

UNIVERSIDADE DE LISBOA
INSTITUTO SUPERIOR TÉCNICO

**Exploiting altered metabolism for cancer treatment:
In search for novel Hexokinase 2 inhibitors**

Sara Cristina Nobre Garcia

Supervisor: Doctor Maria Matilde Soares Duarte Marques

Co-Supervisor: Doctor Rita Alexandra do Nascimento Cardoso Guedes

Thesis approved in public session to obtain the PhD Degree in Chemistry

Jury final classification: Pass with Distinction

2022

UNIVERSIDADE DE LISBOA
INSTITUTO SUPERIOR TÉCNICO

**Exploiting altered metabolism for cancer treatment:
In search for novel Hexokinase 2 inhibitors**

Sara Cristina Nobre Garcia

Supervisor: Doctor Maria Matilde Soares Duarte Marques

Co-Supervisor: Doctor Rita Alexandra do Nascimento Cardoso Guedes

Thesis approved in public session to obtain the PhD Degree in Chemistry

Jury final classification: Pass with Distinction

Jury

Chairperson: Doctor Maria Teresa Nogueira Leal da Silva Duarte, Instituto Superior Técnico, Universidade de Lisboa

Members of the committee:

Doctor Nissim Hay, College of Medicine, University of Illinois at Chicago, EUA

Doctor Rui Ferreira Alves Moreira, Faculdade de Farmácia. Universidade de Lisboa

Doctor Maria Matilde Soares Duarte Marques, Instituto Superior Técnico, Universidade de Lisboa

Doctor Maria Emília da Silva Pereira de Sousa, Faculdade de Farmácia. Universidade do Porto

Doctor Raúl José Silvério Bernardino, Escola Superior de Turismo e Tecnologia do Mar, Politécnico de Leiria

Doctor José Gonçalo Deira Duarte de Campos Justino, Instituto Superior Técnico, Universidade de Lisboa

FUNDING INSTITUTION

Fundação para a Ciência e Tecnologia

2022

Resumo

Exploração do metabolismo central para o tratamento do cancro: à procura de novos inibidores da hexocinase 2

A alteração da via glicolítica é uma característica reconhecida nas células cancerígenas. O aumento da glicólise, e a sua modulação representa assim uma oportunidade no tratamento do cancro. Várias enzimas estão sobreexpressas na via glicolítica de diferentes tipos de células cancerígenas, nomeadamente a hexocinase 2 (HK2). Curiosamente, esta enzima não está apenas envolvida no primeiro, e mais decisivo, passo da glicólise, e consequentemente em outras subvias do metabolismo central, mas também na imortalização das células cancerígenas. Quando a HK2 se encontra cataliticamente ativa, é capaz de interagir com a membrana mitocondrial externa (MOM), prevenindo a normal sinalização pro-apoptótica. O bloqueio desta interação promove a ligação de proteínas pro-apoptóticas à MOM, estimulando a apoptose nas células cancerígenas. Assim, a inibição da HK2 é proposta como uma estratégia para reduzir a maior fonte de energia das células cancerígenas, diminuindo a proliferação celular e promovendo a apoptose.

Com o objetivo de encontrar compostos “*hit*” capazes de interferir com a atividade catalítica da HK2, foi implementada uma estratégia multidisciplinar que envolveu um rastreio virtual baseado na estrutura (SBVS) e a sua validação que incluiu estudos *in silico*, síntese orgânica e uma componente de caracterização química, bioquímica e biológica dos compostos obtidos. Após a preparação do protocolo de SBVS usando como método o docking molecular implementado no programa GOLD V.5.2, foram rastreadas várias bibliotecas de compostos virtuais tais como DrugBank, NCI (National Cancer Institute) e InterBioScreen (natural products database). Deste processo resultaram 2981 moléculas que foram consideradas potenciais inibidoras da HK2. Estas moléculas foram ainda filtradas com base na previsão das suas propriedades drug-like. O protocolo foi validado experimentalmente através de um ensaio bioquímico, utilizando-se 64 das moléculas selecionadas (10 μ M) e a enzima HK2. Os resultados demonstraram que todos os compostos afetaram a atividade da HK2. Destes, 22 compostos evidenciaram ter capacidade para inibir a HK2 na mesma extensão ou mais proeminentemente que o inibidor conhecido 3-bromopiruvato. Deste modo, comprovou-se que o protocolo de SBVS utilizado é capaz de identificar moléculas bioativas cujas estruturas poderão ser otimizadas e/ou submetidas a testes mais elaborados. Neste projeto foi utilizada a síntese orgânica como ferramenta para a obtenção de um inibidor conhecido e de uma família de potenciais inibidores. De acordo com o conhecimento obtido *a priori* relativamente à sensibilidade de vários tipos de linhas celulares à inibição/eliminação da HK2, foi avaliado um conjunto de 62 moléculas selecionadas pelo seu potencial efeito na inibição da proliferação celular de células do carcinoma hepatocelular. As oito moléculas com maior potencial foram avaliadas como inibidores da proliferação de outras linhas celulares, tendo também sido testado o seu efeito na glicólise destas células, paralelamente com dois inibidores da HK2 conhecidos. Foram utilizadas células CHO modificadas para a expressão seletiva da HK1, HK2 ou HK4 para determinar a seletividade de quatro dos inibidores da glicólise descobertos neste estudo. Os quatro compostos inibiram a glicólise a 10 μ M em todos os tipos de células CHO testados. Contudo, resultados preliminares mostraram que um destes compostos é capaz de inibir com elevada seletividade, a uma concentração mais reduzida (1.5 μ M), as células CHO que expressam especificamente a isozima HK2. Embora sejam necessários mais estudos para confirmar o potencial desta molécula, estes resultados encorajam a consideração da HK2 como alvo em futuras campanhas de descoberta de fármacos contra o cancro.

Abstract

Altered central metabolism is a well-established hallmark of cancer cells. The glycolytic pathway has been demonstrated to be consistently increased in these cells, representing a pool of opportunities to be explored for cancer treatment. Several enzymes involved in glycolysis are overexpressed in different types of cancer cells, namely hexokinase 2 (HK2). This enzyme is not only involved in the first and most determinant step of glycolysis, and subsequently in the different branched pathways, but also in the immortalization of cancer cells. When catalytically active, HK2 is able to bind to the mitochondrial outer membrane (MOM), preventing the normal pro-apoptotic signaling. The disruption of this interaction promotes the binding of pro-apoptotic proteins to MOM, stimulating the enhancement of apoptosis in cancer cells. The inhibition of HK2 is proposed as a strategy to reduce the main source of energy to cancer cells, thus significantly decreasing cancer cell proliferation and increasing apoptosis.

As an effort to find hit compounds able to interfere with the HK2 catalytic activity, a multidisciplinary strategy that involved a structure-based virtual screening (SBVS) campaign and its validation was implemented. This effort included *in silico* studies and organic synthesis, complemented with chemical, biochemical and biological characterization of the molecules. After the design of the SBVS protocol using molecular docking calculations with GOLD 5.2 software, several virtual compound libraries such as DrugBank, NCI and InterBioScreen Natural Products were screened. As result, 2981 molecules with the potential to act as new HK2 inhibitors were identified. After filtering those molecules according to the prediction of drug-like properties, biochemical validation of the aforementioned protocol was conducted using 64 selected molecules at 10 μ M against HK2. Preliminary results showed that all test molecules affected HK2 activity to some extent. Twenty-two molecules were found to inhibit the HK2 activity more prominently or in the same range as the known inhibitor 3-bromopyruvate (3BP). Thus, the experimental data supported the predictions of the SBVS procedure, which is able recognize bioactive molecules, ready for structural optimization and/or further testing. Synthetic chemistry was used as tool to synthesize a potent known inhibitor to compare the activity of the newly developed compounds. Additionally, another family of potential HK2 inhibitors found on the SBVS was synthesized. Based on earlier knowledge regarding the sensitivity of several cell lines to HK2 inhibition/ablation, a set of 62 previously selected molecules were firstly evaluated as potential proliferation inhibitors of hepatocellular carcinoma cells. The best eight molecules were then evaluated as proliferation inhibitors of other cell lines and their impact on glycolysis was measured, along with that of two known HK2 inhibitors. Engineered CHO cells overexpressing HK1, HK2 or HK4 were used to determine the isozyme selectivity of four glycolytic inhibitors found in this work. All compounds were able to inhibit glycolysis at 10 μ M in all tested CHO cell lines. However, selective inhibition of HK2-expressing cells was preliminarily observed for one of the compounds at a lower concentration (1.5 μ M). While further studies are required to confirm the potential of this molecule, these results support consideration of HK2 as a target in future cancer drug discovery campaigns.

Palavras-Chave

Metabolismo do cancro; Inibidores da hexocinase 2; Rastreo virtual baseado na estrutura; Síntese orgânica; Avaliação bioquímica e biológica; Via glicolítica.

Keywords

Cancer metabolism; HK2 inhibitors; Structure-based virtual screening (SBVS); Organic Synthesis; Biochemical and biological evaluation; Glycolytic pathway.

"Don't be afraid of hard work. Nothing worthwhile comes easily. Don't let others discourage you or tell you that you can't do it. In my day I was told women didn't go into chemistry. I saw no reason why we couldn't."

Gertrude B. Elion

Nobel Prize in Physiology or Medicine 1988

Pioneer on rational drug discovery

Agradecimentos | Acknowledgements

Em primeiro lugar quero agradecer à minha orientadora Professora Doutora Maria Matilde Marques e co-orientadora Professora Doutora Rita Cardoso Guedes, que tanto admiro, por terem possibilitado o desenvolvimento deste projecto sem nunca limitar as minhas escolhas. Agradeço profundamente a sua orientação, paciência, confiança e disponibilidade durante estes anos. Sempre me permitiram decidir, realizar, errar e por fim aprender. Obrigada por todas as oportunidades de crescer pessoal e cientificamente.

Reconheço também a importância dos meus colegas tanto do iMed.Ulisboa como do CQE, no desenvolvimento das atividades inerentes ao projeto e pelos bons momentos passados. Agradeço particularmente à Doutora Patrícia Serra, à Filipa Ramilo Gomes, ao Pedro Fernandes, à Cátia Marques, ao Pedro Rosado, Doutora Rita Acúrcio, Doutor Pedro Pinheiro e Doutor Gonçalo Justino. Obrigada por toda a ajuda e amizade!

Agradeço ao Doutor Pedro Florindo pelos compostos disponibilizados para os ensaios bioquímicos e biológicos. Quero ainda dirigir os meus agradecimentos à Ana Dias e à Doutora Conceição Oliveira pelas análises de espectrometria de massa.

Agradeço ainda aos meus pais, avó e irmã por me terem ajudado a chegar aqui e pelo apoio incondicional, bem como a todos os restantes familiares e amigos que me apoiaram durante todo o meu percurso académico. Agradeço especialmente ao José Pedro que foi uma presença constante e preciosa nesta caminhada e que a tornou, sem dúvida, muito menos difícil. Obrigada pela paciência, compreensão, todo o apoio e ânimo para continuar.

Finally, I also owe my gratitude to Professor Nissim Hay, Doctor Veronique Nogueira, Soeun Kang, Alexander Terry, and all Hay's Lab for the kindness, technical support, patience, and cooperation during my stay in the University of Illinois at Chicago. A thank you is not enough!

Funding and facilities

I am thankful for the funding provided by the doctoral program MedChemTrain and Fundação para a Ciência e Tecnologia (FCT) through the fellowships PD/BI/128225/2016 and PD/BD/135284/2017, and the Luso-American Development Foundation (FLAD) for the support to my internship at the University of Illinois at Chicago.

The research was supported by FCT through the following projects, funded by national funds through FCT and, when appropriate, co-funded by FEDER under the PT2020 Partnership Agreement: UID/QUI/00100/2013, UID/QUI/00100/2019, UID/QUI/00100/2020, UIDP/00100/2020, LA/P/0056/2020, UID/DTP/04138/2013, SAICTPAC/0019/2015, PTDC/QUI-QAN/32242/2017, PTDC/QEQ-MED/7042/2014, UID/DTP/04138/2019, FCT/CPCA/A2/6972/2020.

This project was physically developed in the facilities of Centro de Química Estrutural (CQE), Institute of Molecular Sciences, Instituto Superior Técnico – Universidade de Lisboa (PT), Research Institute for Medicines (iMed.ULisboa), Faculdade de Farmácia – Universidade de Lisboa (PT) and Department of Biochemistry and Molecular Genetics, College of Medicine, University of Illinois at Chicago (USA).



UIC Department of Biochemistry
and Molecular Genetics

List of publications and communications

Publication in a peer-reviewed journal:

Sara N. Garcia, Rita C. Guedes and M. Matilde Marques, Unlocking the potential of HK2 in cancer metabolism and therapeutics, *Current Medicinal Chemistry*, 2019, 26(41): 7285-7322. DOI: 10.2174/0929867326666181213092652

Oral communication:

Garcia, SN, Marques MM, Guedes, R C. (**February 2019**) Targeting metabolism and apoptosis signaling in cancer cells: a structure-based virtual screening approach toward Hexokinase 2 inhibitors. MuTaLig COST ACTION CA15135 – 3rd WG meeting 2019, Paris, France. - **Travel grant**

Most relevant panel communications:

Garcia, SN, Guedes, R C, Marques MM. (January 2020) Development of hit molecules for hexokinase 2 inhibition: An attempt to target glycolysis and apoptosis in cancer cells. 13th National Organic Chemistry Meeting (13ENQO) and 6th National Medicinal Chemistry Meeting (6ENQT), Aveiro, Portugal.

Garcia, SN, Marques MM, Guedes, R C. (October 2019) Targeting glucose metabolism and mitochondria-induced apoptosis in cancer cells: Discovery of hit molecules for Hexokinase 2 inhibition. 6th Annual meeting of International Society of Cancer Metabolism (ISCaM 2019), Braga, Portugal.

Garcia, SN, Marques MM, Guedes, R C. (September 2019) Targeting glucose metabolism and mitochondria-induced apoptosis in cancer cells: Structure-based virtual screening validation toward Hexokinase 2 inhibitors. 6th EFMC Young Medicinal Chemist Symposium (EFMC-YMCS 2019), Athens, Greece.

Garcia, SN, Marques MM, Guedes, R C. (September 2019) Unlocking the role of HK2 in metabolism and mitochondria-induced apoptosis in cancer cells: A structure-based virtual screening validation VIII EFMC International Symposium on Advances in Synthetic and Medicinal Chemistry (EFMC-ASMC'19), Athens, Greece.

Garcia, SN, Marques MM, Guedes, R C. (July 2019) Inhibiting metabolism and mitochondria-induced apoptosis in cancer cells: A structure-based virtual screening validation toward HK2 inhibitors. XXVI SPQ (Portuguese Society of Chemistry) National Meeting, Porto, Portugal.

Garcia, SN, Marques MM, Guedes, R C. (June 2019), Unlocking the role of HK2 in glucose metabolism and anti-apoptotic signaling in cancer cells: A structure based virtual screening validation. VI SEQT Young Researchers Symposium, Tres Cantos- Madrid, Spain. – **Travel grant**

Garcia, SN, Marques MM, Guedes, R C. (February 2019) Structure-based virtual screening toward Hexokinase 2 inhibitors: targeting metabolism and apoptosis signaling in cancer cells. XXVIth Young Research fellows Meeting (YRFM-SCT 2019), Paris, France.

Garcia, SN, Marques MM, Guedes, R C. (September 2018) Structure-based virtual screening toward Hexokinase 2 inhibitors – Exploiting cancer metabolism for therapy. 5th EFMC Young Medicinal Chemist Symposium (EFMC-YMCS 2018), Ljubljana, Slovenia.

Garcia, SN, Marques MM, Guedes, R C. (September 2018) Targeting glucose metabolism for cancer treatment: a structure-based virtual screening toward Hexokinase 2 inhibitors. XXV EFMC International Symposium on Medicinal Chemistry (EFMC-ISMCS 2018), Ljubljana, Slovenia.

Garcia, S N, Marques MM, Guedes, RC. (July 2018) Hexokinase 2 inhibitors for cancer treatment: an attempt of repurposing. Italian-Spanish-Portuguese Joint Meeting in Medicinal Chemistry - MedChemSicily2018, Palermo, Italy.

Garcia, S N, Marques M M, Guedes, R C. (January 2018) Targeting glucose metabolism for cancer treatment: a virtual screening search for hexokinase II inhibitors. 12th National Organic Chemistry Meeting (12ENQO) and 5th National Medicinal Chemistry Meeting (5ENQT), Coimbra, Portugal.

Garcia, S N, Marques M M, Guedes, R C. (December 2017) Targeting glucose metabolism for cancer treatment: searching hit compounds for hexokinase II inhibition, UL Chemistry PhD Meeting, 2nd Meeting of the CQUL (2ECQUL), Lisboa, Portugal. – **Best poster prize**

Table of Contents

Chapter I: Introduction and Aims	1
Chapter II: State-of-art	7
1. Hallmarks of cancer cells: Reprogramming Energy Metabolism	9
1.1. Impaired glycolysis	9
1.2. Blocking Glycolytic Steps	12
1.3. Blocking multiple pathways	13
1.4. Limited success of current glycolytic inhibitors	13
1.5. HK2 as potential target	14
2. Hexokinases	15
2.1. Roles and distribution of HKs	16
2.1.1. HK1-3	16
2.1.2. HK4/GCK	17
2.1.3. HKDC1	17
2.2. Structures	18
2.3. Expression of HK2	21
2.3.1. HK2 expression in normal cells	21
2.3.2. HK2 upregulation in cancer cells	21
2.3.3. HK2 downregulation	22
2.3.4. HK1 Expression in Cancer Cells	23
2.4. HK anti-apoptotic activity	24
2.4.1. HK-VDAC interaction	24
3. Available HK2 Inhibitors	25
4. Targeting HK2	29
4.1. Proof of concept	30
4.2. Strengths	31
4.3. Limitations	31
4.4. Strategy to target HK2	32
Chapter III: Structure-Based Virtual Screening	35
1. Overview	37
1.1. Insights into the Structure Based Virtual Screening approach	37
1.1.1. Molecular Docking	39
1.2. Prediction of ADMET properties	39
1.3. Biochemical validation of the SBVS	41
1.3.1. ADP detection by HPLC	41
1.3.2. ADP detection by Luminescence - ADP-Glo™ Kinase Assay	41
2. Results and discussion	42
2.1. Molecular docking protocol design and validation	42
2.2. Selection of databases	47
2.3. Structure-based virtual screening (SBVS)	49
2.4. Prediction of ADMET properties and compounds selection	51
2.5. Confirmation of the identity of acquired compounds	54
2.6. SBVS validation – Biochemical assays	58
2.6.1. ATP/ADP detection by HPLC	58
2.6.2. ADP detection by Luminescence - ADP-Glo™ Kinase Assay	59
2.6.3. Evaluation of ruthenium complexes as HK2 inhibitors	61
3. Conclusions	63
4. Experimental Section	64
4.1. Design of Molecular Docking Protocol	64
4.2. Virtual Screening	65
4.3. Prediction of ADMET properties	65
4.4. Characterization of the acquired molecules	66
4.5. Biochemical hexokinase activity assay – ADP-Glo™ kinase assay	66
4.6. Quantitative determination of ATP and ADP by HPLC	67
Chapter IV: Chemical Synthesis	69

1.	Overview	71
2.	Results and discussion	73
2.1.	Synthesis of Ribose-derivatives	73
2.1.1.	Synthesis of the ribose-based core	73
2.1.2.	Synthesis of derivatives of 5	78
2.2.	Synthesis of "cmpd 25" (known inhibitor)	85
2.3.	Biochemical evaluation of the synthesized compounds as HK2 inhibitors	87
3.	Conclusions	87
4.	Experimental Section	88
4.1.	General data for chemicals and instrumentation	88
4.2.	Synthesis of Ribose-derivatives	90
4.2.1.	Synthesis of compound 1 (α + β anomers), (3a <i>R</i> ,4 <i>R</i> ,6 <i>R</i> ,6a <i>R</i>)-6-(hydroxymethyl)-2,2-dimethyltetrahydro-2 <i>H</i> -furo[3,4- <i>d</i>][1,3]dioxol-4-ol	90
4.2.2.	Synthesis of compound 2, ethyl [(3a <i>S</i> ,4 <i>S</i> ,6 <i>R</i> ,6a <i>R</i>)-6-(hydroxymethyl)-2,2-dimethyltetrahydro-2 <i>H</i> -furo[3,4- <i>d</i>][1,3]dioxol-4-yl]acetate	90
4.2.3.	Synthesis of compound 3, ethyl [(3a <i>S</i> ,4 <i>S</i> ,6 <i>R</i> ,6a <i>R</i>)-6-[(methanesulfonyl)oxy]methyl]-2,2-dimethyltetrahydro-2 <i>H</i> -furo[3,4- <i>d</i>][1,3]dioxol-4-yl]acetate	92
4.2.4.	Synthesis of compound 4, ethyl [(3a <i>S</i> ,4 <i>S</i> ,6 <i>R</i> ,6a <i>R</i>)-6-(azidomethyl)-2,2-dimethyltetrahydro-2 <i>H</i> -furo[3,4- <i>d</i>][1,3]dioxol-4-yl]acetate	92
4.2.5.	Synthesis of compound 5 (a + b), ethyl [(2 <i>S</i> ,3 <i>R</i> ,4 <i>S</i> ,5 <i>R</i>)-5-(aminomethyl)-3,4-dihydroxyoxolan-2-yl]acetate + methyl [(2 <i>S</i> ,3 <i>R</i> ,4 <i>S</i> ,5 <i>R</i>)-5-(aminomethyl)-3,4-dihydroxyoxolan-2-yl]acetate	93
4.2.6.	General procedure for the formation of the amide/sulfonamide at C5 (6-11)	94
4.2.7.	General procedure for the hydrolysis of esters (6-11) to the corresponding carboxylic acids (12-17)	98
4.2.8.	General procedure for the synthesis of C2' amides (18-23)	102
4.3.	Synthesis of "cmpd 25" - known inhibitor	107
4.3.1.	Synthesis of 2-Amino-2-deoxy-1,3,4,6-tetrakis-O-(trimethylsilyl)-D-glucopyranose ("cmpd 25a")	108
4.3.2.	Synthesis of 2-[(3-Biphenylcarbonyl)amino]-2-deoxy-D-glucopyranose ("cmpd 25b")	108
4.3.3.	Synthesis of Methyl 2-[(3-biphenylcarbonyl)amino]-2-deoxy-D-glucopyranoside ("cmpd 25c")	109
4.3.4.	Synthesis of Methyl 2-[(3-biphenylcarbonyl)amino]-2-deoxy-6-O-[mesyl]-D-glucopyranoside ("cmpd 25d")	110
4.3.5.	Synthesis of Methyl 6-azido-2-[(3-biphenylcarbonyl)amino]-2,6-dideoxy-D-glucopyranoside ("cmpd 25e")	110
4.3.6.	Synthesis of Methyl 6-amino-2-[(3-biphenylcarbonyl)amino]-2,6-dideoxy-D-glucopyranoside ("cmpd 25f")	111
4.3.7.	Synthesis of Methyl 2-[(3-biphenylcarbonyl)amino]-6-[(2,3-dichlorobenzene-1-sulfonyl)amino]-2,6-dideoxy-D-glucopyranoside ("cmpd 25g")	112
4.3.8.	Synthesis of 2-[(3-Biphenylcarbonyl)amino]-6-[(2,3-dichlorobenzene-1-sulfonyl) amino]-2,6-dideoxy-D-glucopyranose ("cmpd 25")	113
4.4.	Biochemical hexokinase activity assay – ADP-Glo™ kinase assay	113
Chapter V: Biochemical and Biological Evaluation of selected hit molecules		115
1.	Overview	117
1.1.	Effect on cell proliferation and cell death	117
1.2.	Immunoblotting	118
1.3.	Effect on glycolysis	119
1.4.	Biochemical Assay	120
2.	Results and Discussion	121
2.1.	Selection of cell lines	121
2.2.	Cell proliferation and cell death	123
2.2.1.	Cell proliferation in Huh7 cells	123
2.2.2.	Cell death in Huh7 cells	125
2.2.3.	Cell proliferation and cell death in PC3 cells	127
2.2.4.	Cell proliferation and cell death on CHO cells	128
2.2.5.	Effect of PHK2I 21	129
2.3.	Glycolysis stress test	131
2.3.1.	Effect of chronic treatment – GST and CMST	132

2.3.2.	Effect of an acute treatment – GST	135
2.3.3.	Particular effect of PHK2I 10 on GST	137
2.4.	¹³ C isotopic labeling experiments – stable isotope-resolved metabolomics	138
2.5.	Biochemical hexokinase activity assay	142
3.	Conclusions	143
4.	Experimental Section	145
4.1.	Preparation of PHK2I solutions	145
4.2.	Biochemical hexokinase activity assay	145
4.3.	Cell Culture	146
4.4.	CHO cell lines generation	146
4.5.	Effect on cell proliferation	146
4.6.	Effect on cell death	147
4.7.	Immunoblotting	147
4.8.	Metabolomic assays	148
4.8.1.	Measurement of the oxygen consumption rate (OCR) for CMST and the extracellular acidification rate (ECAR) for GST	148
4.8.2.	¹³ C isotopic labeling experiments	149
4.9.	Statistical analysis	150
Chapter VI: Structural and dynamic behavior comparison between HK1 and HK2: a Molecular Dynamics approach to explore potential drug selectivity		151
1.	Overview	153
1.1.	Static structural analysis	153
1.2.	Molecular dynamics	154
2.	Results and Discussion	154
2.1.	Important structure insights from the literature	154
2.2.	Static structural analysis	156
2.3.	MD studies	161
2.3.1.	Design of the experiment	161
2.3.2.	Total energy	162
2.3.3.	Root mean square deviation (RMSD)	164
2.3.4.	Radius of gyration	166
2.3.5.	Clusters of protein conformations	167
2.3.6.	Ligands RMSD	170
2.3.7.	Root mean square fluctuation (RMSF)	170
2.3.8.	Ligands' solvent accessible surface area (SASA)	173
2.3.9.	Distance of ligands to catalytic Asp209/657	174
2.3.10.	Comparing the displacement of helix- α_{13}	177
2.3.11.	Influence of 618,K→T "mutation"	179
2.3.12.	Comparing pockets opening in HK1 and HK2	180
2.3.13.	Notes from the trajectory visualization	183
2.3.14.	Principal component analysis	187
2.3.15.	Protein-ligand interaction fingerprints (PLIFs) over time	191
3.	Conclusions	196
4.	Experimental Section	197
Chapter VII: Final remarks and future work		199
1.	Translation of HK2 inhibition to therapeutics	201
2.	Alternative strategies	201
2.1.	Alternatives to the SBVS approach	202
2.2.	Alternative strategies to obtain more promising test compounds	202
2.3.	Alternatives to the biological evaluation approaches	203
2.4.	Alternatives in the MD study	203
3.	Discussion on the effect of PHK2I 10 and 60 in glycolysis and their physicochemical and pharmacokinetic properties	204
4.	Concluding remarks	206
Chapter VIII: References		207
Annexes		223

List of Figures

Figure 1. Schematic view of the project aims.	5
Figure 2. Summarized scheme of central metabolism focused on glycolysis and the tricarboxylic acid (TCA) cycle..	10
Figure 3. Structures of Glucose (I), FDG (II) and G6P (III).....	11
Figure 4. Structures of glycolytic inhibitors 2DG (IV), LND (V) and 3BP (VI).....	14
Figure 5. A - Common substrates of HKs. B - Simplified mechanism of G6P formation. Besides Asp657 that activates the oxygen for the nucleophilic attack, other residues stabilize the substrates and allow the reaction, such as Ser603 and Arg539 (not shown).....	15
Figure 6. Homology among the catalytic centers of HK from different species. Adapted from [19].	20
Figure 7. HK induced fit.	20
Figure 8. Examples of tissues where cancer cells have shown overexpression of HK2 [5,92–97].	22
Figure 9. Structures of small molecule inhibitors of HK2 expression.	23
Figure 10. Structures of described HK2 inhibitors.	27
Figure 11. Potent and selective HK1 and HK2 inhibitors [43]. The compound designations in quotation marks maintain the numbering used in the original study. The concentrations under each structure represent the corresponding IC ₅₀ values for HK1 (left) and HK2 (right).....	28
Figure 12. Potent and selective HK1 and HK2 inhibitors [43]. Numbers in quotation marks maintain the numbering used in the original study. The concentrations under each structure represent the corresponding IC ₅₀ values for HK2 [149].	28
Figure 13. Structures of the HK2-VDAC interaction inhibitors LV , methyl jasmonate, and LVI , clotrimazole.	29
Figure 14. Examples of diseases/pathways, other than cancer, with HK2 involvement.	29
Figure 15 . Schematic view of the tasks explored in this research project. A corresponds to the path used to find potential HK2 inhibitors and their biological and biochemical evaluation. Path B was used to find differences between HK1 and HK2 to improve selectivity of further inhibitors.	33
Figure 16. Schematic view of the steps taken to conclude the SBVS campaign to find potential HK2 inhibitors. PK: Pharmacokinetics.....	37
Figure 17. A- Simplified drug discovery pipeline. B- Experimental and theoretical approaches for hit discovery after selection of the biological target. In this project a theoretical approach was followed to find small molecule modulators for HK2, using molecular docking as the base technique for VS.	38
Figure 18. Examples of PAINS substructures.	41
Figure 19. Variables tested for the design of the molecular docking protocol.	42
Figure 20. Interactions of HK2 residues (PDBID: 2NZT) with glucose and G6P (depicted at blue, with arrows – H-bonds) in the respective cavities. Nearby residues are also depicted at lighter blue.....	44

Figure 21. Correlations between experimental data (Log[IC ₅₀]) and score (in silico results) of 16 known HK2 inhibitors - PDBID:5HG1, GOLD 5.2, Goldscore scoring function, center: Glu742 (A) and Asn683 (B).....	45
Figure 22. Correlations between experimental data (Log[IC ₅₀]) and score (in silico results) of 16 known HK2 inhibitors – PDBID:5HG1, GOLD 5.2, Goldscore scoring function.	46
Figure 23. Correlations between experimental data (Log[IC ₅₀]) and score (in silico results) of 16 known HK2 inhibitors - PDBID:5HG1, using Autodock4 (left) and Autodock Vina (right).	46
Figure 24. Molecular docking protocol validation.....	47
Figure 25. Results of Drugbank screening: A - from initial 2217 molecules (DBA+DBW), 166 were scored above 80 and 29 were initially selected after visual inspection. B - Main molecules found were cephalosporins and other natural products such as Vitamin B family members. Hydrophilic groups such as phosphates, sugars and sulfate/sulfonamides are the most common.	48
Figure 26. Interactions (2D) between HK2 and “cmpd 30” (A), best inhibitor crystallized (IC ₅₀ = 10 nM) [43], and three of the highest scored molecules from DrugBank VS. (B – cangrelor, score= 115, C - Diosmin, score=95, D - Ceftibuten, score=82). Common residues interacting with the known inhibitor and potential inhibitors are highlighted in yellow. Arrows indicate H-bonds. Dotted lines with benzene an aryl ring in the middle represent H-π interactions.	48
Figure 27. Pharmacophore used for pre-screening of the Ambinter database. Six features were defined. For the pre-screening, molecules must satisfy 4 out of the 6 conditions. Aro – aromatic moiety; Don – H-bond donor; Don&Acc – H-bond donor and acceptor.....	50
Figure 28. A - Criteria involved in the selection of the final set of 111 molecules. B - percentage of the molecules that were rejected (90%), selected by the user (10%) and finally considered suitable for experimental testing (0.37%), from all molecules scored higher than 80 by the molecular docking program. PC: Physicochemical. PK: Pharmacokinetics.....	51
Figure 29. Descriptors and predicted properties of the set of molecules coming from molecular docking (left) and the final set of molecules (right). Predictions and calculations were performed using the FAFdrugs4 platform [204].	53
Figure 30. Results of prediction of Ames test, potential CYPs inhibition and DMSO solubility using the OCHEM server [219], for the final set of 111 molecules.	54
Figure 31. Positive ESI–MS spectra [MW: 100-1000] of four representative samples obtained from the NCI (NSC XXXXX code is the code determined by the NCI, X-NCI is the code given by the author after acquisition). Spectra A and B are examples of spectra in agreement with the expected structures in terms of identity and purity. Spectra C and D are examples of spectra that were not in agreement with the expected structures in terms of identity and purity.	55
Figure 32. Positive (A) and negative (B) ESI–MS spectra [MW: 100-1000] of NSC 31345.....	56
Figure 33. ¹ H NMR (A) and ¹ H- ¹³ C HSQC (B) spectra (300 MHz,75.5 MHz, MeOD) of sample of NSC 31345...	57
Figure 34. A - Structures of ATP (above) and ADP (below). B - Standard curves for ADP “formation” and ATP “conversion”. The graphic shows that the amount of ATP/ADP changes proportionally. Mobile phase: 0.1 M phosphate buffer at 1.2 mL/min, 20 µL injection, isocratic mode for a 15 min run, r.t.. UV detection at 254 nM, C ₁₈ column (ThermoFisher Scientific BDS Hypersil C18 (Ø4.6 x 250 mm, 5 µm) reversed phase column). Error bars correspond to SD.	58

Figure 35. Standard curve for “ATP-to-ADP conversion”.	59
Figure 36. Distribution of rhHK2 inhibition results per ranges of inhibition ability. The number of molecules with higher or comparable effect to 3BP is highlighted. Average % inhibition for each compound resulted from triplicates of at least 2 independent assays.	60
Figure 37. Correlation between experimental data (% HK2 inhibition) and docking score (in silico results) for the 64 molecules tested.	60
Figure 38. Structures of the ruthenium complexes evaluated against rhHK2.	62
Figure 39. Reaction (%) (glucose + ATP → G6P + ADP) catalyzed by rhHK2 that occurred in the absence (control) or presence of either 3BP, ruthenium complexes, or carbohydrate-based ligands at 10 μM (results as average ±SD; * p<0.05).	63
Figure 40. A - The HK2 reaction was prepared with 50 nM HK2, 100 μM ATP and 100 μM Glucose and each of the selected 64 molecules at 10 μM and incubated for 45 min in 384-well white polystyrene microplates. B - After the kinase reaction, the first step was performed by addition of the ADP-Glo™ Reagent that terminated the kinase reaction and depleted any remaining ATP (60-min incubation). Addition of a second reagent converted ADP to ATP and generated light from the newly synthesized ATP using a luciferase/luciferin reaction (60-min incubation). The light generated is proportional to the amount of ADP present and, consequently, kinase or ATPase activity. Adapted [235]. Luminescence was detected with an Anthos Zenyth 3100 Microplate Multimode Detector.	66
Figure 41. A - Core of the ribose family found consistently in the final virtual library of selected molecules and structures of two synthesized molecules. B - Structures of potent known HK1/HK2 inhibitors. The IC ₅₀ for each isozyme is indicated for each molecule as HK1 IC ₅₀ HK2 IC ₅₀ . The main structural features are highlighted: Yellow – sugar moiety; Light green – Amide; Green – Sulfonamide; Red – sulfonamide on the C2 nitrogen of glucosamine.	72
Figure 42. ¹ H NMR spectrum (300 MHz, CDCl ₃) of the C-glycoside 2 . The magnetic non-equivalence of protons 1''a and 1''b and 5a and 5b is clearly displayed, indicating the β-configuration of the molecule, obtained after column chromatography.	75
Figure 43. Most stable pyranose structures: ⁴ C ₁ and ¹ C ₄ chairs.	76
Figure 44. ¹ H NMR spectra (400 MHz, CDCl ₃) of compounds 3 (left) and 4 (right).	77
Figure 45. ¹ H (above) and ¹³ C (below) NMR spectra (400 MHz, 100.6 MHz, MeOD) of compounds 5a (ethyl ester) and 5b (methyl ester). Integration is shown for relevant signals on the ¹ H NMR spectrum.	77
Figure 46. Structures of the final compounds 18-23 .	78
Figure 47. A - Acyl/sulfonyl chloride precursors for the synthesis of compounds 6-11 . B - Primary amines used for the synthesis of compounds 18-23 .	79
Figure 48. ¹ H NMR spectra (400 MHz) of compounds 6 (A) (CDCl ₃) and 12 (B) (MeOD).	80
Figure 49. ¹ H NMR spectra (400 MHz) of esters 5 (black, MeOD), 6 (blue, CDCl ₃) and 12 (orange, CDCl ₃).	81
Figure 50. LC-MS results for the analysis of a sample of compound 18 obtained after semi-preparative HPLC.	84
Figure 51. ¹³ C NMR spectra of the final “ cmpd 25 ” (A , blue) and its methylated precursor “ cmpd 25g ” (B , black).	86

Figure 52. Biochemical evaluation of the synthesized compounds as HK2 inhibitors at 10 μ M. The amount of ADP formed is directly related with the generation of G6P and consequently with the % of reaction occurred in the presence of potential inhibitors (results as average \pm SD, * $p < 0.05$). “Cmpd 25” was able to inhibit the formation of G6P by 97%, contrary to the other molecules that did not significantly reduce the HK2 activity. 87

Figure 53. Screening of 62 PHK2Is yielded eight molecules capable of prominently inhibiting cell proliferation. From those, four were found to inhibit the glycolytic pathway. 117

Figure 54. Detection of untreated PC3 cells using the brightfield of Celigo Adherent Cell Cytometer (Nexcelom Bioscience LLC, Lawrence, MA, USA). Image on the left shows PC3 cells (brighter round white spots) and, on the right, the image shows the ability of the method in identifying (green) those cells. 118

Figure 55. Expected profiles of glycolytic function (**A**) [256] and mitochondrial respiration (**B**) [257]. Arrows indicate the injection of each compound. Changes in pH/[O₂] are detected in real-time. 120

Figure 56. HK2 ablation reduces proliferation by 50% after 7 days treatment in prostate cancer cells LNCaP and PC3 (**A**) and HCC cells (Huh7) (**B**) cells. Effect of HK2 silencing in prostate cancer cells DU145 (**A**) was not significant. Adapted from Nogueira et al (2018) [175] (**A**) and DeWaal et al (2018) [174] (**B**). shScr - cells expressing an inducible control (Scr); shHK2 - HK2 silenced cells; Nt (control) and HK2 (HK2 silenced cells) shRNA indicate Dox-inducible non-targeting and HK2 shRNA, respectively; ### $p < 0.0003$ versus shScr; Error bars correspond to SEM; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ versus shScr; by Student's t-test.. See literature [174,175] for more information. 121

Figure 57. Expression of each HK on PC3, H7 (Huh7), H3B (Hep3B), HG2 (HepG2), and engineered CHO cells monitored by immunoblotting. CHO EV cells possessed a residual amount of GCK. 122

Figure 58. Glycolytic function profile of PC3 (green) and Huh7 (pink) cells based on changes in the extracellular acidification rate (ECAR) in response to injection of 10 mM glucose, 1 μ M oligomycin and 50 mM 2DG. Error bars correspond to SEM. 122

Figure 59. Engineered M15-4 CHO expressing HK2, HK1 or GCK (WT) and EV present different glycolysis rates and glycolytic function profiles. **A** - Glycolytic response of CHO cells to injection of 10 mM glucose. **B** - Glycolytic function profile of CHO cell lines (HK1 –blue, HK2 –orange, GCK – yellow, EV – red), based on changes in the extracellular acidification rate (ECAR) in response to injection of 10 mM glucose, 1 μ M oligomycin and 50 mM 2DG. Error bars correspond to SEM..... 122

Figure 60. A - Effect of “cmpd 25” and “cmpd 34” on Huh7 cells proliferation for 168h (seven days), at 1, 5 or 10 μ M, compared with control cells treated with vehicle (DMSO). **B** - Structures of “cmpd 25” and “cmpd 34” and IC₅₀ reported for the inhibition of HK1 and HK2 with the ADP-Glo™ kit [43]. Error bars correspond to SD. 123

Figure 61. Effect of 62 PHK2Is and “cmpd 25” on Huh7 cells proliferation after 7 days at 10 μ M. Each color corresponds to a measurement at a certain timepoint (0, 24, 48, 72, 120 or 168h). “Cmpd 25” is represented by number 3 (highlighted in red). To compare the effect of each PHK2I with “cmpd 25”, a red dashed line is drawn across all PHK2I bars, marking the maximum relative cell number obtained with cells treated with “cmpd 25”. Error bars correspond to SD. The asterisks represent significant differences compared to the control ($p < 0.05$). Number of cells is normalized according with the cell count on day 0 (first day of treatment with DMSO, “cmpd 25” or PHK2Is)..... 124

Figure 62. Cell death in Huh7 cells after 48h of treatment with PHK2Is in the presence or absence of 5 μ M sorafenib. Cells were treated with IC₅₀ concentrations of PHK2I previously determined for cell proliferation in the same cells. On the left, the total number of cells was measured using Hoechst dye. On the right, the percentage of

dead cells was measured using PI. The number 25 represents “cmpd 25” and 34 represents “cmpd 34”. # - $p < 0.05$ (Student's t) compared to the 0 μM Sorafenib control; * $p < 0.05$ (Student's t) compared to the 5 μM Sorafenib control; Error bars correspond to SD..... 126

Figure 63. Effect of PHK2Is (10 μM) in PC3 cells' proliferation (**A**) and death (**B**). **A** - Effect of PHK2Is in the number of cells after 72h (left) and 144h (right) of treatment. **B** - Results of the cell death experiment, represented by the total number of cells (left) and the number of dead cells (right) compared with control. The number 25 represents “cmpd 25” and 34 represents “cmpd 34”. * $p < 0.05$ (Student's t) compared to the control; Error bars correspond to SD..... 128

Figure 64. Effect of PHK2Is at 10 μM in CHO cell lines' proliferation (**A**) and cell death (**B**)..... 130

Figure 65. PHK2I 21 provoked complete cell death in short periods of time, at low micromolar concentrations. 131

Figure 66. Schematic summary of the protocols applied for GST and CMST experiments..... 131

Figure 67. Glycolytic function and Mitochondrial respiration profiles of Huh7 cells treated with “cmpd 25” and “cmpd 34” (1, 5, and 10 μM) compared with control cells. Error bars correspond to SEM. 133

Figure 68. Glycolysis, assessed as a measure of % change in ECAR compared with control. The number 25 represents “cmpd 25” and 34 represents “cmpd 34”. * $p < 0.05$ (Student's t) compared with the control. Results were obtained from one-time experiment (triplicates or quadruplicates) for PHK2Is 6, 18, 21, and 60. Results with “cmpd 25”, “cmpd34” and PHK2I 40 were obtained from two independent experiments). Error bars correspond to SEM. 133

Figure 69. Glycolytic function and mitochondrial respiration profiles of PHK2Is 6, 18, 21, 40, and 60 at different concentrations. Results were obtained from one-time experiment (triplicates or quadruplicates), except for PHK2I 40 (n=2, triplicates or quadruplicates). Error bars correspond to SEM..... 134

Figure 70. Change in ECAR of Huh7 (**A**), PC3 (**B**), and CHO (**C**) cells after injection of 10 mM glucose, transduced as glycolytic response during GST. Cells underwent prior treatment with known inhibitors and PHK2Is at 10 μM . The number 25 represents “cmpd 25” and 34 represents “cmpd 34”. * $p < 0.05$ (Student's t) regarding control cells (DMSO). Error bars correspond to SEM. 136

Figure 71. Glycolytic response observed during GST with CHO cells treated with compounds 6, 7, 10 and 60 at 1.5 μM . Compound 10 inhibits CHO HK2 and not CHO HK1 cells. * - $p < 0.05$ (Student's t) regarding the corresponding control. Error bars correspond to SEM. Results from one-time experiment..... 137

Figure 72. Effect of PHK2I 10 at 10 and 1.5 μM on glycolytic function, glycolysis and non-glycolytic acidification of CHO CGK (yellow), CHO HK1 (blue) and CHO HK2 cells (orange). C corresponds to control cells. Lighter colors (black outline) in glycolytic function graphs correspond to treated cells. * $p < 0.05$ (Student's t) compared to the corresponding control. Error bars correspond to SEM. Results from one-time experiment..... 138

Figure 73. Schematic summary of protocols applied for ^{13}C isotopic labeling experiments..... 139

Figure 74. ^{13}C isotopic labeling of the glycolytic pathway in Huh7 cells treated with either “cmpd 25” or PHK2Is, 6, 10 and 60. M+0 represents the non-labeled species and M+3 represents labeled species with 3 (lactate, pyruvate) isotopically labeled carbon atoms. 139

Figure 75. ^{13}C isotopic labeling experiments in Huh7 (**A, B, C**) and CHO (**D, E, F**) cells. Labeled glucose was introduced in the culture medium and the incubations lasted for 16h before cell harvesting. Both for PHK2I-treated (5 μM) and control cells, M+0 represents the non-labeled species and M+3 or M+6 represent labeled species with

either 3 (lactate, pyruvate) or 6 (G6P) isotopically labeled carbon atoms (ns – nonsignificant; ** p<0.01; * p<0.05; Student's t).	141
Figure 76. Structures of promising PHK2Is 6, 7, 10 and 60.	145
Figure 77. Studies presented in this chapter, the static structural analysis of HK1 and HK2 and MD studies....	153
Figure 78. Novel regulatory region of NTD activity of HK2, that can be used as target for specific HK2 inhibition (red), by the destabilization of residues Asp447, Ser449 and Lys451. The three-dimensional structure in the figure corresponds to the PDB entry 2NZT [73].	156
Figure 79. Glucose interactions (hydrogen bonds) with residues of HK2 NTD (A) and CTD (B) catalytic pockets.	158
Figure 80. Comparison of HK1 and HK2 CTD pockets, considering the residues in close contact with glucose and the ones nearby (4.5 Å) in the closed (A and C) and open (B and D) state.	160
Figure 81. Structures and IC ₅₀ of HK1 and HK2 inhibitors present in MD simulations.....	162
Figure 82. Plots of system's total energy throughout the simulation.....	163
Figure 83. Plots of RMSD throughout the simulation for systems 2NZT and 2NZT plus crystallographic ligands (A), 4F9O and 4F9O plus crystallographic ligands (B), 1HKC and 1HKC plus inhibitor "cmpd 33" (C), and 5HFU and 5HFU plus inhibitor "cmpd 33" (D).....	165
Figure 84. RMSD (Cα) of each domain (NTD and CTD) for the systems 1HKC plus "cmpd 33" (A) and 5HFU plus "cmpd 33" (B) as a function of time.	166
Figure 85. R _g as a function of time for all systems, with and without ligands..	167
Figure 86. Comparison between the representative structure of the most populated cluster of each system (without ligands): closed versus open state of HK1 (A) and HK2 (C) and respective average RMSD. RMSD by residue is shown in the plots B and D	169
Figure 87. Average RMSD and RMSD by residues between each system with and without crystallographic ligands.	169
Figure 88. Ligands RMSD throughout the simulation. A - Average RMSD (±SD) for all ligands in all systems (1 st Open – 1HKC structure has only the CTD pocket in the open state); B and C - Examples of ligands RMSD over the simulation in HK2 (5HFU, B) and HK1 (1HKC, C).....	170
Figure 89. RMSF of each residue for all the systems in study.....	171
Figure 90. Most relevant RMSF differences among the systems in study.....	172
Figure 91. Structure of HK1 (PDBID: 1HKC) with region Res740-Res826 depicted (cyan), corresponding to the majority of the large subdomain of CTD. Small subdomain is highlighted at yellow.....	173
Figure 92. Main changes in ligands and inhibitors' SASA.	174
Figure 93. Interaction of synthetic co-crystallized inhibitors with HK2 in the catalytic center (CTD).	175
Figure 94. Distance between the center of mass of residue Asp209/657 and the center of mass of inhibitors "cmpd 27" (A/B), "cmpd 30" (C/D) and "cmpd 33" (E/F) in systems 1HKC (green) and 5HFU (pink), throughout the simulation.....	176

Figure 95. Probability Mass Function of the Thr389-Arg468-Ala839angle throughout the simulation for all systems, comparing the absence with the presence of different ligands/inhibitors.	178
Figure 96. Probability Mass Function of the distance between Res389 and Res839 of systems 1HKC and 5HFU in absence or in presence of the inhibitor "cmpd33".....	179
Figure 97. Probability mass function for the distance between Res618-Lys738 for all systems. Each plot compares ligand-free systems with systems displaying crystallographic ligands (A), with inhibitor "cmpd 27" (B), with inhibitor "cmpd 30" (C) and with inhibitor "cmpd 33" (D).....	180
Figure 98. Positioning of residues Thr88 and Thr232 on NTD pocket and Thr536 and Thr680 on CTD pocket and respective distances (Å), measured between the Cα atoms on the crystallographic structure PDBID: 2NZT [73].	181
Figure 99. Probability mass function of the distance between Res88/536 and Thr232/680 in the absence or presence of crystallographic ligands for the four structures.	181
Figure 100. Probability Mass Function of the distance between Res88/536 and Thr232/680 of HK1 (1HKC) and HK2 (5HFU) in the absence and presence of inhibitors "cmpd 27", "cmpd 30", "cmpd 33".....	182
Figure 101. Positioning of residues T536, P605 and Q564 in the small subdomain and T680, L734 and E783 in the large subdomain of CTD pocket, and respective distances (Å, at red), measured between Cα in the CTD pocket of the crystallographic structure PDBID: 2NZT [73]......	184
Figure 102. Plots of distances between key residues in NTD (A, B and C) and CTD (D, E, F) pockets of HK1 (PDBID: 1HKC, green) and HK2 (PDBID:5HFU, pink) in presence of "cmpd 33" throughout the simulation time. Distance was measured between Ser88 and Thr232 (A), Glu116 and Asn335 (B), Pro157 and Leu 286 (C), Thr536 and Thr680 (D), Ile564 and Glu 768 (E) and Pro605-Leu734 (F).....	185
Figure 103. Positioning of the inhibitor "cmpd 33" in the NTD and CTD pockets of HK2 at 0 ns (initial position) (A) and at 200 ns (end of simulation) (B) (system 5HFU + "cmpd 33"). While "cmpd 33" in the NTD pocket aligns the hydrophobic moiety with the helix of the large subdomain, on CTD hydrophobic moiety of "cmpd 33" goes more deeply in the pocket in direction of the small subdomain.	186
Figure 104. Initial structure at 0ns (A) and the distortion caused by "cmpd 33" after 200 ns of simulation in the CTD pocket of HK2 (system 5HFU + "cmpd 33") (B). The distance between Glu564 and Glu783 does not change significantly, but the subdomain is globally affected.....	187
Figure 105. PCA of simulations with HK1 (1HKC) (A) and HK2 systems (5HFU) (B).	187
Figure 106. PCA of simulations with: 5HFU alone (A), 5HFU with "cmpd33" (B), 1HKC alone (C) and 1HKC with "cmpd 33" (D) systems.....	188
Figure 107. Structure of HK1 (1HKC) colored according to the PCA-derived RMSF. Blue means low RMSF and red/pink represents high RMSF. Regions with high RMSF are defined with the starting and ending residue: Gln15-Tyr25; Leu 270-Pro288; Leu310-Thr330; Asn345-Val371.	189
Figure 108. Structure of HK2 (5HFU) colored according to the PCA-derived RMSF. Blue means low RMSF and red/pink represents high RMSF regions. Regions with high RMSF are defined with the starting and ending residue (e.g.: Asp17-Thr35).	190
Figure 109. Crystallographic pose of "cmpd 1" [43] (A) and molecular docking-generated pose of PHK2I 10 (B) and PHK2I 60 (C) into the HK2 catalytic pocket in 2D (left) and 3D (right).	205

Figure 110. Structures of Acid Red 33 and trypan Blue azo dyes.	205
---	-----

List of Tables

Table 1. Characterization and comparative data of human HKs sequences of amino acids.	16
Table 2. HKs tissue distribution and subcellular location.	18
Table 3. Information on the available X-ray structures of HK2 on PDB [78]. The difference between structures is measured through RMSD calculation based on the C α of the residues' backbone.	43
Table 4. Best RMSD and scores for each X-ray structure (PDBID: 2NZT, 5HEX, 5HFU, 5HG1) obtained after molecular docking, using different residues as the center of calculations. A radius = 15Å was used in all described systems, as well as the GoldScore scoring function. Glucose, "cmpd 27", "cmpd 30" and "cmpd 1" are the crystallographic ligands of 2NZT, 5HFU, 5HEX and 5HG1, respectively.	44
Table 5. List of databases used on SBVS together with the number of compounds per database, number of molecules scored above 80 and molecules selected after visual inspection.	49
Table 6. Chemical shifts of H5a and H5b in compounds 5-17	82
Table 7. IC ₅₀ (\pm SD) found for the best inhibitors of Huh7 cells proliferation after 6 days of treatment.	125
Table 8. Comparison of glycolysis inhibition in CHO HK1 and CHO HK2 cells and direct rhHK1/HK2 inhibition with biochemical assays by several PHK2Is (10 μ M). Results are shown as percentage of inhibition.	143
Table 9. Equations used for the calculation of metabolic parameters from OCR and ECAR measurements during CMST and GST, respectively.	149
Table 10. Data on three dimensional structures of wild type hHK1 and hHK2 available on PDB.	157
Table 11. Comparative structure analysis of hHK1 and hHK2 NTD and CTD binding pockets in the open and closed state. "Domain" indicates the location of the pocket; "Residues" indicates the kind of residues evaluated; "N" refers the total number of residues considered for analysis; "RMSD" is used to compare the spatial difference between the two isozymes while identity and similarity compare the difference in the primary sequence. "Different Residues" indicates which residues are in fact different between the two isozymes, where the number corresponds to residue in the sequence, the one letter code corresponds to the HK2 residue and after the arrow the residue present in the HK1 sequence.	159
Table 12. PDBID of HK structures used in the MD studies.	162
Table 13. Average Total Energy (\pm SD) for each system.	164
Table 14. Average RMSD (\pm SD) for the total protein, NTD and CTD of each system.	165
Table 15. Number of conformation clusters for each system and the proportion of structures in cluster 1, 2 and 3 (clusters numbering is inversely proportional to the conformation population of the cluster, i.e., cluster 1 is the most populated. Other clusters are poorly populated.).	168
Table 16. Average Thr389-Arg468-Ala839 angle ($^{\circ}$) \pm SD for each system.	178
Table 17. Observations on systems trajectories regarding the position of the ligand, and the behavior of the NTD and CTD pockets.	183

List of Schemes

Scheme 1. Conversion of luciferin into oxyluciferin catalyzed by luciferase in the presence of ATP and O ₂ .	42
Scheme 2. Types of reactions carried out to obtain the final products, ribose derivatives (A) and “cmpd 25” (B).	73
Scheme 3. Synthetic route for the synthesis of the ribose-based core. (i) 2,2-dimethoxypropane, p-TsOH, acetone, 0 °C to r.t., 1.5 h (quant.); (ii) Ethyl (triphenylphosphoranylidene)acetate, ACN, 2.5 h, 80°C (47%); (iii) MsCl, pyridine, 6h, r.t. (94%); (iv) NaN ₃ , DMF, 75 °C, 1.5h (93%); (v) H ₂ , Pd/C, HCl/EtOH, 18h, r.t., 1 atm (79% - overall yield).	74
Scheme 4. Proposed mechanism for the synthesis of the C-glycoside 2 [240].	75
Scheme 5. Synthetic route for the synthesis of ribose derivatives from compounds 5a and 5b . (vi) Acyl/Sulphonyl Chloride (R-Cl), Et ₃ N, DCM, r.t. to 40 °C, 24h, 20-67%; (vii) 1) LiOH, ACN, 45 min, r.t. 2) HCl (66-99%). (viii) NH ₂ R', DIPEA, HATU, 10:1 DCM/DMF, 18-24h, r.t. to 40°C (3%-39%).	78
Scheme 6. Proposed mechanism for the synthesis of the final products 18-23 from the acid precursors 12-17 and primary amines (4-(Methoxy)benzylamine, 4-Fluorobenzylamine and 2-Methoxyethylamine), catalyzed by HATU in basic conditions.	83
Scheme 7. Synthetic route for the preparation of “cmpd 25”. i) Et ₃ N, TMSCl, pyridine, 115°C, overnight (71%); ii) 3-biphenylcarboxylic acid, DIPEA, HATU, DCM, r.t., overnight (76%); iii) MeOH, HCl, 65°C, overnight (94%); iv) MsCl, pyridine, r.t., 6h (60%); v) NaN ₃ , DMF, 90°C, 2h (94%); vi) Pd/C, EtOH, HCl, r.t., overnight (99%); vii) 2,3-dichlorobenzensulfonyl chloride, pyridine, overnight, r.t. (60%); viii) 2:1 dioxane:water; HCl, 100°C, overnight (9%).	86
Scheme 8. Preparation of compound 1 (anomeric mixture) from ribose: structures and numbering. (i) 2,2-dimethoxypropane, p-TsOH, acetone, r.t., 1.5 h (quant.).	90
Scheme 9. Preparation of compound 2 from compound 1 : structures and numbering . ii) Ethyl (triphenylphosphoranylidene)acetate, ACN, 80 °C, 2.5 h (47%).	90
Scheme 10. Preparation of compound 3 from compound 2 : structures and numbering. iii) MsCl, pyridine, r.t., 6h, (93%).	92
Scheme 11. Preparation of compound 4 : structures and numbering. (iv) NaN ₃ , DMF, 75 °C, 1.5h (93%).	92
Scheme 12. Preparation of compound(s) 5 (shown as the hydrochloride(s)) from compound 4 : structure and numbering. (v) a) H ₂ , Pd/C, HCl/EtOH, 18h, r.t., 1 atm; b) MeOH, r.t. (79%).	93
Scheme 13. General preparation scheme of compounds 6-11 . (vi) Acyl/Sulphonyl Chloride (R-Cl), Et ₃ N, DCM, r.t. – 40 °C, 24h, 20-67%.	94
Scheme 14. Preparation of compound(s) 6 , from compound(s) 5 and 4-(dimethylamino)benzoyl chloride (57%).	94
Scheme 15. Preparation of compound(s) 7 , from compound(s) 5 and 4-(methoxy)benzoyl chloride (51%).	95
Scheme 16. Preparation of compound(s) 8 , from compound(s) 5 and p-toluoyl chloride (67%).	96
Scheme 17. Preparation details of compound 9 , from compound 5 and p-toluenesulfonyl chloride (20%).	96

Scheme 18. Preparation of compound(s) 10 , from compound(s) 5 and 4-chlorobenzenesulfonyl chloride (38%).	97
Scheme 19. Preparation of compounds 11 , from compound(s) 5 and 2,3-dichlorobenzenesulfonyl chloride (51%).	98
Scheme 20. General preparation scheme of compounds 12-17 . vii) 1) LiOH, ACN, 45 min, r.t.; 2) HCl; (66-99%).	98
Scheme 21. Preparation of compound 12 , structures of the precursor(s) (6) and product (99%).	99
Scheme 22. Preparation of compound 13 , structures of the precursor(s) (7) and product (66%).	99
Scheme 23. Preparation of compound 14 , structures of the precursor(s) (8) and product (98%).	100
Scheme 24. Preparation of compound 15 , structures of precursor(s) (9) and product (97%).	101
Scheme 25. Preparation of compound 16 , structures of precursor(s) (10) and product (96%).	101
Scheme 26. Preparation of compound 17 , structures of precursor(s) (11) and product (96%).	102
Scheme 27. General preparation scheme of compounds 18-23 . viii) NH ₂ R', DIPEA, HATU, 10:1 DCM:DMF, 18-24h, r.t.- 40°C (3%-39%).	103
Scheme 28. Preparation of compound 18 from compound 12 and 4-methoxybenzylamine (3%). The obtained mixture contained primarily the N-demethylated compound.	103
Scheme 29. Preparation of compound 19 , from compound 13 and 4-fluorobenzylamine (28%).	104
Scheme 30. Preparation of compound 20 from compound 13 and 4-methoxybenzylamine (35%).	105
Scheme 31. Preparation of compound 21 from compound 14 and 2-methoxyethylamine (21%).	105
Scheme 32. Preparation of compound 22 from compound 15 and 2-methoxyethylamine (10%).	106
Scheme 33. Preparation of compound 23 from compound 17 and 4-methoxybenzylamine (39%).	107
Scheme 34. Preparation of “ cmpd 25a ” from D-2- glucosamine : structures and numbering. i) Et ₃ N, TMSCl, pyridine, 115°C, overnight (71%).	108
Scheme 35. Preparation of “ cmpd 25b ” from “ cmpd 25a ”: structures and numbering. ii) 3-biphenylcarboxylic acid, DIPEA, HATU, DCM, r.t., overnight (76%).	108
Scheme 36. Preparation of “ cmpd 25c ” from “ cmpd 25b ”: structures and numbering. iii) MeOH, HCl, 65°C, overnight (94%).	109
Scheme 37. Preparation of “ cmpd 25d ” from “ cmpd 25c ”: structures and numbering. iv) MsCl, pyridine, r.t., 6h (60%).	110
Scheme 38. Preparation of “ cmpd 25e ” from “ cmpd 25d ”: structures and numbering. v) NaN ₃ , DMF, 90°C, 2h (94%).	110
Scheme 39. Preparation of “ cmpd 25f ” from “ cmpd 25e ”: structures and numbering. vi) Pd/C, EtOH, HCl, r.t., overnight (99%).	111
Scheme 40. Preparation of “ cmpd 25g ” from “ cmpd 25f ”: structures and numbering. vii) 2,3-dichlorobenzensulfonyl chloride, pyridine, overnight, r.t. (60%).	112

Scheme 41. Preparation of “ cmpd 25 ” from “ cmpd 25g ”: structures and numbering. viii) 2:1 dioxane:water; HCl, 100°C, overnight (9%).	113
Scheme 42. Chemiluminescent reaction between luminol and hydrogen peroxide catalyzed by HRP, which is conjugated with the secondary antibody. The reaction generates an excited state product, which decays to a lower energy state releasing photons.	118
Scheme 43. Oxidation of G6P to 6-phosphoglucono- δ -lactone coupled with resazurin reduction to the fluorescent resorufin via NADP ⁺ /NADPH recycling.	120

Abbreviations

Amino acid abbreviation table

Amino acid	3-letter abbreviation	1-letter abbreviation
Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Aspartic acid	Asp	D
Cysteine	Cys	C
Glutamic acid	Glu	E
Glutamine	Gln	Q
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V

General table

¹⁸FDG	2-[¹⁸ F]fluoro-2-deoxy-glucose
2DG	2-Deoxyglucose
2-ME	2-Mercaptoethanol
3BP	3-Bromopyruvate
3PG	3-Phosphoglycerate
ADP	Adenosine diphosphate
ALDO	Aldolase
ATP	Adenosine triphosphate
BSA	Bovine serum albumin
CADD	Computational-aided drug discovery
Calcein-AM	Calcein acetoxymethyl ester
cAMP	Circular adenosine monophosphate
CHO	Chinese hamster ovary
CMST	Cell mito stress test
CRE	cAMP response element
CS	Citrate synthase
CTD	C-terminal domain
DBA	Drugbank Accepted
DBW	Drugbank Withdrawn
DCM	Dichloromethane
DHAP	Dihydroxyacetone phosphate

DIPEA	N,N-diisopropylethylamine
ECAR	Extracellular acidification rate
EDG	Electron donating groups
EDTA	Ethylenediamine tetraacetic acid
EGTA	Ethylene glycol tetraacetic acid
ENO	Enolase
ESI	Electrospray ionization
EtOAc	Ethyl acetate
EtOH	Ethanol
EV	Empty vector
EWG	Electron withdrawing groups
F1,6BP	Fructose-1,6-bisphosphate
F6P	Fructose-6-phosphate
FAS	Fatty acid synthesis
FBDD	Fragment-based drug design
FCCP	Trifluoromethoxy carbonylcyanide phenylhydrazone
FH	Fumarate hydratase
G3P	Glyceraldehyde-3-phosphate
G6P	Glucose-6-phosphate
G6PD	Glucose-6-phosphate dehydrogenase
GADPH	Glyceraldehyde-3-phosphate dehydrogenase
GCK	Glucokinase
GLS	Glutaminase
GLUD	Glutamate dehydrogenase
GLUT	Glucose transporter
GPI	G6P isomerase
GST	Glycolysis stress test
HATU	Hexafluorophosphate azabenzotriazole tetramethyl uronium
HCC	Hepatocellular carcinoma
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HIF-1α	Hypoxia- induced factor-1 α
HK	Hexokinase
HKDC1	HK domain-containing protein-1
HPLC-UV	High-performance liquid chromatography coupled ultra-violet detector
IDH	Isocitrate Dehydrogenase
ITC	Isothermal titration calorimetry
LDH	lactate dehydrogenase
LND	Lonidamide
MCT	Monocarboxylate transporters
MDH	Malate dehydrogenase
MeOH	Methanol
miRNA	Micro RNA
MOM	Mitochondrial outer membrane
MS	Mass spectrometry
MW	Molecular weight
NADP⁺/H	Nicotinamide adenine dinucleotide phosphate (oxidized/reduced)
NFY	Nuclear transcription factor Y
NTD	N-Terminal domain

OAA	Oxaloacetate
OCR	Oxygen consumption rate
OGDH	α -ketoglutarate dehydrogenase
OXPHOS	Oxidative phosphorylation
PAINS	Pan-assay interference compounds
PBS	Phosphate buffered saline
PC	Pyruvate carboxylase
PCA	Principal component analysis
PDBID	Protein data bank identification code
PDGK	Polyphosphate-dependent glucokinase
PDH	Pyruvate dehydrogenase
PEP	Phosphoenolpyruvate
PET	Positron emission tomography
PFK	Phosphofructokinase
PGAM	Phosphoglycerate mutase
PGK	Phosphoglycerate kinase
PHK2I	Potential HK2 inhibitors under study
PI	Propidium iodide
PK	Pyruvate kinase
PLIF	Protein-ligand interaction fingerprints
PMSF	Phenylmethylsulfonyl fluoride
PPP	Pentose phosphate pathway
PVDF	Polyvinylidene fluoride
QTOF	Quadrupole time-of-flight
r.t.	Room temperature
Res	Residue/amino acid
R_g	Radius of gyration
rhHK	Recombinant human hexokinase
RMSD	Root mean square deviation
RMSF	Root mean square fluctuation
RNA	Ribonucleic acid
SAR	Structure-activity relationship
SASA	Solvent accessible surface area
SBDD	Structure-based drug design
SBVS	Structure-based virtual screening
SCS	Succinate synthase
SD	Standard deviation
SDH	Succinate dehydrogenase
SDS-PAGE	Sodium dodecyl sulfate–polyacrylamide gel electrophoresis
SEM	Standard error of the mean
TBST	Tris-buffered saline with 0.1% Tween® 20 Detergent
TCA	Tricarboxylic acid cycle
TLC	Thin layer chromatography
TMS	Tetramethylsilane
TPI	Triosephosphate isomerase
VDAC	Voltage-dependent anion channel
VS	Virtual screening
α-KG	α -Ketoglutarate

Chapter I

Introduction and Aims

Introduction and aims

Cancer is the second most frequent cause of death worldwide. Over 18 million new cancer cases and 9.6 million cancer-related deaths are estimated to have occurred in 2018. By 2040, twenty-seven million new cancer cases per year are anticipated [1].

As a multifaceted disease, cancer is not easy to characterize and target. Cancer cells are heterogeneous and exhibit a deeply complex behavior, involving multiple factors in multiple pathways, and are capable to adapt to various constraints. Indeed, since cancer cell physiology is not fully understood, the possibility of treating the disease by acting upon one target or pathway has long been abandoned and its unlikely success has been clearly recognized [2–4]. Moreover, while many potent drugs have been developed, a considerable number of cancer cell types are known to become therapy-resistant and develop alternative survival routes, as in the case of radiotherapy [5] or chemotherapy with cytotoxic agents [6].

Chemoresistance is considered the key challenge in cancer research. Several molecular mechanisms of how cells avoid the induction of apoptosis commonly instigated by chemotherapeutics have been reported [7,8]. In this regard, rational drug combinations are being planned to achieve the synergy needed to fight these routes [9]. Particularly, blocking specific pathways in altered cancer metabolism seems to attain a potentially safe and efficient approach to overcome drug resistance in combination with chemotherapeutics [10–12]. A common feature of resistant cancer cells is related with the aberrant energy metabolism in which aerobic glycolysis occurs at high rates. Many enzymes in the glycolytic pathway have their expression altered in cancer compared with normal cells, mostly due to the overexpression of several specific isozymes [12,13].

The present research project explores the overexpression of the glycolytic enzyme hexokinase 2 (HK2), envisioning its use as a therapeutic target towards cancer therapy. HK2, often dubbed cancer's double-edged sword, is able to meet the anabolic and catabolic needs of cancer cells but is also involved in anti-apoptotic mechanisms through its binding to mitochondrial outer membrane (MOM), promoting the cancer cells' proliferation and immortalization [14]. HK2 inhibition can lead to cytostatic effects and sensitize cancer cells to several commercial drugs or radiation [5,6,15–17]. Proof-of-concept of the viability of targeting HK2 was previously provided, emphasizing the requirement of HK2 for tumor initiation and maintenance, and the therapeutic effect of HK2 ablation, without systemic constraints in adult mice [18]. However, an obstacle to this strategy is the high amino acid similarity and identity between HK2 and the HK1 isozyme that must be overcome to obtain safer drugs. HK1 is ubiquitously expressed in human cells and its inhibition might lead to harmful events.

Considering the non-availability of specific HK2 inhibitors in the clinic or even in clinical trials, the main aim of this project was the discovery of small molecules able to act as potent and selective HK2 inhibitors, which might allow milder and more effective treatments. Toward this end, a multidisciplinary strategy involving a structure-based virtual screening (SBVS) campaign and its validation was

developed, and it includes *in silico* studies, organic synthesis and a chemical, biochemical and biological characterization of the compounds component. Specifically, the main objectives of this project were:

- Deep understanding of HK2 function and structure (**Chapter 2**);
- Definition of an accurate *in silico* protocol suitable to be used for SBVS on HK2 catalytic pocket (**Chapter 3**);
- Virtual screening of libraries of commercially available molecules and selection of potential hits for HK2 inhibition (**Chapter 3**);
- Acquisition or synthesis of a selection of potential hits (**Chapter 3 and 4**);
- Experimental validation of the SBVS protocol, by biochemical evaluation of the ability of selected molecules to inhibit the catalytic activity of HK2 (**Chapter 3 and 4**);
- Assessment of hit molecules in more biologically complex environments. Evaluation of the effect of hits on cell proliferation, glycolysis, formation of G6P (**Chapter 5**);
- Insight into the structural differences between HK1 and HK2 structures (**Chapter 6**).

This dissertation is organized in 8 chapters according to the subject. Following the present chapter where the theme and objectives are established, and chapter 2, the next four chapters are independently structured. Chapter 3, 4, 5 and 6 are subdivided into an overview, results and discussion, conclusions and experimental procedures. Given the multidisciplinary of the research project, a brief description of the applied methodology is provided in the beginning of each chapter, as part of the general overview. Next paragraphs succinctly describe the content expected in each chapter:

- **Chapter II: State-of-art** – this chapter aims to integrate the objectives of this dissertation in the current state-of-art. Altered glycolysis is summarized and hexokinases (HKs) field is fully explored. This literature review is based on the following review paper [19]: Garcia, S. N.; Guedes, R. C.; Marques, M. M. Unlocking the potential of HK2 in cancer metabolism and therapeutics. *Curr. Med. Chem.* **2019**, 26 (41), 7285-7322. <https://doi.org/10.2174/0929867326666181213092652>
- **Chapter III: Structure Based Virtual Screening (SBVS)** – This chapter describes a SBVS campaign, and discusses its biochemical validation.
- **Chapter IV: Chemical synthesis** – Synthetic routes for the synthesis of a family of potential hits of HK2 inhibition and a known HK2 inhibitor are described. The biochemical evaluation of the synthesized compounds is presented.
- **Chapter V: Biochemical and biological evaluation** – The biological characterization of the selected compounds is presented. Cell proliferation, cell death, interference on glycolysis and mitochondrial respiration were evaluated.

- **Chapter VI: Structural and dynamic behavior comparison between HK1 and HK2: a Molecular Dynamics approach to explore potential drug selectivity**– In this chapter molecular dynamics studies are explored for comparison of HK1 and HK2 structures.
- **Chapter VII: Final remarks and Future work** – Joint conclusions are stated, the relevance of the obtained results, and the future of the project are discussed. Perspectives on cancer therapy mediated by metabolism intervention are also outlined.

Finally, the manuscript ends with the list of bibliography (**Chapter 8 – References**) referenced throughout the manuscript. Moreover, there are some relevant annexes to supplement presented information.

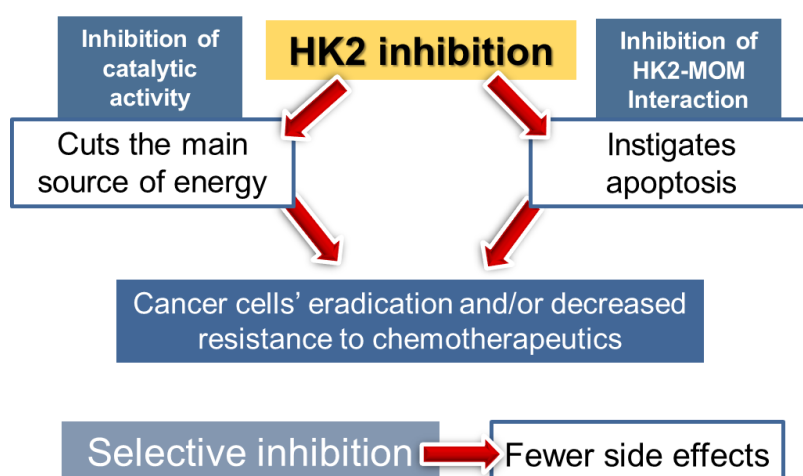


Figure 1. Schematic view of the project aims.

Chapter II¹

State-of-art

¹ The present chapter is based on the following paper: Garcia, S. N.; Guedes, R. C.; Marques, M. M. Unlocking the potential of HK2 in cancer metabolism and therapeutics. *Curr. Med. Chem.* **2019**, 26 (41), 7285-7322. <https://doi.org/10.2174/0929867326666181213092652> [19].

1. Hallmarks of cancer cells: Reprogramming of Energy Metabolism

Cancer hallmarks have been described as “acquired functional capabilities that allow cancer cells to survive, proliferate and disseminate” [20] or, more precisely, as “acquired evolutionary-advantageous characteristics that complementarily promote transformation of phenotypically normal cells into malignant ones, and promote progression of malignant cells while sacrificing/exploiting host tissue” [21]. The establishment of hallmarks provides a more systematic organization of the remarkable features of cancer cells and enables the understanding of their complexity and heterogeneity in order to exploit them in a more rational manner [20]. Currently, eight hallmarks are generally recognized: sustaining proliferative signaling, evading growth suppressors, resisting cell death, enabling replicative immortality, inducing angiogenesis, activating invasion and metastasis, evading immune destruction, and reprogramming energy metabolism. These processes may occur in cancer cells in different extents, allowing cells to be more dependent on one hallmark than another. However, these hallmarks are closely related, as part of a tightly controlled iterative system [20,21].

The recently extended knowledge about the metabolic network of human cancer has consolidated prior information and added important insights, reinforcing the pool of opportunities that this field might offer to achieve the ultimate goal of cancer treatment. The most important aspects of altered cancer metabolism have been well documented and summarized in the last years [13,22–25]. Several specific pathways were noticed to be consistently enhanced, contributing to resistance, namely the glutamine and glucose metabolisms as main energy sources. Moreover, cancer cells require higher levels of catabolite uptake, transfer and utilization when compared with their normal counterparts. **Figure 2** highlights the enhanced steps in central cancer metabolism, contrasting with the regular flux. A strategy involving catabolite deprivation might be a selective and effective anticancer treatment approach. Since glucose (**Figure 3**, I) is considered the main fuel and the most critical nutrient for proliferation and survival of several cancer cell types, interfering with glycolysis could be a key adjuvant in this fight [26,27].

1.1. *Impaired glycolysis*

Otto Warburg’s research work in the first half of the 20th century resulted in the confirmation of a consistently altered metabolism in cancer cells, and almost a hundred years later his theory is still valid in many cases. Warburg postulated that the origin of cancer cells is driven by an irreversible respiratory damage event that leads the cell to find an alternative means of obtaining ATP and essential substrates via fermentation (*i.e.*, glycolysis) rather than through the normal mitochondrial oxidative phosphorylation (OXPHOS). Even in the presence of oxygen, cells are unable to recover from damage that caused them to use glycolysis (“aerobic glycolysis”) as their main pathway to obtain energy and thus cancer cells do not return to a normal status [28]. This effect was later named the “Warburg Effect”, after its discoverer.

Although tumors with normal OXPHOS were quickly discovered, glycolysis is generally still found to be upregulated [13,29].

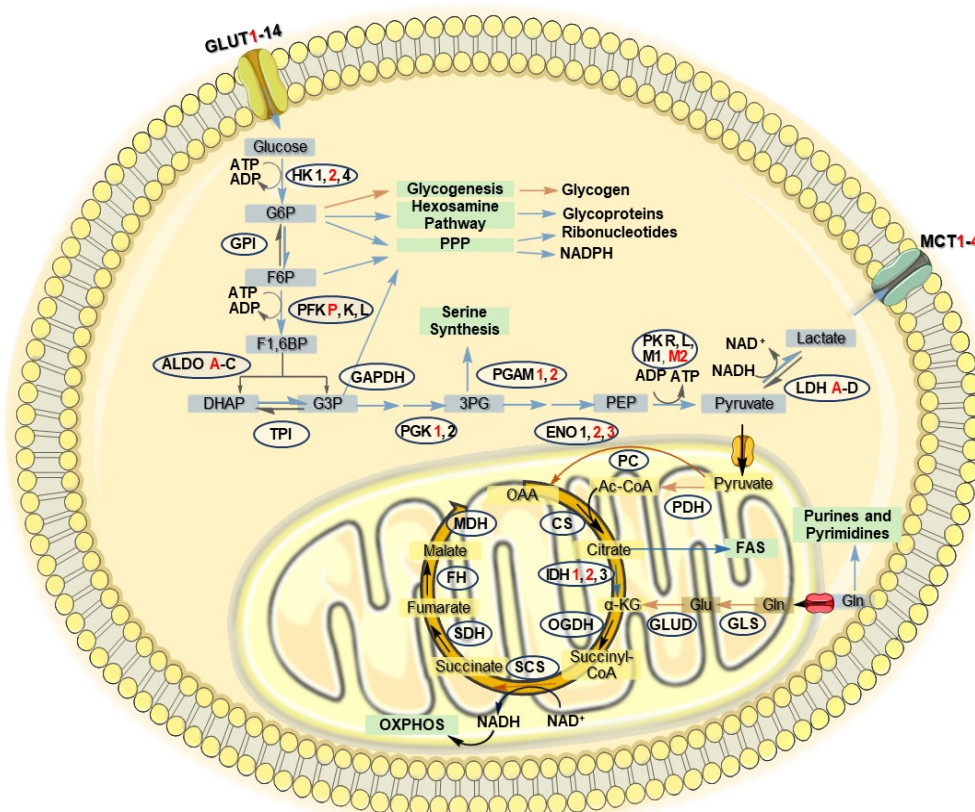


Figure 2. Summarized scheme of central metabolism focused on glycolysis and the tricarboxylic acid (TCA) cycle. Blue arrows represent reactions/pathways significantly upregulated in cancer cells and orange arrows represent the downregulated ones. Highlighted isozymes (red) are consistently upregulated in several cancer cells (adapted from Hu, et al. [11]). α -KG, α -ketoglutarate; 3PG, 3-phosphoglycerate; ADP, adenosine diphosphate; ALDO, aldolase; ATP, adenosine triphosphate; CS, citrate synthase; DHAP, dihydroxyacetone phosphate; ENO, enolase; F1,6BP, fructose-1,6-bisphosphate; F6P, fructose-6-phosphate; FAS, fatty acid synthesis; FH, fumarate hydratase; G3P, glyceraldehyde-3-phosphate; G6P, glucose-6-phosphate; GAPDH, glyceraldehyde-3-phosphate dehydrogenase; Gln, glutamine; GLS, glutaminase; Glu, glutamate; GLUD, glutamate dehydrogenase; GLUT, glucose transporter; GPI, G6P isomerase; HK, hexokinase; IDH, isocitrate dehydrogenase; LDH, lactate dehydrogenase; MCT, monocarboxylate transporter; MDH, malate dehydrogenase; NAD^+/H , nicotinamide adenine dinucleotide (oxidized/reduced); OAA, oxaloacetate; OGDH, α -ketoglutarate dehydrogenase; OXOPHOS, oxidative phosphorylation; PC, pyruvate carboxylase; PDH, pyruvate dehydrogenase; PEP, phosphoenolpyruvate; PFK, phosphofructokinase; PGK, phosphoglycerate kinase; PGAM, phosphoglycerate mutase; PK, pyruvate kinase; PPP, pentose phosphate pathway; SCS, succinate synthase; SDH, succinate dehydrogenase; TPI, triosephosphate isomerase. Adapted from Garcia et al (2019) [19].

Regardless of these early findings, interfering with glycolysis has only started to be seen as a relevant tool for cancer therapy in recent years, in addition to other strategies directly affecting signal transduction pathways and genetic and epigenetic perturbations [13]. Nevertheless, due to the recognition of a consistently enhanced glycolysis pathway, glucose uptake and glycolytic activity of cancer cells have been exploited for tumor detection. Regarding diagnosis, 2-¹⁸F-fluoro-2-deoxy-glucose (¹⁸FDG) remains the most commonly used radiotracer for positron emission tomography (PET). ¹⁸FDG (**Figure 3, II**) is transported into the cell by glucose transporters (GLUT1-14) and is captured by HKs to be metabolized. Increased glycolytic activity is then detected by the accumulation of the positron-labeled glucose analogue [30].

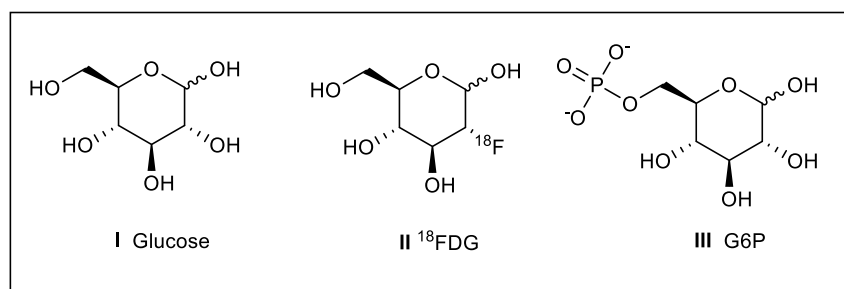


Figure 3. Structures of Glucose (I), FDG (II) and G6P (III)

In order to ensure high proliferation rates, cancer cells must have access to multiple nutrients that allow for a rapid cell growth and further cell division. This is achieved by an increase in anabolic pathways and low dependence on high levels of ATP for cell survival and proliferation. Furthermore, the occurrence of enhanced glycolysis can be explained by kinetic considerations. Indeed, OXPHOS is more efficient, converting one molecule of glucose into 36 ATP molecules while the glycolysis outcome is only 6 ATP molecules, but the latter is 10-100 times faster [31]. Glycolysis is a tightly regulated process in which HKs, phosphofructokinase (PFK) and pyruvate kinase (PK) play a vital role as subjects of allosteric regulation and maintain the necessary homeostasis and adequate levels of metabolic intermediates. HKs are the most influential operative enzymes in setting the flux through glycolysis [32].

Similarly, in normal and cancer cells, the first step is glucose uptake, carried out by GLUT, specifically GLUT1, which is usually overexpressed in cancer cells. By capturing and catalyzing glucose phosphorylation into glucose-6-phosphate (G6P, **Figure 3, III**), HKs are involved in the first and most controlling step of the glycolysis process. With the consumption of ATP, this reaction is highly exergonic and thus considered irreversible in the intracellular medium [32]. Following the glycolytic chain, G6P is reversibly isomerized to fructose-6-phosphate (F6P), which is then irreversibly phosphorylated at C1 by phosphofructokinase (PFK), with ATP consumption. After biotransformation of fructose-1,6-bisphosphate (F1,6BP) into phosphoenolpyruvate (PEP) by a series of reversible reactions, pyruvate kinases (PK – PKM1, PKM2, PKR, and PKL) catalyze its irreversible conversion to pyruvate with concomitant ATP generation. The irreversible reactions catalyzed by HKs, PFK and PKs are then considered controlling checkpoints for glycolysis. After the pyruvate production step, glycolysis is mostly followed by OXPHOS in normal cells, with pyruvate entering the tricarboxylic acid (TCA) cycle, or less likely proceeds to fermentation by catalytic conversion of pyruvate into lactate, mediated by lactate dehydrogenase (LDH). Lactate is rapidly expelled to the extracellular space by monocarboxylate transporters (MCT1-4) to maintain an adequate intracellular pH [32].

Generally, cancer cells behave differently, and the reaction flux is altered, wherein specific isozymes have been demonstrated to be upregulated (highlighted in red in **Figure 2**). Extensive comparative studies between multiple types of primary tumors and healthy tissues enabled the conclusion that all reactions involved in glycolysis are upregulated in cancer cells, as are many reactions involved in the anabolic pathways leading to nucleotides, amino acids and lipids. Among the enzyme families involved

in central metabolism, HK, phosphoglucosmutase (PGM), enolase (ENO), and LDH are the most frequently upregulated in multiple types of cancer cells; typically, at least one isozyme from these families is overexpressed. On the other hand, many catalytic processes required for OXPHOS appear to be downregulated (highlighted with orange arrows in **Figure 2**) [33].

Although blocking an upregulated pathway in cancer may appear a straightforward strategy, the marked heterogeneity among cancer cells precludes establishing a universal pattern. Accordingly, successful therapeutic interventions upon specific metabolic pathways suggest more personalized approaches. A few examples of attempted strategies along the glycolytic pathway and some of their drawbacks follow in the next subsections.

1.2. *Blocking Glycolytic Steps*

GLUT-1 is overexpressed in many cancer cells and is envisioned as a therapeutic target. However its broad expression in healthy tissues precludes its further exploitation [13]. By contrast, certain enzyme families associated with the glycolytic process, such as HK and G6P isomerase (GPI), which are also involved in transcriptional and apoptosis regulation, appear more relevant [34]. Besides its catalytic role, GPI (also known as autocrine motility factor) acts as promoter of proliferation and cancer cells' motility, enabling tumor invasion by inducing the matrix metalloproteinase-3. In addition to its own product, F6P, other GPI inhibitors have been developed, namely 6-phosphogluconic acid [34] (**annex 1 – S-I**). PFK is also a potential target. This enzyme is naturally inhibited by several endogenous molecules, such as PEP, lactate, palmitoyl CoA, ATP or *N*-acetylglucosamine (GlcNAc) (**annex 1 – S-II to S-VI**), shunting F6P to the pentose phosphate (PPP) pathway with NADPH regeneration. However, enhancing PPP still provides a benefit to cancer cells. In addition, the accumulation of F6P will lead to its reversible conversion into G6P, which may also constrain glycolysis by the inhibition of HKs [13]. Aldolase, specifically the isoform A (ALDO A), is also overexpressed in cancer cells. Small molecule ALDO A inhibitors (e.g., 3-deoxyglucose and 3-fluoro-D-glucose (**annex 1 – S-VII, S-VIII**)) have been investigated [13,34]. Likewise, glyceraldehyde-3-phosphate dehydrogenase (GAPDH) inhibitors (e.g., iodoacetate - **annex 1 – S-IX**) have been developed [34]. GAPDH, which is enhanced at low NAD⁺/NADH ratios, contributes to the glycolytic cycle but its inhibition favors the PPP by accumulating upstream products [35]. Phosphoglycerate kinase (PGK) has an interesting potential. Besides its participation in glycolysis, this enzyme is able to reduce disulfide bridges and seems to initiate proteolysis of angiostatin, ultimately increasing angiogenesis [32]. Phosphoglycerate mutase (PGAM) and enolase (ENO) are also being investigated due to their consistent overexpression in cancer cells [34]; nevertheless, no other roles with potential influence in cancer development appear to have been found for these enzymes. Among the PK isozymes, PKM2, which is the most commonly expressed in cancer cells, has been considered a promising target, but its usefulness is controversial in view of inconsistent reports of potential benefits with either activation or inhibition of the enzyme [13,23]. Lactate dehydrogenase (LDH) is generally overexpressed in cancer cells and thus also considered a potential

target. LDH inhibitors are being developed in order to inhibit glycolysis and enhance the use of pyruvate in the TCA cycle. Nevertheless, LDH acts too downstream of the glycolytic chain; due to the accumulation of metabolic substrates for anabolic pathways, its inhibition would probably be insufficient to produce a significant therapeutic effect. Similarly, inhibition of the monocarboxylate transporter (MCT), with ensuing accumulation of lactate and its increased conversion into pyruvate, has been considered to prevent glycolysis; moreover, the concomitant acidification of the intracellular medium can prompt apoptosis [13].

In conclusion, blocking various glycolytic steps in cancer cells is an appealing strategy, but its achievement carries several limitations.

1.3. *Blocking multiple pathways*

In oxidative tumors, which rely on OXPHOS to obtain raw nutrients, the use of anti-glycolytic agents seems, to some extent, fruitless. These cancer cells can obtain substrates from other pathways, including catabolism of lipids, ketone bodies, and amino acids, namely through the oxidation of glutamine, upon which some cancer cells are highly dependent [22]. Another difficulty is the metabolic adaptation of cancer cells, which can switch from one pathway to another at different stages of the disease. Dual inhibition of glycolysis and glutaminolysis has emerged as a potential therapeutic strategy in these cases, with positive *in vitro* results in controlling ovarian cancer cells proliferation [27]. Similarly, combined inhibition of glycolysis, glutaminolysis and *de novo* fatty acid synthesis showed promising results in thirteen diverse cell lines [36]. Indeed, given the possibility of metabolic reprogramming upon glycolysis inhibition, a multitarget approach should be a more efficient therapeutic strategy than simply targeting the glycolytic pathway. In this regard, considering the subcellular location of the processes involved in the central metabolism (**Figure 2**), inhibiting both glycolysis and mitochondria activity might represent the most effective way to fight cancer cells [34].

1.4. *Limited success of current glycolytic inhibitors*

Although, in theory, the glycolysis inhibition strategy appears very attractive, targeting this pathway had met limited success and current glycolysis inhibitors, have been repeatedly put aside [23]. In fact, the glycolytic inhibitors discovered and tested so far are not very effective and carry significant side effects [22,23]. 2-Deoxyglucose (2DG, **IV**, **Figure 4**) is one of the most well-known compounds able to inhibit the glycolytic pathway, specifically for its effect on HK upon conversion into 2DG-6-phosphate, causing inhibition of cell proliferation, and leading to cell death. Nevertheless, this non-selective compound entails low tolerability and excessive side-effects due to the high concentrations required to produce an effect; its benefit has been controversial in clinical trials, some of which have been discontinued [37,38]. Likewise, clinical trials using lonidamine (LND, **V**) were stopped due to lack of positive results [39]; still, LND is approved in some European countries. Moreover, albeit 3-bromopyruvate (3BP, **VI**) seems to

be a promising compound, which is expected to enter clinical trials soon [40], its non-selective alkylating properties [41] raise concerns.

Attempts to selectively inhibit specific isoenzymes involved in cancer glycolysis have gradually come to light, namely those targeting PKM2 [42] and HK2 [43]. In the near future, overcoming the high structural similarity among isoforms is expected to become a major goal (and a reality), ultimately giving rise to an increased availability of selective glycolytic inhibitors. A new generation of glycolytic inhibitors with improved potency and selectivity is undoubtedly needed.

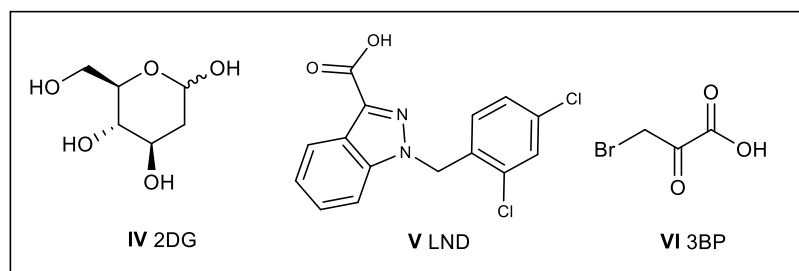


Figure 4. Structures of glycolytic inhibitors 2DG (IV), LND (V) and 3BP (VI)

1.5. HK2 as a potential target

The importance of HKs in glycolysis as the enzymes that catalyze the first (irreversible) step of glycolysis was mentioned above. The HK product, G6P, is not only necessary to proceed to OXPHOS but also to the anabolic processes of PPP; thus, G6P is crucial for the synthesis of nucleotides, glycogenesis, and the hexosamine synthesis pathway. Due to the regeneration of NADPH in PPP, HKs are also ultimately involved in fatty acid synthesis and redox control [32]. Among the four most well-studied human HKs (HK1-4) [44], the HK2 isoform was found to be particularly and consistently overexpressed in cancer cells [33]. Interestingly, HK1 and HK2 have been demonstrated to inhibit apoptosis by a non-catalytic process [45]. Indeed, seeking newly synthesized ATP and an enhancement of the glycolysis rate, these isoforms bind to MOM, most likely to a voltage-dependent anion channel (VDAC) [46]. VDAC-1 is the major protein of MOM in eukaryotes, mediating the permeation of metabolites into/out of the intermembrane space, including the newly synthesised ATP [47], and regulates the intrinsic pathway for apoptosis. Briefly, HK/VDAC interaction avoids the binding of proapoptotic factors, preventing the continuation of the programmed cell death and contributing for the immortalization of cancer cells [16,45,48–50].

Taken together, these observations suggest that HK2 is an excellent potential target for cancer treatment, since a dual role in reprogramming the central metabolism and contributing to avoid senescence of cancer cells can be envisaged. Potent and selective inhibitors for HK2 over HK1 should largely affect cancer cells without compromising systemic homeostasis and normal metabolic functions related to HK1 [43].

2. Hexokinases

HKs are part of the group of anomeric sugar kinases that perform essential roles in carbohydrate metabolism, particularly in glycolysis. Kinases are a diverse class of enzymes (over six hundred human kinases have been identified) able to catalyze the transfer of a phosphate group from a phosphate of a phosphate donor (e.g., ATP), to an acceptor (e.g., sugar) [51]. Specifically, HKs catalyze the phosphorylation of hexoses such as glucose (**I**), mannose (**IV**) and fructose (**V**), generally at the oxygen bound to C6 (**Figure 5 - A**), using ATP as the phosphate donor [44] (**Figure 5 - B**). HK isozymes are widely distributed in nature and are responsible for the first step of glucose usage in all eukaryotic cells and in many prokaryotes [52].

There are five monomeric HK isozymes currently known in the human genome: HK 1-4, named according to their electrophoretic mobility [44,52], and HK domain-containing protein-1 (HKDC1), the one most recently identified [53]. Specifically, HK1-3 and HKDC1 are 100 kDa proteins, while HK4 (or glucokinase – GCK) has approximately half the size [44]. Concerning the evolutionary process, HK2 is suggested as the isozyme most closely related to ancestral HK, with the others resulting from different types of mutations, mostly in the *N*-terminal half, or the complete absence of this domain (in GCK). The retention of catalytic activity in both halves of HK2, rather than the *C*-terminal binding pocket only, such as in HK1 and HK3, is also consistent with this view [54]. Data on the amino acid sequences of human HKs [55], including a comparison of identity and similarity among the different isozymes, are presented in **Table 1**.

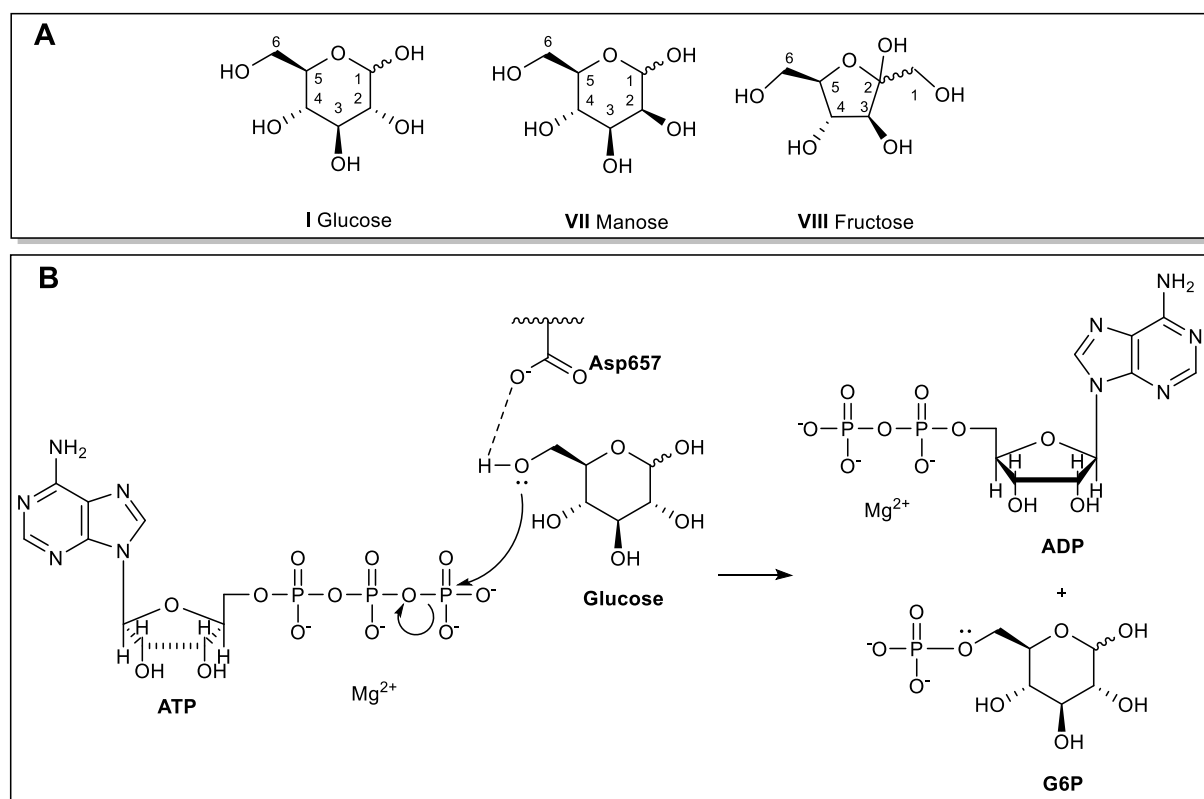


Figure 5. A - Common substrates of HKs. B - Simplified mechanism of G6P formation. Besides Asp657 that activates the oxygen for the nucleophilic attack, other residues stabilize the substrates and allow the reaction, such as Ser603 and Arg539 (not shown).

