivity conditions. Comparative analyses of MjsHSP nanocages from *V. natriegens* and TRAP O-rings/nanocages could provide insights into surface properties and assembly mechanisms. AFM could enhance visualization of individual protein nanocages, investigate mechanical properties like stiffness and elasticity, and study protein-protein interactions to better understand assembly mechanisms. For FCS, improving labeling strategies for TRAP O-rings and nanocages, such as using NHS ester chemistry, could leave cysteines available for gold-mediated assembly. Efficient purification methods, like sequential washes, are essential to remove excess fluorophore. Disassembly studies could explore new reducing agents and conditions, while real-time monitoring of assembly and disassembly using FCS would elucidate reaction kinetics. Exploring alternative fluorescence-based strategies could further characterize these protein-based nanostructures. Additionally, cellular studies using confocal microscopy might assess nanocage internalization and intracellular fate, with drug-loaded protein nanocages tested for cell binding and therapeutic applications, for example, in drug delivery systems.

BIOTECnico Doctoral Program

Completed Curricular Units (36 ECTS)

- Basic Doctoral Formation (6 ECTS);
- Advanced Experimental Techniques and Methodologies in Biotechnology (6 ECTS);
- Bioentrepreneurship (6 ECTS);
- Advanced Topics in Bioengineering and Biological Sciences (6 ECTS);
- Nanobiotechnology (6 ECTS);
- Outreach and Teaching Skills – Department of Bioengineering (6 ECTS).

Extracurricular Courses and Activities

- Workshop "Career development during and after the PhD", organized within the PhD Open Days 2023 – Instituto Superior Técnico – Universidade de Lisboa, November 27, 2023, Lisbon, Portugal (2 hours);
- Workshop "LaTeX 101 – A soft introduction to scientific typesetting", organized within the PhD Open Days 2022 – Instituto Superior Técnico – Universidade de Lisboa, November 16, 2022, Lisbon, Portugal (3 hours);
- Workshop "Writing of the PhD Thesis and preparation of the document", organized within the PhD Open Days 2022 – Instituto Superior Técnico – Universidade de Lisboa, November 16, 2022, Lisbon, Portugal (1.5 hours);
- Course "Introduction to Python and Machine Learning for the Biosciences – 4th Edition", organized in virtual mode by the Instituto de Investigação e Inovação em Saúde (i3S) – Universidade do Porto, October 17-21, 2022 (20 hours);
- Event "ACS on Campus – University of Lisbon", organized by the American Chemical Society (ACS) at Instituto Superior Técnico – Universidade de Lisboa, August 26, 2022, Lisbon, Portugal (4 hours);
- Course "Scientific Tools in Biological Assays: Principles of Cell Cultures and Genetics", organized in virtual mode by Faculdade de Farmácia – Universidade de Lisboa, March 17-19, 2022 (16.5 hours);
- TIMB3 Workshop "Imaging in Biosciences", organized at ITQB NOVA, February 7, 2022, Oeiras, Portugal (9 hours);
- SPICA Virtual Session "The Importance of Chromatography in the Development of Vaccines", organized in virtual mode, October 14, 2021 (2 hours);

- Workshop “Escrita Científica – Da observação da realidade à publicação de um artigo: um processo e um resultado”, organized in virtual mode by the Library of Instituto Superior Técnico – Universidade de Lisboa, June 15-24, 2021 (6 hours);
- Técnico Career Workshop “How to build a CV by BCG”, organized in virtual mode by the Technology Transfer Office – Instituto Superior Técnico – Universidade de Lisboa, April 20, 2021 (2 hours);
- Técnico Career Workshop “LinkedIn (Básico) by Deloitte”, organized in virtual mode by the Technology Transfer Office – Instituto Superior Técnico – Universidade de Lisboa, April 5, 2021 (2 hours);
- Workshop “How to Prepare an Excellent Presentation (or at Least an Acceptable One)?”, organized in virtual mode within the PhD Open Days 2020 – Instituto Superior Técnico – Universidade de Lisboa, October 28, 2020 (3 hours);
- Cell Culture-based Viral Vaccines Course – 5th Edition, organized in virtual mode by the European Society for Animal Cell Technology (ESACT), September 28 – October 2, 2020 (25 hours);
- 17th International PhD Student Symposium “Horizons in Molecular Biology”, organized in virtual mode, September 14-17, 2020 (13 hours);
- Champalimaud Research Symposium 2019 – Tissue Environment in Health and Disease, sponsored by the International Union of Biochemistry and Molecular Biology (IUBMB), October 8-10, 2019, Lisbon, Portugal (18 hours);
- Workshop “Teaching STEM Students – Transformation Guide for Teaching Assistants – 3rd Edition”, organized by the Academic Development Office – Instituto Superior Técnico – Universidade de Lisboa, September 18-19, 2019, Lisbon, Portugal (6 hours).

Publications and Scientific Communications

Articles

- Fonseca SP, Mourato G, Silva RM, João J, Monteiro GA, Silva-Santos AR, Prazeres DMF, *Vibrio natriegens* as a new host for plasmid DNA production, **Journal of Biotechnology**, 2025 (Submitted Manuscript);
- João J, Prazeres DMF, Manufacturing of non-viral protein nanocages for biotechnological and biomedical applications, **Frontiers in Bioengineering and Biotechnology**, 11:1200729, 2023, doi: 10.3389/fbioe.2023.1200729;
- João J, Lampreia J, Prazeres DMF, Azevedo AM, Manufacturing of bacteriophages for therapeutic applications, **Biotechnology Advances**, 49:107758, 2021, doi: 10.1016/j.biotechadv.2021.107758.

Oral Presentations

- João J, Protein nanocages: an innovative platform for vaccine development, **PhD Open Days 2023**, November 27-29, 2023, Instituto Superior Técnico – Universidade de Lisboa, Lisbon, Portugal;
- João J, Prazeres DMF, Biomanufacturing of non-viral protein nanocages with biotechnological and biomedical applications, **3rd iBB – Institute for Bioengineering and Biosciences Workshop**, November 13, 2023, Instituto Superior Técnico – Universidade de Lisboa, Lisbon, Portugal;
- João J, Rosa SS, Paulo PMR, Azevedo AM, Prazeres DMF, Downstream processing of non-viral protein nanocages for biotechnological and biomedical applications: development of chromatography-based purification strategies, **42nd ISPPP – International Symposium on the Separation of Proteins, Peptides and Nucleotides**, November 5-8, 2023, Vienna, Austria;
- João J, Rosa SS, Azevedo AM, Prazeres DMF, Purification of small heat shock protein nanocages: exploring a chromatographic approach, **NInTec – Núcleo de Investigadores Science Days 2022**, October 6-7, 2022, Instituto Superior Técnico – Universidade de Lisboa, Lisbon, Portugal;
- João J, Rosa SS, Azevedo AM, Prazeres DMF, Exploring a chromatographic strategy for the purification of small heat shock protein nanocages, **BPP – Biopartitioning & Purification Conference 2022**, September 25-28, 2022, Aveiro, Portugal;
- João J, Rosa SS, Azevedo AM, Prazeres DMF, Purification of small heat shock protein nanocages: development of a chromatographic strategy, **2nd Chem & Biochem Students Meeting 2022**, July 15, 2022, Faculdade de Ciências – Universidade de Lisboa, Lisbon, Portugal;

- João J, Rosa SS, Azevedo AM, Prazeres DMF, Chromatographic purification of small heat shock protein nanocages, **40th ISPPP – International Symposium on the Separation of Proteins, Peptides and Nucleotides**, November 7-10, 2021, Porto, Portugal.