Table 1. Characterization and comparative data of human HKs sequences of amino acids.

Isozyme	UniProtKB ^a	Weight	Catalytic	Length	Identity/Similarity ^b (%)				
	entry	(kDa)	pockets	(res)	HK1	HK2	HK3	HK4 ^c	HKDC1
HK1	P19367	100	1	917	-	73/87	52/69	53/72	71/85
HK2	P52789	100	2	917	73/87	-	55/72	55/73	68/85
HK3	P52790	100	1	923	52/69	55/72	-	52/72	53/70
HK4	P35557	50	1	465	27/37	28/37	26/36	-	26/36
HKDC1	Q2TB90	100	1	917	71/85	68/85	52/69	52/71	-

^aUniProtKB entries [55] ^b Calculated with MOE2016 software. ^c Obtained by the alignment of HK* C-terminal and HK4; res – number of residues

2.1. Roles and distribution of HKs

HKs in mammals seem to be similarly distributed among tissues, although there are clear differences in the expression of specific isozymes [56]. **Table 2** summarizes information on the most common tissue distribution of the best known HKs and their subcellular location. The next subsections detail more information about each member of the human HK family.

2.1.1. HK1-3

HK1 is commonly expressed in most cells, with particular predominance in the brain [57] and kidneys [56]. It plays a catabolic role, by introducing glucose into the glycolytic metabolism in order to produce ATP (energy) [57]. Likewise, HK2 is broadly expressed, although it is typically present at much lower levels in most tissues; nonetheless, HK2 is the predominant isozyme in insulin-sensitive tissues (*e.g.*, skeletal and cardiac muscles; adipose tissue).

HK1 and HK2 seem to exhibit similar subcellular location, specifically in cytoplasmic fractions and MOM [57]. Interestingly, HK2 was also found to be associated with the nucleus in the presence of inhibitors such as 2DG or clotrimazole [58,59]. These isozymes are recognized as mainly responsible for the general HK activity, although HK2 (as well as HK3) is more likely associated with the anabolic processes of glycogen synthesis and PPP [57,60]. In common, HK1-3 are inhibited by their own product, G6P, and by high concentrations of inorganic phosphate. Remarkably, lower inorganic phosphate levels inhibit HK2 and HK3 activity, but do not affect HK1 [57].

Compared to HK1 and HK2, HK3 exhibits the highest affinity for glucose and the lowest affinity for ATP [61]. Contrasting with HK1 and HK2, HK3 is inhibited by high concentrations of glucose [52,61], which seems to be related with substrate binding to the C-terminal pocket [61]. HK3 lacks the initial hydrophobic moiety present in the N-terminal domain of HK1 and HK2, which leads to a considerable difference in the three-dimensional structure, different subcellular locations and, subsequently, different

roles. For instance, unlike HK1 and HK2, HK3 is unable to bind to mitochondria and is found in the periphery of the cell nucleus [57]. HK3 is a weakly expressed isozyme in mammalian cells, detected primarily in the liver, kidney [56], spleen and lungs [60]; as a result, its role in hexose metabolism does not appear to have been fully characterized [44,60].

2.1.2. HK4/GCK

GCK has been shown to be tissue-specific, restricted to liver and pancreas, where it is expressed in high levels [56,60] and plays critical roles, namely in glycogenesis in the liver, and in pancreatic β -cells, through the control of insulin secretion [62]. The pivotal role of this enzyme in glucose homeostasis has been recognized over the years and different types of diabetes have been related to mutated GCK [44,62–65]. Indeed, GCK activation has been regarded as a valid therapeutic strategy against diabetes and hyperglycemia [66–68]. Interestingly, GCK exhibits lower affinity for the glucose substrate than HK1-3 [52]; however, it does not appear to be inhibited by G6P to the extent observed for HK1-3 under physiological conditions [44,52,62]. Another feature of GCK is its translocation to the nucleus in hepatocytes when complexed with glucokinase regulatory-protein, a process that inhibits its catalytic activity and confers protection from proteolysis in the cytoplasm [69].

Recently, a novel HK4-like mammalian isozyme has been reported. This isozyme is a 55 kDa protein, designated Polyphosphate-Dependent Glucokinase (PDGK), which appears to be structurally related with HK4. PDGK uses inorganic polyphosphates as predominant substrates and was significantly found in liver, heart, muscle and lung of ovine, porcine, bovine and rat species [70]. In the house mouse, the gene was found to be expressed and, contrary to other HKs, ADP is used as substrate instead of ATP; this could be an advantage when energy resources (ATP levels) are lower (e.g., low mitochondrial OXPHOS during aerobic glycolysis, enhanced in cancer states) [71]. Nonetheless, additional studies on the activity and expression of this isozyme are required.

2.1.3. HKDC1

HKDC1 was recently described in the human and other species' genomes [53]. In humans, this isozyme is also widely expressed, namely in thymus, colon, esophagus, pharynx, and eye tissues [13], small and large intestine, pancreas, and lung; however, the retention of HK function and enzymatic activity have not been totally characterized [13]. Recently, enzymatic HKDC1 activity was detected *in vitro*, suggesting a contribution of HKDC1 to the total HK activity in different cell types, albeit reduced when compared with HK1 [67]. High levels of HKDC1 expression in liver have been associated with pregnancy and gestational hyperglycemia [67–70]; it has also been found in hepatocarcinoma [71] and lung cancer tissues [72]. Nevertheless, further structural and biochemical characterization of HKDC1 is still required to clarify its role(s).

Table 2. HKs tissue distribution and subcellular location.

Human isozyme ^a	Tissue Distribution	Subcellular location
HK1	Broadly expressed (high levels)	Cytoplasm MOM
HK2	Broadly expressed (low levels) Adipose, Skeletal and Cardiac Muscle (Higher levels)	Cytoplasm MOM Nucleus
HK3	Little expressed (Kidney, Liver, Spleen, Lungs)	Cytoplasm Nucleus periphery
HK4 (GCK)	Liver, Pancreas (β -cells)	Cytoplasm Nucleus

^a Due to the current limited knowledge about HKDC1 this enzyme is not included in this analysis.

2.2. Structures

Mammalian HKs have long been considered to be substantially different from those of other organisms [52]; however, while sequences share in fact low identity, several examples show that the three-dimensional (3D) structures might be close. For instance, human HK1 (hHK1, PDBID:4F9O [72], 2.65 Å, 2013), hHK2 (PDBID:2NZT [73], 2.45 Å, 2006), rat HK1 (rHK1, PDBID:1BG3 [74], 2.80 Å, 1998), *Schistosoma mansoni* HK (sHK, PDBID:1BDG [74], 2.60 Å, 1998), yeast HKPI (yHKPI, PDBID:3B8A [75], 2.95 Å, 2008) and *Arabidopsis thaliana* HK1 (AtHK1, PDBID:4QS7 [76], R=2.00 Å, 2015) are quite similar (average RMSD= 1.70 Å). When the overlap is directed towards the catalytic pocket, the similarity is even higher (RMSD= 1.20 Å). Analyzing the identity and similarity were analyzed using 29 residues belonging to the glucose cavity, the agreement was 90-100% and 93-100%, respectively (**Figure 6**). Accordingly, HKs present highly (structurally) conserved catalytic pockets between different species, even in HKs sharing low overall identity with hHK1 (rHK1 - 94%, sHK - 23%, yHKP1 - 17%, AtHK1 - 18%).

The first HK 3D structure, yeast (*Saccharomyces cerevisiae*) HKB, was resolved by X-ray crystallography (PDBID:2YHX) in 1978 [77]. Currently, seventy-five HK or HK-related structures are included in the Protein Data Bank (PDB, <https://www.rcsb.org/>) [78], namely forty-six from human origin. From those, thirty-one are GCK, ten represent HK1, four HK2, one is related to the HK3 structure, representing the C-terminal domain [61]. None is associated to HKDC1.

Sharing high identity and similarity, HK1, 2 and 3 are composed by two structurally identical domains, N-terminal (NTD) and C-terminal (CTD), each with a glucose/G6P binding pocket. On the other hand, GCK is constituted by a unique domain. Interestingly, only HK2 has been shown to have two catalytically active pockets; by contrast, in HK1 and HK3 only the pocket located in the C-terminus has catalytic

activity [54,79,80]. Moreover, the *N*-terminal catalytic pocket has been shown to provide the major enzyme activity of HK2 [54].

Structurally, the *N*- and *C*- domains are linked by a long eight-turn α -helix (helix- α^{13}), each formed by a large and a small subdomain. The active site is located in the cleft between the two subdomains [73]. Although several studies [53,81] have contributed with important insights, more investigation on the catalytic/noncatalytic NTD of HK isozymes should be carried out.

According to alignment information from the UniProt database (www.uniprot.org) [55], human HK1 and HK2 have an identity of 73% and a similarity of 87% (the highest among HKs), both with a sequence length of 917 residues [82] (**Table 1**). As aforementioned, only HK1 and HK2 possess a hydrophobic *N*-terminal helix (1-12 aa) able to interact with MOM [57].

Regarding HK4, no G6P was found in any of the reported structures, which is consistent with the lack of inhibition by its own product (*vide supra*). This isozyme has been crystallized mostly with activators and with its substrate (glucose) and/or non-hydrolyzable ATP analogues, such as adenosine-5'-(β - γ -imino)-triphosphate (AMP-PNP) and adenosine-5'-(γ -thiotriphosphate (ATP γ S). When glucose is present, both the α - and β -isomers were found, suggesting weak anomeric selectivity [78].

The induced fit caused by glucose binding to HKs, which involves rigid body rotation and non-rigid body conformational changes, was established long ago [74,75,83]. Basically, a small (non-rigid) subdomain rotates $\sim 17^\circ$ towards the direction of a large (rigid) subdomain, generating a closed state of the protein. The different conformations of several described loops [75,84], as well as changes in the subdomain positions, have become easier to identify by comparison of the most recent structures of HK2 crystallized with strong inhibitors (open state) [43] and the previously obtained structure of HK2 crystallized with glucose (PDBID: 2NZT [73]), in the closed state (**Figure 7**).

As for the ATP pocket, the molecule is thought to overlap with G6P *via* the common phosphate group [75]. ATP has not yet been crystallized with any HK, presumably because it is easily hydrolyzable and/or competes with G6P for the pocket [74,83]. Nonetheless, AMP-PNP and ATP γ S have been crystallized with HK4 (PDBID:3ID8, 3FGU, 3VEY) and found localized close to glucose, overlapping the G6P site, which is consistent with the proposed hypothesis [85,86]. Indeed, this possible competition represents a plausible basis for the G6P inhibition mechanism. In another reported structure, PDBID:1QHA [87], human HK1 was crystallized with AMP-PNP; however, the small molecule seems to bind superficially, quite far from the glucose binding site, suggesting that phosphate transfer to glucose from this site would be hampered. It has been proposed that this non-catalytic pocket, thought to accommodate ATP, may be important for the interaction of HK1 and HK2 with MOM [87]; yet, other phosphorylations, including autophosphorylation, are not excluded [88].

Many functions and activities have been attributed to HKs. Nonetheless, despite important insights and new crystallographic structures, HKs specific mechanisms of action are still not fully understood, including the mechanism of inhibition by G6P, the reason for the existence of non-catalytic pockets whose structure highly resembles the active catalytic pockets, and the binding mode(s) to MOM. Thus, much remains unknown and further research is warranted.

	532	601	620	657	677	679	708	739	742	861	897
hHK1	D L G G T	T F S F P	T K	N D T	I	G T G S N	E	Q	E	D G T	G S
hHK2	D L G G T	T F S F P	T K	N D T	I	G T G S N	E	Q	E	D G T	G S
rHK1	D L G G T	T F S F P	T K	N D T	I	G T G T N	E	Q	E	D G T	G S
sHK	D L G G T	T F S F P	T K	N D T	I	G T G T N	E	Q	E	D G S	G S
yHKP1	D L G G T	T F S Y P	T K	N D T	I	G T G V N	E	Q	E	D G S	G S
AtHK1	D L G G T	T F S F P	T K	N D T	I	G T G T N	E	Q	E	D G G	G S

Figure 6. Homology among the catalytic centers of HK from different species. Adapted from [19].

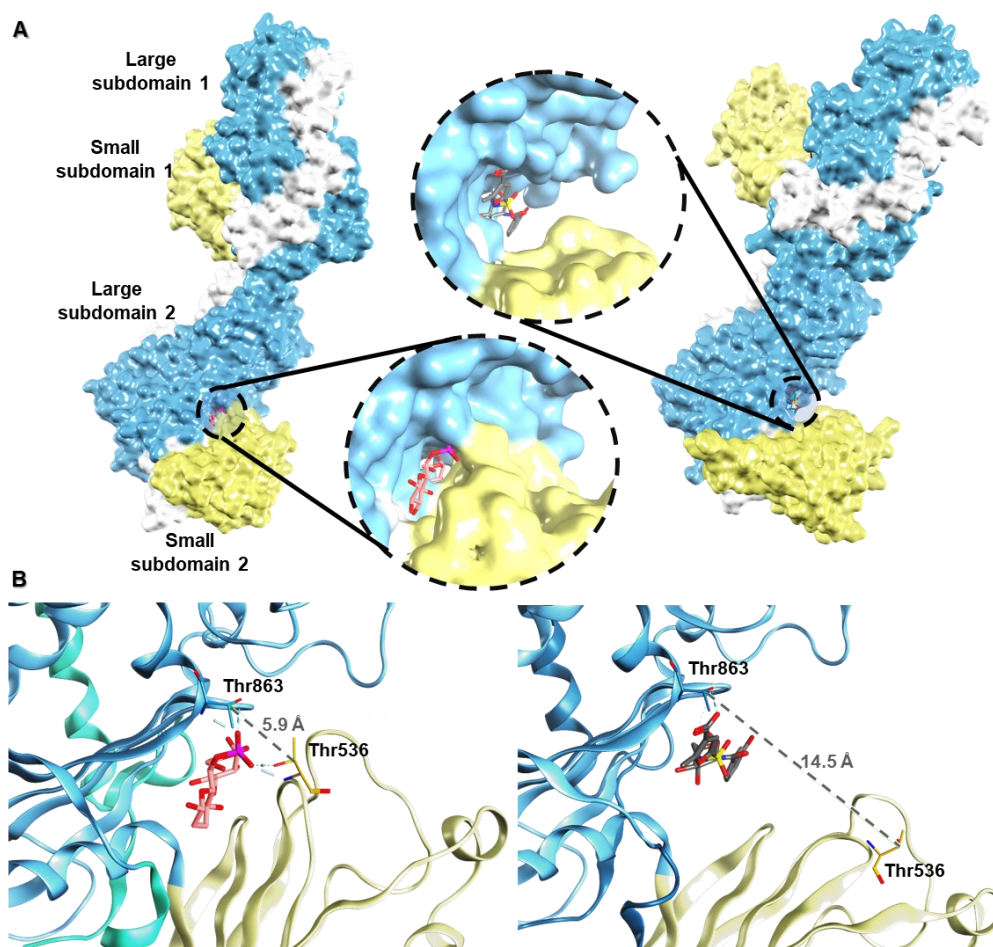


Figure 7. HK induced fit. **A**, molecular surfaces of PDB:2NZT, representing HK2 in the closed state (left), and PDB:5HEX, representing HK2 in the open state (right). The catalytic pocket is expanded, showing different conformations of the enzyme. The small (nonrigid) subdomain (yellow) and the large (rigid) subdomain (blue) are shown. **B**, HK2 interaction with glucose (PDB:2NZT) (left) and a potent inhibitor (PDB:5HEX) (right) in the catalytic pocket. The difference in the distance between Thr863 and Thr536 is highlighted in the two conformational states. Adapted from [19].

2.3. Expression of HK2

2.3.1. HK2 expression in normal cells

HK2 expression is present in normal insulin-sensitive tissues and depends on insulin levels. A phosphorylation cascade is triggered when insulin binds to and induces the intrinsic tyrosine kinase activity of its receptor. Phosphorylation of the insulin receptor substrate (IRS-1) leads to the activation of the phosphatidylinositol 3-kinase (PI3K) pathway, which in turn activates the p70/p85 S6 protein kinase (P70^{s6k}), inducing HK2 mRNA transcription, a rapamycin-sensitive process [89,90]. Catecholamines also induce HK2 expression, through activation of the cyclic AMP (cAMP) pathway, ultimately leading to activation of the cAMP response element (CRE). When active, CRE enhances transcription by binding to promoters of HK2 expression. Activation of nuclear transcription factor Y (NF-Y) also promotes HK2 expression [91].

2.3.2. HK2 upregulation in cancer cells

Numerous recent studies on different cancer types have noted HK2 overexpression [5,92–97] in many different tissues (**Figure 8**). Among these, cholangiocarcinoma [95], oral squamous cell [98], lung [99], cervical [100], colorectal [101], gastric [102], liver [103], breast [104], prostate [105], and laryngeal [5] cancer are well-documented, as are brain astrocytoma, glioblastoma, and oligodendroglioma, cervical squamous cell carcinoma, head and neck squamous cell carcinoma, renal cell carcinoma, diffuse large B-cell lymphoma, Hodgkin's lymphoma, peripheral T-cell lymphoma, pancreatic ductal adenocarcinoma, and papillary thyroid carcinoma [33].

In addition to overexpression in cancer cells, HK2 is upregulated in cancer-associated stromal fibroblasts [106], as well as in severe inflammatory states, consistent with the increased glycolytic activity of immune system cells [99].

The HK2 promoter region was found to be activated by multiple variables, such as low oxygen concentration or the presence of phorbol esters, and has a promiscuous potential to participate in different signal transduction cascades [89]. Genetic and epigenetic mechanisms, related to oncogenes, tumor suppressor genes (e.g. mutated p53) and micro RNAs (miRNA), have been reported to promote HK2 overexpression in cancer cells [5,89,107]. However, the main influential mechanism seems to be the hypoxia-induced factor-1 α (HIF-1 α) activation, which is increased in many cancers. HIF-1 α , the α subunit of the HIF-1 dimer, acts as the major transcriptional regulator of cancer metabolism, increasing the transcription of genes involved in glycolysis, cell survival/proliferation, angiogenesis and metastasis. It is also related with chemo- and radio-resistance of several cancers, since it is capable to adapt to new imposed conditions through changes in cancer metabolism [5]. Upon activation of HIF-1 heterodimer (e.g., by hypoxic, inflammatory, metabolic and oxidative stress) expression of some isoforms of glycolytic enzymes is induced, such as HK2, leading to increased levels of glycolysis but

also decreased OXPHOS and mitochondrial function [29,34,108,109]. Notably, HK1 expression does not seem to be altered by HIF-1 [108].

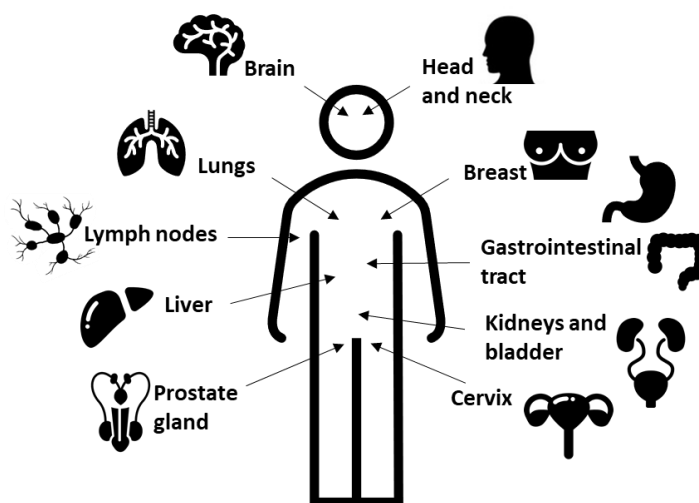


Figure 8. Examples of tissues where cancer cells have shown overexpression of HK2 [5,92–97].

2.3.3. HK2 downregulation

Downregulation of HK2 has also been investigated. One of the most common strategies is through antagonism of the phosphatase and tensin homolog (PTEN) on the PI3K pathway [5,110]. Likewise, taurine upregulated gene 1 (TUG1), a long noncoding RNA, seems to participate in the regulation of early stages of cancer development, as well as the regulation of HK2 expression at the translational level, its ubiquitination and subsequent degradation [111]. Similarly, sirtuin 6 (SIRT6), a mammalian NAD⁺-dependent histone deacetylase, is a negative regulator of HK2 [112]. HK2 has also been reported to be a substrate for chaperone-mediated autophagy (CMA), which could be exploited to eliminate cancer cells by inducing a metabolic catastrophe, through degradation of HK2 [113]. Moreover, several miRNAs have been demonstrated to regulate glucose metabolism and cell proliferation and some have been shown to inhibit HK2 expression and tumor cell growth [114–121].

Exogenous downregulation has also been reported. Several small molecules have been found to reduce HK2 expression, namely the glycolytic inhibitor, 2DG (**IV**, **Figure 4**) [58]. Among these (**Figure 9**), licochalcone (**IX**) has been shown to affect HK2 expression by blocking the Akt signaling pathway in gastric cancer cells [122]. Chrysin (**X**) has also been reported as a glycolysis inhibitor and apoptosis promoter, which reduces HK2 expression, possibly by directly destabilizing HIF-1 α [123]. Resveratrol (**XI**), a known scavenger of reactive oxygen species (ROS) with recognized anticancer properties, has also been shown to inhibit HK2 expression. This molecule was found to inhibit several signaling pathways, including PI3K [17]. Fenofibrate (**XII**) is capable of detaching HK2 from the MOM; it has been shown to reduce HK2 expression and favor the TCA cycle in oral squamous cell carcinoma [98].

Jolkinolide B (**XIII**) is another recently described molecule able to downregulate glycolysis in lung cancer cells, including the expression of HK2, by interfering with the Akt/mTOR pathway; the compound had a moderate overall effect on cell proliferation ($IC_{50} = 60\text{--}80\text{ }\mu\text{M}$) [124]. Halofuginone (**XIV**) also inhibits the Akt/mTOR pathway, leading to diminished HK2 expression; it has been shown to slow glycolysis, inhibit cell proliferation and induce apoptosis, in addition to inhibiting lipid synthesis and reducing NADPH production [125]. Finally, oroxylin A (**XV**) was shown to inhibit the s-Src proto-oncogene pathway (c-Src/Akt/HK2), which resulted in controlled metastasis propagation due to sensitization of lung cancer cells to apoptosis following detachment from the extracellular matrix [126].

2.3.4. HK1 Expression in Cancer Cells

Contrasting with HK2, little is reported about the contribution of HK1 to tumorigenesis. This isoform was occasionally found to be overexpressed in cancer cells, including from breast, bladder, colon, renal and esophageal cancer [89,93,127]. Recent data revealed the involvement of a common growth factor, the platelet-derived growth factor (PDGF), in the overexpression of both HK1 and HK2. PDGF was found to activate the c-Src proto-oncogene, which is involved in many human cancers that affect multiple tissues, namely breast, colon, liver, lung, pancreas and prostate [117].

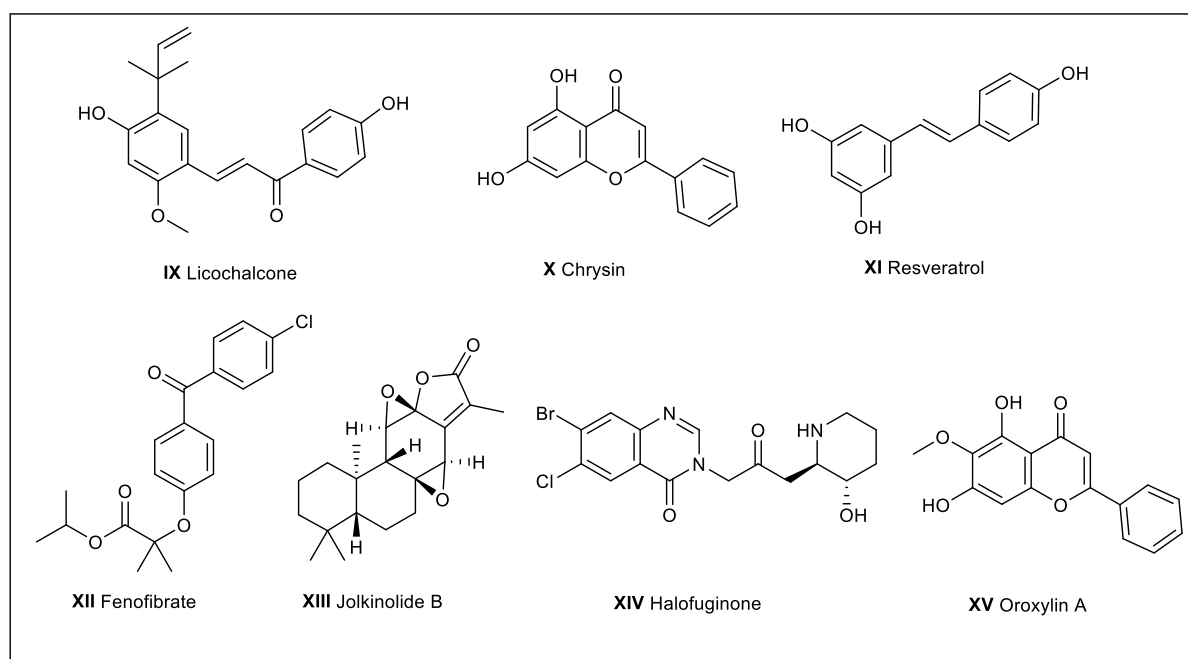


Figure 9. Structures of small molecule inhibitors of HK2 expression.

2.4. HK anti-apoptotic activity

2.4.1. HK-VDAC interaction

Mitochondrial function is crucial for cell survival and its impairment is associated with numerous diseases, such as neurodegenerative and cardiovascular disorders, as well as cancer. VDAC is now documented as a regulator of metabolic and energetic function, acting as a key contributor to the cell death/survival duality, with roles on normal mitochondrial function, cytoprotection, and mitochondria-induced apoptosis [128].

In cancer cells, translocation of HK2 to the MOM is induced, wherein approximately two thirds of the enzyme are bound to VDAC [89]. The serine/threonine-specific protein Akt kinase activates HK2 by phosphorylation of residue Thr473 and enables a stronger interaction, enhancing glycolysis by the access of freshly synthesized ATP and reducing the sensitivity of HK2 to be inhibited by the product G6P [129–131].

The HK/VDAC interaction is thought to stabilize the mitochondrial membrane, control the redox potential and prevent pro-apoptotic factors (Bax and Bak) from binding to MOM [16,45,48], thereby averting destabilization of the mitochondrial membrane and liberation of other pro-apoptotic elements (*e.g.*, cytochrome *c*) that would trigger apoptosis [45,49,50]. As mentioned above, this mechanism is relevant for the immortalization of the cancer cells. The mode of binding between HK2 and VDAC is a subject of discussion and has not been uncovered yet, but the *N*-terminal α -helices of each protein have been recognized to be essential for the interaction [132–134].

Under this scope, HK2 also plays important roles in other pathways. Specifically, HK2 acts as protector against ischemia and ischemia reperfusion injury. During ischemia, OXPHOS is inhibited by the lack of oxygen and glycolysis is the main route to obtain ATP and G6P for anabolic and catabolic pathways. While reperfusion of oxygen leads to oxidative stress and inflammation, the ability of HK1/2 to bind to MOM and prevent apoptosis can preclude premature cell death [135]. In this regard, HK2 inhibition must be taken with precaution towards this type of disorders.

2.4.2. HK2-TIGAR interaction

The TIGAR (TP53-induced glycolysis and apoptosis regulator) protein can function as the fructose-2,6-bisphosphatase domain of PFK2/FBPase-2, converting fructose-2,6-bisphosphate into fructose-6-phosphate (F6P). Since fructose-2,6-bisphosphate is an activator of PFK, TIGAR can downregulate glycolysis and enhance PPP, ultimately leading to the biosynthesis of nucleotides and regeneration of NADPH. This product is then crucial for the regeneration of GSH and ROS deactivation [136].

An HK2-TIGAR complex, capable of modulating HK2 activity, was identified in the mitochondria during hypoxia. The complex appears to be formed by interaction of the last residues of the TIGAR C-terminus with HK2 (the specific HK2 residues involved are unknown). This interaction seems to increase HK2

activity, enhancing the cell's machinery against ROS, and stabilizing the mitochondrial membrane potential, thereby contributing to the anti-apoptotic role of HK2. In terms of metabolism, the association of the two proteins enhances G6P levels, which explains the consequent protection of the cell against oxidative stress, for instance during ischemia reperfusion [137,138]. Moreover, the specific interaction of TIGAR with HK2, but not HK1, may be the reason why HK2 is associated with anabolic processes while HK1 is involved with catabolic processes [57].

3. Available HK2 Inhibitors

Several glycolytic inhibitors have been patented and published since long ago; however, while many were tested in clinical trials, they never reached the market [139,140]. Among these inhibitors are metformin [141–144] (**XXV**), 3BP [145] (**VI**), 2DG [146,147] (**IV**), LND [148] (**V**) and analogs and derivatives from those, claimed as HK inhibitors, but their effect is rarely noticed at concentrations below the millimolar range. A representative number of such structures is shown on **Figure 10 (A and B)**, along with relevant experimental data, obtained with human or mammal HK enzymes, unless noted otherwise.

In more recent years, patents on glucosamine derivatives as HK inhibitors have arisen in a much more promising perspective, with strong HK2 inhibitors in the nanomolar range [149,150]. The effects of glucosamine derivatives on yeast [151,152] and bovine brain [151] HKs have long been recognized. Maley and Lardy, in 1955 [151], and Coats *et al.*, in 1992 [152], synthesized a battery of *N*-acyl-glucosamines (e.g., **XXVIII-XXXI**, **Figure 11**) apparently capable of inhibiting those enzymes. From the new era of glucosamine HK2 inhibitors, the first relevant work was developed by Lin *et al.*, where 2,6-disubstituted glucosamine derivatives with remarkable potency and selectivity were developed [43,150]. This study started from a hit compound ($IC_{50\text{ HK2}} = 6.3\text{ }\mu\text{M}$, $IC_{50\text{ HK1}} = 2.0\text{ }\mu\text{M}$) obtained by high throughput screening, which was then derivatized following a rational design. The objective of targeting the G6P pocket was achieved and compounds with an effect in the very low nanomolar range and highly selective for the HK2 isozyme were found (**Figure 11**), along with other non-selective HK1/HK2 inhibitors [43]. Furthermore, the disclosure of crystallographic structures of the enzymes with inhibitors encouraged the understanding of the main interactions involved in the inhibition. On the other hand, despite the need for further SAR studies to fully comprehend the selectivity mechanism, this work uncovered some key structural features for effective HK inhibition: *a*) presence of a sugar moiety; *b*) ability to occupy the total extension of the catalytic pocket; *c*) ability to mimic phosphate groups and allow H-bond formation; *d*) presence of large substituents, which may provide additional interactions and block the entry of substrates or G6P. Selectivity towards HK1 seems to be improved by large substituents on the 2-position: *o,m*-Cl or *p*- substituents such as CN. These characteristics have provided effective compounds (IC_{50} s in the low nanomolar range), some of which up to 2 orders of magnitude more potent for HK2 than for HK1. While selectivity toward HK2 over HK1 does not appear easy to achieve, in this work it was demonstrated to be possible.

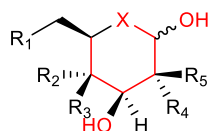
Likewise, another patented work on glucosamine derivatives also reports promising molecules in the nanomolar range of inhibition (**Figure 12**). In contrast to the molecules referred above, in this work glucosamine is mostly substituted in the amine group, forming an amide bound to an indole or oxazole moiety that is further derivatized. Oxazolines fusing the sugar ring by the oxygen on C1 with the amine on C2 were also reported, showing that a free hydroxyl group on C1 is not essential for the activity (**Figure 12, LIII, LIV**) [149].

In the last years few SBDD strategies were followed and several non-sugar like molecules were developed (**XXXII-XXXIV**) [153–155]. Namely, benserazide (**XXXII, Figure 10**), a known “frequent hitter” [156], was found to inhibit HK2 modestly and was used to the development of lead compounds, with better activity profile. The authors claimed the discovery of a new scaffold for HK2 inhibition, despite the still weak *in vitro* results [155]. Despite the potential of these molecules to promiscuity, no other off-target interactions were evaluated, and inhibition of the isozyme HK1 was not investigated.

Other inhibitors of HK have arisen in search for antiparasitic molecules (**XXXV-XXXIX, Figure 10**) [157–161]. Evaluation of these molecules against human HKs could generate useful information about the action of these molecules in the host.

Among known inhibitors, the most common structural features are sugar-like moieties, particularly glucose analogues that maintain several hydroxyl groups. Aromatic (e.g., benzoyl or phenyl) and halogen (e.g., Cl, Br) substituents are also frequent, increasing the lipophilic character of the molecules. Carboxylic acid groups or bioisosteric replacements, such as ester and sulfonamide groups, are also shared by the majority of the structures. A comparison of structures and biological activities suggests that the most potent molecules are long and not significantly branched, with amphipathic character.

Methyl jasmonate (**LV**) derivatives were found to disrupt the interaction between HK and VDAC [162]. More recently, some molecules were developed and found to present satisfactory inhibition in both biochemical and cell-based assays (**Figure 13**) [163]. Regarding the interaction between HK2 and VDAC, other compounds have been shown to disrupt it, promoting apoptosis, namely the glycolytic inhibitors 3BP (**XXIV**) [41] and lonidamine (**XXV**) [164], clotrimazole (**LVI**) [45,59], fenofibrate (**X**) [98] and some peptides derived from the *N*-terminal regions of VDAC [165,166] or HK2 [167].

A

	X	R1	R2	R3	R4	R5		
IV 2DG	O	OH	OH	H	H	H	rHK1	400
III G6P	O	OPO ₃	OH	H	OH	H	mHK1/2	20
XVI 2-DGal	O	OH	H	OH	H	H		
XVII DG6P	O	OPO ₃	OH	H	H	H		
XVIII FDG	O	OH	OH	H	F	H		
XIX FDG	O	OH	OH	H	F	F		
XX FDM	O	OH	OH	H	H	F		
XXI GlucA	O	OH	OH	H	NH ₂	H	rHK1	200
XXII 5-TGluc	S	OH	OH	H	OH	H	rHK1	170
XXIII 5-TG6P	S	OPO ₃	OH	H	OH	H	rHK1	0.40
XXIV FDGal	S	OH	H	OH	F	H		

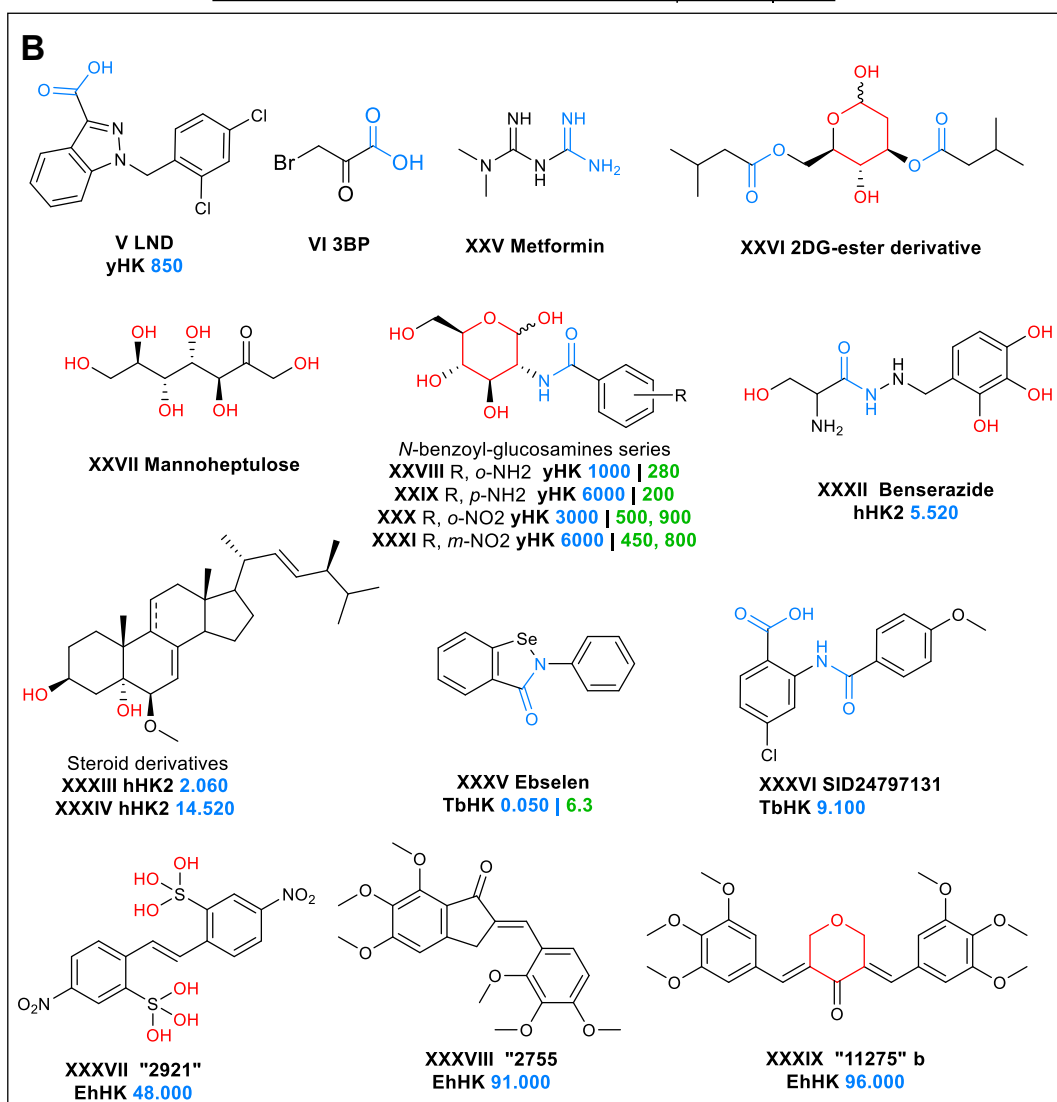
B

Figure 10. Structures of described HK2 inhibitors. A - Reported HK2 inhibitors derivatives/analogs of hexoses. B - Other structures of HK2 inhibitors. *K_i* values are indicated in green. Common features are highlighted in blue and red. When available from *in vitro* biochemical assays, *IC*₅₀ concentrations (in μ M) are indicated in blue and *K_i* values are indicated in green. The relevant references are cited in the text. EhHK, *Entamoeba histolytica* HK; hHK, human HK; mHK, mammal HK; rHK, rat HK; TbHK, *Trypanosoma brucei* HK; yHK, yeast HK.

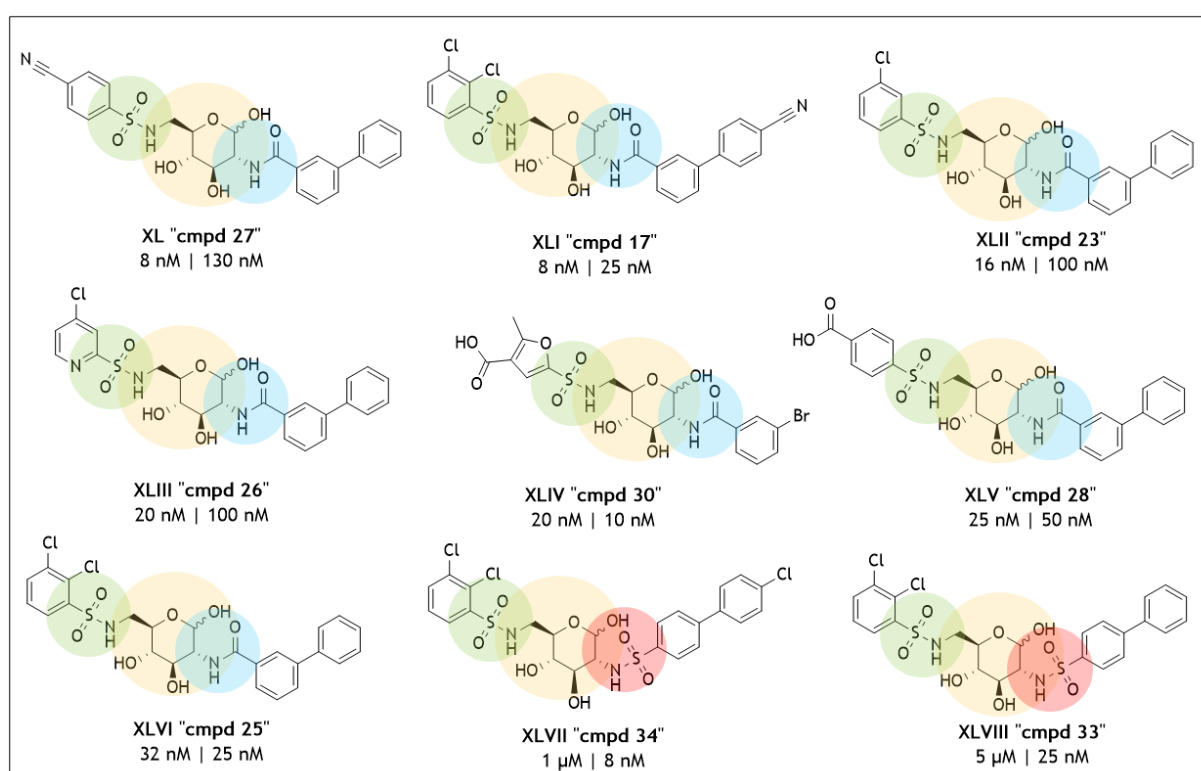


Figure 11. Potent and selective HK1 and HK2 inhibitors [43]. The compound designations in quotation marks maintain the numbering used in the original study. The concentrations under each structure represent the corresponding IC_{50} values for HK1 (left) and HK2 (right).

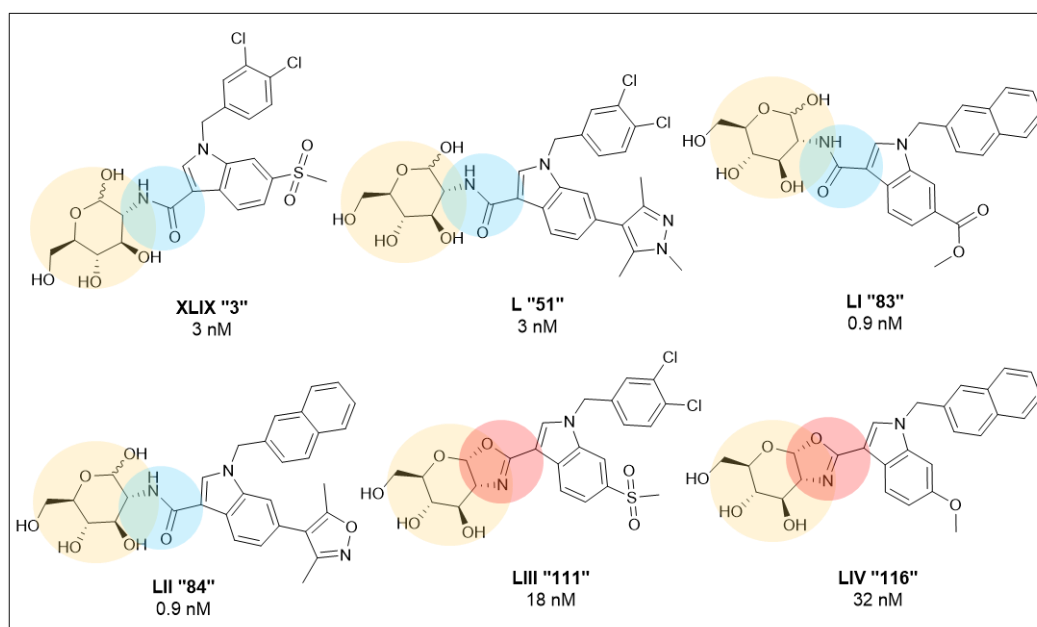


Figure 12. Potent and selective HK1 and HK2 inhibitors [43]. Numbers in quotation marks maintain the numbering used in the original study. The concentrations under each structure represent the corresponding IC_{50} values for HK2 [149].

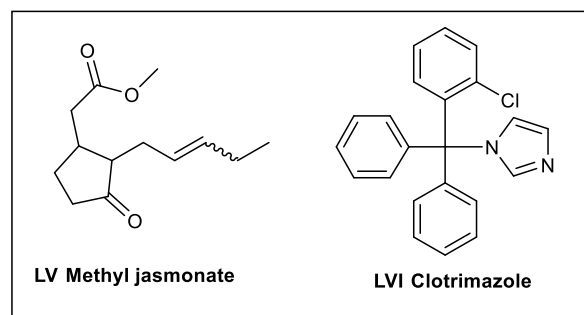


Figure 13. Structures of the HK2-VDAC interaction inhibitors **LV**, methyl jasmonate, and **LVI**, clotrimazole.

4. Targeting HK2

HK2 overexpression has been associated with a myriad of tumors (*vide supra*). Although HK levels may differ according to the specific tumor, and broad conclusions about the ubiquitous overexpression of a particular isoform must be regarded with caution, HK2 has been considered an important cancer biomarker, related with advanced state and aggressive tumors and a poor prognosis [93,168,169].

HK1/HK2 have multiple biological roles in several pathways and diseases other than cancer (Fig. 14), some of which were unveiled very recently [170–172]. Indeed, HKs have been identified as one of the metabolic enzymes types that can act as protein kinases and phosphorylate different protein substrates [88]. Understanding the HK2 interactome is thus essential to comprehend the strengths and limitations of using HK2 inhibition as a therapeutic strategy.

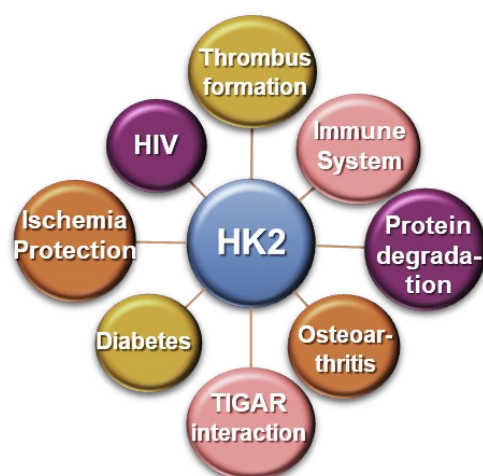


Figure 14. Examples of diseases/pathways, other than cancer, with HK2 involvement.

4.1. *Proof of concept*

The consequences of HK2 gene knockout and HK2 ablation were investigated, and several reports are described below.

In a proof-of-concept study for the HK2 targeting for cancer treatment, Patra *et al.* [18] used a conditional knockout mouse model to demonstrate that deletion of the *Hk2* gene has an impact on tumorigenesis, by showing that HK2 was required for tumor initiation and maintenance in KRas-driven lung cancer, and ErbB2-driven breast cancer, even though HK1 continued to be expressed and did not change significantly. In immortalized embryonic fibroblasts generated from these mice, HK2 deletion caused impaired oncogenic transformation, glucose consumption was reduced by 50% and the total remaining HK activity was reduced by 70%. *In vivo*, *Hk2* deletion impaired tumor growth and development, and restored expression of HK2 led to restored tumorigenesis. Similarly to the mouse model, HK2 ablation inhibited the neoplastic phenotype of human lung and breast cancer cells *in vitro* and *in vivo*. This study also demonstrated that HK2 is critical to divert glucose into metabolic pathways required for anabolic processes in cancer cells (*e.g.*, ribonucleotide synthesis *via* PPP, or serine biosynthesis). Moreover, systemic *Hk2* deletion did not appear to exert adverse physiological consequences in the mice [18]. In line with these results, knocking down HK2 expression significantly attenuated cancer cell growth in mouse and human lung cancer cell lines [97,173].

The required presence of HK2 for tumorigenesis was further confirmed recently in a mouse model for hepatocarcinogenesis, in which deletion of hepatic HK2 inhibited the tumor incidence without severe physiological consequences. In human hepatocellular carcinoma cells, HK2 ablation also reduced proliferation by ~50% and increased cell death. Glucose uptake was reduced by approximately 40%, as well as lactate production and pyruvate secretion, and no effect on glutamine uptake was observed. Combination of HK2 depletion with the chemotherapeutic agent sorafenib and metformin largely increased the therapeutic effect [174]. Likewise, HK2 is also required for prostate cancer development and its deletion induces both cytostatic and cytotoxic effects, with effects both *in vivo* and *in vitro*. Moreover, chemoresistance to etoposide was dramatically decreased in mice [175].

Under this scope, HK2 knockdown was also shown to overcome resistance and sensitize human colon and cervical cancer cells to the actions of 5-fluoruracil and radiotherapy, respectively [176,177]. Furthermore, the combination of HK2 downregulation by resveratrol with sorafenib administration resulted in enhanced anticancer effects, including decreased resistance to the chemotherapeutic agent in mice [17].

4.2. Strengths

HK2 ablation/inhibition should entail a substantial decrease of glycolysis that affects multiple branched pathways of central metabolism such as glycogenesis, the hexosamine pathway (protein glycosylation), the biosynthesis of essential nutrients (e.g., amino acids, nucleotides, and fatty acids), as well as the metabolic energy balance, affecting both fermentation and the TCA cycle [13].

The general effects of HK2 knockout or downregulation have consistently been associated with **1) reduction of cell proliferation**, **2) lower glycolysis levels** and eventually **3) enhanced cell death** related to the detachment of HK2 from MOM. Additionally, the concomitant use of chemotherapeutic agents tends to increase the therapeutic effect, **4) overcoming resistance**. One well-known example is the synergistic enhancement of cisplatin-induced cytotoxicity that results from disruption of the HK2-VDAC association [16]. Moreover, the available evidence is consistent with **5) no enhancement of HK1 expression in cancer cells** to overcome HK2 deficiency [18,108].

Regarding the potential for synergy, **6) HK2 inhibition might be an ally of immunotherapy** [178]. Despite T-cells acquiring energy primarily from oxidative glycolysis and being highly dependent on HKs, a compensation of HK1 overexpression specifically in CD4 and CD8 T-cells in adult mice with systemic deletion of HK2 was reported [171].

Interestingly, **7) the presence of HK1 in normal tissues that mostly express HK2 (insulin-dependent tissues) would be enough to maintain normal glucose homeostasis** under resting conditions and a normal diet in the absence of HK2 activity [179].

Also importantly, the current *in vitro* and *in vivo* data indicate that **8) the absence of HK2 can provide a therapeutic benefit without limiting side effects** [18].

Besides the effect in cancer cells, **9) other diseases and disorders might benefit from HK2 inhibition**. HK2 inhibition was found to be relevant for reduction of neuroinflammation and seems to be a promising strategy to treat hypoxia-associated neural injuries, namely ischemic strokes [180]. Regarding the immune system, the inhibition of HK2 and its relocation to the cytosol seem to induce the inflammasome activation, and stabilize the major histocompatibility (MHC-1) clusters [181,182]. HK2 inhibition has also emerged as a sustainable strategy in the therapy of coagulation disorders, since there is an indication of HK2 overexpression in thrombotic plaques that affects the regular initiation of the coagulation cascade [183].

4.3. Limitations

One of the main drawbacks of the inhibition of HK2 could be the predominance of this HK isozyme in the heart, which could limit the targeting of HK2 for cancer treatment. Nonetheless, under normal aerobic conditions, the adult heart relies on fatty acid oxidation for energy production, and only 5% of cells are estimated to use glycolysis [135]; as such, partial ablation of HK2 did not lead to altered glucose

metabolism [179]. Also noteworthy, adult skeletal muscles generally use fatty acids as alternative energy substrates [18].

Still, as referred above (section 2.7.1.), interaction of HK2 with VDAC confers protection to the cell by precluding apoptosis not only in cancer cells, but also in ischemia and reperfusion injury, **1) avoiding premature cell death**; therefore, HK2 inhibition could provoke damages.

Moreover, **2) HK1/2 are crucial for protection against the early onset of Parkinson's disease**. This effect is due to the fact that Parkinson's disease is closely related with mitochondrial dysfunction, involved with TIGAR, mutated DJ-1 protein, and PINK1/parkin pathway, all affected by HK2 [138,170,184,185].

Diabetes mellitus is characterized by a defective glucose homeostasis. A suggested strategy for type-II diabetes treatment is the use of HK4-activators to enhance glucose uptake and induce insulin secretion. In turn, **3) low HK2 expression and G6P levels were found in non-insulin-dependent diabetes mellitus**. Regarding these observations, and the known HK role in glucose homeostasis, HK2 inhibition should be regarded with caution in therapies aimed at diabetic individuals [66,186–188].

Other canonical and non-canonical roles of HK2 are still being elucidated. More information is expected to emerge soon regarding a better understanding of the pathways in which HK2 is involved, how its modulation may be applied for therapeutic purposes, and what are the likely consequences/side effects of this modulation.

Although the potential of HK2 as a therapeutic target has been recognized before, its druggability has been disputed. The hydrophilicity of residues in the catalytic pocket has been regarded as a limitation to designing molecules with suitable pharmacokinetic properties. Moreover, the catalytic pockets of HK1 and HK2 are virtually indistinguishable, hampering selective inhibition of HK2. Nonetheless, even if selectivity is not fully achieved, a non-selective HK inhibitor would still be expected to preferentially affect cancer cells, due to increased glycolysis compared to normal cells, thereby limiting side effects to some extent.

4.4. Strategy to target HK2

Taken together, the data described in this chapter support the inhibition of HK2 as a promising target for cancer treatment, particularly in combination with cytotoxic compounds. Exploring an inhibitory effect at the catalytic site with small molecules seems to be the most straightforward strategy, as long as selectivity is achieved. Indeed, although gene therapy, involving specific HK2 ablation, has been suggested [189], the applicability of this type of therapy is not trivial yet.

In this regard, this research project envisions the development of potent and selective HK2 inhibitors as potential anticancer agents that prevent resistance mechanisms. The available information on the expression, functions and activities of HKs, their structures and known inhibitors presented herein assisted to perceive the heart of the matter.

Using the structural features of HK2 as a base for the rational design of selective HK2 inhibitors with proper pharmacokinetic and toxicological properties, an SBVS was devised to obtain “virtual” hits with the potential to inhibit HK2. The potential hits were acquired or synthesized and biochemically and biologically evaluated (**Figure 15, A**).

Achieving selectivity against HK2 over HK1 is one of the greatest challenges in the field. Despite the high similarity and identity between HK1 and HK2, certain properties are different such as the response to the presence of inorganic phosphate. While selective molecules have already been achieved, strong rationale for the observed selectivity. Has not been found yet. Herein, a structural comparison between HK1 and HK2 is also reported using sequence/static structures and molecular dynamics, to uncover differences that can potentially be exploited in the design of more selective HK2 inhibitors, assisting in the future development of this type of modulators (**Figure 15, B**).

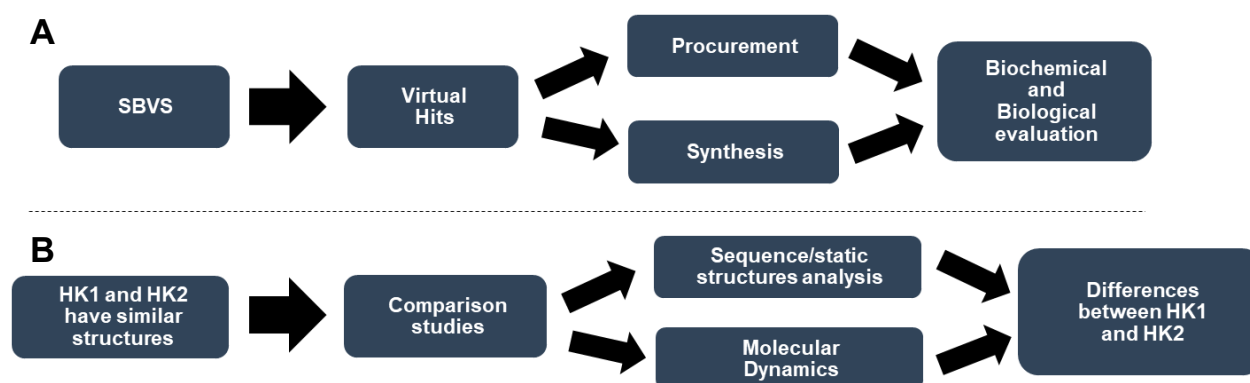


Figure 15 . Schematic view of the tasks explored in this research project. **A** corresponds to the path used to find potential HK2 inhibitors and their biological and biochemical evaluation. Path **B** was used to find differences between HK1 and HK2 to improve selectivity of further inhibitors.

Chapter III

Structure-Based Virtual Screening

1. Overview

Potent HK2 inhibitors were developed by Lin *et al.* after a high-throughput screening (HTS) campaign with millions of molecules wherein the hit compound that later gave rise to more potent molecules was discovered [43]. Despite the increasing progress in HTS campaigns, HTS still does not mean success, as its costs are immeasurable [190,191], and it is usually not accessible to most researchers. Fortunately, other ways to screen readily-available molecules that cover a wide chemical space have emerged, mainly due to the great improvement in the computational field related to biochemistry and medicinal chemistry.

In this chapter, a structure-based virtual screening (SBVS) campaign designed to find HK2 inhibitors is described (**Figure 16**). After developing an appropriate molecular docking protocol, several virtual databases with thousands of molecules were screened. Following identification (and selection) of potential binding of the molecules to the receptor of interest, the molecules were filtered according to their predicted physicochemical and pharmacokinetic properties. Finally, the protocol was biochemically validated after procurement of 64 potential HK2 inhibitors. The identity of the acquired molecules was confirmed by NMR and MS. In addition to the molecules selected from the SBVS campaign, a few metal complexes where ruthenium is coordinated with carbohydrate-containing ligands were also evaluated as potential HK2 inhibitors. The next subsections (1.1. – 1.5.) will enlighten the reader about a few important concepts related to the methodologies explored in this chapter.

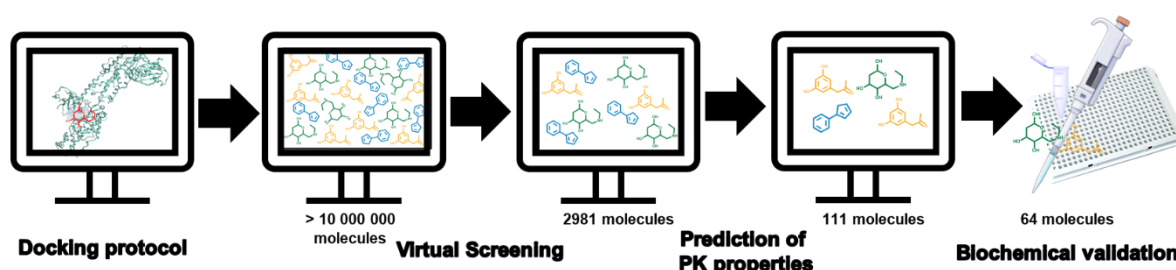


Figure 16. Schematic view of the steps taken to conclude the SBVS campaign to find potential HK2 inhibitors. PK: Pharmacokinetics.

1.1. Insights into the Structure Based Virtual Screening approach

The drug development pathway from hit discovery to commercialization is known to be highly time-consuming and costly [192] (**Figure 17 - A**). Computer-aided drug discovery (CADD) has emerged as a strategy capable of assisting several steps in the drug discovery process in order to reduce time and costs, namely in the hit discovery and hit-to-lead stages. Hit discovery by computational approaches has met “limited” success so far, but due to improved algorithms and computational power, it is becoming more reliable and is foreseen to be increasingly ingrained in the pipeline. Moreover, failure at an early stage is inexpensive, and the new data can be used to refine the method without large investments [193,194]. Following the resolution of the first protein structures by X-ray crystallography, structure-based drug design (SBDD) soon emerged as a widely

used key methodology and as an important branch of the rational drug design process. This structural biology strategy of pursuing the “perfect fit”, coupled with a constant progress in computational techniques, has resulted in the development of numerous drugs [195].

Regarding the druggability of HK2, the difficult access to known inhibitors, analysis of the catalytic pocket, and the availability of recent crystallographic structures with reasonable resolution in interaction with inhibitors, meant that an SBDD strategy would be more plausible to find novel inhibitors, rather than methodologies involving, for instance, ligand-based knowledge. Following this concept, an SBVS strategy was developed using molecular docking (**Figure 17 - B**). SBVS allows a virtual evaluation of millions of potential small molecules HK2 modulators from selected virtual libraries. Using databases of commercial compounds, time and resources can be saved by skipping the synthesis of organic molecules in a first stage. When some of these molecules are not commercially available, at least there is a high chance that the synthetic route has been already published, simplifying the synthesis process. Computer-aided *de novo* design considering the catalytic pocket as a model could be an alternative strategy; however, organic synthesis of potential modulators must be carried out in this case. Since the molecules will tend to be novel, the synthetic route development could be time-consuming and only a small number of molecules would actually be experimentally tested. Virtual screening (VS) using a structure-based pharmacophore is another promising methodology; however, the knowledge about the most important interactions in the catalytic pocket for efficient HK2 inhibition is not fully unveiled, which precluded its use in this project.

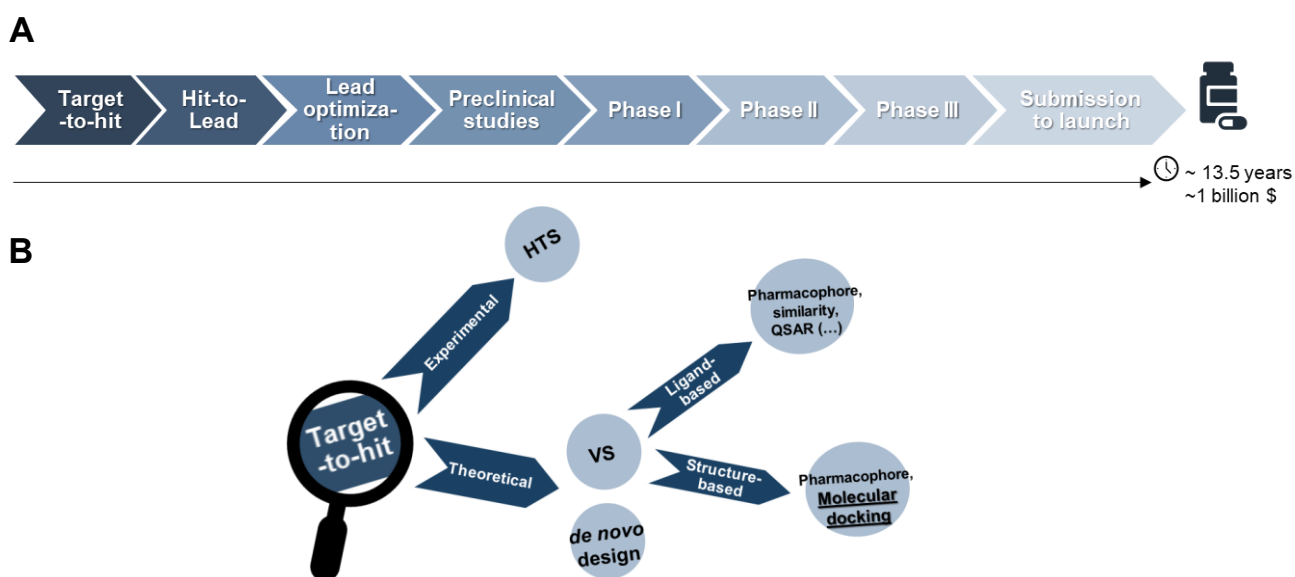


Figure 17. A- Simplified drug discovery pipeline. **B-** Experimental and theoretical approaches for hit discovery after selection of the biological target. In this project a theoretical approach was followed to find small molecule modulators for HK2, using molecular docking as the base technique for VS.

1.1.1. Molecular Docking

Molecular docking, possibly the most widely used method in SBDD, aims to predict as accurately as possible the best pose (orientation, conformation, affinity) of a ligand in a protein binding site. Docking consists of two parts, namely, the accurate prediction of the orientation (pose) of the bioactive conformation into the binding pocket, and the estimation of the tightness of target–ligand interactions (scoring). Continued improvements in docking grant users access to better searching algorithms and scoring functions. While the first step relies on generating multiple conformations that search the entire conformational space, scoring functions rank these conformations based on the calculation of multiple terms related to the interaction energy between the ligand and the receptor, including electrostatic interactions, hydrogen bonds and van der Waals interactions. Altogether, efforts are made to find the most probable binding mode between the receptor and the ligand, which might be a small molecule or another macromolecule [196–198]. In line with the goal of this project, only small molecules will be considered as potential ligands.

The search for suitable conformations involves a geometric and chemical compatibility to avoid the overlapping atoms and favor interactions between the receptor and the ligand. In conformational search, the structural parameters of the ligands, such as torsional (dihedral), translational and rotational degrees of freedom, are incrementally modified. The two critical factors in a search procedure are speed and efficiency in covering the relevant conformational space. On the other hand, the scoring functions should be fast enough to allow it to be applied to a large number of potential solutions and discriminate effectively between active and inactive conformations. Calculations can be performed using any pocket/cavity of the receptor, or in this case, using the catalytic binding site of HK2. Generally, no flexibility is assigned to the receptor structure (semi-flexible docking). This is seen as a limitation of the method in highly dynamic systems. Moreover, some flexibility can be added, for instance, to specific (and flexible) residues in the binding site. To improve the quality and reliability of the predictions, docking can be followed by molecular dynamics calculations in which the full flexibility of the receptor as well as the influence of the solvent is considered. However, the computational power and time required to modulate the interactions for thousands of molecules precludes the use of this technique for VS. Aiming at a balance between accuracy and efficiency, more simplified methods are accepted, considering a rigid receptor and a flexible ligand [196–198].

1.2. *Prediction of ADMET properties*

Until the end of the last century, poor pharmacokinetics accounted for the highest attrition rate in drug discovery. In recent decades, advances in medicinal chemistry and biological pharmacokinetic modeling have significantly reduced the failure due to unsatisfactory absorption, distribution, metabolism, excretion and toxicological (ADMET) properties [192]. The prediction of pharmacokinetic properties has thus become paramount, and methods have been developed along with prediction and simulation of receptor-ligand binding. Besides the important pharmacodynamic properties that a small molecule must achieve, the concern related with physicochemical and ADMET properties is equally important in order to find efficient, safe and inexpensive drugs. In this way, the chance of failure in later stages of the drug development pipeline has been considerably reduced, avoiding wasted resources [199–201]. Development in this field has mostly

arisen in recent years and several free or commercial software, webserver and databases are now available [202]. Overall, the prediction of pharmacokinetic parameters relies on models based on artificial intelligence derived from available *in vitro* data [200,203].

Filtering molecules during a screening process based on the prediction of these properties can be an asset in molecule selection. Nevertheless, finding the appropriate moment to apply these filters can be challenging. On the one hand, using these filters before VS simulations saves computational time, which will only be applied to more drug-like molecules [200]. For instance, this can be especially interesting when a particular property is required, such as crossing the blood-brain barrier. On the other hand, the possibility of removing molecules with potentially interesting pharmacodynamic properties too soon is a risk, since the physicochemical and pharmacokinetic properties of certain molecules may be improved at a later stage.

Another strategy is to apply these filters after VS calculations, the approach chosen for this project. In this case, the risk of losing potential ligands at an early stage was considered too high, due to the nature of the target (a very hydrophilic binding pocket). Furthermore, at the end of the VS, just a small percentage of the initial database is evaluated, which allows a more thorough analysis of the results and a more rational decision on the molecules selected for further steps.

As this is an iterative process, there is no straight path with mandatory steps along the way, but the understanding that ADMET prediction can be useful in different stages, depending on the project itself, has been confirmed to be profoundly important [203].

Following SBVS, the computational calculation of ADMET and physicochemical properties was carried out. To summarize, the well-known webserver FAF-Drugs4 (Free ADME-Tox Filtering Tool) [204,205] was used to generate different predictions/calculations, such as oral bioavailability, solubility, number of Lipinski violations and detection of PAINS (Pan-Assay Interference Compounds – false positive molecules) moieties (**Figure 18**). Different filters can be (and were) applied individually such as a profiler of protein-protein interactions inhibitors (IPPI), the Lilly MedChem Rules [206], or the retrieval of covalent inhibitors. Based on the presence of substructures and moieties known to be involved in toxicity problems, the server can categorize each molecule as ‘accepted’, ‘rejected’, or ‘intermediate’, depending on the structural alerts found. ‘Accepted’ corresponds to molecules without structural alerts and with proper physicochemical properties; an ‘intermediate’ result carries low-risk structural alerts, with a number of occurrences below the threshold; ‘rejected’ molecules are considered as such because they do not meet the physicochemical requests, contain a high-risk structural alert, and/or exceed the acceptable number of low-risk structural alerts [205]. Still, it should be up to the user to carefully evaluate the predictions and decide whether the motives for exclusion are reasonable for the ongoing project. To complement the results, the OCHEM (Online chemical modeling environment) [207] server allowed the prediction of toxicity and interaction with metabolic enzymes.

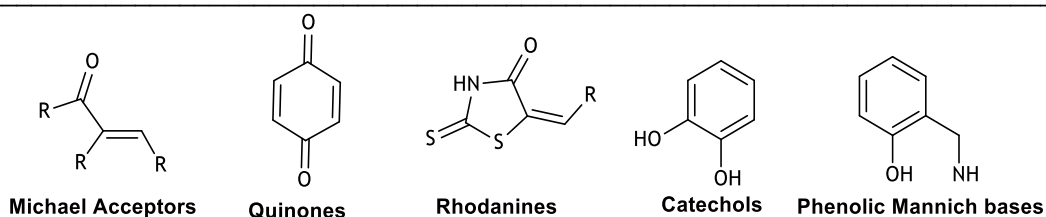


Figure 18. Examples of PAINS substructures.

1.3. Biochemical validation of the SBVS

After the theoretical calculations, the SBVS must be experimentally validated in order to assure that the model used is in fact able to generate adequate ligands for the molecular target. Preferentially, two different methods should be used to safeguard against potentially erroneous results due to interference with the assay. Here, a high-performance liquid chromatography (HPLC) assay, coupled with ultraviolet (UV) detection, and a luminescence-based assay were applied.

1.3.1. ADP detection by HPLC

HPLC-UV is a widely used technique to separate, analyze, and quantify the components of different types of samples able to absorb UV light at certain wavelengths. Quantification of ATP and ADP in different samples using HPLC-UV has been reported since long ago [208–210]. More recently, HK activity in mouse tissue homogenates was determined by this method. The biggest advantage of this method, compared to others, is related to the direct quantification of one of the products of HK activity (ADP), without any coupled reaction [211]. The method requires the construction of a standard curve for ADP and/or ATP to determine the concentration of the metabolite in the sample, using the calculation of the area under the curve (AUC).

1.3.2. ADP detection by Luminescence - ADP-Glo™ Kinase Assay

Luminescence detection is known to be one of the most sensitive and reliable techniques for ATP/ADP measurements, without interference from fluorescent compounds [212]. The ADP-Glo™ Kinase Assay is based on the bioluminescence generated by the biochemical conversion of luciferin into oxyluciferin, catalyzed by luciferase in the presence of ATP and O₂, releasing a yellow-green photon (**Scheme 1**). Briefly, HK2 converts glucose and ATP into G6P and ADP. In a sequence of reactions, the ADP-Glo™ Kinase Assay depletes the remaining ATP (by conversion to AMP) and then uses the previously synthesized ADP to generate ATP, which will be used in the last step for the luminescence generation. In this way, luminescence is proportional to the amount of ADP initially formed.

glucose and G6P, in both pockets of each chain. The remaining structures have complexed HK2 inhibitors: 5HEX, “cmpd 30” (the best inhibitor from that work) ($IC_{50} = 10$ nM), 5HFU, “cmpd 27” ($IC_{50} = 130$ nM), and 5HG1, “cmpd 27” ($IC_{50} = 6.3$ μ M). Altogether, these structures are able to represent different states of the pocket caused by the induced fit of glucose/natural ligands and different inhibitors. 2NZT represents a closed conformation of the pocket, while 5HG1 represents an intermediate structure and 5HEX and 5HFU represent an open structure of the pocket. **Table 3** shows some characteristics of these structures and the RMSD between structures that agrees with the stated above, where the largest differences are found in the comparison of 2NZT with 5HEX and 5HFU. All structures exhibit similar resolutions. To simplify the process, only chain A (of each of the structures) was used for the calculations. Between the *N*- and *C*-termini, the chosen pocket was the *C*-terminus, which is the one with more catalytic information available.

Table 3. Information on the available X-ray structures of HK2 on PDB [78]. The difference between structures is measured through RMSD calculation based on the Ca of the residues' backbone.

PDBID	Resolution (Å)	Year	Chains	RMSD (Å)			
				2NZT	5HEX	5HFU	5HG1
2NZT	2.45	2006	2	-	6.66	6.60	5.12
5HEX	2.73	2016	2	6.66	-	0.44	4.38
5HFU	2.92	2016	2	6.60	0.44	-	4.36
5HG1	2.76	2016	1	5.12	4.38	4.36	-

The pocket used to perform the calculations includes glucose and G6P cavities. **Figure 20** shows the interactions (2D) of HK2 residues (2NZT [73]) with both endogenous ligands, through hydrogen bonds, and with nearby residues. As mentioned in chapter II, section 2, Asp657 is thought to be one of the key residues responsible for facilitating the transfer of a phosphate group to glucose, forming G6P. This residue, as well as a set of seven other residues (**Figure 20**) (Asn656, Gly862, Glu742, Glu708, Asp532, Asn683, and Thr680), were chosen to be tested as the binding center for docking calculations. Water molecules were observed near the catalytic pocket involved in ligand-receptor interactions. Therefore, the influence of water in the pocket was tested by keeping or removing water molecules from the calculation, which revealed that their presence did not produce a relevant effect. All water molecules were thus stripped away. The crystallographic ligands were docked to the four available HK2 crystallographic structures (cross-docking) with GOLD software (Chemscore and GoldScore scoring functions), using the different selected residues as the center, and a radius of 10 and 15 Å for the spherical searching conformational space.

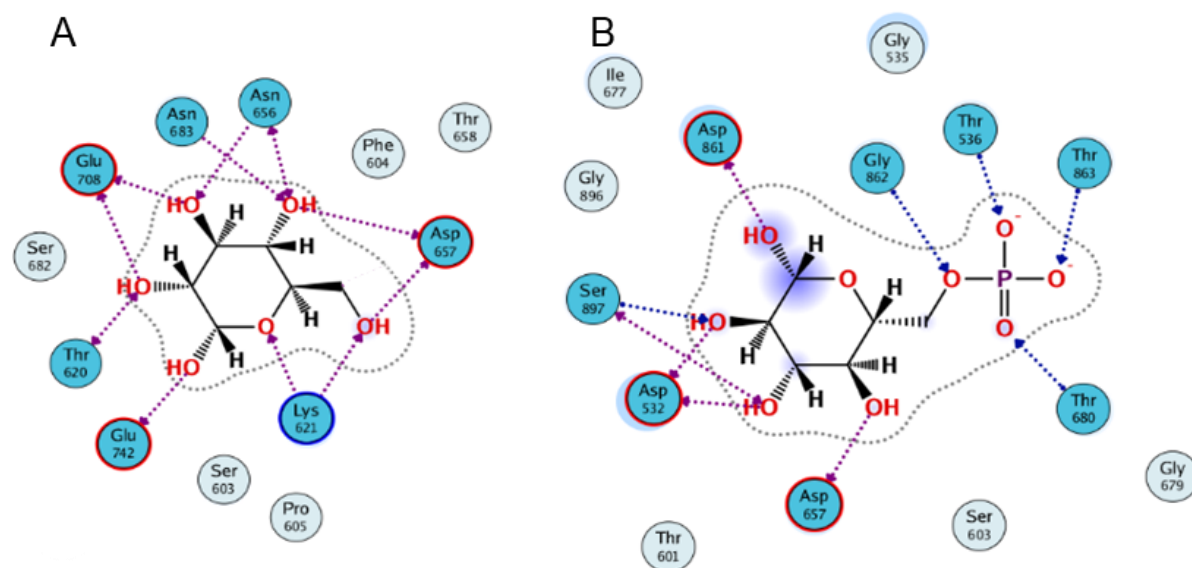


Figure 20. Interactions of HK2 residues (PDBID: 2NZT) with glucose and G6P (depicted at blue, with arrows – H-bonds) in the respective cavities. Nearby residues are also depicted at lighter blue.

The RMSDs between the docked and crystallographic poses were compared. RMSD, the root-mean-square deviation, measures the difference between the crystallographic ligand pose (atom coordinates) and the best pose generated by the molecular docking software. **Table 4** shows the best systems for each X-ray structure. For all systems, the GoldScore scoring function performed best, as did a radius of 15 Å. Among all, the systems with 5HG1 performed better with any residue as the center, being able to more accurately reproduce the crystallographic poses of the ligands. In some instances, the results were considerably close, as **Table 4** describes for the systems with residues Glu742 and Asn683 as centers. Cross-docking allowed the selection of a 15 Å radius and the 3D structure 5HG1 for the docking protocol.

Table 4. Best RMSD and scores for each X-ray structure (PDBID: 2NZT, 5HEX, 5HFU, 5HG1) obtained after molecular docking, using different residues as the center of calculations. A radius = 15Å was used in all described systems, as well as the GoldScore scoring function. Glucose, “cmpd 27”, “cmpd 30” and “cmpd 1” are the crystallographic ligands of 2NZT, 5HFU, 5HEX and 5HG1, respectively.

Ligand	5HG1				5HFU		5HEX		2NZT	
	Glu742		Asn683		Asp657		Asp657		Glu742	
	RMSD	Score	RMSD	Score	RMSD	Score	RMSD	Score	RMSD	Score
Glucose	1.84	51	1.90	50	1.8	37	4.09	37	0.34	63
“cmpd 27”	3.88	90	4.15	93	3.5	79	6.49	82	10.68	-1
“cmpd 30”	2.00	95	1.99	100	3.33	78	3.98	76	4.71	58
“cmpd 1”	0.45	81	0.29	83	1.47	61	7.4	60	11.41	56

Sequentially, the performance of systems with 5HG1 having Glu742 and Asn683 as centers (GOLD 5.2), flexibility of seven catalytic pocket residues (Asn656, Asp657, Lys621, Asn683, Glu742, Ser603, and Thr680), GOLD 5.2, Autodock4 and Autodock Vina software packages were evaluated. With this insight, molecular docking calculations for a set of 16 HK2 inhibitors [43,153] with known experimental data (IC_{50}), were carried out. The structures of the inhibitors were previously prepared and protonated, and the geometry was optimized (energy minimized). Correlations between experimental data ($\text{Log}[IC_{50}]$) and molecular docking score (or predicted binding energy – for Autodock Vina/Autodock4 results) were analyzed to evaluate the protocols' performance.

In fact, 5HG1 with either Glu742 or Asn683 as centers produced very close results in this evaluation as well (**Figure 21**). The correlations obtained had an $R^2 = 0.75$ for Glu742 and $R^2 = 0.71$ for Asn683.

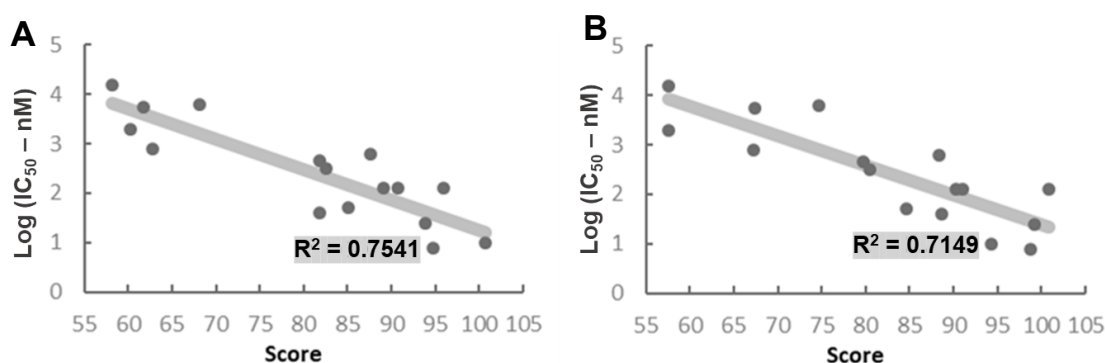


Figure 21. Correlations between experimental data ($\text{Log}[IC_{50}]$) and score (in silico results) of 16 known HK2 inhibitors – PDBID:5HG1, GOLD 5.2, Goldscore scoring function, center: Glu742 (A) and Asn683 (B).

Regarding the flexibility of the catalytic center residues, no improvement was found over previous results (**Figure 21**), neither with simultaneous flexibility assigned to seven residues (**Figure 22-A**) nor with individual assignment of flexibility (**Figure 22-B, C, D**). **Figure 21** shows the best correlations obtained with flexible ligands and none obtained $R^2 > 0.75$.

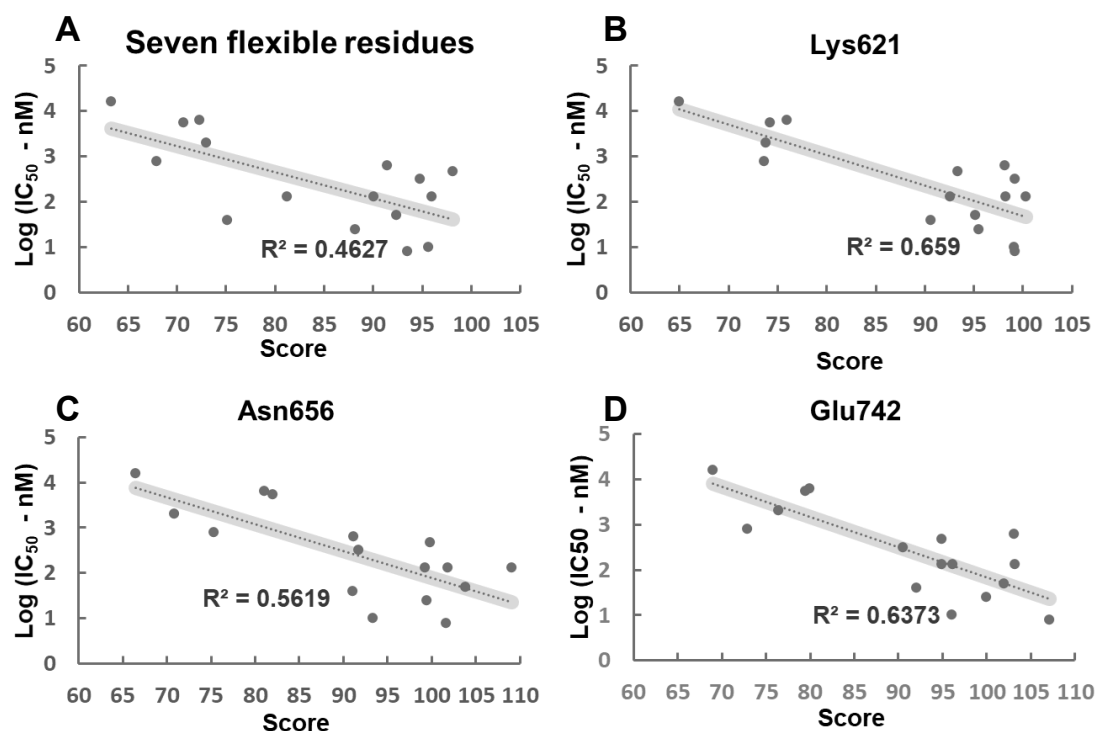


Figure 22. Correlations between experimental data ($\text{Log}[IC_{50}]$) and score (*in silico* results) of 16 known HK2 inhibitors – PDBID:5HG1, GOLD 5.2, Goldscore scoring function. **A** - Evaluation of simultaneous flexibility of seven residues; **B**, **C** and **D** - Best correlations obtained with one residue as a flexible element, respectively Lys621 (**B**), Asn656 (**C**) and Glu742 (**D**).

Autodock4 and Autodock Vina were not able to perform better than GOLD 5.2 with this system and the correlation between experimental and theoretical data was considerably low in both instances (**Figure 23**). RMSD of redocked 5HG1 ligand was also greater than the one obtained previously (Autodock4: RMSD = 1.78 Å; Autodock Vina RMSD = 1.79 Å). Nonetheless, a potentially better performance of these software packages cannot be discarded, should other conditions be tested.

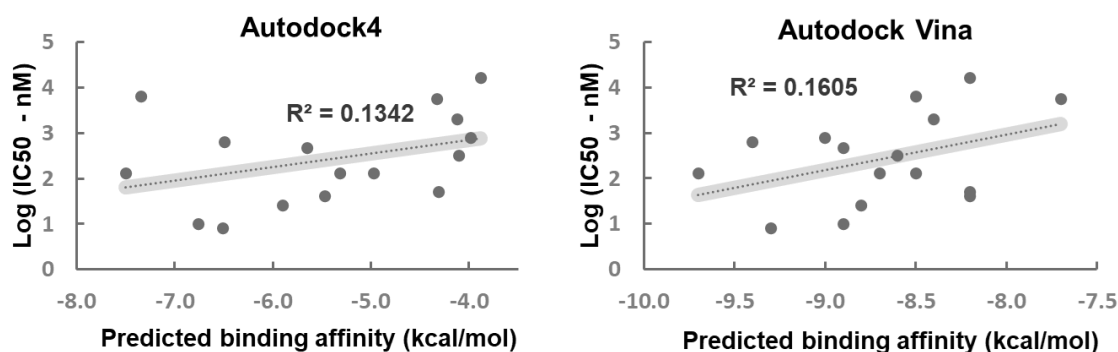


Figure 23. Correlations between experimental data ($\text{Log}[IC_{50}]$) and score (*in silico* results) of 16 known HK2 inhibitors - PDBID:5HG1, using Autodock4 (left) and Autodock Vina (right).

After all the information had been gathered, the best system and software were selected. The elected protocol was able to produce the results with lowest RMSDs between docked and crystallographic structures (**Figure 24 - A**) and the best correlation between experimental (IC_{50}) and *in silico* data (score) (**Figure 20 - B**). The GOLD 5.2 software seemed to offer the best results, in particular with the GoldScore scoring function.

In summary, the final protocol included: CTD of structure PDBID:5HG1 [43], GOLD 5.2 software, GoldScore scoring function, no crystallographic water molecules, no flexible residues, with a spherical docking region centered on Glu742 and with a radius of 15 Å. Besides results pointing to 5HG1, this structure has the particularity of representing an intermediate state of the pocket, which confers some advantages. Within the closed pocket of 2NZT, just a few small molecules will be able to fit, limiting dramatically the chemical space. Using largely open structures such as 5HEX, only bulky and large molecules will be able to interact (with good scores) with several residues to stabilize the complex. With this view, using 5HG1 may lead to a selection of molecules with a more appropriate size and a higher number of predicted interactions.

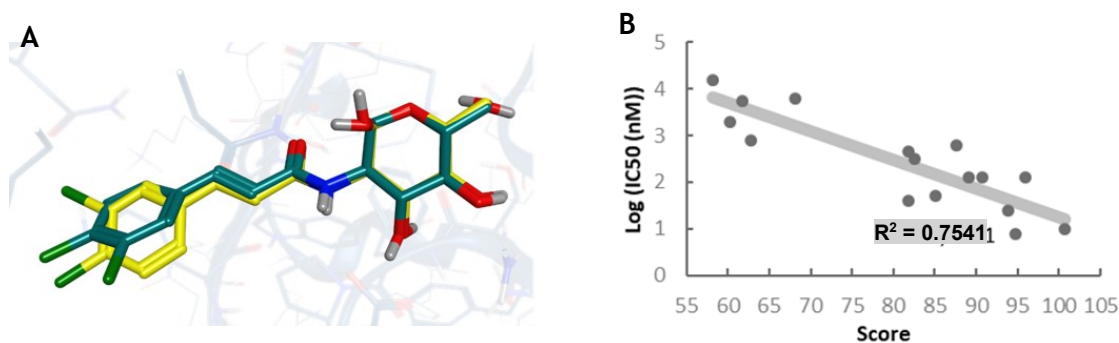


Figure 24. Molecular docking protocol validation: **A**- Overlap of the crystallographic ligand (PDBID:5HG1) [43] (yellow) with its docked conformation (green) obtained with GOLD 5.2 (RMSD= 0.45) - image obtained with MOE2016. **B** - Best correlations between experimental data ($\text{Log}[IC_{50}]$) and score (*in silico* results) of 16 known HK2 inhibitors – PDBID:5HG1, center: Glu742, $r=15$ Å, with the GoldScore scoring function of GOLD 5.2 software.

2.2. Selection of databases

The selected VS protocol was applied to 23 databases. Firstly, a part of the DrugBank library, specifically approved (DBA) and withdrawn (DBW) drugs sub-databases, was considered. Broadly, this set of molecules compiles approved, investigational, illicit and withdrawn drugs. Screening of DBA and DBW molecules allowed us to understand which chemical groups would most likely be relevant for potential interactions with HK2. This was a determinant factor to choose other databases. The results of DrugBank VS are schematically summarized in **Figure 25**. In total, 2117 molecules were docked to the catalytic pocket of HK2. Of these, 166 were scored higher or equal to 80 by the software. As cutoff, a score of 80 was chosen for the analysis of the docking-generated conformations (in line with the scores obtained with the known inhibitors). After all, 29 compounds were considered to fit properly in the catalytic pocket compared to the known inhibitor “cmpd30”, showing the ability to establish important interactions (**Figure 26**). Generally, the selected compounds have hydrophilic groups, with sugar moieties and diverse groups commonly present in natural

products (**Figure 25 - B**). Interestingly, the antibiotic group was the major class of drugs present in the final results, mostly cephalosporin drugs (7 out of 9 antibiotics). In this regard, besides relevant databases such as the NCI, several natural products (NP)-focused databases including NP derivatives were selected for screening.

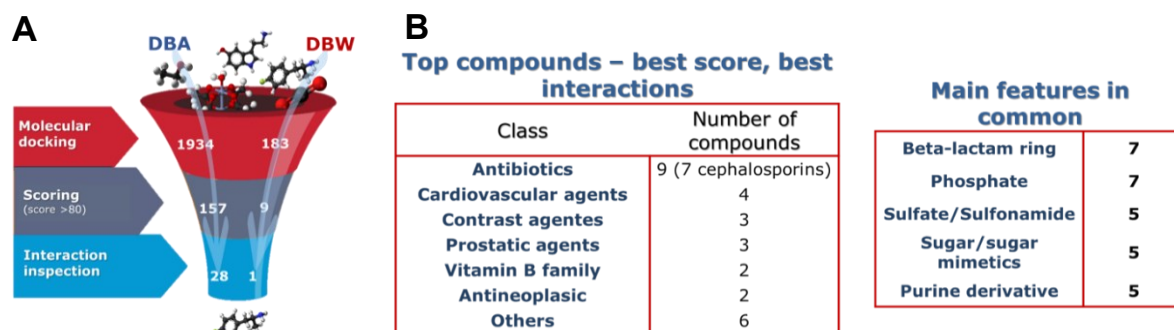


Figure 25. Results of Drugbank screening: **A** - from initial 2217 molecules (DBA+DBW), 166 were scored above 80 and 29 were initially selected after visual inspection. **B** - Main molecules found were cephalosporins and other natural products such as Vitamin B family members. Hydrophilic groups such as phosphates, sugars and sulfate/sulfonamides are the most common.

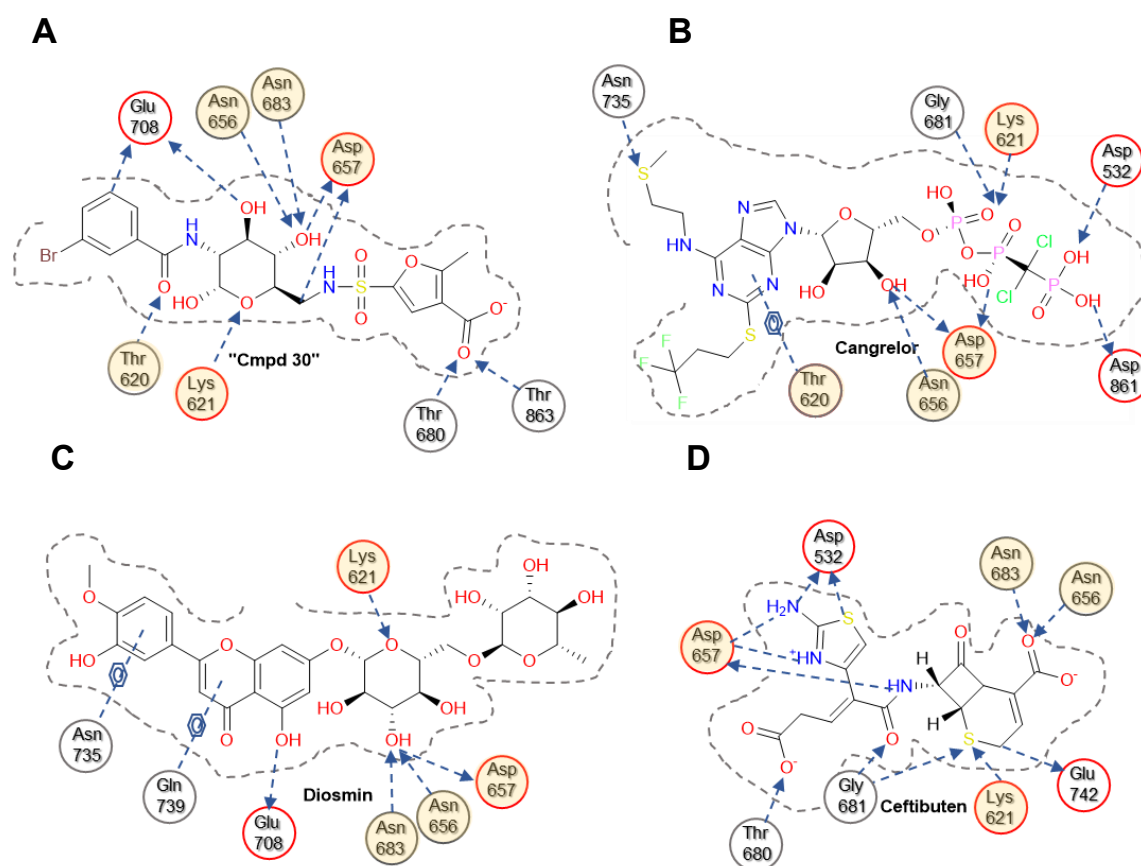


Figure 26. Interactions (2D) between HK2 and "cmpd 30" (**A**), best inhibitor crystallized (IC_{50} = 10 nM) [43], and three of the highest scored molecules from DrugBank VS. (**B** – cangrelor, score= 115, **C** - Diosmin, score=95, **D** - Cefibuten, score=82). Common residues interacting with the known inhibitor and potential inhibitors are highlighted in yellow. Arrows indicate H-bonds. Dotted lines with benzene an aryl ring in the middle represent H- π interactions.

The NCI database itself corresponds to a collection of uniquely diverse synthetic and natural molecules with over 250,000 entries that have been either submitted to the US National Cancer Institute's Developmental Therapeutics Program (DTP) for biological evaluation or, in some cases, synthesized under DTP support [216]. Moreover, these molecules can be obtained free of charge.

Other databases were selected from the ZINC15 repository [217] which contains hundreds of millions of commercially-available compounds subdivided in hundreds of databases, some of which fully dedicated to NPs or NP derivatives. From those, 20 databases were selected, namely INDOFINE, HMDB Plant, HMDB Microbe, HMDB Food, EFI-Phosphate Sugars, UEFS-NP, Herbal Ingredients Targets, Herbal Ingredients in-vivo Metabolism, NuBBE NP, IBScreen Bioactives, IBScreen NP, Specs NP, AfroDB, Keyorganic bioactives, NPACTD, NCGC Pharmaceutical Collection, TimeTec ND, MolMall, AnalytiCon Discovery ND, and Ambinter.

Additionally, the Mu.Ta.Lig Chemotheca database [218] (a repository of synthetic molecules populated by several European researchers, samples of molecules might be requested to the owner for testing) was also screened. The number of molecules *per* database is shown on **Table 5**.

Table 5. List of databases used on SBVS together with the number of compounds per database, number of molecules scored above 80 and molecules selected after visual inspection.

Database	Compounds	Score > 80	Selected	Database	Compounds	Score > 80	Selected
Indofine	64	10	4	Keyorganic bioactives	1149	85	10
HMDB Plant	149	12	1	Herbal Ingredients <i>in-vivo</i> Metabolism	1270	51	6
HMDB Microbe	171	6	1	NPACTD	1574	150	9
EFI-Phosphate Sugars	199	7	7	Drugbank	2117	166	13
UEFS-NP	498	4	0	NCGC Pharmaceutical Collection	2994	153	18
Herbal Ingredients Targets	531	35	9	TimeTec ND	3040	154	33
NuBBE NP	643	31	2	MolMall	15723		30
IBScreen Bioactives	760	53	3	AnalytiCon Discovery ND	24800	984	300
Mu.Ta.Lig Chemotheca	833	40	7	HMDB Food	32518	6995	487
Specs NP	880	30	7	IBScreen NP	56799	3428	128
Afro DB	954	41	3	NCI	265242	6858	519
				Ambinter	10 582 481	10319	1532

2.3. Structure-based virtual screening (SBVS)

The same protocol used for DrugBank VS was used for this large multiple databases screening, with exception of the Ambinter database which required an additional step. Prior to the general procedure, some filters were applied to the Ambinter database (over 10 million compounds), specifically through the elaboration of simple pharmacophore models (with MOE 2016) based on known inhibitors. Four features out of six (**Figure 27**) were selected to be satisfied and Ambinter was screened, resulting in 175137 compounds (1.6%) that proceeded to the molecular docking VS protocol used before. These features are related with key expected hydrogen interactions. One should notice that this technique was used as a way to simplify the screening process, and that the pharmacophore was developed to be used as a general filter.

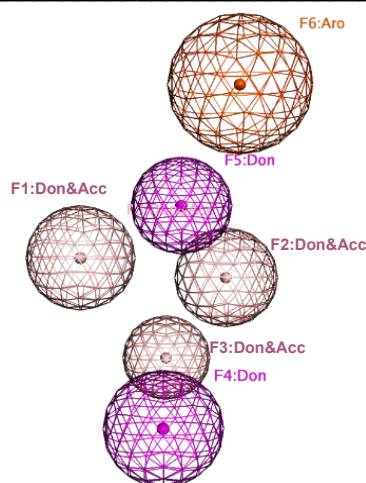


Figure 27. Pharmacophore used for pre-screening of the Ambinter database. Six features were defined. For the pre-screening, molecules must satisfy 4 out of the 6 conditions. Aro – aromatic moiety; Don – H-bond donor; Don&Acc – H-bond donor and acceptor.

According to expected interactions with selected compounds and interactions with known inhibitors, several residues described as essential for inhibition establish common interactions with the (potential) ligands. Hydrogen bonds are privileged, particularly in the center of the catalytic pocket (glucose binding site), and hydrophobic interactions are more likely to appear on the sides, as illustrated in the two-dimensional figure of interactions between the crystallographic ligand “cmpd 30” [43] and residues of HK2 (**Figure 26 - A**).

As indicated earlier, after the molecular docking calculations were completed, conformations scoring 80 or higher were visually inspected. The final results showing the list of all databases screened, the number of initial compounds, the number of compounds scoring 80 or higher, and the number of molecules selected can be found in **Table 5**. To summarize, 2981 molecules were considered as potential hit compounds for HK2 inhibition. Molecules were selected based on interactions previously observed with known inhibitors and those predicted for DrugBank molecules. Molecules that did not fit into the place occupied by glucose or G6P were discarded, as well as long aliphatic chains with high degree of flexibility, among others that did not fit properly.

2.4. Prediction of ADMET properties and compounds selection

Subsequently to the VS campaign, the selected molecules were filtered based on their calculated physicochemical and ADMET properties. These properties were predicted with FAF-Drugs4 [204] and OCHEM [219] online servers. All properties were carefully analyzed and compared.

Cautious analysis of the predicted ADMET properties must be accomplished. Despite the importance of these filters, they cannot be used blindly. For instance, certain structural alerts considered PAINS are present in many marketed drugs (87 small molecules of FDA-approved in 2017) [220] and even frequent hitters can have their properties improved in the hit-to-lead process [155]. Moreover, oral administration is the route that dictates the course of calculation for most properties. Still, these predicted data contributed to prioritize the molecules to be tested experimentally. With the appropriate score/pose and ADMET properties, 454 molecules were considered for the next step. Reselection occurred to reduce the number of potential molecules to acquire, where similarity vs diversity was also the key to this selection, in addition to further improvement of pharmacokinetic properties. Ultimately, a final set of 111 molecules was reached, corresponding to 0.37% of the total number of molecules scoring higher than 80 by the molecular docking program (**Figure 28**).

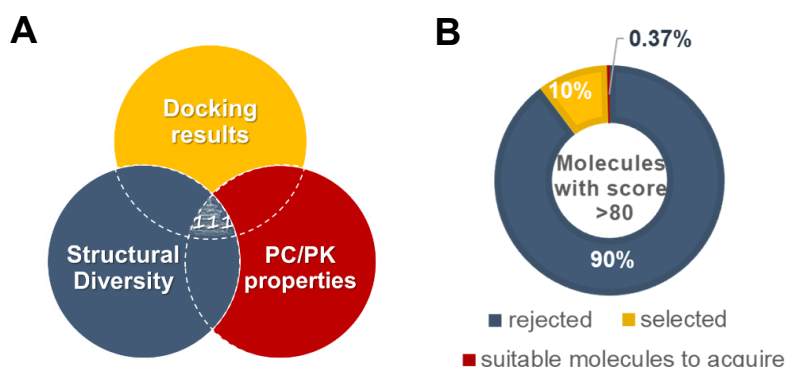


Figure 28. **A** - Criteria involved in the selection of the final set of 111 molecules. **B** - percentage of the molecules that were rejected (90%), selected by the user (10%) and finally considered suitable for experimental testing (0.37%), from all molecules scored higher than 80 by the molecular docking program. PC: Physicochemical. PK: Pharmacokinetics.

Figure 29 summarizes and compares some important descriptors/predicted properties of the initial set of 2981 molecules and the final set of 111 molecules, namely molecular weight (MW), topological polar surface area (TPSA), logP, solubility in water, oral bioavailability, number of Lipinski violations, presence of PAINS alerts, and the final FAF-Drug4 result. Other descriptors were predicted, such as the ability to induce phospholipidosis, and rules such as GSK 4/400 [221] and Pfizer 3/75 [222]. However, these were not used for selection, since the set of 2981 molecules was shown to have residual numbers of molecules that were expected to fail. At first sight, the properties of the two sets (selected and re-selected sets) of molecules do

not appear to be markedly different. However, one should notice that the first set of molecules was chosen manually, and a large amount of obviously unwanted molecules were discarded.

The MW of the selected molecules was substantially high, with 99% of the molecules having MWs higher than 300 g/mol and 66% in the 400-500 g/mol range. After re-selection, 73% of the chosen molecules have MWs of 400-500 g/mol and none have MWs lower than 300 g/mol.

TPSA is one of the most determining descriptors. As anticipated, the selected molecules have a large TPSA, as required by the hydrophilicity of the residues that establish the catalytic pocket. A total of 33% of molecules coming from docking have a TPSA greater than 180 Å², which is not advisable [205]. The 111 selected molecules still present high TPSA, but only 18% have a TPSA higher than 180 Å², while 80% have a TPSA between 100 and 180 Å². This is still not desirable if a molecule is aimed to target the central nervous system, but was considered acceptable for the purpose of this work. As expected, considering the high TPSA, of the selected molecules, the predicted logP is generally low; for 99% of the molecules in the original 2981 set and 97% of the molecules in the 111 molecules set logP (at pH 7.4) is predicted to be equal or lower than 5, as widely recommended [223]. In the end, 60% of the selected molecules are foreseen to have a logP between 1 and 3. Correlated with MW, logP and the number of H-bond donors and acceptors (closely related with TPSA), the number of Lipinski's rule violations is an important descriptor. It was postulated that, to achieve adequate absorption and permeation, molecules should not violate more than one of the components of Lipinski's rule of 5. The two sets of molecules mostly violate zero or one of the components of this rule and also satisfying Veber's rule, thus most of molecules are predicted to be orally bioavailable. Proper solubility is one of the key physicochemical properties for the success of drugs. Fortunately, in both sets of molecules good water solubility was predicted. PAINS filters only retrieved 10% of all molecules in the 2981 set, and only one molecule contained a PAINS substructure in the final set of selected molecules. A great improvement was achieved regarding the acceptance by the FAF-Drugs4 algorithm that includes toxicity potential. While more than 50% of the molecules in the 2981 set were rejected, the rejection rate was only 17% in the final set of 111 molecules. Nonetheless, improving all the relevant descriptors and properties revealed to be a challenge and was not always possible.

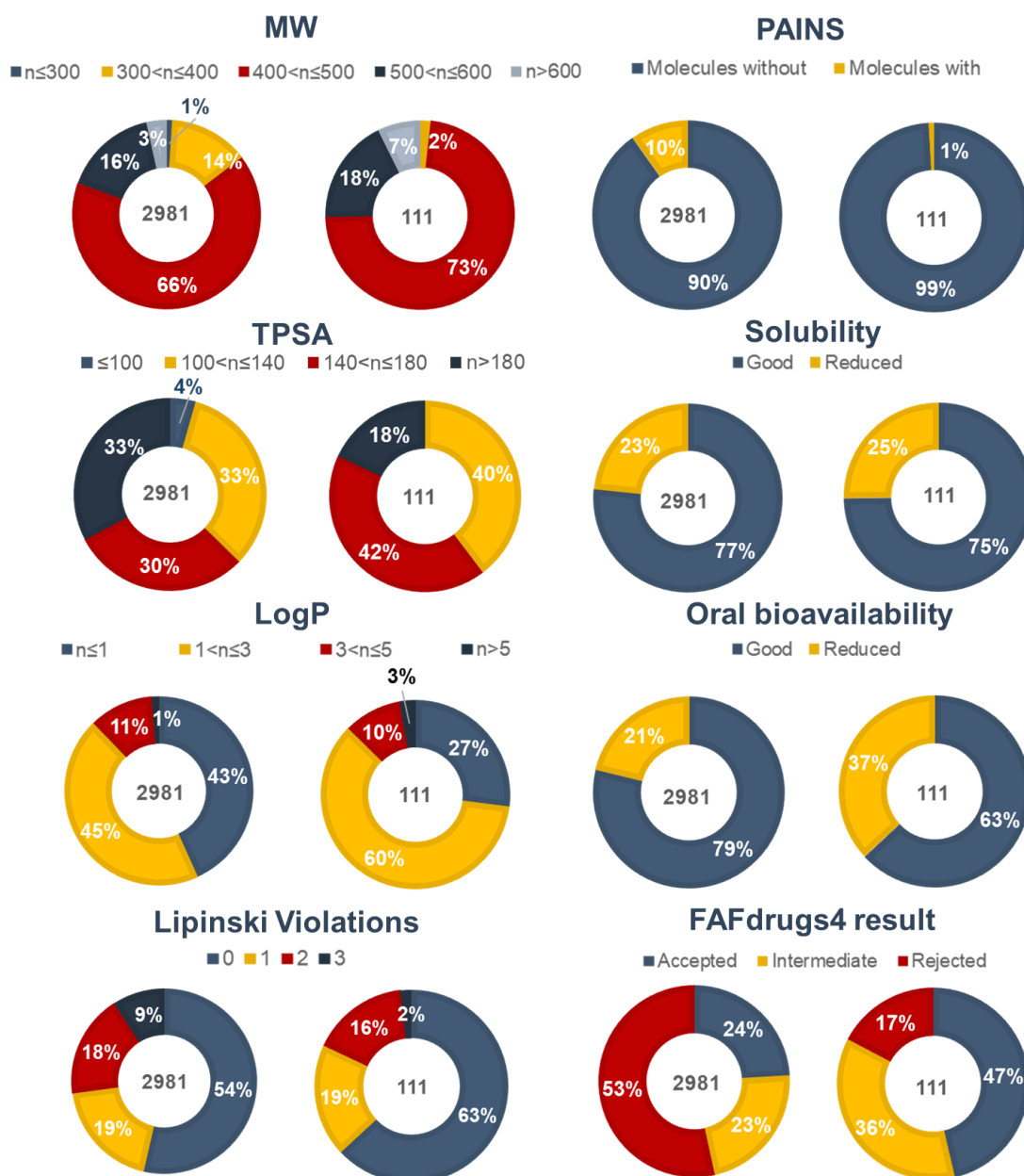


Figure 29. Descriptors and predicted properties of the set of molecules coming from molecular docking (left) and the final set of molecules (right). Predictions and calculations were performed using the FAFdrugs4 platform [204].

Other properties, such as genotoxicity (prediction of Ames test results) and ability to modulate CYPs were also considered relevant (**Figure 30**). Most compounds of the 111 set were predicted to be inactive in the Ames test, thus having a low potential to cause unwanted mutations. Moreover, only 23% of molecules were anticipated to inhibit CYPs. Finally, prediction of solubility in DMSO was important for selection between molecules with similar core structures, although only 11% were predicted not to be soluble in this solvent (**Figure 30**).

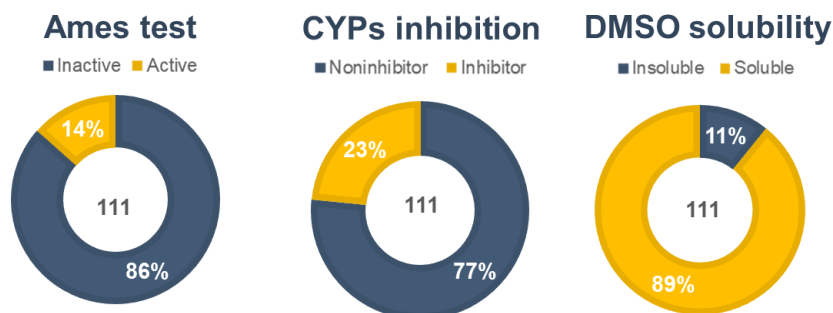


Figure 30. Results of prediction of Ames test, potential CYPs inhibition and DMSO solubility using the OCHEM server [219], for the final set of 111 molecules.

In the selection process, availability and price of the molecules were taken into account. Although ZINC15 is considered a database of commercially available molecules, there are still molecules that cannot be purchased, for different reasons. Furthermore, several compounds present in databases, such as those from the NCI and the Mu.Ta.Lig Chemotheca, were no longer available, which had implications for obtaining the final set of potential HK2 inhibitors. Thus, commercial reasons entailed the acquisition of molecules from both the 2981 and 111 sets of molecules. In the end, 64 molecules were acquired, some from “free-of-charge” sources (National Cancer Institute - NCI, Mu.Ta.Lig Chemotheca) and others purchased from MolPort or Sigma-Aldrich, to be tested experimentally and thereby validate the SBVS protocol. The structures of the acquired molecules, along with their molecular docking score, are in **annex 2**.

2.5. Confirmation of the identity of acquired compounds

Before use for experimental assays, the identity of the 64 acquired molecules plus the glycolytic inhibitors 3BP, Isonitazene and 2DG was tested by LC-MS, or NMR, if the former could not be accomplished due to low ionization potential of the sample. All purchased compounds had spectral data consistent with the expected structures (data not shown). However, several compounds from the NCI database were considered to have purity and/or identity issues. This possibility is disclosed by the NCI itself, admitting that the accuracy of structure and purity of the sample were not confirmed [216], by which its evaluation is recommended.

In total, around 10 molecules were considered to have significant problems regarding their identity. Despite these potential problems, initially those molecules were tested against recombinant human HK2 (rhHK2); however, these issues were considered when deciding upon for further testing. Among molecules with similar structures, the ones with structural issues did not proceed. In the absence of molecules of the same chemical class, molecules with certain structural uncertainties did, nonetheless, proceed for testing, to assess whether or not they displayed promising HK2 inhibition properties.

Figure 31 – **A** and **B** shows representative ESI-MS spectra of molecules that were in agreement with the expected structures, while **Figure 31 - C** and **D** shows representative ESI-MS spectra of samples that do not correspond with the expected structures. The last group was not considered for further testing (after initial biochemical validation).

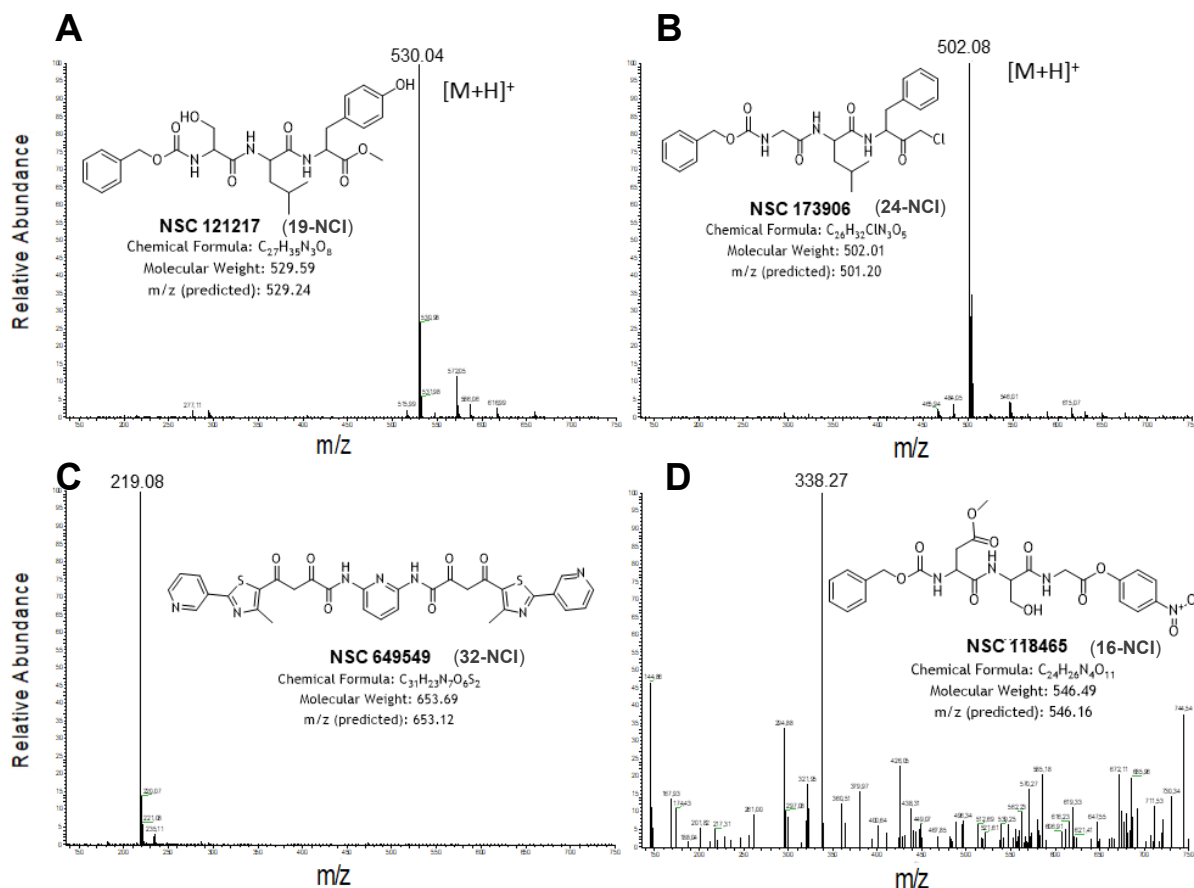


Figure 31. Positive ESI-MS spectra [MW: 100-1000] of four representative samples obtained from the NCI (NSC XXXXX code is the code determined by the NCI, X-NCI is the code given by the author after acquisition). Spectra **A** and **B** are examples of spectra in agreement with the expected structures in terms of identity and purity. Spectra **C** and **D** are examples of spectra that were not in agreement with the expected structures in terms of identity and purity.

Figure 32 shows the ESI-MS spectra of sample NSC 31345 (expected MW=585.60, code: **3-NCI**, annex 2). Apparently, this sample was composed of different compounds. The presence of the expected molecule was confirmed with the positive ESI-MS spectrum; however another non-identified component, more abundant or more easily ionizable, was also found at m/z 474.05, in addition to other impurities (most relevant at m/z 394.14). In the negative ESI-MS spectrum the expected $[M-H]^-$ signal appeared in low abundance compared with other signals. Anions at m/z 471.87 and 392.10, consistent with the cations found in positive ESI mode, were also present but their identification was not accomplished. The presence of other highly abundant anions at m/z 235.69 and 195.83 led to conclusion that these were likely dianions from the 471.85 and 392.10 monoanions, respectively, which suggests a favorable loss of two protons in both instances, possibly from two sulfonic acid substituents. Nonetheless, the difference between 474.05/471.87 and 394.14/392.10 is consistent with the presence of two components differing by one sulfonic acid group, which appears to

contradict the conclusion above. Despite having considered possible impurities from the putative synthesis process of the compound, as well as plausible degradation products, we were unable to identify the sample contaminants.

Despite the small amount available, this sample was also analyzed by NMR. The ^1H NMR spectrum is shown on **Figure 33**. Although some impurities seem to be present, the integration and signals are consistent with the expected structure. Although a good ^{13}C NMR spectrum could not be obtained due to the small amount of compound available, HSQC also indicated the expected correlation between protons and carbons.

Taken together, assessment of NSC 31345 purity issues was unclear and further studies would be necessary. Still, since the presence of the expected compound was demonstrated, this sample was evaluated in biochemical and biological assays.

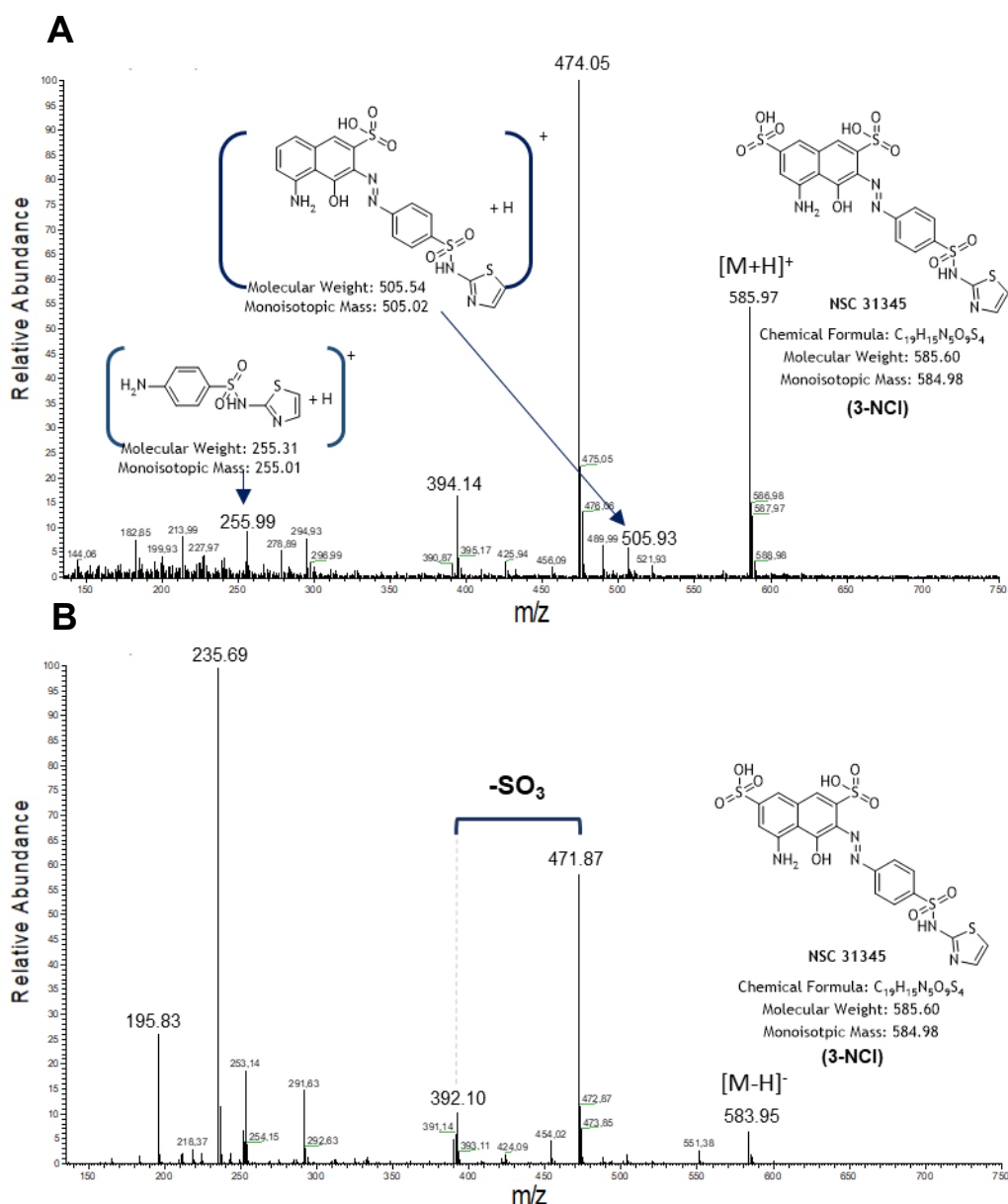


Figure 32. Positive (A) and negative (B) ESI–MS spectra [MW: 100–1000] of NSC 31345.

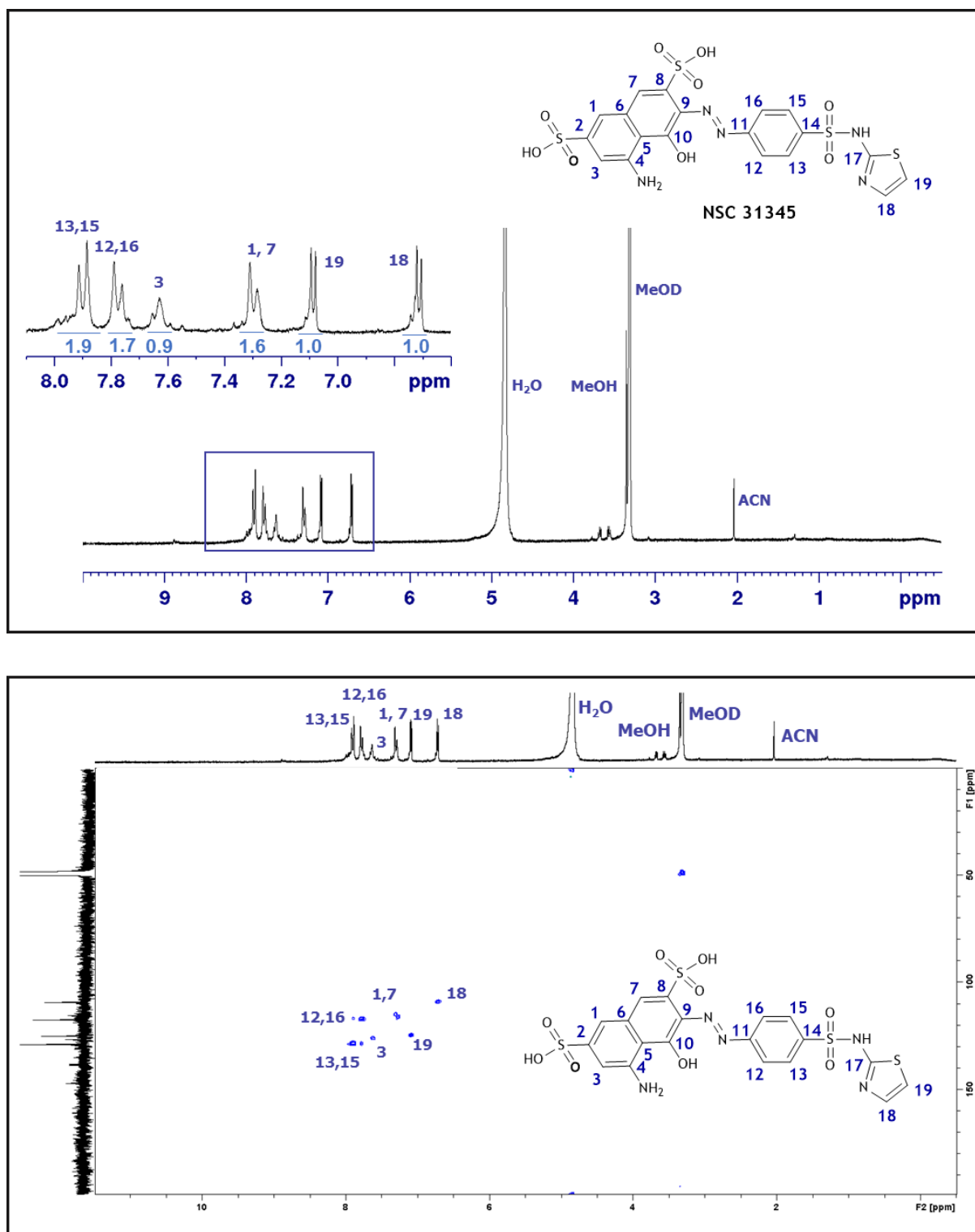


Figure 33. ^1H NMR (A) and ^1H - ^{13}C HSQC (B) spectra (300 MHz, 75.5 MHz, MeOD) of sample of NSC 31345.

2.6. SBVS validation – Biochemical assays

In order to validate the SBVS protocol and experimentally find hit compounds able to inhibit HK2, two biochemical assays were implemented. These assays differ mainly in the ATP/ADP detection method, by HPLC coupled with UV detector (HPLC-UV) or by luminescence. Both experiments started with the incubation of the test compounds with rhHK2 in the same conditions.

2.6.1. ATP/ADP detection by HPLC

Despite the similarity between of ATP and ADP, which present exactly the same chromophore (**Figure 34 - A**), the direct separation and quantification of ATP/ADP using the HPLC technique was expected to be simple, quick and associated with low cost. Based on the work of Santos *et al.* (2019) [211], the design of the best HPLC protocol for accurate separation of ADP from ATP was achieved after several steps of optimization, including the mobile phase, flow rate, elution mode, and run time. Standard curves were successfully obtained (calibration curves are presented in **Figure 34**). However, after numerous tries with different concentrations of enzyme, substrate, co-factor and testing compounds, the sensitivity of the method regarding the small differences between the ATP disappearance and ADP formation during an enzymatic reaction were considered insufficient, contrary to what was reported previously by Santos *et al.* (2019) [211]. Using larger amounts of enzyme might be useful in this approach; however, the high costs would not be manageable at this stage. Thus, evaluation of rhHK2 activity was only performed using the luminescence-based method described below.

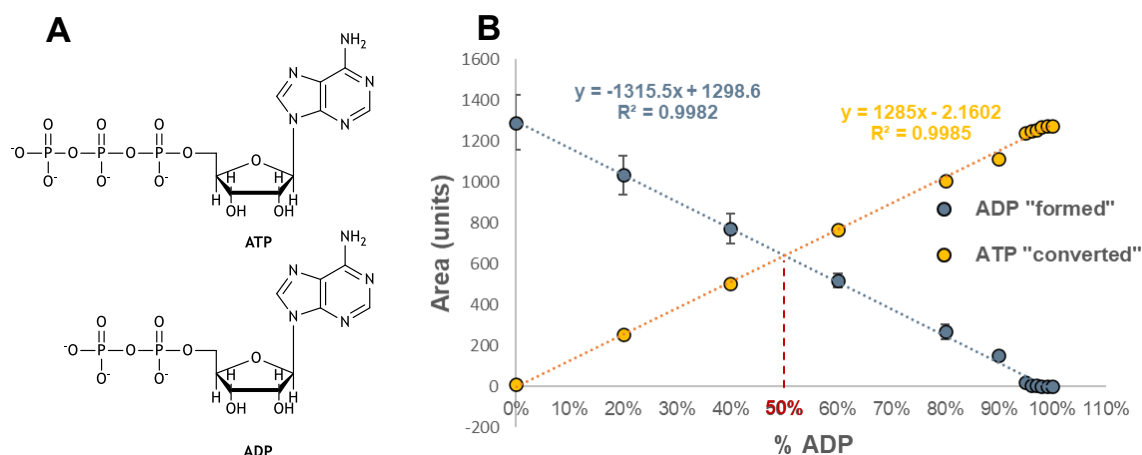


Figure 34. **A** - Structures of ATP (above) and ADP (below). **B** - Standard curves for ADP "formation" and ATP "conversion". The graphic shows that the amount of ATP/ADP changes proportionally. Mobile phase: 0.1 M phosphate buffer at 1.2 mL/min, 20 μ L injection, isocratic mode for a 15 min run, r.t.. UV detection at 254 nM, C₁₈ column (ThermoFisher Scientific BDS Hypersil C18 (Ø4.6 x 250 mm, 5 μ m) reversed phase column). Error bars correspond to SD.

2.6.2. ADP detection by Luminescence - ADP-Glo™ Kinase Assay

Based on the work of Lin *et al.* [43], the protocol for evaluation of rhHK2 activity towards the panel of potential inhibitors was optimized to guarantee the best parameters such as concentrations of substrates and enzyme, plate and the luminescence detector. The final protocol achieved a signal-to-background (S:B) ratio of 15% and the conversion standard curve R^2 was 0.98 (**Figure 35**).

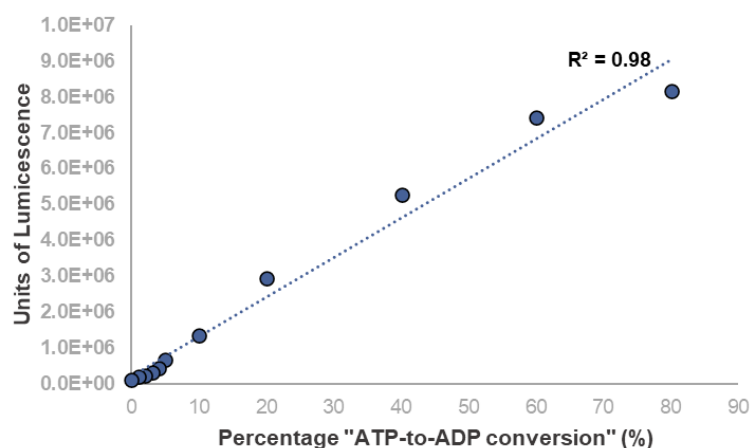


Figure 35. Standard curve for "ATP-to-ADP conversion".

Upon optimization, the previously selected compounds were tested at 10 μ M, along with three known inhibitors, 3BP, 2DG and LND. The results showed that all the 64 molecules inhibit HK2 and only 3 inhibited it by less than 14%. From those 64, 22 molecules presented higher or comparable effect (inhibition $\geq 43\%$) to 3BP. Moreover, one molecule was able to inhibit HK2 by more than 60% (**Figure 36**)². Potential interference of the compounds in the assay was evaluated and none of the molecules substantially affected the output (data not shown). Considering that the 10 μ M concentration is typically used for screening, these 22 molecules were considered hit compounds whose structure could be optimized to improve the activity. Regarding the specific structures, this set is highly diverse (see annex 2). IC_{50} values were not determined since the molecules did not inhibit HK2 much beyond 50%.

² The percentages of rhHK2 inhibition for each molecule are found in **annex 2**.

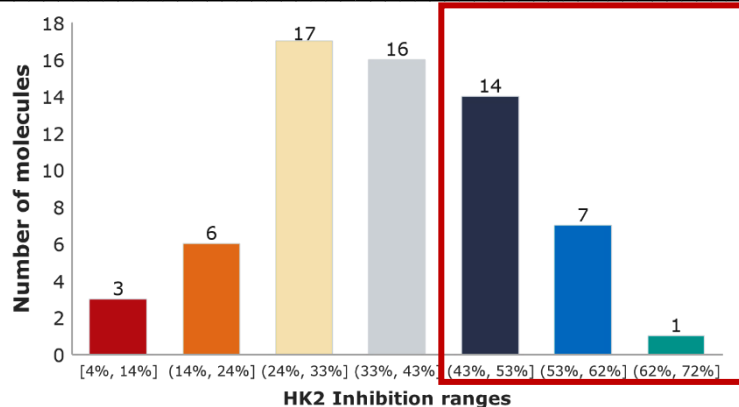


Figure 36. Distribution of rhHK2 inhibition results per ranges of inhibition ability. The number of molecules with higher or comparable effect to 3BP is highlighted. Average % inhibition for each compound resulted from triplicates of at least 2 independent assays.

In order to evaluate the ability of molecular docking to score molecules according to their actual activity, the docking scores of the tested molecules were plotted against their HK2 inhibition ability at 10 μ M (**Figure 37**). A direct correlation cannot be established, and the results are mostly spread along all score ranges in a non-linear fashion. Still, two of the most potent molecules in the biochemical assay were identified by SBVS in the top 5 of best scored molecules, **36-NCI**³ (score:101; rhHK2 inhibition: 66%) and **5-chem**² (score:98; rhHK2 inhibition: 56%). Visual inspection of the molecular docking-generated poses is a determinant step to the success of the docking protocol. For instance, one of the molecules with highest activity against HK2 had a score close to 80 (cut-off score for visual analysis).

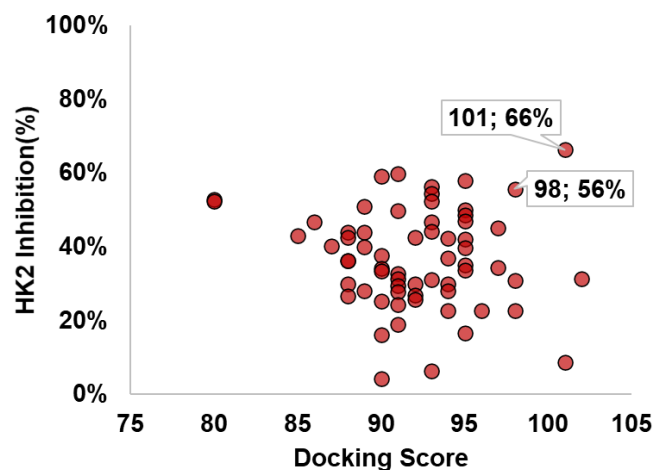


Figure 37. Correlation between experimental data (% HK2 inhibition) and docking score (in silico results) for the 64 molecules tested.

³ See annex 2 for the structure.

2.6.3. Evaluation of ruthenium complexes as HK2 inhibitors

Metal complexes have shown potential in different fields of medicinal chemistry due to their radioactivity, cytotoxicity or, for instance, their photophysical features, and may work as therapeutic or diagnostic agents [224]. The antiproliferative effects of ruthenium complexes have long been reported and some drug candidates have entered clinical trials for cancer treatment, as an attractive alternative to platinum drugs such as cisplatin [225]. The main mechanism of action for many of these complexes is supposed to be related with DNA intercalation; however, multiple effects in other molecular targets have been reported, depending on the coordinated ligands. Remarkably, coordination with existing drugs has shown several times to have synergic effects. This includes the coordination with LND (**V**, **Figure 4**) [225], a glycolytic inhibitor suspected to inhibit HK2 (*cf chapter II, section 3*).

Incorporation of carbohydrate moieties in ruthenium complexes has been carried out. Mostly, these ruthenium complexes have been synthesized in order to obtain catalysts for enantioselective synthesis [226,227]; however, relevant antiproliferative effects have also been reported in some instances [228,229]. Interestingly, some carbohydrate ligands are used as vectors to increase internalization of complexes into the cell, *via* GLUT [228]. Besides the evaluation of cytotoxic activity of these specific complexes, no further assessment of effects on the glycolytic pathway have been addressed. Other organometallic complexes, such as the ones with technetium-99 or rhenium coordinated with glucose, were found to be substrates or to (modestly) inhibit yeast HK [224,230].

In this regard, in collaboration with Doctor Pedro Florindo (iMed.Ulisboa), who kindly supplied two sets of ruthenium complexes with carbohydrate-based ligands synthesized by his group (**RuGly3-10** and **RuGly21-26**, **Figure 38**), these molecules were evaluated as HK2 inhibitors using the same ADPGlo™ method described before. The results are shown in **Figure 39**. The effect of ruthenium complexes was compared to that of the carbohydrate-based ligands and an aglycon (ruthenium complex without sugar moiety), as well as with the known inhibitor 3BP. Ruthenium complexes **RuGly3-10** and the aglycon **TM34** behaved similarly and there was not a great difference between the inhibition ability of ligands and complexes, precluding any specific conclusions. Still, all compounds affected HK2 activity in the same range or to a larger extent than the inhibitor **3BP**. The second set of molecules, **RuGly21-26** and respective ligands behaved differently. Here, only **RuGly21**, **RuGly22** and **RuGly24** inhibited HK2 significantly. More prominently, **RuGly21** was the complex with the highest ability to inhibit HK2 activity comparing with control. Conversely to the previous subset, **RuGly21** and **RuGly22** are then expected to be able to stabilize the protein in a non-catalytic friendly conformation, while the ligand itself could not. Curiously, both complexes are glucosamine derivatives unlike the remaining elements of this subset. Besides the differences related to the structure of the sugar moiety, the other ligands complexed with ruthenium are substantially dissimilar between the two subsets, which presumably explains the differences in the ability to inhibit HK2. It should be noted that both the complexes and the ligands were shown not to interfere with the assay.

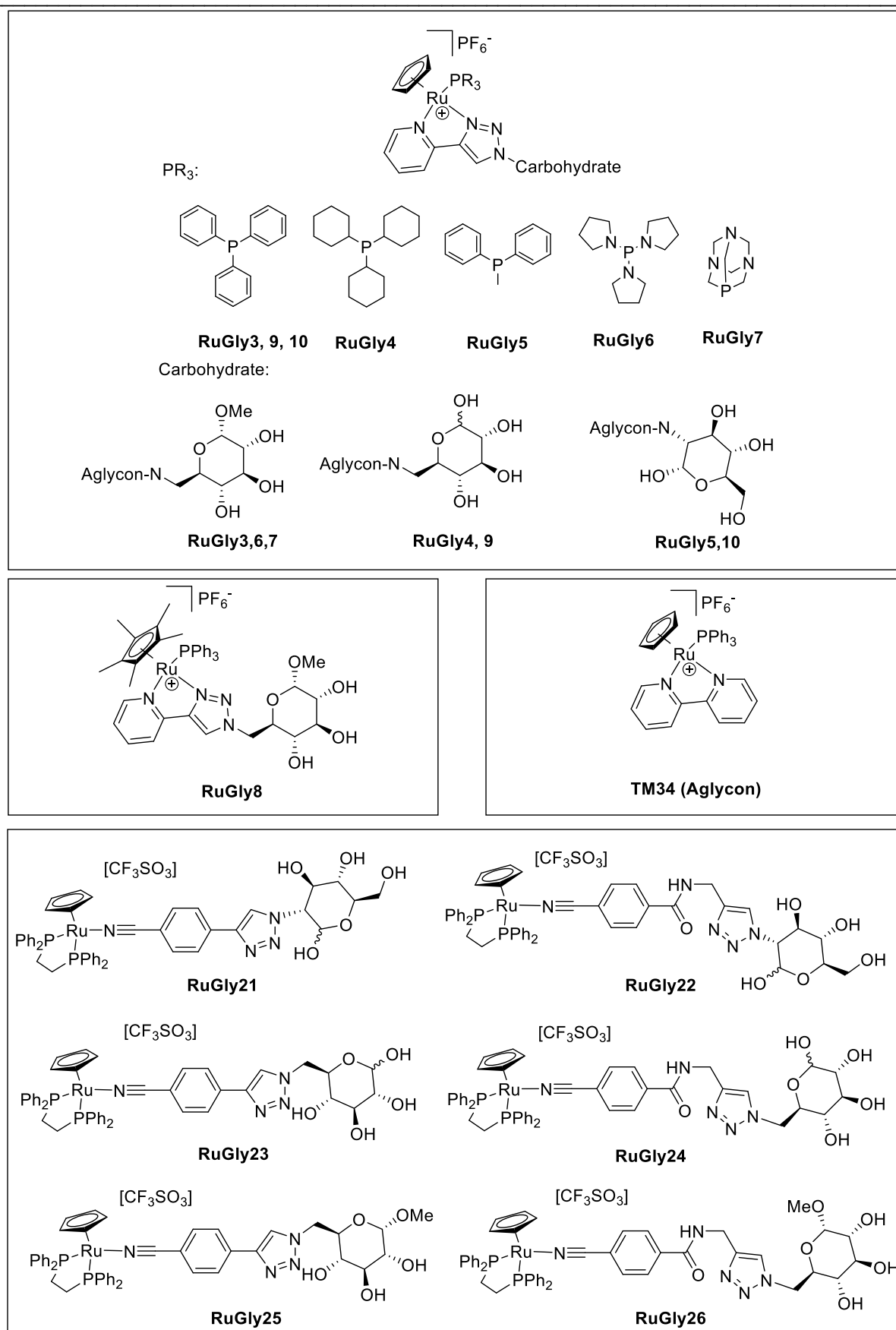


Figure 38. Structures of the ruthenium complexes evaluated against rhHK2.

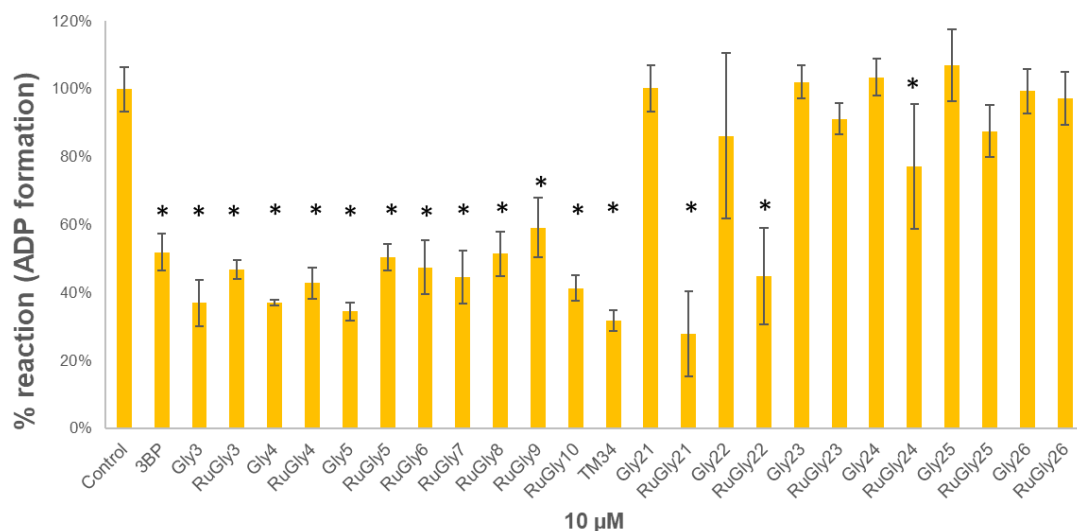


Figure 39. Reaction (%) (glucose + ATP → G6P + ADP) catalyzed by rhHK2 that occurred in the absence (control) or presence of either 3BP, ruthenium complexes, or carbohydrate-based ligands at 10 μ M (results as average \pm SD; * $p < 0.05$).

3. Conclusions

An SBVS protocol was implemented in order to identify potential HK2 inhibitors. This protocol was validated in experimental assays and proved to be successful in identifying new HK2 inhibitors with significant structural diversity. Using molecular docking, 2981 molecules were selected as potential HK2 inhibitors; prediction of ADMET properties reduced the number of potential HK2 inhibitors to 454, and further refinement reduced it to 111. The final acquisition of 64 molecules (appendix 2) included commercially available compounds, preferably from the 111 set, but also from the 2981 set.

Preliminary experimental results showed that all 64 molecules tested affected HK2 activity to some extent. Twenty-two molecules inhibited the HK2 activity more prominently or in the same range of the known inhibitor 3BP (HK2 inhibitor available when the assay was carried out). Thus, experimental data support the predictions obtained from the SBVS procedure, which is able to recognize bioactive molecules, ready for structural optimization and/or further testing. Still, the docking score did not correlate perfectly with the biochemical output. Since known specific inhibitors were not tested experimentally as positive controls for the inhibition, the ultimate significance of these results might be compromised.

Direct measurement of ADP with HPLC-UV to evaluate the activity of rhHK2 was projected; however, issues with sensitivity precluded the use of this technique. To avoid coupled reactions with the luciferase system or G6PD (*cf* **chapter V, section 2.5**), isothermal titration calorimetry (ITC) for a direct measure of the glucose phosphorylation reaction might be a solution for further studies. ITC is a very sensitive method to determine kinetic and thermodynamic parameters of enzymes of biological interest, based on the measurement of heat, absorbed or released. Specifically, ITC has been used to study the catalytic behavior of yeast HK 1 and 2 [231]. Regrettably, time constraints prevented any attempts with this approach.

A rational design strictly directed towards perfect fit could easily have compromised the approach here explored. Visual inspection of poses generated by molecular docking and prediction of ADMET properties are important steps that should not be ignored. As millions of molecules were screened, thousands of molecules were highly scored by the software. Not unexpectedly, most of them had a large number of H-bond donors and acceptors and predicted low permeability, with noticeable poor druglikeness.

The selected molecules belong to a myriad of structural classes, which allows us to cover a large chemical space, enhancing the probability of catching a hit molecule.

Ruthenium complexes have arisen as a class of potential drugs that have quickly developed in recent years [225]. Here several complexes were tested as HK2 inhibitors, revealing surprising results. Depending on the ligands coordinated with ruthenium, these organometallic complexes emerged as a new class of potential HK2 inhibitors. Still, more studies are necessary to draw conclusions in this regard.

4. Experimental Section

4.1. Design of Molecular Docking Protocol

Three-dimensional structures of HK2 PDBID: 2NZT [73], 5HFU [43], 5HG1 [43], and 5HEX [43] were downloaded from the Protein Data Bank (PDB). The procedure started with preparation and protonation at pH 7.4 and calculation of partial charges of ligands and receptor atoms from the crystallographic structures, individually, using MOE2016. Structure errors and clashes were examined and corrected using the same software. Only chain A was used in all crystallographic structures, specifically the CTD. Self- and cross-docking were performed with crystallographic ligands in all structures with GOLD 5.2 software using ChemPLP and GoldScore scoring functions. Parallel calculations were performed with eight different residues as catalytic centers: Asp657, Asn656, Gly862, Glu742, Glu708, Asp532, Asn683, and Thr680. Also in parallel, different radius sizes of 10 and 15 Å were tested for the spherical conformational search region. Self-docking calculations were performed in the presence or in absence of catalytic water molecules. After selecting 5HG1 as the protein structure and a set of HK2 inhibitors with known experimental data (IC_{50}) [43,153], flexibility of seven residues within the catalytic pocket (Ser603, Lys621, Asn656, Asp657, Thr680, Asn683, and Glu742) was tested, together and individually. Moreover, using the structure 5HG1, the performance of Autodock4 and Autodock Vina software packages was also evaluated.

The main non-standard parameters used to run GOLD 5.2 were: 500 runs, GoldScore or ChemPLP scoring functions, center on each of the residues referred above.

The main non-standard parameters used to run Autodock4 were: Lamarckian algorithm, 500 runs, grid size: x = 100; y = 120; z = 100; grid center: x = 82.048; y = 18.960; z = -101.439; spacing = 0.2 Å.

The main non-standard parameters used to run Autodock Vina were: box size: x = 65; y = 65; z = 65; box center: x = 82.048; y = 18.960; z = -101.439; exhaustiveness of the global search = 100.

Final protocol details: Structure: 5HG1; Software: GOLD 5.2; scoring function: GoldScore; radius: 15 Å; docking center: Glu742 (x = 108.32; y = 82.048; z = -101.439). No flexible residues, no water molecules, no receptor energy minimization.

The top three conformations were analyzed for each experiment. All molecules were prepared using MOE2016 [232]; interconversion of formats and calculation of RMSD were done using OpenBabel 2.4.0 [233] software.

4.2. Virtual Screening

After selecting the molecular docking protocol using GOLD 5.2, the databases were downloaded and prepared for VS. DBA and DBW were downloaded from Drugbank 5.09 [234] (<https://go.drugbank.com/releases/5-0-0>). The Mu.Ta.lig. Chemotheca was downloaded from the dedicated website [218] (<http://chemotheca.unicz.it/index.php>) and the remaining databases (see **Table 5** for details) were downloaded from ZINC15 [217] (<https://zinc12.docking.org/>). First, all libraries were curated by removing inorganic molecules, polymers or molecules with MW >750 and <150, mixtures, compounds with heavy metals, duplicates and empty entries. Thereafter, molecules were washed at pH 7.4 and their energy was minimized. This step was performed using MOE2016.

The Ambinter database (>10 000 000 molecules) was pre-screened using a simplified pharmacophore with 6 features based on two known HK2 inhibitors (“cmpd27” and “cmpd34” [43]) using MOE2016 (**Figure 27**). Of the entire library, 175 113 molecules satisfied 4 out of the 6 features and were pre-selected and screened using molecular docking.

Since the VS mode was activated, only 50 conformations were generated during molecular docking (GOLD 5.2) for each molecule of each database. After ranking, the best pose of molecules with a score higher than 80 was visually analyzed. The position of the conformation into the pocket was compared with known crystallographic inhibitors and 2981 molecules were selected.

4.3. Prediction of ADMET properties

FAF-Drugs4 (Free ADME-Tox Filtering Tool) [204] (<https://fafdrugs4.rpbs.univ-paris-diderot.fr/>), and OCHEM (a platform for data storage, model development and publishing of chemical information) [219] (<https://ochem.eu/home/show.do>) online servers were used to predict ADMET properties according to the guidelines found on the respective websites. Information about the calculations performed is available in the websites. All available filters were used in the FAF-Drugs4 server, with exception of the iPPI profiler. OCHEM was used to predict toxicity, including genotoxicity, CYP modulation, and solubility in DMSO. Rejected molecules were carefully analyzed and re-selected or empirically discarded. Among the parameters calculated, Lipinski violations, bioavailability, solubility (water and DMSO), and presence of PAINS moieties were considered.

4.4. Characterization of the acquired molecules

Identity of the acquired molecules was confirmed either by LC-MS and/or NMR. Technical details are found in **Chapter IV, Section 4**.

4.5. Biochemical hexokinase activity assay – ADP-Glo™ kinase assay

Stock solutions of the selected molecules were prepared in DMSO at 10 mM and kept frozen until use. HK activity was measured with a luminescence-based assay using the luciferin-luciferase system, with the ADP-Glo™ Kinase Assay kit (Promega Corporation, Madison, WI, USA). The ADP-Glo™ Kinase kit was used according to Lin *et al.* [43], and recommendations of the manufacturer [235]. The reaction buffer contains 100 mM Hepes, 100 mM KCl, 10 mM MgCl₂, (pH 7.2), 2 mM 1,4-dithiothreitol (DTT), 0.05 mg/mL BSA, and 0.05% 3-((3-cholamidopropyl) dimethylammonio)-1-propanesulfonate (CHAPS). Reaction mixtures (5 µL) of 50 nM rhHK2 (host: *Escherichia coli*), 100 µM ATP and 100 µM Glucose and the selected molecule at 10 µM in DMSO or the corresponding volume of DMSO (control) or were plated in white 384-well microplates and incubated for 45 min at r.t. ADP-Glo™ Reagent (5 µL) was added to terminate the reaction and deplete the remaining ATP (60 min incubation). Kinase Detection Reagent (10 µL) was finally added and after an incubation of 45 min luminescence was detected with an Anthos Zenyth 3100 Microplate Multimode Detector (**Figure 40**). The light generated is proportional to the amount of ADP present and, consequently, HK activity. A blank was prepared using a reaction mixture without HK2. No interference by the selected molecules was detected.

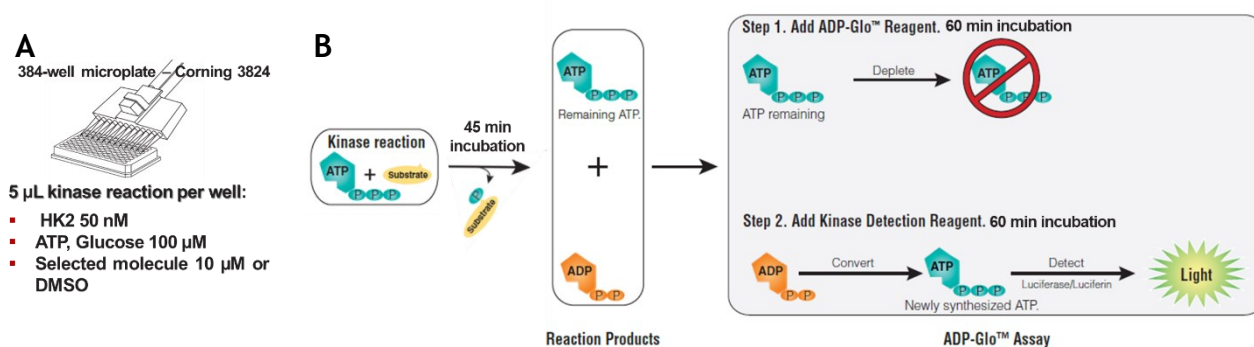


Figure 40. **A** - The HK2 reaction was prepared with 50 nM HK2, 100 µM ATP and 100 µM Glucose and each of the selected 64 molecules at 10 µM and incubated for 45 min in 384-well white polystyrene microplates. **B** - After the kinase reaction, the first step was performed by addition of the ADP-Glo™ Reagent that terminated the kinase reaction and depleted any remaining ATP (60-min incubation). Addition of a second reagent converted ADP to ATP and generated light from the newly synthesized ATP using a luciferase/luciferin reaction (60-min incubation). The light generated is proportional to the amount of ADP present and, consequently, kinase or ATPase activity. Adapted [235]. Luminescence was detected with an Anthos Zenyth 3100 Microplate Multimode Detector.

4.6. *Quantitative determination of ATP and ADP by HPLC*

ADP and ATP solutions (Promega Corporation, Madison, WI, USA) were prepared in the reaction buffer described in 4.5 as sets of sequential dilutions from 100 μM , in the day of the assay. Analyses were performed in a modular HPLC system composed of a Varian ProStar 410 autosampler, two 210-LC chromatography pumps and a ProStar 325 UV detector (Varian, Inc., Palo Alto, CA, USA). Samples (20 μL) were injected onto the column via a Rheodyne injector (Rheodyne LLC, IDEX Corp., Lake Forest, IL, USA) with a 100 μL loop in the μL pickup injection mode. Separations were conducted at room temperature, using a ThermoFisher Scientific BDS Hypersil C18 ($\text{Ø}4.6 \times 250 \text{ mm}$, 5 μm) reversed phase column and a 1 mL/min flow rate. The mobile phase consisted of 50 mM sodium phosphate monobasic buffer pH 5.5 in water (Eluent A). Elution occurred during 15 min isocratic runs at 100% eluent A. Chromatograms were recorded at 254 nm and data acquisition and processing were performed using Varian MS Control 6.9.6 software. Standard and sample concentrations were calculated from sample peak areas and linear equations ($y = ax + b$) (**Figure 35**). Samples were run at least twice in triplicates.

Chapter IV

Chemical Synthesis

1. Overview

Chapter IV describes the organic synthesis of potential HK2 inhibitors found in the molecular docking studies and of a known HK2 inhibitor.

In the structurally diverse library of 454 molecules pre-selected after first filtering of the pharmacokinetic properties, 41 molecules belonged to a family of ribose derivatives (**Figure 41, A**). The synthesis of 6 compounds of this promising family (commercially available but associated with high costs) was the first goal. These compounds would allow the evaluation of the influence of the ring size of the sugar moiety and of sulfonamide, amide, or amine groups in the possible inhibition of HK2. Upon obtaining those results, other compounds were envisioned for the study of structure-activity relationships (SAR). These molecules are in fact similar to known HK2 inhibitors (**Figure 41 A and B**), the most expressive difference being the size of the ring of the central sugar moiety. While known inhibitors likely contain a hexose, primarily in a pyranose form, the synthesized family of compounds acquired a furanose structure.

Upon obtention of the common core of the ribose family, derivatized on positions 1 and 5, some molecules thought to be active against HK2 were generated (**Scheme 2, A**),

Although the moderate activity of commercially available 3BP against HK2 has been amply reported, we opted for synthesizing a molecule (**Scheme 2, B**) known for its high potency at inhibiting HK2 ($IC_{50} = 25$ nM), as positive control to compare the performance of our potential HK2 inhibitors. This molecule, 2-[(3-biphenylcarbonyl)amino]-2,6-dideoxy-6-[[2,3-dichlorophenyl)sulfonyl]amino]-D-glucopyranose, named “cmpd 25” by the authors [43], is not commercially available

In the end, biochemical evaluation of the synthesized molecules as HK2 inhibitors was also performed using the luminescence-based method described in the previous chapter.

Standard organic synthesis reactions were used, such as protection and deprotection, Wittig reaction, SN_2 substitutions, hydrogenation and amide coupling, applied to carbohydrate molecules (**Scheme 2**). Characterization of the molecules followed the typical techniques, i.e., NMR and MS.

Carbohydrates, a group of compounds also known as sugars, are present in all living organisms as part of the essential constituents, being the most abundant and widespread in nature. Also, this class is the most important (in number and availability) of non-nitrogenous compounds of the chiral pool and they are particularly important in chiral synthesis. In general, the monomers of this class are polyhydroxylated structures containing 3 to 9 carbons, typically very functionalized [236,237]. Although carbohydrates are one of most important and diverse class of compounds in nature and act as excellent recognition molecules on the cell surface, just a few carbohydrate-based drugs are on the market. In general, carbohydrates possess high polarity due to the hydroxyl groups (hydrogen bond donors and acceptors), which typically offers poor pharmacokinetic properties, such as i) low tissue permeability, because of their lack of lipophilicity required for passive

absorption through biological membranes, and ii) short serum half-lives, being quickly eliminated by renal excretion. Also, in many cases, the target binding affinities of carbohydrates are in the milli- or micromolar range, which means relatively weak interactions [238,239]. Nonetheless, should the core structures be promising, developing better drug-like molecules is possible through several structural modifications.

Carbohydrate chemistry is not always straightforward, due to the presence of multiple chiral centers with potential to react. Generally, reactions originate side products, sometimes difficult to isolate. In this regard, use of protection groups is indispensable to work in specific positions of the molecule. As the anomeric carbon has the ability to undergo mutarotation in solution, blocking it by forming an acetal or a C-glycoside, instead of keeping free the hydroxyl at C1, tends to stabilize the molecule as a cycle, which is the most plausible via to maintain the α/β -configuration unchanged.

Introduction of large non-polar groups using poorly hydrolysable bonds is a convenient strategy to increase the potential of carbohydrate-based molecules to interact with the desired target, not only by enabling extra van der Waals interactions with the target macromolecule, but also by increasing the chance of the molecule to permeate cell membranes. As an example, 100% of the molecules of the class of HK2 inhibitors derived from glucosamine [43] are predicted to have oral bioavailability, according to the Veber rules.

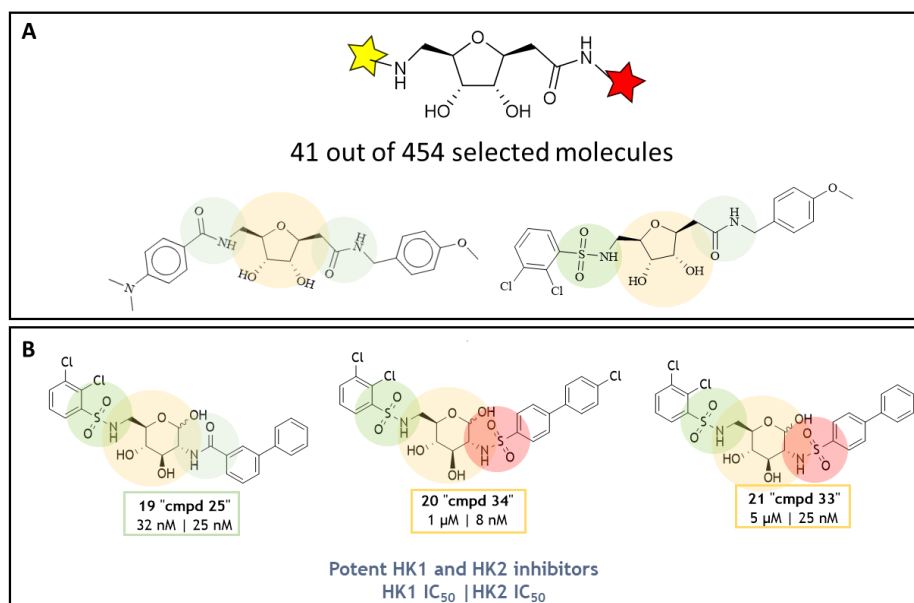
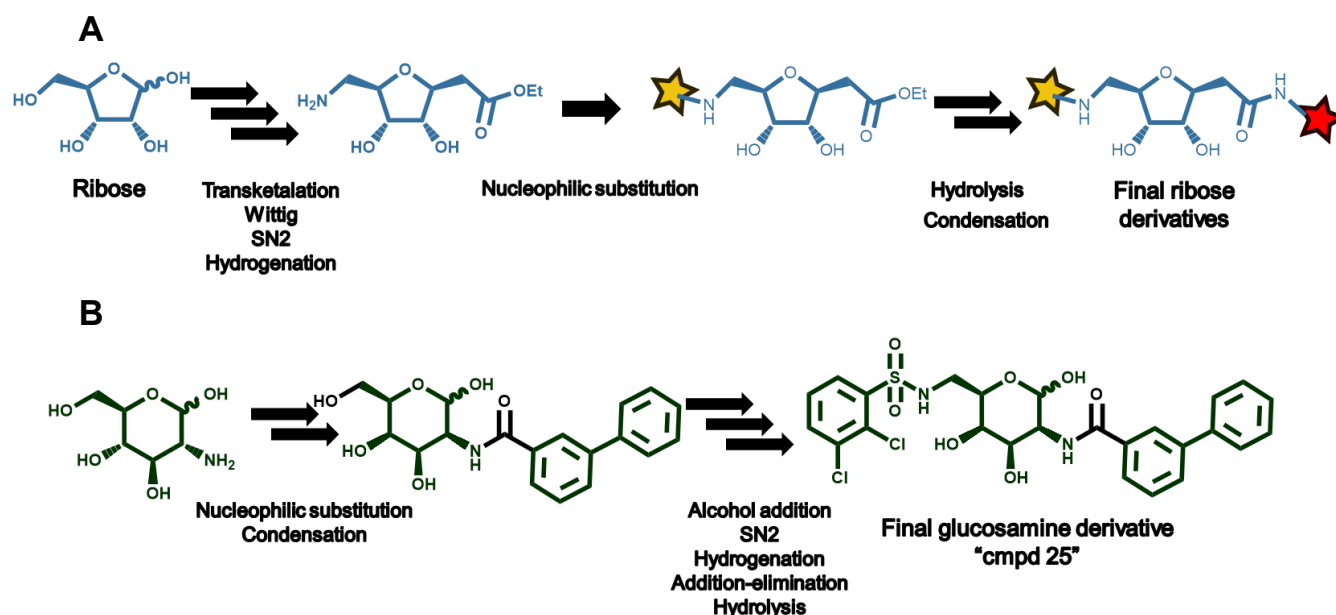


Figure 41. **A** - Core of the ribose family found consistently in the final virtual library of selected molecules and structures of two synthesized molecules. **B** - Structures of potent known HK1/HK2 inhibitors. The IC_{50} for each isozyme is indicated for each molecule as HK1 IC_{50} | HK2 IC_{50} . The main structural features are highlighted: Yellow – sugar moiety; Light green – Amide; Green – Sulfonamide; Red – sulfonamide on the C2 nitrogen of glucosamine.



Scheme 2. Types of reactions carried out to obtain the final products, ribose derivatives (**A**) and "cmpd 25" (**B**).

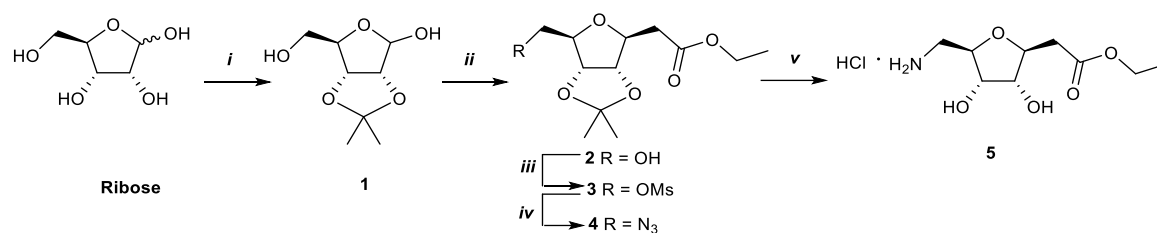
2. Results and discussion

2.1. Synthesis of Ribose-derivatives

2.1.1. Synthesis of the ribose-based core

The synthetic route designed for the synthesis of the main core of the desired ribose family is shown on **Scheme 3**. Albeit all the reactions were described before [240–243], optimization of experimental conditions was enforced.

The route started with the protection of hydroxyl groups on positions 2 and 3 with an isopropylidene group, formed by transketalation from 2,2-dimethoxypropane and the *cis*-2,3-diol in acidic medium. With the aim to overcome the preference of 5-carbon monosaccharides for a pyranose form [244], the introduction of the isopropylidene group not only protects the hydroxyl groups from unwanted reactions but also promotes the recyclization as a furanose, conferring contention and movement restrictions. Other protection groups such as acetyl groups would confer flexibility to the molecule, enabling the formation of the pyranose counterpart. The NMR characterization of **1** (*cf.* Experimental Section) is in full agreement with the literature for the β -configuration as the major species form [245].



Scheme 3. Synthetic route for the synthesis of the ribose-based core. (i) 2,2-dimethoxypropane, *p*-TsOH, acetone, 0 °C to r.t., 1.5 h (quant.); (ii) Ethyl (triphenylphosphoranylidene)acetate, ACN, 2.5 h, 80°C (47%); (iii) MsCl, pyridine, 6h, r.t. (94%); (iv) NaN₃, DMF, 75 °C, 1.5h (93%); (v) H₂, Pd/C, HCl/EtOH, 18h, r.t., 1 atm (79% - overall yield).

Compound **2** was obtained after synthesis of the Wittig reagent ethyl (triphenylphosphoranylidene)acetate. Specific description of the C-glycoside **2** was not found in the literature, but the analogous methyl ester was fully described by Ohruí *et al.* [240], which enabled to confirm the structure as the β -configuration by comparison of the spectroscopic data. In that study, the configuration was established by analysis of the ¹³C NMR data, based on the premise that chemical shifts of carbon atoms are deeply affected by steric crowding, particularly by vicinal oxygen substituents. The chemical shifts of C1, C2 and C3 were found to occur at higher field when a *cis* (rather than a *trans*) relationship occurred between the C2 and C1 substituents, confirmed by X-ray crystallography studies of resembling compounds [240]. Considering the ¹H NMR spectrum, the data of Ohruí *et al.* showed that the most significant features were related with protons **1''a** and **1''b**, which were chemically and magnetically equivalent, as well as protons **5a** and **5b**, in the α -anomer. In the ¹H NMR spectrum obtained from compound **2**, magnetic difference between the abovementioned protons is clear, regarding both chemical shifts and couplings with the neighboring protons (**Figure 42**). **Scheme 4** shows the proposed mechanism for the reaction. Both anomers were found as products; however, the β -anomer was the major compound (kinetic product). Conversion into the α -anomer would be expected to occur in basic reaction medium, as it was found to be the most thermodynamically stable compound [240]. Side-products other than the α -anomer were also found, as well as some unreacted starting material **1**, reducing the yield of the β -anomer. The structures of the impurities were not fully analyzed; however, the presence of a non-cyclic hemiketal with protection of only one of the oxygen atoms (on C2 or C3) is plausible. As α - and β - anomers are present, in total eight compounds derived from ribose would be expected. Together with "triphenylphosphoranylidene derivatives", the purification of the β -anomer was challenging, and no more than 47% yield was obtained.

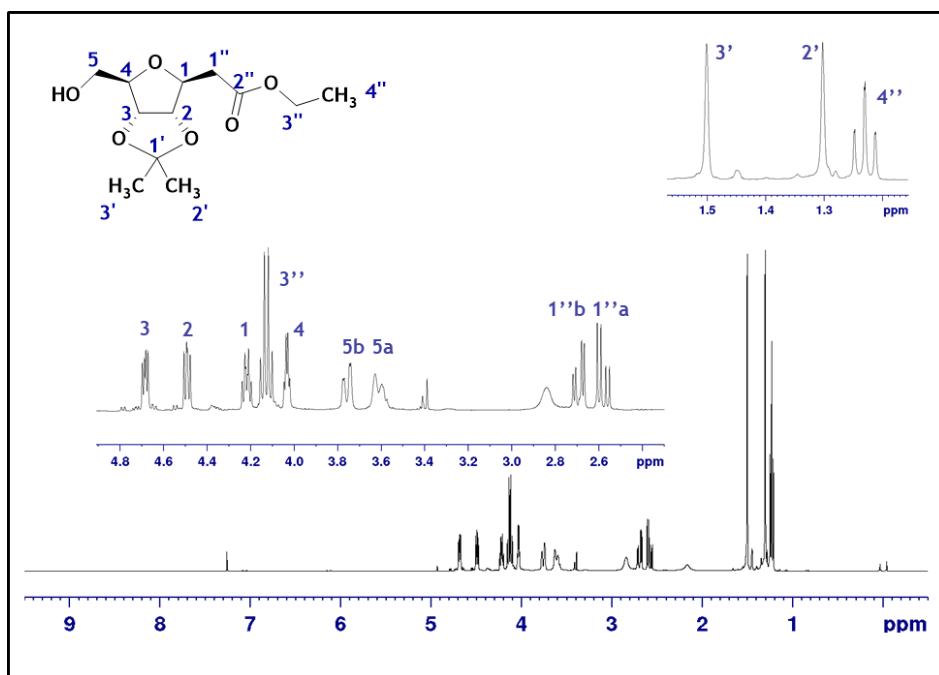
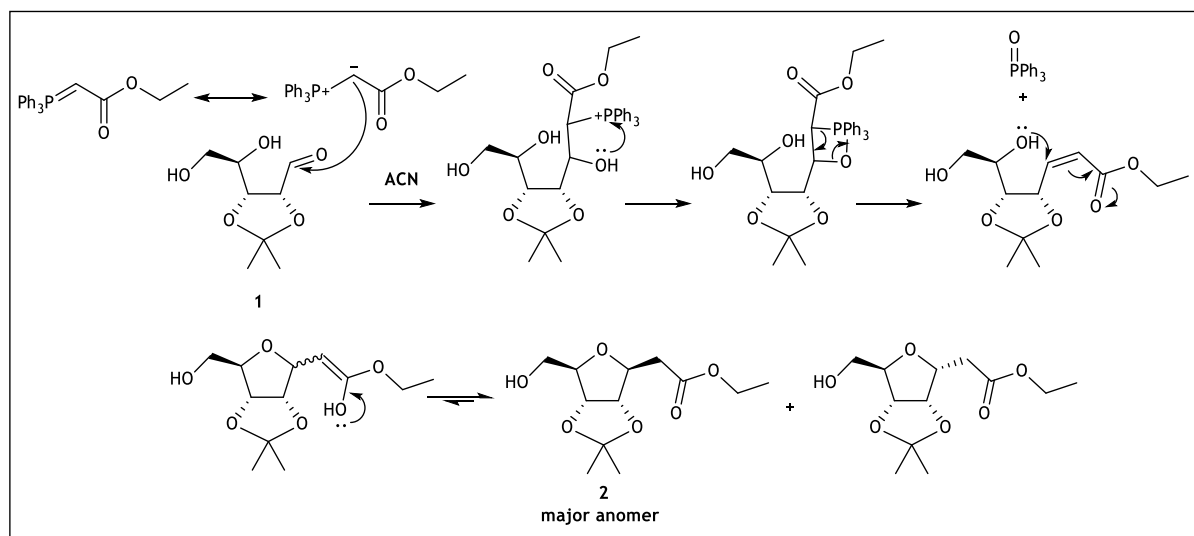


Figure 42. ^1H NMR spectrum (300 MHz, CDCl_3) of the C-glycoside **2**. The magnetic non-equivalence of protons $1''\text{a}$ and $1''\text{b}$ and 5a and 5b is clearly displayed, indicating the β -configuration of the molecule, obtained after column chromatography.



Scheme 4. Proposed mechanism for the synthesis of the C-glycoside **2** [240].

In solution, pyranose structures converge essentially into two stable chairs, $^4\text{C}_1$ generally associated with D-sugars and $^1\text{C}_4$, associated with the L-series (**Figure 43**). This allows a consistent range of coupling constants for each type of proton, including axial-equatorial, equatorial-equatorial and axial-axial coupling constants (J), extremely helpful to decipher the conformation adopted with relatively few NMR studies. In 5-membered rings, such as furanose monosaccharides, barriers to conformations' interconversion are considered lower, leading to an exceptional conformational flexibility. In this way, instead of a preferred conformation, a conformational equilibrium is adopted, characterized by a

pseudo-rotation causing interchanging among the most stable conformations. This dynamic equilibrium implies multiple ^1H - ^1H and ^1H - ^{13}C couplings which are not well categorized, therefore hampering its prediction. Moreover, the protons of the furanoid ring frequently overlap in the 1D spectra, interfering with the full measurement of coupling constants. Furthermore, ribofuranose derivatives' conformation is more sensitive to the substituents, particularly on the anomeric position, than their related pyranose counterparts [244,246–248]. The conformational equilibrium is also affected by the solvent, as demonstrated by differences in the pseudo rotation reported with the use of different solvents [249,250].

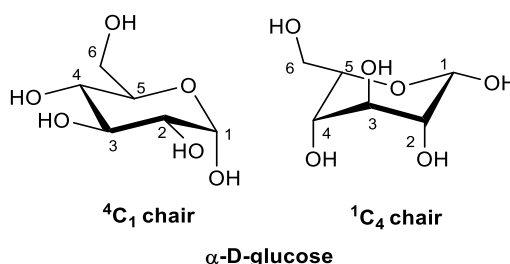


Figure 43. Most stable pyranose structures: $^4\text{C}_1$ and $^1\text{C}_4$ chairs.

Similarly to compound **2**, ethyl esters **3** and **4** were not found to be described in the literature, but their similarity with the corresponding methyl esters [243] allowed the confirmation of the expected β -configuration. Accordingly, the spectroscopic data of common protons and carbons were consistent with those published for the methyl esters. As expected, the substitution reactions were straightforward and no purification steps were necessary, since no significant impurities were found after workup, as confirmed by the ^1H NMR spectra (**Figure 44**). Protons of the mesyl group (H6) are clearly identified in the spectrum as a singlet at 3.03 ppm (**Figure 44**, spectrum on the left). Substitution by the azide group is recognized by the shift of protons 5a and 5b to a higher field compared to the corresponding protons of the precursor (4.3 ppm to 3.5 and 3.3 ppm), and the loss of the signal relative to the methyl group.

The final step to obtain the complete ribose core was the hydrogenation of the azide to primary amine. This reaction occurred smoothly at 1 atm, accompanied by concomitant cleavage of the isopropylidene group. Together with the desired product, partial transesterification occurred during the reaction, due to the use of methanolic HCl. By ^1H NMR, it was estimated that approximately one third of the product mixture was the methyl ester (**5b**) and the remaining two thirds corresponded to the ethyl ester (**5a**). Moreover, the ^{13}C NMR spectrum is in full agreement with the presence of both esters, being possible to find the methyl carbon corresponding to the methyl ester, as well as the existence of two signals consistent with ester carbonyl groups and also two signals for C1' (**Figure 45**). As both esters would be useful for the next steps, purification was not carried out. Given the similarity between the two molecules, separation would not be achieved easily and it was deemed unnecessary to perform a quantitative hydrolysis followed by reesterification.

At this phase, accuracy of yield calculation was not the main concern and for this reason crude mixtures were accepted to proceed to the next reactions without further purification. In this regard, the yields herein presented are approximate. Moreover, optimization of steps such as the synthesis of compound

2 was not fully pursued; it would have been a priority if larger amounts of the compound had been needed.

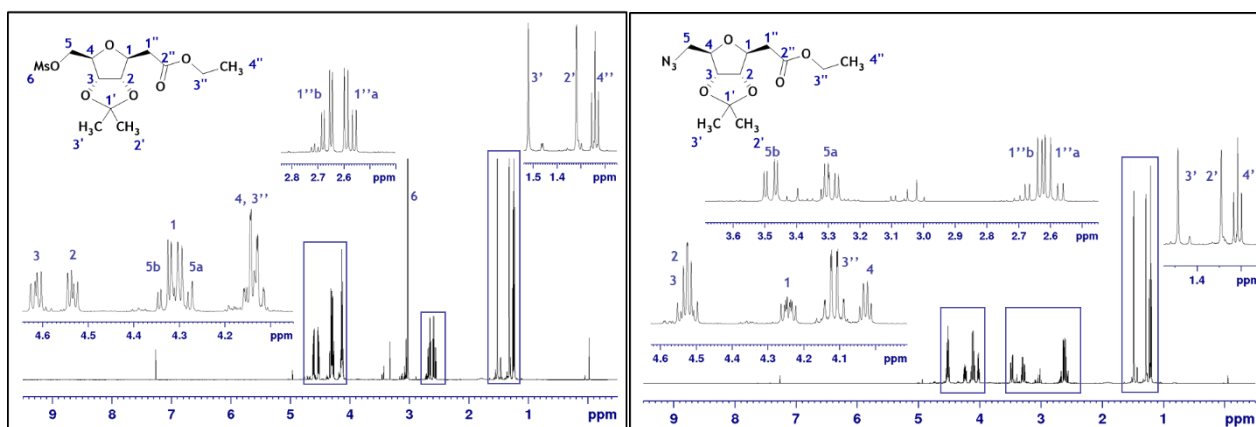


Figure 44. ^1H NMR spectra (400 MHz, CDCl_3) of compounds **3** (left) and **4** (right).

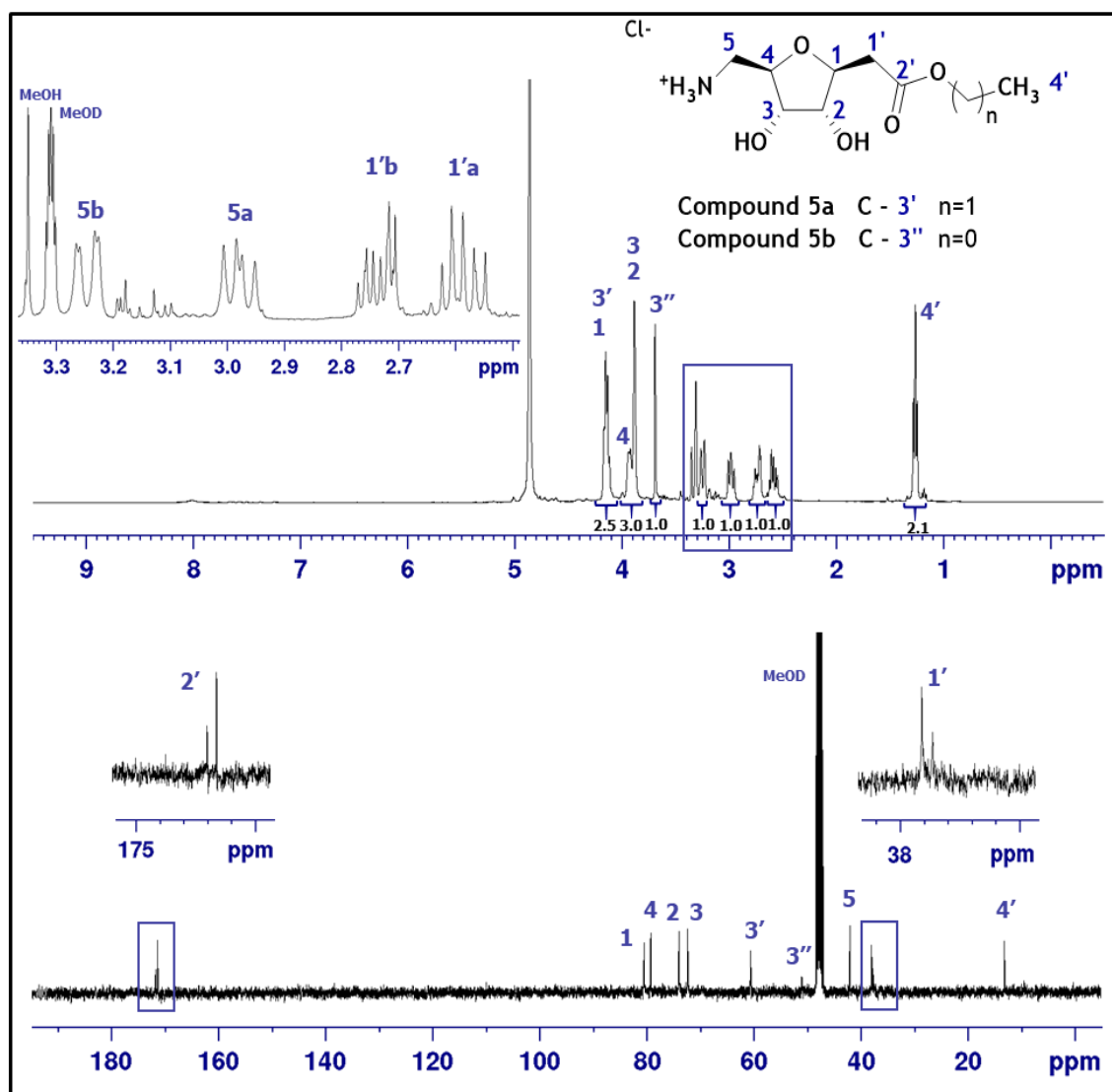
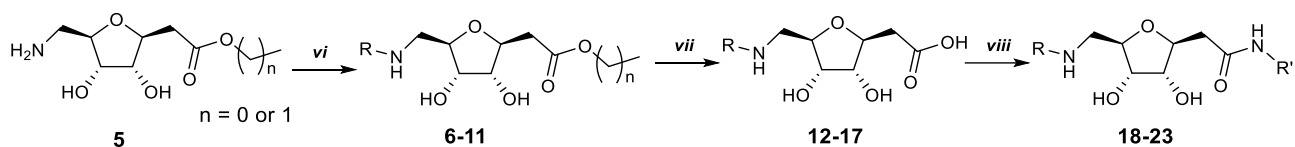


Figure 45. ^1H (above) and ^{13}C (below) NMR spectra (400 MHz, 100.6 MHz, MeOD) of compounds **5a** (ethyl ester) and **5b** (methyl ester). Integration is shown for relevant signals on the ^1H NMR spectrum.

2.1.2. Synthesis of derivatives of 5

When the core **C-glycoside 5** was obtained, derivatization started on position 5 by the formation of amides or sulfonamides. Upon subsequent hydrolysis of the ester, coupling of the resulting carboxylic acid with primary amines resulted in six final ribose derivatives (**Scheme 5, Figure 46**).



Scheme 5. Synthetic route for the synthesis of ribose derivatives from compounds **5a** and **5b**. (vi) Acyl/Sulphonyl Chloride (R-Cl), Et₃N, DCM, r.t. to 40 °C, 24h, 20-67%; (vii) 1) LiOH, ACN, 45 min, r.t. 2) HCl (66-99%). (viii) NH₂R', DIPEA, HATU, 10:1 DCM/DMF, 18-24h, r.t. to 40°C (3%-39%).

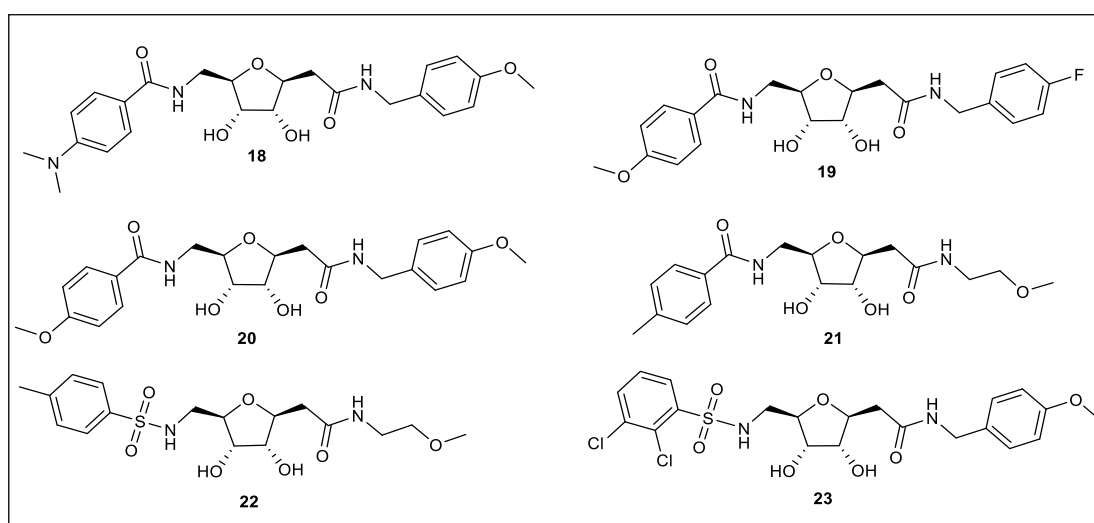


Figure 46. Structures of the final compounds **18-23**.

Acyl/sulfonyl chlorides and primary amines bearing the R and R' groups, respectively, were selected according to their availability, docking results and potential to contribute clues to the establishment of structure-activity relationships (**Figure 47 A and B**). Use of both *p*-toluoyl and *p*-toluenesulfonyl chlorides, would allow the confirmation or disproof of the importance of the sulfonamide group for the inhibitory effect. To compare the effect of electron withdrawing groups (EWG, chlorine, fluorine) with that of electron donating groups (EDG, amino or methoxy groups) these types of groups were incorporated. Also, the impact of a more rigid aromatic benzylamine ring was compared with that of the flexible 2-methoxyethylamine. Due to availability issues, the final compounds did not include derivatives of compound **16** (which should have been prepared from 4-chlorobenzenesulfonyl chloride), but the intermediate products were biochemically evaluated.

As the main objective was to produce just enough compound for biological screening, no more optimization steps were investigated.

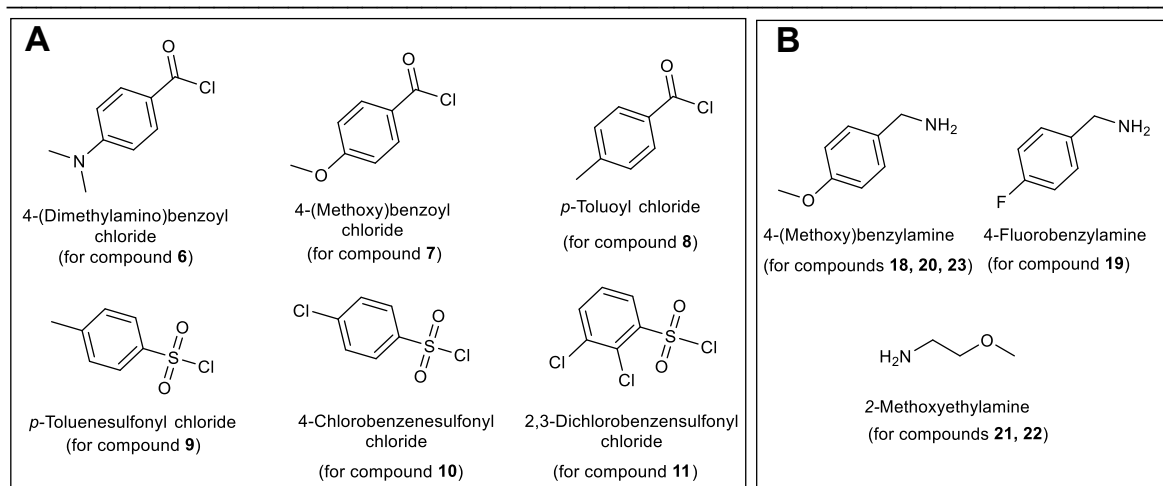


Figure 47. A - Acyl/sulfonyl chloride precursors for the synthesis of compounds 6-11. B - Primary amines used for the synthesis of compounds 18-23.

Amides/sulfonamides (compounds **6-11**) were first obtained in the synthetic sequence by reaction of the primary amine on the ribose C-5 with acyl or sulfonyl chlorides in basic conditions. These reactions were expected to be straightforward; however, they were not complete and poor yields were obtained after purification by preparative TLC (20-67%). As an attempt to achieve better results, different conditions such as temperature, reaction time and solvents were tested (data not shown). The reactions were mostly not complete, regardless of the reaction time, and no benefit was found in allowing the reaction to proceed for more than 24h prior to workup. Slight changes in temperature were also tried, from r.t. to 40°C (DCM boiling point) and warmer conditions were preferred. The reactions were also attempted using ACN (no reaction) and pyridine as solvents, but no benefit was observed. As the main objective was to produce just enough compound for biological screening, no more optimization steps were investigated.

As both compounds **5a** and **5b** were used, ethyl and methyl esters were found as products (**6-11**). The next step was the ester hydrolysis and both products were used without further separation. NMR confirmed the hydrolysis, with loss of the ethyl and methyl groups (compounds **12-17**). The multiplicity of the ^1H NMR signals of the esters was generally unclear due to the presence of both methyl and ethyl groups. In contrast, the ^1H NMR signals of the resulting carboxylic acids were well defined, since only one product was present. Because superposition of different proton signals was observed multiple times for ribose protons, the presence of both esters is mostly noted by the multiplicity observed for H1'a, H1'b and H5a, H5b (**Figure 45**). Each of these protons should appear as a doublet of doublets (dd); however, in the esters the multiplicity of these protons was easily confused with a double doublet of doublets (ddd) or a doublet of triplets (dt) or was not even defined (**6-11**). With the exception of H1'a of carboxylic acid **15** that shares the same chemical shift with a methyl group the multiplicity of H1'a, H1'b of all carboxylic acids was dd, as expected. **Figure 48** exemplifies the typical signals obtained in the ^1H NMR spectra of precursor esters and the product carboxylic acid. Clear signals of protons H5a, H5b were not always obtained due to signals superposition.

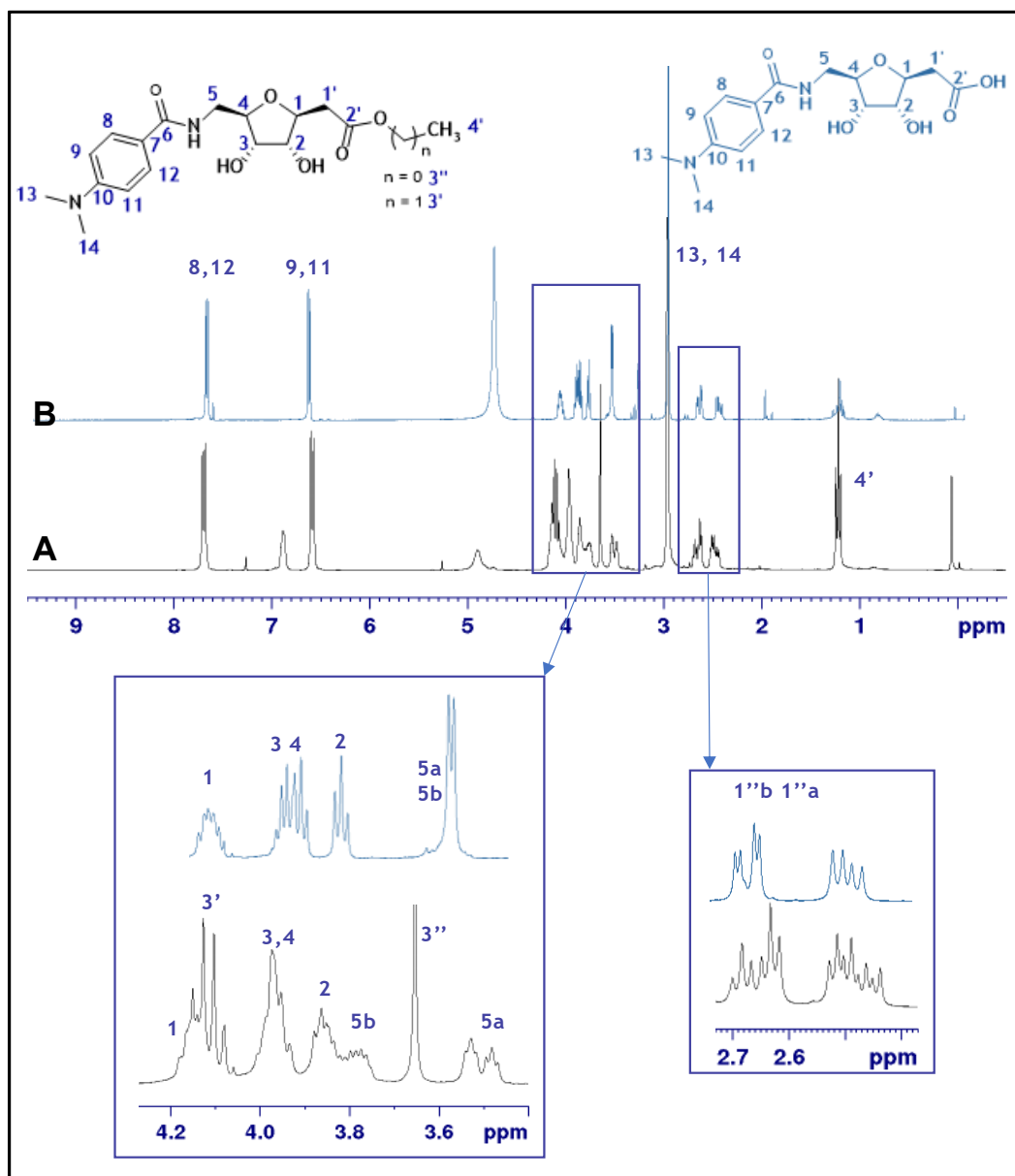


Figure 48. ^1H NMR spectra (400 MHz) of compounds **6** (A) (CDCl_3) and **12** (B) (MeOD).

Comparing the ^1H NMR spectrum of compound(s) **5** (Figure 45) with those of compounds **6-11**, the most noticeable chemical shift change was related with H5a and H5b, the protons closer to the reaction site. Amide formation (**6,7,8**) brought those protons downfield (~ 0.5 ppm), due to the electrowithdrawing effect of the carbonyl group. By contrast, the presence of the weaker electrowithdrawing sulfonamide did not influence much the average chemical shift, although the H5a and H5b chemical shifts were closer to each other (Figure X). When the carboxylic acid was free at C2', the magnetic nonequivalence of H5a and H5b was slim or undetectable in both C5-sulfonamides (**15, 16, 17**) and -amides (**12, 13, 14**) (Figure X, Table 6). Not unexpectedly, the H1'a and H1'b protons in the carbon located α to the carbonyl group were virtually not affected by the changes of functional groups (ester to carboxylic acid to amide) as no major differences should occur in the deshielding ability of the carbonyl.