Poster Presentations

- João J, Prazeres DMF, Biomanufacturing of non-viral protein nanocages for biotechnological and biomedical applications, **PhD Open Days 2023**, November 27-29, 2023, Instituto Superior Técnico – Universidade de Lisboa, Lisbon, Portugal;
- João J, Rosa SS, Paulo PMR, Azevedo AM, Prazeres DMF, Biomanufacturing of non-viral protein nanocages with biotechnological and biomedical applications, **3rd iBB – Institute for Bioengineering and Biosciences Workshop**, November 13, 2023, Instituto Superior Técnico – Universidade de Lisboa, Lisbon, Portugal;
- João J, Rosa SS, Paulo PMR, Azevedo AM, Prazeres DMF, Downstream processing of non-viral protein nanocages for biotechnological and biomedical applications: development of chromatography-based purification strategies, **42nd ISPPP – International Symposium on the Separation of Proteins, Peptides and Nucleotides**, November 5-8, 2023, Vienna, Austria;
- João J, Rosa SS, Paulo PMR, Azevedo AM, Prazeres DMF, Development of biomanufacturing processes for non-viral protein nanocages with biotechnological and biomedical applications, **3rd Chem & Biochem Students Meeting 2023**, July 13, 2023, Faculdade de Ciências – Universidade de Lisboa, Lisbon, Portugal;
- João J, Prazeres DMF, Biomanufacturing of protein nanocages for biomedical applications, **PhD Open Days 2022**, November 14-16, 2022, Instituto Superior Técnico – Universidade de Lisboa, Lisbon, Portugal;
- João J, Rosa SS, Azevedo AM, Prazeres DMF, Purification of small heat shock protein nanocages for biomedical applications, **2nd iBB – Institute for Bioengineering and Biosciences Workshop**, September 23, 2022, Instituto Superior Técnico – Universidade de Lisboa, Lisbon, Portugal;
- João J, Rosa SS, Azevedo AM, Prazeres DMF, Development of a chromatographic approach for the purification of small heat shock protein nanocages, **3rd BiolberoAmerica – Ibero-American Congress on Biotechnology**, April 7-9, 2022, Braga, Portugal;
- João J, Rosa SS, Azevedo AM, Prazeres DMF, A chromatographic strategy for the purification of small heat shock protein nanocages, **MICROBIOTEC'21 – Congress of Microbiology and Biotechnology 2021**, November 23-26, 2021, virtual mode;
- João J, Prazeres DMF, Purification of small heat shock protein nanocages through a chromatographic approach, **PhD Open Days 2021**, November 22-24, 2021, Instituto Superior Técnico – Universidade de Lisboa, Lisbon, Portugal;

- João J, Rosa SS, Azevedo AM, Prazeres DMF, Chromatographic purification of small heat shock protein nanocages, **15th SBCN – International Symposium on Biochromatography and Nanoseparations**, May 17-20, 2021, virtual mode (recognized with the Poster Prize – VIJI Awards SBCN 2021).

References

1. Lee EJ. Recent advances in protein-based nanoparticles. Korean J Chem Eng. 2018;35(9):1765–78.

2. Bhaskar S, Lim S. Engineering protein nanocages as carriers for biomedical applications. NPG Asia Mater. 2017;9(4):e371–e371.

3. Schreiber A, Schiller SM. Nanobiotechnology of protein-based compartments: steps toward nanofactories. Bioinspired, Biomimetic and Nanobiomaterials. 2013;2(4):154–72.

4. Diekmann Y, Pereira-Leal JB. Evolution of intracellular compartmentalization. Biochem J. 2013;449(2):319–31.

5. Lim S. Protein nanocages: the versatile molecular shell. Asia Pacific Biotech News. 2013;17(1):39.

6. DasSarma S, DasSarma P. Gas vesicle nanoparticles for antigen display. Vaccines. 2015;3(3):686–702.

7. Andar AU, Karan R, Pecher WT, DasSarma P, Hedrich WD, Stinchcomb AL, et al. Microneedle-assisted skin permeation by nontoxic bioengineerable gas vesicle nanoparticles. Mol Pharmaceutics. 2017;14(3):953–8.

8. Hill AM, Salmond GPC. Microbial gas vesicles as nanotechnology tools: exploiting intracellular organelles for translational utility in biotechnology, medicine and the environment. Microbiology. 2020;166(6):501–9.

9. Kim JM, Kim YS, Kim YR, Choi MJ, DasSarma P, DasSarma S. Bioengineering of Halobacterium sp. NRC-1 gas vesicle nanoparticles with GvpC fusion protein produced in E. coli. Appl Microbiol Biotechnol. 2022;106(5–6):2043–52.

10. Pfeifer F. Recent advances in the study of gas vesicle proteins and application of gas vesicles in biomedical research. Life. 2022;12(9):1455.

11. Karan R, Renn D, Nozue S, Zhao L, Habuchi S, Allers T, et al. Bioengineering of air-filled protein nanoparticles by genetic and chemical functionalization. J Nanobiotechnol. 2023;21(1):108.

12. Pieters BJGE, van Eldijk MB, Nolte RJM, Mecinović J. Natural supramolecular protein assemblies. Chem Soc Rev. 2016;45(1):24–39.

13. Lee EJ, Lee NK, Kim IS. Bioengineered protein-based nanocage for drug delivery. Advanced Drug Delivery Reviews. 2016;106:157–71.

14. Diaz D, Care A, Sunna A. Bioengineering strategies for protein-based nanoparticles. Genes. 2018;9(7):370.

15. Roldão A, Mellado MCM, Castilho LR, Carrondo MJ, Alves PM. Virus-like particles in vaccine development. Expert Review of Vaccines. 2010;9(10):1149–76.

16. Hassani-Mehraban A, Creutzburg S, van Heereveld L, Kormelink R. Feasibility of Cowpea chlorotic mottle virus-like particles as scaffold for epitope presentations. BMC Biotechnol. 2015;15(1):80.

17. Wen AM, Shukla S, Saxena P, Aljabali AAA, Yildiz I, Dey S, et al. Interior engineering of a viral nanoparticle and its tumor homing properties. Biomacromolecules. 2012;13(12):3990–4001.

18. Li H, Onbe K, Liu Q, Iijima M, Tatematsu K, Seno M, et al. Synthesis and assembly of Hepatitis B virus envelope protein-derived particles in Escherichia coli. Biochemical and Biophysical Research Communications. 2017;490(2):155–60.

19. Fu Y, Li J. A novel delivery platform based on bacteriophage MS2 virus-like particles. Virus Research. 2016;211:9–16.

20. Patterson DP, Prevelige PE, Douglas T. Nanoreactors by programmed enzyme encapsulation inside the capsid of the bacteriophage P22. ACS Nano. 2012;6(6):5000–9.

21. Smith MT, Varner CT, Bush DB, Bundy BC. The incorporation of the A2 protein to produce novel Qβ virus-like particles using cell-free protein synthesis. Biotechnol Progress. 2012;28(2):549–55.

22. Lee EJ, Lee SJ, Kang YS, Ryu JH, Kwon KC, Jo E, et al. Engineered proteinticles for targeted delivery of siRNA to cancer cells. Adv Funct Mater. 2015;25(8):1279–86.

23. Zhen Z, Tang W, Chen H, Lin X, Todd T, Wang G, et al. RGD-modified apoferritin nanoparticles for efficient drug delivery to tumors. ACS Nano. 2013;7(6):4830–7.

24. Jeon JO, Kim S, Choi E, Shin K, Cha K, So IS, et al. Designed nanocage displaying ligand-specific peptide bunches for high affinity and biological activity. ACS Nano. 2013;7(9):7462–71.

25. Lee W, Seo J, Kwak S, Park EJ, Na DH, Kim S, et al. A double-chambered protein nanocage loaded with thrombin receptor agonist peptide (TRAP) and γ-carboxyglutamic acid of protein C (PC-Gla) for sepsis treatment. Adv Mater. 2015;27(42):6637–43.

26. Toita R, Murata M, Abe K, Narahara S, Piao JS, Kang JH, et al. A nanocarrier based on a genetically engineered protein cage to deliver doxorubicin to human hepatocellular carcinoma cells. Chem Commun. 2013;49(67):7442.

27. Ren D, Dalmau M, Randall A, Shindel MM, Baldi P, Wang SW. Biomimetic design of protein nanomaterials for hydrophobic molecular transport. Adv Funct Mater. 2012;22(15):3170–80.