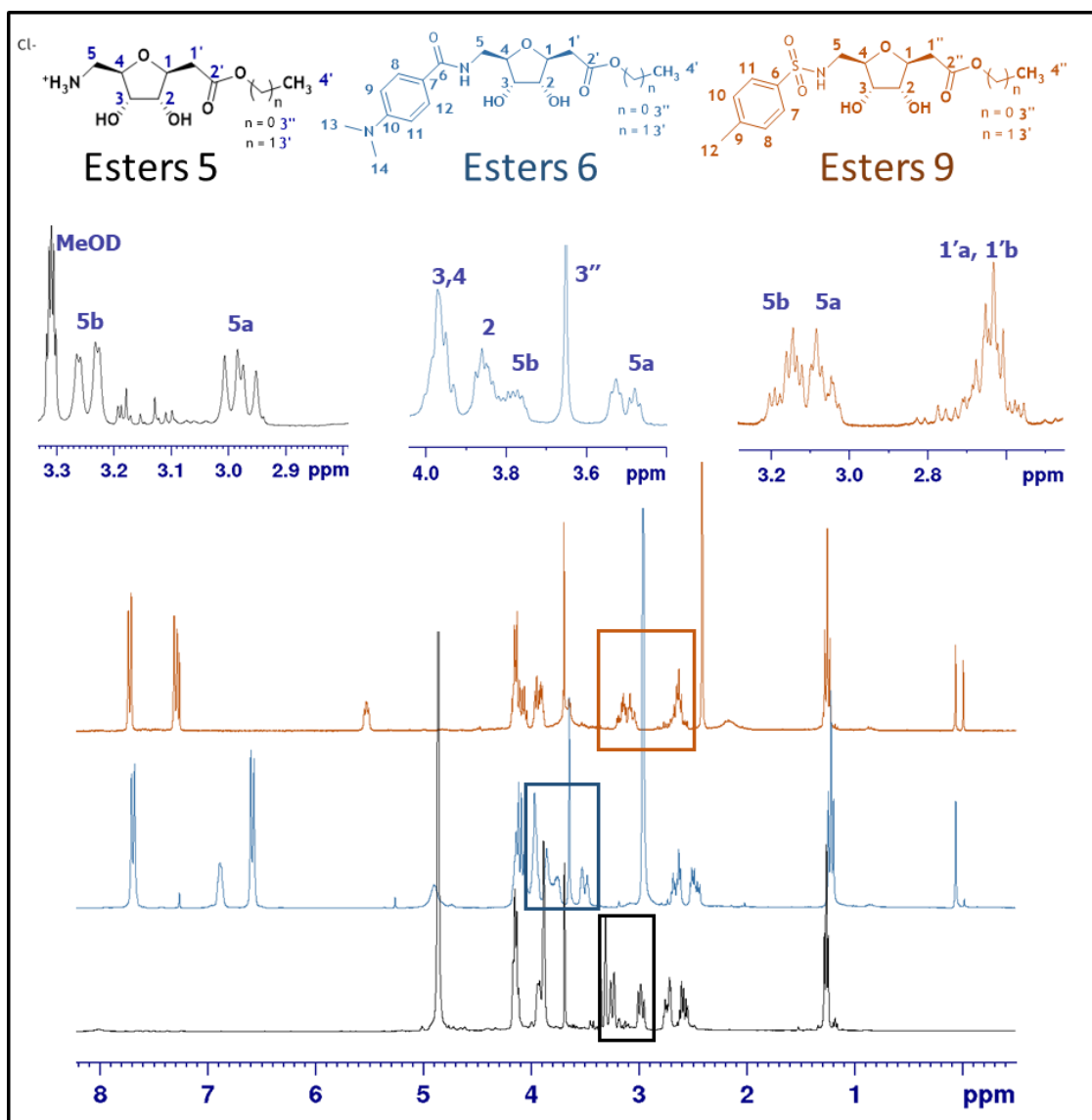
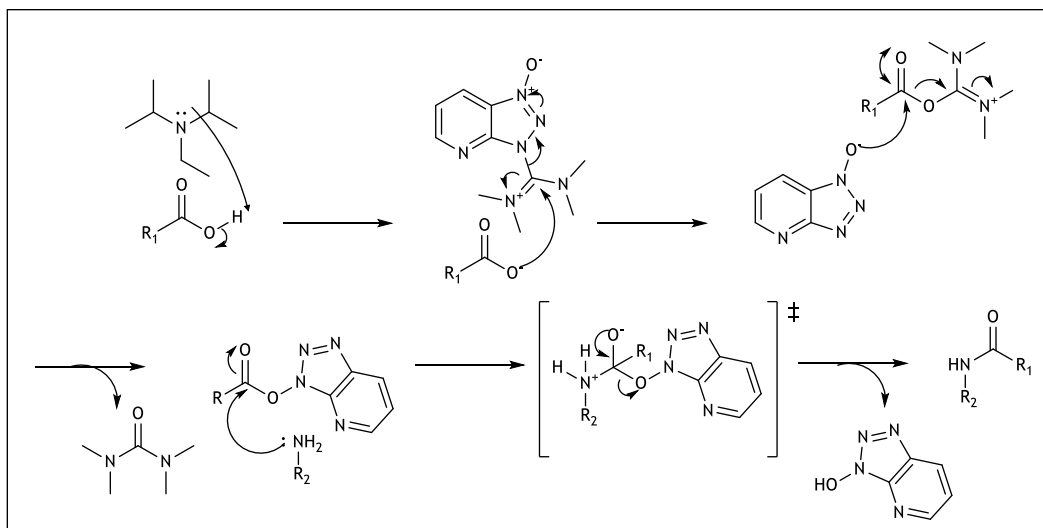


Figure 49. ^1H NMR spectra (400 MHz) of esters 5 (black, MeOD), 6 (blue, CDCl_3) and 12 (orange, CDCl_3).

Table 6. Chemical shifts of H5a and H5b in compounds 5-17.

Compound	Ester/ Carboxylic Acid	R1	Average chemical shift (ppm)	
			H5a	H5b
5	Ester	-	2.98	3.25
6	Ester	4-(Dimethylamino) benzoyl	3.50	3.80
12	Acid	4-(Dimethylamino) benzoyl	3.56	3.60
7	Ester	4-(Methoxy)benzoyl	3.54	3.91
13	Acid	4-(Methoxy)benzoyl	3.54	3.59
8	Ester	<i>p</i> -(Toluoyl)	3.55	3.85
14	Acid	<i>p</i> -(Toluoyl)	3.59	3.59
9	Ester	<i>p</i> -(Toluenesulfonyl)	3.06	3.16
15	Acid	<i>p</i> -(Toluenesulfonyl)	2.96	3.02
10	Ester	4-Chlorobenzene-sulfonyl	3.10	3.10
16	Acid	4-Chlorobenzene-sulfonyl	3.00	3.07
11	Ester	2,3-Dihlorobenzene-sulfonyl	3.12	3.12
17	Acid	2,3-Dihlorobenzene-sulfonyl	3.01	3.18

Coupling of the carboxylic acids with primary amines was carried out using hexafluorophosphate azabenzotriazole tetramethyl uronium (HATU) and *N,N*-diisopropylethylamine (DIPEA) in DCM (**Scheme 6**). This reaction was not complete, probably due to the multiple intermediate compounds that might have been stabilized over time. The final products (**Figure 46**) were mostly isolated by reverse phase HPLC (**18**, **21**, **22** and **23**) in very low to low yields (3-39%). Precipitation after HCl addition only occurred with **19** (29%) and **20** (35%), the ones derived from **13**, which contained the 4-methoxybenzoyl group on the C5 nitrogen. Precipitation of the remaining compounds was attempted with different solvents, with no success. Due to the characteristics of the resulting final products, the best chromatographic strategy to isolate the desired products was using reverse phase HPLC, which allowed the purification of most of them. However, despite of all the efforts, the desired product **18** was obtained as a mixture, containing the intended product and its *N*-demethylated derivative, as indicated by NMR and LC-MS analysis (**Figure 50**). *N*-demethylation took place due to the acidic aqueous media where the compound was prior to extraction with ethyl acetate. The extent of demethylation is unclear. LC-MS analysis (**Figure 50**) suggested that the major component was compound **18** but this the *N*-methyl protons were not detected in the ¹H NMR spectrum. However, the amount of product was very limited to obtain high quality NMR spectra and therefore the uncertainty persisted.



Scheme 6. Proposed mechanism for the synthesis of the final products 18-23 from the acid precursors 12-17 and primary amines (4-(Methoxy)benzylamine, 4-Fluorobenzylamine and 2-Methoxyethylamine), catalyzed by HATU in basic conditions.

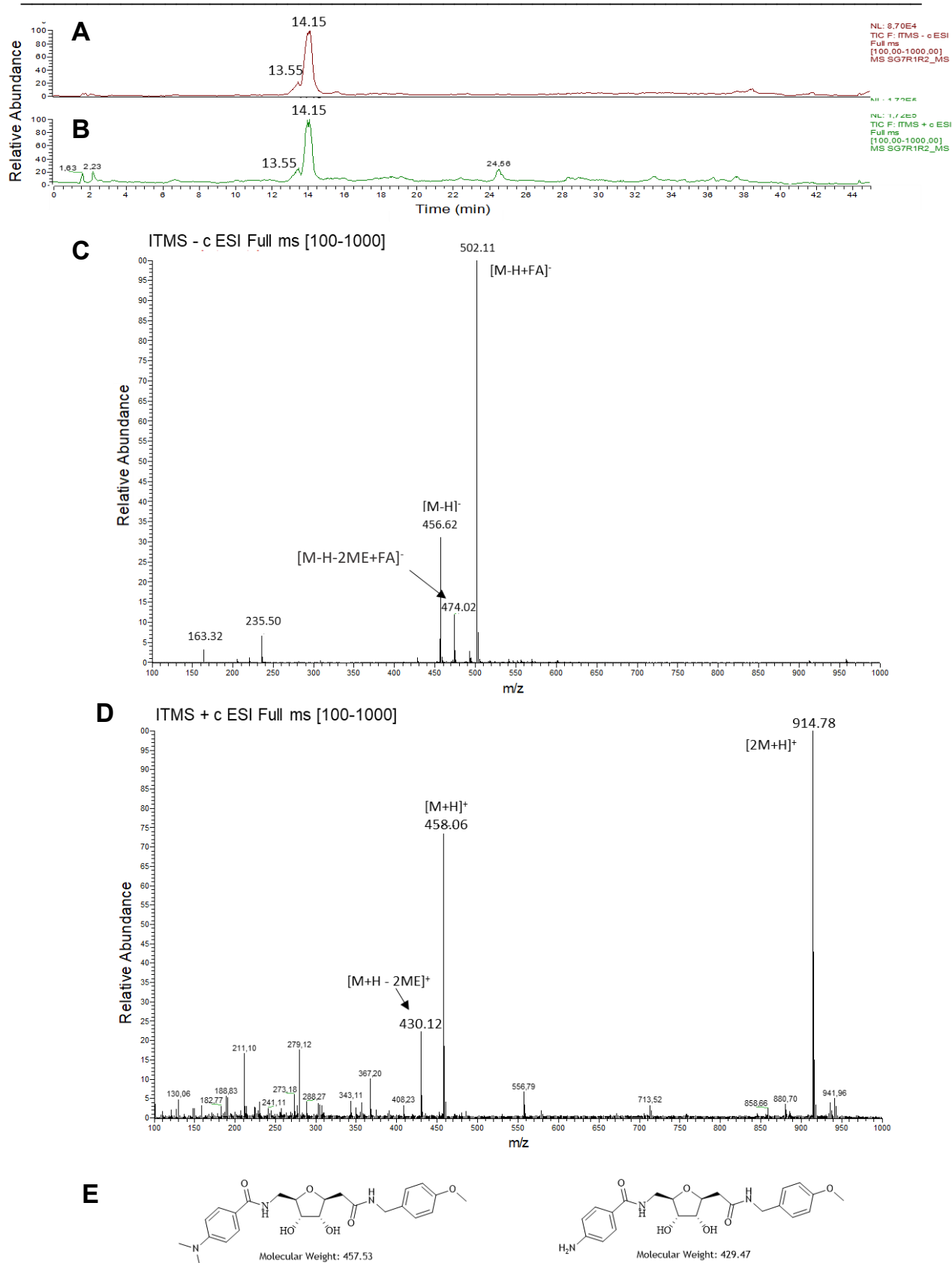


Figure 50. LC-MS results for the analysis of a sample of compound **18** obtained after semi-preparative HPLC. **A** - Chromatogram of the sample in negative ESI mode. **B** - Chromatogram of the sample in positive ESI mode. **C** - Negative ESI-MS spectrum [MW: 100-1000], retention time 0-45min. **D** - Positive ESI-MS spectrum [MW: 100-1000], retention time 0-45min. **E** - Structures and MW of compound **18** and its N-demethylated derivative. The main ESI-MS signals are identified, with M referenced to compound **18**. FA - Formic Acid; Me - Methyl group.

2.2. Synthesis of “cmpd 25” (known inhibitor)

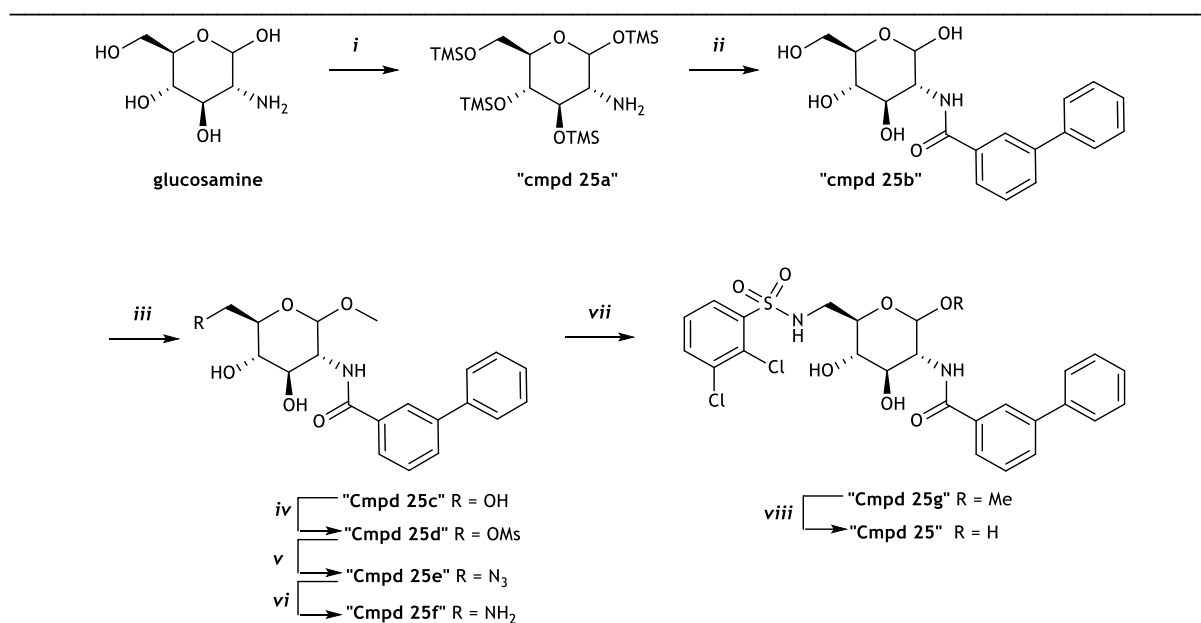
The synthetic route followed to obtain “cmpd 25” (**Scheme 7**) was mostly described by Lin *et al.* [43], along with its activity profile against HK1 and HK2.

The first reaction started with the protection of the free hydroxyl groups of glucosamine with TMS. TMS is particularly interesting in this step, enhancing the solubility of the initial molecule in organic solvents and allowing the selective protection of the hydroxyls while maintaining the amine free for subsequent reaction. The synthesis of the amide, “**cmpd 25b**”, took place in basic conditions (DIPEA), with the coupling agent HATU in DCM. Double acidification of the reaction mixture permitted the easy displacement of the silyl groups, the protonation of HATU, and then the selective precipitation of the product, “**cmpd 25b**”.

Although C6, as a carbon of a primary alcohol, is the most reactive position, in order to work on this position in the next steps proper protection of the anomeric carbon was required to block the ring configuration, eliminate the equilibrium with the open-chain structure, and thereby avoid unwanted side products. The methoxy group was easily introduced upon reflux in acidic methanol, while the secondary hydroxyl groups remained free (“**cmpd 25c**”). One should notice that this was the formation of an acetal from a hemiacetal and not a methylation, which explains why only the C1 position was affected.

The primary oxygen on position 6 was then mesylated (“**cmpd 25d**”) to allow a straightforward SN₂ reaction, substituting this group by an azide (“**cmpd 25e**”). Hydrogenation at 1 atm occurred overnight to give the free amine (“**cmpd 25f**”). The following sulfonamide formation was obtained in a modest yield of 60% (“**cmpd 25g**”), by the addition of the appropriate benzenesulfonyl chloride to a solution of “**cmpd 25f**” in pyridine. Deprotection of the anomeric carbon was the most critical step, since the dissolution of “**cmpd 25g**” in H₂O was not possible, even at reflux temperature. The starting material was recovered, and the reaction was repeated in a mixture of 2:1 dioxane:water, refluxed overnight. The crude mixture was purified by reverse phase HPLC and the product, “**cmpd 25**”, was obtained in a low yield (9%) and fully characterized

The change on the chemical shift of C1 from 99.68 to 92.46 ppm was essential to confirm the deprotection, along with the disappearance of the methyl carbon of the acetal (**Figure 51**). A shift on H1 from 4.63 to 5.01 ppm was also found (not shown); however, using the crude ¹H spectrum to identify the presence or absence of the Me singlet was not reliable, due to overlap of multiple resonances in that region, including the strong signal from the solvent (MeOD).



Scheme 7. Synthetic route for the preparation of "cmpd 25". i) Et₃N, TMSCl, pyridine, 115°C, overnight (71%); ii) 3-biphenylcarboxylic acid, DIPEA, HATU, DCM, r.t., overnight (76%); iii) MeOH, HCl, 65°C, overnight (94%); iv) MsCl, pyridine, r.t., 6h (60%); v) NaN₃, DMF, 90°C, 2h (94%); vi) Pd/C, EtOH, HCl, r.t., overnight (99%); vii) 2,3-dichlorobenzensulfonyl chloride, pyridine, overnight, r.t. (60%); viii) 2:1 dioxane:water; HCl, 100°C, overnight (9%).

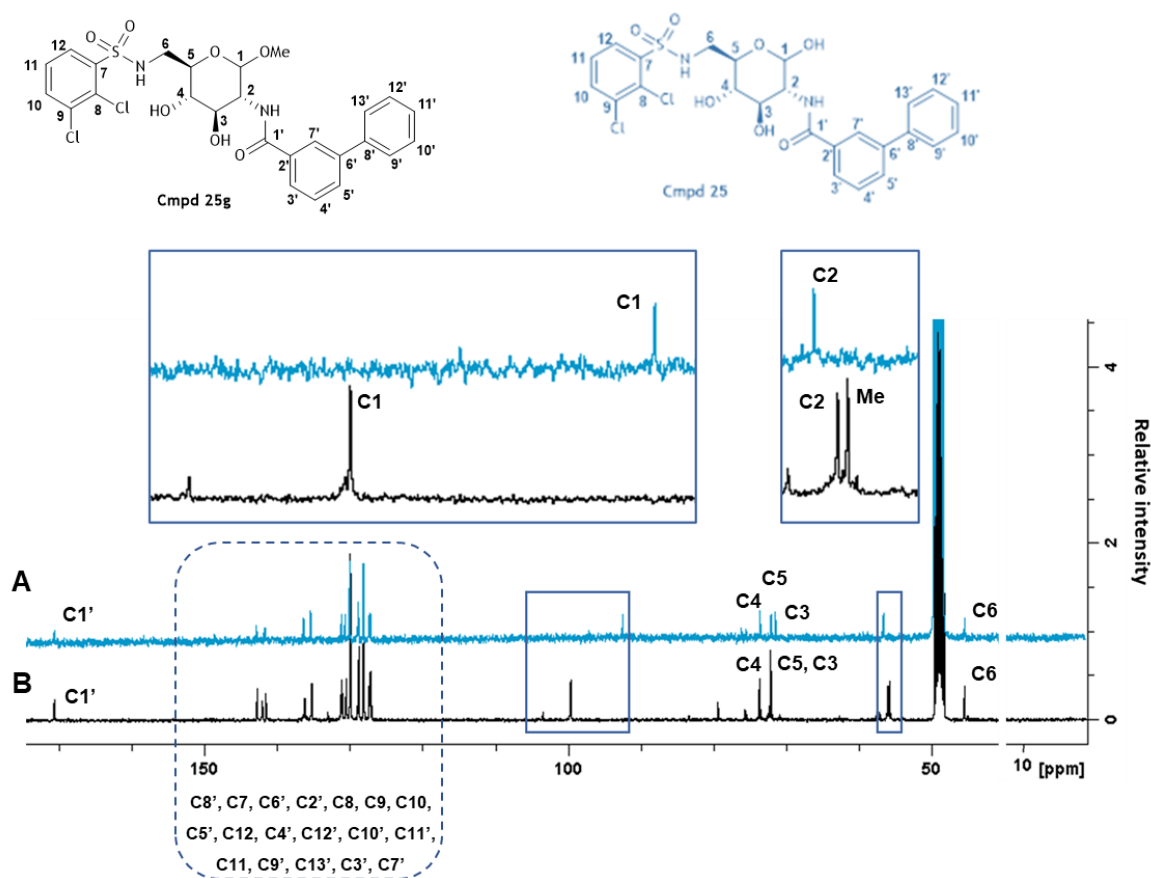


Figure 51. ¹³C NMR spectra of the final "cmpd 25" (A, blue) and its methylated precursor "cmpd 25g" (B, black).

2.3. Biochemical evaluation of the synthesized compounds as HK2 inhibitors

Using the same method described in the previous chapter (Chapter III, section 4.6.), the potential of the synthesized compounds as HK2 inhibitors was evaluated. The results are shown on **Figure 52**. None of the ribose derivatives was able to reduce the activity of HK2 substantially. However, as expected, “cmpd 25” inhibited HK2 by 97%, showing the reliability of the assay.

These results seem to highlight the importance of the pyranose ring for the enzyme inhibition, since the substituents of some of compounds were similar to those of known inhibitors, as shown previously on **Figure 41**. Still, this feature is not required for the activity of the enzyme, as fructose, another HK substrate usually assumes a furanose configuration, which is required for the phosphorylation on position 6.

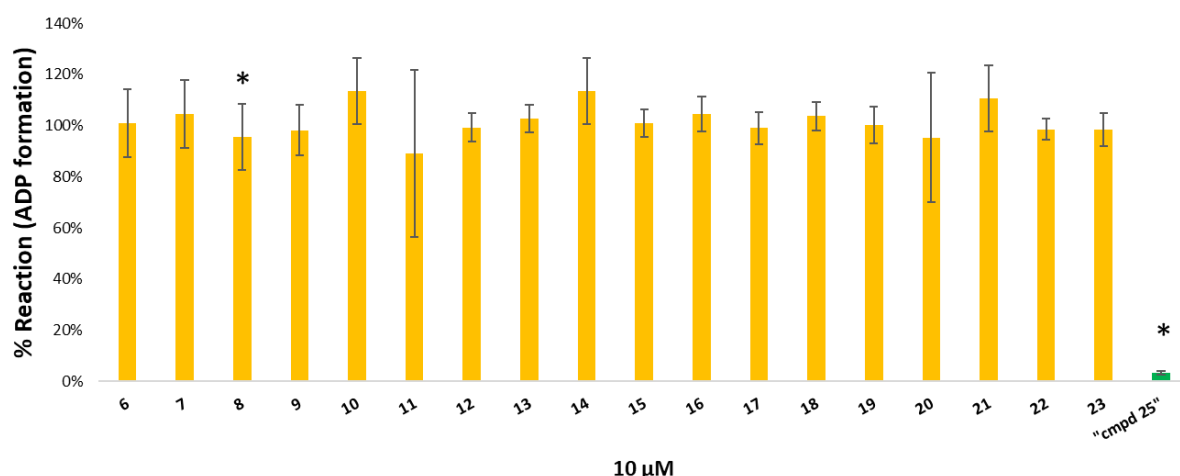


Figure 52. Biochemical evaluation of the synthesized compounds as HK2 inhibitors at 10 μ M. The amount of ADP formed is directly related with the generation of G6P and consequently with the % of reaction occurred in the presence of potential inhibitors (results as average \pm SD, * $p < 0.05$). “Cmpd 25” was able to inhibit the formation of G6P by 97%, contrary to the other molecules that did not significantly reduce the HK2 activity.

3. Conclusions

The synthesis of several molecules from a family of compounds found *in silico* as promising HK2 inhibitors was accomplished in low yields. Unfortunately, biochemical evaluation of the synthesized molecules as modulators of HK2 activity indicates that none of these molecules were able to inhibit hHK2 at 10 μ M.

The synthesis of the known inhibitor “cmpd 25” was successfully achieved despite the low yield, mostly due to the last step, the hydrolysis of the O-glycoside.

Both synthetic routes are suitable for optimization. Improvement of the yields in the first steps would be essential to obtain the final products in larger quantity. Starting with larger amounts of substrates could

also help to manage the diverse synthetic steps in a more profitable way. Since ribose derivatives showed no impact on the activity of HK2, optimization of the synthesis process was not pursued further

4. Experimental Section

4.1. General data for chemicals and instrumentation

All the reagents and solvents were purchased from commercial sources and used without further purification. For reactions requiring dry solvents, mostly standard drying methods were used [251]; DCM and ACN was pre-treated with CaCl_2 and distilled over CaH_2 under vacuum, THF was pre-treated and distilled over sodium wire under vacuum, and DMF was dried over molecular sieves 4A.

HPLC grade acetonitrile was used as eluent in the HPLC systems. As for LC-HRMS systems LC/MS grade solvents were used (Optima™ grade; ThermoFisher Scientific, Waltham, MA, USA). Anhydrous deuterated DMSO ($\text{DMSO-}d_6$, 99.96% D), chloroform (CDCl_3 , 99.80% D), methanol (CD_3OD , 99.95% D) and water (D_2O , >99.90% D) were acquired from Eurisotop, Saint-Aubin, France, and used in all NMR experiments. Anhydrous N_2 (99.999%, Alphagaz™ 1) and anhydrous H_2 (99.999% Alphagaz™ 2) were obtained from Air Liquide Portugal (Algés, Portugal).

Thin-layer chromatography (TLC) was performed on aluminum coated Silica Gel 60Å F254 Merck (0.25 mm) plates, with visualization by UV light ($\lambda = 254 \text{ nm}$) and by carrying with 10% H_2SO_4 in ethanol, KMnO_4 or vanillin stain. Preparative column chromatography procedures were carried out using silica gel 60 Å 70-200 mesh (VWR, Radnor, PA, USA) as the stationary phase. The mobile phase is specified for each particular case.

High performance liquid chromatography (HPLC-UV) analysis and purification of were carried out in an Ultimate 3000 system with a Dionex Ultimate 3000 pump and Ultimate 3000 photodiode module (DAD) with a manual injector (2 mL loop) (Dionex, Sunnyvale, California, US). Phenomenex reverse phase RP-18e (Luna C18(2), 250x4.6 mm, 5 μm beads) analytical columns were used at 1 mL/min. For semi-preparative HPLC experiments reverse phase Phenomenex RP-18e (Luna C18(2), 250x10 mm, 5 μm beads) were used at 3 mL/min. Both types of columns were preceded by a Phenomenex Column-guard (Phenomenex, Sartrouville, France). HPLC systems were operated at constant flow rate using solvent gradients containing acidified water (0.1% v/v formic acid) and/or acetonitrile. Gradient programs are specified below for each type of separation. UV absorbance was monitored at 254 nm. Chromeleon software was used for data acquisition and analysis.

^1H and ^{13}C NMR spectra were recorded on Bruker Avance II+ 300 or 400 spectrometers, equipped with broad-band probes (BBO), operating at 300 MHz (75.5 MHz for ^{13}C -NMR) and 400 MHz (100.6 MHz for ^{13}C -NMR), respectively. Chemical shifts (δ) are reported in ppm downfield from tetramethylsilane (TMS) and coupling constants (J) are reported in Hz; $J_{x,x}$ correspond to geminal or vicinal couplings. Bidimensional COSY (^1H - ^1H correlation spectroscopy), ^1H - ^{13}C HSQC (heteronuclear

single quantum coherence), and ^1H - ^{13}C HMBC (heteronuclear multiple bond correlation) experiments were used to assist on NMR assignments. All spectra were analyzed using Bruker Topspin v3.6.1.

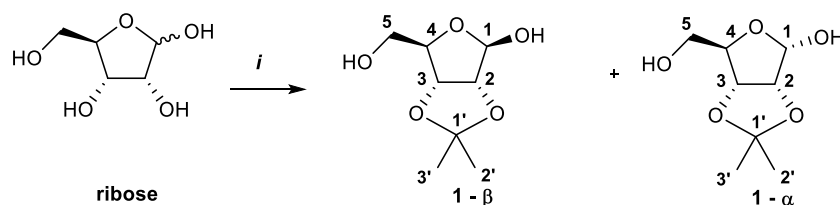
For low resolution MS, samples were analyzed in an LCQ Fleet® ion trap mass spectrometer equipped with an ESI (electrospray ionization) source (Thermo Scientific™, San Jose, CA, USA). The mass spectrometer was operated in the ESI positive and negative ion modes, with the following optimized parameters: ion spray voltage, ± 4.5 kV; capillary voltage, 16/-18 V; tube lens offset, -70/58 V; sheath gas (N_2), 40 arbitrary units; auxiliary gas (N_2), 20 arbitrary units; capillary temperature, 300 °C. MS spectra typically corresponded to an average of 20–35 scans, and were recorded in the full acquisition mode in the range between 100-1000 Da. Tandem mass spectra (MS^n , $n=2-4$) were obtained with an isolation window of 2 m/z units, 25-35% relative collision energy, and with excitation time of 30 msec. LC-MS analyses were performed using an UHPLC Ultimate 3000 RSLC nanosystem (ThermoFisher Scientific) coupled with the above described mass spectrometer. Chromatographic separations were carried out with a Kinetex C18 column (2.6 μm , 150 \times 2.1 mm; Phenomenex, Torrance, CA, USA). The eluent was H_2O (0.1% v/v formic acid) as solvent A and acetonitrile as solvent B, in the following elution conditions: 0-15 min 5-15% B; 15-30 min 15-20% B; 30-35 min 50% B; 35-40 min 100% B; 45 min 5% B, at a flow rate of 0.350 mL min^{-1} . The DAD was monitored in a range from 250 to 370 nm. Data acquisition and processing were performed using the Xcalibur 2.2 software.

High-resolution mass spectra were acquired with a quadrupole time-of-flight (QTOF) Impact II mass spectrometer with an electrospray ion source (Bruker Daltonics, Billerica, MA, USA). Analyses were performed in positive or negative electrospray ionization (ESI+/ESI-) mode operating in the high-resolution mode over the m/z range of 50–1000 Da with an acquisition rate of 3 Hz. The optimized mass spectrometer parameters were as follows: ion spray voltage, 2.5 kV; end plate offset, -500 V; nebulizer gas (N_2), 2.8 bars; dry gas (N_2), 8 L/min; dry heater, 200°C. Sodium formate (10 mM) was used for internal calibration on the high-precision calibration mode (HPC). Data were recorded and processed using Bruker DataAnalysis 4.6 software.

In the structures represented below, atom numbering of each compound starts always with the anomeric carbon as C1, to facilitate comparison between structures. In each instance, unless otherwise noticed, the designations according to the IUPAC nomenclature are included for the β anomers; these do not reflect the numbering system shown in the Schemes.

4.2. Synthesis of Ribose-derivatives

4.2.1. Synthesis of compound 1 ($\alpha + \beta$ anomers), (3a*R*,4*R*,6*R*,6a*R*)-6-(hydroxymethyl)-2,2-dimethyltetrahydro-2*H*-furo[3,4-*d*][1,3]dioxol-4-ol



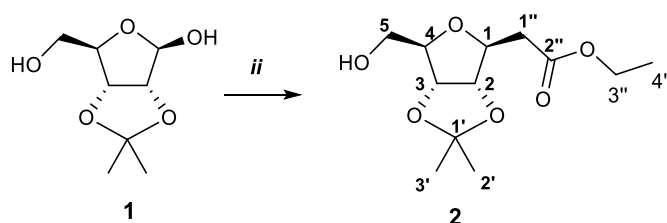
Scheme 8. Preparation of compound **1** (anomeric mixture) from ribose: structures and numbering. (i) 2,2-dimethoxypropane, *p*-TsOH, acetone, r.t., 1.5 h (quant.).

To a solution of D-ribose (2g, 13.3 mmol) in acetone (10 mL), *p*-toluenesulfonic acid (95 mg, 0.665 mmol) and 2,2-dimethoxypropane (2.45 mL, 20 mmol, 1.5 eq) were added. The reaction was stirred for 1.5 h at r.t., and then neutralized with solid NaHCO₃. The suspension was filtered, and the filtrate was concentrated at reduced pressure to give the a colorless syrup of product **1** as an anomeric mixture (quantitative yield) [241].

NMR data for the β anomer: ¹H NMR (400 MHz, CDCl₃): δ = 1.28 (s, C2', 3H), 1.44 (s, C3', 3H), 3.66 (m, H5a, H5b, 2H), 4.34 (br t, H4, 1H), 4.53 (d, H2, $J_{2,3}$ = 5.9 Hz, 1H), 4.76 (d, H3, 1H), 5.36 (s, H1, 1H).

¹³C NMR (100.6 MHz, CDCl₃): δ = 24.75 (C2'), 26.40 (C3'), 63.47 (C5), 81.68(C3), 86.74(C2), 87.63 (C4), 102.70 (C1), 112.23 (C1').

4.2.2. Synthesis of compound 2, ethyl [(3a*S*,4*S*,6*R*,6a*R*)-6-(hydroxymethyl)-2,2-dimethyltetrahydro-2*H*-furo[3,4-*d*][1,3]dioxol-4-yl]acetate



Scheme 9. Preparation of compound **2** from compound **1**: structures and numbering . ii) Ethyl (triphenylphosphoranylidene)acetate, ACN, 80 °C, 2.5 h (47%).

The Wittig reagent, ethyl (triphenylphosphoranylidene) acetate was prepared according to Dong *et al.* [252]. Ethyl chloroacetate (2.5 mL, 24 mmol, 1.3 eq) diluted in EtOAc (4 mL) was added dropwise into

a solution of triphenylphosphine (5g, 19 mmol) in EtOAc (6 mL). The reaction mixture was stirred overnight at r.t. The white precipitate was filtered off, washed with ethyl ether and dried under vacuum to give **(2-ethoxy-2-oxoethyl)(triphenyl)phosphonium chloride** (75%).

^1H NMR (400 MHz, CDCl_3): δ = 1.06 (t, J = 7.2 Hz, 3H), 4.03 (q, J = 7.2 Hz, 2H), 5.70 (m, 2H), 7.64-7.69 (m, 6H), 7.77 (t, J = 8.0 Hz, 3H), 7.88-7.90 (m, 6H).

^{13}C NMR (100.6 MHz, CDCl_3): δ = 13.75, 32.54 (d, J = 56.2 Hz), 62.88, 118.04 (d, J = 90.0 Hz), 130.28, 130.15, 133.90, 134.00, 135.22, 135.25, 164.64.

To the solution of the phosphonium salt (2.5 g, 6.5 mmol) in DCM (8.5 mL) at 0°C, cold 1M NaOH (6.5 mL, 1 eq) was added. The reaction mixture was stirred for 15 min. The organic layer was dried over Mg_2SO_4 and concentrated to give the **ethyl (triphenylphosphoranylidene) acetate** as a white solid mixture of *Z* and *E* isomers (88%).

^1H NMR (400 MHz, CDCl_3): δ = 0.59 (t, J = 7.2 Hz, 1H), 1.12 (t, J = 7.2 Hz, 2H), 2.68 (d, J = 21.7 Hz, 0.4H), 2.88 (d, J = 22.1 Hz, 0.6H), 3.75 (q, J = 7.2 Hz, 0.7H), 3.95 (q, J = 7.2 Hz, 1.3H), 7.31-7.33 (m, 6H), 7.39 (t, J = 6.4 Hz, 3H), 7.53-7.59 (m, 6H).

^{13}C NMR (100.6 MHz, CDCl_3): δ = 14.12, 14.80, 29.69 (d, J = 29.0 Hz), 30.16 (d, J = 119.6 Hz), 57.14, 57.57, 127.41 (d, J = 91.63 Hz), 127.48 (d, J = 91.63 Hz), 128.14, 128.24, 128.48, 131.57, 131.63, 131.65, 132.54, 132.64, 171.23, 171.38.

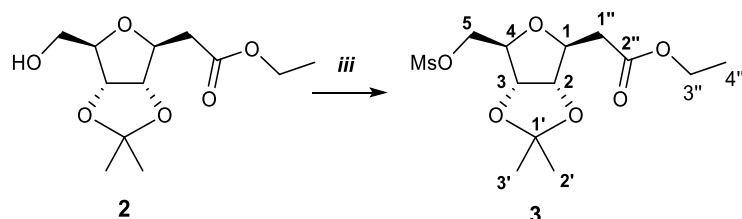
Without further purification, compound **1** (2.8 g, 14.7 mmol) was coevaporated with toluene (3 X 7 mL) and dissolved in acetonitrile (20 mL). Ethyl (triphenylphosphoranylidene) acetate (5.63 g, 16.2 mmol, 1.1 eq) was added and the reaction mixture was stirred under reflux for 2.5h. After cooling to r.t., the solution was concentrated, and the residue was purified by silica gel chromatography (eluent: EtOAc:Hex 1:2→3:1 to afford compound **2** as the β -anomer in 47% yield (colorless syrup) [240].

^1H NMR (400 MHz, CDCl_3): δ = 1.23 (t, C4'' , $J_{4'',3''}$ = 7.1 Hz, 3H), 1.30 (s, C2' , 3H), 1.50 (s, C3' , 3H), 2.58 (dd, H1''a , $J_{1''a,1}$ = 6.8 Hz, $J_{1''a,1''b}$ = 15.7 Hz, 1H), 2.69 (dd, H1''b , $J_{1''b,1}$ = 4.9 Hz, 1H), 2.84 (s, OH, 1H), 3.57-3.63 (m, H5a), 3.76 (dd, H5b , $J_{5b,4}$ = 2.4 Hz, $J_{5b,5a}$ = 12.1 Hz, 1H), 4.03 (br q, H4 , 1H), 4.13 (q, H3'' , 2H), 4.22 (ddd, H1 , $J_{1,2}$ = 4.9 Hz, 1H), 4.49 (dd, H2 , $J_{2,3}$ = 6.5 Hz, 1H), 4.68 (dd, H3 , $J_{3,4}$ = 3.5 Hz, 1H).

^{13}C NMR (100.6 MHz, CDCl_3): δ = 14.19 (C4''), 25.51 (C2'), 27.46 (C3'), 37.89 (C1''), 60.98 (C3''), 62.71 (C5), 80.81 (C1), 81.64 (C3), 84.03 (C2), 84.84 (C4), 114.43 (C1'), 170.97 (C2'').

HRMS calcd for $[\text{C}_{12}\text{H}_{20}\text{O}_6 + \text{H}]^+$: 261.1333. Found: 261.1334.

4.2.3. Synthesis of compound 3, ethyl [(3a*S*,4*S*,6*R*,6a*R*)-6-[(methanesulfonyl)oxy]methyl]-2,2-dimethyltetrahydro-2*H*-furo[3,4-*d*][1,3]dioxol-4-yl]acetate



Scheme 10. Preparation of compound **3** from compound **2**: structures and numbering. iii) MsCl , pyridine, r.t., 6h, (93%).

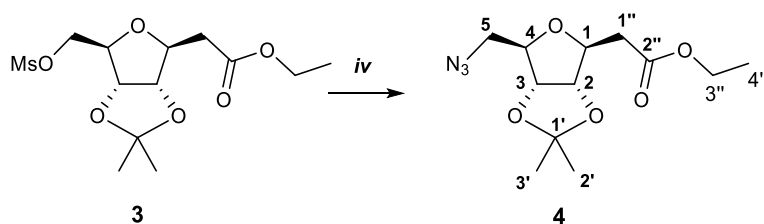
Compound **2** (1.7 g, 6.6 mmol) was co-evaporated (3 X 7 mL) and dissolved in pyridine (30 mL). Methanesulfonylchloride (600 μL , 7.9 mmol 1.2 eq) was added and the reaction mixture was stirred for 6h. The solvent was removed at reduced pressure and the residue was dissolved in EtOAc (50 mL) and washed with H_2O (25 mL), sat. aq. NaHCO_3 (3 X 25 mL) and brine (25 mL). After drying over Mg_2SO_4 the residue was concentrated under vacuum to give the crude product **3** as a colorless oil (94% - crude).

^1H NMR (500 MHz, CDCl_3): δ = 1.24 (t, $\text{H}_{4''}$, $J_{4'',3''} = 7.2$ Hz, 3H), 1.32 (s, $\text{H}_{2'}$, 3H), 1.52 (s, $\text{H}_{3'}$, 3H), 2.57 (dd, $\text{H}_{1''\text{a}}$, $J_{1''\text{a},1} = 7.1$ Hz, $J_{1''\text{a},1''\text{b}} = 15.8$ Hz, 1H), 2.66 (dd, $\text{H}_{1''\text{b}}$, $J_{1''\text{b},1} = 5.3$ Hz, 1H), 3.03 (s, Ms CH_3 , 3H), 4.11-4.16 (m, H_4 , $\text{H}_{3''}$, 3H), 4.27-4.35 (m, H_1 , 1H), 4.28 (dd, $\text{H}_{5\text{a}}$, $J_{5\text{a},5\text{b}} = 11.2$ Hz, $J_{5\text{a},4} = 4.7$ Hz, 1H), 4.33 (dd, $\text{H}_{5\text{b}}$, $J_{5\text{b},4} = 3.5$ Hz, 1H), 4.53 (dd, H_2 , $J_{2,1} = 4.2$ Hz, $J_{2,3} = 6.7$ Hz 1H), 4.61 (dd, H_3 , $J_{3,4} = 4.5$ Hz, 1H).

^{13}C NMR (100.6 MHz, CDCl_3): δ = 14.23 ($\text{C}_{4''}$), 25.51 ($\text{C}_{2'}$), 27.42 ($\text{C}_{3'}$), 37.64 (Ms CH_3), 38.35 ($\text{C}_{1''}$), 60.90 ($\text{C}_{3''}$), 69.24 (C_5), 81.23 (C_1, C_3), 81.99 (C_4), 84.17 (C_2), 115.09 ($\text{C}_{1'}$), 170.28 ($\text{C}_{2''}$).

HRMS calcd for $[\text{C}_{13}\text{H}_{22}\text{O}_8\text{S} + \text{H}]^+$: 339.1108. Found: 339.1105.

4.2.4. Synthesis of compound 4, ethyl [(3a*S*,4*S*,6*R*,6a*R*)-6-(azidomethyl)-2,2-dimethyltetrahydro-2*H*-furo[3,4-*d*][1,3]dioxol-4-yl]acetate



Scheme 11. Preparation of compound **4**: structures and numbering. (iv) NaN_3 , DMF, 75 $^\circ\text{C}$, 1.5h (93%).

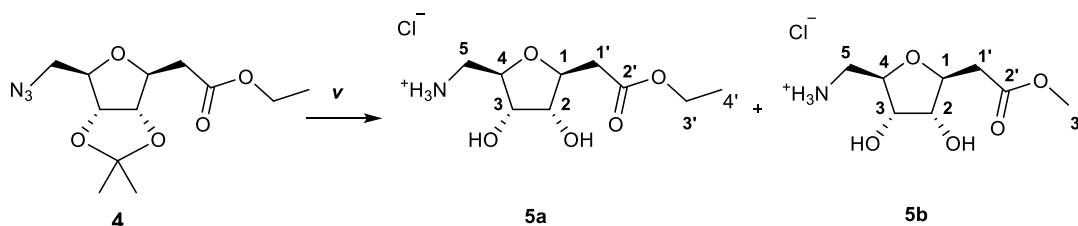
Without further purification, product **3** (2.2 g, 6.5 mmol) was co-evaporated with toluene (3 X 7 mL) and dissolved in DMF (20 mL). Sodium azide (1 g, 16.2 mmol, 2.5 eq) was added and the reaction mixture was stirred at 75°C for 1.5h. The reaction mixture was then allowed to cool down to r.t. and diluted with water (10 mL). The aqueous phase was extracted with EtOAc (3 X 25 mL). The combined organic phases were washed with brine (20 mL) and dried over Mg₂SO₄. The solvent was removed at reduce pressure to afford compound **4** as a slightly yellow oil with 93% yield (crude).

¹H NMR (400 MHz, CDCl₃): δ= 1.21 (t, H4'', J_{4'',3''} = 7.1 Hz, 3H), 1.29 (s, H2', 3H), 1.49 (s, H3', 3H), 2.59 (dd, H1''a, J_{1''a,1} = 6.9 Hz, J_{1''a,1''b} = 15.8 Hz, 1H), 2.65 (dd, H1''b, J_{1''b,1} = 5.5 Hz, 1H), 3.29 (dd, H5a, J_{5a,5b} = 13.1 Hz, J_{5a,4} = 4.8 Hz, 1H), 3.48 (dd, H5b, J_{5b,4} = 3.8 Hz, 1H), 4.03 (br q, H4, 1H), 4.11 (2 q H3'' (two rotamers), 2H), 4.24 (ddd, H1, J_{1,2} = 4.1 Hz, 1H), 4.50-4.55 (m, H2, H3, 2H).

¹³C NMR (100.6 MHz, CDCl₃): δ= 14.14 (C4''), 25.47 (C2'), 27.38 (C3'), 38.23 (C1''), 52.24 (C5), 60.74 (C3''), 80.87 (C1), 82.02 (C3), 83.07 (C4), 84.19 (C2), 114.89 (C1'), 170.31 (C2'').

HRMS calcd for [C₁₃H₂₂N₃O₈ + H]⁺: 286.1397. Found: 286.1397.

4.2.5. Synthesis of compound **5** (a + b), ethyl [(2*S*,3*R*,4*S*,5*R*)-5-(aminomethyl)-3,4-dihydroxyoxolan-2-yl]acetate + methyl [(2*S*,3*R*,4*S*,5*R*)-5-(aminomethyl)-3,4-dihydroxyoxolan-2-yl]acetate



Scheme 12. Preparation of compound(s) **5** (shown as the hydrochloride(s)) from compound **4**: structure and numbering. (v) a) H₂, Pd/C, HCl/EtOH, 18h, r.t., 1 atm; b) MeOH, r.t. (79%).

A solution of compound **4** (1.7 g, 5.8 mmol) in ethanol (23 mL) containing 3 M HCl (2.3 mL) was degassed, Pd/C (10%) (162 mg) was added, and the mixture was degassed once more. The reaction was stirred overnight at r.t. under H₂, 1 atm. The reaction mixture was filtered over celite and the filtrate was concentrated under reduced pressure to give the crude of compounds **5** as a slightly yellow oil (79% - considering a 2:1 ethyl:methyl acetate ratio).

¹H NMR (400 MHz, MeOD): δ= 1.26 (t, H4', J_{4',3'} = 7.1 Hz, 2H), 2.58 (m, H1'a, 1H), 2.74 (m, H1'b, 1H), 2.98 (dd, H5a, J_{5a,4} = 9.0 Hz, J_{5a,5b} = 13.1 Hz, 1H), 3.25 (dd, H5b, J_{5b,4} = 2.7 Hz, 1H), 3.69 (s, H3'', 1H), 3.88-3.89 (m, H2, H3, 2H), 3.92-3.95 (m, H4, 1H), 4.12-4.18 (m, H1, H3', 2.5H).

¹³C NMR (100.6 MHz, MeOD): δ= 14.45 (C4'), 38.97, 39.25 (C1'), 43.31 (C5), 52.33 (C3''), 61.87 (C3'), 73.62 (C3), 75.24 (C2), 80.68 (C4), 81.92 (C1), 172.84, 173.24 (C2').

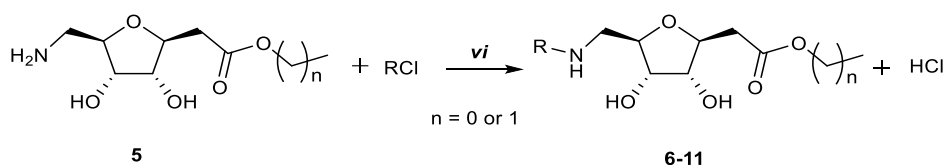
HRMS calcd for $[C_9H_{17}NO_5 + H]^+$: 220.1179. Found: 220.1181.

HRMS calcd for $[C_8H_{15}NO_5 + H]^+$: 206.1023. Found: 206.1024.

4.2.6. General procedure for the formation of the amide/sulfonamide at C5 (6-11)

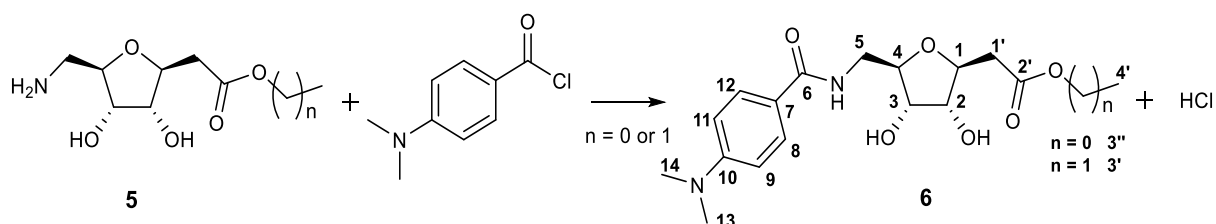
Compound(s) **5** (0.25 mmol, 1 eq) was added to a solution of Et_3N (3 eq) in dry DCM (750 μ L), followed by the acyl/sulfonyl chloride (0.25 mmol, 1 eq) (**Scheme 13-Scheme 19**). The reaction mixture was stirred for 18-48h at r.t. to 40°C.

When the total consumption of substrate **5** by TLC (eluent: EtOAc) and product formation were observed, the reaction was quenched with water (2.5 mL). The layers were separated and the aqueous layer was extracted with EtOAc (3 X 10mL). The combined organic layers were washed with sat. aq. $NaHCO_3$ (15 mL), dried over $MgSO_4$ and concentrated under vacuum. The crude mixture was purified by preparative column chromatography (Eluent: EtOAc), to give the expected product in 20-51% yield.



Scheme 13. General preparation scheme of compounds **6-11**. (vi) Acyl/Sulphonyl Chloride ($R-Cl$), Et_3N , DCM, r.t. – 40 °C, 24h, 20-67%.

4.2.6.1. Synthesis of compound(s) **6**, ethyl/methyl {(2S,3R,4S,5R)-3,4-dihydroxy-5-[(4-(dimethylamino)benzamido)methyl]oxolan-2-yl}acetate



Scheme 14. Preparation of compound(s) **6**, from compound(s) **5** and 4-(dimethylamino)benzoyl chloride (57%).

Compound(s) **5**: 64 mg, 0.25 mmol; Acyl Chloride: 4-(dimethylamino)benzoyl chloride: 46 mg, 0.25 mmol; Triethylamine: 105 μ L; DCM: 1mL; Temperature: r.t.; Reaction time: 18h; Yield: 57%.

1H NMR (300 MHz, $CDCl_3$): δ = 1.21 (t, $H_{4'}$, $J_{4',3'} = 7.1$ Hz, 2H), 2.44-2.53 (m, $H_{1'a}$, 1H), 2.62-2.70 (m, $H_{1'b}$, 1H), 2.97 (s, H_{13} , H_{14} , 6H), 3.47-3.54 (m, H_{5a} , 1H), 3.65 (s, $H_{3''}$, 1H), 3.75-3.86 (m, H_{5b} , H_2 ,

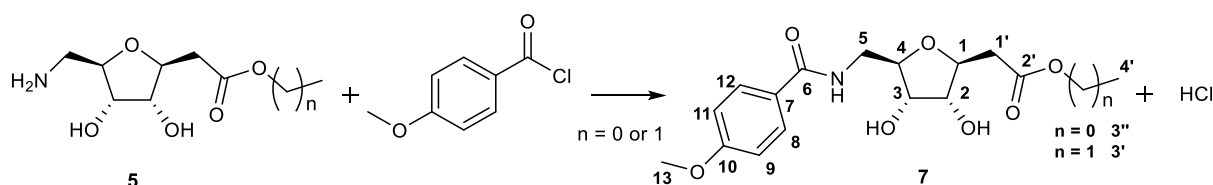
2H), 3.93-3.97 (m, H3, H4, 2H), 4.06-4.15 (m, H1, H3', 2.5H), 4.90 (s, OH, 1H), 6.59 (d, H9, H11, $J = 8.8$ Hz, 2H), 6.89 (br s, NH, 1H), 7.69 (d, H8, H12, 2H).

^{13}C NMR (75.5 MHz, CDCl_3): $\delta = 14.19$ (C4'), 38.09, 38.33 (C1'), 40.08 (C13, C14), 41.04 (C5), 51.88 (C3''), 60.80 (C3'), 71.66 (C3), 74.37, 74.40 (C2), 79.76, 79.85 (C1), 82.50, 82.55 (C4), 111.00 (C9, C11), 120.27 (C10), 128.74 (C8, C12), 152.60 (C7), 168.48 (C6), 171.23, 171.63 (C2').

HRMS calcd for $[\text{C}_{18}\text{H}_{26}\text{N}_2\text{O}_6 + \text{H}]^+$: 367.1864. Found 367.1863.

HRMS calcd for $[\text{C}_{17}\text{H}_{24}\text{N}_2\text{O}_6 + \text{H}]^+$: 353.1707. Found 353.1704.

4.2.6.2. Synthesis of compound(s) 7, ethyl/methyl {(2S,3R,4S,5R)-3,4-dihydroxy-5-[(4-methoxybenzamido)methyl]oxolan-2-yl}acetate



Scheme 15. Preparation of compound(s) 7, from compound(s) 5 and 4-(methoxy)benzoyl chloride (51%).

Compound(s) 5: 64 mg, 0.25 mmol; Acyl Chloride: 4-(methoxy)benzoyl chloride: 46 mg, 0.25 mmol; Triethylamine: 105 μL ; DCM: 1mL; Temperature: r.t.; Reaction time: 18h. Yield: 51%.

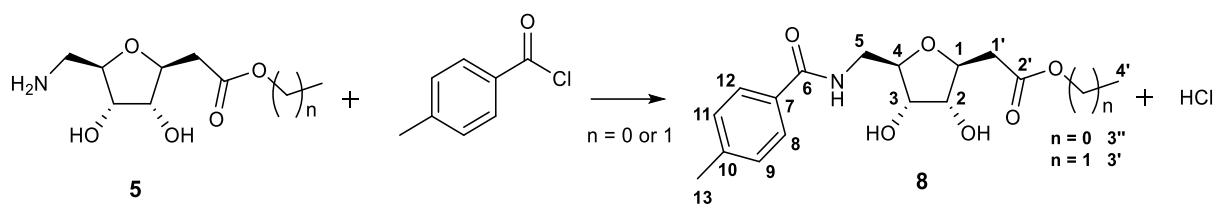
^1H NMR (300 MHz, CDCl_3): $\delta = 1.24$ (t, H4', $J_{4',3'} = 7.2$ Hz, 2H), 2.47-2.56 (m, H1'a, 1H), 2.63-2.72 (m, H1'b, 1H), 3.52-3.56 (m, H5a, 1H), 3.64 (s, H3'', 1H), 3.83 (s, H13, 3H), 3.86-3.96 (m, H2, H5'b, 2H), 3.96-4.04 (m, H3, H4, 2H), 4.09-4.19 (m, H1, H3', 2.5H), 4.53 (s, OH, 1H), 6.89 (d, H9, H11, $J = 8.8$ Hz, 2H), 6.94 (s, NH, 1H) 7.79 (d, H8, H12, 2H).

^{13}C NMR (75.5 MHz, CDCl_3): $\delta = 14.27$ (C4'), 38.20 (C1'), 41.11 (C5), 51.97 (C3''), 55.53 (C13), 61.01 (C3'), 71.67 (C3), 74.44 (C2), 80.02 (C1), 82.45 (C4), 113.86 (C9, C11), 126.05 (C7), 129.19 (C8, C12), 162.53 (C10), 168.15 (C6), 171.38 (C2').

HRMS calcd for $[\text{C}_{17}\text{H}_{23}\text{NO}_7 + \text{H}]^+$: 354.1547. Found 354.1546.

HRMS calcd for $[\text{C}_{16}\text{H}_{21}\text{NO}_7 + \text{H}]^+$: 340.1391. Found 340.1387.

4.2.6.3. Synthesis of compound(s) 8, ethyl/methyl {(2S,3R,4S,5R)-3,4-dihydroxy-5-[(4-methylbenzamido)methyl]oxolan-2-yl}acetate



Scheme 16. Preparation of compound(s) 8, from compound(s) 5 and *p*-toluoyl chloride (67%).

Compound(s) 5: 128 mg, 0.5 mmol; Acyl Chloride: *p*-toluoyl chloride: 66 μ L, 0.5 mmol; Triethylamine: 210 μ L; DCM: 1.5mL; Temperature: r.t.- 40°C; Reaction time: 24h; Yield: 67%.

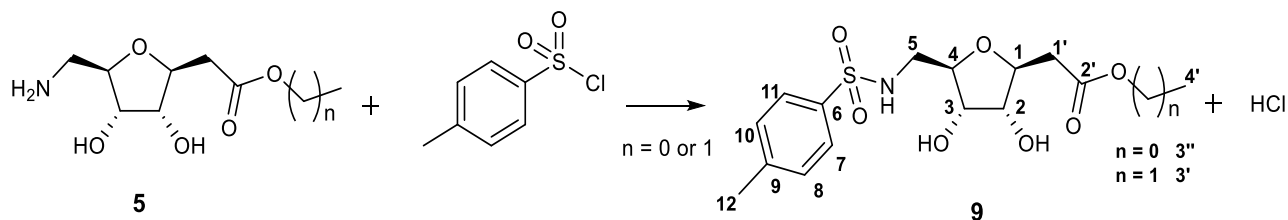
^1H NMR (300 MHz, CDCl_3): δ = 1.23 (t, $\text{H}_{4'}$, $J_{4',3'} = 7.1\text{ Hz}$, 2H), 2.36 (s, H_{13} , 3H), 2.47-2.56 (m, $\text{H}_{1'a}$, 1H), 2.63-2.72 (m, $\text{H}_{1'b}$, 1H), 3.51-3.59 (m, H_{5a} , $J_{5a,5b} = 14.2\text{ Hz}$, 1H), 3.66 (s, $\text{H}_{3''}$, 1H), 3.80-3.90 (m, H_2 , $\text{H}_{5'b}$, 2H), 3.95-4.02 (m, H_3 , H_4 , 2H), 4.08-4.17 (m, H_1 , $\text{H}_{3'}$, 2.5H), 4.67 (s, OH, 1H), 7.04 (br, NH, 1H), 7.19 (d, H_8 , H_{12} , $J = 8.1\text{ Hz}$, 2H), 7.70 (d, H_9 , H_{11} , 2H).

^{13}C NMR (75.5 MHz, CDCl_3): δ = 14.24 ($\text{C}_{4'}$), 21.56 (C_{13}), 38.01, 38.24 ($\text{C}_{1'}$), 41.14 (C_5), 52.01 ($\text{C}_{3''}$), 60.99 ($\text{C}_{3'}$), 71.63 (C_3), 74.38, 74.41 (C_2), 79.95, 80.02 (C_1), 82.35, 82.38 (C_4), 127.31 (C_9 , C_{11}), 129.32 (C_8 , C_{12}), 130.95 (C_7), 142.37 (C_{10}), 168.61 (C_6), 171.37, 171.76 ($\text{C}_{2'}$).

HRMS calcd for $[\text{C}_{17}\text{H}_{23}\text{NO}_6 + \text{H}]^+$: 338.1598. Found 338.1599.

HRMS calcd for $[\text{C}_{16}\text{H}_{21}\text{NO}_6 + \text{H}]^+$: 324.1442. Found 324.1442.

4.2.6.4. Synthesis of compound(s) 9, ethyl/methyl [(2S,3R,4S,5R)-3,4-dihydroxy-5-[[4-methylbenzene-1-sulfonyl]amino]methyl]oxolan-2-yl]acetate



Scheme 17. Preparation details of compound 9, from compound 5 and *p*-toluenesulfonyl chloride (20%).

Compound(s) **5**: 128 mg, 0.5 mmol; Sulfonyl Chloride: *p*-toluenesulfonyl chloride (97%): 96 mg, 0.5 mmol; Triethylamine: 210 μ L; DCM: 1.5mL; Temperature: r.t. to 40°C; Reaction time: 24h; Yield: 20%.

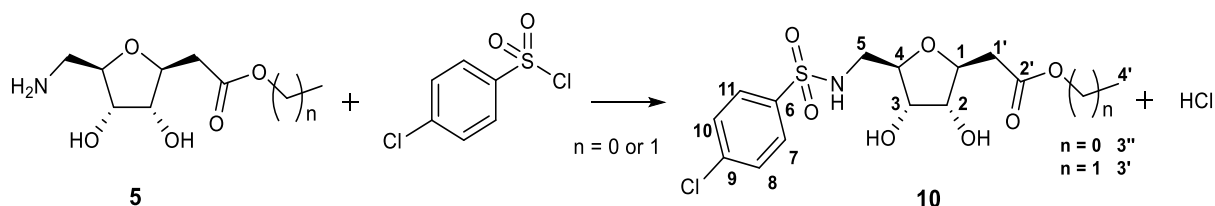
^1H NMR (300 MHz, CDCl_3): δ = 1.25 (t, $\text{H}_{4'}$, $J_{4',3'} = 7.1$ Hz, 2H), 2.41 (s, H_{12} , 3H), 2.55-2.77 (m, $\text{H}_{1'a}$, $\text{H}_{1'b}$, 2H), 3.06-3.09 (m, H_{5a} , 1H), 3.12-3.20 (m, H_{5b} , 1H), 3.69 (s, $\text{H}_{3''}$, 1H), 3.89-3.97 (m, H_2 , H_4 , 2H), 4.05-4.18 (m, H_1 , H_3 , $\text{H}_{3'}$, 3.5H), 5.53 (br, NH, 1H), 7.30 (d, H_8 , H_{10} , $J = 8.2$ Hz, 2H), 7.72 (d, H_7 , H_{11} , 2H).

^{13}C NMR (75.5 MHz, CDCl_3): δ = 14.23 ($\text{C}_{4'}$), 21.65 (C_{12}), 37.67, 37.93 ($\text{C}_{1'}$), 44.44 (C_5), 52.25 ($\text{C}_{3''}$), 61.26 ($\text{C}_{3'}$), 71.85, 71.90 (C_3), 74.35, 74.40 (C_2), 79.33, 79.37 (C_1), 82.18, 82.22 (C_4), 127.13 (C_7 , C_{11}), 129.93 (C_8 , C_{10}), 136.60 (C_6), 143.75 (C_9), 171.82, 172.18 (C_2').

HRMS calcd for $[\text{C}_{16}\text{H}_{23}\text{NO}_7\text{S} + \text{H}]^+$: 374.1268. Found 374.1272.

HRMS calcd for $[\text{C}_{15}\text{H}_{21}\text{NO}_7\text{S} + \text{H}]^+$: 360.1111. Found 360.1113.

4.2.6.5. Synthesis of compound(s) **10**, ethyl/methyl [(2*S*,3*R*,4*S*,5*R*)-5-[(4-chlorobenzene-1-sulfonyl)amino]methyl]-3,4-dihydroxyoxolan-2-yl]acetate



Scheme 18. Preparation of compound(s) **10**, from compound(s) **5** and 4-chlorobenzenesulfonyl chloride (38%).

Compound(s) **5**: 128 mg, 0.5 mmol; Sulfonyl Chloride: 4-chlorobenzenesulfonyl chloride (97%): 109 mg, 0.5 mmol; Triethylamine: 210 μ L; DCM: 1.5mL; Temperature: r.t.; Reaction time: 24h; Yield: 38%

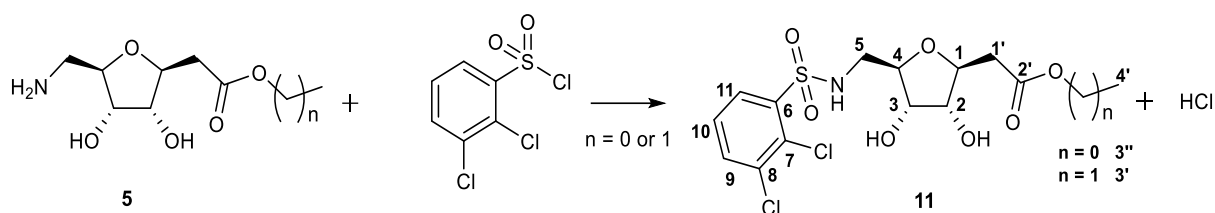
^1H NMR (400 MHz, CDCl_3): δ = 1.22 (t, $\text{H}_{4'}$, $J_{4',3'} = 7.1$ Hz, 2H), 2.50-2.58 (m, $\text{H}_{1'a}$, 1H), 2.63-2.71 (m, $\text{H}_{1'b}$, 1H), 3.04-3.16 (m, H_{5a} , H_{5b} , 2H), 3.66 (s, $\text{H}_{3''}$, 1H), 3.88-3.95 (m, H_2 , H_4 , 2H), 4.05-4.13 (m, H_1 , H_3 , $\text{H}_{3'}$, 3.5H), 6.04 (br, NH, 1H), 7.45 (d, H_7 , H_{11} , $J = 8.6$ Hz, 2H), 7.77 (d, H_8 , H_{10} , 2H).

^{13}C NMR (100.6 MHz, CDCl_3): δ = 14.16 ($\text{C}_{4'}$), 37.70, 37.94 ($\text{C}_{1'}$), 44.56 (C_5), 52.22 ($\text{C}_{3''}$), 61.24 ($\text{C}_{3'}$), 71.91 (C_3), 74.27 (C_2), 79.31, 79.38 (C_1), 80.02, 80.07 (C_4), 128.54 (C_8 , C_{10}), 129.56 (C_7 , C_{11}), 138.15 (C_6), 139.28 (C_9), 171.90, 172.29 (C_2').

HRMS calcd for $[\text{C}_{15}\text{H}_{20}^{35}\text{ClNO}_7\text{S} + \text{H}]^+$: 394.0722. Found 394.0722.

HRMS calcd for $[\text{C}_{14}\text{H}_{18}^{35}\text{ClNO}_7\text{S} + \text{H}]^+$: 380.0565. Found 380.0564.

4.2.6.6. Synthesis of compound(s) **11** ethyl/methyl [(2*S*,3*R*,4*S*,5*R*)-5-[(2,3-dichlorobenzene-1-sulfonyl)amino]methyl]-3,4-dihydroxyoxolan-2-yl]acetate



Scheme 19. Preparation of compounds **11**, from compound(s) **5** and 2,3-dichlorobenzenesulfonyl chloride (51%).

Compound(s) **5**: 104 mg, 0.4 mmol; Sulfonyl Chloride: 2,3-dichlorobenzenesulfonyl chloride 150 mg, 0.6 mmol; Triethylamine: 167 μ L; DCM: 1.2mL; Temperature: r.t.; Reaction time: 24h; Yield: 51%.

^1H NMR (400 MHz, CDCl_3): δ = 1.26 (t, $\text{H}_{4'}$, $J_{4',3'} = 7.1$ Hz, 2.5H), 2.59-2.71 (m, $\text{H}_{1'a}$, $\text{H}_{1'b}$, 2H), 3.04-3.20 (m, H_{5a} , H_{5b} , 2H), 3.71 (s, $\text{H}_{3''}$, 0.5 H) 3.89 (br q, H_4 , 1H), 3.98 (br t, H_2 , 1H), 4.06-4.19 (m, H_1 , H_3 , $\text{H}_{3'}$, 3.7H), 5.96 (br, NH, 1H), 7.36 (t, H_{10} , $J_{10,9} = J_{10,11} = 8.0$ Hz, 1H), 7.67 (d, H_9 , 1H), 8.00 (d, H_{11} , 1H).

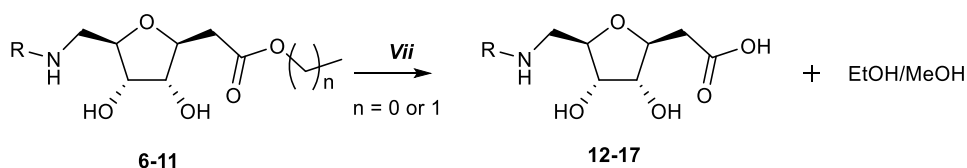
^{13}C NMR (100.6 MHz, CDCl_3): δ = 14.18 ($\text{C}_{4'}$), 37.55, 37.81 ($\text{C}_{1'}$), 44.50 (C_5), 61.25 ($\text{C}_{3'}$), 71.81 (C_3), 74.21 (C_2), 79.56 (C_1), 81.88 (C_4), 127.62 (C_{10}), 129.66 (C_{11}), 130.01 (C_7), 134.59 (C_9), 135.58 (C_8), 139.08 (C_6), 171.66, 172.04 ($\text{C}_{2'}$).

HRMS calcd for $[\text{C}_{15}\text{H}_{19}^{35}\text{Cl}_2\text{NO}_7\text{S} + \text{H}]^+$: 428.0332. Found 428.0314.

HRMS calcd for $[\text{C}_{14}\text{H}_{17}^{35}\text{Cl}_2\text{NO}_7\text{S} + \text{H}]^+$: 414.0176. Found 414.0176.

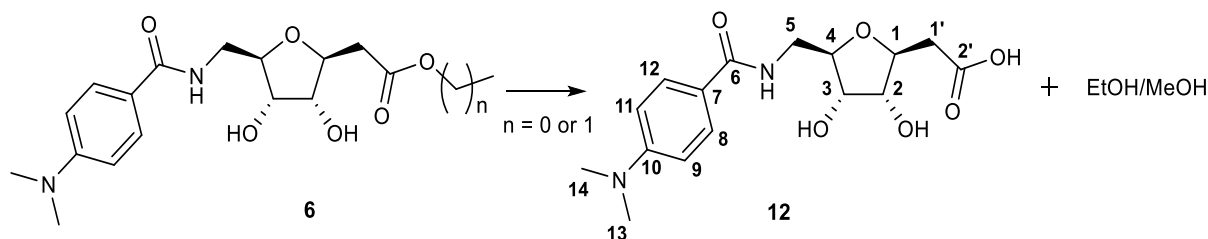
4.2.7. General procedure for the hydrolysis of esters (**6-11**) to the corresponding carboxylic acids (**12-17**)

To a solution of each ester (**6-11**, 1 eq) in ACN, LiOH (1M, 3 eq) was added. The reaction mixture was stirred for 45 min at r.t., and then quenched with HCl (3.6 eq). The reaction was allowed to cool down and the mixture was extracted with EtOAc (3 x 10 mL). The combined organic phases were washed with brine, dried over Mg_2SO_4 and concentrated to give the respective acid as a white powder in 66-99% yield (**Scheme 20-Scheme 26**).



Scheme 20. General preparation scheme of compounds **12-17**. vii) 1) LiOH, ACN, 45 min, r.t.; 2) HCl; (66-99%).

4.2.7.1. Synthesis of compound 12, [(2S,3R,4S,5R)-3,4-dihydroxy-5-([4-(dimethylamino)benzamido]methyl)oxolan-2-yl]acetic acid



Scheme 21. Preparation of compound **12**, structures of the precursor(s) (**6**) and product (99%).

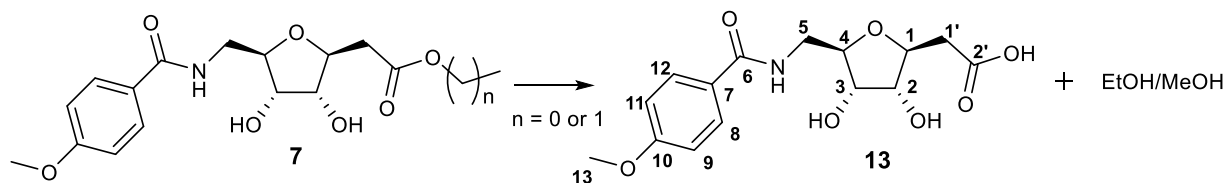
Compound(s) **6**: 85 mg, 0.23 mmol; ACN: 2.3 mL; LiOH (1 M): 735 μ L; HCl (1 M): 854 μ L; Yield: 99%.

^1H NMR (400 MHz, CDCl_3): δ = 2.48 (dd, H1'a, $J_{1'a,1} = 8.2$ Hz, $J_{1'a,1'b} = 15.5$ Hz, 1H), 2.68 (dd, H1'b, $J_{1'b,1} = 4.2$ Hz, 1H), 3.00 (s, H13, H14, 6H), 3.56 (d, H5, $J_{5,4} = 4.7$ Hz, 2H), 3.82 (t, H2, $J_{2,1} = J_{2,3} = 5.6$ Hz, 1H), 3.90-3.96 (m, H3, H4, 2H), 4.09-4.14 (m, H1, 1H), 6.68 (d, H9, H11, $J = 8.9$ Hz, 2H), 7.72 (d, H8, H12, 2H).

^{13}C NMR (100.6 MHz, CDCl_3): δ = 38.33 (C1'), 39.53 (C13, C14), 41.57 (C5), 72.13 (C3), 74.36 (C2), 79.65 (C1), 82.81 (C4), 111.00 (C9, C11), 120.29 (C10), 128.71 (C8, C12), 152.93 (C7), 169.29 (C6), 173.68 (C2').

HRMS calcd for $[\text{C}_{16}\text{H}_{22}\text{N}_2\text{O}_6 + \text{H}]^+$: 339.1551. Found 339.1557.

4.2.7.2. Synthesis of compound 13, {(2S,3R,4S,5R)-3,4-dihydroxy-5-[(4-methoxybenzamido)methyl]oxolan-2-yl}acetic acid



Scheme 22. Preparation of compound **13**, structures of the precursor(s) (**7**) and product (66%).

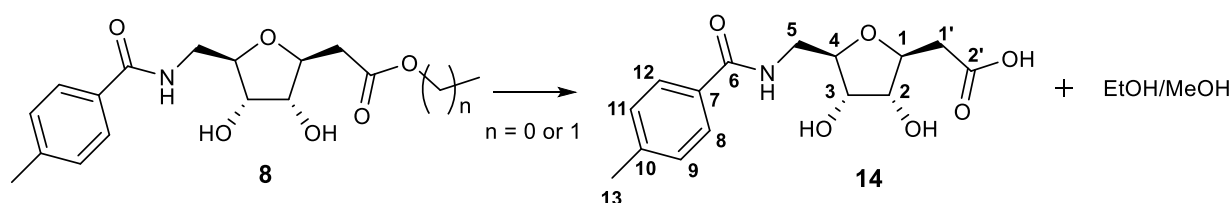
Compound(s) **7**: 67 mg, 0.19 mmol; ACN: 1.5 mL; LiOH (1 M): 570 μ L; HCl (1 M): 690 μ L; Yield: 66%

^1H NMR (300 MHz, CDCl_3): δ = 2.48 (dd, $\text{H1}'\text{a}$, $J_{1'\text{a},1} = 8.3$ Hz, $J_{1'\text{a},1'\text{b}} = 15.6$ Hz, 1H), 2.69 (dd, $\text{H1}'\text{b}$, $J_{1'\text{b},1} = 4.2$ Hz, 1H), 3.54 (m, H5a , H5b , 2H), 3.82 (t, H2 , $J_{2,1} = J_{2,3} = 5.7$ Hz, 1H), 3.85 (s, H13 , 3H), 3.91-3.98 (m, H3 , H4 , 2H), 4.11 (ddd, H1 , 1H), 6.98 (d, H9 , H11 , $J = 8.9$ Hz, 2H), 7.82 (d, H8 , H12 , 2H).

^{13}C NMR (75.5 MHz, CDCl_3): δ = 39.37 ($\text{C1}'$), 43.11 (C5), 55.90 (C13), 73.71 (C3), 75.69 (C2), 80.71 (C1), 84.03 (C4), 114.70 (C9 , C11), 127.50 (C7), 130.25 (C8 , C12), 163.97 (C10), 169.98 (C6), 174.99 ($\text{C2}'$) .

HRMS calcd for $[\text{C}_{15}\text{H}_{19}\text{NO}_7 + \text{H}]^+$: 326.1234. Found 326.1236.

4.2.7.3. Synthesis of compound **14**, {(2*S*,3*R*,4*S*,5*R*)-3,4-dihydroxy-5-[(4-methylbenzamido)methyl]oxolan-2-yl}acetic acid



Scheme 23. Preparation of compound **14**, structures of the precursor(s) (**8**) and product (98%).

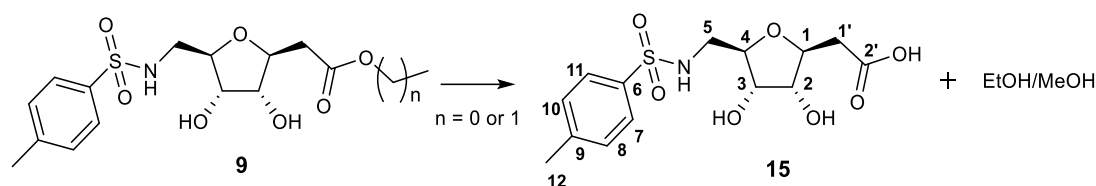
Compound(s) **8**: 45 mg, 0.135 mmol; ACN: 1.5 mL; LiOH (1M): 470 μL ; HCl (1 M): 546 μL ; Yield: 98%.

^1H NMR (300 MHz, CDCl_3): δ = 2.38 (s, H13 , 3H), 2.47 (dd, $\text{H1}'\text{a}$, $J_{1'\text{a},1} = 8.4$ Hz, $J_{1'\text{a},1'\text{b}} = 15.6$ Hz), 2.69 (dd, $\text{H1}'\text{b}$, $J_{1'\text{b},1} = 4.2$ Hz), 3.55-3.63 (m, H5a , H5b , 2H), 3.86 (t, H2 , $J_{2,1} = J_{2,3} = 5.1$ Hz, 1H), 3.94-4.04 (m, H3 , H4 , 2H), 4.09-4.15 (m, H1 , 1H), 7.27 (d, H8 , H12 , $J = 8.1$ Hz, 2H), 7.73 (d, H9 , H11 , 2H).

^{13}C NMR (100.6 MHz, CDCl_3): δ = 21.54 (C13), 39.27, ($\text{C1}'$), 43.15 (C5), 73.65, 73.70 (C3), 75.55 (C2), 80.54 (C1), 83.72 (C4), 128.46 (C9 , C11), 130.24 (C8 , C12), 132.19 (C7), 143.70 (C10), 170.73 (C6), 173.54 ($\text{C2}'$) .

HRMS calcd for $[\text{C}_{15}\text{H}_{19}\text{NO}_6 + \text{H}]^+$: 310.1285. Found 310.1279.

4.2.7.4. Synthesis of compound 15, [(2S,3R,4S,5R)-3,4-dihydroxy-5-[(4-methylbenzene-1-sulfonyl)amino]methyl]oxolan-2-yl]acetic acid



Scheme 24. Preparation of compound 15, structures of precursor(s) (9) and product (97%).

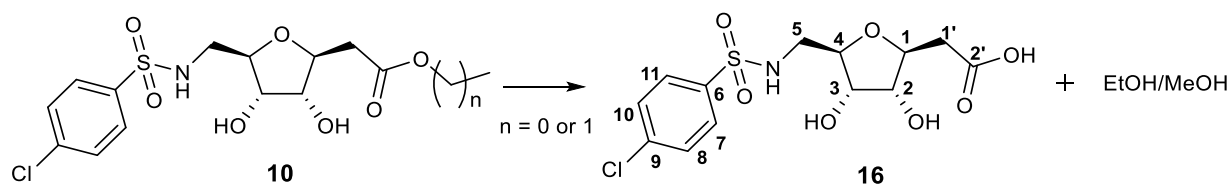
Compound(s) 9: 22 mg, 0.06 mmol; ACN: 1 mL; LiOH (1M): 191 mL; HCl (1 M): 223 μ L; Yield: 97%

^1H NMR (300 MHz, CDCl_3): δ = 2.39-2.47 (m, H1'a, H12, 4H), 2.63 (dd, H1'b, $J_{1'b,1} = 4.2$ Hz, $J_{1'b,1'a} = 15.6$ Hz, 1H), 2.96 (m, H5a, H5b, 2H), 3.75-3.81 (m, H2, H4, 2H), 3.93 (dd, H3, $J_{3,2/4} = 4.8$ Hz, $J_{3,2/4} = 5.5$ Hz, 1H), 4.04 (ddd, H1, $J_{1,1'a} = 8.4$ Hz, $J_{1,2} = 6.3$ Hz, 1H), 7.37 (d, H8, H10, $J = 8.1$ Hz, 2H), 7.73 (d, H7, H11, 2H).

^{13}C NMR (75.5 MHz, CDCl_3): δ = 21.43 (C12), 39.07 (C1'), 46.06 (C5), 73.16 (C3), 75.55 (C2), 80.59 (C1), 83.97 (C4), 128.01 (C7, C11), 130.71 (C8, C10), 138.76 (C6), 144.65 (C9), 175.01 (C2').

HRMS calcd for $[\text{C}_{14}\text{H}_{19}\text{NO}_7\text{S} + \text{H}]^+$: 346.0955. Found 346.0962.

4.2.7.5. Synthesis of compound 16, [(2S,3R,4S,5R)-5-[(4-chlorobenzene-1-sulfonyl)amino]methyl]-3,4-dihydroxyoxolan-2-yl]acetic acid



Scheme 25. Preparation of compound 16, structures of precursor(s) (10) and product (96%).

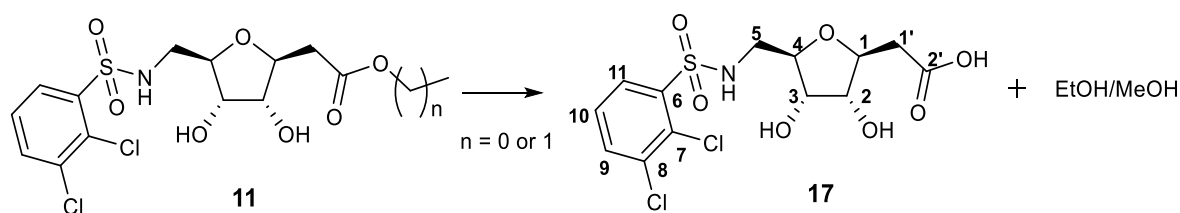
Compound(s) 10: 32 mg, 0.08 mmol; ACN: 1 mL; LiOH (1 M): 258 μ L; HCl (1 M): 310 μ L; Yield: 96%.

^1H NMR (300 MHz, CDCl_3): δ = 2.43 (dd, H1'a, $J_{1'a,1} = 8.3$ Hz, $J_{1'a,1b} = 15.6$ Hz, 1H), 2.69 (dd, H1'b, $J_{1'b,1} = 4.2$ Hz, 1H), 3.00 (dd, H5a, $J_{5a,4} = 5.2$ Hz, $J_{5a,5b} = 13.5$ Hz, 1H), 3.07 (dd, H5b, $J_{5b,4} = 4.1$ Hz, 1H), 3.76-3.81 (m, H2, H4, 2H), 3.92 (t, H3, $J_{3,2} = J_{3,4} = 5.3$ Hz, 1H), 4.04 (ddd, H1, $J_{1,2} = 6.1$ Hz, 1H), 7.55 (d, H7, H11, $J = 8.6$ Hz, 2H), 7.83 (d, H8, H10, 2H).

^{13}C NMR (75.5 MHz, CDCl_3): δ = 39.20 (C1'), 46.10 (C5), 73.16 (C3), 75.54 (C2), 80.72 (C1), 83.85 (C4), 129.70 (C8, C10), 130.40 (C7, C11), 139.74 (C6), 140.66 (C9), 175.10 (C2').

HRMS calcd for $[\text{C}_{13}\text{H}_{16}^{35}\text{ClNO}_7\text{S} + \text{H}]^+$: 366.0409. Found 366.0411.

4.2.7.6. Synthesis of compound 17, [(2S,3R,4S,5R)-5-[(2,3-dichlorobenzene-1-sulfonyl)amino]methyl]-3,4-dihydroxyoxolan-2-yl]acetic acid



Scheme 26. Preparation of compound 17, structures of precursor(s) (11) and product (96%).

Compound(s) 11: 32 mg, 0.08 mmol; ACN: 1mL; LiOH (1 M): 258 μL ; HCl (1 M): 310 μL ; Yield: 96%

^1H NMR (400 MHz, MeOD): δ = 2.39 (dd, H1'a, $J_{1'a,1} = 8.2$ Hz, $J_{1'a,1b} = 15.6$ Hz, 1H), 2.60 (dd, H1'b, $J_{1'b,1} = 4.6$ Hz, 1H), 3.01 (dd, H5a, $J_{5a,4} = 6.10$ Hz, $J_{5a,5b} = 14.0$ Hz, 1H), 3.18 (dd, H5b, $J_{5b,4} = 4.1$ Hz, 1H), 3.75-3.80 (m, H2, H4, 2H), 4.90 (t, H3, $J_{3,4} = J_{3,2} = 5.6$ Hz, 1H), 4.01 (ddd, H1, $J_{1,2} = 5.2$ Hz, 1H), 7.46 (t, H10, $J_{10,9} = J_{10,11} = 8.0$ Hz, 1H), 7.78 (d, H9, 1H), 8.02 (d, H11, 1H).

^{13}C NMR (100.6 MHz, MeOD): δ = 37.9 (C1'), 46.15 (C5), 73.14 (C3), 75.42 (C2), 80.91 (C1), 83.70 (C4), 128.93 (C10), 130.51 (C11), 131.03 (C7), 135.38 (C9), 136.31 (C8), 141.57 (C6), 174.75 (C2').

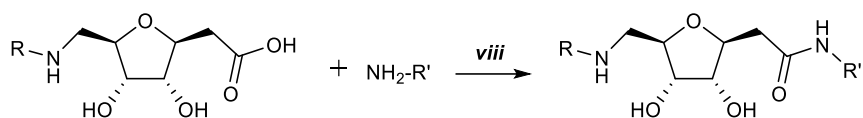
HRMS calcd for $[\text{C}_{13}\text{H}_{15}^{35}\text{Cl}_2\text{NO}_7\text{S} + \text{H}]^+$: 400.0019. Found 400.0028.

4.2.8. General procedure for the synthesis of C2' amides (18-23)

To a solution of each carboxylic acid (12-17) in dry 10:1 DCM:DMF, DIPEA was added, followed by HATU and the primary amine (Scheme 27-Scheme 33). The reaction mixture was stirred for 18-24h at r.t. to 40°C.

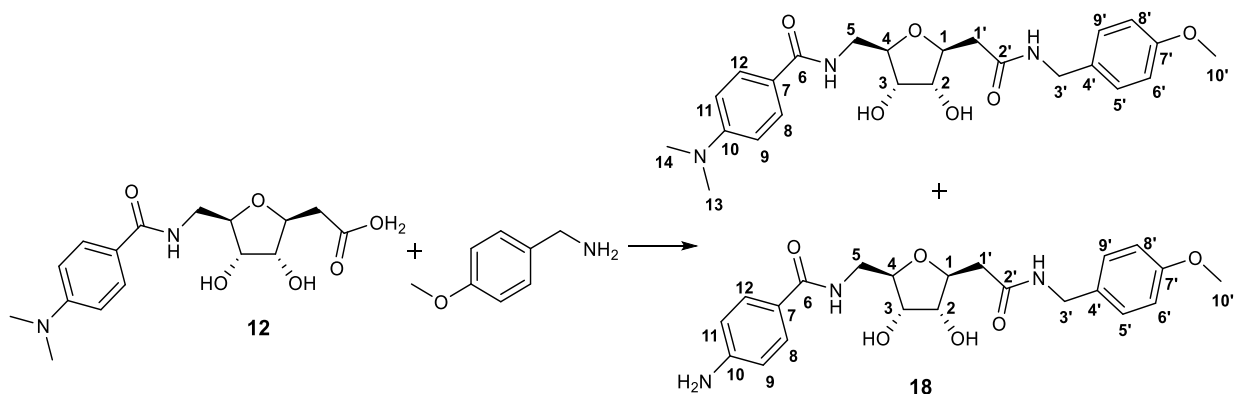
When consumption of the acid and product formation was observed by TLC (eluent: DCM:MeOH:H₂O 30:10:1), the reaction mixture was washed with 0.1 M HCl. The organic layer was concentrated. The residue was dissolved in DCM and washed with 1M HCl. The white precipitate formed was filtered. When precipitation did not occur, the mixture was extracted with 3 X 20 mL of EtOAc, washed with brine, dried over Mg_2SO_4 and concentrated. The crude mixtures were purified by semi-preparative reverse phase

HPLC (Program: solvent gradient starting at 90% acidified water (0.1% v/v formic acid) and 10% acetonitrile for one minute, followed by a linear increase of the acetonitrile proportion to 100% in 10 minutes. From minutes 11 to 13.5 the acetonitrile was kept at 100%, before returning to the initial (10%) proportion in a 2.5 min linear slope).



Scheme 27. General preparation scheme of compounds **18-23**. viii) $\text{NH}_2\text{R}'$, DIPEA, HATU, 10:1 DCM:DMF, 18-24h, r.t. - 40°C (3%-39%).

4.2.8.1. Synthesis of compound 18, 2-[(2S,3R,4S,5R)-3,4-dihydroxy-5-({[4-(dimethylamino)benzamido]]methyl)oxolan-2-yl]-N-[(4-methoxyphenyl)methyl]acetamide / 2-[(2S,3R,4S,5R)-5-({[4-(amino)benzamido]]methyl)-3,4-dihydroxyoxolan-2-yl]-N-[(4-methoxyphenyl)methyl]acetamide



Scheme 28. Preparation of compound **18** from compound **12** and 4-methoxybenzylamine (3%). The obtained mixture contained primarily the N-demethylated compound.

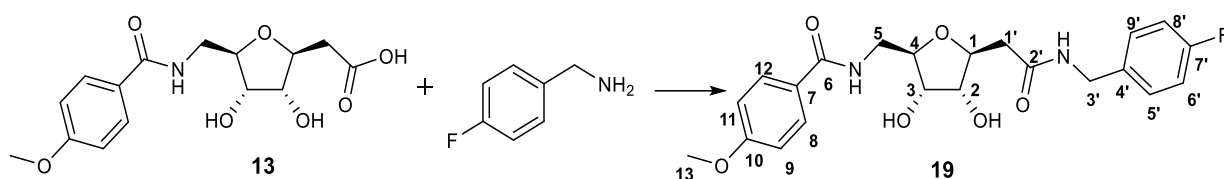
Compound **12**: 30 mg, 0.09 mmol; Primary amine: 4-methoxybenzylamine: 12 μL , 12.35 mg, 0.09 mmol; DIPEA: 22.5 μL ; HATU: 30 mg; DCM: 1.4 mL; Temperature: r.t.; Reaction time: 72h; Yield: ~3% (Mixture of the desired compound and the N-demethylated compound).

^1H NMR (400 MHz, CDCl_3): δ = 2.44 (dd, $\text{H}_{1'a}$, $J_{1'a,1} = 8.2$ Hz, $J_{1'a,1'b} = 14.2$ Hz, 1H), 2.60 (dd, $\text{H}_{1'b}$, $J_{1'b,1} = 4.1$ Hz, 1H), 3.55 (d, H_{5a} , H_{5b} , $J_{5,4} = 5.3$ Hz, 2H), 3.73 (s, $\text{H}_{10'}$, 3H), 3.83 (t, H_2 , $J_{2,1} = J_{2,3} = 6.1$ Hz, 1H), 3.91 (bt, H_3 , $J_{2,1} = J_{2,3} = 6.1$ Hz, 1H), 3.90-3.98 (m, H_2 , H_4 , 2H), 4.06-4.11 (m, H_1 , 1H), 4.29 (s, $\text{H}_{3'}$, 2H), 6.81 (d, $\text{H}_{6'}$, $\text{H}_{8'}$, $J = 8.5$ Hz, 2H), 7.18 (d, $\text{H}_{5'}$, $\text{H}_{9'}$, 2H), 7.66 (d, H_9 , H_{11} , $J = 8.6$ Hz, 2H), 7.80 (d, H_8 , H_{12} , 2H).

^{13}C NMR (100.6 MHz, CDCl_3): δ = 41.32 (C1'), 43.43, 43.61 (C5, C3'), 55.68 (C10'), 73.92 (C3), 75.97 (C2), 81.18 (C1), 84.05 (C4), 114.91 (C9, C11), 120.35 (C6', 8'), 129.43, 129.78 (C8, C12, C5', C9'), 142.13 (C4'), 161.74 (C7'), 173.03 (C2'). C13, C14, C6, and C10, not found. As this fraction was primarily *N*-demethylated, H13 and H14 and C13 and C14 were not found. A limitation of the amount of compound obtained prevented further analysis to find C6 and C10.

HRMS calcd for $[\text{C}_{24}\text{H}_{31}\text{N}_3\text{O}_6 + \text{H}]^+$: 458.2286. Found 458.192; HRMS calcd for $[\text{C}_{22}\text{H}_{27}\text{N}_3\text{O}_6 + \text{H}]^+$: 430.1973. Found 430.1973.

4.2.8.2. Synthesis of compound 19, 2-[(2*S*,3*R*,4*S*,5*R*)-3,4-dihydroxy-5-[(4-methoxybenzamido)methyl]oxolan-2-yl]-*N*-[(4-fluorophenyl)methyl]acetamide



Scheme 29. Preparation of compound **19**, from compound **13** and 4-fluorobenzylamine (28%).

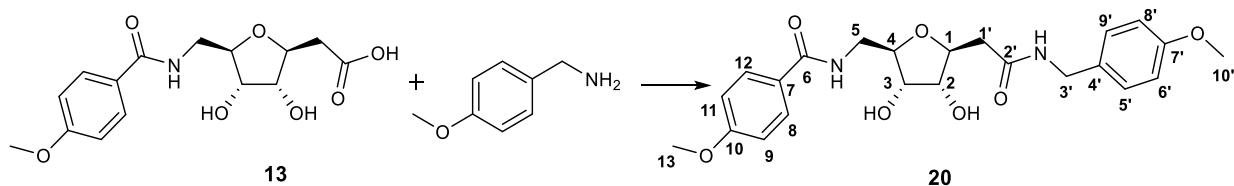
Compound **13**: 30 mg, 0.09 mmol; Primary amine: 4-fluorobenzylamine: 12 mg, 0.09 mmol; DIPEA: 19.2 μL ; HATU: 35 mg; DCM: 1mL; Temperature: r.t.; Reaction time: 18h; Yield: 28%.

^1H NMR (300 MHz, MeOD): δ = 2.44 (dd, H1'a, $J_{1'a,1} = 8.5$ Hz, $J_{1'a,1b} = 14.2$ Hz, 1H), 2.61 (dd, H1'b, $J_{1'b,1} = 4.0$ Hz, 1H), 3.53-3.55 (m, H5a, H5b, 2H), 3.80-3.84 (m, H2, 1H), 3.84 (s, H13, 3H), 3.90-3.99 (m, H3, H4, 2H), 4.09 (ddd, H1, $J_{1,2} = 6.1$ Hz, 1H), 4.34 (2d, H3'a, H3'b, $J_{3'a,3'b} = 15.4$ Hz, 2H), 6.93-7.00 (m, H9, H11, H6', H8', 4H), 7.27 (dd, H5', H9', $J_{5'/9', 6'/8'} = 8.4$ Hz, $J_{5'/9', F} = 5.2$ Hz, 2H), 7.78 (d, H8, H12, $J_{8/12, 9/11} = 8.9$ Hz, 2H).

^{13}C NMR (75.5 MHz, MeOD): δ = 41.34 (C1'), 43.29, 43.35 (C5, C3'), 55.87 (C13), 73.87 (C4), 76.00 (C2), 81.01 (C1), 84.12 (C3), 114.67 (C9, C11), 115.83 (d, C6', C8', $J_{6'/8', F} = 21.6$ Hz), 127.45 (C7), 130.20, 130.20 (C8, C12), 130.25 (d, C5', C9', $J_{5'/9', F} = 7.9$ Hz), 135.86 (d, C4', $J_{4', F} = 3.15$ Hz), 163.33 (d, C7', $J_{7', F} = 243.4$ Hz), 163.94 (C10), 169.98 (C6), 173.20 (C2').

HRMS calcd for $[\text{C}_{22}\text{H}_{25}\text{FN}_2\text{O}_6 + \text{H}]^+$: 433.1769. Found 433.1768.

4.2.8.3. Synthesis of compound 20, 2-[(2S,3R,4S,5R)-3,4-dihydroxy-5-[(4-methoxybenzamido)methyl]oxolan-2-yl]-N-[(4-methoxyphenyl)methyl]acetamide



Scheme 30. Preparation of compound 20 from compound 13 and 4-methoxybenzylamine (35%).

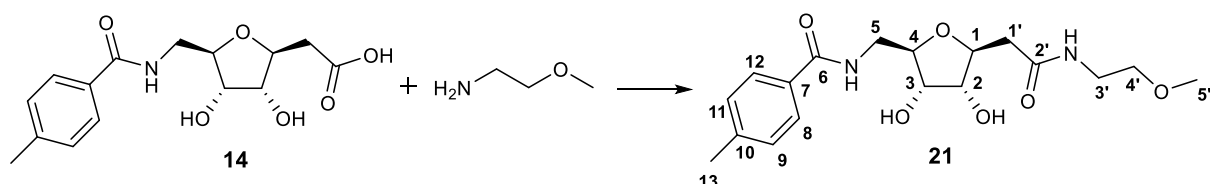
Compound 13: 30 mg, 0.08 mmol; Primary amine: 4-methoxybenzylamine: 21.9 mg, 0.16 mmol; DIPEA: 27.9 μ L; HATU: 37 mg; DCM: 1.2 mL; Temperature: r.t.; Reaction time: 24h; Yield: 35%.

^1H NMR (400 MHz, MeOD): δ = 2.44 (dd, H1'a, $J_{1'a,1} = 8.5$ Hz, $J_{1'a,1b} = 14.2$ Hz, 1H), 2.60 (dd, H1'b, $J_{1'b,1} = 4.0$ Hz, 1H), 3.53-3.55 (m, H5a, H5b, 2H), 3.73 (s, H10', 3H), 3.81-3.84 (m, H2, 1H), 3.84 (s, H13, 3H), 3.89-3.98 (m, H3, H4), 4.06-4.01 (m, H1, 1H), 4.29-4.30 (m, H3'a, H3'b, 2H), 6.80 (d, H6', H8', $J_{6',5'} = J_{8',9'} = 8.6$ Hz, 2H), 6.95 (d, H9, H11, $J_{9,8} = J_{11,12} = 8.8$ Hz, 2H), 7.18 (d, H5', H9', 2H), 7.78 (d, H8, H12, 2H).

^{13}C NMR (100.6 MHz, MeOD): δ = 41.34 (C1'), 43.37, 43.59 (C5, C3'), 55.66, 55.92 (C13, C10'), 73.91 (C4), 75.99 (C2), 81.15 (C1), 84.12 (C3), 114.72, 114.90 (C9, C11, C6', C8'), 127.54 (C7), 129.77, 130.25 (C8, C12, C5', C9'), 131.83 (C4'), 160.35 (C7'), 163.97 (C10), 170.00 (C6), 173.06 (C2').

HRMS calcd for $[\text{C}_{23}\text{H}_{28}\text{N}_2\text{O}_7 + \text{H}]^+$: 445.1969. Found 445.1966.

4.2.8.4. Synthesis of compound 21, 2-[(2S,3R,4S,5R)-3,4-dihydroxy-5-[(4-methoxybenzamido)methyl]oxolan-2-yl]-N-(2-methoxyethyl)acetamide



Scheme 31. Preparation of compound 21 from compound 14 and 2-methoxyethylamine (21%).

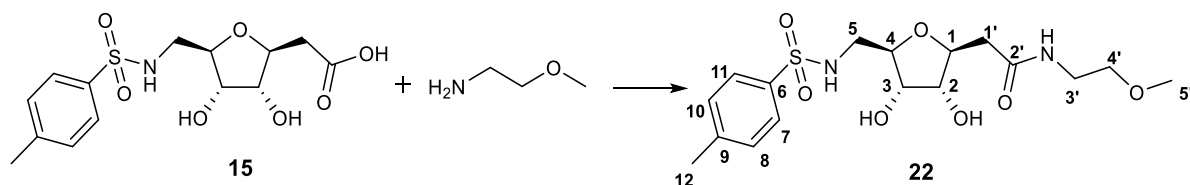
Compound 14: 90 mg, 0.29 mmol; Primary amine: 2-methoxyethylamine: 27 mg, 32 μ L, 0.36 mmol, 1.2 eq; DIPEA: 101 μ L; HATU: 137mg; DCM: 4.5 mL; Temperature: r.t.; Reaction time: 48h; Yield: 21%.

^1H NMR (400 MHz, CDCl_3): δ = 2.39 (s, H13, 3H), 2.45 (dd, H1'a, $J_{1'a,1} = 8.1$ Hz, $J_{1'a,1'b} = 14.3$ Hz, 1H), 2.60 (dd, H1'b, $J_{1'b,1} = 4.2$ Hz, 1H), 3.34-3.38 (m, H3', 2H), 3.37 (s, H5', 3H), 3.44 (t, H4', $J_{4',3'} = 5.2$ Hz, 2H), 3.55-3.61 (m, H5a, H5b, 2H), 3.82 (t, H2, $J_{2,1} = J_{2,3} = 5.5$ Hz, 1H), 3.93-3.98 (m, H3, H4, 2H), 4.06 (ddd, H1, 1H), 7.28 (d, H9, H11, $J = 8.0$ Hz, 2H), 7.74 (d, H8, H12 2H).

^{13}C NMR (100.6 MHz, CDCl_3): δ = 21.40 (C13), 40.26, 41.04 (C1', C3'), 43.32 (C5), 58.83 (C5'), 71.77 (C4'), 73.80 (C3), 75.87 (C2), 80.98 (C1), 83.99 (C4), 128.39 (C8, C12), 130.13 (C9, C11), 132.57 (C7), 143.46 (C10), 170.44 (C6), 173.46 (C2') .

HRMS calcd for $[\text{C}_{18}\text{H}_{26}\text{N}_2\text{O}_6 + \text{H}]^+$: 367.1864. Found 367.1864.

4.2.8.5. Synthesis of compound 22, [(2S,3R,4S,5R)-3,4-dihydroxy-5-[(4-methylbenzene-1-sulfonyl)amino]methyl]oxolan-2-yl]-N-(2-methoxyethyl)acetamide



Scheme 32. Preparation of compound 22 from compound 15 and 2-methoxyethylamine (10%).

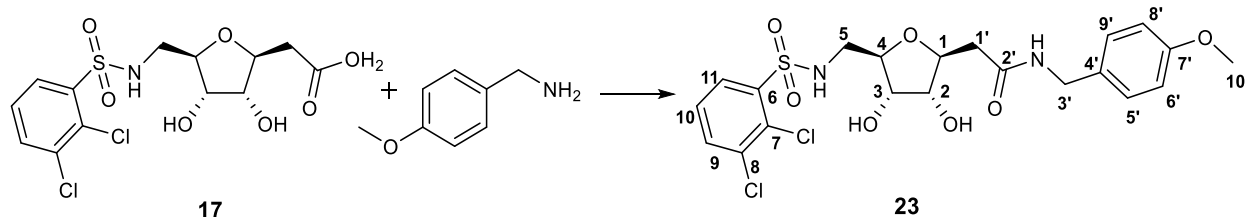
Compound **15**: 20 mg, 0.06 mmol; Primary amine: 2-methoxyethylamine: 0.45 mg, 5.2 μL , 0.06 mmol; DIPEA: 15 μL ; HATU: 23 mg; DCM: 0.93 mL; Temperature: r.t.; Reaction time: 18h; Yield: 10%.

^1H NMR (400 MHz, CDCl_3): δ = 2.37 (dd, H1'a, $J_{1'a,1} = 8.2$ Hz, $J_{1'a,1'b} = 14.2$ Hz, 1H), 2.43 (s, H12, 3H), 2.53 (dd, H1'b, $J_{1'b,1} = 4.1$ Hz, 1H), 2.97 (dd, H5a, $J_{5a,4} = 5.4$ Hz, $J_{5a,5b} = 13.5$ Hz, 1H), 3.04 (dd, H5b, $J_{5b,4} = 3.8$ Hz, 1H), 3.33 (s, H5', 3H), 3.33-3.37 (m, H3', 2H), 3.45 (t, H4', $J_{4',3'} = 5.5$ Hz, 2H), 3.75-3.79 (m, H2, H4, 2H), 3.91 (dd, H3, $J_{3,2} = J_{3,4} = 5.3$ Hz, 1H), 3.99 (ddd, H1, $J_{1,2} = 5.3$ Hz, 1H), 7.38 (d, H8, H10, $J = 8.0$ Hz, 2H), 7.75 (d, H7, H11, 2H).

^{13}C NMR (100.6 MHz, CDCl_3): δ = 21.42 (C12), 40.18, 41.01 (C1', C3'), 46.08 (C5), 58.87 (C5'), 71.91 (C4'), 73.28 (C3), 75.75 (C2), 81.19 (C1), 83.94 (C4), 128.03 (C7, C11), 130.74 (C8, C10), 138.94 (C6), 144.72 (C9), 173.37 (C2') .

HRMS calcd for $[\text{C}_{17}\text{H}_{26}\text{N}_2\text{O}_7\text{S} + \text{H}]^+$: 403.1533. Found 403.1545.

4.2.8.6. Synthesis of compound 23, [(2S,3R,4S,5R)-5-[(2,3-dichlorobenzene-1-sulfonyl)amino]methyl]-3,4-dihydroxyoxolan-2-yl]-N-[(4-methoxyphenyl)methyl]acetamide



Scheme 33. Preparation of compound **23** from compound **17** and 4-methoxybenzylamine (39%).

Compound **17**: 23 mg, 0.06 mmol; Primary amine: 4-methoxybenzylamine: 15.7 μ L, 0.12 mmol; DIPEA: 21 μ L; HATU: 28.3 mg DCM: 0.9 mL; Temperature: r.t. Reaction time: 18h; Yield: 39%.

^1H NMR (300 MHz, MeOD): δ = 2.37 (dd, H1'a, $J_{1'a,1} = 8.2$ Hz, $J_{1'a,1b} = 14.3$ Hz, 1H), 2.54 (dd, H1'b, $J_{1'b,1} = 4.2$ Hz, 1H), 3.02 (dd, H5a, $J_{5a,4} = 5.9$ Hz, $J_{5a,5b} = 13.9$ Hz, 1H), 3.15 (dd, H5b, $J_{5b,4} = 4.1$ Hz, 1H), 3.76 (s, H10', 3H) 3.75-3.80 (m, H2, H4, 2H), 5.4 (t, H3, $J_{3,4} = J_{3,2} = 5.3$ Hz, 1H), 4.01 (ddd, H1, $J_{1,2} = 5.1$ Hz, 1H), 4.31 (s, H3'a, H3'b, 2H), 6.86 (d, H6', H8', $J_{6',5'} = J_{8',9'} = 8.6$ Hz, 2H), 7.21 (d, H5', H9', 2H), 7.44 (t, H10, $J_{10,9} = J_{10,11} = 8.1$ Hz, 1H), 7.78 (d, H9, 1H), 7.99 (d, H11, 1H).

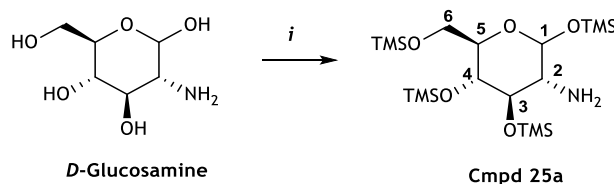
^{13}C NMR (75.5 MHz, MeOD): δ = 41.09 (C1'), 43.62 (C3'), 46.21(C5), 55.70 (C10'), 73.28 (C3), 75.69 (C2), 81.34 (C1), 83.78 (C4), 114.93 (C6', C8'), 128.96 (C10), 129.92 (C5', C9'), 130.55 (C11), 131.04 (C7), 131.83 (C4'), 135.42 (C9), 136.34 (C8), 141.56 (C6), 160.39 (C7'), 172.91(C2').

HRMS calcd for $[\text{C}_{21}\text{H}_{24}^{35}\text{Cl}_2\text{N}_2\text{O}_7\text{S} + \text{H}]^+$: 519.0754. Found 519.0763.

4.3. Synthesis of “cmpd 25” - known inhibitor

The synthetic route used to synthesize 2-[(3-Biphenyl)carbonyl]amino]-6-[(2,3-dichlorobenzene-1-sulfonyl)amino]-2,6-dideoxy-D-glucopyranose (“cmpd 25”) was based on the one described in Lin *et al* [43]. For convenience, the designations below are based on D-glucose. The anomeric carbon configuration is not specified.

4.3.1. Synthesis of 2-Amino-2-deoxy-1,3,4,6-tetrakis-O-(trimethylsilyl)-D-glucopyranose (“**cmpd 25a**”)



Scheme 34. Preparation of “**cmpd 25a**” from D-2-glucosamine: structures and numbering. i) Et_3N , TMSCl, pyridine, 115°C , overnight (71%).

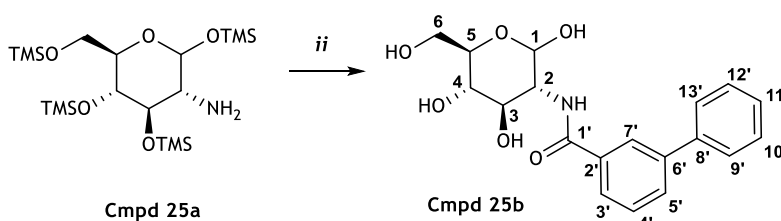
Triethylamine (7 mL, 50 mmol, 5 eq) and TMSCl (6.39 mL, 50 mmol, 5 eq) were added to a solution of 2-amino-2-deoxy-D-glucopyranose (HCl salt, 2.156 g, 10 mmol) in pyridine (25 mL). The mixture was refluxed overnight. The reaction mixture was then let to cool down and filtered. The filtrate was poured into water (100 mL) and was extracted with hexane (3 X 150 mL). The organic layers were combined and washed with water (50 mL x 2) and brine. The organic layer was dried over MgSO_4 and concentrated under vacuum to give “**cmpd 25a**” as an off-white solid (71%).

^1H NMR (400 MHz, CDCl_3): δ = 0.15-0.21 (m, H_{TMS} , 36H), 2.54 (dd, H2, $J_{1,2}$ = 3.3 Hz, $J_{2,3}$ = 9.2 Hz, 1H), 3.43-3.50 (m, H5, 1H), 3.55 (t, H3, $J_{3,4}$ = 9.2 Hz, 1H), 3.68-3.74 (m, H4, H6, 3H), 5.15 (d, H1, 1H).

^{13}C NMR (100.6 MHz, CDCl_3): δ = 0.05, 0.97, 1.43 (TMS), 57.68 (C2), 61.95 (C6), 72.04 (C5), 72.52 (C4), 77.34 (C3), 94.87 (C1).

HRMS calcd for $[\text{C}_{18}\text{H}_{45}\text{NO}_5\text{Si}_4 - \text{TMS} + \text{H}]^+$: 396.2052. Found 396.2058.

4.3.2. Synthesis of 2-[(3-Biphenyl)carbonyl]amino-2-deoxy-D-glucopyranose (“**cmpd 25b**”)



Scheme 35. Preparation of “**cmpd 25b**” from “**cmpd 25a**”: structures and numbering. ii) 3-biphenylcarboxylic acid, DIPEA, HATU, DCM, r.t., overnight (76%).

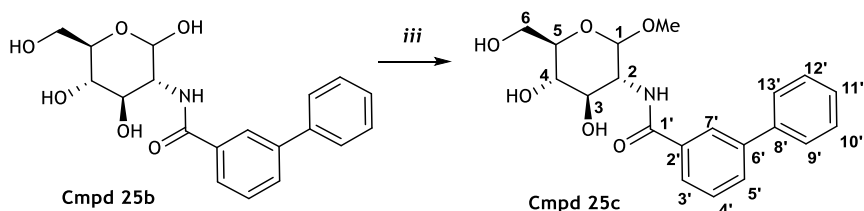
DIPEA (316 μ L, 1.8 mmol, 1.2 eq) was added to a solution of “**cmpd 25a**” (708 mg, 1.5 mmol, 1 eq) in dichloromethane (DCM) (15 mL), followed by the addition of 3-biphenylcarboxylic acid (300 mg, 1.5 mmol, 1 eq) and HATU (575 mg, 1.5 mmol, 1 eq). The reaction mixture was stirred overnight at r.t. The mixture was washed with 0.1 M HCl, and the organic layer was concentrated. The residue was dissolved in DCM, followed by the addition of 1 M HCl (25 mL). The white solid was formed and filtered to give the product “**cmpd 25b**” (76%).

^1H NMR (300 MHz, MeOD): δ = 3.38-3.49 (m, H4, 1H), 3.67 (m, H6a, 1H), 3.82-3.96 (m, H3, H5, H6b, 3H), 4.12 (dd, H2, $J_{2,1}$ = 3.6 Hz, $J_{2,3}$ = 10.7 Hz, 1H), 5.29 (d, H1, 1H), 7.30 (t, H11', $J_{11',10'} = J_{11',12'} = 7.4$ Hz, 1H), 7.39 (t, H10', H12', $J_{10',9'} = J_{12',13'} = 7.4$ Hz, 2H), 7.53 (t, H4', $J_{4',3'} = J_{4',5'} = 7.8$ Hz, 1H), 7.67 (d, H9', H13', 2H), 7.78 (d, H5', 1H), 7.84 (d, H3', 1H), 8.13 (s, H7', 1H).

^{13}C NMR (75.5 MHz, MeOD): δ = 56.75 (C2), 62.84 (C6), 72.36, 72.55, 73.21 (C3, C4, C5), 92.59 (C1), 127.14 (C7'), 127.34 (C3'), 128.12 (C9', C13'), 128.79 (C11'), 129.97 (C10', C12'), 130.05 (C4'), 131.15 (C5'), 136.30 (C2'), 141.60 (C6'), 142.80 (C8'), 170.67 (C1').

HRMS calcd for $[\text{C}_{19}\text{H}_{21}\text{NO}_6 + \text{H}]^+$: 360.1442. Found 360.1441.

4.3.3. Synthesis of Methyl 2-[(3-biphenyl)carbonyl]amino]-2-deoxy-D-glucopyranoside (“**cmpd 25c**”)



Scheme 36. Preparation of “**cmpd 25c**” from “**cmpd 25b**”: structures and numbering. iii) MeOH, HCl, 65°C, overnight (94%).

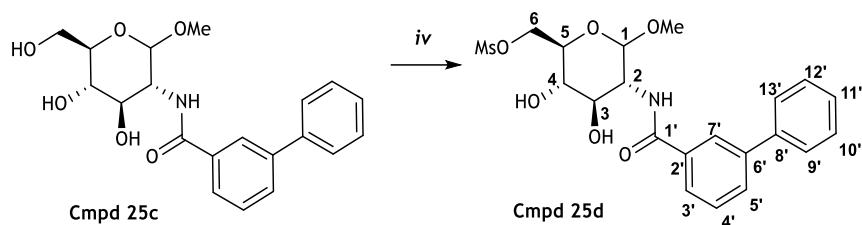
To a solution of “**cmpd 25b**” (230 mg, 0.64 mmol, 1 eq) in methanol (5 mL), 37% HCl (100 μ L, 1.28 mmol, 2 eq) was added. The reaction mixture was stirred at reflux overnight. The mixture was let to cool down, concentrated and washed with DCM to give the product as a white powder/needle crystals (94%).

^1H NMR (300 MHz, MeOD): δ = 33.41 (s, Me, 3H), 3.41-3.50 (m, H4, 1H), 3.59-3.64 (m, H5, 1H), 3.73 (dd, H6a, $J_{6a,5}$ = 5.6 Hz, $J_{6a,6b}$ = 11.8 Hz, 1H), 3.84-3.90 (m, H3, H6b, 2H), 4.16 (dd, H2, $J_{2,1}$ = 3.6 Hz, $J_{2,3}$ = 10.7 Hz, 1H), 4.85 (d, H1, 1H), 7.37 (t, H11', $J_{11',10'} = J_{11',12'} = 7.4$ Hz, 1H), 7.47 (t, H10', H12', $J_{10',9'} = J_{12',13'} = 7.4$ Hz, 2H), 7.55 (t, H4', $J_{4',3'} = J_{4',5'} = 7.8$ Hz, 1H), 7.69 (d, H9', H13', 2H), 7.79-7.85 (m, H3', H5', 2H), 8.13 (s, H7', 1H).

^{13}C NMR (75.5 MHz, MeOD): δ = 55.59 (Me), 56.33 (C2), 62.76 (C6), 72.39 (C3), 72.51 (C4), 73.70 (C5), 99.77 (C1), 127.21 (C7'), 127.41 (C3'), 128.14 (C9', C13'), 128.80 (C11'), 129.99, 130.05 (C10', C12', C4'), 131.20 (C5'), 136.25 (C2'), 141.56 (C8'), 142.78 (C6'), 170.99 (C1').

HRMS calcd for $[C_{20}H_{23}NO_6 + H]^+$: 374.1598. Found 374.1592.

4.3.4. Synthesis of Methyl 2-[(3-biphenylcarbonyl)amino]-2-deoxy-6-O-[mesyl]-D-glucopyranoside ("cmpd 25d")



Scheme 37. Preparation of "cmpd 25d" from "cmpd 25c": structures and numbering. iv) MsCl, pyridine, r.t., 6h (60%).

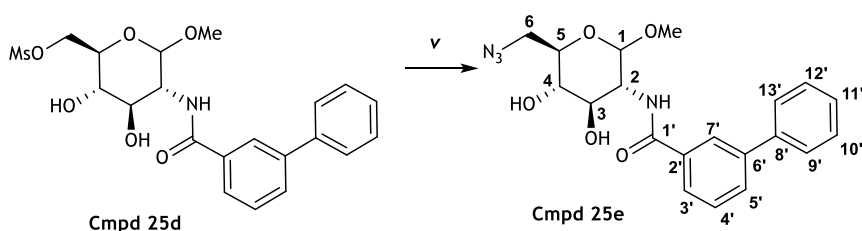
Mesylchloride (24 μ L, 0.31 mmol, 1.2 eq) was added to a solution of "cmpd 25c" (97 mg, 0.26 mmol, 1 eq) in pyridine (2 mL). The mixture was stirred for 6h at r.t. The reaction mixture was concentrated, and the residue poured into water (10 mL). The mixture was extracted with EtOAc (3 X 20 mL). The combined organic layers were washed with brine and dried over $MgSO_4$. The solvent was removed to give a yellowish powder as the crude "cmpd 25d" (60%).

1H NMR (400 MHz, MeOD): δ = 3.12 (s, Ms-CH₃, 3H), 3.38 (s, OMe, 3H), 3.47-3.49 (m, H4, 1H), 3.85-3.94 (m, H3, H5, 2H), 4.20-4.23 (m, H2, 1H), 4.45-4.47 (m, H6a, 1H), 4.54-4.60 (m, H6b, 1H), 4.86 (m, H1, 1H), 7.31-7.36 (m, H11', 1H), 7.40-7.45 (m, H10', H12', 2H), 7.47-7.53 (m, H4', 1H), 7.60-7.65 (m, H9', H13', 2H), 7.72-7.76 (m, H5', 1H), 7.82-7.84 (m, H3', 1H), 8.12 (s, H7', 1H),

^{13}C NMR (100.6 MHz, MeOD): δ = 37.34 (Ms-CH₃), 55.84, 55.92 (OMe, C2), 70.72 (C6), 71.23 (C5), 71.84 (C4), 72.40 (C3), 99.79 (C1), 127.11 (C7'), 127.32 (C3'), 128.01 (C9', C13'), 128.72 (C11'), 129.90 (C10', C12'), 129.98 (C4'), 131.13 (C5'), 135.98 (C2'), 141.37 (C6'), 142.59 (C8'), 170.51 (C1').

HRMS calcd for $[C_{21}H_{25}NO_8S + H]^+$: 452.1374. Found 452.1370.

4.3.5. Synthesis of Methyl 6-azido-2-[(3-biphenylcarbonyl)amino]-2,6-dideoxy-D-glucopyranoside ("cmpd 25e")



Scheme 38. Preparation of "cmpd 25e" from "cmpd 25d": structures and numbering. v) NaN_3 , DMF, 90°C, 2h (94%).

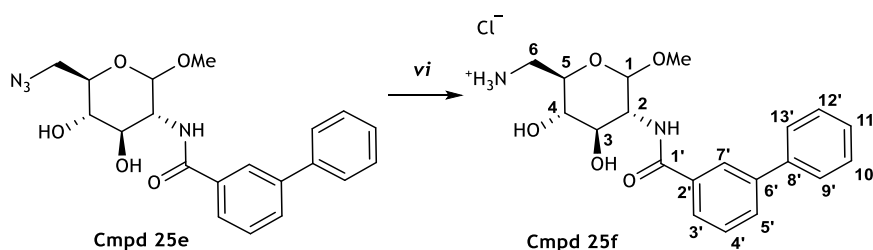
Sodium azide (29 mg, 0.3 mmol, 2 eq) was added to a solution of “**cmpd 25d**” (70 mg, 0.155 mmol) in DMF (2 mL). The solution was heated at 90 °C for 2 hr and then cooled down to r.t. The mixture was partitioned between EtOAc and water and the organic layer was dried over MgSO₄. The yellowish crude product was obtained after concentration under reduced pressure (94%).

¹H NMR (400 MHz, CDCl₃): δ= 3.44 (s, Me, 3H), 3.24-3.48 (m, H4, H6a, H6b, 3H), 3.65-3.71 (m, H3, 1H), 3.77-3.85 (m, H5, 1H), 4.28-4.36 (m, H2, 1H), 4.67-4.72 (m, H1, 1H), 7.24-7.40 (m, H4', H10', H11', H12', 4H), 7.48-7.53 (m, H9', H13', 2H), 7.61 (d, H5', J_{5',4'} = 7.8 Hz, 1H), 7.64 (d, H3', J_{3',4'} = 7.8 Hz, 1H), 7.94-8.02 (s, H7', 1H).

¹³C NMR (100.6 MHz, CDCl₃): δ= 51.55 (C6), 54.07, 55.33 (Me, C2), 71.15 (C5), 71.86 (C4), 73.12 (C3), 98.43 (C1), 125.90 (C7'), 126.27 (C3'), 127.19 (C9', C13'), 127.81 (C11'), 128.91 (C10', C12'), 129.01 (C4'), 130.47 (C5'), 134.35 (C2'), 140.07 (C6'), 141.65 (C8'), 168.64 (C1').

HRMS calcd for [C₂₀H₂₂N₄O₅ + H]⁺: 399.1663. Found 399.1670.

4.3.6. Synthesis of Methyl 6-amino-2-[(3-biphenyl)carbonyl]amino]-2,6-dideoxy-D-glucopyranoside (“**cmpd 25f**”)



Scheme 39. Preparation of “**cmpd 25f**” from “**cmpd 25e**”: structures and numbering. vi) Pd/C, EtOH, HCl, r.t., overnight (99%).

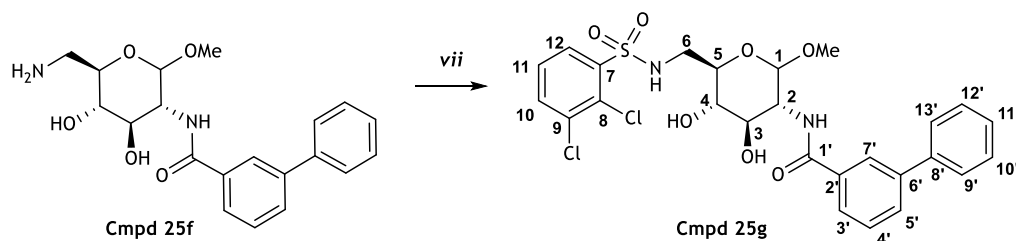
Pd/C (10%; 10 mg) was added to a solution of “**cmpd 25e**” in EtOH (1.39 mL) and HCl (3 M, 139 μ L). After degassing, the reaction mixture was stirred at r.t. overnight under H₂, 1 atm. After filtration over celite, the filtrate was concentrated to give the product “**cmpd 25f**” (99%).

¹H NMR (400 MHz, MeOD): δ= 3.10-3.18 (m, H6a, 1H), 3.44 (s, Me, 3H), 3.33-3.53 (m, H4, H6b, 2H), 3.81-3.95 (m, H3, H5, 2H), 4.15-4.23 (m, H2, 1H), 4.90-4.93 (m, H1, 1H), 7.33-7.40 (m, H11', 1H), 7.42-7.49 (m, H10', H12', 2H), 7.51-7.57 (m, H4', 1H), 7.68 (d, H9', H13', J_{9',10'} = J_{13',12'} = 7.4 Hz, 2H), 7.76-7.86 (m, H3', H5', 2H), 8.12 (s, H7', 1H).

¹³C NMR (100.6 MHz, MeOD): δ= 42.22 (C6), 55.87, 56.39 (Me, C2), 69.33, 71.82 (C3, C5), 73.89 (C4), 99.88 (C1), 127.06 (C7'), 127.31 (C3'), 127.97 (C9', C13'), 128.68 (C11'), 129.87 (C10', C12'), 129.98 (C4'), 131.15 (C5'), 135.86 (C2'), 141.32 (C6'), 142.59 (C8'), 170.56 (C1').

HRMS calcd for [C₂₀H₂₅N₂O₅]⁺: 373.1758. Found 373.1768.

4.3.7. Synthesis of Methyl 2-[(3-biphenylcarbonyl)amino]-6-[(2,3-dichlorobenzene-1-sulfonyl)amino]-2,6-dideoxy-D-glucopyranoside ("cmpd 25g")



Scheme 40. Preparation of "cmpd 25g" from "cmpd 25f": structures and numbering. vii) 2,3-dichlorobenzene-1-sulfonyl chloride, pyridine, overnight, r.t. (60%).

To a solution of "cmpd 25f" (142 mg, 0.38 mmol) in pyridine (5 mL), 2,3-dichlorobenzenesulfonyl chloride (187 mg, 0.76 mmol, 2 eq) was added, and the reaction mixture was stirred overnight at r.t. The crude product obtained after concentration at reduced pressure was dissolved in DCM and washed with 1M HCl and brine. The product (60%) was precipitated with ethyl ether and collected by filtration.

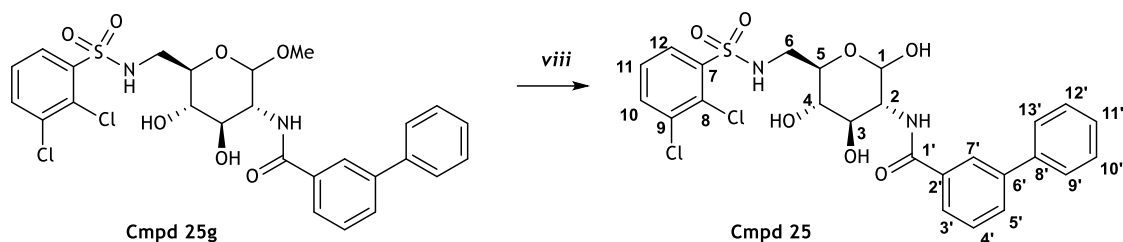
^1H NMR (400 MHz, MeOD): δ = 3.14-3.29 (m, H4, H6a), 3.41-3.60 (m, H5, H6b, Me), 3.80 (dd, H3, $J_{3,2}$ = 10.7 Hz, $J_{3,4}$ = 8.8 Hz, 1H), 4.02-4.07 (m, H2, 1H), 4.36 (d, H1 β , $J_{1\beta,2}$ = 8.4 Hz, 0.3H), 4.63 (d, H1 α , $J_{1\alpha,2}$ = 3.5 Hz, 0.7H), 7.37 (t, H11', $J_{11',10'} = J_{11',12'} = 7.5$ Hz, 1H), 7.45-7.57 (m, H10', H12', H11, H4', 4H), 7.67-7.69 (m, H9', H13', 2H), 7.78-7.83 (m, H10, H3', H5', 3H), 8.04-8.13 (m, H12, 7', 2H).

^{13}C NMR (100.6 MHz, MeOD): δ = 45.54 (C6), 55.83 (Me), 56.06 (C2), 72.18 (C5, C3), 73.74 (C4), 99.68 (C1), 127.16 (C7'), 127.32 (C3'), 128.09 (C9', C13'), 128.77, 128.84 (C11', C11), 129.95 (C10', C12'), 130.00 (C4'), 130.51 (C12), 131.16 (C5'), 135.24 (C10), 136.16, 136.24 (C2', C8, C9), 141.56 (C6'), 142.11 (C7), 142.77 (C8'), 170.64 (C1').

HRMS calcd for $[\text{C}_{26}\text{H}_{26}^{35}\text{Cl}_2\text{N}_2\text{O}_7\text{S} + \text{H}]^+$: 581.0911. Found 581.0986.

^{13}C NMR (100.6 MHz, MeOD): δ = 45.54 (C6), 55.83 (Me), 56.06 (C2), 72.18 (C5, C3), 73.74 (C4), 99.68 (C1), 127.16 (C7'), 127.32 (C3'), 128.09 (C9', C13'), 128.77, 128.84 (C11', C11), 129.95 (C10', C12'), 130.00 (C4'), 130.51 (C12), 131.16 (C5'), 135.24 (C10), 136.16, 136.24 (C2', C8, C9), 141.56 (C6'), 142.11 (C7), 142.77 (C8'), 170.64 (C1').

4.3.8. Synthesis of 2-[(3-Biphenylcarbonyl)amino]-6-[(2,3-dichlorobenzene-1-sulfonyl)amino]-2,6-dideoxy-D-glucopyranose (“**cmpd 25**”)



Scheme 41. Preparation of “**cmpd 25**” from “**cmpd 25g**”: structures and numbering. *viii*) 2:1 dioxane:water; HCl, 100°C, overnight (9%).

The mixture of “**cmpd 25g**” anomers (57 mg, 0.1 mmol) in 2:1 dioxane:water (3 ml) containing 0.5 mL of 6M HCl was heated to reflux overnight. The crude mixture obtained after solvent evaporation under reduced pressure was purified by reverse phase HPLC to give the final product, “**cmpd 25**” (9%). HPLC program: in solvent gradient, starting with 90% acidified water (0.1% v/v formic acid) and 10% acetonitrile for one minute, followed by a linear increase of the acetonitrile fraction to 90% in 9 minutes. From minute 10 to 13 the acetonitrile fraction was kept at 90%. Then, acetonitrile was increased to 100% in one minute and maintained for 3 minutes, before returning to the initial (10%) proportion in 3-min linear gradient.

^1H NMR (400 MHz, MeOD): δ = 3.06-3.27 (m, H4, H6a, 2H), 3.40 (dd, H6b, $J_{6b,6a}$ = 2.8 Hz, $J_{6b,5}$ = 13.7 Hz, 1H), 3.80-3.85 (m, H3, H5, 2H), 3.40 (dd, H2, $J_{1\alpha,2}$ = 3.3 Hz, $J_{2,3}$ = 10.5 Hz, 1H), 4.68 (d, H1 β , $J_{1\beta,2}$ = 8.4 Hz, 0.3H), 5.01 (d, H1 α , 0.7H), 7.37 (t, H11', $J_{11',10'}$ = $J_{11',12'}$ = 7.3 Hz, 1H), 7.44-7.50 (m, H10', H12', H11, 3H), 7.54 (t, H4', $J_{4',3'}$ = $J_{4',5'}$ = 7.8 Hz, 1H), 7.68 (d, H9', H13', $J_{9',10'}$ = $J_{13',12'}$ = 7.3 Hz, 2H), 7.78-7.83 (m, H10, H3', H5', 3H), 8.03-8.07 (m, H12, 1H), 8.12 (s, H7', 1H).

^{13}C NMR (100.6 MHz, MeOD): δ = 45.43 (C6), 56.60 (C2), 71.45 (C5), 72.07 (C3), 73.62 (C4), 92.46 (C1), 127.12 (C7'), 127.31 (C3'), 128.12 (C9', C13'), 128.79, 128.83 (C11', C11), 129.97, 130.04 (C4', C10', C12'), 130.62 (C12), 131.14 (C5'), 135.37 (C10), 136.33 (C2', C9, C8), 141.61 (C6'), 141.76 (C7), 142.82 (C8'), 170.63 (C1').

HRMS calcd for $[\text{C}_{25}\text{H}_{24}^{35}\text{Cl}_2\text{N}_2\text{O}_7\text{S} + \text{H}]^+$: 567.0754. Found 567.0766.

4.4. Biochemical hexokinase activity assay – ADP-GloTM kinase assay

Evaluation of the potential of the synthesized compounds as HK2 inhibitors was performed using the same methodology described in the Chapter III, section 4.6, where the technical details are found.

Chapter V

Biochemical and Biological Evaluation of selected hit molecules⁴

⁴ Chapter V reports the experiments performed during a 5-month internship in Professor Nissim Hay's laboratory, at the Department of Biochemistry and Molecular Biology, the University of Illinois at Chicago, USA.

1. Overview

As a world leader research group in the investigation of HK2 roles, Hay *et al*/proved that HK2 is required for tumor initiation and maintenance, and that its ablation can be therapeutic and does not elicit adverse physiological consequences in mouse models of breast and lung cancer [18]. Moreover, HK2 ablation inhibited tumorigenesis and increased apoptosis in HCC [174] and prostate cancer [175].

Broadly, the set of experiments described throughout this chapter aimed to evaluate several potential HK2 inhibitors (PHK2I⁵) in more biologically complex environments, specifically using immortalized cell lines. Since the behavior of known potent rhHK2 inhibitors in cells is still not completely unveiled, the previously reported “cmpd 25” and “cmpd 34” [43] were tested in parallel. Toward this end, effects on cell proliferation, cell death, and glycolysis were evaluated. Also, an additional biochemical assay was performed to confirm the results previously obtained (*cf.* **Chapter 3**) concerning the inhibition of HK activity.

The next sub-sections (**1.1-1.4**) succinctly explore the fundamentals of the methods used to screen PHK2Is and find those that are effectively capable of inhibiting glycolysis (**Figure 53**).

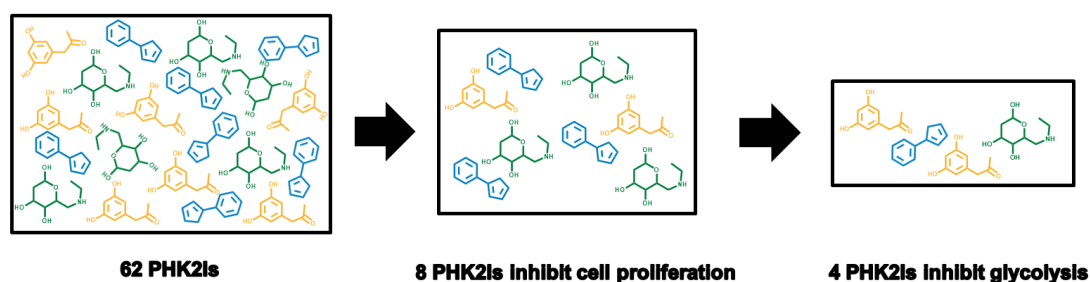


Figure 53. Screening of 62 PHK2Is yielded eight molecules capable of prominently inhibiting cell proliferation. From those, four were found to inhibit the glycolytic pathway.

1.1. Effect on cell proliferation and cell death

Basic cell culture procedures were used to maintain all cell lines. Cell proliferation was measured using an adherent cell cytometer, a labeled-free and non-destructive approach that uses a brightfield to identify cells (**Figure 54**), allowing to keep the same plate during the whole experiment, regardless of its duration. This is one of the great advantages of the method, ensuring a reliable normalization over time. However, the analysis requires customization of the parameters for each cell type/focus conditions. Using the same cytometer, cell death was detected based on the ability of cells to permeate Hoechst stain and propidium iodide (PI). Independently of the integrity of the membrane, Hoechst will

⁵ In order to simplify writing and reading, from this point on the set of potential HK2 inhibitors under study will be briefly referred as PHK2I.

be able to get into cells and intercalate DNA, generating fluorescence. However, since PI is not able to permeate viable cells, DNA is only intercalated by this dye when the cell membrane is compromised [253].

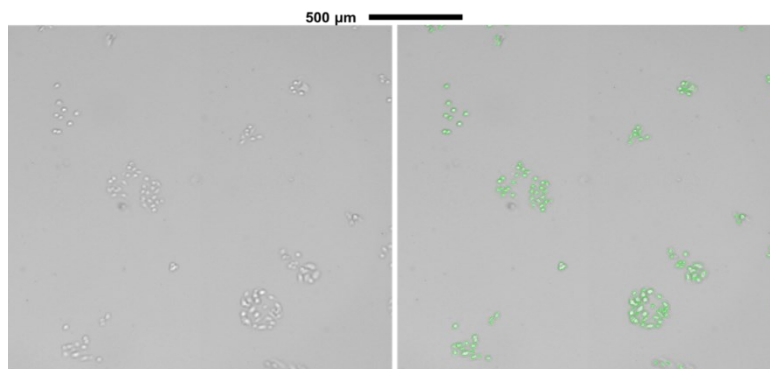
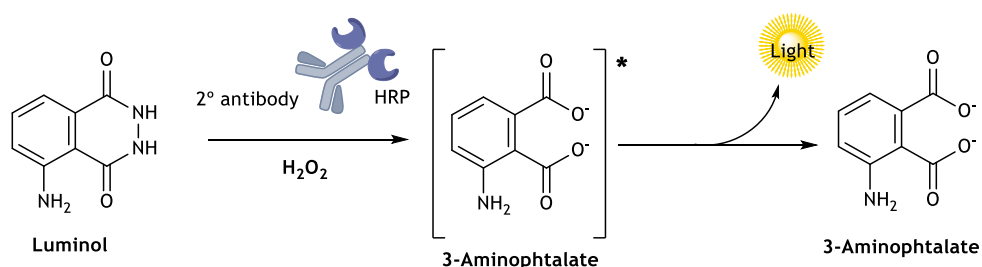


Figure 54. Detection of untreated PC3 cells using the brightfield of Celigo Adherent Cell Cytometer (Nexcelom Bioscience LLC, Lawrence, MA, USA). Image on the left shows PC3 cells (brighter round white spots) and, on the right, the image shows the ability of the method in identifying (green) those cells.

1.2. Immunoblotting

Immunoblotting was done using standard methods to analyze qualitatively the expression of several HK isozymes in different cell lines. Sodium dodecyl sulfate (SDS)–polyacrylamide gel electrophoresis (PAGE) was used to separate macromolecules by MW, and specific antibodies were applied to probe proteins of interest. The detection was carried out by chemiluminescence, after reaction of hydrogen peroxide with luminol, catalyzed by horseradish peroxidase (HRP) (conjugated with the secondary antibody) (**Scheme 42**). This method is highly sensitive and can detect proteins even in low the picogram range [254,255].



Scheme 42. Chemiluminescent reaction between luminol and hydrogen peroxide catalyzed by HRP, which is conjugated with the secondary antibody. The reaction generates an excited state product, which decays to a lower energy state releasing photons.

1.3. *Effect on glycolysis*

Efficient inhibition of the glycolytic pathway was the ultimate objective of this research project. Impact on glycolysis was measured by two different methods. At first, an extracellular flux analyzer was used to measure glycolysis as a function of extracellular pH change – extracellular acidification rate (ECAR). Additionally, the oxygen consumption rate (OCR), closely related with mitochondrial respiration, was measured concomitantly.

The Glycolysis Stress Test (GST) allows the measurement of several metabolic parameters in real-time, namely the glycolysis, glycolytic capacity, glycolytic reserve, and the non-glycolytic acidification, depending on the injection of glucose, oligomycin and 2DG (**Figure 55 - A**). Subsequently to the change of regular medium for glucose-free medium and the registration of the baseline for non-glycolytic acidification, glucose is injected to promote glycolysis. Oligomycin, an inhibitor of complex V of the electron transport chain, inhibits the oxygen consumption to generate ATP, which stimulates glycolysis. Finally, 2-DG inhibits glycolysis, dropping ECAR to the baseline [256].

Regarding the mitochondrial respiration, parameters such as the basal respiration, ATP production, proton leak, maximal respiration and spare capacity can be calculated from the cell mito stress test (CMST) data, using the cells' response to oligomycin, carbonyl cyanide-4-(trifluoromethoxy)phenylhydrazone (FCCP), and a mix of rotenone and antimycin A (**Figure 55 - B**). Briefly, oligomycin is injected first, to inhibit completely the ATP production, FCCP disrupts the mitochondrial potential allowing a maximal respiration, and finally, rotenone (inhibitor of complex I) and antimycin A (inhibitor of complex III) inhibit mitochondrial respiration completely [257].

Additionally, gas chromatography-mass spectrometry (GC/MS)-based metabolomics studies using isotopically labeled glucose were performed. Stable isotope-resolved metabolomics can provide information on the activity of a myriad of metabolic pathways and the contribution of certain molecules to specific metabolite pools [258]. Here, cells were incubated with [U-¹³C]glucose-containing medium for a known period. Using GC-MS, fractional enrichment of the intracellular metabolites, G6P, pyruvate and lactate, were compared between treated and control cells. Considering that labeled glucose allows the examination of stable ¹³C-enriched metabolites generated exclusively from glucose metabolism, a reduced percentage of labeled metabolites compared with control is correlated with inhibition of the glycolytic pathway.

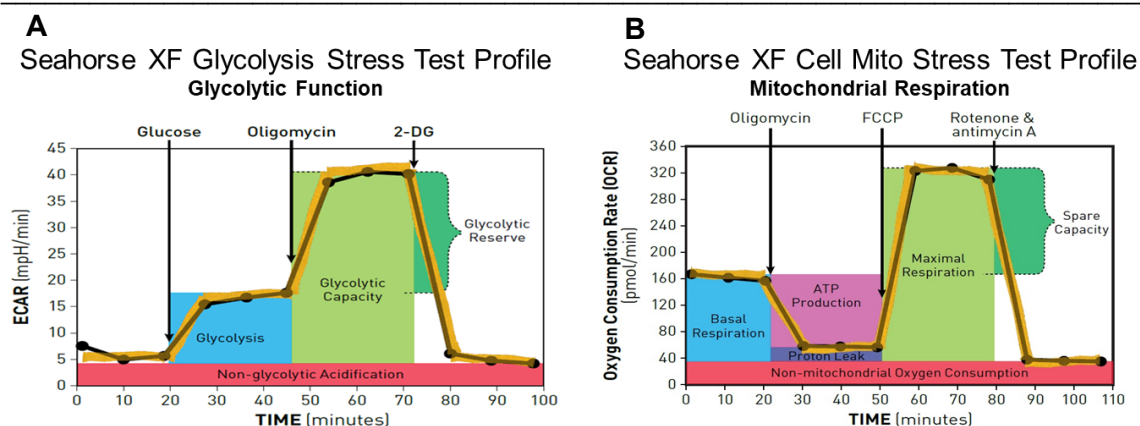
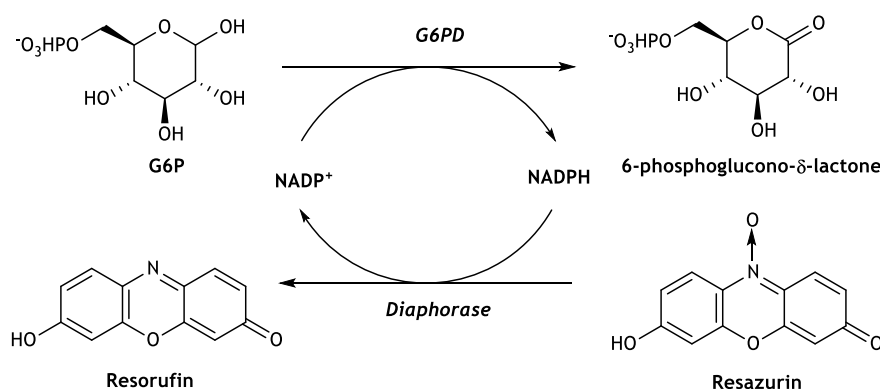


Figure 55. Expected profiles of glycolytic function (A) [256] and mitochondrial respiration (B) [257]. Arrows indicate the injection of each compound. Changes in pH/[O₂] are detected in real-time.

1.4. Biochemical Assay

HK activity in presence or absence of PHK2I was evaluated by means of a fluorescence-based assay using the resazurin-diaphorase system. The detection occurs after a couple of reactions subsequent to the reaction of interest. HK catalyzes the conversion of glucose into G6P. To detect the amount of G6P formed, glucose-6-phosphate dehydrogenase (G6PD) is used to reduce NADP⁺ to NADPH in the presence of G6P that is oxidized to 6-phosphoglucono- δ -lactone. NADPH is then used by the resazurin-diaphorase system to generate the highly fluorescent resorufin. This fluorescence is proportional to the amount of G6P formed and can be quantified (**Scheme 43**).



Scheme 43. Oxidation of G6P to 6-phosphoglucono- δ -lactone coupled with resazurin reduction to the fluorescent resorufin via NADP⁺/NADPH recycling.

2. Results and Discussion

2.1. Selection of cell lines

Herein PC3, Huh7 and M15-4 CHO cells were used to evaluate PHK2Is. Previous results on HK2 silencing have shown a 50% inhibition of cell proliferation of Huh7 and PC3 cells after 7 days [174,175] (**Figure 56**). HCC cell lines have an interesting characteristic that might be extremely useful when HK2 is targeted. As mentioned in chapter II, section 2.1, liver and pancreas express the tissue-specific GCK. Along with cancer initiation, expression of GCK is exchanged for HK2. **Figure 57** shows the expression of HK1, HK2 and GCK in different cell lines, including three HCC cell lines: Huh7, HepG2 and Hep3B. Only residual amounts of GCK were detected, contrary to HK2 that is visibly expressed in the HCC cell lines. Moreover, HK1 was not detected. Thus, Huh7 cells were selected as a representative HCC cell line to start the screening.

Highly glycolytic PC3 cells (**Figure 58**) from metastatic prostate cancer greatly express HK, specifically both HK1 and HK2 (**Figure 57**). The effect of PHK2Is on this cell line was also tested, although a lower effect was to be expected, should selectivity for the HK2 isozyme be confirmed.

In regular conditions M15-4 CHO cells present low glycolytic rates, due to their lack of hexokinase activity, while other glycolytic enzymes are still expressed in normal levels [259]. For this reason, different cell lines were generated to express selectively the desired HK [260], allowing the study of the glycolytic behavior of each cell type under different conditions. In this way, available engineered M15-4 CHO cells expressing HK1, HK2 or GCK wild type (WT) were used to evaluate the selectivity of test compounds to inhibit glycolysis. CHO cells with an empty vector (EV) were used as control. **Figure 57** shows the expression of each isozyme in the four cell lines and **figure 59** shows their glycolytic profile. The CHO EV cell line was confirmed to have a low glycolytic activity compared to the other generated cell lines. CHO GCK showed an intermediate glycolytic capacity, with inability to respond to oligomycin, and CHO HK2 and CHO HK1 presented high (and similar) glycolytic behavior.

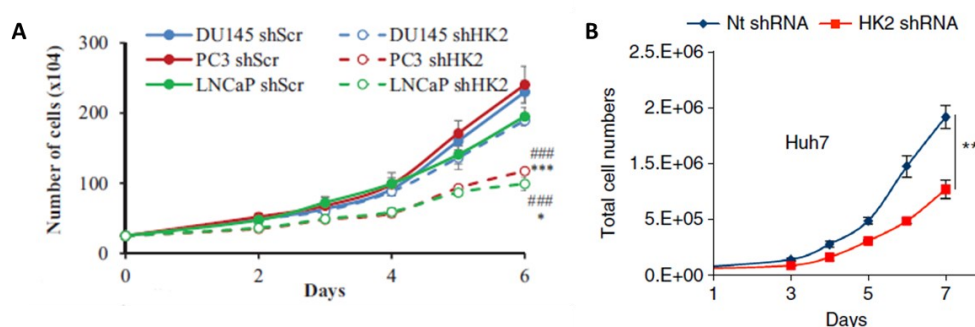


Figure 56. HK2 ablation reduces proliferation by 50% after 7 days treatment in prostate cancer cells LNCaP and PC3 (**A**) and HCC cells (Huh7) (**B**) cells. Effect of HK2 silencing in prostate cancer cells DU145 (**A**) was not significant. Adapted from Nogueira et al (2018) [175] (**A**) and DeWaal et al (2018) [174] (**B**). shScr - cells expressing an inducible control (Scr); shHK2 - HK2 silenced cells; Nt (control) and HK2 (HK2 silenced cells) shRNA indicate Dox-inducible non-targeting and HK2 shRNA, respectively; ### $p < 0.0003$ versus shScr; Error bars correspond to SEM; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ versus shScr; by Student's t-test.. See literature [174,175] for more information.

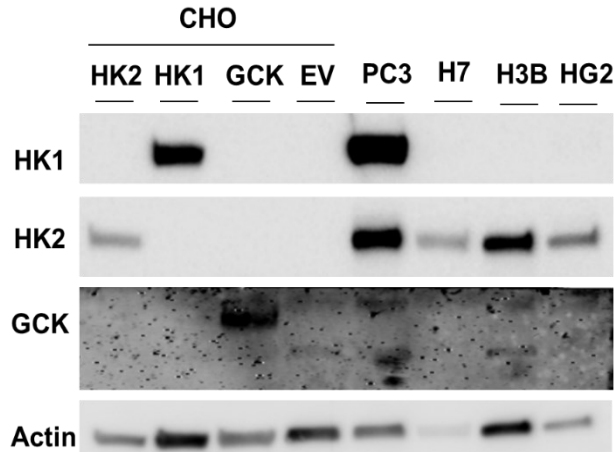


Figure 57. Expression of each HK on PC3, H7 (Huh7), H3B (Hep3B), HG2 (HepG2), and engineered CHO cells monitored by immunoblotting. CHO EV cells possessed a residual amount of GCK.

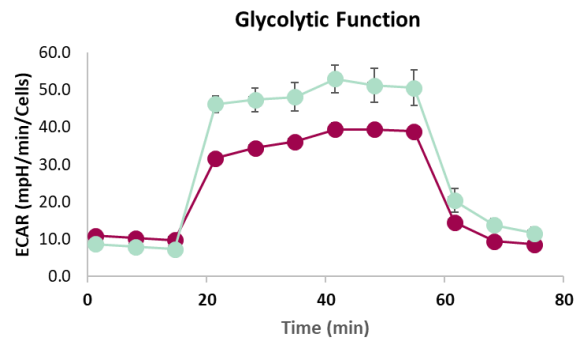


Figure 58. Glycolytic function profile of PC3 (green) and Huh7 (pink) cells based on changes in the extracellular acidification rate (ECAR) in response to injection of 10 mM glucose, 1 μ M oligomycin and 50 mM 2DG. Error bars correspond to SEM.

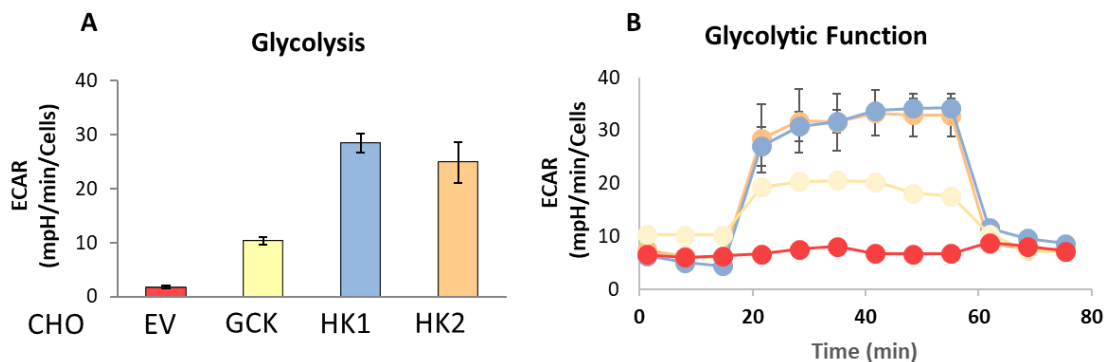


Figure 59. Engineered M15-4 CHO expressing HK2, HK1 or GCK (WT) and EV present different glycolysis rates and glycolytic function profiles. **A** - Glycolytic response of CHO cells to injection of 10 mM glucose. **B** - Glycolytic function profile of CHO cell lines (HK1 –blue, HK2 –orange, GCK – yellow, EV – red), based on changes in the extracellular acidification rate (ECAR) in response to injection of 10 mM glucose, 1 μ M oligomycin and 50 mM 2DG. Error bars correspond to SEM.

2.2. Cell proliferation and cell death

2.2.1. Cell proliferation in Huh7 cells

In order to compare the effects of HK2 ablation and possible HK2 inhibition on cell proliferation, Huh7 cells were incubated with “cmpd 25” and “cmpd 34” [43] at 1, 5 and 10 μM and kept for 7 days after treatment. A long incubation period was elected due to the previous information on the effect of HK2 silencing on Huh7 cell proliferation [174], highlighting that the effect of HK2 inhibition is not expected to be immediate. Results and structures of “cmpd 25” and “cmpd 34” are shown on **figure 60**, showing that small changes in the structure can affect the response significantly. While “cmpd 25” was able to inhibit proliferation by 40% at 10 μM , “cmpd 34” only inhibited by 25%, after seven days. As found before with silencing of HK2 [174], the effect only started to be noticeable after five days.

Since HK1 expression was not detected in Huh7 cells, the better behavior of “cmpd 25” compared with “cmpd 34” cannot be attributed to its ability to inhibit both HK1 and HK2 at low nanomolar level. Then, the impact of the minor structural change on pharmacokinetics might be the key for the observed differences in cell proliferation. Regarding the structure of “cmpd 34”, the chlorine group is speculated to block metabolism on the second ring of the biphenyl group. The corresponding position in “cmpd 25” is available to undergo metabolism, which may generate active metabolites that could better bind HK2 or even other unknown targets.

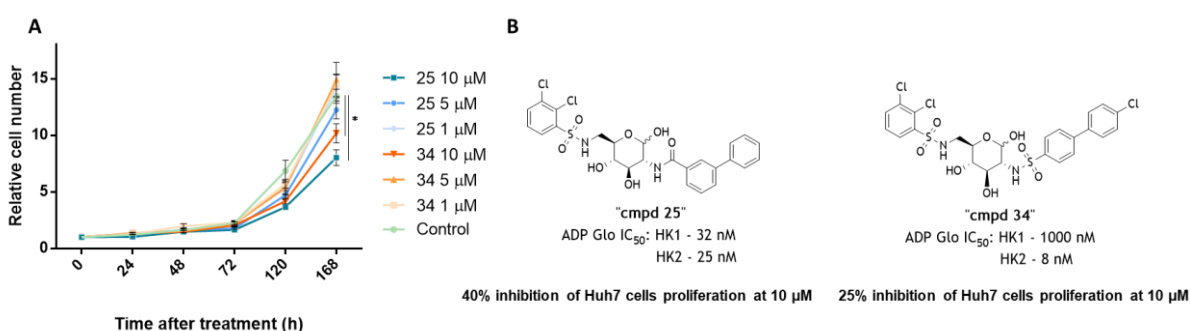


Figure 60. **A** - Effect of “cmpd 25” and “cmpd 34” on Huh7 cells proliferation for 168h (seven days), at 1, 5 or 10 μM , compared with control cells treated with vehicle (DMSO). **B** - Structures of “cmpd 25” and “cmpd 34” and IC_{50} reported for the inhibition of HK1 and HK2 with the ADP-GloTM kit [43]. Error bars correspond to SD.

Considering these observations, Huh7 cells were incubated with 62 PHK2Is and “cmpd 25” for seven days at 10 μM (**Figure 61**). Compounds were labeled from 1 to 63 to facilitate all procedures and analysis⁶. The selection of PHK2Is among all molecules available was based mostly on the purity of the compound assessed earlier and also on structural diversity.

⁶ See annex 3 for information about the correspondence between code number here implemented and the molecule code and respective structure of all the test molecules.

In this set, “cmpd 25” is represented by number 3. From the whole set, 13 compounds were considered to inhibit proliferation of Huh7 cells with comparable or stronger effect than “cmpd 25”. With few exceptions (compounds 6, 7, 41 and 60), consistent reduction of cell proliferation was not observed before 72h after treatment. From those 13 compounds, eight molecules showing a stronger effect than “cmpd 25” were selected for further studies, starting with the calculation of IC₅₀. Since culture medium was changed every 72h, IC₅₀s were calculated after six days of treatment and not seven, with the purpose of saving compounds, some of them available in very limited amounts. **Table 7** shows the IC₅₀ values found for each selected PHK2I and “cmpd 34” along with corresponding SD⁷. Due to availability constraints, the IC₅₀ of “cmpd 25” was not determined and “cmpd 34” was included in the set to compare the concentration-response curves of PHK2Is with a known inhibitor. Different ranges of IC₅₀ were found: PHK2I 6, 7, 21, 41 and 60 were shown to inhibit proliferation by 50% at concentrations < 3.6 μ M; an intermediate IC₅₀ (4.77 μ M) was found for PHK2I 40; and the IC₅₀s of PHK2I 10 and 18 and “cmpd34” were found to be higher than 10 μ M (13.77, 11.50 and 12.91 μ M, respectively). These values were particularly useful to plan subsequent experiments, namely cell death measurements. Inhibition of cell proliferation can be directly related with cell death caused by multiple pathways, implying cytotoxicity by the PHK2I and off-target interactions, since total ablation of HK2 slowly reduces cell proliferation over time and does not have an immediate effect. On one hand, compounds with low micromolar IC₅₀s might have high toxicity and a hypothetical short therapeutic window. On the other hand, if the molecule is highly selective for the desired target, even lower concentrations can be used to obtain a solid therapeutic response. Molecules associated with higher IC₅₀s are probably safer; however, the desired effect might be reduced or take longer (implying multiple dosages) to occur.

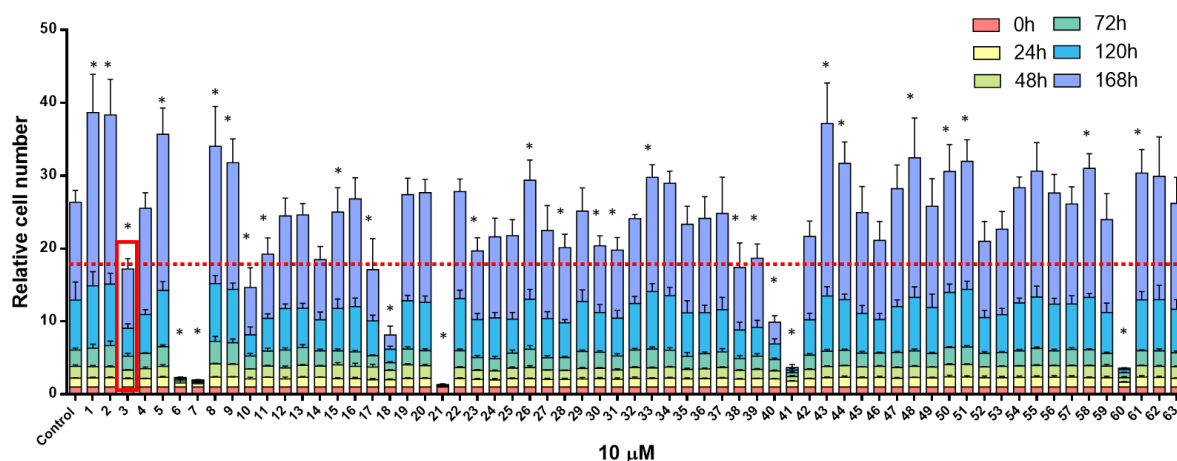


Figure 61. Effect of 62 PHK2Is and “cmpd 25” on Huh7 cells proliferation after 7 days at 10 μ M. Each color corresponds to a measurement at a certain timepoint (0, 24, 48, 72, 120 or 168h). “Cmpd 25” is represented by number 3 (highlighted in red). To compare the effect of each PHK2I with “cmpd 25”, a red dashed line is drawn across all PHK2I bars, marking the maximum relative cell number obtained with cells treated with “cmpd 25”. Error bars correspond to SD. The asterisks represent significant differences compared to the control ($p < 0.05$). Number of cells is normalized according with the cell count on day 0 (first day of treatment with DMSO, “cmpd 25” or PHK2Is).

⁷ See annex 4 for the concentration-response curves.

Table 7. IC_{50} ($\pm SD$) found for the best inhibitors of Huh7 cells proliferation after 6 days of treatment.

PHK2I	IC_{50} (μM)
“Cmpd 34”	12.91 ± 1.50
6	3.17 ± 1.14
7	3.56 ± 1.08
10	13.77 ± 1.13
18	11.50 ± 1.09
21	3.42 ± 1.64
40	4.77 ± 1.47
41	1.62 ± 1.52
60	1.93 ± 1.06

2.2.2. Cell death in Huh7 cells

As previously reported, HK2 ablation or inhibition is considered a potential coadjuvant in cancer chemotherapy [17,38,174,176,261]. With this in mind, Huh7 cells' death was evaluated in the presence or absence of sorafenib, the standard of care for patients with advanced hepatocellular carcinoma [262], with concomitant treatment with PHK2Is. Selected molecules for this experiment included the ones most effective at inhibiting cell proliferation and also “cmpd 25”, “cmpd 34” and compounds 4 and 5, two PHK2Is that have shown high potential to inhibit rhHK2 [43].

Huh7 cells were treated for 48h with the selected PHK2Is at the IC_{50} concentrations previously obtained for cell proliferation, except for “cmpd 25” - 13 μM (IC_{50} was not determined and IC_{50} found for “cmpd 34” was used) and PHK2Is 4 and 5 – 10 μM (IC_{50} was not determined, but those compounds had already shown potential to inhibit rhHK2 (**chapter III, section 2.6.2.; annex 2**), interestingly PHK2I 5 showed higher cell growth than the control (**Figure 61**). In parallel, another experiment was performed using 5 μM sorafenib, in addition to the treatment with each PHK2I. After 48h, the cells were incubated with Hoechst and PI dyes and the total number of cells and the number of dead cells were registered (**Figure 62**). While the total number of cells treated with different PHK2I exhibited noticeable differences in the absence of sorafenib, the numbers were approximately the same in the presence of sorafenib. Regarding cell death, sorafenib alone was able to cause ~20% cell death. Moreover, in the presence of sorafenib, only PHK2Is 10 and 18 did not significantly increase cell death compared with the sorafenib control, with most test molecules seemingly showing a cumulative effect of the concomitant treatment. A synergetic effect was clearly observed with “cmpd 34” and to some extent with “cmpd 25”. As a synergy between PHK2Is 4 and 5 and sorafenib was also observed, effect of these molecules should be further explored.

Notwithstanding, these results should be regarded with caution, as the method showed a considerable drawback. Dead cells detach from the bottom of the well, and sometimes the adherent cell cytometer is not able to detect fluorescence from different planes. In this way, to find reliable results for the number of dead cells throughout the experiments was a challenge that was not completely overcome, as the error bars in **Figure 62** evince. Catching cell death in the earliest possible stage would be essential to perform a better analysis of the results but comparing the effects of different molecules is always tricky, as the onset surely occurs at different timepoints. An alternative to PI would be the use of viability dyes, such as calcein acetoxymethyl ester (calcein-AM), to identify live cells, with concomitant use of Hoechst stains. Cell death would then be estimated by the subtraction of live cells from the total number of cells. Unfortunately, time constraints prevented further experiments using this method.

Nonetheless, most of the test molecules can be considered reasonably non-toxic in these conditions, not provoking cell death in high rates and having little effect on cell proliferation after 48h, with the exception of compound 6, 7 and 60 (*cf.* **Figure 61**). These molecules are presumably inducing cell death by other unknown pathways.

Interestingly, “cmpd25” showed a stronger effect on Huh7 cells’ viability than the one demonstrated in the cell proliferation experiments. Nevertheless, as discussed before, it was not possible to calculate IC_{50} for “cmpd 25” and the concentration used for this assay was 13 μM (the IC_{50} obtained for “cmpd, 34”), which may have led to a more pronounced effect.

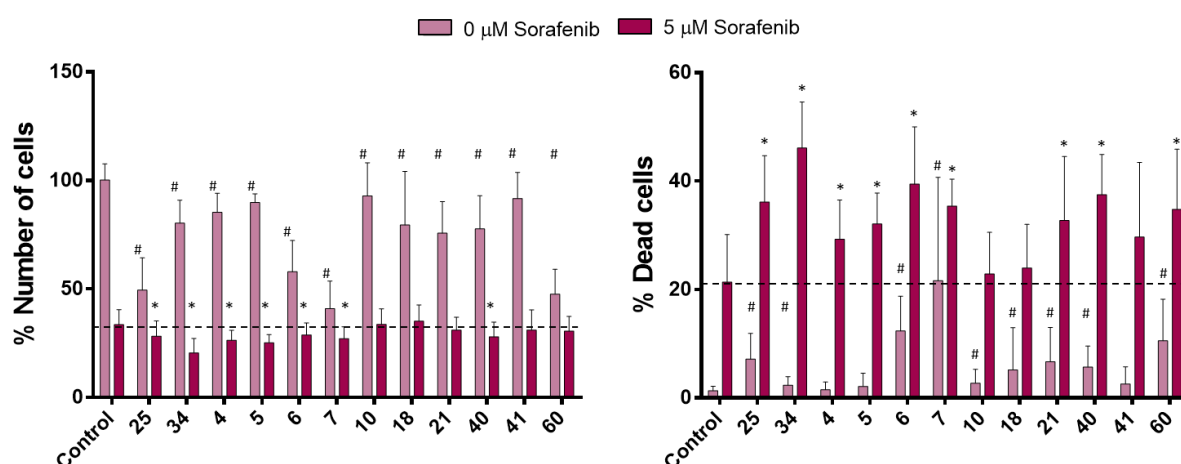


Figure 62. Cell death in Huh7 cells after 48h of treatment with PHK2Is in the presence or absence of 5 μM sorafenib. Cells were treated with IC_{50} concentrations of PHK2I previously determined for cell proliferation in the same cells. On the left, the total number of cells was measured using Hoechst dye. On the right, the percentage of dead cells was measured using PI. The number 25 represents “cmpd 25” and 34 represents “cmpd 34”. # - $p < 0.05$ (Student’s t) compared to the 0 μM Sorafenib control; * $p < 0.05$ (Student’s t) compared to the 5 μM Sorafenib control; Error bars correspond to SD.

2.2.3. Cell proliferation and cell death in PC3 cells

As abovementioned, PC3 cells express both HK1 and HK2, whereby a lower effect on cell proliferation and cell death by PHK2I is expected. Here PC3 cells were incubated for 6 days with PHK2I, at 10 μ M, and cells were counted after 72h and 144h of treatment. For this experiment, the set of test molecules consisted of the same molecules used to evaluate cell death in Huh7 cells plus PHK2I 1 and 2. These latter molecules were synthesized, rather than procured, and the objective of their inclusion was to confirm the lack of significant biological effects in these cells, since there was no evidence of their ability to inhibit rhHK2 or cell proliferation in Huh7 cells.

Figure 63 illustrates the results. In fact, PHK2Is 1, 2, 4, and 5 did not show any ability to inhibit PC3 cells proliferation after either 72 or 144h. Likewise, known HK2 inhibitors (“cmpd 25” – 25, “cmpd 34” – 34) did not affect cell proliferation either. Still as part of this group, PHK2I 10 did not reduce cell proliferation, and even significantly increased it. Generally, the compounds that affected proliferation in the first 72h were the same after 144h, while in Huh7 cells some compounds only showed their impact after the longer period. Showing a considerable effect on Huh7 cells, PHK2Is 18 and 40 affected PC3 cell proliferation although by less than 50%. Once more (*cf. section 2.2.1*), PHK2Is 6, 7, 21, 41 and 60 have shown to be the most potent molecules, and thus are probably not selective for HK2, at the test concentration.

To further characterize the effect of these molecules, cell death was measured after 48h of treatment with PHK2Is at 10 μ M. Despite of lack of inhibitory effect on cell proliferation, in this assay a slight reduction of the number of cells compared with the control was found on the wells treated with “cmpd 25”, “cmpd 34”, and PHK2Is 1, 2 and 4. On the other hand, after 48h, PHK2Is 18 and 40 had a relative minor effect in the number of cells. PHK2I 6, 7, 21 and 60 were able to reduce the number of cells by more than 50%, while, interestingly, 41 did not dramatically affected the relative number of cells. Regarding the number of dead cells, almost all samples registered values lower than 5%, with the exception of 21 that reached more than 10% of cell death.

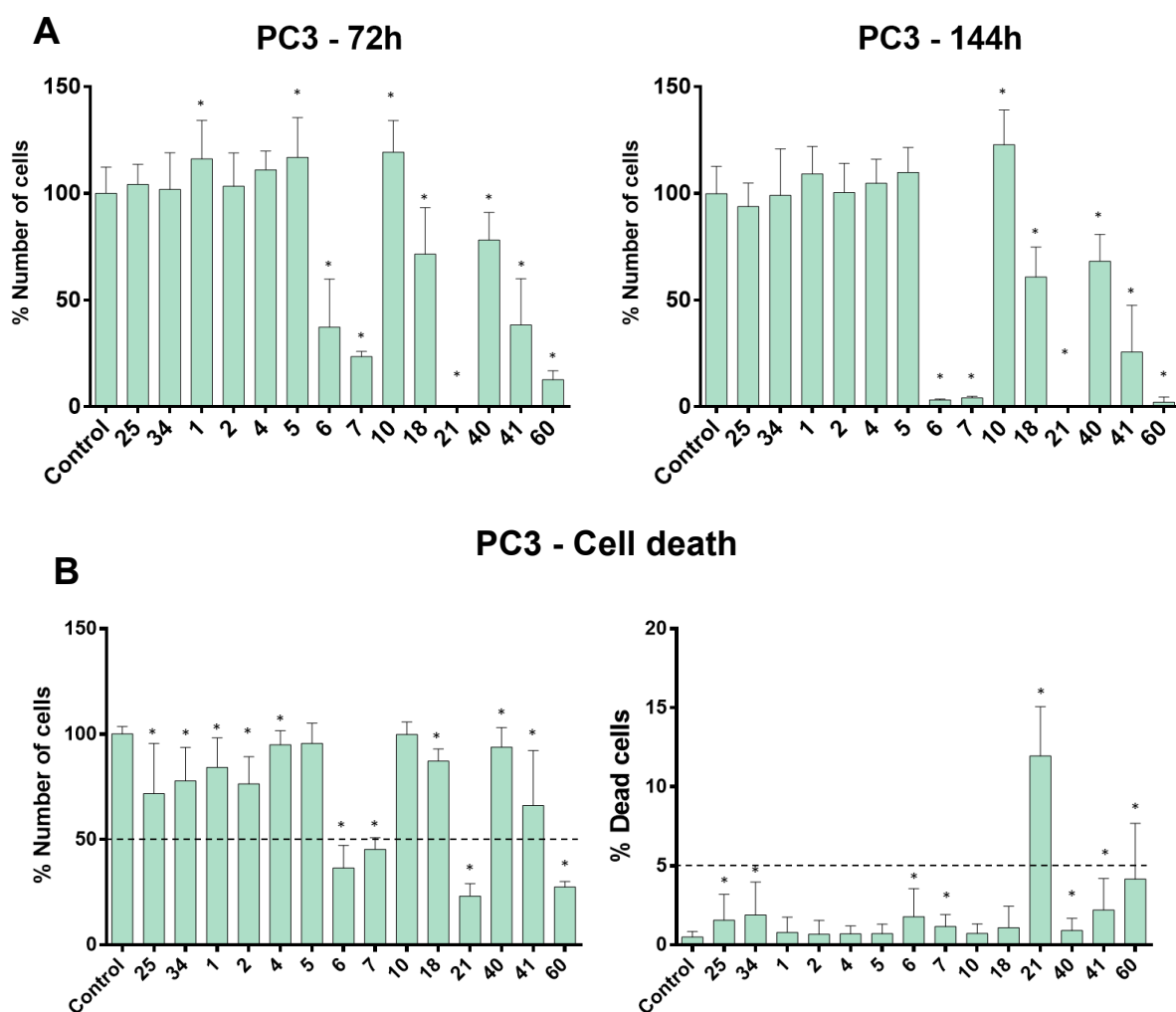


Figure 63. Effect of PHK2Is (10 μ M) in PC3 cells' proliferation (**A**) and death (**B**). **A** - Effect of PHK2Is in the number of cells after 72h (left) and 144h (right) of treatment. **B** - Results of the cell death experiment, represented by the total number of cells (left) and the number of dead cells (right) compared with control. The number 25 represents "cmpd 25" and 34 represents "cmpd 34". * $p < 0.05$ (Student's t) compared to the control; Error bars correspond to SD.

2.2.4. Cell proliferation and cell death on CHO cells

The same procedures and test molecules mentioned above for PC3 cells were used to evaluate CHO cell proliferation and cell death, and the results obtained were very similar as well (**Figure 64**). The main difference between the results from two types of cells was the effect caused by compounds 18 and 40. On one hand, PHK2I 18 exhibited a relative minor effect on cell proliferation, and on the other hand PHK2I 40 inhibited cell proliferation by more than 50%, a result comparable to that obtained with PHK2I 41, which was more effective in the experiments described in the previous section. Considering the cell death experiment, the total number of cells was mostly comparable to results found for PC3 cells, with exception of PHK2Is 40. PHK2I 40 showed to decrease the total number of cells to some extent. Generally, the number of dead CHO cells was higher than the number of dead PC3 cells, being more

pronounced in cells treated with PHK2Is 6, 7, 21 and 60 (12-25%); remarkably, 21 caused death in more than 70% of CHO HK1 cells. Once again, it should be noted that the percentage of cell death should be further confirmed, due to the high variance found in the experiments.

Overall, CHO cell lines behaved similarly in both experiments (**Figure 64**). CHO HK1 cells seemed slightly more sensitive than the other cell lines to cell proliferation effects, except for PHK2Is 10, 18, 41 and 60; still, the differences are hardly noteworthy. Considering that CHO cells are normal ovary cells and that both cell proliferation and cell death were mostly affected by the same compounds that affected PC3 cells (prostate cancer cells), it can be concluded that PHK2Is 6, 7, 21, 40, 41 and 60 did not display any selectivity for specific cancer cells at the test concentration and might be associated with potential general cytotoxicity. These molecules might interfere with other targets or pathways to produce cell death or reduction in cell proliferation, since CHO survival and proliferation are known not to rely on glycolysis. From the entire set of molecules, only compound 10, “cmpd 25” and “cmpd 34” did not show cytotoxicity. Compound 18 also showed evidence of low toxicity and some selectivity for cancer cells.

2.2.5. Effect of PHK2I 21

PHK2I 21 (**Figure 65 - A**) showed to be a highly toxic molecule and behaved similarly in all cell lines in study, killing most of the cells treated at 10 μ M after 24h, and virtually all cells by 72h, when only cell debris were detected (**Figure 65 - B**). With the aim to determine the IC₅₀ of compound 21 regarding Huh7 cells' proliferation, a dose-response experiment was performed using 6 μ M as the highest test concentration. Interestingly, cell proliferation was maintained unchanged until the concentration was raised to 3 μ M, where a decline was observed, with almost 100% cell death at 4 μ M (**Figure 65 - C**). The downhill slope was very sharp (-16.94) in this concentration range, representing a small window to work with this compound, for which even minor pipetting errors could have a huge impact on the results obtained. Despite PHK2I 21 being part of the NCI database, no findings have been reported on this molecule. The effect of 21 is particularly intriguing, and its mechanism of action merits investigation, particularly regarding its potential as a cytotoxic agent. This molecule could have interesting uses as a positive control for cell death or related roles that might be of use to improve other research methodologies.

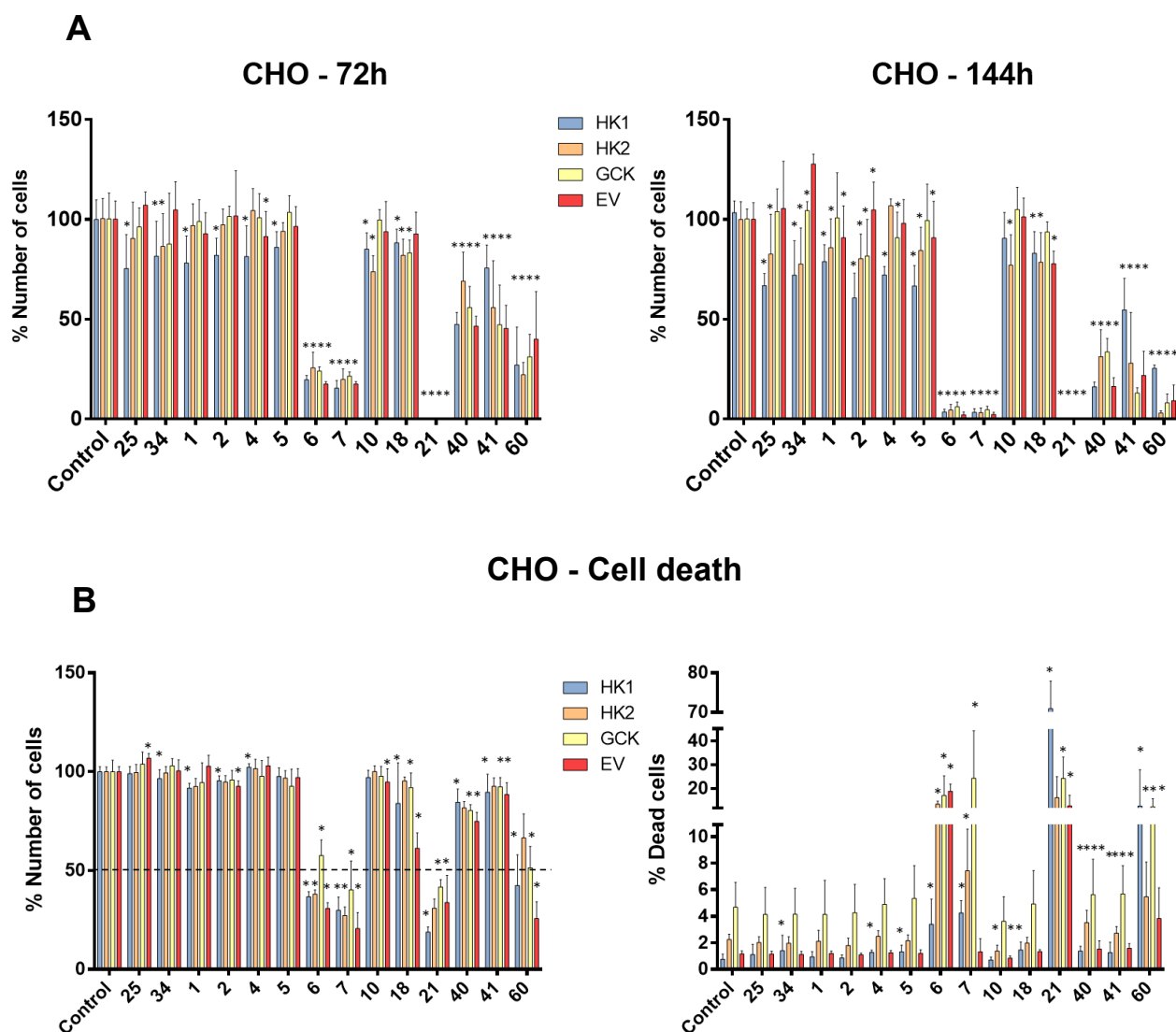


Figure 64. Effect of PHK2Is at 10 μ M in CHO cell lines' proliferation (**A**) and cell death (**B**). **A** - Effect of PHK2Is in the number of cells after 72h (left) and 144h (right) of treatment. **B** - Results of the cell death experiment represented by the total number of cells (left) and the number of dead cells (right) compared with control. In each group, bars correspond to CHO HK1 (blue), CHO HK2 (orange), CHO GCK (yellow) and CHO EV (red) cells, respectively. The number 25 represents "cmpd 25" and 34 represents "cmpd 34". * $p < 0.05$ (Student's t) compared to the control; Error bars correspond to SD.

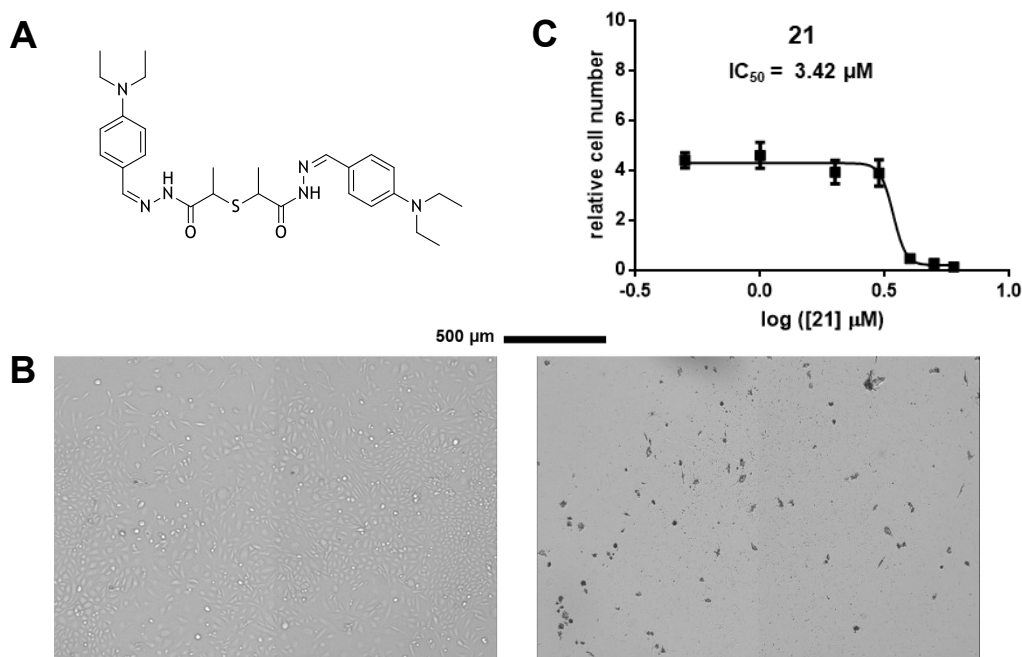


Figure 65. PHK2I 21 provoked complete cell death in short periods of time, at low micromolar concentrations. **A** - 2D structure of PHK2I 21. **B** - CHO HK2 cells after 72h of treatment with PHK2I 21 (10 μ M) (right) and control CHO HK2 cells without treatment (left). **C** - Concentration-response curve for Huh7 cells' proliferation treated with PHK2I 21. Error bars correspond to SEM.

2.3. Glycolysis stress test

With the aim to measure the effect of selected PHK2Is on glycolysis, Huh7, PC3 and CHO cells underwent a glycolysis stress test (GST). Firstly, the effect of a chronic treatment was tested, and then an acute response was evaluated, with a total of three conditions tested (**Figure 66**).

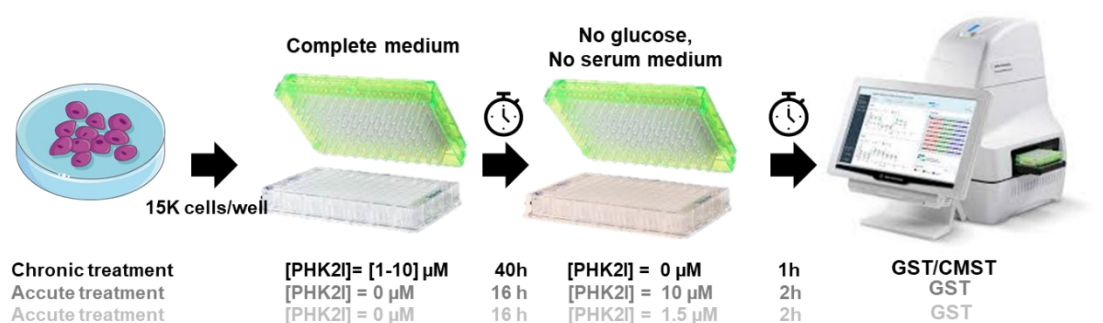


Figure 66. Schematic summary of the protocols applied for GST and CMST experiments.

2.3.1. Effect of chronic treatment – GST and CMST

Preliminary assays were performed with the Huh7 cells used for proliferation assays, after 9 days of treatment. After the 7 days required for proliferation measurements, cells were re-plated, and approximately 40 hours later they underwent GST and CMST in parallel, without the presence of PHK2I. Thus, a cumulative cell treatment response instead of an acute response was measured for a few selected compounds in the Huh7 cells, seeking to evaluate the impact of each PHK2I on the cells' behavior over time, using a suitable range of concentrations considering the IC₅₀ of each PHK2I for cell proliferation. In this assay, ECAR and OCR are measured while several injections with glycolysis or respiration modulators occur at specific time points.

Figure 67 show that ECAR and OCR profile of cells treated with “cmpd 25” and “cmpd 34” at 1, 5 and 10 μ M. Despite considerable variation between replicates found for “cmpd 25”, the glycolytic function and mitochondrial respiration had similar profiles. At the lower concentrations, 1 and 5 μ M, the inhibitory effect of compounds was absent, and they even showed to slightly enhance ECAR and OCR compared with control cells, as a signal of higher glycolytic and respiratory capacity.

Although glycolysis seems to have been inhibited in cells treated with 10 μ M of each compound, the difference between baselines should be noted, since glycolysis is calculated as the difference between the highest ECAR value measured before oligomycin injection and the last measurement before glucose injection. **Figure 68** illustrates ECAR change as a measure of glycolysis occurring in cells treated with different PHK2Is. “Cmpd 34” seems to inhibit glycolysis to some extent but the trend is not statistically significant. Regarding mitochondrial respiration, basal respiration was indeed diminished in cells treated with “cmpd 25” and “cmpd 34” at 10 μ M. Basal respiration corresponds to the difference between the last measurement before oligomycin injection and the lowest measurement after rotenone/antimycin A injection, which is mostly the same for all samples.

In the same conditions, a few PHK2I with previously found ability to reduce Hu7 cells proliferation were evaluated (*cf section 2.2.1.*), using a suitable range of concentrations considering the IC₅₀ found for cell proliferation for each molecule. **Figure 69** shows the glycolytic function and mitochondrial respiration profile for cells treated with PHK2Is 6, 18, 21, 40 and 60. In this set, only experiments with PHK2I 40 were repeated due to technical constraints. Generally, the PHK2Is changed the mitochondrial respiration profile, by reducing the OCR during the whole test. As an exception, PHK2I 18 did not reduce the OCR to a great extent, even at 10 μ M. GST revealed that, as found with the known inhibitors mentioned above (*cf. Figure 67*), these molecules reduced the baseline ECAR, specially at higher concentrations. This change was basically negligible with PHK2Is 18 and 21. Looking at **Figure 68**, the change in ECAR is translated into an impact on glycolysis, which was only reduced with PHK2I 40 at 10 μ M and 60 at 2 μ M. Interestingly, at 5 μ M PHK2I 40 was not able to reduce ECAR or OCR, and at 1 μ M it was even able to increase those rates. At 1 μ M, PHK2I 60 was incapable to decrease ECAR, but still impacted OCR. Despite the preliminary character of this experiment, 60 showed a potential to inhibit glycolysis and mitochondrial respiration at low micromolar concentrations.

Since the assay is performed in the absence of PHK2I, the residual PHK2I present inside the cells is postulated not to have been sufficient to inhibit HK2 during GST for most compounds, as happened before with the known inhibitors “cmpd 25” and “cmpd34”. Still, an evidence of change in cells behavior was observed upon treatment with these molecules.

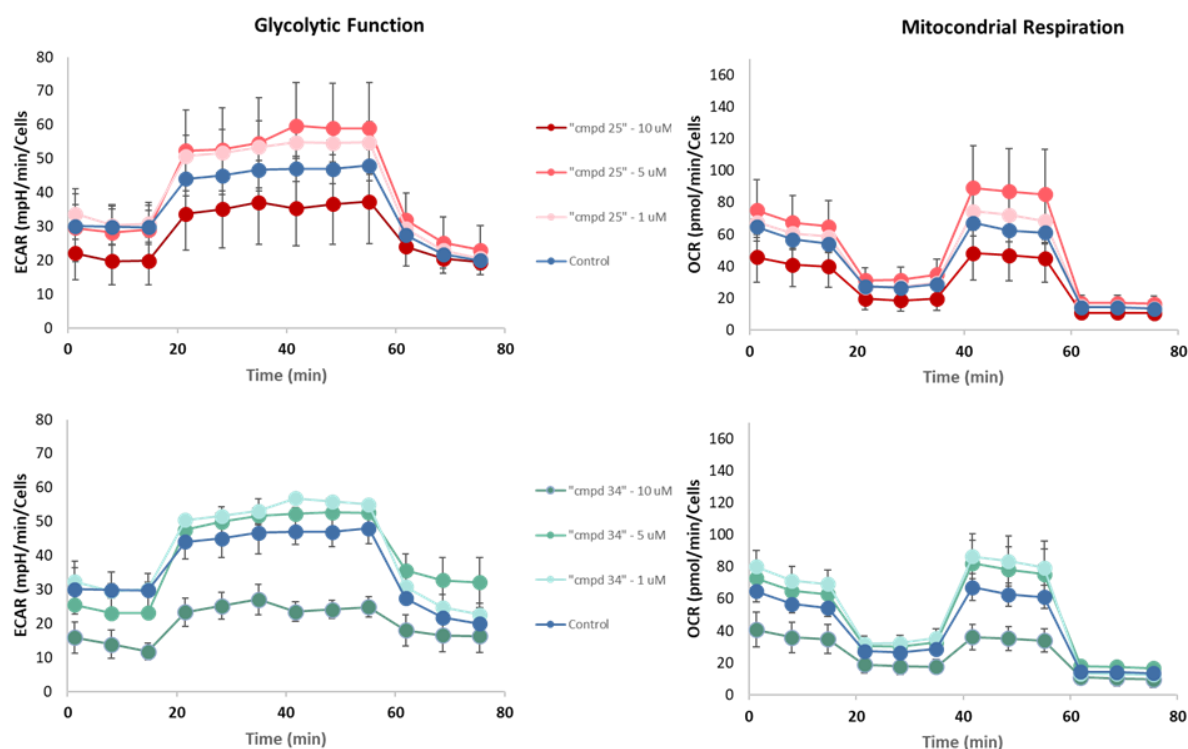


Figure 67. Glycolytic function and Mitochondrial respiration profiles of Huh7 cells treated with “cmpd 25” and “cmpd 34” (1, 5, and 10 μ M) compared with control cells. Error bars correspond to SEM.

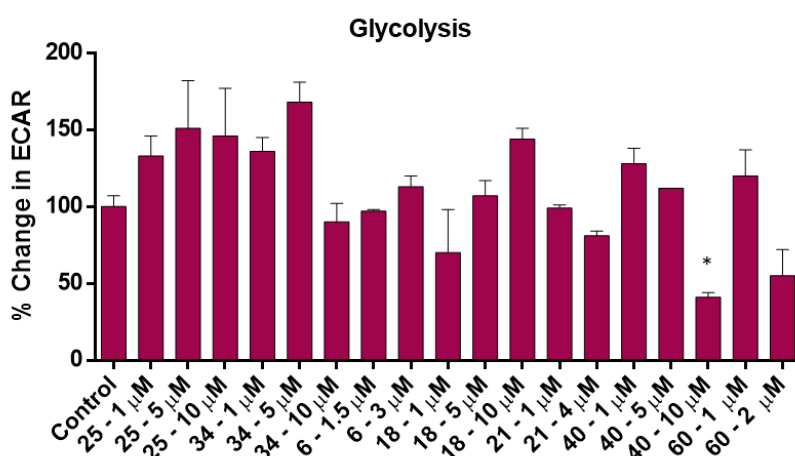


Figure 68. Glycolysis, assessed as a measure of % change in ECAR compared with control. The number 25 represents “cmpd 25” and 34 represents “cmpd 34”. * $p < 0.05$ (Student’s t) compared with the control. Results were obtained from one-time experiment (triplicates or quadruplicates) for PHK2Is 6, 18, 21, and 60. Results with “cmpd 25”, “cmpd34” and PHK2I 40 were obtained from two independent experiments). Error bars correspond to SEM.

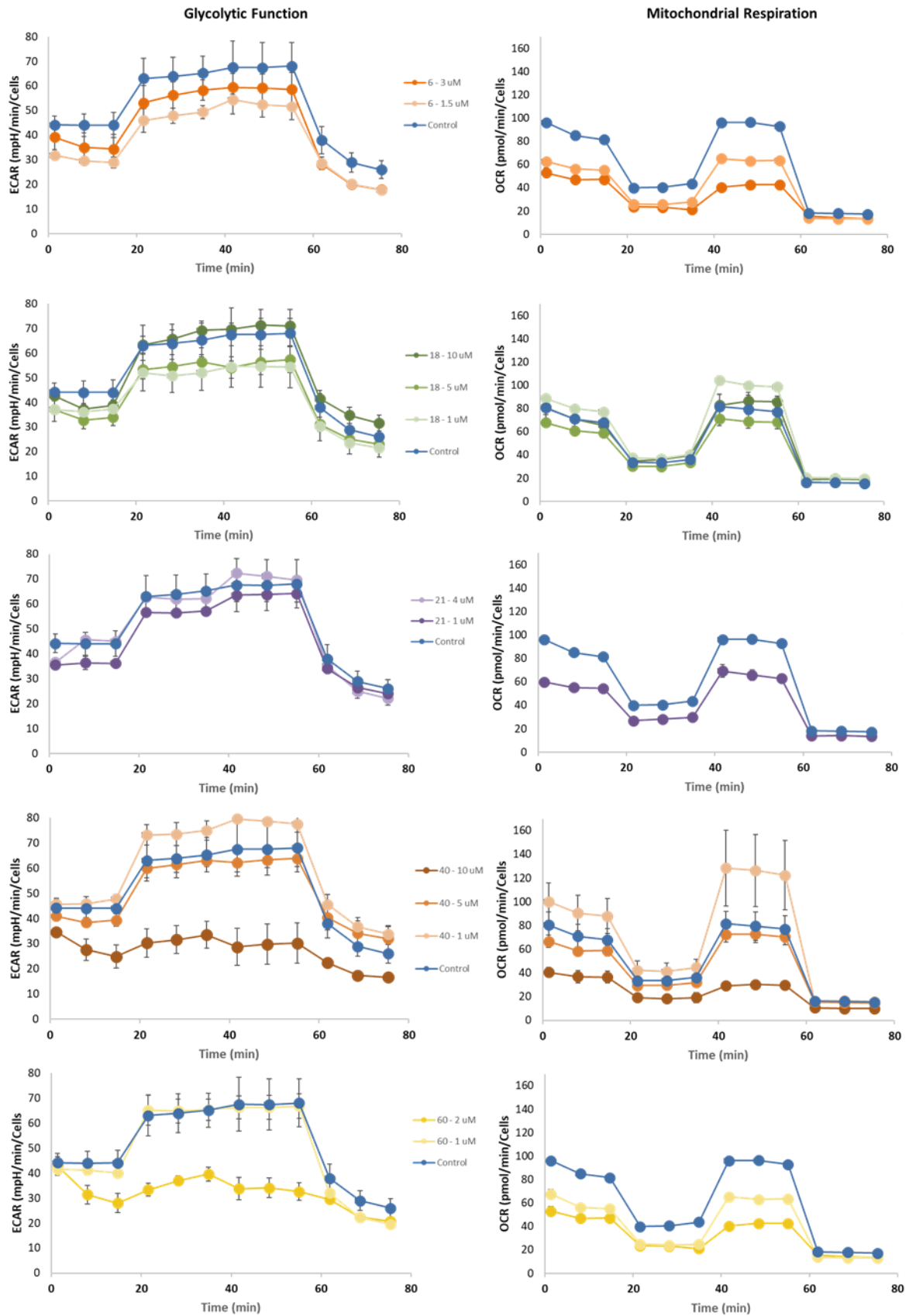


Figure 69. Glycolytic function and mitochondrial respiration profiles of PHK2Is 6, 18, 21, 40, and 60 at different concentrations. Results were obtained from one-time experiment (triplicates or quadruplicates), except for PHK2I 40 ($n=2$, triplicates or quadruplicates). Error bars correspond to SEM.