28. Lai CY, Wiethoff CM, Kickhoefer VA, Rome LH, Nemerow GR. Vault nanoparticles containing an adenovirus-derived membrane lytic protein facilitate toxin and gene transfer. ACS Nano. 2009;3(3):691–9.

29. Suci P, Kang S, Gmür R, Douglas T, Young M. Targeted delivery of a photosensitizer to Aggregatibacter actinomycetemcomitans biofilm. Antimicrob Agents Chemother. 2010;54(6):2489–96.

30. Moon H, Lee J, Min J, Kang S. Developing genetically engineered encapsulin protein cage nanoparticles as a targeted delivery nanoplatform. Biomacromolecules. 2014;15(10):3794–801.

31. Ra JS, Shin HH, Kang S, Do Y. Lumazine synthase protein cage nanoparticles as antigen delivery nanoplatforms for dendritic cell-based vaccine development. Clin Exp Vaccine Res. 2014;3(2):227.

32. Vázquez E, Villaverde A. Engineering building blocks for self-assembling protein nanoparticles. Microb Cell Fact. 2010;9(1):101.

33. Theil EC. Ferritin protein nanocages-the story. Nanotechnol Percept. 2012;8(1):7–16.

34. Molino NM, Wang SW. Caged protein nanoparticles for drug delivery. Current Opinion in Biotechnology. 2014;28:75–82.

35. Corsi F, Mazzucchelli S. The potential of protein-based nanocages for imaging and drug delivery. Therapeutic Delivery. 2016;7(3):149–51.

36. Uchida M, Klem MT, Allen M, Suci P, Flenniken M, Gillitzer E, et al. Biological containers: protein cages as multifunctional nanoplatforms. Adv Mater. 2007;19(8):1025–42.

37. Giessen TW. Encapsulins: microbial nanocompartments with applications in biomedicine, nanobiotechnology and materials science. Curr Opin Chem Biol. 2016;34:1–10.

38. Chen H, Tan X, Fu Y, Dai H, Wang H, Zhao G, et al. The development of natural and designed protein nanocages for encapsulation and delivery of active compounds. Food Hydrocolloids. 2021;121:107004.

39. Murata M, Narahara S, Kawano T, Hamano N, Piao JS, Kang JH, et al. Design and function of engineered protein nanocages as a drug delivery system for targeting pancreatic cancer cells via neuropilin-1. Mol Pharmaceutics. 2015;12(5):1422–30.

40. Choi SH, Kwon IC, Hwang KY, Kim IS, Ahn HJ. Small heat shock protein as a multifunctional scaffold: integrated tumor targeting and caspase imaging within a single cage. Biomacromolecules. 2011;12(8):3099–106.

41. Kang HJ, Kang YJ, Lee YM, Shin HH, Chung SJ, Kang S. Developing an antibody-binding protein cage as a molecular recognition drug modular nanoplatform. Biomaterials. 2012;33(21):5423–30.

42. Kickhoefer VA, Han M, Raval-Fernandes S, Poderycki MJ, Moniz RJ, Vaccari D, et al. Targeting vault nanoparticles to specific cell surface receptors. ACS Nano. 2009;3(1):27–36.

43. Domingo GJ, Orru' S, Perham RN. Multiple display of peptides and proteins on a macromolecular scaffold derived from a multienzyme complex. J Mol Biol. 2001;305(2):259–67.

44. Zhang Y, Orner BP. Self-assembly in the ferritin nano-cage protein superfamily. IJMS. 2011;12(8):5406–21.

45. Doyle CM, Rumfeldt JA, Broom HR, Broom A, Stathopulos PB, Vassall KA, et al. Energetics of oligomeric protein folding and association. Arch Biochem Biophys. 2013;531(1–2):44–64.

46. Lv C, Zhang X, Liu Y, Zhang T, Chen H, Zang J, et al. Redesign of protein nanocages: the way from 0D, 1D, 2D to 3D assembly. Chem Soc Rev. 2021;50(6):3957–89.

47. Kim M, Rho Y, Jin KS, Ahn B, Jung S, Kim H, et al. pH-dependent structures of ferritin and apoferritin in solution: disassembly and reassembly. Biomacromolecules. 2011;12(5):1629–40.

48. Ferrer-Miralles N, Rodríguez-Carmona E, Corchero JL, García-Fruitós E, Vázquez E, Villaverde A. Engineering protein self-assembling in protein-based nanomedicines for drug delivery and gene therapy. Crit Rev Biotechnol. 2015;35(2):209–21.

49. Dalmau M, Lim S, Wang SW. Design of a pH-dependent molecular switch in a caged protein platform. Nano Lett. 2009;9(1):160–6.

50. Peng T, Lim S. Trimer-based design of pH-responsive protein cage results in soluble disassembled structures. Biomacromolecules. 2011;12(9):3131–8.

51. Sánchez-Sánchez L, Cadena-Nava RD, Palomares LA, Ruiz-Garcia J, Koay MST, Cornelissen JJMT, et al. Chemotherapy pro-drug activation by biocatalytic virus-like nanoparticles containing cytochrome P450. Enzyme and Microbial Technology. 2014;60:24–31.

52. Shen L, Zhou J, Wang Y, Kang N, Ke X, Bi S, et al. Efficient encapsulation of Fe3O4 nanoparticles into genetically engineered Hepatitis B core virus-like particles through a specific interaction for potential bioapplications. Small. 2015;11(9–10):1190–6.

53. Belval L, Hemmer C, Sauter C, Reinbold C, Fauny J, Berthold F, et al. Display of whole proteins on inner and outer surfaces of grapevine fanleaf virus-like particles. Plant Biotechnol J. 2016;14(12):2288–99.

54. Swift J, Butts CA, Cheung-Lau J, Yerubandi V, Dmochowski IJ. Efficient self-assembly of Archaeoglobus fulgidus ferritin around metallic cores. Langmuir. 2009;25(9):5219–25.

55. Lai YT, King NP, Yeates TO. Principles for designing ordered protein assemblies. Trends in Cell Biology. 2012;22(12):653–61.

56. Guimaraes CP, Witte MD, Theile CS, Bozkurt G, Kundrat L, Blom AEM, et al. Site-specific C-terminal and internal loop labeling of proteins using sortase-mediated reactions. Nat Protoc. 2013;8(9):1787–99.

57. Tamura A, Fukutani Y, Takami T, Fujii M, Nakaguchi Y, Murakami Y, et al. Packaging guest proteins into the encapsulin nanocompartment from Rhodococcus erythropolis N771: packaging guest proteins into the encapsulin nanocompartment. Biotechnol Bioeng. 2015;112(1):13–20.

58. Kar UK, Srivastava MK, Andersson Å, Baratelli F, Huang M, Kickhoefer VA, et al. Novel CCL21-vault nanocapsule intratumoral delivery inhibits lung cancer growth. Pandey S, editor. PLoS ONE. 2011;6(5):e18758.

59. Seebeck FP, Woycechowsky KJ, Zhuang W, Rabe JP, Hilvert D. A simple tagging system for protein encapsulation. J Am Chem Soc. 2006;128(14):4516–7.

60. Chen H, Zhang S, Xu C, Zhao G. Engineering protein interfaces yields ferritin disassembly and reassembly under benign experimental conditions. Chem Commun. 2016;52(46):7402–5.

61. Peng T, Lee H, Lim S. Design of a reversible inversed pH-responsive caged protein. Biomater Sci. 2015;3(4):627–35.

62. Flenniken ML, Willits DA, Harmsen AL, Liepold LO, Harmsen AG, Young MJ, et al. Melanoma and lymphocyte cell-specific targeting incorporated into a heat shock protein cage architecture. Chemistry & Biology. 2006;13(2):161–70.

63. Laplagne DA, Zylberman V, Ainciart N, Steward MW, Sciutto E, Fossati CA, et al. Engineering of a polymeric bacterial protein as a scaffold for the multiple display of peptides. Proteins. 2004;57(4):820–8.

64. Phippen SW, Stevens CA, Vance TDR, King NP, Baker D, Davies PL. Multivalent display of antifreeze proteins by fusion to self-assembling protein cages enhances ice-binding activities. Biochemistry. 2016;55(49):6811–20.