2.3.2. Effect of an acute treatment – GST

As the conditions for GST were refined to analyze an acute response, 12 PHK2Is and the known inhibitors “cmpd 25” and “cmpd 34” at a fixed concentration (10 μ M) were evaluated in Huh7 and PC3 cells after 2h incubation, in the presence of the test molecules. This set of molecules was the one previously evaluated in PC3 and CHO cells’ proliferation and death. In order to be able to screen a larger number of PHK2Is for glycolysis inhibition, CMST was no longer performed in this phase.

Figure 70 (A and B) shows the results for glycolysis inhibition in Huh7 and PC3 cells. As can be easily concluded, PC3 cells were much more resistant to inhibition by these PHK2Is.

In Huh7 cells, seven out of the 14 molecules reduced glycolysis, including “cmpd 25” and “cmpd 34” (as expected). Here, “cmpd 34” and PHK2I 60 had the highest percentages of glycolysis inhibition (~90%). PHK2I 6, 7 and 10 inhibited by more than 50% and finally “cmpd 25” and PHK2I 2 only affected glycolysis to a small extent. As opposed to the large effect on glycolysis inhibition observed in the chronic treatment experiment, PHK2I 40 was not able to produce an acute response. Future testing of different time-points using this compound might therefore be convenient.

Regarding PC3 cells, only 4 PHK2Is inhibited glycolysis significantly, albeit to a lower extent than in Huh7 cells. Here, PHK2I 10 inhibited by more than 50%, while PHK2Is 6, 7 and 41 only exerted a slight change in ECAR compared with control cells. In theory, PC3 should in fact be less affected by selective HK2 inhibitors, due to the extensive expression of HK1.

GST was equally performed in the four CHO cell lines; however, due to no evidence of consistent glycolysis inhibition in Huh7 and PC3 cells, compounds 1, 18, 21 and 40 were not tested. The other compounds were tested simultaneously in the four CHO cell lines to avoid extra variability and enhance the reliability on the comparison. As discussed above, the CHO EV glycolytic function profile was basically non-existent. Nevertheless, observation of differences in ECAR with treated cells was still possible. However, due to the features required for this assay, and the profile observed previously (**Figure 61**), the significance of the results for this specific cell line might be compromised.

The four CHO cell lines were basically affected by the same compounds, PHK2Is 6, 7, 10, and 60, although to different extents (**Figure 70 – C**). Since there is not much glycolysis without the artificially expressed HKs, lack of selectivity toward HK2 at the test concentration is a reality. Still, comparing the results in CHO HK1, HK2, and GCK cells, PHK2I 6 seems to be the only molecule that preferentially inhibited HK2. As observed with PC3 cells, the known inhibitors, “cmpd 25” and “cmpd 34”, did not affect any of the CHO cell lines significantly.

In order to test the possible selectivity of PHK2Is 6, 7, 10, and 60 when assessed under lower concentrations, CHO HK1, CHO HK2 and CHO GCK cell lines underwent GST with those PHK2Is at 1.5 μ M. The results of a one-time experiment are shown on **Figure 71**. As opposed to the results at higher concentration, PHK2I 6 displayed no selectivity against CHO HK2 cells in this experiment. On the other hand, compound 10 seemed to selectively inhibit CHO HK2 cells, potentially representing an

excellent lead compound for further studies. Unfortunately, due to time constraints, this experiment was not repeated, and its potential selectivity must be confirmed in the future.

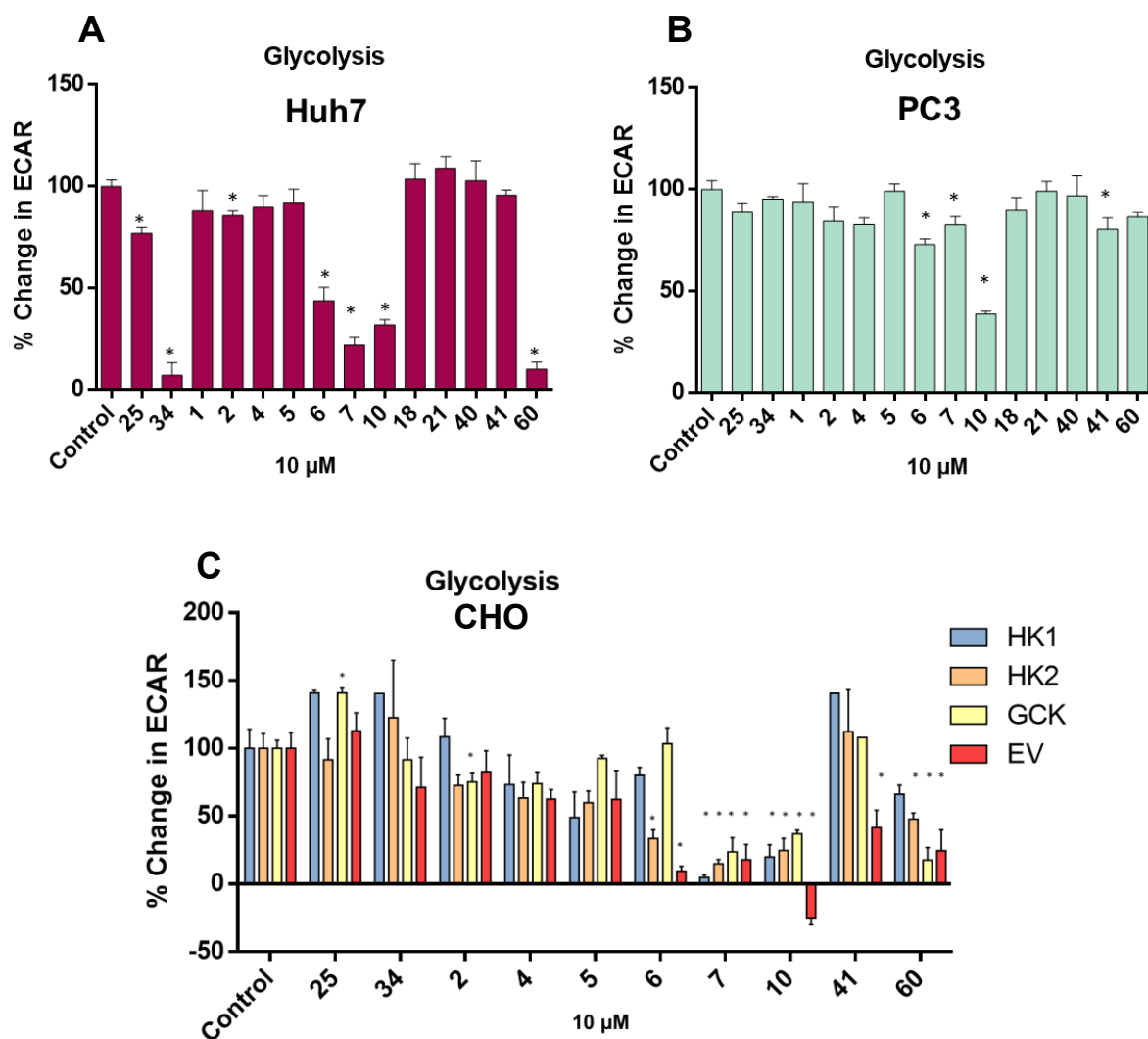


Figure 70. Change in ECAR of Huh7 (A), PC3 (B), and CHO (C) cells after injection of 10 mM glucose, transduced as glycolytic response during GST. Cells underwent prior treatment with known inhibitors and PHK2Is at 10 μ M. The number 25 represents "cmpd 25" and 34 represents "cmpd 34". * $p < 0.05$ (Student's t) regarding control cells (DMSO). Error bars correspond to SEM.

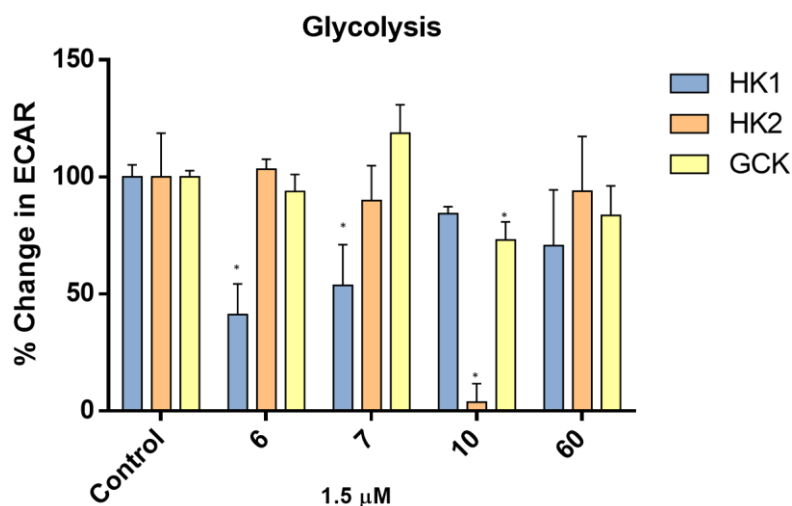


Figure 71. Glycolytic response observed during GST with CHO cells treated with compounds 6, 7, 10 and 60 at 1.5 μ M. Compound 10 inhibits CHO HK2 and not CHO HK1 cells. * - $p < 0.05$ (Student's t) regarding the corresponding control. Error bars correspond to SEM. Results from one-time experiment.

2.3.3. Particular effect of PHK2I 10 on GST

PHK2I 10 was found to decrease the glycolytic activity of all tested cell lines. Interestingly, this compound was also able to decrease dramatically the baseline levels of ECAR, *i.e.*, the non-glycolytic acidification, at 10 μ M (**Figure 72**). Considering this result, other sources of extracellular acidification not related with glycolysis could presumably be reduced in these cells, thereby explaining that no change in ECAR was detected. However, a discrepant difference in pH between treated and control cells was found, which suggests a compound-specific effect. In wells treated with PHK2I 10 the medium pH was around 6 (5.95-6.08), and in control cells it was 7 (6.80-7.02) at the time of the third measurement before the glucose injection. By contrast, the medium of cells treated with other compounds was usually unaffected (6.80-7.1), and no other significant changes were found. To confirm that the medium pH was not changed by the molecule itself, pH was measured before and after addition of PHK2 10 and no change was found after 10 min. The mechanism of acidification prior to GST remained unclear; however, it might result from an effect of compound 10 on yet unidentified targets/pathways. However, when incubated with PHK2I 10 at 1.5 μ M, CHO cells behaved differently, and the acidification was not detected but the anti-glycolytic effect was still found for CHO HK2 cells (**Figure 72**).

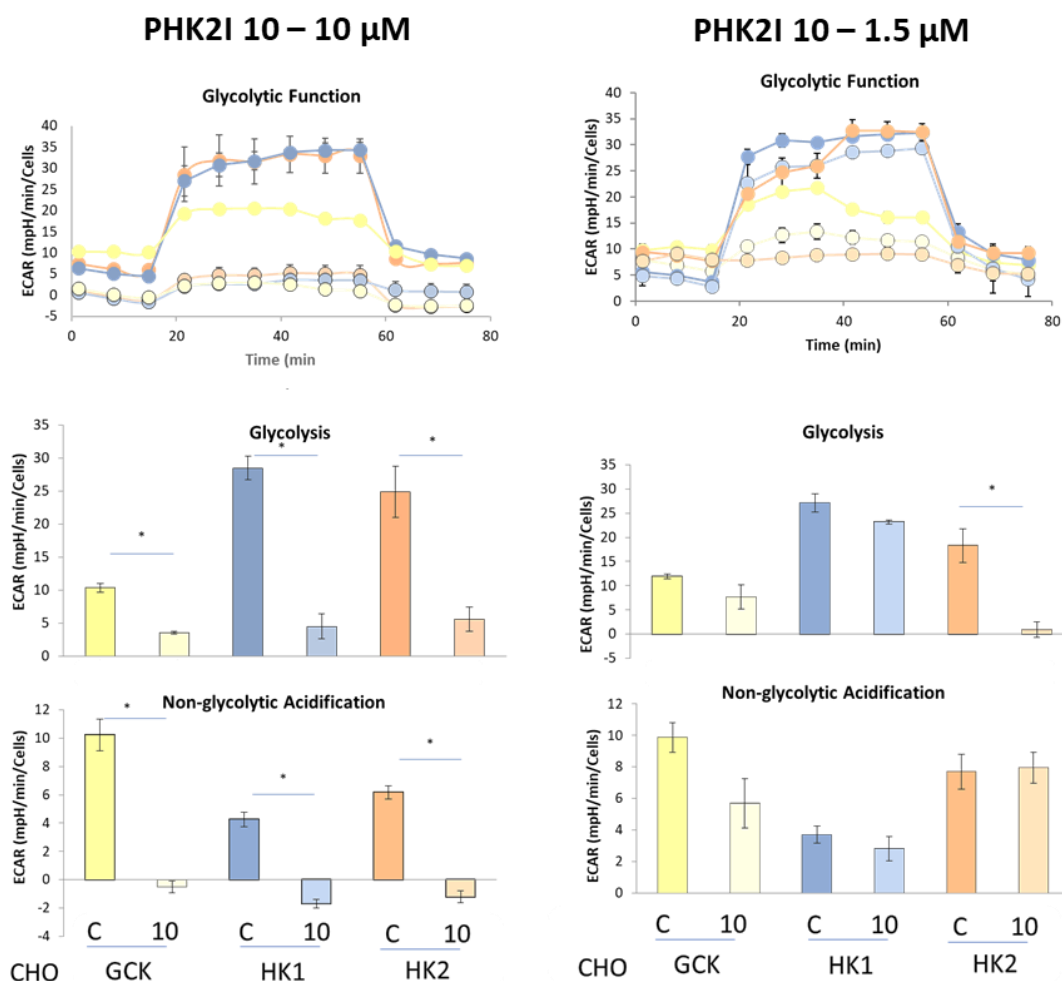


Figure 72. Effect of PHK2I 10 at 10 and 1.5 μM on glycolytic function, glycolysis and non-glycolytic acidification of CHO CGK (yellow), CHO HK1 (blue) and CHO HK2 cells (orange). C corresponds to control cells. Lighter colors (black outline) in glycolytic function graphs correspond to treated cells. * $p < 0.05$ (Student's t) compared to the corresponding control. Error bars correspond to SEM. Results from one-time experiment.

2.4. ^{13}C isotopic labeling experiments – stable isotope-resolved metabolomics

With the aim to investigate if the change in extracellular pH measured during GST was related with the concentration of lactate produced, and to exclude other external factors, labeled $[\text{U-}^{13}\text{C}]$ glucose was used to monitor the evolution of the glycolytic pathway in PHK2I-treated and non-treated cells by the measurement of labeled G6P, pyruvate and lactate variations, assessed by mass spectrometry. The first experiment tested the effect of “cmpd 25” at 13 μM on Huh7 cells, starting with a pre-incubation of 16 hours before the labeling, which lasted for either 4 or 24h (1st protocol – **Figure 73**). In both conditions “cmpd 25” affected the fractional enrichment of pyruvate and lactate (proportion of ^{13}C labeled metabolites in the metabolite pool) significantly reducing labeled metabolites (M+3) compared with control cells (**Figure 74 - A and B**). Under these conditions it was not possible to properly analyze the accumulation of labeled G6P.

To test different conditions, Huh7 cells were pre-incubated with compounds 6, 10 and 60 at 10 μM for two hours before labeling with $[\text{U-}^{13}\text{C}]\text{glucose}$ for 4h (2nd protocol – **Figure 41**). These conditions were chosen because the selected test compounds were shown in the GST measurements to have an effect on glycolysis just after a 2h incubation, which was not found consistently for “cmpd 25”. Reduced labeling of lactate and pyruvate compared with the control was confirmed to be significant for PHK2Is 6, 10 and 60 (**Figure 66 - C and D**); again, a proper analysis of G6P was not feasible. Indeed, the accumulation of G6P was limited and non-labeled G6P was the major species, which was consistent with prior reports [258]. G6P is either directed to glycolysis, PPP, or its accumulation might lead to glycogenesis. On the other hand, glycogen breakdown generates G6P, which contributed to an increase in the relative amount of the non-labeled species. All those biochemical implications are translated into difficulties to analyze and interpret the results for this metabolite.

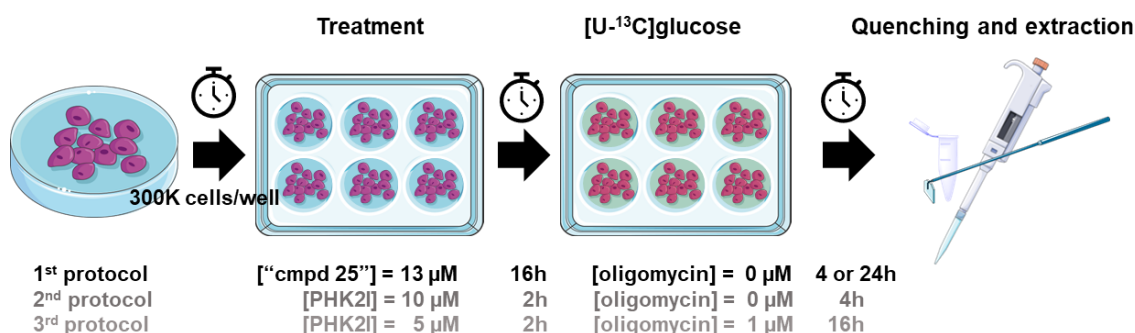


Figure 73. Schematic summary of protocols applied for ^{13}C isotopic labeling experiments.

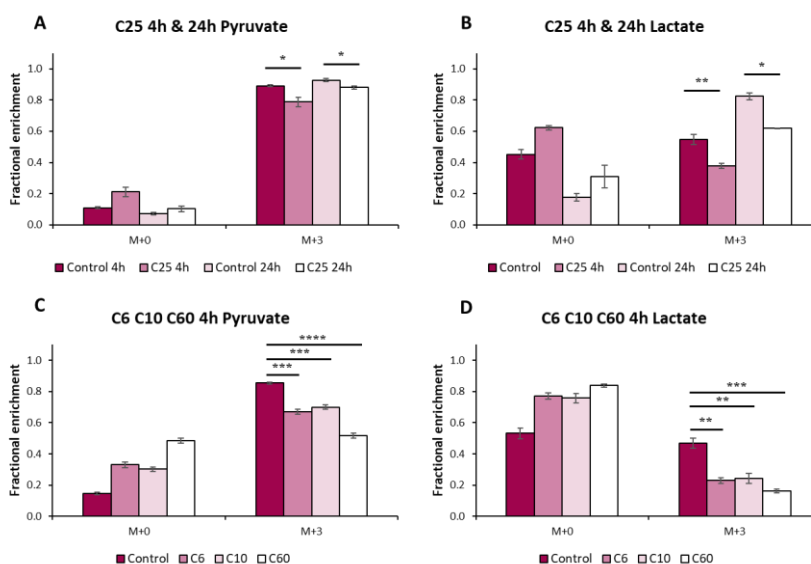


Figure 74. ^{13}C isotopic labeling of the glycolytic pathway in Huh7 cells treated with either “cmpd 25” or PHK2Is, 6, 10 and 60. M+0 represents the non-labeled species and M+3 represents labeled species with 3 (lactate, pyruvate) isotopically labeled carbon atoms. **A** and **B** correspond to fractional enrichment of pyruvate and lactate (respectively) found in the experiments with 4 and 24h labeling incubation times, when “cmpd 25” (13 μM) was tested. **C** and **D** represent fractional enrichment of pyruvate and lactate (respectively) of cells treated with 10 μM 6, 10 or 60, for a 4h labeling period. C25 represents “cmpd 25”; Error bars correspond to SEM; **** $p < 0.0001$; *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; (Student’s *t*).

Oligomycin, as well as other electron transport chain inhibitors, has been shown to increase the ^{13}C enrichment in the G6P pool, after a 12h incubation time with $[\text{U-}^{13}\text{C}]\text{glucose}$ [258]. With the aim to increase the accumulation of G6P, and specifically increase the fractional isotopic enrichment, compounds 6, 10 and 60 were tested in Huh7 cells in the presence of 1 μM oligomycin. As for the previous experiment, these three PHK2Is were pre-incubated with Huh7 cells for 2h prior to incubation with $[\text{U-}^{13}\text{C}]\text{glucose}$; however a longer labeling period, 16h, was applied to reach the isotopic steady state, since the isotopic enrichment (M+3) in control cells was clearly deficient (**Figure 74 - D**). As a longer incubation time and the use of a respiration inhibitor were demanded, a lower concentration of PHK2I (5 μM) was chosen to prevent significant cell death. Whereas a high fractional enrichment of labeled G6P (~ 1) was achieved, the analysis of these data is difficult, as no change between the fractional enrichment among samples was detected, with the exception of a minor reduction trend in cells treated with compound 60. Fractional isotopic enrichment of lactate was also increased from 0.45 to $\sim 0.7\text{--}0.8$ (M+3), while pyruvate fractional enrichment reached 0.8-0.9 (M+3) for control cells (comparison between **Figure 74 - C and D** and **Figure 75 - A, B, C**). Towards a lower concentration of PHK2Is (5 μM) relative to the previous experiment, fractional enrichment of lactate was substantially increased, while pyruvate fractional enrichment was not deeply affected.

Under these conditions, compound 6 did not show to affect pyruvate fractional enrichment (M+3) relative to control cells (M+3) (**Figure 75 - B**). By contrast, compound 10 kept caused an enrichment similar to that observed at 10 μM , while the effect of compound 60 seemed to be consistently reduced. It should be noted, though, that as a consequence of oligomycin action, glycolysis was enhanced, which may have influenced the results obtained.

On GST, compound 6 (10 μM), seemed to preferentially inhibit glycolysis in CHO HK2 over CHO HK1 cells. However, at 1.5 μM , compound 6 inhibited CHO HK1 cells and not CHO HK2. To confirm the potential of 6 to inhibit glycolysis more prominently in one cell line than in the other, CHO HK1 and CHO HK2 cells were also submitted to the abovementioned ^{13}C isotopic labeling procedure. Cells were pre-incubated for 2h with compound 6 (5 μM) and then labeled with $[\text{U-}^{13}\text{C}]\text{glucose}$ for 16h in presence of either compound 6 or DMSO for control cells, and 1 μM oligomycin.

With exception of the G6P pool, isotopic enrichment efficiency was modest, and the isotopic steady state does not seem to have been achieved, as considerable amount of M+0 isotopes of control cells was found. Still, a significant reduction on the fractional enrichment compared to each control was found for pyruvate and lactate. Fractional enrichment of pyruvate and lactate found in CHO HK1 cells treated with compound 6 was 17% and 41% lower than in control cells, respectively. As for CHO HK2 cells treated with compound 6, fractional enrichment was 37% and 60% lower, respectively. A difference that suggests selectivity was found indeed, and apparently compound 6 inhibits CHO HK2 cells more prominently. Nevertheless, the relevance of this modest difference is debatable, considering the preliminary character of these experiments, although the observed trend warrants further investigation (**Figure 75 - D, E, F**).

No difference was found in the accumulation of G6P for any of the cells. Once more, technical issues were at play in this analysis, and assuming that these molecules are responsible for the inhibition of another step-in glycolysis that is not related with HKs is premature. To evaluate the hypothetical inhibition of HK activity more efficiently by this method, the use of 2DG or other non-readily metabolizable HK substrates might be a strategy. Differences in the accumulation of the product 2DG-6-phosphate and/or 6-phospho-2-deoxyglucuronic acid⁸, in treated and non-treated cells, could indicate the extent of HK inhibition; however, to avoid allosteric inhibition by the product, quenching and extraction would have to be performed at an early time point. The accumulation of these products can avoid the limitations related to the very low concentrations of G6P naturally occurring in cytosol and the sensitivity of current techniques for the analysis of sugar phosphates. Dependence on derivatization efficiency is also an issue that must be considered [263,264].

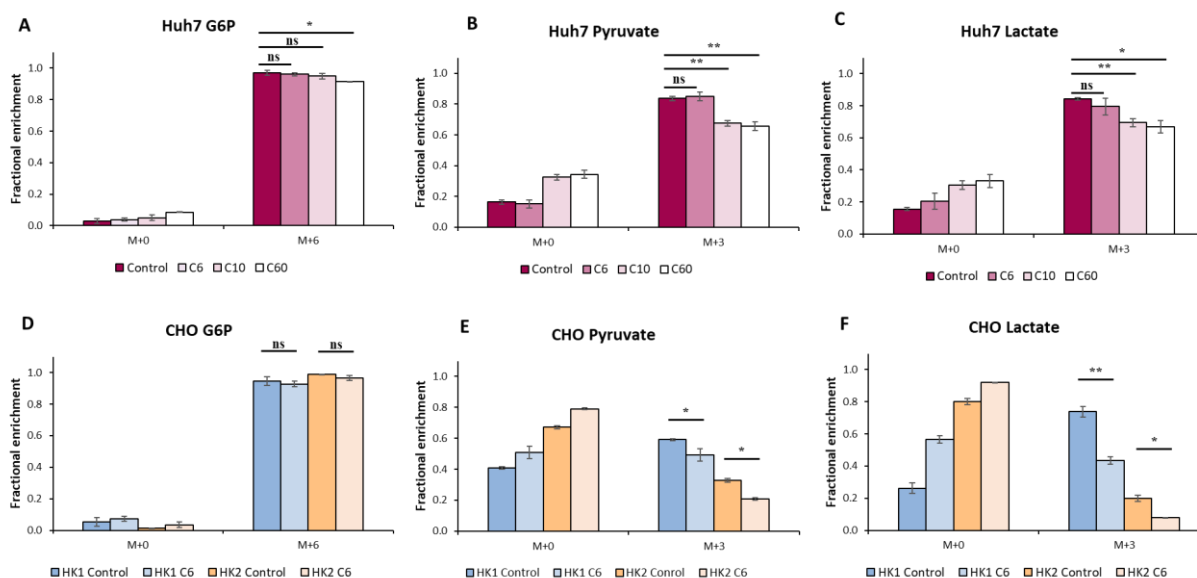


Figure 75. ¹³C isotopic labeling experiments in Huh7 (A, B, C) and CHO (D, E, F) cells. Labeled glucose was introduced in the culture medium and the incubations lasted for 16h before cell harvesting. Both for PHK2I-treated (5 μ M) and control cells, M+0 represents the non-labeled species and M+3 or M+6 represent labeled species with either 3 (lactate, pyruvate) or 6 (G6P) isotopically labeled carbon atoms (ns – nonsignificant; ** $p < 0.01$; * $p < 0.05$; Student's t).

⁸ 6-Phospho-2-deoxyglucuronic acid is the product from the reaction between 2-DG-6-phosphate and NADP⁺, catalyzed by G6PD.

2.5. Biochemical hexokinase activity assay

Our previous results regarding the hexokinase inhibition found with ADP-Glo™ and the inhibition of glycolysis in cells were not mutually consistent. Compounds able to inhibit glycolysis to a large extent did not always show great potential in the biochemical assay (e.g., PHK2I 60). The reasons are yet to be disclosed, but the inhibition of other steps in the glycolytic pathway cannot be excluded. On the other hand, certain molecules that have shown promising results in the biochemical assay were not able to inhibit glycolysis in cells. This might be related with their pharmacokinetic properties and consequent ability to reach the target.

To better comprehend those results, another biochemical assay was carried out to provide additional information. In this experiment, fluorescence was used to detect indirectly the product of reaction catalyzed by HK2 – G6P. **Table 8** shows the ability of a set of PHK2Is to inhibit glycolysis in CHO HK2 and CHO HK1 cell lines, and the results of both biochemical assays as percentage of inhibition, all using a final concentration of 10 µM of each compound.

Apart from for known inhibitor “cmpd 25” whose inhibition of HK2 was pronounced, as expected, none of the compounds matched the inhibition percentage between the two biochemical assays. Moreover, the highest % inhibition among the test PHK2Is observed in the fluorescence assay was 38%, for HK1 inhibition by compound 7. To confirm the adequate performance of this assay, “cmpd 34” was also tested, and it was found to inhibit both isozymes, specially HK2 (87%).

Compounds 1 and 2 confirmed their behavior and did not interfere in HK activity and subsequently glycolysis. Compounds 4 and 5 have shown good ability to inhibit HK2 with ADP-Glo™ test; however, this was not confirmed in the fluorescence-based biochemical assay. Still, these two compounds were found to partially inhibit glycolysis in both CHO HK1 and CHO HK2 cells.

Compound 6, which had previously shown to inhibit glycolysis in CHO HK2 and not much in CHO HK1 cells, was conversely found inhibit HK1 (by 67%) and not HK2 in the fluoresce-based experiment. By contrast, the same compound inhibited HK2 by 73% when using the ADP-Glo™ assay. Interestingly, at 5 µM, stable isotope-resolved metabolomics showed an inhibition of glycolysis in both cell lines, although more prominently in CHO HK2 cells.

Compound 7 inhibited glycolysis to a large extent in most of the tested cell lines. However, while this potential was consistent with the results from the ADP-Glo™ assay, the fluorescence-based assay did not show HK2 inhibition and only detected 31% inhibition of HK1 activity.

Previously, compound 10 showed inhibitory capacity in both GST (all cell lines) and ADP-Glo™ experiments; nonetheless, its ability to inhibit HK2 and HK1 in the fluorescence assay was not as high as expected.

Compounds 18, 21, 40 and 41 have demonstrated a potential to inhibit cell proliferation in all tested cell lines. However, compounds 18, 21 and 40 did not show a potential to inhibit glycolysis in both Huh7 and PC3 cells, while compound 41 did not reveal a capacity to inhibit glycolysis in CHO cells. None of

these compounds was able to inhibit HK1 and HK2 consistently in the resazurin test and only compounds 18 and 40 were shown to inhibit HK2 by about 50% with ADP-Glo™.

Compound 60, with a moderate ability to inhibit HK2 in the ADP-Glo™ assay failed to show the same potential in the fluorescence assay; however, its capacity to inhibit glycolysis in different cell lines was highly pronounced.

In summary, the results did not show consistency between biochemical assays and glycolysis inhibition in cells. As such, the available data are insufficient to postulate that the inhibition of glycolysis by compounds 6, 7, 10 and 60 occurs via inhibition of HK2 activity. More studies are necessary to uncover the real mechanism of action of these compounds.

Table 8. Comparison of glycolysis inhibition in CHO HK1 and CHO HK2 cells and direct rhHK1/HK2 inhibition with biochemical assays by several PHK2Is (10 µM). Results are shown as percentage of inhibition.

PHK2I	CHO HK2 Glycolysis inhibition	CHO HK1 Glycolysis inhibition	HK2 inhibition (Fluorescence)	HK1 inhibition (Fluorescence)	HK2 inhibition (Luminescence)
Cmpd 25	8%	<0%	90%	51%	97%
Cmpd 34	<0%	<0%	87%	17%	ND ¹
1	ND ²	ND ²	<0%	10%	0%
2	27%	<0%	3%	9%	2%
4	37%	27%	4%	17%	52%
5	40%	51%	27%	24%	66%
6	67%	19%	7%	67%	73%
7	85%	95%	<0%	31%	56%
10	75%	80%	23%	38%	60%
18	ND ²	ND ²	11%	7%	51%
21	ND ²	ND ²	3%	10%	23%
40	ND ²	ND ²	6%	19%	50%
41	<0%	<0%	1%	29%	37%
60	52%	34%	<0%	12%	44%

^{1,2} ND – not determined; <0% - enhancement of the HK activity.

3. Conclusions

This chapter contains a compilation of preliminary results. Still, the experiments described herein allowed important insights into the behavior of selected PHK2Is in more complex biological environments. Time constraints limited a full approach with optimization of the conditions of each experiment and more robust results. Still, a starting point was established, and many experiments can now be planned for the near future.

Four PHK2Is (**6, 7, 10, 60** – **Figure 76**) were found to inhibit glycolysis, and potentially HK2, in different cell lines. Those molecules were often more efficient than known inhibitors of HK2. The potential of three of these molecules (PHK2Is 6, 10, 60) was confirmed by two different methods. Moreover, at low micromolar range, these PHK2Is showed evidence of selectively inhibiting glycolysis in cells expressing HK2. Additional metabolomics studies must be performed to assess the accumulation of G6P and confirm whether or not the formation of this metabolite is the step inhibited in glycolysis, rather than another later step in the pathway.

To confirm the hexokinase inhibition in cell extracts, the abovementioned biochemical assays might be performed without rhHK2. The pre-incubation of cells with 2-DG as a surrogate for glucose would enable the accumulation of unmetabolized 2-DG-6-phosphate, favoring an accurate measurement of HK activity in cells at the time of extraction. In this way, the results would not be affected by possible inhibition of downstream glycolytic steps. Since the major source of cytosolic NADPH comes from the reaction catalyzed by G6PD [265], remaining NADPH sources would not be expected to significantly affect the outcome of this experiment.

The interference of PHK2Is in other targets/pathways cannot be excluded and should be characterized. The potential selectivity of any of these compounds has yet to be confirmed, and this should be ascertained soon, since the biochemical assays with recombinant enzymes did not produce reliable results.

Despite the inability of PHK2Is 21, 40 and 41 to inhibit glycolysis significantly, their effect on cell proliferation and/or change in the glycolytic and respiratory functions over time could be meaningful, and further characterization could be fruitful if other roles in cell biology were to be found.

Although a great potential was shown against a urothelial carcinoma cell line (UM-UC-3) [43], known HK2 inhibitors “cmpd 25” and “cmpd 34” fell far short of the expectations in the cell lines tested herein. Glycolysis was only inhibited prominently in the Huh7 cell line by “cmpd 34”, and proliferation was not greatly affected. “Cmpd 25” showed higher potential to inhibit cell proliferation, but the anti-glycolytic effect was basically negligible. Specific transporters and/or metabolic machinery in Huh7 and UM-UC-3 cell lines that could explain the disparity of the results should be investigated.

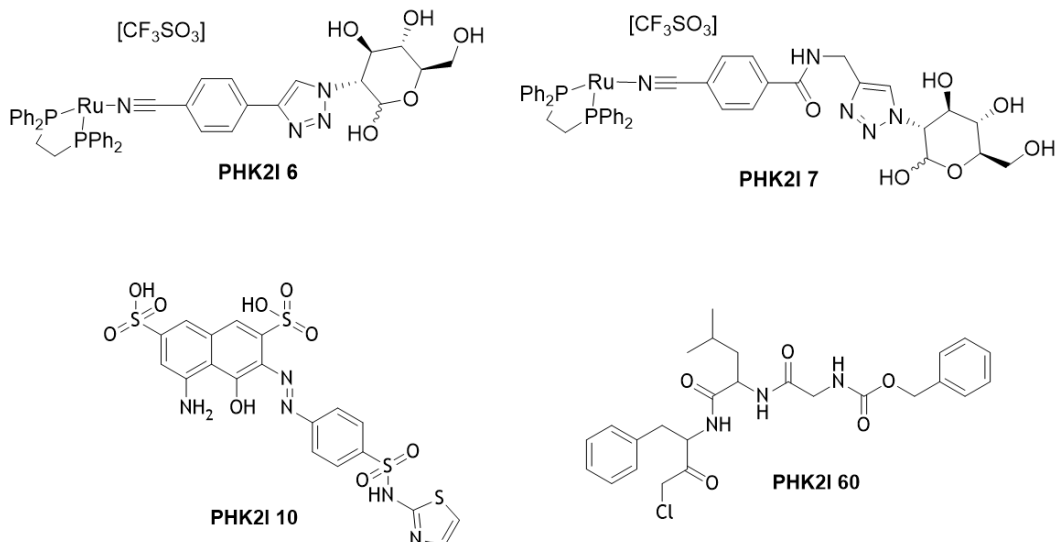


Figure 76. Structures of promising PHK2Is 6, 7, 10 and 60.

4. Experimental Section

4.1. Preparation of PHK2I solutions

PHK2Is were all dissolved in dimethyl sulfoxide (DMSO) to a 10 mM stock concentration, aliquoted and stored at -80 °C. The various PHK2I solutions in different concentrations were prepared by adequate dilutions in the complete culture medium before each experiment. The DMSO concentration in the studies was ≤1%. All experiments were compared with the same volume of vehicle (DMSO) as the concentrated PHK2I solution.

4.2. Biochemical hexokinase activity assay

RhHK activity inhibition was measured using a G6PDH-coupled enzymatic fluorometric assay. Briefly, 10 µL of HK dilution buffer (20 mM KH₂PO₄, 100 mM KCl, 1mM MgCl₂, 1mM EDTA, 1.4 mM 2-mercaptoethanol (2-ME), 60 g/L glycerol, 1g/L BSA) containing rhHK1 or rhHK2 and DMSO or PHK2I were plated in black 96-well microplates (clear bottom) and incubated for 30 min. Then 90 µL of reaction cocktail (100 mM triethanolamine hydrochloride pH 7.6; 22 mM monothioglycerol, 15 mM MgCl₂, 1 mM EDTA, 1.4 mM 2-ME) containing 1 mM glucose, 6.6 mM ATP, 0.5 mM NADP⁺, 4.5 U/mL G6PD, 0.2 U/mL diaphorase and 10 µM resazurin were added. Catalysed by G6PD, G6P formed is oxidized with conversion of NADP⁺ in NADPH. NADPH is then used by the resazurin-diaphorase system to generate the highly fluorescent resorufin. Fluorescence emitted at 590 nm was recorded every 5 min until 20 min reaction, using excitation at 530 nm. Background fluorescence was corrected subtracting the value of

the blank (no HK) for each sample. None of the compounds was found to interfere with the assay. Fluorescence was measured using a Cytation™ 1 Cell Imaging Multi-Mode Reader (BioTek, Agilent Technologies, Santa Clara, CA, USA).

4.3. Cell Culture

The PC3 and Huh7 cell lines were purchased from ATCC (Manassas, VA). MI5-4 CHO cells were kindly provided by Dr. John Wilson (Department of Biochemistry and Molecular Biology, Michigan State University, MI, USA). Huh7 cells were grown in Dulbecco's Modified Essential Media (DMEM, Corning), PC3 in Ham's F-12K Nutrient Mixture Kaighn's Modified (F12K, Corning), and CHO cells in Minimum Essential Media Alpha (α-MEM, Corning) supplemented with 2 mM glutamine, all in 10% fetal bovine serum (FBS, Atlanta Biologicals) and 1× Pen/Strep (Fisher), at 37°C in a 5% CO₂ atmosphere. Glucose concentrations were kept at 25 mM.

4.4. CHO cell lines generation⁹

The pLenti-6-D-TOPO-Blast lentiviral vector (Invitrogen) was used for the overexpression of hexokinases in MI5-4 CHO cells: HA-tagged rat HK2, V5-tagged Human HK1 (pancreas variant) and HA-tagged human GCK (liver variant). An empty vector was used for CHO EV cell line [260].

4.5. Effect on cell proliferation

Cells (3.5 x 10³/well) were seeded on 24 well plates in triplicate and treated with vehicle (DMSO), "cmpd 25", "cmpd 34" or PHK2I after 24h. Huh7 cells were counted on day 0, 1, 2, 3, 5, and 7, PC3 and CHO was counted on day 0, 3 and 6. Media were changed each 72h to ensure continuous natural growth. Cells were counted using the bright field on Celigo Adherent Cell Cytometer (Nexcelom Bioscience LLC, Lawrence, MA, USA).

For IC₅₀ calculations, cells were treated with 7 different concentrations of each compound. The number of cells was recorded on day 0, 3 and 6 of the treatment. Calculations were performed using GraphPad Prism® 6.0 (GraphPad Software Inc., La Jolla, CA, USA) through the following equation:

$$Y = \text{Bottom} + (\text{Top} - \text{Bottom}) / (1 + 10^{((\text{LogIC}_{50} - X) * \text{HillSlope})})$$

⁹ Cell lines generation was not performed by the author.

4.6. Effect on cell death

Cells (3×10^4 /well) were seeded on 48 well plates in triplicate and 24h later they were treated with vehicle (DMSO), "cmpd 25", "cmpd 34" or selected PHK2Is. The number of cells was recorded on day 0 and 2. After 48h of treatment, cells were incubated with Hoescht 33342 dye (for all cells) and propidium iodide (for dead cells) for 30 min. Celigo Adherent Cell Cytometer (Nexcelom Bioscience LLC, Lawrence, MA, USA), was used to acquire and analyze images. Markers were identified in each fluorescent channel.

4.7. Immunoblotting

After harvesting with PBS, on ice, cells were centrifuged at 4000 rpm for 5min. Pellets were frozen overnight at -80°C and then lysed with 5 freeze/thaw cycles on dry ice, with lysis buffer: 20 mM HEPES, 150 mM NaCl, 1% Triton X-100, 1 mM EDTA, 1mM EGTA, 10 mM sodium pyrophosphate, 100 mM NaF, 5 mM iodoacetic acid, 20 nM okadaic acid, 0.2 mM phenylmethylsulfonyl flouride (PMSF), complete protease inhibitor cocktail and phosphatase inhibitor cocktail tablets (Roche Diagnostics). After centrifugation (8000 rpm, 5 min) the supernatant was collected and used for the determination of protein concentration on each sample with the Bradford method.

In this method, 4 μL of 1:10 diluted samples are plated on a 96-well clear plate in duplicates, as well as a set of BSA samples at 0, 0.2, 0.4, 0.8 and 1.2 $\mu\text{g}/\mu\text{L}$. After incubation for 15 min at r.t. with 1 X Bio-Rad Protein Assay Dye Reagent, fluorescence was read at 581 nm with Cytation™ 1 Cell Imaging Multi-Mode Reader (BioTek, Agilent Technologies, Santa Clara, CA, USA).

Then, 20 μg protein suspended in SDS loading buffer, was run on 12-25% SDS polyacrylamide gels and electrotransferred to PVDF membranes. After blocking with 4% milk in TBST, membranes were exposed to the primary antibody on 2% milk/TBST solutions overnight at 4°C .

Antibodies used included: anti-HK2 (Cell Signaling Technology, 2867- dilution: 1:2000), anti-HK1 (Cell Signaling Technology, 2024), anti-GCK, 1:500 (Santa Cruz, 7908), anti-HA (Convance, MMS101R), anti- β -actin, 1:2 K (Sigma, A5441), anti-V5 Tag (ThermoFisher, R960-25).

Membranes were incubated with appropriate HRP secondary antibodies (1:3000 dilution: anti-mouse, anti-rabbit, 1:2000 dilution: anti-goat) for 1h30min at r.t., and the Pierce ECL Western Blotting Substrate (Thermo Scientific) was used for development of immunoreactive bands, with Azure Imaging System (Azure Biosystems Inc., Dublin, CA, USA).

4.8. Metabolomic assays

4.8.1. Measurement of the oxygen consumption rate (OCR) for CMST and the extracellular acidification rate (ECAR) for GST

OCR and ECAR were measured using the XF96e Extracellular Flux analyzer (Agilent Technologies, Santa Clara, CA, USA). Cells were plated at 1.5×10^4 cells/well on XF96 cell culture plates (Agilent Technologies) and incubated in a humidified 37°C incubator with 5% CO₂ in normal culture medium.

For chronic treatment, cells were previously treated with PHK2I or DMSO for 7 days and then plated. After ~40 h, and 1 hour prior to the assay, the growth medium in the wells was replaced by XF assay medium (XF base medium lacking bicarbonate and HEPES containing 10 mM glucose, 1 mM sodium pyruvate, and 2 mM glutamine for OCR measurements, and 2 mM glutamine only for ECAR measurements), and the plate was transferred to a 37°C CO₂-free incubator.

To evaluate the acute response to PHK2Is, cells were plated and after 16-18h the growth medium in the wells was replaced by XF assay medium as described above, containing either vehicle (DMSO) or PHK2I and the plate was incubated for 2h at 37°C in a CO₂-free incubator.

For OCR measurement, successive injection of compounds measured ATP-coupled respiration (1 µM oligomycin), maximal respiration (1 µM FCCP), and non-mitochondrial respiration (0.5 µM rotenone/antimycin A). Basal respiration, proton leakage, and spare respiratory capacity were then calculated according to the change in the concentration of O₂ prior/after the injections (**Figure 55 - B, Table 9**).

As for ECAR measurement, successive injection of compounds measured glycolysis (10 mM glucose), glycolytic capacity (1 µM oligomycin), and non-glycolytic acidification (50 mM 2DG). The glycolytic reserve was then calculated according to the change in the pH prior/after the injections (**Figure 48 - A, Table 9**).

In this experiment, three baseline measurements were taken prior to the addition of any compound, and three response measurements (every 3 min) were taken after the addition of each compound. The OCR and ECAR are reported as being normalized against cell counts (pmoles/min/10⁴ cells for OCR and mpH/min/10⁴ cells for ECAR). The baseline OCR or ECAR refers to the starting rates prior to the addition of a compound. Unless referred otherwise in figure captions, each experiment was performed at least twice in quadruplicate.

Table 9. Equations used for the calculation of metabolic parameters from OCR and ECAR measurements during CMST and GST, respectively.

Measurement	Metabolic parameters	Equation
OCR	Non-mitochondrial oxygen Consumption	= Minimum rate measurement after Rotenone/antimycin A injection
	Basal respiration	= (Last rate measurement before first injection) – (Non-mitochondrial respiration)
	Maximal Respiration	= (Maximum rate measurement after FCCP injection) – Non-mitochondrial respiration
	Proton leakage	= (Minimum rate measurement after oligomycin injection) – (Non-mitochondrial respiration)
	ATP production	= (Last rate measurement before oligomycin injection) – (Minimum rate measurement after oligomycin injection)
	Spare respiratory capacity	= Maximal respiration – Basal respiration
ECAR	Glycolysis	= (Maximum rate measurement before oligomycin injection) – (Last rate measurement before glucose injection)
	Glycolytic Capacity	= (Maximum rate measurement after oligomycin injection) – (Last rate measurement before = glucose injection)
	Non-glycolytic acidification	= Last rate measurement prior to glucose injection
	Glycolytic reserve	= Glycolytic Capacity – Glycolysis

4.8.2. ¹³C isotopic labeling experiments¹⁰

For labeling experiments, 3×10^5 cells were plated onto 6-cm plates and incubated with normal complete culture medium. On the next day cells were treated with either vehicle (DMSO), or “cmpd 25”, 6, 10, 60 in triplicates. After incubation at 37 °C (2h or 16h) medium was removed, cells were washed with 2 X PBS and new medium (glucose free, dialyzed FBS) containing the vehicle or PHK2I and 25 mM [U-¹³C]glucose (Cambridge Isotope Laboratories) was added for labelling. After 4, 16 or 24h the medium was removed and cells were washed with 2 X 0.9% NaCl (ice cold). Then, 300 uL of 5:3:2 MeOH/ACN/DW (HPLC grade) (MAD) containing the internal standard Norvaline (2mg/ mL) were added, and each well was scrapped. The cell suspension was transferred to a 1.5 mL tube and 300 uL of MAD were additionally used to wash the well and collected into the same tube. Samples were kept at -80°C until GC/MS analysis. The dried intracellular metabolites were dissolved in 50 µL of 2 wt% methoxylamine hydrochloride in pyridine and incubated at 37 °C for 90 min on a heating block. Next, 80 µL of *N*-methyl-*N*-(tert-butyldimethylsilyl)-trifluoroacetamide (MTBSTFA) + 1% tert-butyldimethylchlorosilane (TBDMCS) (Thermo Scientific, Bellefonte, PA) were added and the samples

¹⁰ After cells harvesting, samples were treated and run by Soeun Kang (Department of Biochemistry and Molecular Genetics, College of Medicine, University of Illinois at Chicago, IL, USA), who also analyzed the results.

were incubated for 30 min at 60°C. After an overnight incubation at room temperature, the derivatized samples were briefly centrifuged and the clear liquid was transferred into GC vials for GC/MS analysis. GC/MS analysis was performed on an Agilent 7890B GC system equipped with a DB-5MS capillary column (30 m, 0.25 mm i.d., 0.25- μ m phase thickness; Agilent J&W Scientific), connected to an Agilent 5977 A mass spectrometer operating under ionization by electron impact (EI) at 70 eV. Mass isotopomer distributions were obtained by integration and corrected for natural isotope abundances.

4.9. Statistical analysis.

Raw experimental data was processed using GraphPad Prism® 6.0 (GraphPad Software Inc., La Jolla, CA, USA) or Microsoft Excel 2019 (v16). Results are expressed as the mean \pm standard deviation (SD) or standard error of the mean (SEM), as indicated in the figure caption. All experiments were repeated at least twice, at least in triplicates, unless otherwise indicated. To determine statistical significance ($p < 0.05$), *Student's t* test was performed using the abovementioned software.

Chapter VI

*Structural and dynamic behavior
comparison between HK1 and HK2: a
Molecular Dynamics approach to explore
potential drug selectivity*

1. Overview

One of the biggest obstacles to targeting HK2 is its high similarity to HK1, which prevents selective inhibitors from been obtained. Without a selective inhibition of HK2, side effects might preclude its use for cancer treatment. The present chapter describes a detailed comparison between the structures of HK1 and HK2 (**Figure 77**). First, the static crystallographic structures were carefully analyzed, with and without ligands, and the differences were annotated. Unfortunately, together with literature data, the differences could not properly explain the behavior of each isozyme towards selective inhibitors. In this regard, molecular dynamics (MD) studies were performed, simulating both structures in solution at physiological pH. This allows for a natural evolution of structure over time, either in the presence or absence of several small molecules such as endogenous ligands and inhibitors. If the catalytic pocket of the static crystallographic structure cannot be a source of knowledge for understanding selectivity of some inhibitors for an isozyme, the key might be related with different motifs away from the catalytic pocket that might profoundly influence the resulting structure. It should be noted that the HK2 structure was resolved using the HK1 structure a model [73], and that could influence the resulting structure. MD simulations can overcome this issue and predict a more likely three-dimensional conformation. In the end, the differences between the X-ray determined structure and the MD-optimized structure can be used for future selectivity studies.

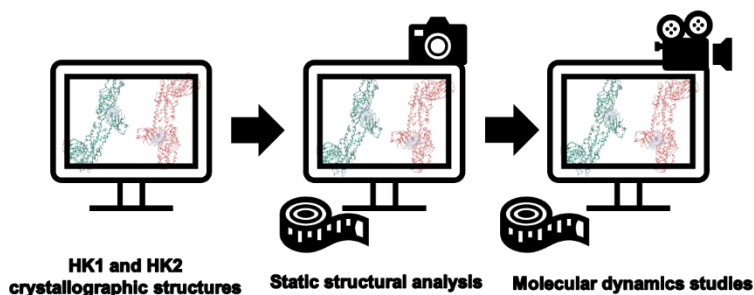


Figure 77. Studies presented in this chapter, the static structural analysis of HK1 and HK2 and MD studies.

1.1. Static structural analysis

The analysis of the static structure of a protein can provide a considerable amount of information. The distances between residues/atoms can be measured and compared with other structures that have been crystallized in distinct conditions when different polypeptides are being compared, different residues might be found to be placed in strategic places, and it allows, for example, the study of the effect of mutations on structure.

1.2. *Molecular dynamics*

A single rigid structure is unlikely to adequately predict the binding modes of different compounds, as it is nothing more than a photograph taken under very specific conditions that are usually not comparable to the physiological ones. In biochemistry, MD simulations aim to observe conformational changes of relevant biological molecules in solution, mostly involving at least one macromolecule such as a protein.

Classic MD studies follow Newton's laws of motion to calculate the trajectory of protein conformations as a function of time. Calculations are performed for each particle (atom) to determine the position and velocity of each one at a later time. A force field is applied to mathematically calculate the forces between atoms and molecules. The use of MD simulations has led, for instance, to the development of antivirals [266,267].

Over the years, simulations have become more accurate with improved forcefields and ensemble techniques, leading to a more realistic description of a system. An MD user is able to choose a set of initial molecular configurations, run the simulation, and make observations from the trajectory. Throughout the trajectory, multiple conformations are likely to form; differences in these conformations, obtained with different simulation conditions, can drive to important conclusions. While animations of a trajectory can give important information about the motion of the molecule, a less subjective analysis, using precise measurements and calculations (distances, angles, areas, volumes, etc.) is preferred. As more data are obtained, the predictions will be more accurate [268,269].

In this project, MD simulations were performed using GROMACS software (v2016) [270]. GROMACS (GROninge Machine for Chemical Simulation) is one of the most widely used free and open-source software packages for MD simulations of biomolecules, and its use is reported in thousands of publications every year. GROMACS provides versatile tools to perform different types of calculations, systems preparation and analysis and is being coupled with other tools and force fields such as plugins for PyMOL software or MARTINI forcefield [268–270].

2. Results and Discussion

2.1. Important structure insights from the literature

In chapter II (section 2), Table 1 shows the identity and similarity between HK1 and HK2 to be the highest among HKs. Because the highest levels of similarity between HKs from humans and other species are related to extremely conserved glucose and G6P binding pockets, human HK1 and HK2 share the exact residues in binding sites (Chapter II, Section 2.2. **Figure 6**). However, as mentioned above, only HK2 is catalytically active in both *N*- and *C*-terminal binding pockets, unlike HK1 and HK3, whose *N*-catalytic pocket plays exclusively regulatory functions [54,79,80]. This ability to catalyze the

conversion of glucose to G6P in both catalytic pockets is one of the features that make HK2 unique and a characteristic that can be targetable.

The reasons for this discrepancy appear to be related with specific mutations near the catalytic pockets, which affect the chemical environment (charge, size, and hydrophobicity). The results of an early study of site-directed mutagenesis suggested that the G174R and D447S double mutations were primarily responsible for reducing the catalytic activity of HK1 NTD, whereas a single mutation (G231D) appeared to be sufficient for a 95% inactivation of the *N*-terminal moiety in HK3 [81]; phylogenetic analysis is consistent with this interpretation [53]. While these studies provided important insights, the availability of solved crystallographic structures in recent years has become an increasingly valuable tool for exploring how conformational differences between isozymes may impact catalytic activity. In this regard, the crystallographic structure of HK2, combined with computational molecular dynamics data, detected a difference in the linking region between the NTD and CTD of HK1 and HK2, confirming that the catalytic activity of the NTD of HK2 is regulated by the linker, helix- α_{13} . In HK1, helix- α_{13} undergoes a deviation of 16°, resulting in a bent, and stiff structure that is associated with the loss of catalytic activity [73].

Connected with this last finding, a new regulatory region of NTD activity was discovered [271]. In the NTD catalytic pocket region three important substructures interact closely: the helix- α_5 , where D209 is placed (an essential residue for catalytic activity); the loop⁴⁴⁴⁻⁴⁴⁷; and helix- α_{13} , the linker that binds NTD and CTD. In that study, several single site-directed mutations were carried out in these regions and HK activity was measured. Interestingly, only two residues differ between HK1 and HK2, R444 and D447 in HK2 and L444 and S447 in HK1. Despite the conflicting properties of these residues, HK2 single mutants R444L and D447S still exhibited at least 50% of catalytic activity. These findings suggest that a single residue is not likely to be responsible for the inhibition of HK1 NTD catalytic activity, and site-directed double mutations with these residues would be comprehensive strategy to pursue. Yet, the mutants D447A, K451A and S449A and S449T completely abolish the HK2 NTD activity, attesting to the importance of these residues for the activity. These residues were confirmed to establish important interactions with residues in the helix- α_5 and internally in the helix- α_{13} itself, contributing to a more stable three-dimensional structure around the catalytic pocket. This is postulated to improve the binding stability of glucose for subsequent alignment with ATP, enabling the catalytic reaction. Thus, the region found in the beginning of helix- α_{13} has been proposed as a regulatory site that could be targeted for specific HK2 inhibition by disrupting these interactions, preventing the stabilization necessary for glucose binding [271] (**Figure 78**).

To the best of our knowledge, no other structural evidence has been proved to differentiate HK1 from HK2.

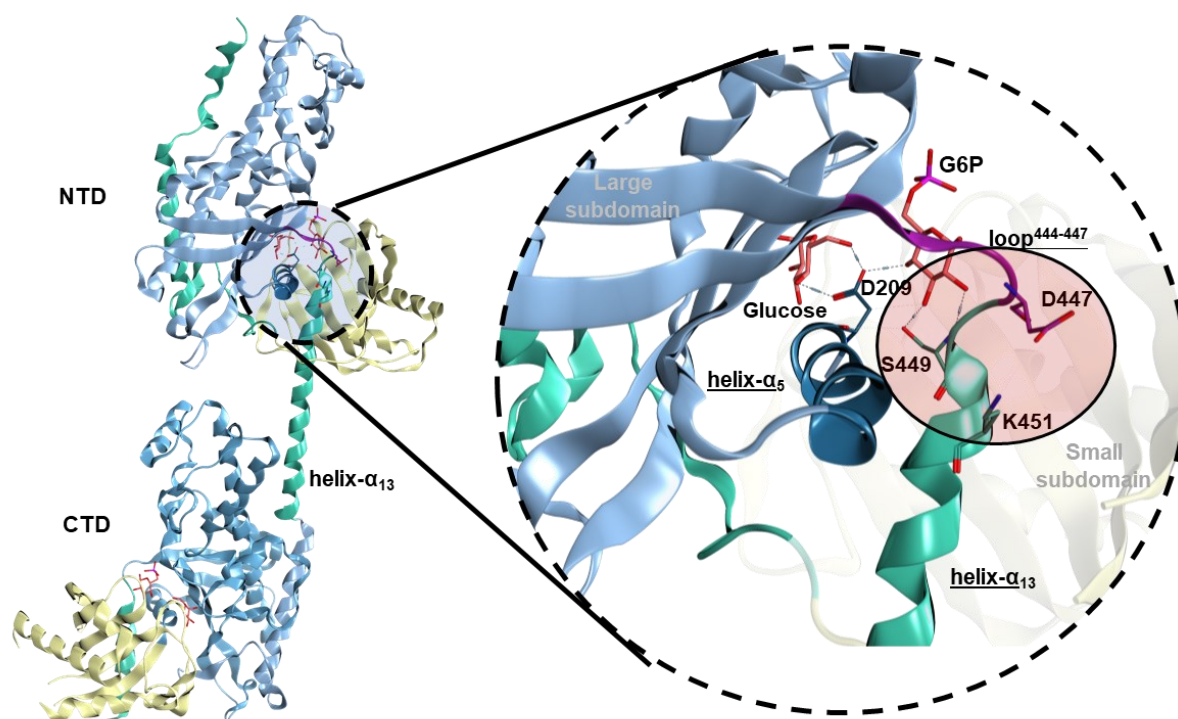


Figure 78. Novel regulatory region of NTD activity of HK2, that can be used as target for specific HK2 inhibition (red), by the destabilization of residues Asp447, Ser449 and Lys451. The three-dimensional structure in the figure corresponds to the PDB entry 2NZT [73].

2.2. Static structural analysis

Static structural analysis was performed using resolved crystallographic structures as valuable tools to explore conformational and identity differences among human HK1 and HK2 isozymes and elucidate their possible impact on the selectivity of certain inhibitors. Structure analysis, comparison studies and distance measurements were performed with Molecular Operating Environment 2016 (MOE2016) software. Global structures, catalytic pockets (or binding pocket for the *N*-terminal pocket of HK1) were compared, as well as “neighboring residues” (less than 4.5 Å apart from catalytic pocket residues).

To date, six wild-type hHK1 and four hHK2 structures were resolved and are available on PDB [78] (**Table 10**). None of available HK1 and HK2 structures contain the total sequence of residues (914), yet they are considered representative of the complete structures except for the initial α -helix (residues 1-12). This hydrophobic moiety is known to be unstable, denaturing rapidly [54]. Structures were obtained at a reasonable resolution ($2.25 \leq R \leq 3 \text{ \AA}$) between the years 1998 and 2016. Except for 1HKC [84] and 5HFU [43], each entry is composed by two chains, establishing a dimer.

Alignment and superposition of HK1 structures revealed very similar conformations (average RMSD = 2.58 Å, RMSD ranging from 0.31 to 4.55 Å). Alignment of the catalytic pocket (27 residues) and catalytic pocket plus neighboring residues (91 residues) showed similar and minimal differences (RMSD from

VI - Structural comparison between HK1 and HK2: a Molecular Dynamics approach to investigate potential drug selectivity

0.13 to 0.34) among structures with glucose (1HKC is not included), in the closed state. However, without glucose and inorganic phosphate present in the CTD, structure 1HKC remains in the open state, presenting an RMSD of the catalytic pocket above 2.50 Å when compared with the other five structures. Interestingly, this structure exhibits a glucose molecule bound to the regulatory NTD, changed to the closed state. In this domain, the *N*-terminal pocket (27 residues) and the pocket plus neighboring residues (96 residues) display similar RMSDs ranging only from 0.16 to 0.62 Å, with all structures in the closed state.

Table 10. Data on three dimensional structures of wild type hHK1 and hHK2 available on PDB.

Human isozyme	Structure (PDBID)	Chains	Resolution (Å)	Year	Residues	Ligands
HK1	1HKB	2	2.80	1998	16-914 (899/chain)	β -D-Glucose; β -D-G6P
	1HKC	1	2.80	1998	16-914 (899)	β -D-Glucose; inorganic phosphate
	1QHA	2	2.25	1999	12-914 (903/chain)	α -D-Glucose; AMP-PNP; α -D-G6P
	4F9O	2	2.65	2013	16-914 (899/chain)	β -D-Glucose; β -D-2-deoxy-G6P
	4FOE	2	2.70	2013	16-914 (899/chain)	β -D-Glucose; β -D-Mannose-6-phosphate
	4FPB	2	3.00	2013	16-914 (899/chain)	β -D-Glucose; 1,5-Anhydroglucitol-6-phosphate
HK2	2NZT	2	2.45	2006	17-913 (871/chain)	α -D-Glucose; β -D-G6P
	5HEX	2	2.73	2016	17-911 (880/chain)	"cmpd 30"
	5HFU	2	2.92	2016	17-911 (895/chain)	"cmpd 27"
	5HG1	1	2.76	2016	17-913 (867)	"cmpd 1", β -D-G6P

PDBID references: 1HKB, 1HKC [84]; 4F9O, 4FOE, 4FPB [72]; 2NZT [73]. 5HEX, 5HFU, 5HG1 [43].

The available HK2 structures, when aligned and superposed, present a more pronounced difference among them (average RMSD = 5.15 Å). 5HEX [43] and 5HFU [43] structures closely resemble each other (RMSD= 0.44 Å), wherein both interact with synthetic inhibitors, showing an open state of the structure, when no glucose is bound. Conversely, 2NZT [73] shows the most closed (tight) state, interacting with native ligands, glucose and G6P. Structure 5HG1 [43] presents an intermediary state, more closed than 5HEX and 5HFU (RMSD= \sim 4.40 Å) and less tight than 2NZT (RMSD=5.28). The structure 5HG1 is associated with a weak and non-selective inhibitor [43]. Comparing only catalytic pockets (27 residues), the structural resemblance of the three-dimensional poses is higher, with an average RMSD of 2.32 Å and 2.54 Å, respectively, for NTD and CTD pockets. When comparing 5HEX and 5HFU, RMSD values do not go beyond 1.20 Å for the NTD pocket and 0.69 Å for the CTD pocket. Due to the presence of both the synthetic inhibitor and G6P, the *N*-terminal pocket of 5HG1 is more closely related with the remaining open state HK2 structures (RMSD= 1.60 Å for 5HFU, RMSD= 2.04 Å for 5HEX, RMSD= 2.10 for 2NZT). On the other hand, the CTD pocket, when bound to the weak

synthetic inhibitor, is more closely related with 2NZT (RMSD= 2.80 Å for 5HFU and 5HEX, RMSD= 1.08 for 2NZT).

When comparing the glucose cavities of both *N*- and *C*-domains of the HK2 structure (PDB:2NZT [73]), it is possible to understand that they are identical, wherein glucose and G6P interact in the same way with approximately the same residues in both domains [Asn208/656, Asn235/683, Asp209/657 (with glucose and G6P), Glu260/708, Glu294/742, Lys173/621, Thr172/620 and Asp84/532, Asp413/861, Gly414/862, Ser449/897, Thr88/536, Thr232/680] (**Figure 79**). The equivalent happens with pockets in HK1 (not shown).

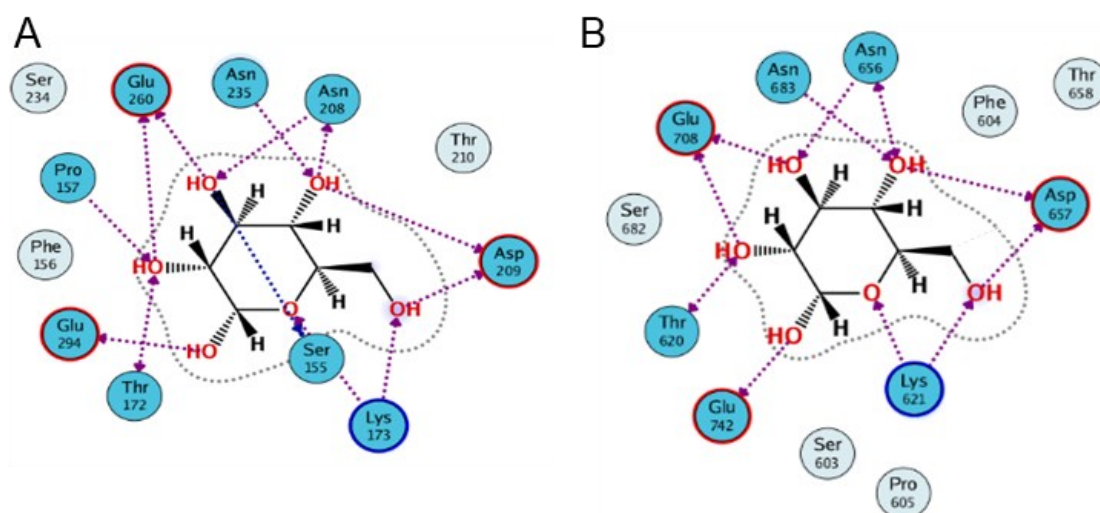


Figure 79. Glucose interactions (hydrogen bonds) with residues of HK2 NTD (A) and CTD (B) catalytic pockets.

Aiming to compare the binding pockets' structures of HK1 and HK2, one structure representative of the closed state and one of the open state were elected for each isozyme and the main results are shown on **Table 11**. Since there is also a possible choice for the closed state of the HK2 isozyme, PDBID:2NZT [73] was selected. The closed structure of HK1, as well as the open state structure of HK2, were elected based on the RMSD among other structures (to serve as an “average” structure), type of ligands (interaction with added value inhibitors for the open state analysis), resolution and year of deposit in PDB database. After examining these aspects, the 4F9O [72] structure of HK1 and the 5HEX for HK2 were selected for the comparison studies. The analysis of open state HK1 is more delicate, since the unique structure without the induced fit boosted by glucose is 1HKC and specifically at the *C*-terminus, and thus *N*-terminal open state pocket will not be examined. Moreover, no inhibitor has been found in this structure and therefore it is not possible to fully understand how it would fit.

Table 11. Comparative structure analysis of hHK1 and hHK2 NTD and CTD binding pockets in the open and closed state. "Domain" indicates the location of the pocket; "Residues" indicates the kind of residues evaluated; "N" refers the total number of residues considered for analysis; "RMSD" is used to compare the spatial difference between the two isozymes while identity and similarity compare the difference in the primary sequence. "Different Residues" indicates which residues are in fact different between the two isozymes, where the number corresponds to the residue in the sequence, the one letter code before the arrow corresponds to the HK2 residue and the one after the arrow corresponds to the residue present in the HK1 sequence.

Domain (state)	Residues	N	RMSD (Å)	Identity (&)	Similarity (%)	Different Residues
NTD (closed state ^a)	Pocket	27	0.33	88.9	100	88,T→S; 234,S→T
	Pocket + near residues	95	0.47	82.1	91.6	60,A→T; 88,T→S; 111,I→V, 159,H→Q; 167,F→I; 169,V→I, 170,S→T; 174,G→R; 234,S→T; 259,M→T, 297,I→V; 386,C→V, 416,V→L; 419,K→T; 447,D→S
CTD (closed state ^a)	Pocket	26	0.28	100	100	-
	Pocket + near residues	92	0.38	95.7	97.8	508,P→V; 617,L→I; 618,K→T; 741,→Y
CTD (open state ^b)	Pocket	16	0.70	100	100	-
	Pocket + near residues	58	0.91	98.3	100	741,F→Y

^aHK1, PDBID: 4F9O [72], HK2, PDBID: 2NZT [73]; ^bHK1, PDBID: 1HKC, HK2, PDBID: 5HEX ;

From the alignment and superposition of HK2 (2NZT) and HK1 (4F9O), an average RMSD of 4.20 Å was obtained. When the superposition is restricted to the pocket in NTD, a conserved structure can be recognized with RMSD = 0.33 Å, with an 88.9% identity and 100% similarity between the 27 residues in the glucose cavity. In this selection two different residues are found in the sequence of HK2 and HK1, namely 88,T→S and 234,S→T. These alterations were analyzed and do not lead to structural changes, as the three-dimensional orientation remains the same. Including the neighboring residues of the pocket (95 residues), RMSD increases to 0.47 Å and identity and similarity are still high (82.1% and 91.6%, respectively). In this selection, 16 residues were found to be different in both sequences, but, once more, no structural distortion was observed. From those, only the 170,S→T change can alter the structure, since in the HK2 structure there is no serine interaction with Gly181 by H-bonding, which is formed by threonine in HK1 structure.

In the CTD, a conserved structure of the catalytic pocket can be recognized with a RMSD = 0.28 Å, with 100% identity between the 26 residues of the glucose cavity. When neighboring residues are also considered (92 residues), RMSD increases to 0.38 Å and identity and similarity are just slightly affected (95.7% and 97.8%, respectively). Here, 4 residues are dissimilar 508,P→V; 617,L→I; 618,K→T; and 741,F→Y. From careful analysis of the effect of these mutations on the protein structures, 508,P→V, 617,L→I and 741,F→Y do not seem to affect the overall isozyme structure. However, 618,L→T seems to give rise to structural differences, since in HK1 Thr618 establishes H-bonds to Lys738 and Trp619, located very close to the pocket and to other important residues, namely Thr620 and Lys621. In HK2, the substitution of a threonine by a lysine leads to a conformational change, in which no H-bond is found

between Lys618 and other residues. The pose of Lys738 is perturbed and oriented in another direction. Visual analysis may indicate a potential influence of this mutation in the ability to accommodate a ligand and an associated induced fit (**Figure 80**).

With respect to the open state, the difference between the three-dimensional poses is more prominent. Initially, an overall RMSD difference between structures was 8.40 Å. Conversely, the C-terminal catalytic pocket (16 residues – RMSD = 0,70 Å, identity = 100%) and the catalytic pocket and nearby residues (58 residues – RMSD = 0,91 Å, identity = 98.3%) showed a rather conserved pose, even in the 1HKC structure, which does not accommodate any inhibitor, such as the HK2 structure (5HEX), undergoing an induced fit. Among the 58 residues under analysis, only the “mutation” 741,F→Y was found. Since the lateral chain of both residues conserves strongly similar properties, this change accounts for 100% similarity (**Table 11**).

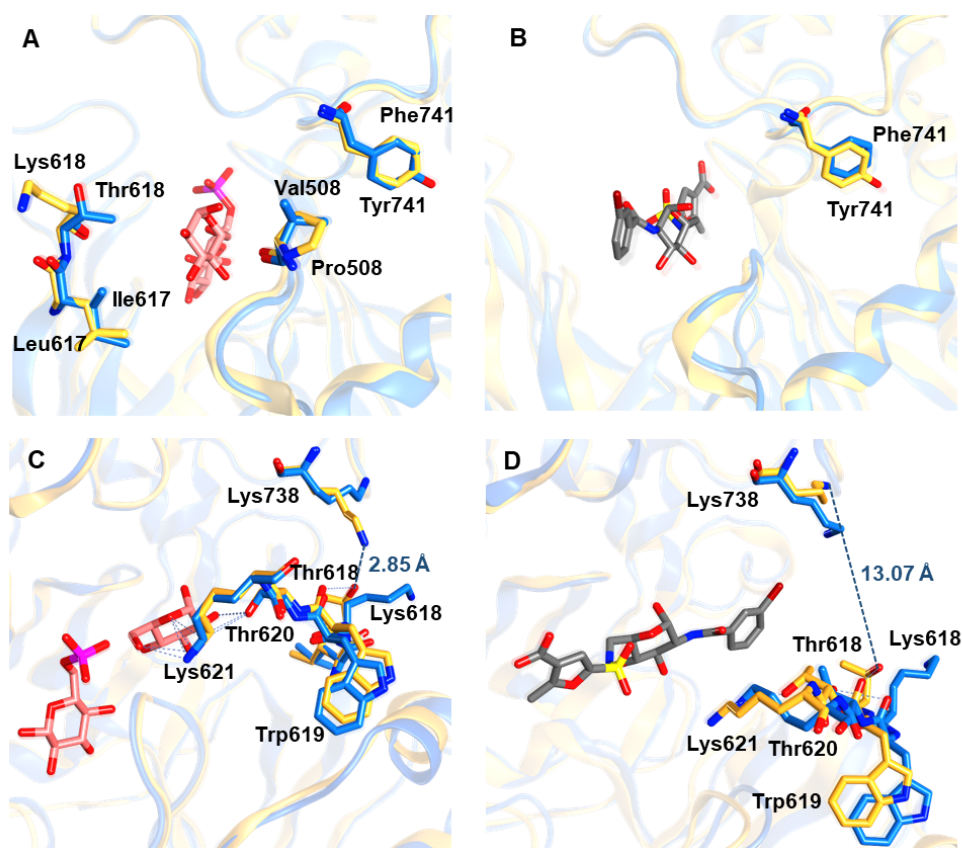


Figure 80. Comparison of HK1 and HK2 CTD pockets, considering the residues in close contact with glucose and the ones nearby (4.5 Å) in the closed (**A** and **C**) and open (**B** and **D**) state. **A** - superposition of HK2 PDBID: 2NZT (yellow) and HK1 PDBID: 4F9O (blue) with focus on the different residues in two isozyms. **B** - Superposition of HK2 PDBID: 5HEX (blue) and HK1 PDB: 1HKC (yellow) with focus on the different amino acids in the two isozyms. In **C** and **D**, a K618T change is shown for the two states, along with important nearby residues. The opening of the cavity is visually denoted by the distance between Ca of Thr618 and Lys738 (**D**). Glucose and G6P (pink) and the HK2 inhibitor, “cmpd30” [43] (grey) indicate the proximity with the ligand cavity. Adapted from [19].

To sum up, only a few subtle differences in the catalytic pocket structures were discovered that can hardly explain the differences in the HK1 and HK2 activities toward the same inhibitors. Structural resemblance represents a major difficulty in the rational design of selective inhibitors and the key for success may lie in the exploitation of differences in induced fit tuning constraints. Here, the 618,K→T “mutation” is suggested to influence this mechanism. Nevertheless, other residues, even if substantially remote from the catalytic pocket, may also be determinant. This hypothesis is supported by the abovementioned loss of activity of the *N*-terminal pocket of HK1, associated with the orientation of helix- α_{13} even though this structure is far from the glucose cavity [73]. Site-directed mutagenesis studies focusing on the CTD and the linkage between the two domains, as well as a dynamic analysis of both isozymes, are clearly required to elucidate this issue.

Surprisingly, several highly potent selective inhibitors have indeed been developed [43] (**Chapter II, section 3**). Nevertheless, the need for more SAR studies to fully understand the selectivity mechanism that will allow the rational development of new selective inhibitors of HK2 is evident. In fact, selectivity would not clearly be easy to achieve, even if it has been proven to be possible.

2.3. MD studies

2.3.1. *Design of the experiment*

This section describes the MD simulations performed to compare the evolution of HK1 and HK2 structures in aqueous solution at physiological pH over 200 ns in order to obtain a more accurate and realistic receptor/ligand-receptor model.

The first goal was to predict whether the closed and open state structures would converge to similar conformations in the absence of crystallographic ligands when allowed to relax in solution. The presence of the crystallographic ligands was also tested. The second and the major goal was to compare the behavior of HK1 and HK2 in the presence of potent inhibitors, with special focus on a highly selective HK2 inhibitor, “cmpd 33” (HK1 IC_{50} = 5 μ M; HK2 IC_{50} = 25 nM). The differences here found can be explored in future targeting studies.

Initially a set of structures were selected based on their features of interest and the quality of the structure, specifically regarding internal clashes (**Table 12**). For HK1 representatives, 4F9O and 1HKC were selected for the closed and open state, respectively. As mentioned above, 1HKC is the only HK1 structure in the open state, yet it just happens in CTD, which may have implications for the structure relaxation. Representing HK2, 2NZT for the closed state and 5HFU for the open state were preferred. Information on these structures and their ligands can be found in **Table 10** in the previous section. In addition to the crystallographic ligands, “cmpd 27” (also crystallographic ligand of 5HFU), “cmpd 30” and “cmpd 33” (**Figure 81**) were selected to be docked and simulated into the binding pockets of HK1 and HK2 open states (1HKC and 5HFU, respectively).

The analysis of the results is composed by different calculations performed throughout the simulation time such as the total energy of the system, radius of gyration, root mean square deviation (RMSD) and fluctuation (RMSF), ligands' solvent accessible surface area (SASA), distances and angles between key residues, principal component analysis (PCA) and protein-ligand interaction fingerprints (PLIF).

Table 12. PDBID of HK structures used in the MD studies.

State	HK1	HK2
Closed	4F9O	2NZT
Open	1HKC ¹	5HFU

¹ The 1HKC structure has only the CTD pocket in the open state. PDBID: 4F9O [72], 2NZT [73], 1HKC [84], 5HFU [43].

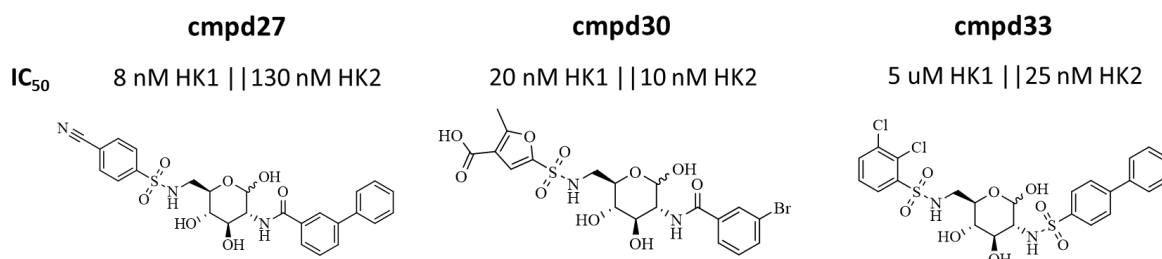


Figure 81. Structures and IC₅₀ of HK1 and HK2 inhibitors used in MD simulations.

2.3.2. Total energy

The total energy of the systems was trending constant throughout all simulations. Systems with proteins alone showed lower overall energy compared to their counterparts with ligands, whether endogenous or inhibitors (**Figure 82**). No significant difference was detected between systems with different ligands. Systems with 1HKC, the HK1 isozyme whose CTD pocket is found in the open state, present considerably lower associated energy. **Table 13** shows the mean energy and respective SD for all systems.

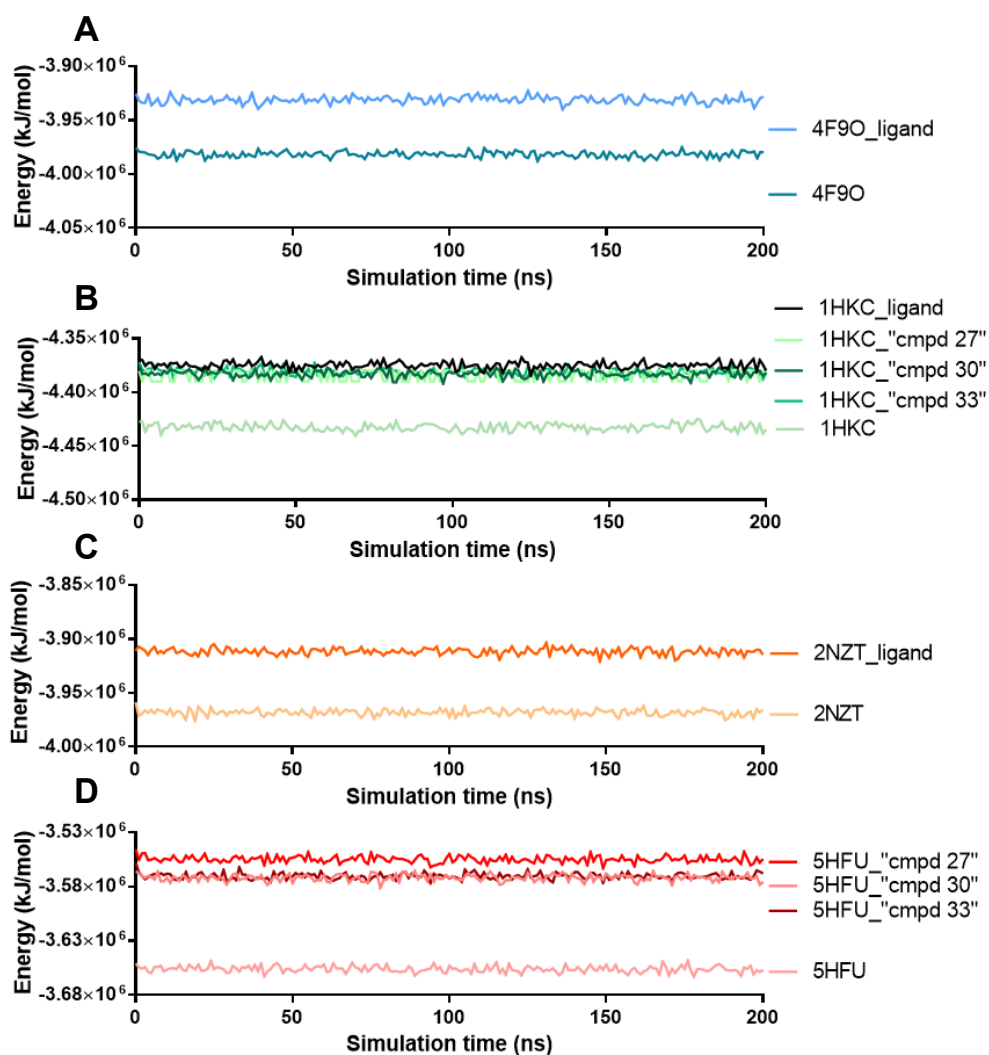


Figure 82. Plots of system's total energy throughout the simulation: **A** - HK1 (PDBID: 4F9O), closed state, with or without ligands; **B** - HK1 (PDBID: 1HKC) open state (CTD), with or without ligands or inhibitors; **C** - HK2 (PDBID: 2NZT), closed state, with and without ligands; **D** - HK2 (PDBID: 5HFU), open state, with or without inhibitors.

Table 13. Average Total Energy (\pm SD) for each system.

Isozyme	State	Ligand	PDBID	Aver. Total energy (kJ/mol)
HK1	Closed	-	4F9O	-3.98E+06 \pm 2.74E+03
	Closed	Cryst. ¹	4F9O	-3.93E+06 \pm 3.01E+03
	Open ²	-	1HKC	-4.43E+06 \pm 3.18E+03
	Open ²	Cryst. ¹	1HKC	-4.38E+06 \pm 3.19E+03
	Open ²	"Cmpd 27"	1HKC	-4.38E+06 \pm 5.02E+03
	Open ²	"Cmpd 30"	1HKC	-4.38E+06 \pm 3.18E+03
	Open ²	"Cmpd 33"	1HKC	-4.38E+06 \pm 3.18E+03
HK2	Closed	-	2N2T	-3.97E+06 \pm 2.98E+03
	Closed	Cryst. ¹	2N2T	-3.91E+06 \pm 3.00E+03
	Open	-	5HFU	-3.66E+06 \pm 2.82E+03
	Open	"Cmpd 27" ³	5HFU	-3.56E+06 \pm 2.86E+03
	Open	"Cmpd 30"	5HFU	-3.57E+06 \pm 2.84E+03
	Open	"Cmpd 33"	5HFU	-3.57E+06 \pm 2.96E+03

¹Cryst. – Crystallographic ligands; ²Open – 1HKC structure has only the CTD pocket in the open state; ³"Cmpd 27" – "cmpd 27" is the crystallographic ligand of 5HFU.

2.3.3. Root mean square deviation (RMSD)

The root mean square deviations were calculated to understand whether the system converges to a stable group of conformations and when does it occur. **Table 14** shows the average RMSD and respective SD for each system (calculated considering the backbone C α atoms). All systems present an average RMSD value lower than 10 Å. However, unexpectedly, structure 2N2T (HK2, closed state), with its crystallographic ligands, show a considerably higher RMSD profile than the protein alone (2N2T) (**Figure 83 - A**). In the absence of ligands, the 4F9O system (HK1, closed state), also experiments a higher average RMSD and an atypical change throughout the simulation, but the RMSD in the last 50 ns tends to stabilize and approximate to the RMSD of the same structure but in presence of crystallographic ligands (**Figure 83 - B**).

Regarding the open state of isozymes, the highest changes in RMSD compared with the stable profile of the structures without ligands occur with the isozymes in the presence of "cmpd 33". In HK1, the 1HKC system (CTD open state) with "cmpd33", tends to vary more from the starting conformation comparing with 1HKC alone (CTD open state) (**Figure 83 - C**). This difference is higher in 5HFU (HK2, open state), since "cmpd 33" does not let the structure to fully stabilize during the 200 ns of simulation (**Figure 83 - D**).

The RMSD for each domain of the isozymes, NTD and CTD was also calculated in order to find out which part of the protein deviates the most from the initial structure. For HK1 systems, both domains diverge similarly. In the 4F9O systems with crystallographic ligands and the 1HKC systems with

crystallographic ligand a slightly higher CTD deviation was found. The 1HKC system with inhibitor “cmpd 33” has a slightly increased NTD RMSD compared to CTD. As for HK2, the RMSD of the 2NZT and 5HFU systems without ligands did not change considerably. In the presence of glucose and G6P, 2NZT showed an increased RMSD in NTD, while 5HFU in presence of inhibitors “cmpd 27” and “cmpd 33” showed a higher deviation from CTD. The highest increase in RMSD occurred with “cmpd 33” in the last 50 ns of simulation, which may explain the increase in the overall structure RMSD. No pattern was found; however, “cmpd 33” destabilized the NTD of HK1 and the CTD of HK2 the most (**Figure 84**).

Table 14. Average RMSD (\pm SD) for the total protein, NTD and CTD of each system.

Isozyme	State	Ligand	PDBID	Aver. RMSD (Å)	RMSD NTD (Å)	RMSD CTD (Å)
HK1	Closed	-	4F9O	6.88 \pm 1.88	3.51 \pm 0.36	4.01 \pm 0.53
	Closed	Cryst. ¹	4F9O	4.92 \pm 1.06	3.90 \pm 0.53	4.04 \pm 0.42
	Open ²	-	1HKC	5.07 \pm 1.18	3.47 \pm 0.43	3.79 \pm 0.36
	Open ²	Cryst. ¹	1HKC	5.48 \pm 1.24	3.10 \pm 0.32	3.61 \pm 0.41
	Open ²	“Cmpd 27”	1HKC	5.80 \pm 0.84	3.95 \pm 0.50	4.08 \pm 0.54
	Open ²	“Cmpd 30”	1HKC	4.91 \pm 1.21	3.09 \pm 0.29	3.08 \pm 0.28
	Open ²	“Cmpd 33”	1HKC	7.12 \pm 1.33	4.22 \pm 0.69	3.56 \pm 0.36
HK2	Closed	-	2NZT	4.53 \pm 1.06	3.84 \pm 0.56	3.37 \pm 0.33
	Closed	Cryst. ¹	2NZT	9.32 \pm 1.70	3.80 \pm 0.47	3.71 \pm 0.39
	Open	-	5HFU	4.41 \pm 0.83	4.23 \pm 0.62	4.32 \pm 0.51
	Open	“Cmpd 27” ³	5HFU	5.37 \pm 0.81	3.94 \pm 0.41	4.26 \pm 0.54
	Open	“Cmpd 30”	5HFU	5.75 \pm 0.93	3.75 \pm 0.56	3.71 \pm 0.39
	Open	“Cmpd 33”	5HFU	8.69 \pm 3.09	4.13 \pm 0.63	4.56 \pm 1.08

¹Cryst. – Crystallographic ligands; ²Open – 1HKC structure has only the CTD pocket in the open state; ³“Cmpd 27” – “cmpd 27” is the crystallographic ligand of 5HFU.

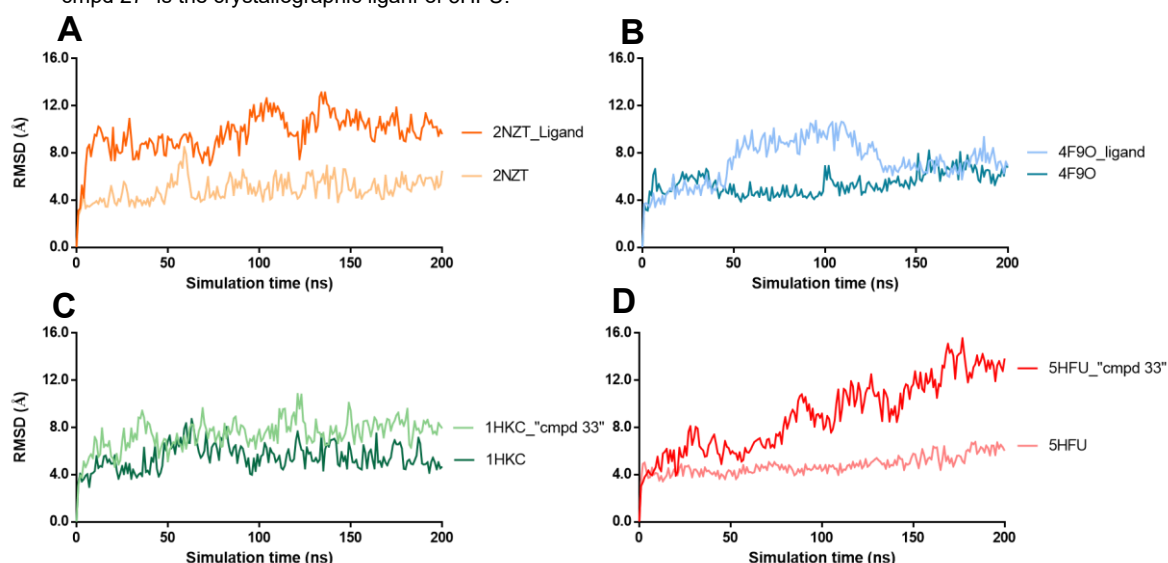


Figure 83. Plots of RMSD throughout the simulation for systems 2NZT and 2NZT plus crystallographic ligands (**A**), 4F9O and 4F9O plus crystallographic ligands (**B**), 1HKC and 1HKC plus inhibitor “cmpd 33” (**C**), and 5HFU and 5HFU plus inhibitor “cmpd 33” (**D**).

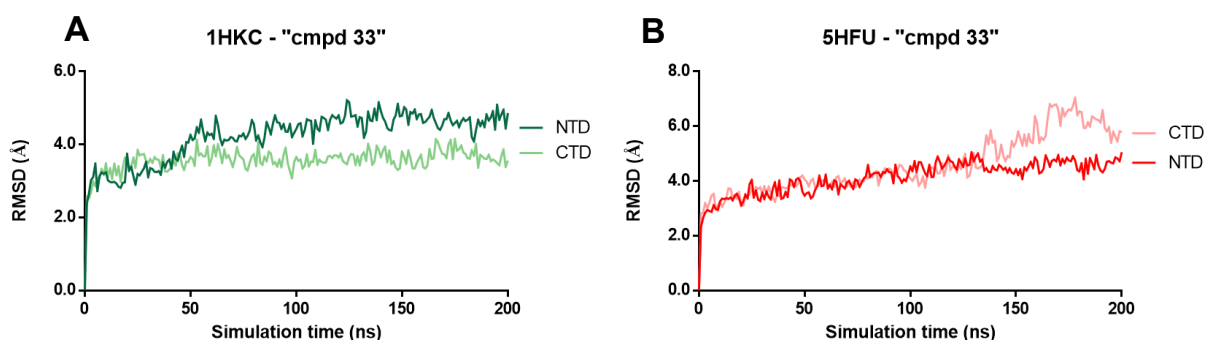


Figure 84. RMSD (C α) of each domain (NTD and CTD) for the systems 1HKC plus "cmpd 33" (A) and 5HFU plus "cmpd 33" (B) as a function of time.

2.3.4. Radius of gyration

Together with RMSD, the radius of gyration (R_g) also provides information about the stability of the protein across the simulation. R_g is described as the overall spread of the molecule. **Figure 85** shows this descriptor as a function of time for all systems. While the closed state HK1 (4F9O) behaves similarly with and without ligands and R_g decreases trendily over time (**Figure 85 - A**), the closed state HK2 (2NZT) with ligand presents a lower R_g during 90% of the simulation compared with the counterpart without any ligand (**Figure 85 - C**). Regarding the open state isozymes (HK1: 1HKC; HK2: 5HFU), the R_g values in systems with endogenous ligands, "cmpd 27" and "cmpd 30" seems to be considerably stable and similar to systems without any ligand. Interestingly, with "cmpd 33" the two isozymes behaved differently. Whilst R_g in the system with HK1 is decreased compared with the other HK1 systems, R_g of HK2 with "cmpd 33" consistently increases after 74 ns (**Figure 85 – B and D**). Observing the last 50 ns of simulation, stabilization of R_g seems to be achieved for all systems.

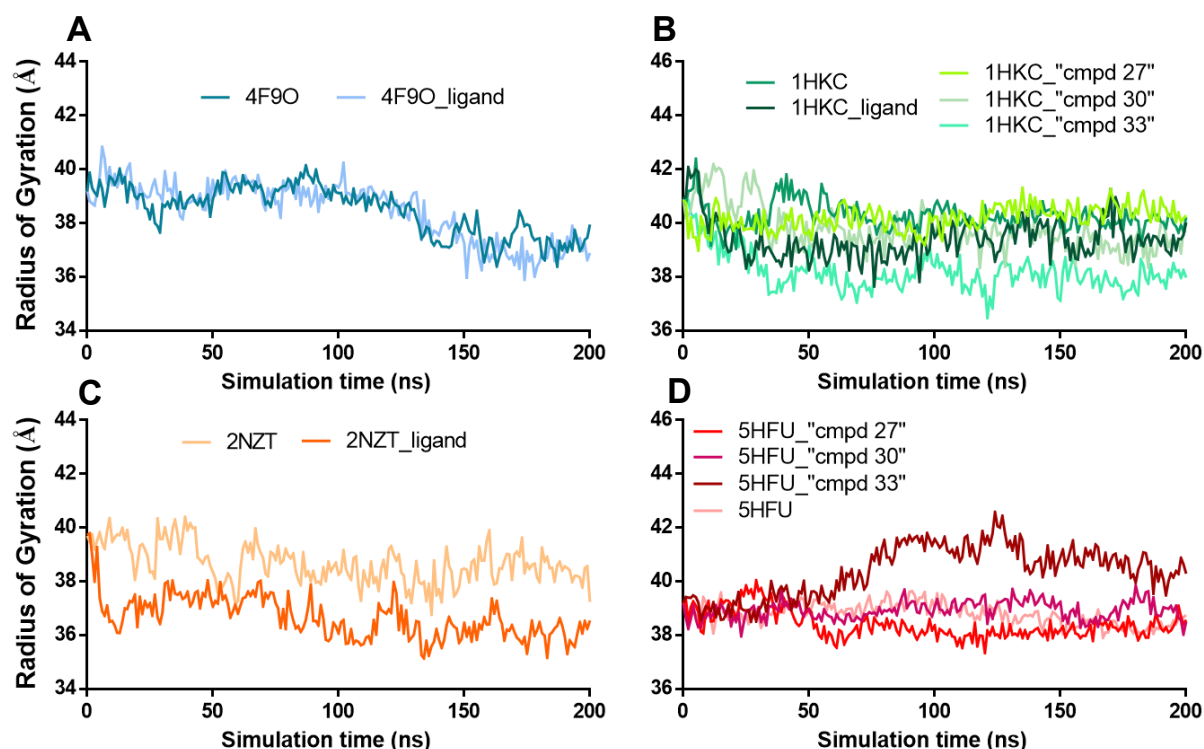


Figure 85. R_g as a function of time for all systems, with and without ligands. **A** - HK1 (PDBID: 4F9O) closed state; **B** - HK1 (PDBID: 1HKC) open state (CTD); **C** - HK2 (PDBID: 2NZT) closed state; **D** - HK2 (PDBID: 5HFU) open state.

2.3.5. Clusters of protein conformations

The most populated structures from each simulation were obtained using the gmx cluster function of GROMACS, applying a 0.4 nm cutoff as RMSD to 201 structures coming from the 200 ns of simulation plus the initial structure. **Table 15** shows the number of clusters found for each system and the population of structures (%) in the first three clusters. The higher the number of clusters, the more diverse are the conformations found throughout the simulation. The distribution of structures across the different clusters indicates the stability of each conformation over time (more structures in a cluster suggests higher stability of the conformation represented by the cluster).