65. Falvo E, Tremante E, Fraioli R, Leonetti C, Zamparelli C, Boffi A, et al. Antibody–drug conjugates: targeting melanoma with cisplatin encapsulated in protein-cage nanoparticles based on human ferritin. Nanoscale. 2013;5(24):12278.

66. Tanaka H, Kato K, Yamashita E, Sumizawa T, Zhou Y, Yao M, et al. The structure of rat liver vault at 3.5 angstrom resolution. Science. 2009;323(5912):384–8.

67. Kong LB, Siva AC, Rome LH, Stewart PL. Structure of the vault, a ubiquitous cellular component. Structure. 1999;7(4):371–9.

68. Goldsmith LE, Pupols M, Kickhoefer VA, Rome LH, Monbouquette HG. Utilization of a protein “shuttle” to load vault nanocapsules with gold probes and proteins. ACS Nano. 2009;3(10):3175–83.

69. Kickhoefer VA, Garcia Y, Mikyas Y, Johansson E, Zhou JC, Raval-Fernandes S, et al. Engineering of vault nanocapsules with enzymatic and fluorescent properties. Proc Natl Acad Sci USA. 2005;102(12):4348–52.

70. Yu K, Yau YH, Sinha A, Tan T, Kickhoefer VA, Rome LH, et al. Modulation of the vault protein-protein interaction for tuning of molecular release. Sci Rep. 2017;7(1):14816.

71. Ryu SJ, An HJ, Oh YS, Choi HR, Ha MK, Park SC. On the role of major vault protein in the resistance of senescent human diploid fibroblasts to apoptosis. Cell Death Differ. 2008;15(11):1673–80.

72. Poderycki MJ, Kickhoefer VA, Kaddis CS, Raval-Fernandes S, Johansson E, Zink JI, et al. The vault exterior shell is a dynamic structure that allows incorporation of vault-associated proteins into its interior. Biochemistry. 2006;45(39):12184–93.

73. Kickhoefer VA, Vasu SK, Rome LH. Vaults are the answer, what is the question? Trends in Cell Biology. 1996;6(5):174–8.

74. Anderson DH, Kickhoefer VA, Sievers SA, Rome LH, Eisenberg D. Draft crystal structure of the vault shell at 9-Å resolution. Petsko GA, editor. PLoS Biol. 2007;5(11):e318.

75. San BH, Kim S, Moh SH, Lee H, Jung DY, Kim KK. Platinum nanoparticles encapsulated by aminopeptidase: a multifunctional bioinorganic nanohybrid catalyst. Angew Chem Int Ed. 2011;50(50):11924–9.

76. Kim D, San BH, Moh SH, Park H, Kim DY, Lee S, et al. Structural basis for the substrate specificity of PepA from Streptococcus pneumoniae, a dodecameric tetrahedral protease. Biochemical and Biophysical Research Communications. 2010;391(1):431–6.

77. Kang S, Lucon J, Varpness ZB, Liepold L, Uchida M, Willits D, et al. Monitoring biomimetic platinum nanocluster formation using mass spectrometry and cluster-dependent H2 production. Angew Chem Int Ed. 2008;47(41):7845–8.

78. Haikarainen T, Papageorgiou AC. Dps-like proteins: structural and functional insights into a versatile protein family. Cell Mol Life Sci. 2010;67(3):341–51.

79. Ilari A, Stefanini S, Chiancone E, Tsernoglou D. The dodecameric ferritin from Listeria innocua contains a novel intersubunit iron-binding site. Nat Struct Mol Biol. 2000;7:38–43.

80. Hempstead PD, Yewdall SJ, Fernie AR, Lawson DM, Artymiuk PJ, Rice DW, et al. Comparison of the three-dimensional structures of recombinant human H and horse L ferritins at high resolution. Journal of Molecular Biology. 1997 May 2;268(2):424–48.

81. Wang Z, Gao H, Zhang Y, Liu G, Niu G, Chen X. Functional ferritin nanoparticles for biomedical applications. Front Chem Sci Eng. 2017;11(4):633–46.

82. He D, Marles-Wright J. Ferritin family proteins and their use in bionanotechnology. N Biotechnol. 2015;32(6):651–7.

83. Kuruppu AI, Zhang L, Collins H, Turyanska L, Thomas NR, Bradshaw TD. An apoferritin-based drug delivery system for the tyrosine kinase inhibitor gefitinib. Adv Healthcare Mater. 2015;4(18):2816–21.

84. Palombarini F, Di Fabio E, Boffi A, Macone A, Bonamore A. Ferritin nanocages for protein delivery to tumor cells. Molecules. 2020;25(4):825.

85. Zhang C, Zhang X, Zhao G. Ferritin nanocage: a versatile nanocarrier utilized in the field of food, nutrition, and medicine. Nanomaterials. 2020;10(9):1894.

86. Lawson DM, Artymiuk PJ, Yewdall SJ, Smith JMA, Livingstone JC, Treffry A, et al. Solving the structure of human H ferritin by genetically engineering intermolecular crystal contacts. Nature. 1991;349:541–4.

87. Kim KK, Kim R, Kim SH. Crystal structure of a small heat-shock protein. Nature. 1998;394(6693):595–9.

88. Bova MP, Huang Q, Ding L, Horwitz J. Subunit exchange, conformational stability, and chaperone-like function of the small heat shock protein 16.5 from Methanococcus jannaschii. J Biol Chem. 2002;277(41):38468–75.

89. Kim KK, Yokota H, Santoso S, Lerner D, Kim R, Kim SH. Purification, crystallization, and preliminary X-ray crystallographic data analysis of small heat shock protein homolog from Methanococcus jannaschii, a hyperthermophile. Journal of Structural Biology. 1998;121(1):76–80.

90. Kim R, Kim KK, Yokota H, Kim SH. Small heat shock protein of Methanococcus jannaschii, a hyperthermophile. Proc Natl Acad Sci USA. 1998;95(16):9129–33.

91. Kim R, Lai L, Lee HH, Cheong GW, Kim KK, Wu Z, et al. On the mechanism of chaperone activity of the small heat-shock protein of Methanococcus jannaschii. Proc Natl Acad Sci USA. 2003;100(14):8151–5.

92. Flenniken ML, Willits DA, Brumfield S, Young MJ, Douglas T. The small heat shock protein cage from Methanococcus jannaschii is a versatile nanoscale platform for genetic and chemical modification. Nano Lett. 2003;3(11):1573–6.

93. Abedin MJ, Liepold L, Suci P, Young M, Douglas T. Synthesis of a cross-linked branched polymer network in the interior of a protein cage. J Am Chem Soc. 2009;131(12):4346–54.

94. Varpness Z, Peters JW, Young M, Douglas T. Biomimetic synthesis of a H2 catalyst using a protein cage architecture. Nano Lett. 2005;5(11):2306–9.

95. Izard T, Ævarsson A, Allen MD, Westphal AH, Perham RN, de Kok A, et al. Principles of quasi-equivalence and Euclidean geometry govern the assembly of cubic and dodecahedral cores of pyruvate dehydrogenase complexes. Proc Natl Acad Sci U S A. 1999;96(4):1240–5.

96. Jung HI, Cooper A, Perham RN. Identification of key amino acid residues in the assembly of enzymes into the pyruvate dehydrogenase complex of Bacillus stearothermophilus: a kinetic and thermodynamic analysis. Biochemistry. 2002;41(33):10446–53.

97. Dalmau M, Lim S, Chen HC, Ruiz C, Wang SW. Thermostability and molecular encapsulation within an engineered caged protein scaffold. Biotechnol Bioeng. 2008;101(4):654–64.

98. Milne JLS, Wu X, Borgnia MJ, Lengyel JS, Brooks BR, Shi D, et al. Molecular structure of a 9-MDa icosahedral pyruvate dehydrogenase subcomplex containing the E2 and E3 enzymes using cryoelectron microscopy. Journal of Biological Chemistry. 2006;281(7):4364–70.

99. Sutter M, Boehringer D, Gutmann S, Günther S, Prangishvili D, Loessner MJ, et al. Structural basis of enzyme encapsulation into a bacterial nanocompartment. Nat Struct Mol Biol. 2008;15(9):939–47.

100. Rahmanpour R, Bugg TDH. Assembly in vitro of Rhodococcus jostii RHA1 encapsulin and peroxidase DypB to form a nanocompartment. FEBS J. 2013;280(9):2097–104.

101. Cassidy-Amstutz C, Oltrogge L, Going CC, Lee A, Teng P, Quintanilla D, et al. Identification of a minimal peptide tag for in vivo and in vitro loading of encapsulin. Biochemistry. 2016;55(24):3461–8.

102. Ritsert K, Huber R, Turk D, Ladenstein R, Schmidt-Bäse K, Bacher A. Studies on the lumazine synthase/riboflavin synthase complex of Bacillus subtilis: crystal structure analysis of reconstituted, icosahedral β-subunit capsids with bound substrate analogue inhibitor at 2.4 Å resolution. Journal of Molecular Biology. 1995;253(1):151–67.

103. Zhang X, Meining W, Fischer M, Bacher A, Ladenstein R. X-ray structure analysis and crystallographic refinement of lumazine synthase from the hyperthermophile Aquifex aeolicus at 1.6 Å resolution: determinants of thermostability revealed from structural comparisons. Journal of Molecular Biology. 2001;306(5):1099–114.

104. Schott K, Ladenstein R, König A, Bacher A. The lumazine synthase-riboflavin synthase complex of Bacillus subtilis. Crystallization of reconstituted icosahedral beta-subunit capsids. Journal of Biological Chemistry. 1990;265(21):12686–9.

105. Shenton W, Mann S, Cölfen H, Bacher A, Fischer M. Synthesis of nanophase iron oxide in lumazine synthase capsids. Angew Chem Int Ed. 2001;40(2):442–5.

106. Ladenstein R, Fischer M, Bacher A. The lumazine synthase/riboflavin synthase complex: shapes and functions of a highly variable enzyme system. The FEBS Journal. 2013;280(11):2537–63.

107. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, et al. The Protein Data Bank. Nucleic Acids Research. 2000;28(1):235–42.

108. Sehnal D, Bittrich S, Deshpande M, Svobodová R, Berka K, Bazgier V, et al. Mol* Viewer: modern web app for 3D visualization and analysis of large biomolecular structures. Nucleic Acids Research. 2021;49(W1):W431–7.

109. Lai YT, Cascio D, Yeates TO. Structure of a 16-nm cage designed by using protein oligomers. Science. 2012;336(6085):1129–1129.

110. Gu C, Zhang T, Lv C, Liu Y, Wang Y, Zhao G. His-mediated reversible self-assembly of ferritin nanocages through two different switches for encapsulation of cargo molecules. ACS Nano. 2020;14(12):17080–90.

111. Malay AD, Miyazaki N, Biela A, Chakraborti S, Majsterkiewicz K, Stupka I, et al. An ultra-stable gold-coordinated protein cage displaying reversible assembly. Nature. 2019;569(7756):438–42.

112. Heddle JG, Fujiwara I, Yamadaki H, Yoshii S, Nishio K, Addy C, et al. Using the ring-shaped protein TRAP to capture and confine gold nanodots on a surface. Small. 2007;3(11):1950–6.

113. Heddle JG, Yokoyama T, Yamashita I, Park SY, Tame JRH. Rounding up: engineering 12-membered rings from the cyclic 11-mer TRAP. Structure. 2006;14(5):925–33.

114. Naskalska A, Borzęcka-Solarz K, Różycki J, Stupka I, Bochenek M, Pyza E, et al. Artificial protein cage delivers active protein cargos to the cell interior. Biomacromolecules. 2021;22(10):4146–54.

115. Majsterkiewicz K, Biela AP, Maity S, Sharma M, Piette BMAG, Kowalczyk A, et al. Artificial protein cage with unusual geometry and regularly embedded gold nanoparticles. Nano Lett. 2022;22(8):3187–95.

116. Stupka I, Azuma Y, Biela AP, Imamura M, Scheuring S, Pyza E, et al. Chemically induced protein cage assembly with programmable opening and cargo release. Sci Adv. 2022;8(1):eabj9424.

117. McCarthy S, Gonen S. Improved interface packing and design opportunities revealed by CryoEM analysis of a designed protein nanocage. Heliyon. 2022;8(12):e12280.

118. Hsia Y, Bale JB, Gonen S, Shi D, Sheffler W, Fong KK, et al. Design of a hyperstable 60-subunit protein icosahedron. Nature. 2016;535(7610):136–9.

119. Bruun TUJ, Andersson AMC, Draper SJ, Howarth M. Engineering a rugged nanoscaffold to enhance plug-and-display vaccination. ACS Nano. 2018;12(9):8855–66.

120. Votteler J, Ogohara C, Yi S, Hsia Y, Nattermann U, Belnap DM, et al. Designed proteins induce the formation of nanocage-containing extracellular vesicles. Nature. 2016;540(7632):292–5.

121. Gao R, Tan H, Li S, Ma S, Tang Y, Zhang K, et al. A prototype protein nanocage minimized from carboxysomes with gated oxygen permeability. Proceedings of the National Academy of Sciences. 2022;119(5):e2104964119.

122. Obata J, Kawakami N, Tsutsumi A, Nasu E, Miyamoto K, Kikkawa M, et al. Icosahedral 60-meric porous structure of designed supramolecular protein nanoparticle TIP60. Chem Commun. 2021;57(79):10226–9.

123. Ohara N, Kawakami N, Arai R, Adachi N, Moriya T, Kawasaki M, et al. Reversible assembly of an artificial protein nanocage using alkaline earth metal ions. J Am Chem Soc. 2023;145(1):216–23.

124. Kawakami N, Kondo H, Matsuzawa Y, Hayasaka K, Nasu E, Sasahara K, et al. Design of hollow protein nanoparticles with modifiable interior and exterior surfaces. Angew Chem Int Ed. 2018;57(38):12400–4.

125. Bellini M, Mazzucchelli S, Galbiati E, Sommaruga S, Fiandra L, Truffi M, et al. Protein nanocages for self-triggered nuclear delivery of DNA-targeted chemotherapeutics in cancer cells. J Control Release. 2014;196:184–96.

126. Guan X, Chang Y, Sun J, Song J, Xie Y. Engineered Hsp protein nanocages for siRNA delivery. Macromol Biosci. 2018;18(5):1800013.

127. Ji P, Wang X, Yin J, Mou Y, Huang H, Ren Z. Selective delivery of curcumin to breast cancer cells by self-targeting apoferritin nanocages with pH-responsive and low toxicity. Drug Delivery. 2022;29(1):986–96.

128. SIB Swiss Institute of Bioinformatics. ExPASy - ProtParam Tool [Internet]. [cited 2025 Jan 5]. Available from: https://web.expasy.org/protparam/

129. Wang XY, Chen X, Manjili MH, Repasky E, Henderson R, Subjeck JR. Targeted immunotherapy using reconstituted chaperone complexes of heat shock protein 110 and melanoma-associated antigen gp100. Cancer Res. 2003;63(10):2553–60.

130. Kar UK, Jiang J, Champion CI, Salehi S, Srivastava M, Sharma S, et al. Vault nanocapsules as adjuvants favor cell-mediated over antibody-mediated immune responses following immunization of mice. Rodrigues MM, editor. PLoS ONE. 2012;7(7):e38553.

131. Molino NM, Anderson AKL, Nelson EL, Wang SW. Biomimetic protein nanoparticles facilitate enhanced dendritic cell activation and cross-presentation. ACS Nano. 2013;7(11):9743–52.

132. Han JA, Kang YJ, Shin C, Ra JS, Shin HH, Hong SY, et al. Ferritin protein cage nanoparticles as versatile antigen delivery nanoplatforms for dendritic cell (DC)-based vaccine development. Nanomedicine: Nanotechnology, Biology and Medicine. 2014;10(3):561–9.

133. Fan K, Cao C, Pan Y, Lu D, Yang D, Feng J, et al. Magnetoferritin nanoparticles for targeting and visualizing tumour tissues. Nature Nanotech. 2012;7(7):459–64.