In large contrast with the counterpart without any ligand, system 5HFU plus "compd 33" had the highest number of clusters, and the ones more evenly distributed. This means that a great variety of considerably different conformations can be found and there are three that are the most plausible. These data indicate a lower stability of the system. On the other hand, conformations of 5HFU without ligands can be divided in only 2 clusters, where the first one contains the large majority of structures (99%). This indicates that this structure did not change substantially over time, compared with the other structures.

With the exception of the closed structure of HK1 (4F9O), whose structures are distributed in more than one cluster, most systems have at least two thirds of the total number of conformations in the first cluster, showing more stability overtime.

Table 15. Number of conformation clusters for each system and the proportion of structures in cluster 1, 2 and 3 (clusters numbering is inversely proportional to the conformation population of the cluster, i.e., cluster 1 is the most populated. Other clusters are poorly populated.).

Isozyme	State	Ligand	PDBID	Clusters	Structures on cluster 1	Structures on cluster 2	Structures on cluster 3
HK1	Closed	-	4F9O	5	42%	35%	19%
	Closed	Cryst. ¹	4F9O	8	53%	20%	15%
	Open ²	-	1HKC	7	69%	18%	7%
	Open ²	Cryst. ¹	1HKC	6	89%	6%	2%
	Open ²	"Cmpd 27"	1HKC	7	71%	6%	3%
	Open ²	"Cmpd 30"	1HKC	5	79%	10%	6%
	Open ²	"Cmpd 33"	1HKC	4	89%	8%	2%
HK2	Closed	-	2NZT	9	69%	12%	8%
	Closed	Cryst. ¹	2NZT	5	74%	16%	7%
	Open	-	5HFU	2	99%	1%	-
	Open	"Cmpd 27" ³	5HFU	4	88%	10%	1%
	Open	"Cmpd 30"	5HFU	6	86%	9%	3%
	Open	"Cmpd 33"	5HFU	10	32%	29%	25%

¹Cryst. – Crystallographic ligands; ²Open – 1HKC structure has only the CTD pocket in the open state; ³"Cmpd 27" – "cmpd 27" is the crystallographic ligand of 5HFU.

In order to conclude whether the closed and open state structures of each isozyme mainly converge to a similar structure, the structures of the most populated cluster of 4F9O and 1HKC (HK1) and 2NZT and 5HFU (HK2) were compared and RMSDs were measured (**Figure 86**). These results suggest that the HK1 and HK2 structures are very fluid and do not converge to similar conformations. Interestingly, a considerable difference was found between RMSD of the CTD and the NTD of HK2 (**Figure 86 C and D**): the higher RMSD values are clearly attributed to residues in the first half of the isozyme, which possibly indicates a higher mobility of the NTD.

RMSDs between the structure of the first cluster of each system with and without the crystallographic ligand were also measured (**Figure 87**). While the RMSDs between structures with and without glucose and G6P as crystallographic ligands were considerably high, the lowest RMSDs were found between 5HFU and 5HFU with "cmpd 27" (crystallographic ligand).

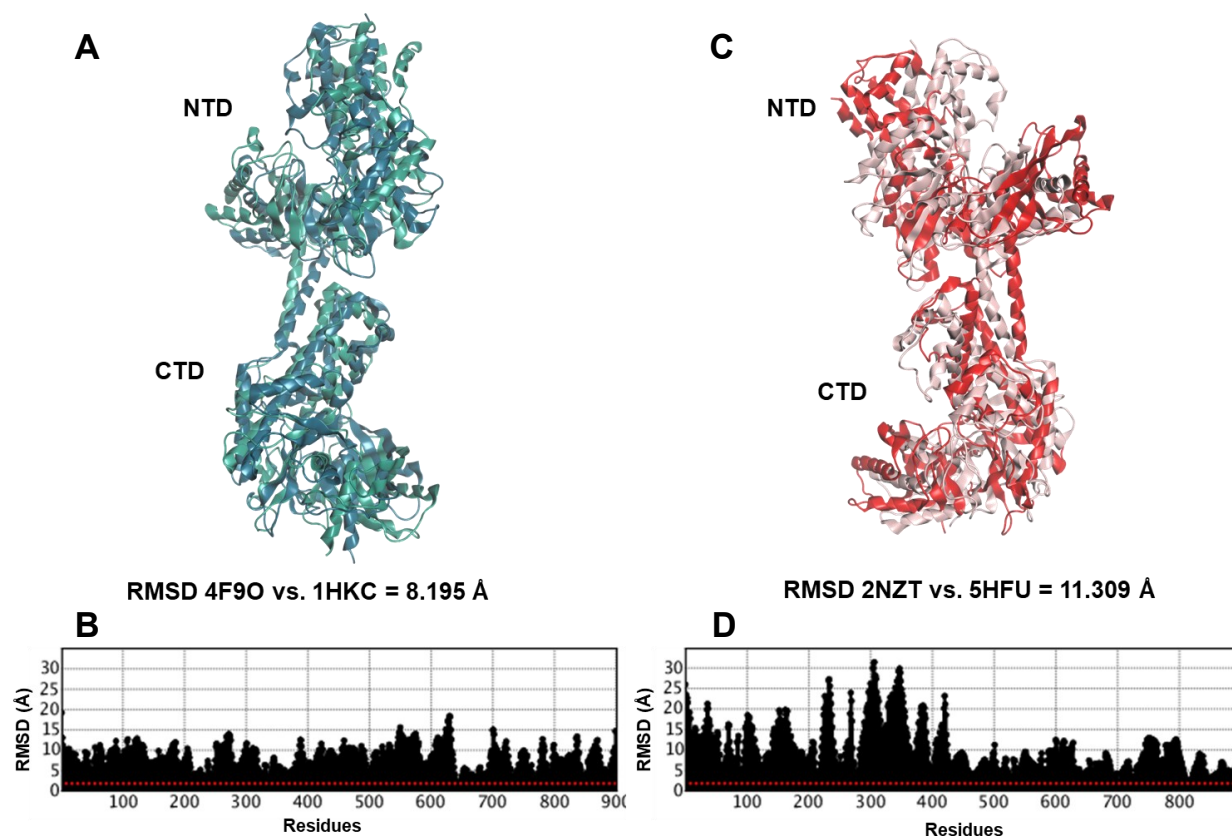


Figure 86. Comparison between the representative structure of the most populated cluster of each system (without ligands): closed versus open state of HK1 (A) and HK2 (C) and respective average RMSD. RMSD by residue is shown in the plots B and D.

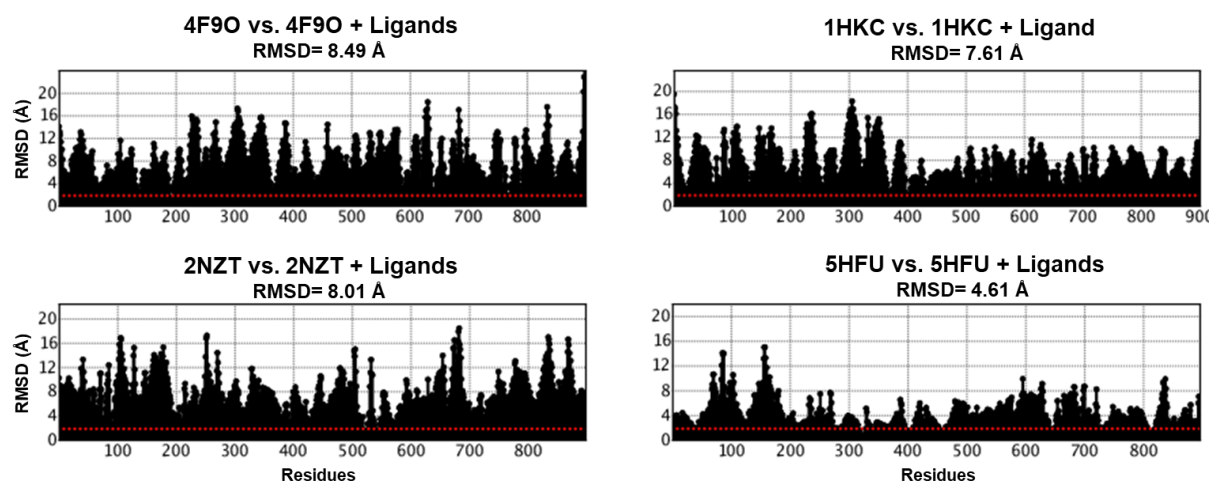


Figure 87. Average RMSD and RMSD by residues between each system with and without crystallographic ligands.

2.3.6. Ligands RMSD

To discover how the conformation of the ligands, crystallographic or inhibitors, evolves from the initial structure throughout the simulation, the RMSD was calculated for each molecule. **Figure 88 – A** shows the average RMSD (plus SD) for each ligand. The endogenous crystallographic ligands display lower RMSDs when compared to the inhibitors. Unexpectedly, inhibitor “cmpd27”, the crystallographic ligand of 5HFU, caused the highest RMSD in both HK1 and HK2. As a crystallographic ligand, its conformation in the pocket was expected to be similar to that in equilibrium in solution. On the other hand, inhibitor “cmpd 30” exhibit one of the lowest RMSDs (2.90 Å) in the CTD pocket. It should be noticed that the initial conformations of inhibitors “cmpd 30”, “cmpd 33” and “cmpd27” (in HK1) were found through molecular docking. Generally, RMSDs were kept stable after the beginning of the simulation (see example plots on **Figure 88 - B and C**).

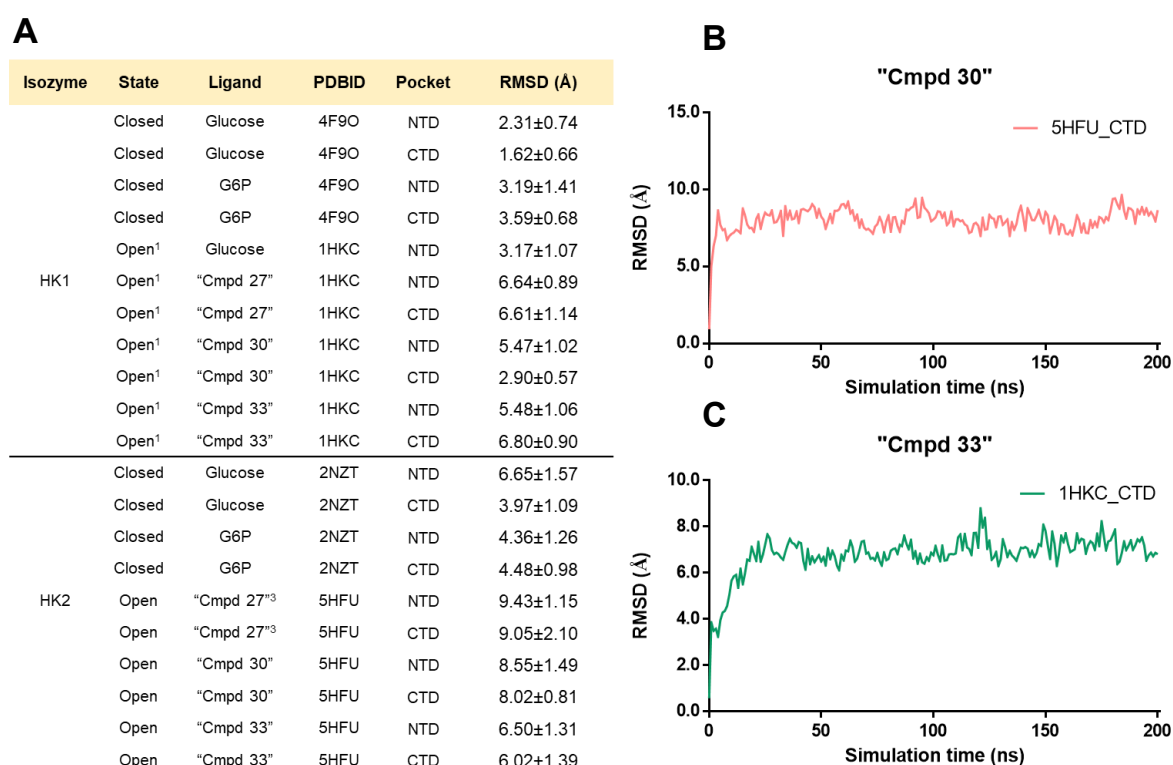


Figure 88. Ligands RMSD throughout the simulation. **A** - Average RMSD (\pm SD) for all ligands in all systems (¹Open – 1HKC structure has only the CTD pocket in the open state); **B** and **C** - Examples of ligands RMSD over the simulation in HK2 (5HFU, **B**) and HK1 (1HKC, **C**).

2.3.7. Root mean square fluctuation (RMSF)

The root mean square fluctuation (RMSF) was calculated (using the GROMACS module gmx rmsf) to observe the residues deviations compared to the initial structures. A summary of the results is shown in **Figure 89**. **Figure 89-A** illustrates the RMSF plots of all systems, showing how dynamic these systems are. Still, approximately the same pattern is observed for all cases, and generally the residues

VI - Structural comparison between HK1 and HK2: a Molecular Dynamics approach to investigate potential drug selectivity

that fluctuate the most are the same, changing, however in different ranges (reddish plots – higher fluctuations; blueish plots - lower fluctuations). In **Figure 89-B** the HK1 structure (1HKC) displays the regions that fluctuate the most in all systems throughout the simulations, showing that regions more exposed to the solvent are the ones more susceptible to undergo more variations in their positions. As expected, these include the initial and final helices- α . To understand which systems fluctuate the most, **Figure 89 - C** presents the sum of RMSF for each system together with its discrimination for the NTD and CTD regions.

The system that fluctuates the least is the ligand-free open structure of HK2 (5HFU). Conversely, the system that shows higher RMSF residue fluctuations is the system composed equally by cmpd 33-bound 5HFU ("cmpd 33" is the inhibitor with higher selectivity for HK2, displaying an HK1 IC_{50} = 5 μ M and HK2 IC_{50} = 25 nM) (**Figure 89 - A**). Interestingly, in the presence of this compound, 1HKC experiments the smallest overall variation (**Figure 89 - B**). The presence of other inhibitors ("cmpd 27" and "cmpd 30") and crystallographic ligands did not markedly affect residue fluctuation relative to the ligand-free protein structures.

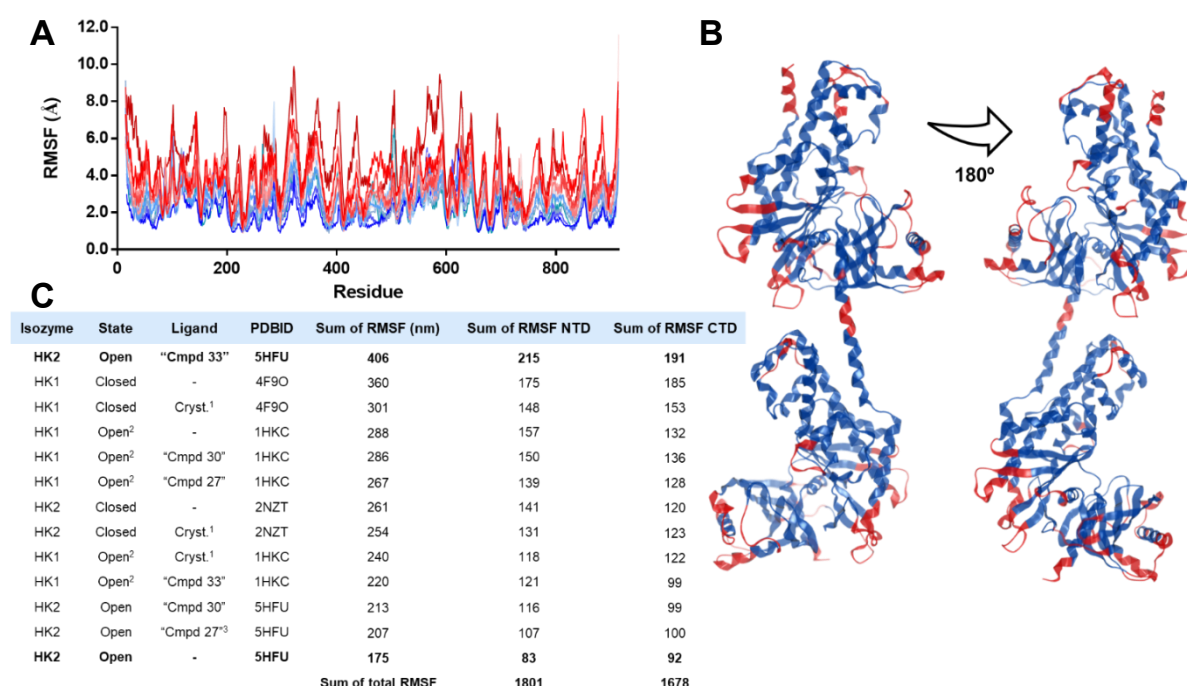


Figure 89. RMSF of each residue for all the systems in study. **A** - RMSF plot of all systems together, showing that differences occur in the same residues but in different ranges; **B** - Red regions in the structure of HK1 (PDBID: 1HKC) indicate the most affected residues throughout all simulations. The open state of the CTD pocket is also visible; **C** - The table indicates the sum of total, NTD and CTD RMSF's for each system. ¹ Cryst. - Crystallographic ligands; ² Open - 1HKC structure has only the CTD pocket in the open state; ³ "Cmpd 27" - "cmpd 27" is the crystallographic ligand of 5HFU.

NTD is the domain that presents higher residue oscillations (**Figure 90- C**). For instance, **Figure 90-C** shows the RMSF plots of 1HKC and 5HFU, evidencing a slightly higher fluctuation on NTD. The region 449-479 corresponds to the linker helix- α_{13} (linker between NTD and CTD), whose variation in the 1HKC system stands out. Structure 1HKC originally displays an endogenous ligand in the NTD pocket, presenting a closed conformation in this domain. This result indicates that, in solution without ligands, HK2 tends to display a more open conformation. A considerable higher fluctuation appears on the region composed by residues 740-826, part of the CTD large subdomain (**Figure 91**). The fluctuation of this region in HK1 and HK2 in presence of “cmpd33” stands out (more dramatically in HK2) since it does not match the fluctuation pattern observed for the same region of the systems without ligands. Interestingly, the large subdomain was previously described as a rigid region comparing with other regions such as the small subdomain (non-rigid) that should be the one which fluctuates the most [74,75,83] (*cf chapter II, section 2.2.*). However, the small subdomain region fluctuation’s pattern does not change substantially from the systems without any ligand, despite the general increase in the fluctuation of all regions in HK2 with “cmpd 33” (**Figure 90 – A and B**).

The closed structure of HK1 (4F9O) seems to behave differently and CTD residues seems to fluctuate more in comparison with the CTD of the HK2 closed structure (**Figure 90- D**), and this starts right at the beginning of helix- α_{13} .

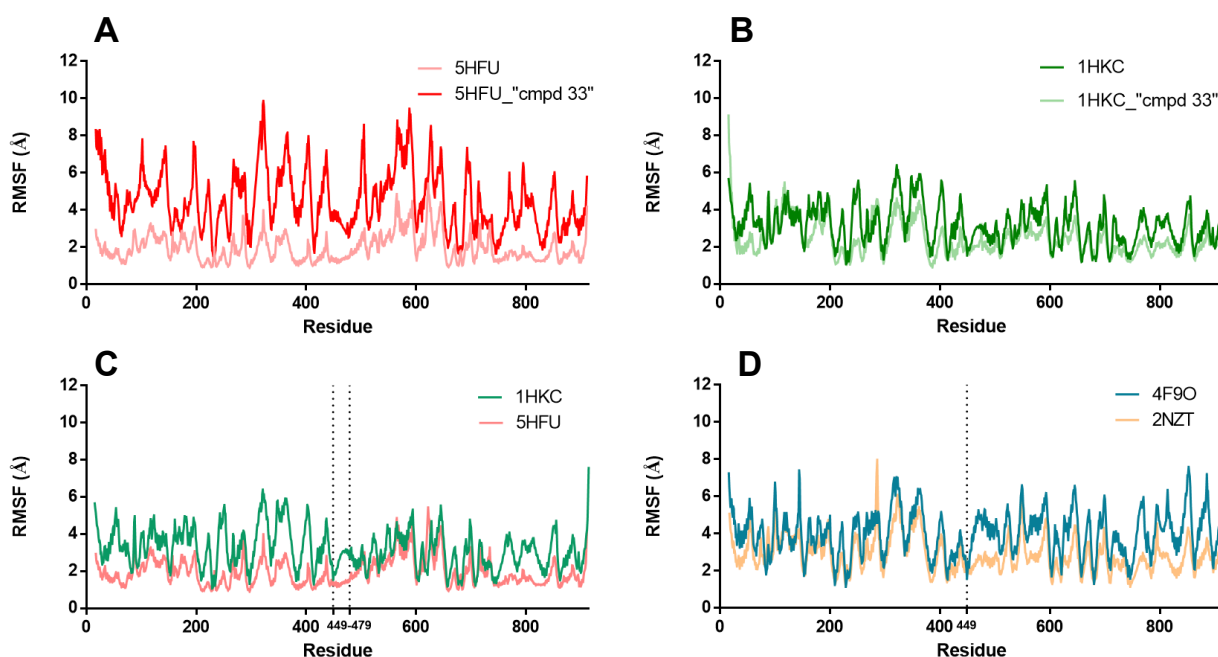


Figure 90. Most relevant RMSF differences among the systems in study. **A** - RMSF plots of the lowest and highest fluctuation patterns of all systems, 5HFU (light salmon) and 5HFU + “cmpd33” (red), respectively; **B** - RMSF plots of 1HKC (green) and 1HKC + “cmpd 33” (light green). **C** - RMSF plots of open state HK2 (5HFU -light salmon) and HK1(1HKC - green); 449-479 indicates de region of the linker helix- α_{13} . **D** - RMSF plots of closed state HK1 (4F9O - blue) and HK2 (2NZT- orange).



Figure 91. Structure of HK1 (PDBID: 1HKC) with region Res740-Res826 depicted (cyan), corresponding to the majority of the large subdomain of CTD. The small subdomain at highlighted in yellow.

2.3.8. Ligands' solvent-accessible surface area (SASA)

The solvent-accessible surface area (SASA) of ligands was calculated (using the GROMACS module gmx sasa) in order to understand how the contact with solvent changes overtime, and consequently how the structure embraces the ligand and closes the conformation. The amplitude of SASA does not seem to change considerably throughout the simulation for both endogenous ligands and inhibitors, independently of the pocket. When normalized by the number of heavy atoms, endogenous ligands' SASA is even higher than inhibitors' SASA (**Figure 92- A**), which does not agree with the closed state of the protein when bound to glucose and open state in presence of inhibitors. Differences between glucose and G6P in the different pockets and systems (4F9O, 2NZT and 1HKC-glucose only) were found to be insignificant. Except for “cmpd 27” and “cmpd 33” in 5HFU systems, no significant differences were found between ligands/inhibitors in different pockets.

The statistical difference between “cmpd 27” SASA distribution in the two pockets of 5HFU was verified, however, when the plot is analyzed (**Figure 92- B**) it is confirmed that after ~60 ns both molecules converge to a similar SASA until the end of the simulation.

Regarding the difference between “cmpd 33” in the different pockets, it might be significant, since this statistical difference is also verified between the two CTDs with “cmpd 33” of 1HKC and 5HFU. In **Figure 92 – C** the plot of “cmpd 33” SASA for NTD and CTD pockets in 5HFU and 1HKC systems indicates a consistently higher SASA for CTD “cmpd 33” in the 5HFU system, which is confirmed by the probability mass function in **Figure 92 – D**.

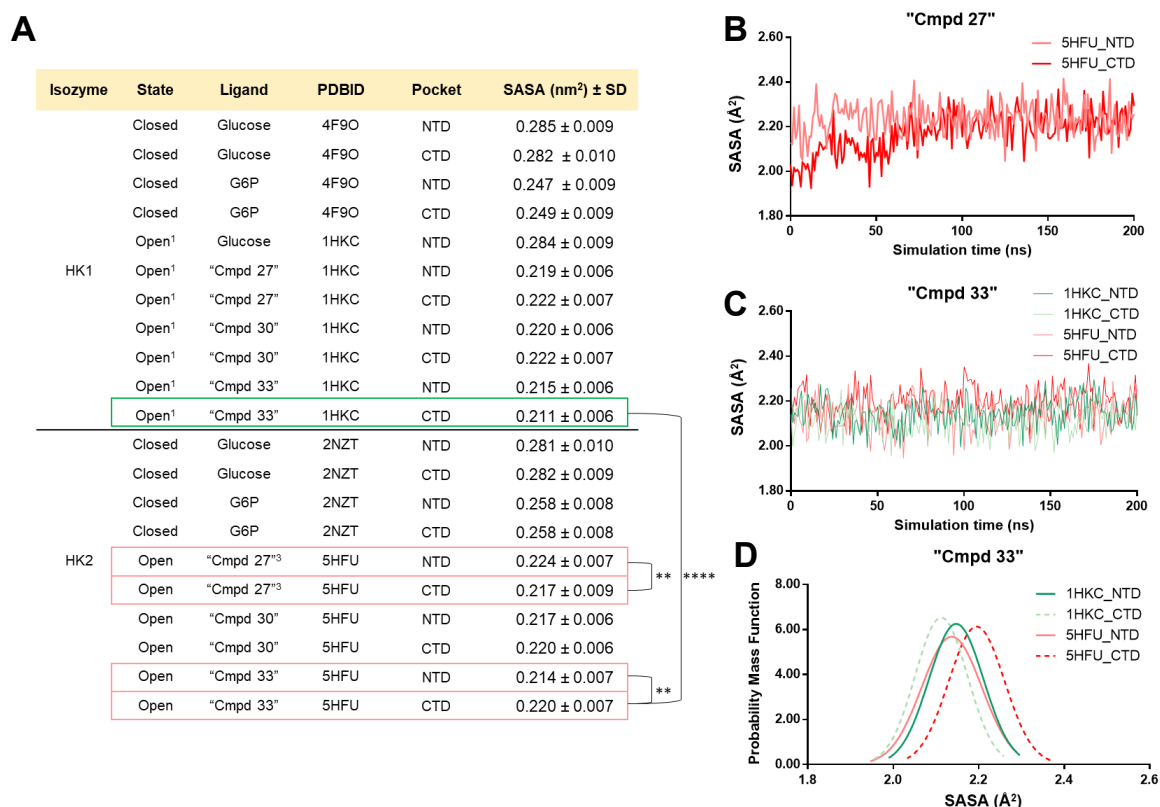


Figure 92. Main changes in ligands and inhibitors' SASA. **A** - Normalized average SASA (\pm SD) for each ligand/inhibitor in each pocket. The statistical significance (Kruskal-Wallis test) between the distribution of SASA for relevant comparisons is depicted. ** $p < 0.001$, **** $p < 0.0001$. **B** - Plot of "cmpd 27" SASA throughout the simulation in the NTD and CTD pockets, for the 5HFU system. **C** and **D** - Plot (**C**) and probability mass function (**D**) of "cmpd 33" SASA throughout the simulation in the NTD and CTD pockets, for the 5HFU and 1HKC systems.

2.3.9. Distance of ligands to the catalytic Asp209/657

The Asp209 residue in the NTD pocket, and its CTD-pocket Asp657 counterpart, have been demonstrated to be key catalytic residues responsible for the facilitation of the catalysis process of HK1/2, i.e., the transfer of a phosphate group from ATP to glucose, forming G6P (*cf* **chapter II, section 2**). In the crystallographic structures of HK2 5HFU, 5HEX and 5HG1, inhibitors "cmpd 27", "cmpd 30" and "cmpd 1" interact with Asp209/657 (**Figure 93**). In this regard, the distance between the center of mass of the inhibitors and the center of mass of the Asp residues was measured throughout the simulations (**Figure 94**) to find out how the position of the ligands changes towards the catalytic pocket.

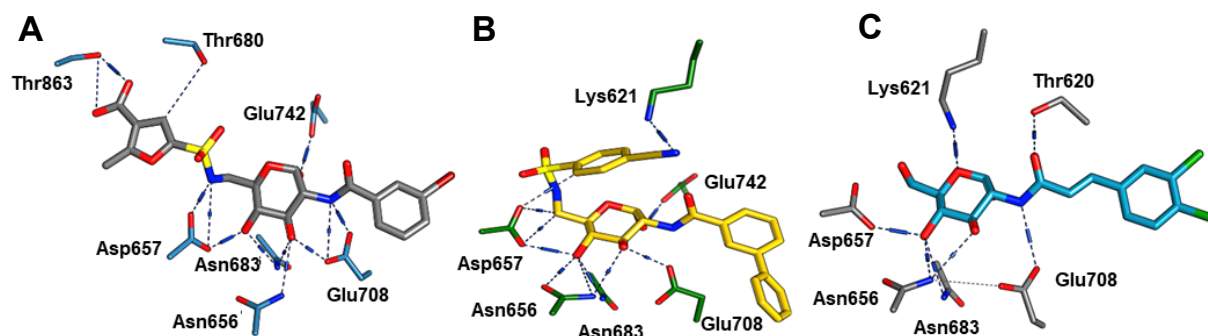


Figure 93. Interaction of synthetic co-crystallized inhibitors with HK2 in the catalytic center (CTD): **A** – PDBID: 5HEX with “cmpd 30” ($IC_{50} = 10$ nM); **B** - PDBID: 5HFU with “cmpd 27” ($IC_{50} = 130$ nM); **C** - PDB:5HG1 with “cmpd 1” ($IC_{50} = 6.3$ μ M). Hydrogen bonds between the lateral chain of surrounding residues and the inhibitor are depicted in dark blue.

The distances between inhibitors and Asp209 are similar between HK1 and HK2 in the presence of “cmpd 27” and did not change significantly since the initial conformation (**Figure 94 – A**). Unexpectedly, the distance that increases the most from the starting of the simulation is the one between 5HFU Asp657 (CTD) and “cmpd 27”, the crystallographic ligand of 5HFU (**Figure 94 – B**). The distances between residue Asp209 and “cmpd 30” or “cmpd 33” are much larger in 5HFU than in 1HKC, throughout the entire simulation (**Figure 94 – C and E**). The same is valid for the distance between Asp657 and “cmpd 33” (**Figure 94 – F**); however, with “cmpd 30” the distance between the inhibitor and Asp657 is very similar and is kept stable until the end of the simulation (**Figure 94 – D**).

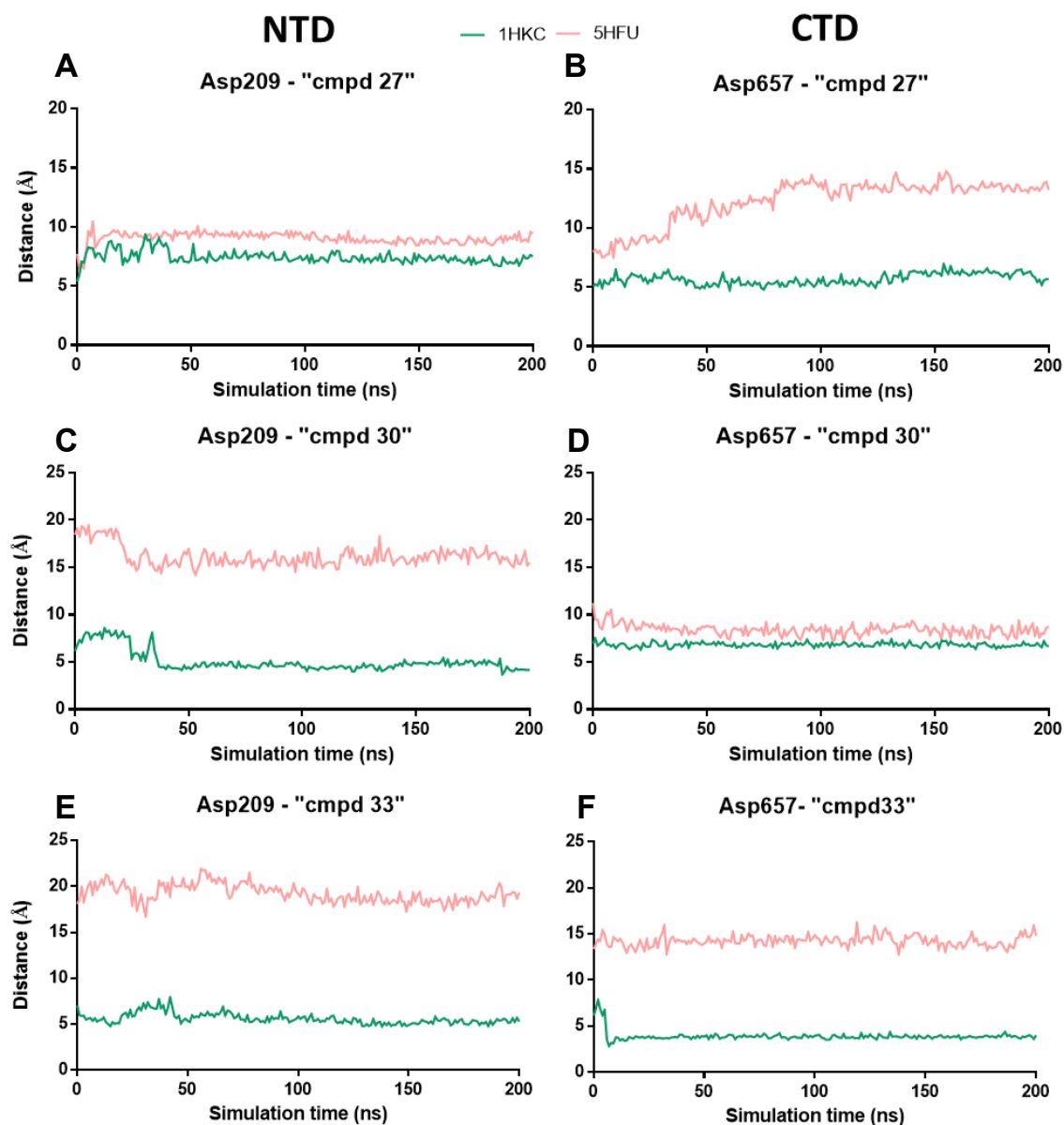


Figure 94. Distance between the center of mass of residue Asp209/657 and the center of mass of inhibitors "cmpd 27" (A/B), "cmpd 30" (C/D) and "cmpd 33" (E/F) in systems 1HKC (green) and 5HFU (pink), throughout the simulation.

2.3.10. Comparing the displacement of helix- α_{13}

Recently, Ferreira et al. [271] reported that the displacement of the linker helix- α_{13} impacted the relative orientation of the NTD and CTD in an HK2 mutant with loss of *N*-terminal catalytic activity, where a wider angle between the two domains was found to be the most probable in the MD simulation. Herein the angle and distance between NTD and CTD were also compared in all systems to verify if the presence of inhibitors affects the displacement of helix- α_{13} as well. The angle formed by C α of Thr389 (NTD), Arg468 (helix- α_{13}) and Ala839 (CTD) was measured throughout the simulation. **Table 16** shows the average angle for each system (\pm SD). Contrary to what was found previously [271], the average angles of all systems were kept between 120-130°, with one exception. The closed state of HK2, in the presence of endogenous ligands, seems to display both domains much closed to each other, presenting a much lower average angle ($95.2^\circ \pm 7.3$). The same does not happen with HK1 in the same state (4F9O+ligands), which presents a higher angle ($130.4^\circ \pm 8.8$) than in the absence of ligands ($118.3^\circ \pm 11.8$).

Apart from 5HFU, generally without any endogenous ligand or in their presence, this angle varies considerably throughout the simulation as demonstrated by the probability mass function graphs (broad curves, **Figure 95**). 5HFU, the open structure of HK2, keeps the probability mass function quite sharp in absence of ligands and in presence of “cmpd 27”, showing that during the simulation the structure presents an angle very similar to the average, *i.e.*, stability. The probability mass function for the inhibitor “cmpd 30” in 5HFU is wider, and similar to that with the same compound in 1HKC (HK1). The biggest difference is associated with the presence of “cmpd 33” in 5HFU, where the probability mass function curve is broader, showing instability or higher movement of the structure throughout the simulation in the presence of this inhibitor. On the contrary, HK1 (1HKC) shows similar curves with the parent system without ligand.

Regarding the angle, 1HKC tends to show a slightly lower angle than the 1HKC system without ligands. In average, the mean Thr389-Arg468-Ala839 angle in 5HFU with inhibitors does not change much from that of the 5HFU system without any ligand.

As expected, the probability mass function distances between Thr389-Ala839 follow the same distribution pattern for all systems, still, the behavior of 5HFU and 1HKC seems to be slightly different in the presence of “cmpd 33”, independently of the angle (**Figure 96**). The distance between the two residues in the 1HKC system with “cmpd 33” is substantially reduced when compared to the structure without the inhibitor. On the opposite side, this distance is consistently larger in 5HFU with “cmpd 33” relative to the system without ligands.

In conclusion, “cmpd33” seems to destabilize the HK2 structure more prominently than the HK1 structure. This destabilization, as well as the increased distance between NTD and CTD, is in agreement with the increased angle (Thr389-Arg468-Ala839) found in the MD simulations of Ferreira *et al.* (2020) [271] for a non-NTD catalytically active HK2 mutant.

Table 16. Average Thr389-Arg468-Ala839 angle (°) ± SD for each system.

Isozyme	State	Ligand	PDBID	Angle 339-468-839 (°)
HK1	Closed	-	4F9O	118.3 ± 11.8
	Closed	Cryst. ¹	4F9O	130.4 ± 8.8
	Open ²	-	1HKC	129.7 ± 5.4
	Open ²	Cryst. ¹	1HKC	125.9 ± 6.3
	Open ²	"Cmpd 27"	1HKC	121.6 ± 4.4
	Open ²	"Cmpd 30"	1HKC	124.6 ± 7.5
	Open ²	"Cmpd 33"	1HKC	125.7 ± 4.0
HK2	Closed	-	2NZT	125.0 ± 6.4
	Closed	Cryst. ¹	2NZT	95.2 ± 7.3
	Open	-	5HFU	122.8 ± 2.9
	Open	"Cmpd 27" ³	5HFU	119.7 ± 3.3
	Open	"Cmpd 30"	5HFU	122.4 ± 5.3
	Open	"Cmpd 33"	5HFU	126.9 ± 9.4

¹Cryst. – Crystallographic ligands; ²Open – 1HKC structure has only the CTD pocket in the open state; ³"Cmpd 27" – "cmpd 27" is the crystallographic ligand of 5HFU.

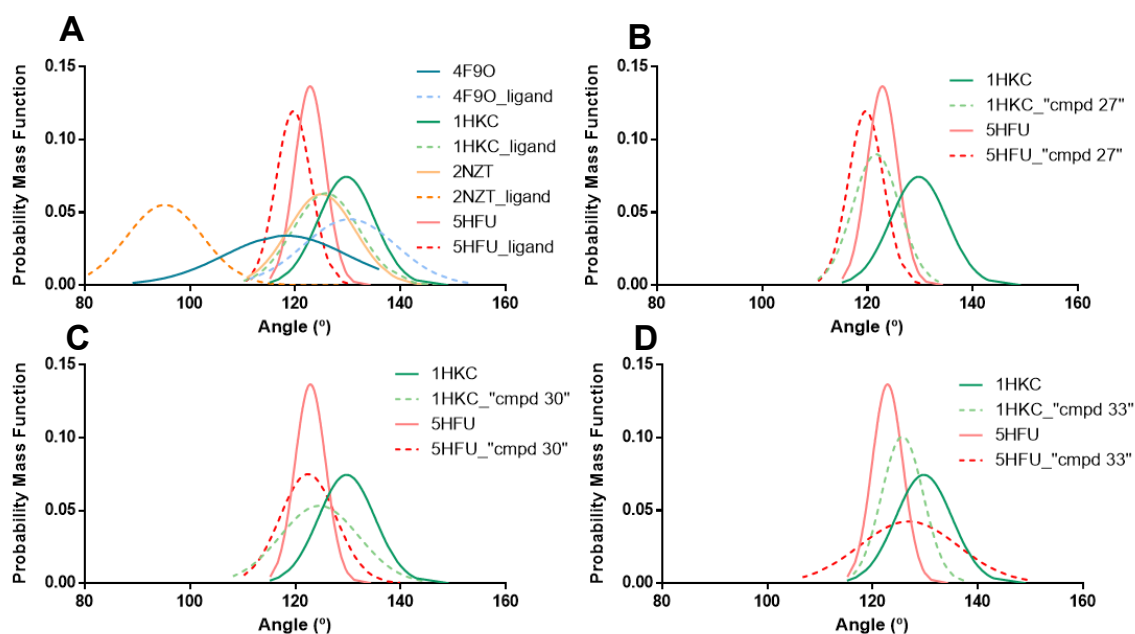


Figure 95. Probability Mass Function of the Thr389-Arg468-Ala839angle throughout the simulation for all systems, comparing the absence with the presence of different ligands/inhibitors: **A** - Crystallographic ligands; **B** - "Cmpd 27"; **C** - "Cmpd 30"; **D** - "Cmpd 33".

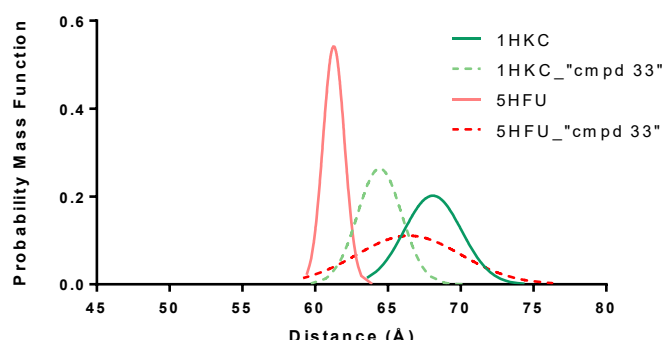


Figure 96. Probability Mass Function of the distance between Res389 and Res839 of systems 1HKC and 5HFU in absence or in presence of the inhibitor “cmpd33”.

2.3.11. Influence of the 618,K→T “mutation”

In section 2.2, the static analysis of HK1 and HK2 structures suggested the 618,K→T “mutation” (HK2→HK1) to be a probable player in the selectivity for certain inhibitors. In HK1, Thr618 interacts by H-bonding with Lys738 effectively closing the catalytic pocket, while in HK2 Lys618 does not interact with Lys738 (**Figure 80**). The distance between residue 618 and Lys738 was measured in order to evaluate the impact of the 618,K→T “mutation” throughout the simulations. **Figure 97** shows the respective probability mass function of this distance for all systems. Together, the information obtained seems to lead to the conclusion that a different Res618 does not justify the difference of selectivity among the distinct inhibitors. A pattern could not be identified and the distribution of the distance between Res618-Lys738 for the most selective inhibitor, “cmpd 33”, is almost identical in HK1 and HK2 (**Figure 97 - D**). Moreover, potent inhibitors of both HK1 and HK2 (“cmpd 27” and “cmpd 30”) tend to cause different changes in the distance distribution. While “cmpd 27” considerably induced lower distances between the residues in HK2 and higher distances in HK1 (similar to HK2 without any ligand), “cmpd 30” did not change the distance in HK2 and slightly increased that distance in HK1 (**Figure 97 - B and C**).

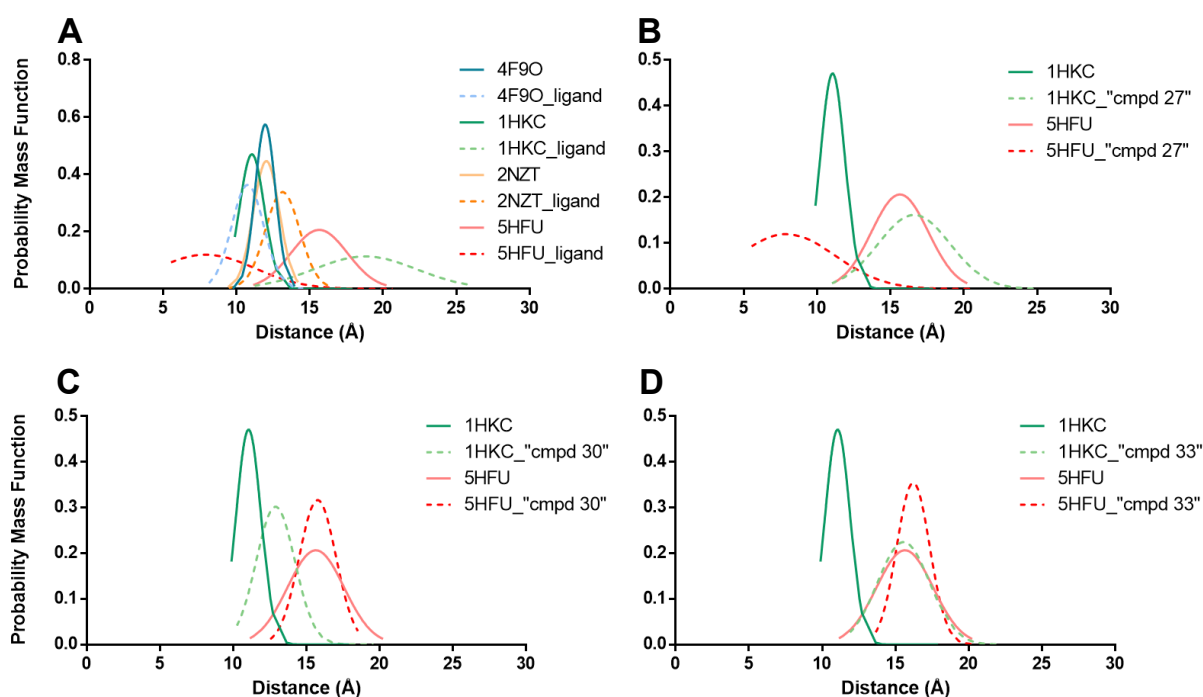


Figure 97. Probability mass function for the distance between Res618-Lys738 for all systems. Each plot compares ligand-free systems with systems displaying crystallographic ligands (A), with inhibitor "cmpd 27" (B), with inhibitor "cmpd 30" (C) and with inhibitor "cmpd 33" (D).

2.3.12. Comparing pocket opening in HK1 and HK2

As previously reported [271], the distance between C α of Res88 (small subdomain, Ser88 in HK1, Thr88 in HK2) and Thr232 (large subdomain) was used as an indicator of the opening of the NTD pocket. Likewise, the distance between the corresponding residues in CTD (Thr536 and Thr680) was used to measure the opening of the CTD active site (**Figure 98**).

Recalling previous information, the 1HKC crystallographic structure displays an NTD in the closed state, just like 4F9O and 2NZT, but the CTD presents an open state pocket. 5HFU possesses both pockets in the open state. Throughout the simulation, distances between the residues referred above are the reflection of these initial structures. The probability mass function of the Res88-Thr232 distance for 2NZT, 4F9O, and 1HKC in the presence or absence of crystallographic ligands is very similar, while 5HFU, either in presence or absence of its crystallographic ligand, displays much larger distances for the same pair of residues (**Figure 99 – A**). On the CTD pocket, this difference is much smaller. 4F9O and 2NZT present approximately the same distribution throughout the simulation in presence or absence of crystallographic ligands, whereas 1HKC and 5HFU share a slightly larger Thr536-Thr680 distance. In this pocket, in the presence of the crystallographic ligand (in the NTD), the Thr536-T680 distance in 1HKC is considerably longer than in the ligand-free system (**Figure 99 – B**). From these data, it is possible to conclude that the 200 ns simulation

might not be sufficiently long to observe similar opening of the glucose/G6P pockets, especially in the NTD case, where 5HFU shows a wider opening of the pocket compared with the other structures and the CTD pocket. It might be assumed that the opening of the CTD pocket might be limited by any non-detected constrain.

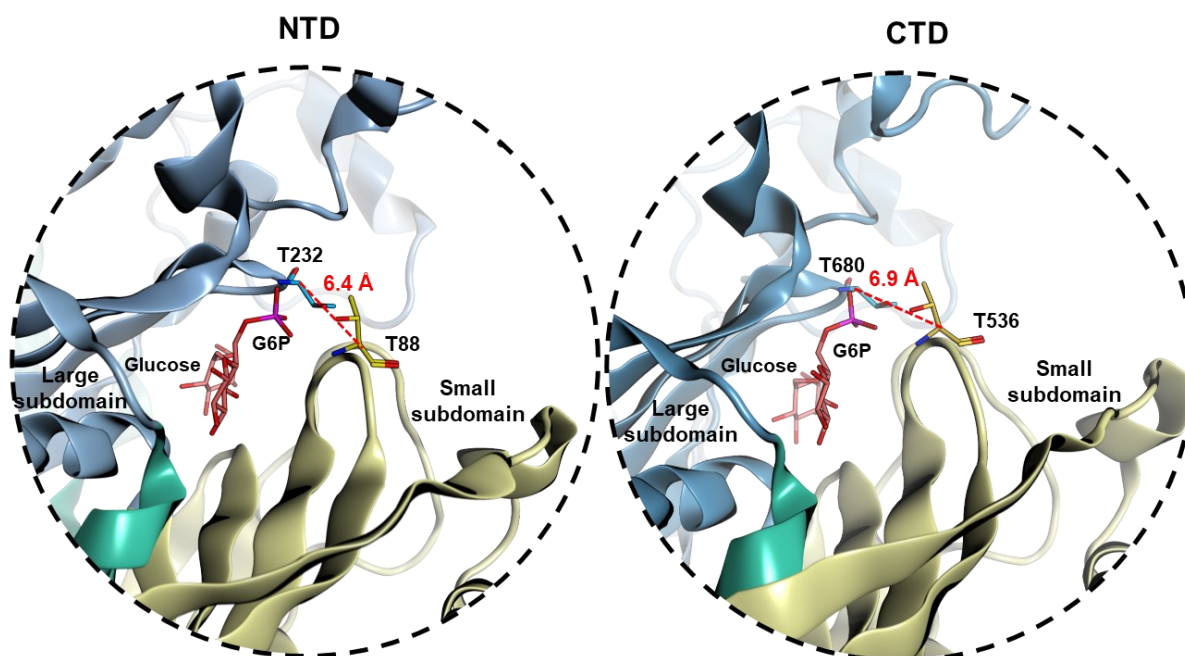


Figure 98. Positioning of residues Thr88 and Thr232 on the NTD pocket and Thr536 and Thr680 on the CTD pocket and respective distances (Å), measured between the Ca atoms on the crystallographic structure PDBID: 2NZT [73].

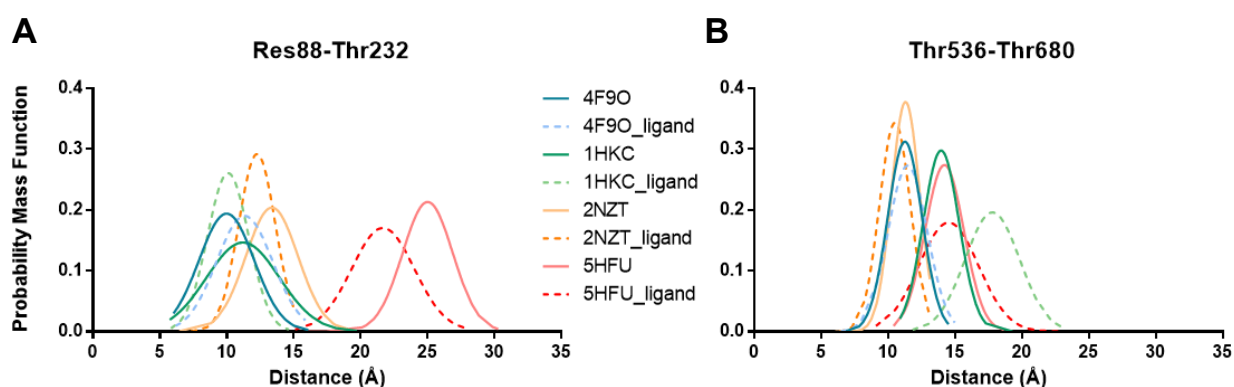


Figure 99. Probability mass function of the distance between Res88/536 and Thr232/680 in the absence or presence of crystallographic ligands for the four structures.

Figure 100 shows the probability mass function for Res88-Thr232 and Thr536-Thr680 distances in HK1 (1HKC) and HK2 (5HFU) systems in the absence or presence of inhibitors. In the presence of inhibitors “cmpd 27” and “cmpd 30”, the Res88-Thr232 distance in 1HKC was kept in the same range than in the system without ligands. Regarding the inhibitor “cmpd33”, the presence of this molecule increased the distance in analysis. As for 5HFU, the presence of inhibitors led to a smaller most probable distance.

This decrease in the distance was mostly detected in the presence of “cmpd 30”. With “cmpd 33”, the most probable distances between Res88-Thr232 are closer in the two isozyms, halfway between those of the systems 1HKC and 5HFU without any ligand (**Figure 100**).

Considering the distance between Thr536-T680 in 5HFU, the distribution almost did not change from the system without ligands. For 1HKC, a slight increase compared with the system without ligands was found with all inhibitors, more prominent with “cmpd 33”.

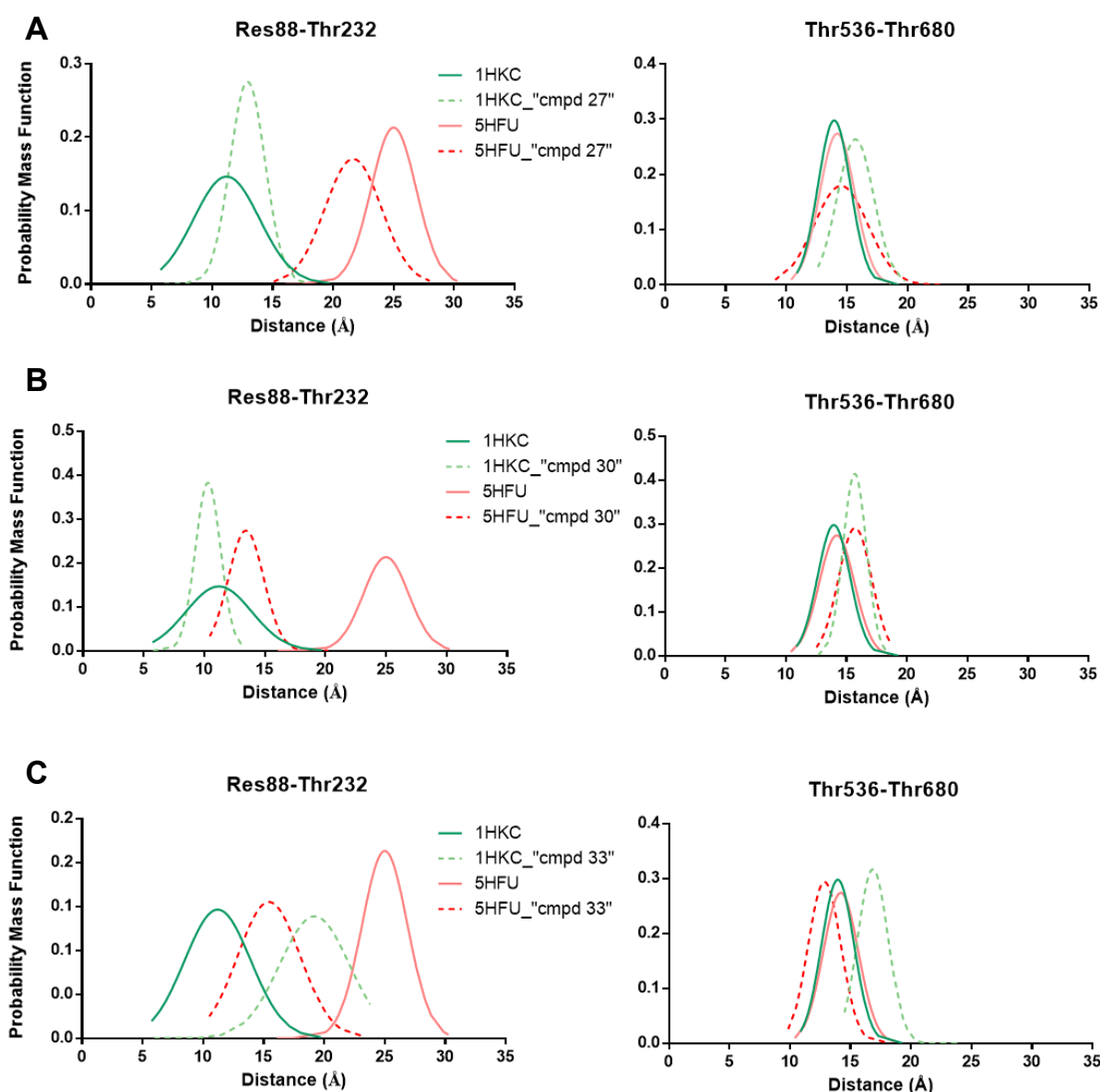


Figure 100. Probability Mass Function of the distance between Res88/536 and Thr232/680 of HK1 (1HKC) and HK2 (5HFU) in the absence and presence of inhibitors “cmpd 27”, “cmpd 30”, “cmpd 33”.

2.3.13. Notes from the trajectory visualization

Upon visualization of the 200 ns trajectories of each simulated system with VMD software, several tendencies of movement were observed, comparing with the starting structure (**Table 17**). Opening and closure movement of the two catalytic pockets is notorious; however it is more prominent or smooth depending on each system. Supporting the RMSF results, movement of the NTD pocket seems to be more significant than that of the CTD in the majority of systems.

Table 17. Observations on systems trajectories regarding the position of the ligand, and the behavior of the NTD and CTD pockets.

Isozyme	State	Ligand	PDBID	Observations		
				Ligand	NTD pocket	CTD pocket
HK1	Closed	-	4F9O	-	Leads to a slight opening of the pocket;	Leads to pocket opening and closure.
	Closed	Cryst. ¹	4F9O	Does not change its position. Binds in the cleft between the large and the small subdomain;	Does not change significantly;	Closes overtime.
	Open ²	-	1HKC	-	Leads to pocket opening and closure;	Closes overtime.
	Open ²	Cryst. ¹	1HKC	Does not change its position significantly; Binds in the cleft between the large and the small subdomain;	Does not change significantly;	Leads to pocket opening and closure.
	Open ²	"Cmpd 27"	1HKC	NTD - Does not change its position. Stays between the large and the small subdomain; CTD - Binds deeper in the cleft between the large and the small subdomain;	Leads to pocket opening;	Does not change considerably.
	Open ²	"Cmpd 30"	1HKC	Does not change its position. Stays between the large and the small subdomain;	Leads to pocket opening;	Slightly closes overtime.
	Open ²	"Cmpd 33"	1HKC	Slightly changes its position; Binds in the cleft between the large and the small subdomain;	Leads to pocket opening;	Closes after 10 ns.
HK2	Closed	-	2NZT	-	Leads to pocket opening;	Leads to a slight opening of the pocket.
	Closed	Cryst. ¹	2NZT	Does not change its position. Binds in the cleft between the large and the small subdomain;	Opens and then closes; Leads to a slight opening of the pocket;	Leads to a slight opening of the pocket.
	Open	-	5HFU	-	Leads to a slight opening of the pocket;	Slightly closes overtime
	Open	"Cmpd 27" ³	5HFU	Changes considerably its position. Binds in the cleft between the large and the small subdomain;	Leads to pocket opening, and then closure;	Slightly closes overtime
	Open	"Cmpd 30"	5HFU	Changes considerably its position; Binds do both subdomains in both pockets;	Opening/closure of the pocket is blocked;	Slightly closes overtime
	Open	"Cmpd 33"	5HFU	Binds to the large subdomain in both pockets;	Closes overtime;	Does not change considerably.

¹Cryst. – Crystallographic ligands; ²Open – 1HKC structure has only the CTD pocket in the open state; ³Cmpd 27" – "cmpd 27" is the crystallographic ligand of 5HFU.

Distribution of the Res88-232 (NTD) and Res536-680 (CTD) distance was not indicative of the evolution of the opening/closure process of the systems. This distance was used to measure the pocket opening in Ferreira et al. (2020) [271]. However, those residues are placed deep in the cleft between the two subdomains, and cannot be used to properly describe the movement of the two subdomains. Analysis of more than one distance in three dimensional systems is required for an accurate description of such conformational changes. Toward this goal, other distances between residues placed further away from

the active site were chosen for the analysis, supporting prior observations. In the end, distances between three residues from the large subdomain (Res232[NTD]/680[CTD], Res286/734, Res335/783) and three from the small subdomain (Res88/536, Res157/605, Res116/564) of each pocket were measured throughout the simulation (**Figure 101**). **Figure 102** exemplifies as those distances help to rationalize certain observations, in this case, relative to 1HKC in presence of “cmpd 33” (green). Together, distances explored in plots **A**, **B** and **C**, tendentially increase overtime, showing the opening of the NTD pocket. Regarding the CTD pocket, plots **D**, **E** and **F** show a slight decrease in the distances analyzed, indicating closure of the pocket. These observations relative to the ligands can be rationalized using the ligands RMSD previously covered in **section 2.3.6**. and the protein-ligand interaction fingerprint (PLIF) further covered in **section 2.3.15**.

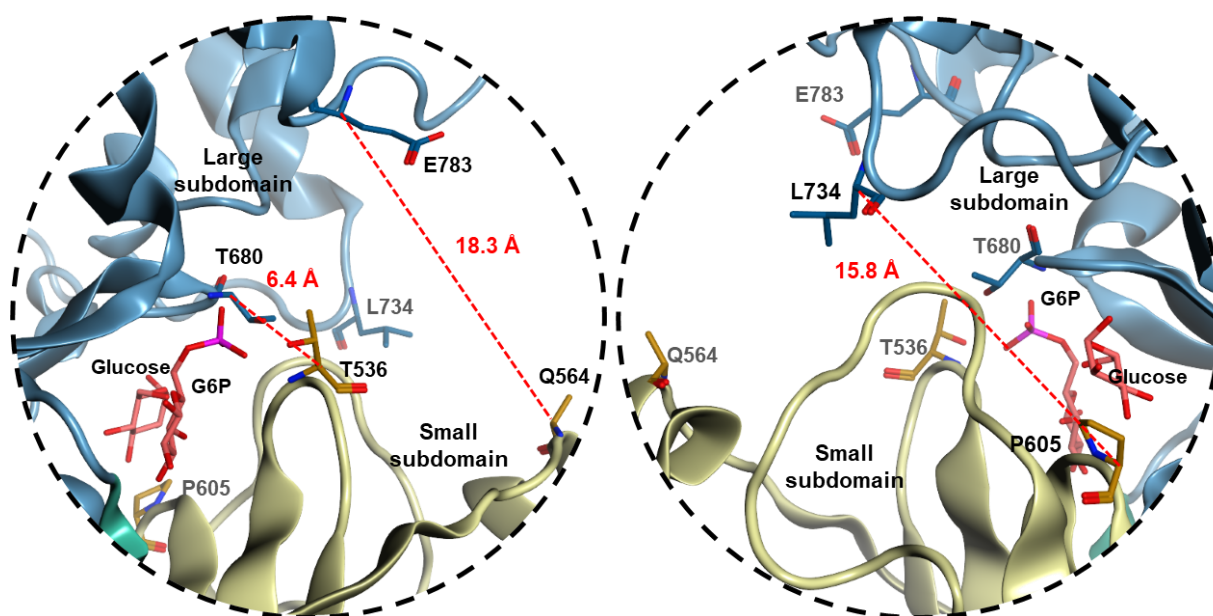


Figure 101. Positioning of residues T536, P605 and Q564 in the small subdomain and T680, L734 and E783 in the large subdomain of the CTD pocket, and respective distances (Å, in red), measured between Cα in the CTD pocket of the crystallographic structure PDBID: 2NZT [73].

In general, the NTD of HK1 in the presence of inhibitors opens or keeps approximately the same opening throughout the simulation, while in HK2 this pocket closes (**Figure 102** – pink). Interestingly, despite of the decrease pocket width, in presence of inhibitors “cmpd 27” and “cmpd 30”, the pocket of HK2 (5HFU) is slightly wider or in the same range of HK1 (1HKC), while in presence of “cmpd 33” 5HFU the trend is to have a pocket much closed than the presented by 1HKC.

The HK2 CTD pocket usually did not change its amplitude considerably (**Figure 102** – pink). However, prior measurements, such as RMSD calculations, indicate a large change in the CTD pocket in 5HFU with “cmpd 33”. Careful observation of the CTD (large and small subdomains) indicates that the presence of the inhibitor in different conformations (**Figure 103**) greatly impacts the overall structure, without having a significant impact in the pocket opening. In fact, a backbone torsion seems to be induced that might destabilize the entire protein structure. This torsion does not occur in the HK1 system or with any other inhibitor. The NTD pocket is not as deeply affected, since the inhibitor binds differently

and does not cause the same distortion. **Figure 104** compares the initial structure and the conformation obtained after the 200 ns simulation and shows the distortion caused in the structure without a significant change in the opening amplitude of the pocket

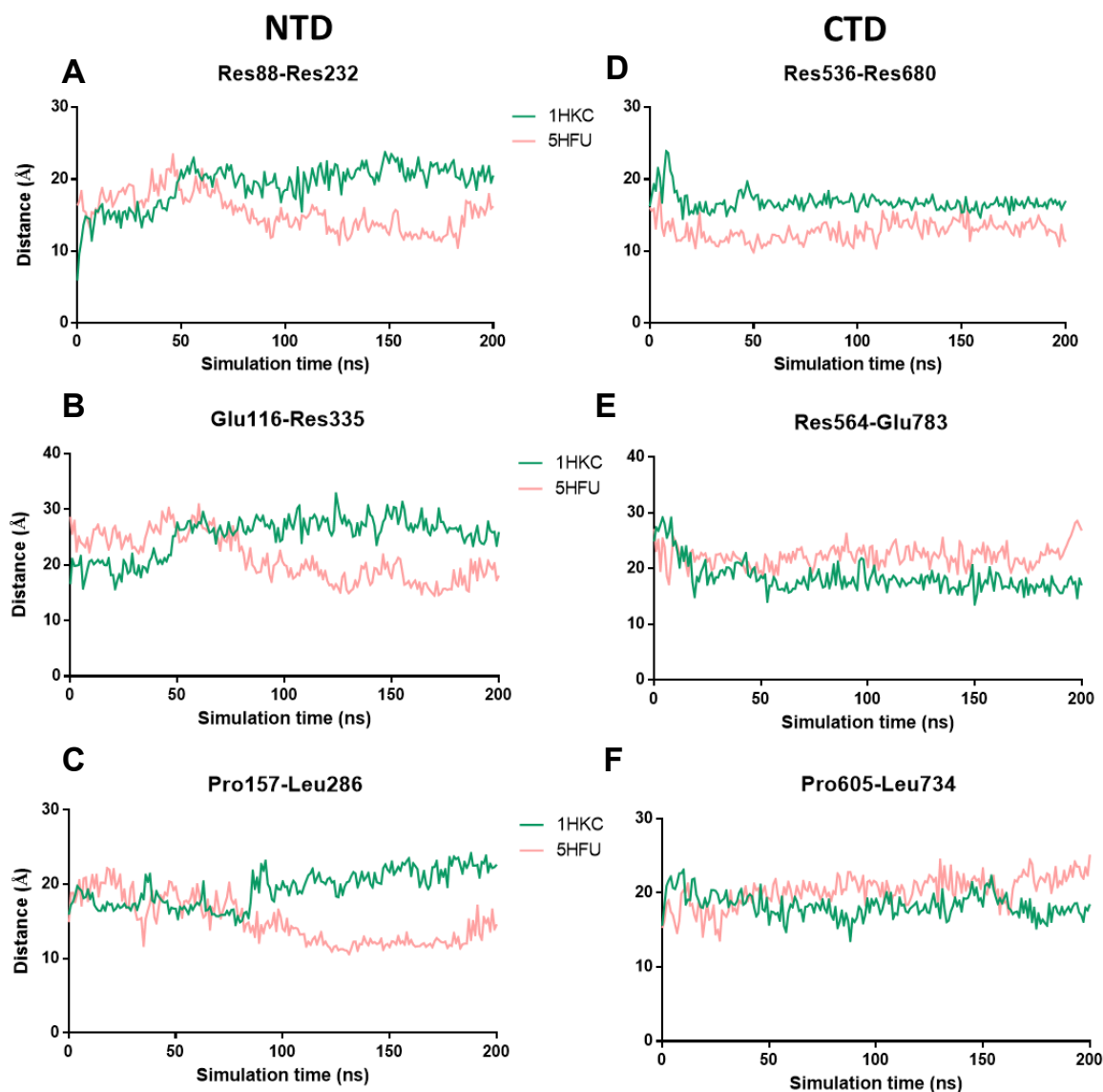


Figure 102. Plots of distances between key residues in NTD (**A**, **B** and **C**) and CTD (**D**, **E**, **F**) pockets of HK1 (PDBID: 1HKC, green) and HK2 (PDBID:5HFU, pink) in presence of “cmpd 33” throughout the simulation time. Distance was measured between Ser88 and Thr232 (**A**), Glu116 and Asn335 (**B**), Pro157 and Leu 286 (**C**), Thr536 and Thr680 (**D**), Ile564 and Glu 768 (**E**) and Pro605-Leu734 (**F**).

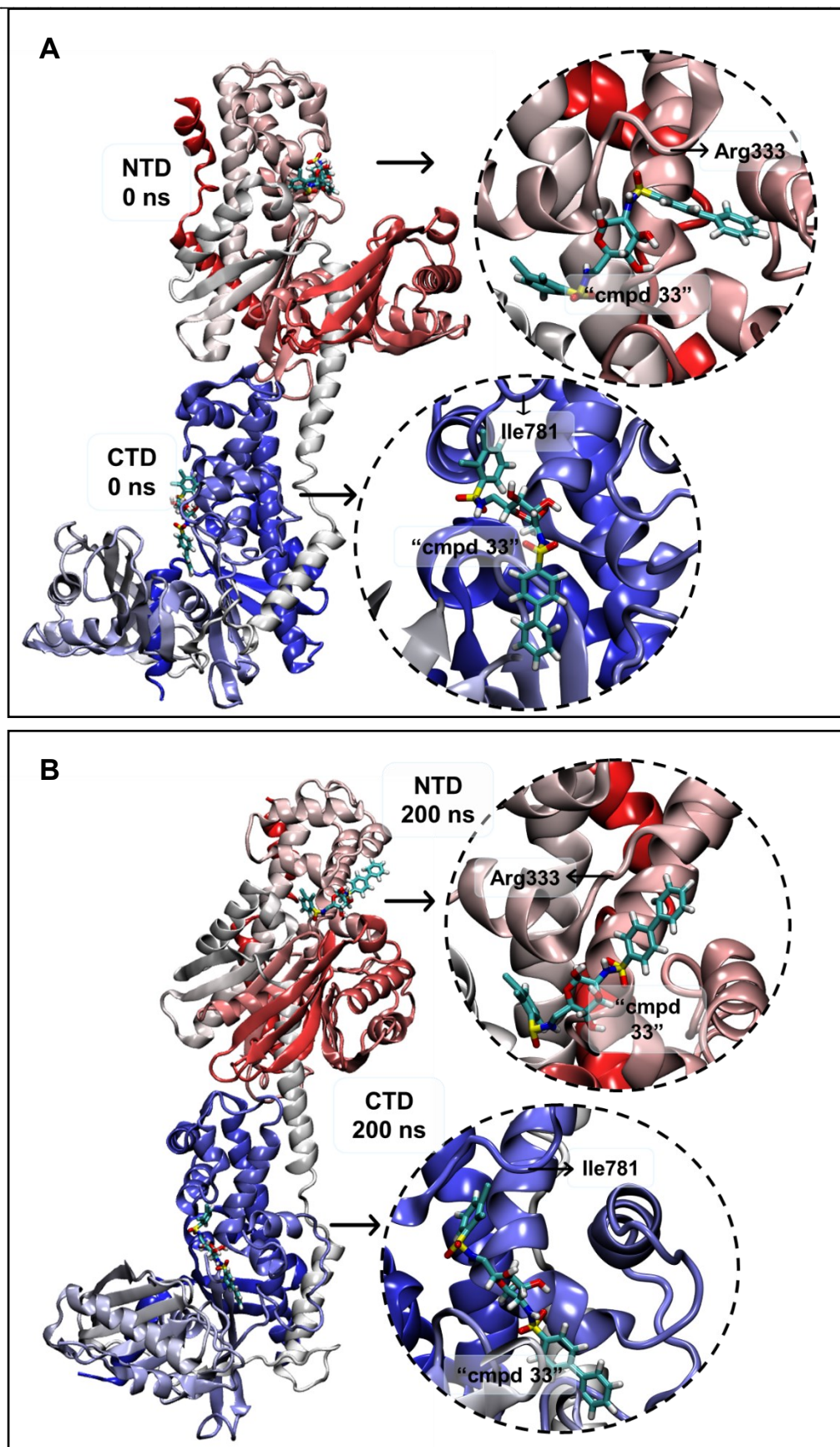


Figure 103. Positioning of the inhibitor "cmpd 33" in the NTD and CTD pockets of HK2 at 0 ns (initial position) (A) and at 200 ns (end of simulation) (B) (system 5HFU + "cmpd 33"). While "cmpd 33" in the NTD pocket aligns the hydrophobic moiety with the helix of the large subdomain, the CTD hydrophobic moiety of "cmpd 33" goes more deeply into the pocket in direction of the small

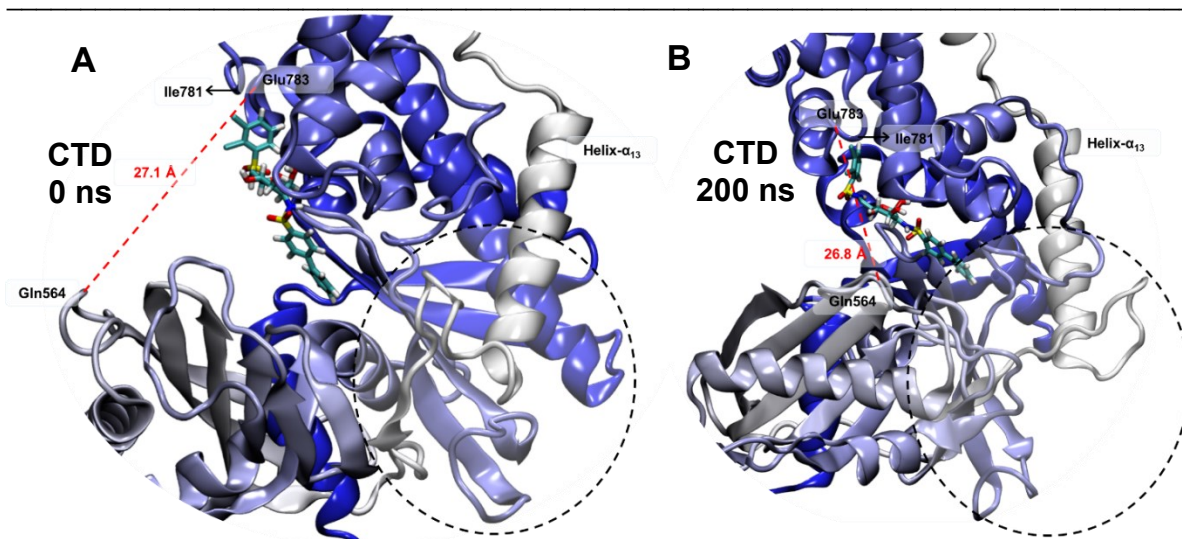


Figure 104. Initial structure at 0 ns (A) and the distortion caused by “cmpd 33” after 200 ns of simulation in the CTD pocket of HK2 (system 5HFU + “cmpd 33”) (B). The distance between Gln564 and Glu783 does not change significantly, but the subdomain is globally affected.

2.3.14. Principal component analysis

Principal component analysis (PCA) was performed (using the GROMACS modules gmx covar and gmx anaeig) to determine the variance among the set of conformations generated throughout the simulation. Scatter plots of eigenvector 1 vs. eigenvector 2 projections show the configuration space explored in the simulation defined by the selected set of modes.

Figure 105 A and B present the PCA for the systems involving HK1 (1HKC) and HK2 (5HFU), respectively. Overall, HK1 conformations seem to be more closely related to each other, since highly correlated conformations cluster together. The same is true for HK2 without any ligand and with “cmpd 27” and “cmpd 30”. However, projecting data points for “cmpd 33” are spread all over the plot, indicating that this structure is highly affected and different conformations are found throughout the simulation. This is in agreement with the previous findings in this MD study.

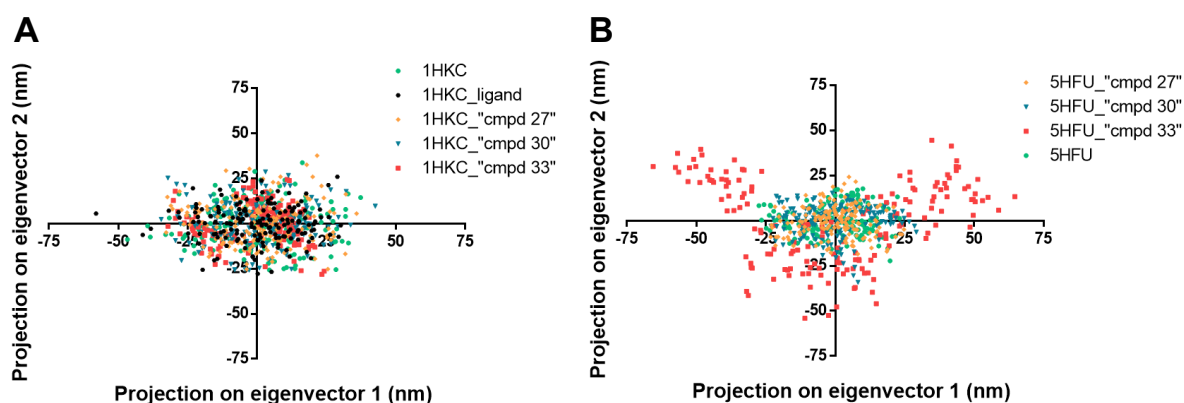


Figure 105. PCA of simulations with HK1 (1HKC) (A) and HK2 systems (5HFU) (B).

Figure 106 specifically shows the PCA for 5HFU alone (**A**) and 5HFU with “cmpd 33” (**B**) to evince the differences. The MD trajectory seems to sample multiple states of the protein and approximately 3 clusters can be defined (C1, C2 and C3), which agrees with the number of clusters found using an cutoff an RMSD of 0.4 nm as cutoff (**section 2.3.5**).

Interestingly, HK1 with cmpd 33 also behaved differently from HK1 alone (**Figure 106 - C and D**). Two defined clusters (C1 and C2) were found with “cmpd 33” (**C**), suggesting that the MD trajectory sampled two states of the protein, although C2 seems to be more populated.

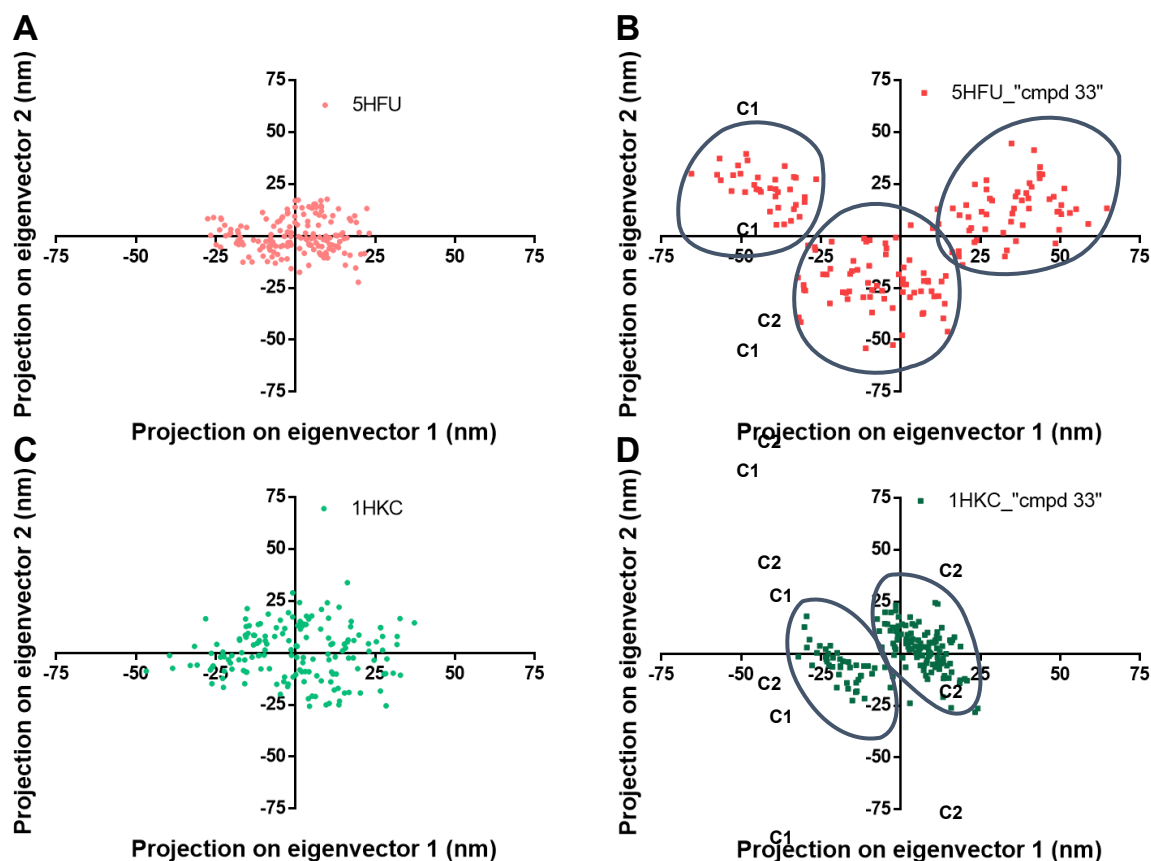


Figure 106. PCA of simulations with: 5HFU alone (**A**), 5HFU with “cmpd33” (**B**), 1HKC alone (**C**) and 1HKC with “cmpd 33” (**D**) systems.

The RMSF derived from the covariance analysis allowed the identification of the regions responsible for the majority of changes in the protein conformations. In **Figure 107**, most changing regions of HK1, when in the presence of “cmpd 33”, are placed in the large subdomain of the NTD. Not unexpectedly, the most affected region is the *N*-terminal helix, one end of the protein. On the other hand, changes in HK2 when in presence of “cmpd 33” were more evenly distributed, but with more weight on the CTD part (**Figure 108**). These results confirm that the residues that fluctuate the most are not in the catalytic pocket itself but rather in its vicinity, and that no changes in the pocket opening were observed. More exact causes for this destabilization should be further investigated.

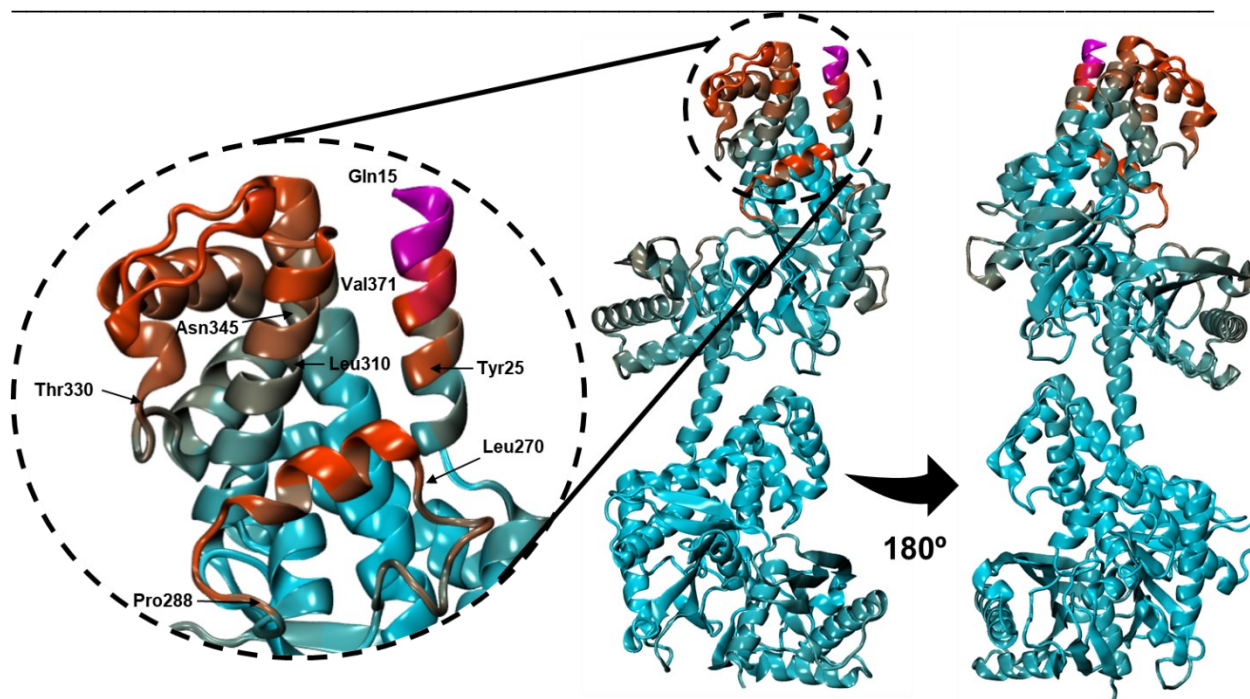


Figure 107. Structure of HK1 (1HKC) colored according to the PCA-derived RMSF. Blue means low RMSF and red/pink represents high RMSF. Regions with high RMSF are defined with the starting and ending residue: Gln15-Tyr25; Leu 270-Pro288; Leu310-Thr330; Asn345-Val371.

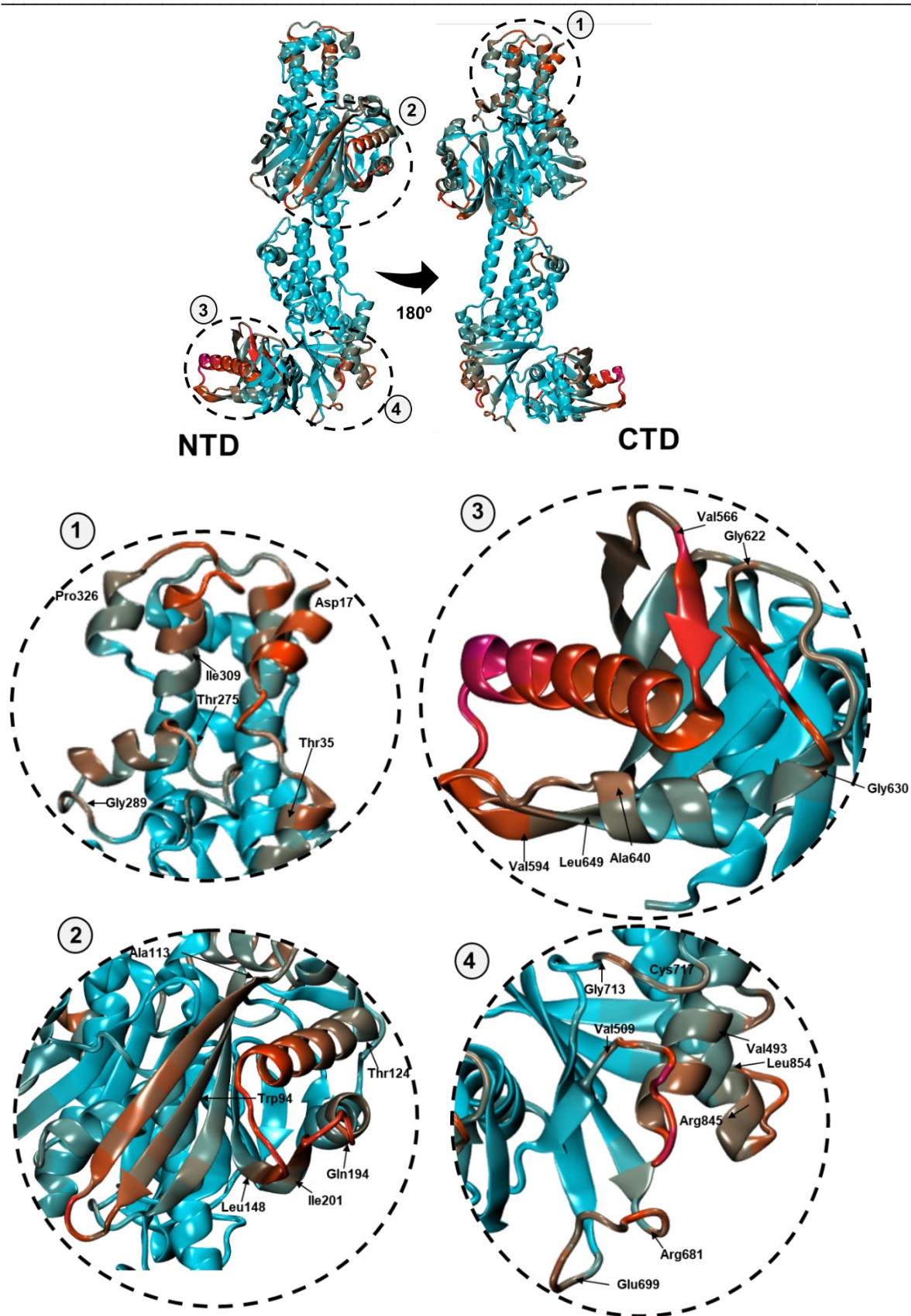


Figure 108. Structure of HK2 (5HFU) colored according to the PCA-derived RMSF. Blue means low RMSF and red/pink represents high RMSF regions. Regions with high RMSF are defined with the starting and ending residue (e.g.: Asp17-Thr35).

2.3.15. Protein-ligand interaction fingerprints (PLIFs) over time

Protein-ligand interaction fingerprints (PLIFs) were determined throughout the simulation [272] for the three inhibitors “cmpd 27”, “cmpd 30 and cmpd 33” in both HK1 and HK2, in the NTD and CTD. In this section the most common interactions with each ligand in each domain are compared, with greater focus on the “cmpd33” interactions in both isozymes. Complete PLIF plots (0-200 ns) for each system are shown on annex 5.

2.3.15.1. “Cmpd27”

HK1

During the simulation, HK1 NTD interacted more frequently with “cmpd 27” through residues Lys173 (salt bridge, H-bond), Arg174 (H-bond, hydrophobic, salt bridge), Asp209 (H-bond, hydrophobic), Ile229 (hydrophobic), Gly231 (H-bond), Thr232 (H-bond), Gly233 (H-bond), Thr234 (H-bond), Asn235 (H-bond), and Glu260 (H-bond, hydrophobic).

In the CTD, the most common contacts of “cmpd 27” were with residues Leu512 (hydrophobic), Cys606 (H-bond), Asn656 (H-bond), Asp657 (H-bond, hydrophobic), Ile677 (hydrophobic), and Glu708 (H-bond).

In common, NTD and CTD contacted with “cmpd 27” very often through Glu260/708, Ile229/677 and Asp209/Asp657.

HK2

In the NTD “cmpd27” interacted mostly with residues Leu64 (hydrophobic, H-bond), Pro65 (hydrophobic, H-bond), Ser155 (H-bond), Phe156 (H-bond), Cys158 (H-bond), Thr172 (H-bond), Lys173 (H-bond, salt bridge, π -cation), Asn208 (H-bond), Asp209 (H-bond), Thr210 (H-bond, hydrophobic), Asn235 (H-bond, hydrophobic), Val248 (hydrophobic), and Asn258 (hydrophobic).

The CTD residues that interacted most with “cmpd27” were Pro508 (hydrophobic, H-bond), Lys510 (H-bond, π -cation, hydrophobic), Cys606 (H-bond), Gln607 (H-bond, hydrophobic), Gln608 (H-bond, hydrophobic), Lys621 (H-bond), Asn656 (H-bond, hydrophobic), Asn683 (H-bond), Glu708 (H-bond), and Gln739 (H-bond).

The common interactions found in both NTD and CTD were through residues Cys158/606, Lys173/Lys621, Asn208/656, and Asn235/683, suggesting a similar binding mode in the two domains.

HK1 and HK2 with “cmpd 27”

During the simulations, none of the residues that interacted with “cmpd27” were the same in the two isozymes HK1 and HK2 in both domains. However, Lys173 and Asn235 of both isozymes maintained contact with “cmpd27” in the NTD, as did residues Cys606, Asn656 and Glu708 in the CTD.

2.3.15.2. “Cmpd30”

HK1

In the NTD the most frequent interactions occurred with residues Asp84 (H-bond), Arg91 (H-bond, π -cation, salt bridge), Ser155 (H-bond), Lys173 (H-bond, salt bridge, π -cation), Asn208 (H-bond), Asp209 (H-bond), Thr210 (hydrophobic), Ile229 (hydrophobic), Asn235 (H-bond, hydrophobic), Asp413 (H-bond), Gly414 (H-bond), Lys418 (H-bond, salt bridge), and Ser449 (H-bond).

“Cmpd 30” interactions were very consistent from the beginning until the end of the simulation in both NTD and CTD.

In the CTD, the most common contacts occurred with residues Lys510 (hydrophobic), Leu512 (hydrophobic), Ser603 (H-bond), Pro605 (hydrophobic), Cys606 (H-bond), Ile617 (hydrophobic), Lys621 (H-bond, salt bridge), Asn656 (H-bond), Asp657 (H-bond), Thr658 (H-bond), Ile677 (hydrophobic), Thr680 (H-bond), Asn683 (H-bond), Asn706 (H-bond), Glu708 (H-bond, hydrophobic) and Thr863 (H-bond).

Several contacts were found in both domains (Ser155/630, Lys173/621, Asn208/656, Asp209/657, Thr210/658, Ile229/677 and Asn235/683), indicating a similar binding mode in the two domains.

HK2

When compared with the CTD, fewer consistent interactions between “cmpd30” and HK2 were found in the NTD: Lys173 (H-bond, hydrophobic), Phe175 (hydrophobic, π -stacking), Met302 (H-bond), Gly303 (H-bond), Glu304 (H-bond, salt bridge) and Arg333 (H-bond, π -cation, salt bridge).

In the CTD, main interactions between “cmpd30” and HK2 are mediated by residues Phe602 (H-bond, hydrophobic), Ser603 (H-bond), Lys621 (H-bond, hydrophobic, salt bridge, π -cation), Asn656 (H-bond, hydrophobic), Asn683 (H-bond), Asn706 (H-bond), Glu708 (H-bond), Ser733 (H-bond), Leu734 (H-bond), Asn735 (H-bond) and Gln739 (H-bond).

Except for Lys173/621, none of the interacting residues interacting with toward “cmpd 30”, in one domain matched with the other.

HK1 and HK2 with “cmpd 30”

Throughout the simulations, Lys173/621 was the only residue that interacted with the two isozymes in the two domains. Other residues interacted with “cmpd30” in both isozymes in the CTD (Ser603, Asn706 and Glu708) but not in the NTD.

2.3.15.3. “Cmpd33”

HK1 - NTD

At the beginning of the simulation “cmpd 33” interacted with residues Asp84, Ser88, Arg91, Asp209, Ile229, Thr232, Asp413, Gly414, Ser415, Lys418 and Ser449. At 200 ns, “cmpd 33” maintained the contact with Asp209, Gly414 and Ser415, besides new contacts with Leu64, Lys173, Arg 174, Thr232, Gly233, Asn235 and Glu260.

In summary, the most common interactions throughout the simulation were to residues Asn208 (hydrophobic), Asp209 (H-bond, hydrophobic), Thr232 (H-bond), Gly233 (H-Bond), Asn235 (hydrophobic), Asp413 (H-bond, salt bridge), Gly414 (H-bond), and Lys418 (H-bond, hydrophobic, π -cation interaction).

Some contacts were lost throughout the simulation (Leu93, Arg91, Ser88, Pro157, Thr172, Ile230, Ser449, and others were formed (Leu64, Asn208, Asn258, Glu260).

HK1- CTD

Many interactions with “cmpd 33” were formed after the first nanoseconds. At 0 ns, “cmpd 33” had contact with Gln608, Val654, Asn656, Gly681, Ser682, Asn683, Glu708, Gln739 and Glu742. By the end of the simulation interactions with Gln608, Asn656, Asn683 and Glu708 were kept, and contacts with Thr620, Lys621, Asp657, Thr658, Thr661, Asn706, Asp861, Gly862 and Thr863 were formed.

The most common and consistent interactions were found with Leu512 (hydrophobic), Phe602 (hydrophobic), Ser603 (H-bond), Pro605 (hydrophobic), Gln608 (hydrophobic), Thr620 (H-bond), Lys621 (H-bond, π -cation interaction), Val654 (hydrophobic), Asn656 (hydrophobic), Asp657 (hydrophobic, H-bond, salt bridge), Thr658 (hydrophobic), Thr661(hydrophobic), Asn683 (H-bond), Asn706 (H-bond, hydrophobic), Glu708 (H-bond, halogen bond), Asp861 (H-bond), Gly862 (H-bond), Thr863 (H-bond, hydrophobic).

HK1 - “cmpd33” key results:

Interactions in the CTD were more consistent throughout the simulation, with many residues maintaining the contact with the inhibitor when compared with the NTD.

Regarding the most frequent contacts, "cmpd33" interacted with 6 common residues in both NTD and CTD (Lys173/621, Asn208/656, Asp209/657, Asn235/683, Asp413/861, Gly414/862), revealing similar contacts in both domains.

HK2 - NTD

At the beginning of the simulations "cmpd 33" interacted with Ile281, Gly299, Tyr301, Gly303, Glu304, Glu335, Thr336, Lys337, Val416 and Lys419. At 200 ns, interactions with residues Glu304 (halogen bond), Glu335, Lys337, Val416 were maintained, while interactions with Ser415 and Thr232 arose.

The most consistent interactions between "cmpd 33" and HK2 are through the residue Thr232 (H-bond, hydrophobic), Glu304 (salt bridge), Glu335 (H-bond, hydrophobic, salt bridge), Thr336 (H-bond, hydrophobic), Lys337 (H-bond, hydrophobic, salt bridge), Ser415 (H-bond), Val416 (hydrophobic) and Lys419 (hydrophobic).