134. Wang Z, Huang P, Jacobson O, Wang Z, Liu Y, Lin L, et al. Biomineralization-inspired synthesis of copper sulfide–ferritin nanocages as cancer theranostics. ACS Nano. 2016;10(3):3453–60.

135. Kawano T, Murata M, Kang JH, Piao JS, Narahara S, Hyodo F, et al. Ultrasensitive MRI detection of spontaneous pancreatic tumors with nanocage-based targeted contrast agent. Biomaterials. 2018;152:37–46.

136. Allen M, Willits D, Mosolf J, Young M, Douglas T. Protein cage constrained synthesis of ferrimagnetic iron oxide nanoparticles. Adv Mater. 2002;14(21):1562–5.

137. Ensign D, Young M, Douglas T. Photocatalytic synthesis of copper colloids from Cu(II) by the ferrihydrite core of ferritin. Inorg Chem. 2004;43(11):3441–6.

138. João J, Lampreia J, Prazeres DMF, Azevedo AM. Manufacturing of bacteriophages for therapeutic applications. Biotechnology Advances. 2021;49:107758.

139. Owczarek B, Gerszberg A, Hnatuszko-Konka K. A brief reminder of systems of production and chromatography-based recovery of recombinant protein biopharmaceuticals. BioMed Research International. 2019;2019:1–13.

140. Puetz J, Wurm FM. Recombinant proteins for industrial versus pharmaceutical purposes: a review of process and pricing. Processes. 2019;7(8):476.

141. Tripathi NK, Shrivastava A. Recent developments in bioprocessing of recombinant proteins: expression hosts and process development. Front Bioeng Biotechnol. 2019;7:420.

142. Wang M, Kickhoefer VA, Rome LH, Foellmer OK, Mahendra S. Synthesis and assembly of human vault particles in yeast. Biotechnology and Bioengineering. 2018;115(12):2941–50.

143. Champion CI, Kickhoefer VA, Liu G, Moniz RJ, Freed AS, Bergmann LL, et al. A vault nanoparticle vaccine induces protective mucosal immunity. Unutmaz D, editor. PLoS ONE. 2009;4(4):e5409.

144. Wang M, Abad D, Kickhoefer VA, Rome LH, Mahendra S. Vault nanoparticles packaged with enzymes as an efficient pollutant biodegradation technology. ACS Nano. 2015;9(11):10931–40.

145. Stephen AG, Raval-Fernandes S, Huynh T, Torres M, Kickhoefer VA, Rome LH. Assembly of vault-like particles in insect cells expressing only the major vault protein. Journal of Biological Chemistry. 2001;276(26):23217–20.

146. Tripathi NK. Production and purification of recombinant proteins from Escherichia coli. ChemBioEng Reviews. 2016;3(3):116–33.

147. Zou W, Liu X, Chen D, Wang J, Zhao X, Li J, et al. Expression, purification, and characterization of recombinant human H-chain ferritin. Preparative Biochemistry & Biotechnology. 2016;46(8):833–7.

148. Zou W, Liu X, Zhao X, Wang J, Chen D, Li J, et al. Expression, purification, and characterization of recombinant human L-chain ferritin. Protein Expression and Purification. 2016;119:63–8.

149. Cristie-David AS, Marsh ENG. Metal-dependent assembly of a protein nano-cage. Protein Science. 2019;28(9):1620–9.

150. Santambrogio P, Cozzi A, Levi S, Rovida E, Magni F, Albertini A, et al. Functional and immunological analysis of recombinant mouse H- and L-ferritins from Escherichia coli. Protein Expression and Purification. 2000;19(1):212–8.

151. Kawano T, Murata M, Piao J, Narahara S, Hamano N, Kang JH, et al. Systemic delivery of protein nanocages bearing CTT peptides for enhanced imaging of MMP-2 expression in metastatic tumor models. IJMS. 2014;16(1):148–58.

152. Palombarini F, Ghirga F, Boffi A, Macone A, Bonamore A. Application of crossflow ultrafiltration for scaling up the purification of a recombinant ferritin. Protein Expression and Purification. 2019;163:105451.

153. Martín F, Carreño A, Mendoza R, Caruana P, Rodriguez F, Bravo M, et al. All-in-one biofabrication and loading of recombinant vaults in human cells. Biofabrication. 2022;14(2):025018.

154. Mrazek J. Cell-free methods of producing vault particles and vault particles resulting therefrom [Internet]. WO2016049122A1, 2016 [cited 2025 Jan 5]. Available from: https://patents.google.com/patent/WO2016049122A1/en

155. Colant N, Melinek B, Teneb J, Goldrick S, Rosenberg W, Frank S, et al. A rational approach to improving titer in Escherichia coli-based cell-free protein synthesis reactions. Biotechnol Progress. 2021;37(1):e3062.

156. Des Soye BJ, Davidson SR, Weinstock MT, Gibson DG, Jewett MC. Establishing a high-yielding cell-free protein synthesis platform derived from Vibrio natriegens. ACS Synth Biol. 2018;7(9):2245–55.

157. Failmezger J, Scholz S, Blombach B, Siemann-Herzberg M. Cell-free protein synthesis from fast-growing Vibrio natriegens. Front Microbiol. 2018;9(1146):1–10.

158. Gregorio NE, Levine MZ, Oza JP. A user's guide to cell-free protein synthesis. MPs. 2019;2(1):24.

159. Kwon YC, Jewett MC. High-throughput preparation methods of crude extract for robust cell-free protein synthesis. Sci Rep. 2015;5(1):8663.

160. Tran K, Gurramkonda C, Cooper MA, Pilli M, Taris JE, Selock N, et al. Cell-free production of a therapeutic protein: expression, purification, and characterization of recombinant streptokinase using a CHO lysate. Biotechnol Bioeng. 2018;115(1):92–102.

161. Silva F, Sitia L, Allevi R, Bonizzi A, Sevieri M, Morasso C, et al. Combined method to remove endotoxins from protein nanocages for drug delivery applications: the case of human ferritin. Pharmaceutics. 2021;13(2):229.

162. Wang Z, Zhao Y, Zhang S, Chen X, Sun G, Zhang B, et al. Re-engineering the inner surface of ferritin nanocage enables dual drug payloads for synergistic tumor therapy. Theranostics. 2022;12(4):1800–15.

163. Sana B, Johnson E, Sheah K, Poh CL, Lim S. Iron-based ferritin nanocore as a contrast agent. Biointerphases. 2010;5(3):FA48–52.

164. Johnson E, Cascio D, Sawaya MR, Gingery M, Schröder I. Crystal structures of a tetrahedral open pore ferritin from the hyperthermophilic archaeon Archaeoglobus fulgidus. Structure. 2005;13(4):637–48.

165. Kanekiyo M, Wei CJ, Yassine HM, McTamney PM, Boyington JC, Whittle JRR, et al. Self-assembling influenza nanoparticle vaccines elicit broadly neutralizing H1N1 antibodies. Nature. 2013;499(7456):102–6.

166. Tiainen P, Gustavsson PE, Ljunglöf A, Larsson PO. Superporous agarose anion exchangers for plasmid isolation. Journal of Chromatography A. 2007;1138(1):84–94.

167. Prazeres DM de F. Chromatographic separation of plasmid DNA using macroporous beads. In: Macroporous Polymers: Production, Properties and Biotechnological/Biomedical Applications. New York: CRC Press; 2009. p. 335–61.

168. Kim J won, Heu W, Jeong S, Kim HS. Genetically functionalized ferritin nanoparticles with a high-affinity protein binder for immunoassay and imaging. Analytica Chimica Acta. 2017;988:81–8.

169. de Turris V, Cardoso Trabuco M, Peruzzi G, Boffi A, Testi C, Vallone B, et al. Humanized archaeal ferritin as a tool for cell targeted delivery. Nanoscale. 2017;9(2):647–55.

170. Franco E, Garcia-Recio V, Jiménez P, Garrosa M, Girbés T, Cordoba-Diaz M, et al. Endotoxins from a pharmacopoeial point of view. Toxins. 2018;10(8):331.

171. Schwarz H, Schmittner M, Duschl A, Horejs-Hoeck J. Residual endotoxin contaminations in recombinant proteins are sufficient to activate human CD1c+ dendritic cells. Li L, editor. PLoS ONE. 2014;9(12):e113840.

172. Boyton I, Goodchild SC, Diaz D, Elbourne A, Collins-Praino LE, Care A. Characterizing the dynamic disassembly/reassembly mechanisms of encapsulin protein nanocages. ACS Omega. 2022;7(1):823–36.

173. Stühn L, Auernhammer J, Dietz C. pH-depended protein shell dis- and reassembly of ferritin nanoparticles revealed by atomic force microscopy. Sci Rep. 2019;9(1):17755.

174. Heinze K, Sasaki E, King NP, Baker D, Hilvert D, Wuite GJL, et al. Protein nanocontainers from nonviral origin: testing the mechanics of artificial and natural protein cages by AFM. J Phys Chem B. 2016;120(26):5945–52.

175. Llauró A, Guerra P, Kant R, Bothner B, Verdaguer N, de Pablo PJ. Decrease in pH destabilizes individual vault nanocages by weakening the inter-protein lateral interaction. Sci Rep. 2016;6(1):34143.

176. Schmitt S, Nuhn L, Barz M, Butt H, Koynov K. Shining light on polymeric drug nanocarriers with fluorescence correlation spectroscopy. Macromol Rapid Commun. 2022;43(12):2100892.

177. Effio CL, Oelmeier SA, Hubbuch J. High-throughput characterization of virus-like particles by interlaced size-exclusion chromatography. Vaccine. 2016;34(10):1259–67.

178. Steppert P, Burgstaller D, Klausberger M, Tover A, Berger E, Jungbauer A. Quantification and characterization of virus-like particles by size-exclusion chromatography and nanoparticle tracking analysis. Journal of Chromatography A. 2017;1487:89–99.

179. Moleirinho MG, Silva RJS, Alves PM, Carrondo MJT, Peixoto C. Current challenges in biotherapeutic particles manufacturing. Expert Opinion on Biological Therapy. 2020;20(5):451–65.

180. Papathanasiou MM, Kontoravdi C. Engineering challenges in therapeutic protein product and process design. Current Opinion in Chemical Engineering. 2020;27:81–8.

181. Aguilar PP, González-Domínguez I, Schneider TA, Gòdia F, Cervera L, Jungbauer A. At-line multi-angle light scattering detector for faster process development in enveloped virus-like particle purification. J Sep Sci. 2019;(42):2640–9.

182. Patel BA, Gospodarek A, Larkin M, Kenrick SA, Haverick MA, Tugcu N, et al. Multi-angle light scattering as a process analytical technology measuring real-time molecular weight for downstream process control. mAbs. 2018;1–6.

183. Kasyutich O, Ilari A, Fiorillo A, Tatchev D, Hoell A, Ceci P. Silver ion incorporation and nanoparticle formation inside the cavity of Pyrococcus furiosus ferritin: structural and size-distribution analyses. J Am Chem Soc. 2010;132(10):3621–7.

184. Ren D, Kratz F, Wang SW. Protein nanocapsules containing doxorubicin as a pH-responsive delivery system. Small. 2011;7(8):1051–60.

185. Hopcroft NH, Manfredo A, Wendt AL, Brzozowski AM, Gollnick P, Antson AA. The interaction of RNA with TRAP: the role of triplet repeats and separating spacer nucleotides. Journal of Molecular Biology. 2004;338(1):43–53.

186. João J, Prazeres DMF. Manufacturing of non-viral protein nanocages for biotechnological and biomedical applications. Front Bioeng Biotechnol. 2023;11:1200729.

187. Francis DM, Page R. Strategies to optimize protein expression in E. coli. Curr Protoc Protein Sci. 2010;61(1):5241–52429.

188. Bhatwa A, Wang W, Hassan YI, Abraham N, Li XZ, Zhou T. Challenges associated with the formation of recombinant protein inclusion bodies in Escherichia coli and strategies to address them for industrial applications. Front Bioeng Biotechnol. 2021;9.

189. Kormanová Ľ, Rybecká S, Levarski Z, Struhárňanská E, Levarská L, Blaško J, et al. Comparison of simple expression procedures in novel expression host Vibrio natriegens and established Escherichia coli system. Journal of Biotechnology. 2020 Sep 10;321:57–67.

190. Gibson B, Wilson DJ, Feil E, Eyre-Walker A. The distribution of bacterial doubling times in the wild. Proc Biol Sci. 2018;285(1880):20180789.

191. Studier FW, Rosenberg AH, Dunn JJ, Dubendorff JW. Use of T7 RNA polymerase to direct expression of cloned genes. Methods Enzymol. 1990;185:60–89.

192. Rosano GL, Ceccarelli EA. Recombinant protein expression in Escherichia coli: advances and challenges. Front Microbiol. 2014;5.

193. Rosano GL, Morales ES, Ceccarelli EA. New tools for recombinant protein production in Escherichia coli: A 5-year update. Protein Sci. 2019 Aug;28(8):1412–22.

194. Eagon RG. Pseudomonas natriegens, a marine bacterium with a generation time of less than 10 minutes. J Bacteriol. 1962;83(4):736–7.

195. Thoma F, Blombach B. Metabolic engineering of Vibrio natriegens. Essays in Biochemistry. 2021;65(2):381–92.

196. Hoffart E, Grenz S, Lange J, Nitschel R, Müller F, Schwentner A, et al. High substrate uptake rates empower Vibrio natriegens as production host for industrial biotechnology. Drake HL, editor. Appl Environ Microbiol. 2017;83(22):e01614-17.

197. Long CP, Gonzalez JE, Cipolla RM, Antoniewicz MR. Metabolism of the fast-growing bacterium Vibrio natriegens elucidated by 13C metabolic flux analysis. Metabolic Engineering. 2017 Nov 1;44:191–7.

198. Weinstock MT, Hesek ED, Wilson CM, Gibson DG. Vibrio natriegens as a fast-growing host for molecular biology. Nature Methods. 2016;13(10):849–51.

199. Örencik C, Müller S, Kirner T, Amann E. An analysis and optimization of growth condition requirements of the fast-growing bacterium Vibrio natriegens. bioRxiv. 2019;775437.

200. Xu J, Dong F, Wu M, Tao R, Yang J, Wu M, et al. Vibrio natriegens as a pET-compatible expression host complementary to Escherichia coli. Front Microbiol. 2021;12:627181.

201. Lee HH, Ostrov N, Wong BG, Gold MA, Khalil AS, Church GM. Vibrio natriegens, a new genomic powerhouse. bioRxiv. 2016;058487.

202. Liu X, Han X, Peng Y, Tan C, Wang J, Xue H, et al. Rapid production of L-DOPA by Vibrio natriegens, an emerging next-generation whole-cell catalysis chassis. Microb Biotechnol. 2022;15(5):1610–21.

203. Becker W, Wimberger F, Zangger K. Vibrio natriegens: an alternative expression system for the high-yield production of isotopically labeled proteins. Biochemistry. 2019;58(25):2799–803.

204. Tschirhart T, Shukla V, Kelly EE, Schultzhaus Z, NewRingeisen E, Erickson JS, et al. Synthetic biology tools for the fast-growing marine bacterium Vibrio natriegens. ACS Synth Biol. 2019;8(9):2069–79.

205. Skulj M, Okrslar V, Jalen S, Jevsevar S, Slanc P, Strukelj B, et al. Improved determination of plasmid copy number using quantitative real-time PCR for monitoring fermentation processes. Microb Cell Fact. 2008;7:6.

206. Schneider CA, Rasband WS, Eliceiri KW. NIH image to ImageJ: 25 years of image analysis. Nat Methods. 2012;9(7):671–5.

207. Monod J. The growth of bacterial cultures. Annual Review of Microbiology. 1949;3(Volume 3, 1949):371–94.

208. Powell EO. Growth rate and generation time of bacteria, with special reference to continuous culture. Journal of General Microbiology. 1956;15(3):492–511.

209. Khalilzadeh R, Shojaosadati SA, Bahrami A, Maghsoudi N. Fed-batch cultivation of recombinant Escherichia coli producing human interferon-γ under controlled specific growth rate. Iranian Journal of Biotechnology. 2004;2(2):113–22.

210. Han MJ, Jeong KJ, Yoo JS, Lee SY. Engineering Escherichia coli for increased productivity of serine-rich proteins based on proteome profiling. Appl Environ Microbiol. 2003;69(10):5772–81.

211. Khalilzadeh R, Shojaosadati SA, Bahrami A, Maghsoudi N. Over-expression of recombinant human interferon-gamma in high cell density fermentation of Escherichia coli. Biotechnology Letters. 2003;25(23):1989–92.

212. Mairhofer J, Scharl T, Marisch K, Cserjan-Puschmann M, Striedner G. Comparative transcription profiling and in-depth characterization of plasmid-based and plasmid-free Escherichia coli expression systems under production conditions. Applied and Environmental Microbiology. 2013;79(12):3802–12.

213. Xu J, Yang S, Yang L. Vibrio natriegens as a host for rapid biotechnology. Trends in Biotechnology. 2022;40(4):381–4.

214. Ellis GA, Tschirhart T, Spangler J, Walper SA, Medintz IL, Vora GJ. Exploiting the feedstock flexibility of the emergent synthetic biology chassis Vibrio natriegens for engineered natural product production. Marine Drugs. 2019;17(12).

215. Mojica N, Kersten F, Montserrat-Canals M, Huhn III GR, Tislevoll AM, Cordara G, et al. Using Vibrio natriegens for high-yield production of challenging expression targets and for protein perdeuteration. Biochemistry. 2024;63(5):587–98.

216. Rogers-Broadway KR, Karteris E. Amplification efficiency and thermal stability of qPCR instrumentation: Current landscape and future perspectives. Exp Ther Med. 2015;10(4):1261–4.

217. Hou Y, Zhang H, Miranda L, Lin S. Serious overestimation in quantitative PCR by circular (supercoiled) plasmid standard: microalgal pcna as the model gene. PLOS ONE. 2010 Mar;5(3):1–7.

218. del Solar G, Giraldo R, Ruiz-Echevarría MJ, Espinosa M, Díaz-Orejas R. Replication and control of circular bacterial plasmids. Microbiol Mol Biol Rev. 1998;62(2):434–64.

219. Guss AM, Olson DG, Caiazza NC, Lynd LR. Dcm methylation is detrimental to plasmid transformation in Clostridium thermocellum. Biotechnology for Biofuels. 2012 May 6;5(1):30.

220. Rouches MV, Xu Y, Cortes LBG, Lambert G. A plasmid system with tunable copy number. Nature Communications. 2022;13(1):3908.

221. Kram KE, Finkel SE. Rich medium composition affects Escherichia coli survival, glycation, and mutation frequency during long-term batch culture. Appl Environ Microbiol. 2015;81(13):4442–50.

222. Klumpp S, Zhang Z, Hwa T. Growth rate-dependent global effects on gene expression in bacteria. Cell. 2009;139(7):1366–75.

223. James J, Yarnall B, Koranteng A, Gibson J, Rahman T, Doyle DA. Protein over-expression in Escherichia coli triggers adaptation analogous to antimicrobial resistance. Microbial Cell Factories. 2021;20(1):13.

224. Xu JM, Wu ZS, Zhao KJ, Xi ZJ, Wang LY, Cheng F, et al. IPTG-induced high protein expression for whole-cell biosynthesis of L-phosphinothricin. Biotechnol J. 2023;18(9):e2300027.

225. Farewell A, Neidhardt FC. Effect of temperature on in vivo protein synthetic capacity in Escherichia coli. J Bacteriol. 1998;180(17):4704–10.

226. Sadeghi HMM, Rabbani M, Rismani E, Moazen F, Khodabakhsh F, Dormiani K, et al. Optimization of the expression of reteplase in Escherichia coli. Res Pharm Sci. 2011;6(2):87–92.

227. Rajacharya GH, Sharma A, Yazdani SS. Proteomics and metabolic burden analysis to understand the impact of recombinant protein production in E. coli. Scientific Reports. 2024;14(1):12271.

228. Azam TA, Iwata A, Nishimura A, Ueda S, Ishihama A. Growth phase-dependent variation in protein composition of the Escherichia coli nucleoid. Journal of Bacteriology. 1999;181(20):6361–70.

229. Takahashi M, Aoyagi H. Analysis and effect of conventional flasks in shaking culture of Escherichia coli. AMB Express. 2020;10(1):77.

230. McDaniel LE, Bailey EG, Zimmerli A. Effect of oxygen-supply rates on growth of Escherichia coli. Appl Microbiol. 1965;13(1):109–14.

231. Ganjave SD, Dodia H, Sunder AV, Madhu S, Wangikar PP. High cell density cultivation of E. coli in shake flasks for the production of recombinant proteins. Biotechnology Reports. 2022;33:e00694.

232. Zulkifly NAH, Selas Castiñeiras T, Overton TW. Optimisation of recombinant TNFα production in Escherichia coli using GFP fusions and flow cytometry. Front Bioeng Biotechnol. 2023;11:1171823.

233. Smith AD, Tschirhart T, Compton J, Hennessa TM, VanArsdale E, Wang Z. Rapid, high-titer biosynthesis of melanin using the marine bacterium Vibrio natriegens. Front Bioeng Biotechnol. 2023;11:1239756.

234. Eichmann J, Oberpaul M, Weidner T, Gerlach D, Czermak P. Selection of high producers from combinatorial libraries for the production of recombinant proteins in Escherichia coli and Vibrio natriegens. Frontiers in Bioengineering and Biotechnology. 2019;7.

235. Biener R, Horn T, Komitakis A, Schendel I, König L, Hauenstein A, et al. High-cell-density cultivation of Vibrio natriegens in a low-chloride chemically defined medium. Appl Microbiol Biotechnol. 2023;107(23):7043–54.

236. Nguyen TH, Tran TVA, Dam TH, Pham TA, Le TH, Le T. Trends in biotechnology: Vibrio natriegens as potential micro-factory for valorization of crustacean waste. Academia Journal of Biology. 2024;46(4):71–89.

237. Hoff J, Daniel B, Stukenberg D, Thuronyi BW, Waldminghaus T, Fritz G. Vibrio natriegens: an ultrafast-growing marine bacterium as emerging synthetic biology chassis. Environ Microbiol. 2020;22(10):4394–408.

238. Wiegand DJ, Lee HH, Ostrov N, Church GM. Establishing a cell-free Vibrio natriegens expression system. ACS Synth Biol. 2018;7(10):2475–9.

239. Smith M, Hernández JS, Messing S, Ramakrishnan N, Higgins B, Mehalko J, et al. Producing recombinant proteins in Vibrio natriegens. Microb Cell Fact. 2024;23(1):208.

240. Gemünde A, Gail J, Holtmann D. Anodic respiration of Vibrio natriegens in a bioelectrochemical system. ChemSusChem. 2023;16(16):e202300181.

241. Mendes BB, Conniot J, Avital A, Yao D, Jiang X, Zhou X, et al. Nanodelivery of nucleic acids. Nature Reviews Methods Primers. 2022;2(1):24.

242. Stetefeld J, McKenna SA, Patel TR. Dynamic light scattering: a practical guide and applications in biomedical sciences. Biophys Rev. 2016;8(4):409–27.

243. Thermo Fisher Scientific. Alexa Fluor 647 Protein Labelling Kit - User Guide. 2017.

244. Valeur B. Molecular Fluorescence: Principles and Applications. Weinheim: Wiley-VCH; 2002.

245. Lakowicz JR. Principles of Fluorescence Spectroscopy. 3rd Edition. New York: Springer; 2006. 954 p.

246. Kapusta P, PicoQuant GmbH. Absolute Diffusion Coefficients: Compilation of Reference Data for FCS Calibration - Application Note. 2010.