Interaction with Gly299 was found very often at the start, however this contact was lost after 56 ns of simulation.

HK2 - CTD

At 0 ns, "cmpd 33" interacted with Thr784 (hydrophobic), Glu783 (hydrophobic), Tyr749 (H-bond), Met748 (H-bond), Asn683 (hydrophobic), Thr680 (hydrophobic) and Ile677 (hydrophobic). By the end of the simulation (200ns), interactions with Thr784, Tyr479 and Met748 were kept and new interactions with Lys621 (hydrophobic, H-bond), Glu742 (salt bridge) and Ser746 (H-bond).

In the C-terminal domain the most common interactions throughout the simulation were through Thr620 (hydrophobic), Lys621 (hydrophobic, H-bond, π -cation), Glu742 (hydrophobic, H-bond, salt bridge), Met748 (H-bond), Tyr749 (H-bond), Glu783 (hydrophobic, salt bridge) and Thr784 (hydrophobic, H-bond).

Some interactions were significantly different in the last 50 ns of simulation (151-200) compared to the first 50 ns. For instance, Thr863, which did not sustainably participate in any interaction throughout the first nanoseconds of simulation, interacted via H-bond in almost every nanosecond of the last 50 (41/50). Moreover, "cmpd 33" only started to interact with Ala711 (hydrophobic) in the last 80 ns.

By contrast, interaction with Thr620 (hydrophobic), although frequent in three fourths of the simulation, only occurred in 10 of the 50 last nanoseconds. Also, contact with Pro605 which appeared often in the first 50 ns, was lost throughout the simulation.

Interestingly, some hydrophobic interactions were found mostly from 40 to 110 ns and then lost (Thr680, Leu787, Leu864).

HK2 - “cmpd33” key results:

“Cmpd 33” was not found to interact with key residues for catalysis Asp209 or Asp 657 at any time. Throughout the simulation, only two residues commonly interact with “cmpd 33” in both NTD and CTD (Glu335/783; Thr336/784); however, those interactions were not with the same moieties. In the NTD Thr336 is facing the glucose ring and Glu335 is closer to the biphenyl moiety. In the CTD the portion facing those residues is the phenyl substituent.

In agreement with the visual inspection of the simulation, the remaining interactions in NTD are mostly with residues of the large subdomain (Ser415, Val416, Lys419), while most interactions of “cmpd 33” in the CTD occurred with residues of the cleft in the opposite side of the domain.

HK1 and HK2 with “cmpd 33”

None of the most frequent contacts were found in both HK1 and HK2 in both domains, however regular contact with Lys173/621 occurred in the CTD of both HK1 and HK2 and in the NTD of HK1. Thr232 had contact with both HK1 and HK2 in the NTD, but not with the corresponding residue in the CTD (Thr680). On the other hand, contact with Thr620 and Lys621 was frequent in both HK1 and HK2 but not with the corresponding residue (Thr172 and Lys173) in NTD.

In summary, despite the initial spatial position being similar (molecular docking conditions to obtain the initial position had the same center [Glu742] and volume for the conformation search [15 Å]), “cmpd 33” interacts with quite different residues in both enzymes. Moreover “cmpd 33” interacted with residues substantially different in each domain of HK2.

2.3.15.4. General key messages

As anticipated, most interactions occurred via H-bonding or salt bridge, although hydrophobic contacts were also observed.

The PLIF plots analysis indicated Lys173/621 as a key residue for the inhibition with all three test compounds in both enzymes (exception systems: HK2 NTD-2 “cmpd33” and HK1 CTD-“cmpd27”).

Asn208/656 and Asp209/Asp657 are suggested as key residues for the inhibition of HK1 and HK2 by “cmpd 30” and “cmpd 27” and for HK1 by “cmpd33”. Interestingly, these residues did not contact with “cmpd 33” in both domains in any of the 200 ns of simulation. The same was found for Asn235/683. In the CTD, contact of Glu708 with “cmpd 33” in HK2 was not frequent, but it was the main common interaction for the remaining systems (HK1- “cmpd33”, HK1 and HK2 with both “cmpd 30” and “cmpd 27” in both isozymes). Accordingly, these residues were already known to perform crucial interactions with glucose or G6P (*vide supra* **Figure 20**, Section 2.1, Chapter III).

As described above, two residues were shown to maintain contact with “cmpd 33” only in HK2 in both domains, including Glu335/783 that interacts with “cmpd 33” mostly via salt bridge. Noteworthy, in HK1 NTD the residue with opposite charge (at pH7.4) Arg335 is in place of Glu335. This change might affect the way the molecule interacts with the isozyme.

Sustained interactions with Asn335/Gly335/783 and Thr336/784 were not found in any of the other systems, neither were the ones previously found to interact with glucose or G6P (*vide supra* **Figure 20**, Section 2.1, Chapter III).

3. Conclusions

Static structural analysis and recent studies in the field [73,271] gave important insights about the resemblances and differences between HK1 and HK2. However, the few structural differences found did not indicate strong potential to be responsible for the changes in the activity of the two isozymes in presence of certain inhibitors. As MD has the potential to predict the behavior of enzymes in aqueous medium, MD studies were expected to give more information.

From the data here analyzed, HK1 and HK2, especially HK2 (5HFU), are mostly affected by “cmpd 33”. This compound is the most selective inhibitor towards HK2 and it seems to deeply affect the structure as RMSD, RMSF and PCA indicate (**section 2.3.3, 2.3.7., 2.3.14**), more prominently in the CTD. This is shown not only by CTD RMSD but also in the significative increase of “cmpd 33” SASA.

Residue 618, a Thr in HK1 but a Lys in HK2, was suspected to have a role in the difference between HK1 and HK2 activity. This was analyzed during the simulation, and although the results were not conclusive, this “mutation” dose not seem likely to cause the difference.

Analysis of PLIF plots revealed that interactions of “cmpd 33” with Gly335/783 and Thr336/784 were exclusive to HK2 and no contact was made with the key residue Asp209/Asp657 which might indicate a unique inhibition mechanism driven by “cmpd33”, that should be explored.

Other simulations with “cmpd 33” using different starting systems (different protein structure or different docking poses) should be performed to understand if the drastic change in the structure was an artifact or if indeed the inhibitor is able to deeply affect the structure of the protein. Once the effect of “cmpd 33” is confirmed, the mentioned destabilization should be investigated.

4. Experimental Section

4.1. Structural analysis

Structure analysis, comparison studies and distance measurements were performed with Molecular Operating Environment 2016 (MOE2016) [232] software. Three-dimensional structures of HKs PDBID: 4F9O [72], 2NZT [73], 5HEX [43], 1BG3 [74], 1BDG [74], 3B8A [75], 4QS7 [76] were downloaded from Protein Data Bank (PDB) [78]. Global structures, catalytic pockets (or binding pocket for *N*-terminal pocket of HK1) were compared, as well as “neighbouring residues”. The catalytic pocket comprehends all the residues containing at least one atom within 4.5 Å of glucose and G6P atoms; any residue containing at least one atom within 4.5 Å of any atom of the catalytic pocket is considered a “near residue”. For the analyses, measurements, alignments and superpositions, water was removed and only chain A was considered for all structures. Metal ions as Na⁺, Mg²⁺ and K⁺ were removed, since they did not seem to interact with the catalytic pocket. Root mean square deviations (RMSDs) were calculated according to the deviation of C α atoms.

4.2. Molecular Dynamics Simulations

As for structural analysis, three-dimensional structures of HKs PDBID: 4F9O [72], 2NZT [73], 5HFU [43], 1HKC [84] were downloaded from Protein Data Bank (PDB) [78] and corrected using MOE2016 [232] software. Structure clashes, inconsistencies and residue naming issues were resolved, and protonation at pH 7.4 was performed. Protein and residue charges (300 K, pH 7.4) were recorded. Structures of exogenous ligands for each system were obtained from the first ranking result of molecular docking using GOLD 5.2 software according to the settings described in chapter III, section 4.2. Ligand topologies were generated using Automated Topology Builder (ATB) [273]. The MD GROMACS v2016.4 package was used for the steps of MD simulation preparation, running and analysis, unless mentioned otherwise. After generating .gro files containing the atomic system, octahedral boxes large enough to avoid interaction with its periodic image were created and systems were solvated following a simple point charge (SPC) water model. Charge neutralization (using the necessary number of Na⁺ ions) and energy minimization (1000 steps steepest descent algorithm) to remove the unreasonable atomic contacts and stereochemical conflicts) were performed before a 1 ns temperature equilibration under the canonical ensemble (NVT) and a 1 ns pressure equilibration under the isothermal-isobaric ensemble (NPT) (300 K, 1 atm). MD simulations were carried out at 300 K and 1 atm., using the GROMOS 54A7 force field [274] for 200 ns. A cut-off of 1.0 nm was used to compute the short-range electrostatic and van der Waals interactions. The Particle Mesh Ewald method [275,276] was used to calculate the long-range electrostatic interactions.

After rotation removal, translation removal, and recentering, results (structures and trajectories) were visualized using Visual Molecular Dynamics (VMD) software [277]. PLIF plots were obtained from the

PDB simulation files with 201 snapshots (corresponding to the 200 ns of simulation + initial structure) using the Python module protein–ligand interaction profiler (PLIP) [272]. The PLIP parameters and considered interactions (hydrophobic, π -stacking, π -cation, H-bonds, halogen bonds, and salt bridges) were previously described [272].

The Kruskal-Wallis test was performed to evaluate the statistical significance of the differences between distributions (probability mass functions; **Figure 92**) using GraphPad Prism® 6.0 (GraphPad Software Inc., La Jolla, CA, USA).

Chapter VII

Final remarks and future work

1. Translation of HK2 inhibition to therapeutics

Inhibition of HK2 by itself is expected to not be enough to produce a full therapeutic effect, and combination therapy with other agents as part of a targeted therapy strategy is the most likely scenario for the use of HK2 inhibitors.

Advances in cancer research allowed the development of several therapeutical strategies in recent years, such as cell therapy or molecular targeted therapy. At present, more personalized targeted therapeutic options, according to specific features of the disease (e.g., metabolic profile, affected tissues, or stage) have yet to meet their full potential, both in terms of intrinsic efficacy and cost-benefit considerations [278]. Molecular targeted therapy works by inhibiting specific biomolecular entities that are essential for cancer progression and particularly expressed in the cancer cells of interest. This implies genome sequencing of patient cancer cells for biomarker determination [279], to be as specific as possible. Moreover, achievement of better diagnosis and follow-up procedures are thus fundamental approaches towards faster, more cost-effective, and successful therapeutic options. With the view to overcome cancer cells resistance, which is the greatest challenge for an effective cancer treatment, combination therapy involving targeted therapy and chemotherapy or radiotherapy is seen as the most promising strategy [7,8]. In this regard, targeted therapy envisioned to block specific pathways in cancer metabolism is a potentially safe and efficient strategy [10].

As previously mentioned, HK2 has been shown to be a consistent biomarker in several cancers, whose many reports have arisen recently. If efficient HK2 inhibitors were found, probably many patients would benefit from targeted therapy. Unfortunately, eligibility for targeted therapy is still very low (~15% of US patients in 2018 [280]). Thus, contributing to the enlargement of the pool of targetable biomarkers will enhance the coverage of more patients with this therapeutic strategic, thereby increasing the rates of success [279].

2. Alternative strategies

Finding novel HK2 inhibitors was the major goal driving this research project. Many obstacles were anticipated, including the poor druggability of the target and the selectivity issues against the desired isozyme. A hydrophilic catalytic pocket, extremely flexible, requires a precise balance between structural features that could improve the binding affinity, along with the potential to prevent an active conformation of the macromolecule, and others that could maintain adequate pharmacokinetic properties, namely the permeability across biological membranes. Another, more unlikely, hypothesis could be the use of known transporters that might facilitate the translocation of the molecule.

In this project different paths could be taken. In the next subsections alternative methodologies and directions are suggested.

2.1. Alternatives to the SBVS approach

For a single and fixed conformation to be considered a proper representation of the protein, the system must be extensively rigid, to correspond perfectly to the “lock-and-key” concept of binding [281]. The flexibility of the catalytic pocket, which accommodates glucose and ATP, was possibly one of the biggest barriers to the success of SBVS, a problem that may apply to most kinases. As a picture of a certain conformation in a certain moment, crystallographic structures and subsequently molecular docking do not consider the flexibility of entire subdomains that could limit the binding affinity of potential ligands. In contrast to other SBVS studies reported in the literature [153,155], instead of the closed conformation of HK2 (PDBID: 2NZT [73]), a more open conformation (PDBID: 5HFU [43]) was explored in this work and this is likely to have influenced the results. This difference might be explored in the future, as an attempt to evaluate the importance of the free space available for docking. Notably, our results showed that known potent inhibitors do not fit inside the closed state and thus would never have been highly scored. Moreover, different additional databases might be included. As an example, Dark Chemical Matter databases would be interesting choices to avoid off-target interactions and increase the novelty on the hit molecules potentially found [282]. Another approach to improve the results would be the selection process of PHK2Is. First, filtering databases based on druglike properties could be applied before the molecular docking. This could save a lot of calculation time; nevertheless, the parameters used for cutoff must be carefully analyzed to not exclude molecules untimely. Furthermore, the virtual selection based on the best poses/interactions could be improved using the knowledge acquired so far to prioritize, for instance, molecules with a central hexose moiety or another 6-membered ring substituted with H-bond donors/acceptors.

In addition to those alternative strategies, a different SBDD technique might be applied instead of SBVS, specifically, fragment-based *de novo* design (FBND). FBND might be applied to guarantee the interaction with key residues, using different building blocks to achieve new molecules with potential ability to interfere with HK2 activity. This might increase the specificity of the molecule to the target.

2.2. Alternative strategies to obtain more promising test compounds

The SBVS strongly indicated that ribose derivatives would be HK2 inhibitors. Thus, because SBVS showed that all compounds eventually selected for testing were expected to be able to interfere with HK2 activity to some extent, ribose derivatives were synthesized. The experimental results

ultimately showed that the effort to synthesize those compounds did not add any value to the set of PHK2Is already reported. Thus, synthesis efforts directed to improve the potency of molecules already found to be active would be a better strategy. Preparation of a set of derivatives of a known potent inhibitor to provide comprehensive SAR data could improve the quality of the overall project.

2.3. Alternatives to the biological evaluation approaches

The major aim of the biological evaluation of PHK2Is and known inhibitors was to understand how these molecules affect the proliferation of cancer cells and confirm their activity against the desired pathway and macromolecular target. Other approaches than the ones illustrated in Chapter V might be applied. For instance, if resources were not an issue and HTS could be an option, PHK2Is could be evaluated directly as glycolytic inhibitors against the different CHO cell lines. In this way, confirmation of the effect on the desired target and the selectivity against HK2 could be elucidated simultaneously.

Following the approach chosen for this project, additional studies could be carried out in the future, such as the calculation of IC_{50} for glycolysis inhibition for different cell lines, in order to measure properly the effect of PHK2Is. Moreover, improvement of certain methods would lead to more insightful results. For instance, as previously mentioned, accumulation of labeled G6P in the metabolomic studies was hard to achieve. To overcome this issue, using labeled 2DG or FDG, which are metabolized by HKs but not by enzymes in subsequent pathways, might lead to detectable amounts of labeled phosphate derivatives. This could allow a better perception of the effect of small molecules on the inhibition of the desired target.

PHK2I **10** has shown an ability to inhibit glycolysis; however, significant inhibition of cell proliferation during a period of seven days was barely achieved. As the effect on cell proliferation might not be immediate, treating cells with PHK2I **10** for a longer period (*e.g.*, 2 weeks) could be a better strategy to evaluate this molecule.

Promising results were found for PHK2Is **6**, **7**, **10** and **60**. Therefore they should be further probed. To evaluate potential metabolic obstacles to the use of these molecules, metabolism studies are envisioned. For instance, incubation with human and/or rodent liver S9 fraction (9000g supernatant of a liver homogenate) associated with MS can provide insights into the metabolic stability of the compounds and the potential phase 1 and phase 2 metabolic reactions [283].

2.4. Alternatives in the MD study

MD studies can provide important mechanistic insights into biological processes. The NTD pocket of HK1 has been proven to regulate the enzyme activity. This mechanism could be further explored by

simulation of HK1 and HK2 isozymes in the presence of ligands in one pocket at a time to investigate the potential differences in the overall structure. The presence of endogenous molecules might be included.

Herein, the behavior of HK2 in the presence of the selective inhibitor “cmpd 33” was found to be very particular; however, the exact mechanism of the change in the structure was not fully understood and more studies should be carried out, such as the simulation of other docking solutions. Moreover, simulation of other known inhibitors such “cmpd 34”, one the most selective inhibitors toward HK2, could provide more insights in the behavior of HK1 and HK2

3. Discussion on the effect of PHK2I **10** and **60** in glycolysis and their physicochemical and pharmacokinetic properties

Regarding HK2 inhibition, PHK2I **10** was the most promising compound discovered in this project. However, several concerns are associated with this molecule. This molecule has a PAINS substructure related with the azo function, violates 3 out of the 4 components of Lipinski's rule of 5, has one of the highest TPSA (301 Å²) among the compounds selected for study, and contains two (considered low-risk) free sulfonic acids. Other substructures, such as the 8-aminonaphtol, can also be a reason for worry related with the reactivity or metabolism of the molecule (**Figure 109 –B**). Surprisingly, this molecule was able to affect glycolysis without inducing cell death (48h) or significantly reducing cell proliferation (6 days), showing no signs of toxicity or induction of differences in cell morphology. Still, a long-term evaluation on cell proliferation is advisable. Interestingly, a substantial difference in pH between control and PHK2I **10**-treated cells (10 µM), specifically a decreased pH in the extracellular medium, was found, which suggests an unexpected cellular response to the molecule. This molecule is a dark reddish-brown powder, pink in diluted solutions, which may be considered an azo dye family member with a structure quite similar to well-known dyes such as Acid Red 33 and Trypan blue (**Figure 110**).

PHK2I **60** also showed a meaningful inhibition of glycolysis. However, this molecule displayed a strong ability to reduce proliferation and increased cell death in cancer and non-cancer cells. Previously predicted ADMET properties unveiled several problems with this molecule, namely predicted poor solubility, the presence of a large number of rotatable bonds and, most importantly, a high-risk ketone with a halogen substituted alpha carbon, available for alkylation.

These molecules were procured willfully, based on their positioning on the HK2 pocket (**Figure 109**). Moreover, their availability from a free-of-charge source also impacted the decision to acquire and test PHK2Is **10** and **60**.

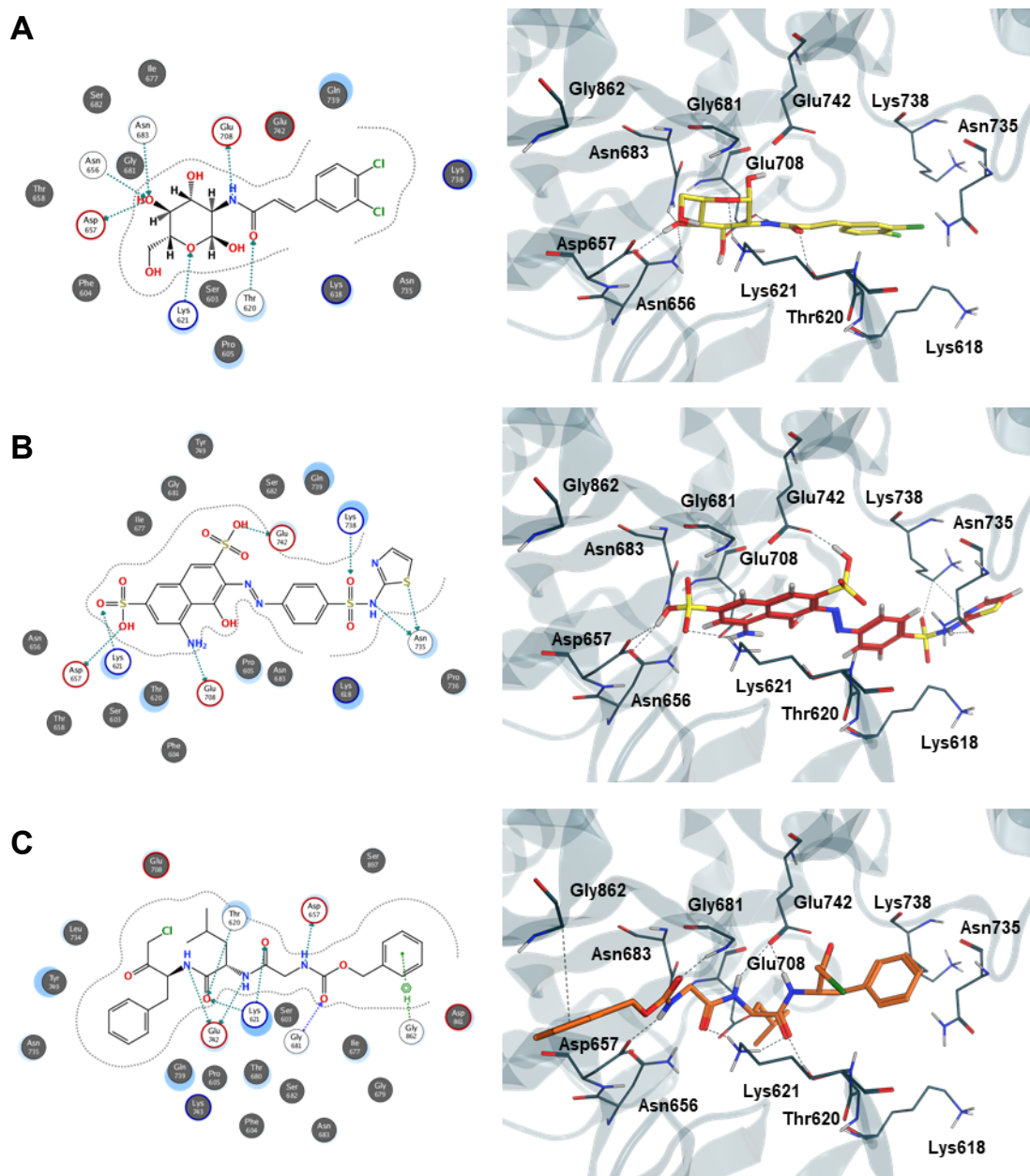


Figure 109. Crystallographic pose of “*cmpd 1*” [43] (A) and molecular docking-generated pose of PHK2I 10 (B) and PHK2I 60 (C) into the HK2 catalytic pocket in 2D (left) and 3D (right).

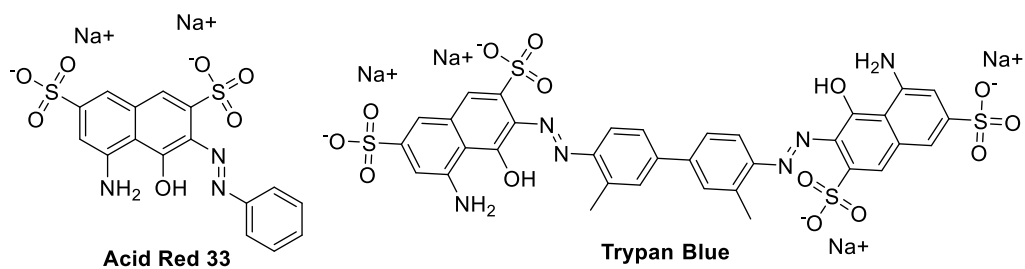


Figure 110. Structures of Acid Red 33 and trypan Blue azo dyes.

4. Concluding remarks

In summary, an SBVS strategy was implemented in this project. After virtual hit selection and biological evaluation, PHK2Is **6**, **7**, **10**, and **60** were found to inhibit glycolysis in several cell lines. However, the molecular target of these molecules was not unequivocally confirmed to be HK2, as dubious results from biochemical tests were obtained. Future work is required to clarify their mechanism of action and metabolic profile. Derivatives of these molecules might be designed and synthesized in order to obtain improved activity and drug-like properties.

Up to now, little information was available regarding HK2 targeting. From a medicinal chemistry perspective, this project allowed starting an exploratory approach to HK2. Structural, biochemical and biological insights in the field can provide tools to improve the approach followed herein and drive better decisions to stimulate the development of new potent and selective HK2 inhibitors, with the ability to be effective in combination therapy for cancer treatment.

Chapter VIII

References

- [1] Christopher, P. W.; Weiderpass, E.; Bernard, W. S. *World Cancer Report:T: Cancer Research for Cancer Prevention*; Lyon, France, **2020**.
- [2] Petrelli, A.; Giordano, S. From Single- to Multi-Target Drugs in Cancer Therapy: When Aspecificity Becomes an Advantage. *Curr. Med. Chem.* **2008**, *15* (5), 422–432.
- [3] Raghavendra, N. M.; Pingili, D.; Kadasi, S.; Mettu, A.; Prasad, S. V. U. M. Dual or Multi-Targeting Inhibitors: The next Generation Anticancer Agents. *Eur. J. Med. Chem.* **2018**, *143*, 1277–1300.
- [4] Ma, X.; Lv, X.; Zhang, J. Exploiting Polypharmacology for Improving Therapeutic Outcome of Kinase Inhibitors (KIs): An Update of Recent Medicinal Chemistry Efforts. *Eur. J. Med. Chem.* **2018**, *143*, 449–463.
- [5] Zhong, J.-T.; Zhou, S.-H. Warburg Effect, Hexokinase-II, and Radioresistance of Laryngeal Carcinoma. *Oncotarget* **2017**, *8* (8), 14133–14146.
- [6] Zhang, X. Y.; Zhang, M.; Cong, Q.; Zhang, M. X.; Zhang, M. Y.; Lu, Y. Y.; Xu, C. J. Hexokinase 2 Confers Resistance to Cisplatin in Ovarian Cancer Cells by Enhancing Cisplatin-Induced Autophagy. *Int. J. Biochem. Cell Biol.* **2018**, *95*, 9–16.
- [7] Zheng, H. C. The Molecular Mechanisms of Chemoresistance in Cancers. *Oncotarget* **2017**, *8* (35), 59950–59964.
- [8] Yeldag, G.; Rice, A.; Hernández, A. del R. Chemoresistance and the Self-Maintaining Tumor Microenvironment. *Cancers (Basel)*. **2018**, *10* (12).
- [9] Holohan, C.; Van Schaeybroeck, S.; Longley, D. B.; Johnston, P. G. Cancer Drug Resistance: An Evolving Paradigm. *Nat. Rev. Cancer* **2013**, *13* (10), 714–726.
- [10] Cairns, R. A.; Mak, T. W. The Current State of Cancer Metabolism. *Nat. Rev. Cancer* **2016**, *16* (10), 613–614.
- [11] Ma, L.; Zong, X. Metabolic Symbiosis in Chemoresistance: Refocusing the Role of Aerobic Glycolysis. *Front. Oncol.* **2020**, *10* (January), 1–8.
- [12] Zaal, E. A.; Berkers, C. R. The Influence of Metabolism on Drug Response in Cancer. *Front. Oncol.* **2018**, *8* (NOV), 1–15.
- [13] Hay, N. Reprogramming Glucose Metabolism in Cancer: Can It Be Exploited for Cancer Therapy? *Nat. Rev. Cancer* **2016**, *16* (10), 635–649.
- [14] Mathupala, S. P. S. P.; Ko, Y. H.; Pedersen, P. L. P. L. Hexokinase II: Cancer's Double-Edged Sword Acting as Both Facilitator and Gatekeeper of Malignancy When Bound to Mitochondria. *Oncogene* **2006**, *25* (34), 4777–4786.
- [15] Liu, Y.; Murray-Stewart, T.; Casero, R. A.; Kagiampakis, I.; Jin, L.; Jiawen, Z.; Huihui, W.; Che, Q.; Huan, T.; Ke, J.; et al. Targeting Hexokinase 2 Inhibition Promotes Radiosensitization in HPV16 E7-Induced Cervical Cancer and Suppresses Tumor Growth. *Int. J. Oncol.* **2017**, *50* (6), 2011–2023.
- [16] Shulga, N.; Wilson-smith, R.; Pastorino, J. G. Hexokinase II Detachment from the Mitochondria Potentiates Cisplatin Induced Cytotoxicity through a Caspase-2 Dependent Mechanism. *Cell Cycle* **2009**, *8* (20), 3355–3364.
- [17] Dai, W.; Wang, F.; Lu, J.; Xia, Y.; He, L.; Chen, K.; Li, J.; Li, S.; Liu, T.; Zheng, Y.; et al. By Reducing Hexokinase 2, Resveratrol Induces Apoptosis in HCC Cells Addicted to Aerobic Glycolysis and Inhibits Tumor Growth in Mice. *Oncotarget* **2015**, *6* (15), 13703–13717.
- [18] Patra, K. C.; Wang, Q.; Bhaskar, P. T.; Miller, L.; Wang, Z.; Wheaton, W.; Chandel, N.; Laakso, M.; Muller, W. J.; Allen, E. L.; et al. Hexokinase 2 Is Required for Tumor Initiation and Maintenance and Its Systemic Deletion Is Therapeutic in Mouse Models of Cancer. *Cancer Cell* **2013**, *24* (2), 213–228.
- [19] Garcia, S. N.; Guedes, R. C.; Marques, M. M. Unlocking the Potential of HK2 in Cancer Metabolism and Therapeutics. *Curr. Med. Chem.* **2019**, *26* (41), 7285–7322.
- [20] Hanahan, D.; Weinberg, R. A. Hallmarks of Cancer: The next Generation. *Cell* **2011**, *144* (5), 646–674.
- [21] Fouad, Y. A.; Aanei, C. Revisiting the Hallmarks of Cancer. *Am. J. Cancer Res.* **2017**, *7* (5), 1016–1036.
- [22] Amoedo, N. D.; Obre, E.; Rossignol, R. Drug Discovery Strategies in the Field of Tumor Energy Metabolism: Limitations by Metabolic Flexibility and Metabolic Resistance to Chemotherapy. *Biochim.*

-
- Biophys. Acta - Bioenerg.* **2017**, 1858 (8), 674–685.
- [23] Vander Heiden, M. G.; DeBerardinis, R. J. Understanding the Intersections between Metabolism and Cancer Biology. *Cell* **2017**, 168 (4), 657–669.
- [24] Tran, Q.; Lee, H.; Park, J.; Kim, S. H.; Park, J. Targeting Cancer Metabolism - Revisiting the Warburg Effects. *Toxicol. Res.* **2016**, 32 (3), 177–193.
- [25] Martinez-Outschoorn, U. E.; Peiris-Pagès, M.; Pestell, R. G.; Sotgia, F.; Lisanti, M. P. Cancer Metabolism : A Therapeutic Perspective. *Nat. Rev. Clin. Oncol.* **2017**, 14 (1), 11–31.
- [26] Han, J.; Zhang, L.; Guo, H.; Wysham, W. Z.; Roque, D. R.; Willson, A. K.; Sheng, X.; Zhou, C.; Bae-Jump, V. L. Glucose Promotes Cell Proliferation, Glucose Uptake and Invasion in Endometrial Cancer Cells via AMPK/MTOR/S6 and MAPK Signaling. *Gynecol. Oncol.* **2015**, 138 (3), 668–675.
- [27] Sun, L.; Yin, Y.; Clark, L. H.; Sun, W.; Sullivan, S. A.; Tran, A.-Q.; Han, J.; Zhang, L.; Guo, H.; Madugu, E.; et al. Dual Inhibition of Glycolysis and Glutaminolysis as a Therapeutic Strategy in the Treatment of Ovarian Cancer. *Oncotarget* **2017**, 8 (38), 63551–63561.
- [28] Warburg, O. On the Origin of Cancer Cells. *Science* **1956**, 123 (3191), 309–314.
- [29] Kroemer, G.; Pouyssegur, J. Tumor Cell Metabolism: Cancer's Achilles' Heel. *Cancer Cell* **2008**, 13 (6), 472–482.
- [30] Zaidi, H.; Karakatsanis, N. Nuclear Medicine : Physics Special Feature Review Article - Towards Enhanced PET Quantification in Clinical Oncology. *Br. J. Radiol.* **2018**, 91, 20170508.
- [31] Liberti, M. V.; Locasale, J. W. The Warburg Effect : How Does It Bene Fi t Cancer Cells ? *Trends Biochem. Sci.* **2016**, 41 (3), 211–218.
- [32] Nelson, D. L.; Cox, M. M. *Lehninger, Principles of Biochemistry*, 4th ed.; Freeman and Company: New York, **2005**.
- [33] Hu, J.; Locasale, J. W.; Bielas, J. H.; O'Sullivan, J.; Sheahan, K.; Cantley, L. C.; Heiden, M. G. V.; Vitkup, D. Heterogeneity of Tumor-Induced Gene Expression Changes in the Human Metabolic Network. *Nat. Biotechnol.* **2013**, 31 (6), 522–529.
- [34] Pathania, D.; Millard, M.; Neamati, N. Opportunities in Discovery and Delivery of Anticancer Drugs Targeting Mitochondria and Cancer Cell Metabolism. *Adv. Drug Deliv. Rev.* **2009**, 61 (14), 1250–1275.
- [35] Singh, D.; Arora, R.; Kaur, P.; Singh, B.; Mannan, R.; Arora, S. Overexpression of Hypoxia-Inducible Factor and Metabolic Pathways: Possible Targets of Cancer. *Cell Biosci.* **2017**, 7 (1), 1–9.
- [36] Cervantes-Madrid, D.; Dueñas-González, A. Antitumor Effects of a Drug Combination Targeting Glycolysis, Glutaminolysis and de Novo Synthesis of Fatty Acids. *Oncol. Rep.* **2015**, 34 (3), 1533–1542.
- [37] Dwarakanath, B.; Jain, V. Targeting Glucose Metabolism with 2-Deoxy- D -Glucose for Improving Cancer Therapy. *Futur. Oncol.* **2009**, 5 (5), 581–585.
- [38] Raez, L. E.; Papadopoulos, K.; Ricart, A. D.; Chiorean, E. G.; Dipaola, R. S.; Stein, M. N.; Rocha Lima, C. M.; Schlesselman, J. J.; Tolba, K.; Langmuir, V. K.; et al. A Phase i Dose-Escalation Trial of 2-Deoxy-d-Glucose Alone or Combined with Docetaxel in Patients with Advanced Solid Tumors. *Cancer Chemother. Pharmacol.* **2013**, 71 (2), 523–530.
- [39] Threshold Pharmaceuticals Inc. Press Release: Phase 2 and Phase 3 Clinical Trials of TH-070 in Benign Prostatic Hyperplasia (BPH) Do Not Meet Primary Endpoint. **2006**.
- [40] Business Wire: PreScience Closes on Institutional Round of Financing <https://www.businesswire.com/news/home/20170426005840/en/PreScience-Closes-Institutional-Financing> (accessed Mar 13, **2018**).
- [41] Chen, Z.; Zhang, H.; Lu, W.; Huang, P. Role of Mitochondria-Associated Hexokinase II in Cancer Cell Death Induced by 3-Bromopyruvate. *Biochim. Biophys. Acta - Bioenerg.* **2009**, 1787 (5), 553–560.
- [42] Ning, X.; Qi, H.; Li, R.; Jin, Y.; McNutt, M. A.; Yin, Y. Synthesis and Antitumor Activity of Novel 2,3-Dithiocarbamate Substituted Naphthoquinones as Inhibitors of Pyruvate Kinase M2 Isoform. *J. Enzyme Inhib. Med. Chem.* **2017**, 33 (1), 126–129.
- [43] Lin, H.; Zeng, J.; Xie, R.; Schulz, M. J.; Tedesco, R.; Qu, J.; Erhard, K. F.; Mack, J. F.; Raha, K.; Rendina, A. R.; et al. Discovery of a Novel 2,6-Disubstituted Glucosamine Series of Potent and Selective Hexokinase 2 Inhibitors. *ACS Med. Chem. Lett.* **2016**, 7 (3), 217–222.

-
- [44] Wilson, J. E. Hexokinases. *Rev. Physiol. Biochem. Pharmacol.* **1995**, 126, 65–198.
- [45] Pastorino, J. G.; Shulga, N.; Hoek, J. B. Mitochondrial Binding of Hexokinase II Inhibits Bax-Induced Cytochrome c Release and Apoptosis. *J. Biol. Chem.* **2002**, 277 (9), 7610–7618.
- [46] Arora, K. K.; Pedersen, P. L. Functional Significance of Mitochondrial Bound Hexokinase in Tumor Cell Metabolism. *J. Biol. Chem.* **1988**, 263 (33), 17422–17428.
- [47] Mazure, N. M. VDAC in Cancer. *BBA - Bioenerg.* **2017**, 1858 (8), 665–673.
- [48] Gall, J. M.; Wong, V.; Pimental, D. R.; Havasi, A.; Wang, Z.; Pastorino, J. G.; Bonegio, R. G. B.; Schwartz, J. H.; Borkan, S. C. Hexokinase Regulates Bax-Mediated Mitochondrial Membrane Injury Following Ischemic Stress. *Kidney Int.* **2011**, 79 (11), 1207–1216.
- [49] Pastorino, J. G.; Hoek, J. B. Hexokinase II: The Integration of Energy Metabolism and Control of Apoptosis. *Curr. Med. Chem.* **2003**, 10 (16), 1535–1551.
- [50] Krasnov, G. S.; Dmitriev, A. A.; Lakunina, V. A.; Kirpiy, A. A.; Kudryavtseva, A. V. Targeting VDAC-Bound Hexokinase II: A Promising Approach for Concomitant Anti-Cancer Therapy. *Expert Opin. Ther. Targets* **2013**, 17 (10), 1221–1233.
- [51] Conway, L. P.; Voglmeir, J. Functional Analysis of Anomeric Sugar Kinases. *Carbohydr. Res.* **2016**, 432, 23–30.
- [52] Cárdenas, M. L.; Cornish-Bowden, A.; Ureta, T. Evolution and Regulatory Role of the Hexokinases. *Biochim. Biophys. Acta - Mol. Cell Res.* **1998**, 1401 (3), 242–264.
- [53] Irwin, D. M.; Tan, H. Molecular Evolution of the Vertebrate Hexokinase Gene Family: Identification of a Conserved Fifth Vertebrate Hexokinase Gene. *Comp. Biochem. Physiol. - Part D Genomics Proteomics* **2008**, 3 (1), 96–107.
- [54] Ahn, K. J.; Kim, J.; Yun, M.; Park, J. H.; Lee, J. D. Enzymatic Properties of the N- and C-Terminal Halves of Human Hexokinase II. *BMB Rep.* **2009**, 42 (6), 350–355.
- [55] Bateman, A.; Martin, M. J.; O'Donovan, C.; Magrane, M.; Alpi, E.; Antunes, R.; Bely, B.; Bingley, M.; Bonilla, C.; Britto, R.; et al. UniProt: The Universal Protein Knowledgebase. *Nucleic Acids Res.* **2017**, 45, D158–D169.
- [56] Ureta, T. The Comparative Isozymology of Vertebrate Hexokinases. *Comp. Biochem. Physiol. -- Part B Biochem.* **1982**, 71 (4), 549–555.
- [57] Wilson, J. E. Isozymes of Mammalian Hexokinase: Structure, Subcellular Localization and Metabolic Function. *J. Exp. Biol.* **2003**, 206 (12), 2049–2057.
- [58] Neary, C. L.; Pastorino, J. G. Nucleocytoplasmic Shuttling of Hexokinase II in a Cancer Cell. *Biochem. Biophys. Resarch Commun.* **2010**, 394 (4), 1075–1081.
- [59] Neary, C. L.; Pastorino, J. G. Akt Inhibition Promotes Hexokinase 2 Redistribution and Glucose Uptake in Cancer Cells. *J. Cell. Physiol.* **2013**, 228 (9), 1943–1948.
- [60] Robey, R. B.; Hay, N. Mitochondrial Hexokinases, Novel Mediators of the Antiapoptotic Effects of Growth Factors and Akt. *Oncogene* **2006**, 25 (34), 4683–4696.
- [61] Tsai, H. J.; Wilson, J. E. Functional Organization of Mammalian Hexokinases: Characterization of the Rat Type III Isozyme and Its Chimeric Forms, Constructed with the N- and C-Terminal Halves of the Type I and Type II Isozymes. *Arch. Biochem. Biophys.* **1997**, 338 (2), 183–192.
- [62] Postic, C.; Shiota, M.; Magnuson, M. A. Cell-Specific Roles of Glucokinase in Glucose Homeostasis. *Recent Prog. Horm. Res.* **2001**, 56, 195–217.
- [63] Sternisha, S. M.; Liu, P.; Marshall, A. G.; Miller, B. G. Mechanistic Origins of Enzyme Activation in Human Glucokinase Variants Associated with Congenital Hyperinsulinism. *Biochemistry* **2018**, 57 (10), 1632–1639.
- [64] Chakera, A. J.; Steele, A. M.; Gloyn, A. L.; Shepherd, M. H.; Shields, B.; Ellard, S.; Hattersley, A. T. Recognition and Management of Individuals with Hyperglycemia Because of a Heterozygous Glucokinase Mutation. *Diabetes Care* **2015**, 38 (7), 1383–1392.
- [65] Haeusler, R. A.; Camastra, S.; Astiarraga, B.; Nannipieri, M.; Anselmino, M.; Ferrannini, E. Decreased Expression of Hepatic Glucokinase in Type 2 Diabetes. *Mol. Metab.* **2015**, 4 (3), 222–226.
- [66] Lei, L.; Liu, S.; Li, Y.; Song, H.; He, L.; Liu, Q.; Sun, S.; Li, Y.; Feng, Z.; Shen, Z. The Potential Role of

- Glucokinase Activator SHP289-04 in Anti-Diabetes and Hepatic Protection. *Eur. J. Pharmacol.* **2018**, 826, 17–23.
- [67] Matschinsky, F. M. Assessing the Potential of Glucokinase Activators in Diabetes Therapy. *Nat. Rev. Drug Discov.* **2009**, 8 (5), 399–416.
- [68] Cheruvallath, Z. S.; Gwaltney, S. L.; Sabat, M.; Tang, M.; Wang, H.; Jennings, A.; Hosfield, D.; Lee, B.; Wu, Y.; Halkowycz, P.; et al. Discovery of Potent and Orally Active 1,4-Disubstituted Indazoles as Novel Allosteric Glucokinase Activators. *Bioorganic Med. Chem. Lett.* **2017**, 27 (12), 2678–2682.
- [69] Toyoda, Y.; Tsuchida, A.; Iwami, E.; Shironoguchi, H.; Miwa, I. Regulation of Hepatic Glucose Metabolism by Translocation of Glucokinase between the Nucleus and the Cytoplasm in Hepatocytes. *Horm. Metab. Res.* **2001**, 33 (6), 329–336.
- [70] Ali, A.; Wathes, D. C.; Swali, A.; Burns, H.; Burns, S. A Novel Mammalian Glucokinase Exhibiting Exclusive Inorganic Polyphosphate Dependence in the Cell Nucleous. *Biochem. Biophys. Reports* **2017**, 12, 151–157.
- [71] Ronimus, R. S.; Morgan, H. W. Cloning and Biochemical Characterization of a Novel Mouse ADP-Dependent Glucokinase. *Biochem. Biophys. Res. Commun.* **2004**, 315 (3), 652–658.
- [72] Shen, L.; Gao, Y.; Honzatko, R. B. Inhibitor Sites of Unequal Affinity Linked by Binding Synergism in Mutant Forms of Recombinant Human Hexokinase Type-I. RSCB PDB - 4F9O, 4FOE, 4FPB. *Released on July 3rd 2013*.
- [73] Nawaz, M. H.; Ferreira, J. C.; Nedyalkova, L.; Zhu, H.; Carrasco-López, C.; Kirmizialtin, S.; Rabeh, W. M. The Catalytic Inactivation of the N-Half of Human Hexokinase 2 and Structural and Biochemical Characterization of Its Mitochondrial Conformation. *Biosci. Rep.* **2018**, 38 (1), 1–17.
- [74] Mulichak, A. M.; Wilson, J. E.; Padmanabhan, K.; Garavito, R. M. The Structure of Mammalian Hexokinase-1. *Nat. Struct. Biol.* **1998**, 5 (7), 555–560.
- [75] Kuser, P.; Cupri, F.; Bleicher, L.; Polikarpov, I. Crystal Structure of Yeast Hexokinase PI in Complex with Glucose: A Classical “Induced Fit” Example Revised. *Proteins Struct. Funct. Genet.* **2008**, 72 (2), 731–740.
- [76] Feng, J.; Zhao, S.; Chen, X.; Wang, W.; Dong, W.; Chen, J.; Shen, J. R.; Liu, L.; Kuang, T. Biochemical and Structural Study of Arabidopsis Hexokinase 1. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2015**, D71, 367–375.
- [77] Anderson, C. M.; Stenkamp, R. E.; Steitz, T. A. Sequencing a Protein by X-Ray Crystallography. *J. Mol. Biol.* **1978**, 123 (1), 15–33.
- [78] Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucleic Acids Res.* **2000**, 28 (1), 235–242.
- [79] Ardehali, H.; Yano, Y.; Printz, R. L.; Koch, S.; Whitesell, R. R.; May, J. M.; Granner, D. K. Functional Organization of Mammalian Hexokinase II - Retention of Catalytic and Regulatory Functions in Both the NH₂- and COOH-Terminal Halves. *J. Biol. Chem.* **1996**, 271 (4), 1849–1852.
- [80] Tsai, H. J.; Wilson, J. E. Functional Organization of Mammalian Hexokinases: Both N-and C-Terminal Halves of the Rat Type II Isozyme Possess Catalytic Sites. *Arch. Biochem. Biophys.* **1996**, 329 (1), 17–23.
- [81] Tsai, H. J. Functional Organization and Evolution of Mammalian Hexokinases: Mutations That Caused the Loss of Catalytic Activity in N-Terminal Halves of Type I and Type III Isozymes. *Arch. Biochem. Biophys.* **1999**, 369 (1), 149–156.
- [82] McWilliam, H.; Li, W.; Uludag, M.; Squizzato, S.; Park, Y. M.; Buso, N.; Cowley, A. P.; Lopez, R. Analysis Tool Web Services from the EMBL-EBI. *Nucleic Acids Res.* **2013**, 41, 597–600.
- [83] Steitz, T. A.; Shoham, M.; Bennett, W. S. Structural Dynamics of Yeast Hexokinase During Catalysis. *Philos. Trans. R. Soc. B Biol. Sci.* **1981**, 293 (1063), 43–52.
- [84] Aleshin, A. E.; Zeng, C.; Bartunik, H. D.; Fromm, H. J.; Honzatko, R. B. Regulation of Hexokinase I: Crystal Structure of Recombinant Human Brain Hexokinase Complexed with Glucose and Phosphate. *J. Mol. Biol.* **1998**, 282 (2), 345–357.
- [85] Petit, P.; Antoine, M.; Ferry, G.; Boutin, J. A.; Lagarde, A.; Gluais, L.; Vincentelli, R.; Vuillard, L. The Active Conformation of Human Glucokinase Is Not Altered by Allosteric Activators. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2011**, 67 (11), 929–935.

- [86] Liu, S.; Ammirati, M. J.; Song, X.; Knafels, J. D.; Zhang, J.; Greasley, S. E.; Pfefferkorn, J. A.; Qiu, X. Insights into Mechanism of Glucokinase Activation: Observation of Multiple Distinct Protein Conformations. *J. Biol. Chem.* **2012**, *287* (17), 13598–13610.
- [87] Rosano, C.; Sabini, E.; Rizzi, M.; Deriu, D.; Murshudov, G.; Bianchi, M.; Serafini, G.; Magnani, M.; Bolognesi, M. Binding of Non-Catalytic ATP to Human Hexokinase I Highlights the Structural Components for Enzyme-Membrane Association Control. *Structure* **1999**, *7* (11), 1427–1437.
- [88] Lu, Z.; Hunter, T. Metabolic Kinases Moonlighting as Protein Kinases. *Trends Biochem. Sci.* **2018**, *43* (4), 301–310.
- [89] Smith, T. A. D. Mammalian Hexokinases and Their Abnormal Expression in Cancer. *Br. J. Biomed. Sci.* **2000**, *57* (2), 170–178.
- [90] Osawa, H.; Sutherland, C.; Brooks Robey, R.; Printz, R. L.; Granner, D. K. Analysis of the Signalling Pathway Involved in the Regulation of Hexokinase II Gene Transcription by Insulin. *J. Biol. Chem.* **1996**, *271* (28), 16690–16694.
- [91] Osawa, H.; Robey, R. B.; Printz, R. L.; Granner, D. K. Identification and Characterization of Basal and Cyclic AMP Response Elements in the Promoter of the Rat Hexokinase II Gene. *J Biol Chem* **1996**, *271* (29), 17296–17303.
- [92] Katagiri, M.; Karasawa, H.; Takagi, K.; Nakayama, S.; Yabuuchi, S.; Fujishima, F.; Naitoh, T.; Watanabe, M.; Suzuki, T.; Unno, M.; et al. Hexokinase 2 in Colorectal Cancer: A Potent Prognostic Factor Associated with Glycolysis, Proliferation and Migration. *Histol. Histopathol.* **2017**, *32* (4), 351–360.
- [93] Wu, J.; Hu, L.; Wu, F.; Zou, L.; He, T. Poor Prognosis of Hexokinase 2 Overexpression in Solid Tumors of Digestive System : A Meta-Analysis. **2017**, *8* (19), 32332–32344.
- [94] Zhang, Z.; Feng, X.; Chen, H.; Duan, Z.; Wang, L.; Yang, D.; Liu, P. Prognostic Significance of Synergistic Hexokinase-2 and Beta2-Adrenergic Receptor Expression in Human Hepatocellular Carcinoma after Curative Resection. *BMC Gastroenterol.* **2016**, *16* (57), 1–10.
- [95] Thamrongwarangoon, U.; Seubwai, W.; Phoomak, C.; Sangkhamanon, S.; Cha'on, U.; Boonmars, T.; Wongkham, S. Targeting Hexokinase II as a Possible Therapy for Cholangiocarcinoma. *Biochem. Biophys. Res. Commun.* **2017**, *484* (2), 409–415.
- [96] Kharitonov, S.; Zikriahodzhaev, A.; Ermoshchenkova, M.; Sukhot'ko, A.; Fedorova, M.; Pudova, E.; Alekseev, B.; Kaprin, A.; Kudryavtseva, A. Hexokinases in Breast Cancer. *Int. J. Biosci. Biotechnol.* **2017**, *4* (2), 110–116.
- [97] Wang, H.; Wang, L.; Zhang, Y.; Wang, J.; Deng, Y.; Lin, D. Inhibition of Glycolytic Enzyme Hexokinase II (HK2) Suppresses Lung Tumor Growth. *Cancer Cell Int.* **2016**, *16* (1), 1–11.
- [98] Huang, Y.-P.; Chang, N. W. Proteomic Analysis of Oral Cancer Reveals New Potential Therapeutic Targets Involved in the Warburg Effect. *Clin. Exp. Pharmacol. Physiol.* **2017**, *44*, 880–887.
- [99] Mamede, M.; Higashi, T.; Kitaichi, M.; Ishizu, K.; Ishimori, T.; Nakamoto, Y.; Yanagihara, K.; Li, M.; Tanaka, F.; Wada, H.; et al. [18F]FDG Uptake and PCNA, Glut-1, and Hexokinase-II Expressions in Cancers and Inflammatory Lesions of the Lung. *Neoplasia* **2005**, *7* (4), 369–379.
- [100] Huang, X.; Liu, M.; Sun, H.; Wang, F.; Xie, X.; Chen, X.; Su, J.; He, Y.; Dai, Y.; Wu, H.; et al. HK2 Is a Radiation Resistant and Independent Negative Prognostic Factor for Patients with Locally Advanced Cervical Squamous Cell Carcinoma. *Int J Clin Exp* **2015**, *8* (4), 4054–4063.
- [101] Hamabe, A.; Yamamoto, H.; Konno, M.; Uemura, M.; Nishimura, J.; Hata, T.; Takemasa, I.; Mizushima, T.; Nishida, N.; Kawamoto, K.; et al. Combined Evaluation of Hexokinase 2 and Phosphorylated Pyruvate Dehydrogenase-E1 α in Invasive Front Lesions of Colorectal Tumors Predicts Cancer Metabolism and Patient Prognosis. *Cancer Sci.* **2014**, *105* (9), 1100–1108.
- [102] Qiu, M. Z.; Han, B.; Luo, H. Y.; Zhou, Z. W.; Wang, Z. Q.; Wang, F. H.; Li, Y. H.; Xu, R. H. Expressions of Hypoxia-Inducible Factor-1 α and Hexokinase-II in Gastric Adenocarcinoma: The Impact on Prognosis and Correlation to Clinicopathologic Features. *Tumor Biol.* **2011**, *32* (1), 159–166.
- [103] Gong, L.; Cui, Z.; Chen, P.; Han, H.; Peng, J.; Leng, X. Reduced Survival of Patients with Hepatocellular Carcinoma Expressing Hexokinase II. *Med. Oncol.* **2012**, *29* (2), 909–914.
- [104] Palmieri, D.; Fitzgerald, D.; Shreeve, S. M.; Hua, E.; Bronder, J. L.; Weil, R. J.; Davis, S.; Stark, A. M.; Merino, M. J.; Kurek, R.; et al. Analyses of Resected Human Brain Metastases of Breast Cancer Reveal the Association between Up-Regulation of Hexokinase 2 and Poor Prognosis. *Mol. Cancer Res.* **2009**, *7* (9), 1438–1445.

- [105] Deng, Y.; Lu, J. Targeting Hexokinase 2 in Castration-Resistant Prostate Cancer. *Mol. Cell. Oncol.* **2015**, *2* (3), e974465.
- [106] Hu, J. wei; Sun, P.; Zhang, D. xiang; Xiong, W. jun; Mi, J. Hexokinase 2 Regulates G1/S Checkpoint through CDK2 in Cancer-Associated Fibroblasts. *Cell. Signal.* **2014**, *26* (10), 2210–2216.
- [107] Kim, J. -w.; Gao, P.; Liu, Y. -C.; Semenza, G. L.; Dang, C. V. Hypoxia-Inducible Factor 1 and Dysregulated c-Myc Cooperatively Induce Vascular Endothelial Growth Factor and Metabolic Switches Hexokinase 2 and Pyruvate Dehydrogenase Kinase 1. *Mol. Cell. Biol.* **2007**, *27* (21), 7381–7393.
- [108] Gwak, G.-Y.; Yoon, J.-H.; Kim, K. M.; Lee, H.-S.; Chung, J. W.; Gores, G. J. Hypoxia Stimulates Proliferation of Human Hepatoma Cells through the Induction of Hexokinase II Expression. *J. Hepatol.* **2005**, *42* (3), 358–364.
- [109] Yasuda, S.; Arii, S.; Mori, A.; Isobe, N.; Yang, W.; Oe, H.; Fujimoto, A.; Yonenaga, Y.; Sakashita, H.; Imamura, M. Hexokinase II and VEGF Expression in Liver Tumors: Correlation with Hypoxia-Inducible Factor-1 α and Its Significance. *J. Hepatol.* **2004**, *40* (1), 117–123.
- [110] Iadecola, C.; Anrather, J.; Medical, W. C.; Wang, L.; Xiong, H.; Wu, F.; Zhang, Y. Y.; Wang, J.; Zhao, L.; Guo, X.; et al. Hexokinase 2-Mediated Warburg Effect Is Required for PTEN and P53-Deficiency Driven Prostate Cancer Growth. *Cell Rep.* **2014**, *8* (5), 1461–1474.
- [111] Lin, Y. H.; Wu, M. H.; Huang, Y. H.; Yeh, C. T.; Cheng, M. L.; Chi, H. C.; Tsai, C. Y.; Chung, I. H.; Chen, C. Y.; Lin, K. H. Taurine Up-Regulated Gene 1 Functions as a Master Regulator to Coordinate Glycolysis and Metastasis in Hepatocellular Carcinoma. *Hepatology* **2018**, *67* (1), 188–203.
- [112] Singh, A.; Sen, E. Reciprocal Role of SIRT6 and Hexokinase 2 in the Regulation of Autophagy Driven Monocyte Differentiation. *Exp. Cell Res.* **2017**, *360* (2), 365–374.
- [113] Xia, H. guang; Najafov, A.; Geng, J.; Galan-Acosta, L.; Han, X.; Guo, Y.; Shan, B.; Zhang, Y.; Norberg, E.; Zhang, T.; et al. Degradation of HK2 by Chaperone-Mediated Autophagy Promotes Metabolic Catastrophe and Cell Death. *J. Cell Biol.* **2015**, *210* (5), 705–716.
- [114] Xiao, M.; Lou, C.; Xiao, H.; Yang, Y.; Cai, X.; Li, C.; Jia, S.; Huang, Y. MiR-128 Regulation of Glucose Metabolism and Cell Proliferation in Triple-Negative Breast Cancer. *Br. J. Surg.* **2018**, *105* (1), 75–85.
- [115] Liu, G.; Li, Y.; Gao, X. Overexpression of MicroRNA-133b Sensitizes Non-Small Cell Lung Cancer Cells to Irradiation through the Inhibition of Glycolysis. *Oncol. Lett.* **2016**, *11* (4), 2903–2908.
- [116] Zhao, X.; Lu, C.; Chu, W.; Zhang, B.; Zhen, Q.; Wang, R.; Zhang, Y.; Li, Z.; Lv, B.; Li, H.; et al. MicroRNA-124 Suppresses Proliferation and Glycolysis in Non-Small Cell Lung Cancer Cells by Targeting AKT–GLUT1/HKII. *Tumor Biol.* **2017**, *39* (5), 1–9.
- [117] Tao, T.; Chen, M.; Jiang, R.; Guan, H.; Huang, Y.; Su, H.; Hu, Q.; Han, X.; Xiao, J. Involvement of EZH2 in Aerobic Glycolysis of Prostate Cancer through MiR-181b/HK2 Axis. *Oncol. Rep.* **2017**, *37* (3), 1430–1436.
- [118] Li, L. Q.; Yang, Y.; Chen, H.; Zhang, L.; Pan, D.; Xie, W. J. MicroRNA-181b Inhibits Glycolysis in Gastric Cancer Cells via Targeting Hexokinase 2 Gene. *Cancer Biomarkers* **2016**, *17* (1), 75–81.
- [119] Qin, Y.; Cheng, C.; Lu, H.; Wang, Y. MiR-4458 Suppresses Glycolysis and Lactate Production by Directly Targeting Hexokinase2 in Colon Cancer Cells. *Biochem. Biophys. Res. Commun.* **2016**, *469* (1), 37–43.
- [120] Jiang, S.; Yan, W.; Wang, S. E.; Baltimore, D. Let-7 Suppresses B Cell Activation through Restricting the Availability of Necessary Nutrients. *Cell Metab.* **2018**, *27* (2), 393–403.
- [121] Liu, C.; Cai, L.; Li, H. MiR-185 Regulates the Growth of Osteosarcoma Cells via Targeting Hexokinase 2. *Mol. Med. Rep.* **2019**, *20* (3), 2774–2782.
- [122] Wu, J.; Zhang, X.; Wang, Y.; Sun, Q.; Chen, M.; Liu, S.; Zou, X. Licochalcone A Suppresses Hexokinase 2-Mediated Tumor Glycolysis in Gastric Cancer via Downregulation of the Akt Signaling Pathway. *Oncol. Rep.* **2017**, *39*, 1181–1190.
- [123] Xu, D.; Jin, J.; Yu, H.; Zhao, Z.; Ma, D.; Zhang, C.; Jiang, H. Chrysin Inhibited Tumor Glycolysis and Induced Apoptosis in Hepatocellular Carcinoma by Targeting Hexokinase-2. *J. Exp. Clin. Cancer Res.* **2017**, *36* (1), 44.
- [124] Gao, X.; Han, H. Jolkinolide B Inhibits Glycolysis by Downregulating Hexokinase 2 Expression through Inactivating the Akt/MTOR Pathway in Non-Small Cell Lung Cancer Cells. *J. Cell. Biochem.* **2018**, *119* (6), 4967–4974.

- [125] Chen, G.-Q.; Tang, C.-F.; Shi, X.-K.; Lin, C.-Y.; Fatima, S.; Pan, X.-H.; Yang, D.-J.; Zhang, G.; Lu, A.-P.; Lin, S.-H.; et al. Halofuginone Inhibits Colorectal Cancer Growth through Suppression of Akt/MTORC1 Signaling and Glucose Metabolism. *Oncotarget* **2015**, *6* (27), 24148–24162.
- [126] Wei, L.; Dai, Q.; Zhou, Y.; Zou, M.; Li, Z.; Lu, N.; Guo, Q. Oroxylin A Sensitizes Non-Small Cell Lung Cancer Cells to Anoikis via Glucose-Deprivation-like Mechanisms: C-Src and Hexokinase II. *Biochim. Biophys. Acta - Gen. Subj.* **2013**, *1830* (6), 3835–3845.
- [127] Zhang, J.; Wang, S.; Jiang, B.; Huang, L.; Ji, Z.; Li, X.; Zhou, H.; Han, A.; Chen, A.; Wu, Y.; et al. C-Src Phosphorylation and Activation of Hexokinase Promotes Tumorigenesis and Metastasis. *Nat. Commun.* **2017**, *8* (13732), e13732.
- [128] Camara, A. K. S.; Zhou, Y. F.; Wen, P. C.; Tajkhorshid, E.; Kwok, W. M. Mitochondrial VDAC1: A Key Gatekeeper as Potential Therapeutic Target. *Front. Physiol.* **2017**, *8*, 1–18.
- [129] Rosano, C.; Einstein, A. Molecular Model of Hexokinase Binding to the Outer Mitochondrial Membrane Porin (VDAC1): Implication for the Design of New Cancer Therapies. *Mitochondrion* **2011**, *11* (3), 513–519.
- [130] Roberts, D. J.; Tan-Sah, V. P.; Smith, J. M.; Miyamoto, S. Akt Phosphorylates HK-II at Thr-473 and Increases Mitochondrial HK-II Association to Protect Cardiomyocytes. *J. Biol. Chem.* **2013**, *288* (33), 23798–23806.
- [131] Kolar, D.; Gresikova, M.; Waskova-Arnostova, P.; Elsnicova, B.; Kohutova, J.; Hornikova, D.; Vebr, P.; Neckar, J.; Blahova, T.; Kasparova, D.; et al. Adaptation to Chronic Continuous Hypoxia Potentiates Akt/HK2 Anti-Apoptotic Pathway during Brief Myocardial Ischemia/Reperfusion Insult. *Mol. Cell. Biochem.* **2017**, *432*, 99–108.
- [132] Shuvo, S. R.; Ferens, F. G.; Court, D. A. The N-Terminus of VDAC: Structure, Mutational Analysis, and a Potential Role in Regulating Barrel Shape. *Biochim. Biophys. Acta - Biomembr.* **2016**, *1858* (6), 1350–1361.
- [133] Abu-Hamad, S.; Arbel, N.; Calo, D.; Arzoine, L.; Israelson, A.; Keinan, N.; Ben-Romano, R.; Friedman, O.; Shoshan-Barmatz, V. The VDAC1 N-Terminus Is Essential Both for Apoptosis and the Protective Effect of Anti-Apoptotic Proteins. *J. Cell Sci.* **2009**, *122* (11), 1906–1916.
- [134] Geula, S.; Ben-Hail, D.; Shoshan-Barmatz, V. Structure-Based Analysis of VDAC1: N-Terminus Location, Translocation, Channel Gating and Association with Anti-Apoptotic Proteins. *Biochem. J.* **2012**, *444* (3), 475–485.
- [135] Calmettes, G.; Ribalet, B.; John, S.; Korge, P.; Ping, P.; James, N. Hexokinases and Cardioprotection. *J. Mol. Cell Cardiol.* **2016**, 107–115.
- [136] Bensaad, K.; Tsuruta, A.; Selak, M. A.; Vidal, M. N. C.; Nakano, K.; Bartrons, R.; Gottlieb, E.; Vousden, K. H. TIGAR, a P53-Inducible Regulator of Glycolysis and Apoptosis. *Cell* **2006**, *126* (1), 107–120.
- [137] Cheung, E. C.; Ludwig, R. L.; Vousden, K. H. Mitochondrial Localization of TIGAR under Hypoxia Stimulates HK2 and Lowers ROS and Cell Death. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109* (50), 20491–20496.
- [138] Lopez, K. L. R. The Role of TIGAR in Parkinson 's Disease. PhD Thesis, The University of Sheffield: Sheffield, December. **2017**.
- [139] Sheng, H.; Tang, W. Glycolysis Inhibitors for Anticancer Therapy: A Review of Recent Patents. *Recent Pat. Anticancer. Drug Discov.* **2016**, *11*, 297–308.
- [140] Fortunato, S.; Bononi, G.; Granchi, C.; Minutolo, F. An Update on Patents Covering Agents That Interfere with the Cancer Glycolytic Cascade. *ChemMedChem* **2018**, *13* (21), 2251–2265.
- [141] Salani, B.; Marini, C.; Rio, A. Del; Ravera, S.; Massollo, M.; Orengo, A. M.; Amaro, A.; Passalacqua, M.; Maffioli, S.; Pfeffer, U.; et al. Metformin Impairs Glucose Consumption and Survival in Calu-1 Cells by Direct Inhibition of Hexokinase-II. *Sci. Rep.* **2013**, *3*, 2070.
- [142] Marini, C.; Salani, B.; Massollo, M.; Amaro, A.; Esposito, A. I.; Orengo, A. M.; Capitanio, S.; Emionite, L.; Riondato, M.; Bottoni, G.; et al. Direct Inhibition of Hexokinase Activity by Metformin at Least Partially Impairs Glucose Metabolism and Tumor Growth in Experimental Breast Cancer. *Cell Cycle* **2013**, *12* (22), 3490–3499.
- [143] Kang, Y.; Hsu, W.; Wu, C.; Hsin, I.; Wu, P. Metformin Alleviates Nickel-Induced Autophagy and Apoptosis via Inhibition of Hexokinase-2 , Activating Lipocalin-2 , in Human Bronchial Epithelial Cells. **2017**, *8* (62), 105536–105552.

- [144] Pernicova, I.; Korbonits, M. Metformin-Mode of Action and Clinical Implications for Diabetes and Cancer. *Nat. Rev. Endocrinol.* **2014**, *10* (3), 143–156.
- [145] Ko, Y. H.; Pedersen, P. L.; Geschwind, J. F. Glucose Catabolism in the Rabbit VX2 Tumor Model for Liver Cancer: Characterization and Targeting Hexokinase. *Cancer Lett.* **2001**, *173* (1), 83–91.
- [146] Kurtoglu, M.; Maher, J. C.; Lampidis, T. J. Differential Toxic Mechanisms of 2-Deoxy-D-Glucose *versus* 2-Fluorodeoxy-D -Glucose in Hypoxic and Normoxic Tumor Cells. *Antioxid. Redox Signal.* **2007**, *9* (9), 1383–1390.
- [147] Tidmarsh, G. Treatment of Cancer with 2-Deoxyglucose. US6979675B2, **2005**.
- [148] Chambers, J. W.; Fowler, M. L.; Morris, M. T.; Morris, J. C. The Anti-Trypanosomal Agent Lonidamine Inhibits Trypanosoma Brucei Hexokinase 1. *Mol. Biochem. Parasitol.* **2008**, *158* (2), 202–207.
- [149] Mjalli, A. M. M.; Gaddam, B.; Gohimukkula, D. R.; Poliseti, D. R.; Rao, M.; Guzel, M.; Singh, N.; Hajjo, R.; Andrews, R. C.; Xie, R.; et al. Inhibitors of Hexokinase and Methods of Use Thereof, **2016**.
- [150] Lin, H.; Luengo, J. I.; Schulz, M.; Zeng, J. Glucosamine Derivatives. WO2012083145 A1, **2012**.
- [151] Maley, F.; Lardy, H. A. Synthesis of N-Substitued Glucosamines and Their Effect on Hexokinase. *J. Biol. Chem.* **1955**, *214* (2), 765–773.
- [152] Coats, E. A.; Skau, K. A.; Caperelli, C. A.; Solomacha, D. Exploring the Hexokinase Glucose Binding Site through Correlation Analysis and Molecular Modeling of Glucosamine Inhibitors. *J. Enzyme Inhib. Med. Chem.* **1993**, *6* (4), 271–282.
- [153] Li, W.; Zheng, M.; Wu, S.; Gao, S.; Yang, M.; Li, Z.; Min, Q.; Sun, W.; Chen, L.; Xiang, G.; et al. Benserazide, a Dopadecarboxylase Inhibitor, Suppresses Tumor Growth by Targeting Hexokinase 2. *J. Exp. Clin. Cancer Res.* **2017**, *36*:58, 1–12.
- [154] Bao, F.; Yang, K.; Wu, C.; Gao, S.; Wang, P.; Chen, L.; Li, H. New Natural Inhibitors of Hexokinase 2 (HK2): Steroids from Ganoderma Sinense. *Fitoterapia* **2018**, *125*, 123–129.
- [155] Liu, Y.; Li, M.; Zhang, Y.; Wu, C.; Yang, K.; Gao, S.; Zheng, M.; Li, X.; Li, H.; Chen, L. Structure Based Discovery of Novel Hexokinase 2 Inhibitors. *Bioorg. Chem.* **2020**, *96* (December 2019), 103609.
- [156] Roche, O.; Schneider, P.; Zuegge, J.; Guba, W.; Kansy, M.; Alanine, A.; Bleicher, K.; Danel, F.; Gutknecht, E. M.; Rogers-Evans, M.; et al. Development of a Virtual Screening Method for Identification of “Frequent Hitters” in Compound Libraries. *J. Med. Chem.* **2002**, *45* (1), 137–142.
- [157] Flaherty, D. P.; Harris, M. T.; Schroeder, C. E.; Khan, H.; Kahney, E. W.; Hackler, A. L.; Patrick, S. L.; Weiner, W. S.; Aubé, J.; Sharlow, E. R.; et al. Optimization and Evaluation of Antiparasitic Benzamidobenzoic Acids as Inhibitors of Kinetoplastid Hexokinase 1. *ChemMedChem* **2017**, *12* (23), 1994–2005.
- [158] Sharlow, E. R.; Lyda, T. A.; Dodson, H. C.; Mustata, G.; Morris, M. T.; Leimgruber, S. S.; Lee, K. H.; Kashiwada, Y.; Close, D.; Lazo, J. S.; et al. A Target-Based High Throughput Screen Yields Trypanosoma Brucei Hexokinase Small Molecule Inhibitors with Antiparasitic Activity. *PLoS Negl. Trop. Dis.* **2010**, *4* (4), e659.
- [159] Gordhan, H. M.; Patrick, S. L.; Swasy, M. I.; Hackler, A. L.; Anayee, M.; Golden, J. E.; Morris, J. C.; Whitehead, D. C. Evaluation of Substituted Ebselen Derivatives as Potential Trypanocidal Agents. *Bioorganic Med. Chem. Lett.* **2017**, *27* (3), 537–541.
- [160] Saucedo-Mendiola, M. L.; Salas-Pacheco, J. M.; Nájera, H.; Rojo-Domínguez, A.; Yépez-Mulia, L.; Avitia-Domínguez, C.; Téllez-Valencia, A. Discovery of *Entamoeba Histolytica* Hexokinase 1 Inhibitors through Homology Modeling and Virtual Screening. *J. Enzyme Inhib. Med. Chem.* **2014**, *29* (3), 325–332.
- [161] Tielens, A. G. M.; Houweling, M.; Van Den Bergh, S. G. The Effect of 5-Thioglucose on the Energy Metabolism of Schistosoma Mansoni in Vitro. *Biochem. Pharmacol.* **1985**, *34* (18), 3369–3373.
- [162] Goldin, N.; Arzoin, L.; Heyfets, A.; Israelson, A.; Zaslavsky, Z.; Bravman, T.; Bronner, V.; Notcovich, A.; Flescher, E.; Shoshan-Barmatz, V.; et al. Methyl Jasmonate Binds to and Detaches Mitochondria-Bound Hexokinase. *Oncogene* **2008**, *27* (34), 4636–4643.
- [163] Sucu, B. O.; Ipek, O. S.; Kurtulus, S. O.; Yazici, B. E.; Karakas, N.; Guzel, M. Synthesis of Novel Methyl Jasmonate Derivatives and Evaluation of Their Biological Activity in Various Cancer Cell Lines. *Bioorg. Chem.* **2019**, *91* (April), 103146.
- [164] Tidmarsh, G. Combination Therapies for the Treatment of Cancer. US 20060276527A1, **2006**.

- [165] Arzoine, L.; Zilberberg, N.; Ben-Romano, R.; Shoshan-Barmatz, V. Voltage-Dependent Anion Channel 1-Based Peptides Interact with Hexokinase to Prevent Its Anti-Apoptotic Activity. *J. Biol. Chem.* **2009**, *284* (6), 3946–3955.
- [166] Prezma, T.; Shteinifer, A.; Admoni, L.; Raviv, Z.; Sela, I.; Levi, I.; Shoshan-Barmatz, V. VDAC1-Based Peptides: Novel pro-Apoptotic Agents and Potential Therapeutics for B-Cell Chronic Lymphocytic Leukemia. *Cell Death Dis.* **2013**, *4*, e809.
- [167] Woldetsadik, A. D.; Vogel, M. C.; Rabeh, W. M.; Magzoub, M. Hexokinase II-Derived Cell-Penetrating Peptide Targets Mitochondria and Triggers Apoptosis in Cancer Cells. *FASEB J.* **2017**, *31* (5), 2168–21874.
- [168] Wang, Z.; Chen, J.; Yang, L.; Cao, M.; Yu, Y.; Zhang, R.; Quan, H.; Jiang, Q.; Hua, Y.; Wei, W.; et al. Single-Cell Sequencing-Enabled Hexokinase 2 Assay for Noninvasive Bladder Cancer Diagnosis and Screening by Detecting Rare Malignant Cells in Urine. *Anal. Chem.* **2020**, *92* (24), 16284–16292.
- [169] Sato-Tadano, A.; Suzuki, T.; Amari, M.; Takagi, K.; Miki, Y.; Tamaki, K.; Watanabe, M.; Ishida, T.; Sasano, H.; Ohuchi, N. Hexokinase II in Breast Carcinoma: A Potent Prognostic Factor Associated with Hypoxia-Inducible Factor-1 α and Ki-67. *Cancer Sci.* **2013**, *104* (10), 1380–1388.
- [170] Hauser, D. N.; Mamais, A.; Conti, M. M.; Primiani, C. T.; Kumaran, R.; Dillman, A. A.; Langston, R. G.; Beilina, A.; Garcia, J. H.; Diaz-Ruiz, A.; et al. Hexokinases Link DJ-1 to the PINK1/Parkin Pathway. *Mol. Neurodegener.* **2017**, *12*:70, 1–17.
- [171] Varanasi, S. K.; Jaggi, U.; Hay, N.; Rouse, B. T. Hexokinase II May Be Dispensable for CD4 T Cell Responses against a Virus Infection. *PLoS One* **2018**, *13* (1), e0191533.
- [172] Wang, C.; Silverman, R. M.; Shen, J.; O'Keefe, R. J. Distinct Metabolic Programs Induced by TGF-B1 and BMP2 in Human Articular Chondrocytes with Osteoarthritis. *J. Orthop. Transl.* **2018**, *12*, 66–73.
- [173] Xi, F.; Ye, J. Inhibition of Lung Carcinoma A549 Cell Growth by Knockdown of Hexokinase 2 In Situ and In Vivo. *Oncol. Res.* **2016**, *23* ((1-2)), 53–59.
- [174] DeWaal, D.; Nogueira, V.; Terry, A. R.; Patra, K. C.; Jeon, S.-M.; Guzman, G.; Au, J.; Long, C. P.; Antoniewicz, M. R.; Hay, N. Hexokinase-2 Depletion Inhibits Glycolysis and Induces Oxidative Phosphorylation in Hepatocellular Carcinoma and Sensitizes to Metformin. *Nat. Commun.* **2018**, *9* (1), e446.
- [175] Nogueira, V.; Patra, K. C.; Hay, N. Selective Eradication of Cancer Displaying Hyperactive Akt by Exploiting the Metabolic Consequences of Akt Activation. *Elife* **2018**, *7*, 1–23.
- [176] Peng, Q.-P.; Zhou, J.-M.; Zhou, Q.; Pan, F.; Zhong, D.-P.; Liang, H.-J. Downregulation of the Hexokinase II Gene Sensitizes Human Colon Cancer Cells to 5-Fluorouracil. *Chemotherapy* **2008**, *54* (5), 357–363.
- [177] Liu, Y.; Murray-Stewart, T.; Casero, R. A.; Kagiampakis, I.; Jin, L.; Jiawen, Z.; Huihui, W.; Che, Q.; Huan, T.; Ke, J.; et al. Targeting Hexokinase 2 Inhibition Promotes Radiosensitization in HPV16 E7-Induced Cervical Cancer and Suppresses Tumor Growth. *Int. J. Oncol.* **2017**, *50* (6), 2011–2023.
- [178] Zappasodi, R.; Serganova, I.; Cohen, I. J.; Maeda, M.; Shindo, M.; Senbabaoglu, Y.; Watson, M. L. J.; Leftin, A.; Maniyar, R.; Verma, S.; et al. CTLA-4 Blockade Drives Loss of Treg Stability in Glycolysis-Low Tumours. *Nature* **2021**, *591* (November 2019).
- [179] Nederlof, R.; Eerbeek, O.; Hollmann, M. W.; Southworth, R.; Zuurbier, C. J. Targeting Hexokinase II to Mitochondria to Modulate Energy Metabolism and Reduce Ischaemia-Reperfusion Injury in Heart. *Br. J. Pharmacol.* **2014**, *171* (8), 2067–2079.
- [180] Li, Y.; Lu, B.; Sheng, L.; Zhu, Z.; Sun, H.; Zhou, Y.; Yang, Y.; Xue, D.; Chen, W.; Tian, X.; et al. Hexokinase 2-Dependent Hyperglycolysis Driving Microglial Activation Contributes to Ischemic Brain Injury. *J. Neurochem.* **2018**, *144* (2), 186–200.
- [181] Ghosh, S.; Gupta, P.; Sen, E. TNF α Driven HIF-1 α -Hexokinase II Axis Regulates MHC-I Cluster Stability through Actin Cytoskeleton. *Exp. Cell Res.* **2016**, *340* (1), 116–124.
- [182] Wolf, A. J.; Reyes, C. N. R.; Liang, W.; Becker, C.; Shimada, K.; Wheeler, M. L.; Cho, H. C.; Popescu, N. I.; Coggeshall, M. K.; Arditi, M.; et al. Hexokinase Is an Innate Immune Receptor for the Detection of Bacterial Peptidoglycan. *Cell* **2016**, *166* (3), 624–636.
- [183] Okuyama, N.; Matsuda, S.; Yamashita, A.; Moriguchi-Goto, S.; Sameshima, N.; Iwakiri, T.; Matsuura, Y.; Sato, Y.; Asada, Y. Human Coronary Thrombus Formation Is Associated With Degree of Plaque Disruption and Expression of Tissue Factor and Hexokinase II. *Circ. J.* **2015**, *79*, 2430–2438.

- [184] Okatsu, K.; Iemura, S. ichiro; Koyano, F.; Go, E.; Kimura, M.; Natsume, T.; Tanaka, K.; Matsuda, N. Mitochondrial Hexokinase HK1 Is a Novel Substrate of the Parkin Ubiquitin Ligase. *Biochem. Biophys. Res. Commun.* **2012**, *428* (1), 197–202.
- [185] McCoy, M. K.; Kaganovich, A.; Rudenko, I. N.; Ding, J.; Cookson, M. R. Hexokinase Activity Is Required for Recruitment of Parkin to Depolarized Mitochondria. *Hum. Mol. Genet.* **2014**, *23* (1), 145–156.
- [186] Xu, J.; Lin, S.; Myers, R. W.; Addona, G.; Berger, J. P.; Campbell, B.; Chen, H. shen; Chen, Z.; Eiermann, G. J.; Elowe, N. H.; et al. Novel, Highly Potent Systemic Glucokinase Activators for the Treatment of Type 2 Diabetes Mellitus. *Bioorganic Med. Chem. Lett.* **2017**, *27* (9), 2069–2073.
- [187] Malkki, M.; Laakso, M.; Deeb, S. S. The Human Hexokinase II Gene Promoter: Functional Characterization and Detection of Variants among Patients with NIDDM. *Diabetologia* **1997**, *40* (12), 1461–1469.
- [188] Lis, P.; Dylag, M.; Niedzwiecka, K.; Ko, Y. H.; Pederson, P. L.; Goffeau, A.; Ulaszewski, S. The HK2 Dependent “Waburg Effect” and Mitochondrial Oxidative Phosphorylation in Cancer : Targets for Effective Therapy with 3-Bromopyruvate. *Molecules* **2016**, *21* (12), e1730.
- [189] Peng, Q.; Zhou, Q.; Zhou, J.; Zhong, D.; Pan, F.; Liang, H. Stable RNA Interference of Hexokinase II Gene Inhibits Human Colon Cancer LoVo Cell Growth in Vitro and in Vivo. *Cancer Biol. Ther.* **2008**, *7* (7), 1128–1135.
- [190] Bender, A.; Bojanic, D.; Davies, J. W.; Crisman, T. J.; Mikhailov, D.; Scheiber, J.; Jenkins, J. L.; Deng, Z.; Hill, W. A. G.; Popov, M.; et al. Which Aspects of HTS Are Empirically Correlated with Downstream Success? *Curr. Opin. Drug Discov. Dev.* **2008**, *11* (3), 327–337.
- [191] Mayr, L. M.; Fuerst, P. The Future of High-Throughput Screening. *J. Biomol. Screen.* **2008**, *13* (6), 443–448.
- [192] Mohs, R. C.; Greig, N. H. Drug Discovery and Development : Role of Basic Biological Research. **2017**, *3*, 651–657.
- [193] Cavasotto, C. N.; Aucar, M. G.; Adler, N. S. Computational Chemistry in Drug Lead Discovery and Design. *Int. J. Quantum Chem.* **2019**, *119* (2), 1–19.
- [194] Slater, O.; Kontoyianni, M. The Compromise of Virtual Screening and Its Impact on Drug Discovery. *Expert Opin. Drug Discov.* **2019**, *14* (7), 619–637.
- [195] Van Montfort, R. L. M.; Workman, P. Structure-Based Drug Design: Aiming for a Perfect Fit. *Essays Biochem.* **2017**, *61*, 431–437.
- [196] Ferreira, L. G.; Dos Santos, R. N.; Oliva, G.; Andricopulo, A. D. Molecular Docking and Structure-Based Drug Design Strategies. *Molecules* **2015**, *20* (7), 13384–13421.
- [197] Ramírez, D. Computational Methods Applied to Rational Drug Design. *Open Med. Chem. J.* **2016**, *10* (1), 7–20.
- [198] Mavromoustakos, T.; Durdagi, S.; Koukoulitsa, C.; Simcic, M.; G. Papadopoulos, M.; Hodosecek, M.; Golig Grdadolnik, S. Strategies in the Rational Drug Design. *Curr. Med. Chem.* **2011**, *18* (17), 2517–2530.
- [199] Ghosh, J.; Lawless, M. S.; Waldman, M.; Gombar, V.; Fraczkiwicz, R. Chapter 4- Modeling Admet. In *In Silico Methods for Predicting Drug Toxicity, Methods in Molecular Biology*; **2016**; Vol. 1425, pp 63–83.
- [200] Fraczkiwicz, R.; Zhuang, D.; Zhang, J.; Miller, D.; Woltosz, W. S.; Bolger, M. B. Busting the Black Box Myth: Designing Out Unwanted ADMET Properties with Machine Learning Approaches. *CICSJ Bull.* **2009**, *27* (4), 96.
- [201] Wang, J.; Hou, T. *Chapter 5 - Recent Advances on in Silico ADME Modeling*; Elsevier, **2009**; Vol. 5.
- [202] Click2Drug - Swiss Institute of Bioinformatics (SIB). Directory of computer-aided Drug Design tools - ADME Toxicity [https://www.click2drug.org/index.html#ADME Toxicity](https://www.click2drug.org/index.html#ADME%20Toxicity) (accessed May 29, **2019**).
- [203] Bhhatari, B.; Walters, W. P.; Hop, C. E. C. A.; Lanza, G.; Ekins, S. Opportunities and Challenges Using Artificial Intelligence in ADME/Tox. *Nat. Mater.* **2019**, *18*, 410–427.
- [204] University of Paris Diderot. FAFdrugs4 Characterization of stress degradation products of amodiaquine dihydrochloride by liquid chromatography with high-re (accessed Jun 19, 2019).
- [205] Lagorce, D.; Bouslama, L.; Becot, J.; Miteva, M. A.; Villoutreix, B. O.; Brion, H. Structural Bioinformatics FAF-Drugs4 : Free ADME-Tox Filtering Computations for Chemical Biology and Early Stages Drug Discovery. **2017**, *33* (July), 3658–3660.

- [206] Bruns, R. F.; Watson, I. A. Rules for Identifying Potentially Reactive or Promiscuous Compounds. *J. Med. Chem.* **2012**, *55* (22), 9763–9772.
- [207] Sushko, I.; Salmina, E.; Potemkin, V. A.; Poda, G.; Tetko, I. V. ToxAlerts: A Web Server of Structural Alerts for Toxic Chemicals and Compounds with Potential Adverse Reactions. *J. Chem. Inf. Model.* **2012**, *52* (8), 2310–2316.
- [208] Anderson, F. S.; Murphy, R. C. Isocratic Separation of Some Purine Nucleotide, Nucleoside and Base Metabolites from Biological Extracts by High-Performance Liquid Chromatography. *J. Chromatogr.* **1976**, *121*, 251–262.
- [209] Ryder, J. M. Determination of Adenosine Triphosphate and Its Breakdown Products in Fish Muscle by High-Performance Liquid Chromatography. *J. Agric. Food Chem.* **1985**, *33* (4), 678–680.
- [210] Veciana-Nogues, M. T.; Izquierdo-Pulido, M.; Vidal-Carou, M. C. Determination of ATP Related Compounds in Fresh and Canned Tuna Fish by HPLC. *Food Chem.* **1997**, *59* (3), 467–472.
- [211] Brito Santos, R.; Pereira da Silva, R.; Akihiro Melo Otsuka, F.; de Jesus Trindade, D.; Costa Santos, A.; Reis Matos, H. An HPLC Method for the Determination of Adenosine Diphosphate: An Important Marker of Hexokinase Activity in Metabolic Diseases. *Biomed. Chromatogr.* **2019**, *33* (4), 1–9.
- [212] Taylor, P.; Ma, H.; Deacon, S.; Horiuchi, K. The Challenge of Selecting Protein Kinase Assays for Lead Discovery Optimization The Challenge of Selecting Protein Kinase Assays for Lead Discovery Optimization. *Expert Opin. Drug Discov.* **2008**, *3* (6), 607–621.
- [213] Morris, G. M.; Huey, R.; Lindstrom, W.; Sanner, M. F.; Belew, R. K.; Goodsell, D. S.; Olson, A. J. AutoDock4 and AutoDockTools4: Automated Docking with Selective Receptor Flexibility. **2010**, *30* (16), 2785–2791.
- [214] Trott, O.; Olson, A. J. AutoDock Vina: Improving the Speed and Accuracy of Docking with a New Scoring Function, Efficient Optimization and Multithreading. *J. Comput. Chem.* **2010**, *31*, 455–461.
- [215] Jones, G.; Willett, P.; Glen, R. C.; Leach, A. R.; Taylor, R. Development and Validation of a Genetic Algorithm for Flexible Docking. *J. Mol. Biol.* **1997**, *267* (3), 727–748.
- [216] Developmental Therapeutics Program (DTP) - National Cancer Institute. NCI/DTP Open Chemicals. Assessed on February **2021**.
- [217] Irwin, J. J.; Shoichet, B. K. ZINC – A Free Database of Commercially Available Compounds for Virtual Screening. *J. Chem. Inf. Model.* **2005**, *45* (1), 177–182.
- [218] Ortuso, F.; Bagetta, D.; Maruca, A.; Talarico, C.; Bolognesi, M. L.; Haider, N.; Borges, F.; Bryant, S.; Langer, T.; Senderowitz, H.; et al. The Mu.Ta.Lig. Chemothea: A Community-Populated Molecular Database for Multi-Target Ligands Identification and Compound-Repurposing. *Front. Chem.* **2018**, *6* (APR), 1–6.
- [219] Sushko, I.; Novotarskyi, S.; Körner, R.; Pandey, A. K.; Rupp, M.; Teetz, W.; Brandmaier, S.; Abdelaziz, A.; Prokopenko, V. V.; Tanchuk, V. Y.; et al. Online Chemical Modeling Environment (OCHEM): Web Platform for Data Storage, Model Development and Publishing of Chemical Information. *J. Comput. Aided. Mol. Des.* **2011**, *25* (6), 533–554.
- [220] Capuzzi, S. J.; Muratov, E. N.; Tropsha, A. Phantom PAINS: Problems with the Utility of Alerts for P an- A Ssay in Terference Compound S. *J. Chem. Inf. Model.* **2017**, *57* (3), 417–427.
- [221] Gleeson, M. P. Generation of a Set of Simple, Interpretable ADMET Rules of Thumb. *J. Med. Chem.* **2008**, *51* (4), 817–834.
- [222] Hughes, J. D.; Blagg, J.; Price, D. A.; Bailey, S.; DeCrescenzo, G. A.; Devraj, R. V.; Ellsworth, E.; Fobian, Y. M.; Gibbs, M. E.; Gilles, R. W.; et al. Physiochemical Drug Properties Associated with in Vivo Toxicological Outcomes. *Bioorganic Med. Chem. Lett.* **2008**, *18* (17), 4872–4875.
- [223] Lipinski, C. A. Lead- and Drug-like Compounds: The Rule-of-Five Revolution. *Drug Discov. Today Technol.* **2004**, *1* (4), 337–341.
- [224] Gottschaldt, M.; Schubert, U. S. Prospects of Metal Complexes Peripherally Substituted with Sugars in Biomedical Applications. *Chem. - A Eur. J.* **2009**, *15* (7), 1548–1557.
- [225] Kenny, R. G.; Marmion, C. J. Toward Multi-Targeted Platinum and Ruthenium Drugs - A New Paradigm in Cancer Drug Treatment Regimens? *Chem. Rev.* **2019**, *119* (2), 1058–1137.
- [226] Schmollinger, D.; Kraft, J.; Ewald, C.; Ziegler, T. Synthesis of Ruthenium and Palladium Complexes from Glycosylated 2,2'-Bipyridine and Terpyridine Ligands. *Tetrahedron Lett.* **2017**, *58* (37), 3643–3645.

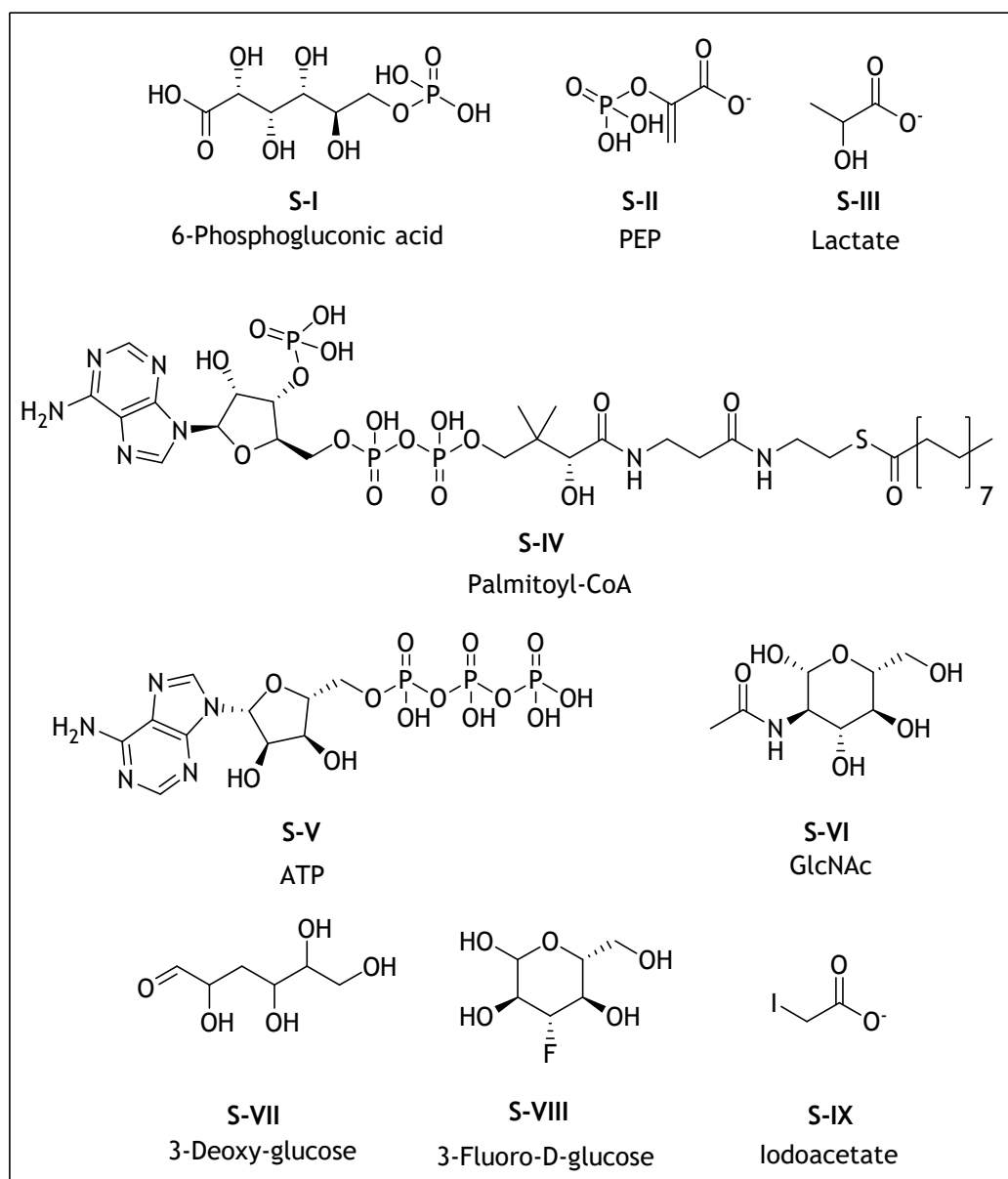
- [227] Byrne, J. P.; Musembi, P.; Albrecht, M. Carbohydrate-Functionalized N-Heterocyclic Carbene Ru(II) Complexes: Synthesis, Characterization and Catalytic Transfer Hydrogenation Activity. *Dalt. Trans.* **2019**, 48 (31), 11838–11847.
- [228] Fernandes, A. C. Synthesis, Biological Activity and Medicinal Applications of Ruthenium Complexes Containing Carbohydrate Ligands. *Curr. Med. Chem.* **2019**, 26 (35), 6412–6437.
- [229] Berger, I.; Hanif, M.; Nazarov, A. A.; Hartinger, C. G.; John, R. O.; Kuznetsov, M. L.; Groessler, M.; Schmitt, F.; Zava, O.; Biba, F.; et al. In Vitro Anticancer Activity and Biologically Relevant Metabolization of Organometallic Ruthenium Complexes with Carbohydrate-Based Ligands. *Chem. - A Eur. J.* **2008**, 14 (29), 9046–9057.
- [230] Schibli, R.; Dumas, C.; Petrig, J.; Spadola, L.; Scapozza, L.; Garcia-Garayoa, E.; Schubiger, P. A. Synthesis and in Vitro Characterization of Organometallic Rhenium and Technetium Glucose Complexes against Glut 1 and Hexokinase. *Bioconjug. Chem.* **2005**, 16 (1), 105–112.
- [231] Bianconi, M. L. Calorimetric Determination of Thermodynamic Parameters of Reaction Reveals Different Enthalpic Compensations of the Yeast Hexokinase Isozymes. *J. Biol. Chem.* **2003**, 278 (21), 18709–18713.
- [232] Chemical Computing Group ULC. Molecular Operating Environment (MOE). Montreal, QC **2017**.
- [233] O'Boyle, N. M.; Banck, M.; James, C. A.; Morley, C.; Vandermeersch, T.; Hutchison, G. R. Open Babel: An Open Chemical Toolbox. *J. Cheminform.* **2011**, 3 (1), 33.
- [234] Wishart, D. S.; Knox, C.; Guo, A. C.; Shrivastava, S.; Hassanali, M.; Stothard, P.; Chang, Z.; Woolsey, J. DrugBank: A Comprehensive Resource for in Silico Drug Discovery and Exploration. *Nucleic Acids Res.* **2006**, 34 (90001), D668–D672.
- [235] Kinase/ATPase Activity Assays - Protocols & Applications Guide, rev.7/18 www.promega.com.
- [236] Davis, B. G.; Fairbanks, A. J. *Carbohydrate Chemistry*; Oxford University Press: New York, **2002**.
- [237] Buckingham, J. *Dictionary of Natural Products. Version 9.2 on CD-ROM*; John Buckingham, Ed.; CRC Press: London, **1993**.
- [238] Galan, M. C.; Benito-Alifonso, D.; Watt, G. Carbohydrate Chemistry in Drug Discovery. *Org. Biomol. Chem.* **2011**, 9, 3598–3610.
- [239] Magnani, J. L. Glycomimetic Drugs - A New Source of Therapeutic Opportunities. *Discov. Med.* **2009**, 8 (43), 247–252.
- [240] Ohri, H.; Jones, G. H.; Moffatt, J. G.; Maddox, M. L.; Christensen, A. T.; Byram, S. K. C- Glycosyl Nucleosides. V. Some Unexpected Observations on the Relative Stabilities of Compounds Containing Fused Five-Membered Rings with Epimerizable Substituents. *J. Am. Chem. Soc.* **1975**, 97 (16), 4602–4613.
- [241] Kim, Y. A.; Sharon, A.; Chu, C. K.; Rais, R. H.; Safarjalani, O. N. Al; Naguib, F. N. M.; Kouni, M. H. Structure - Activity Relationships of 7-Deaza-6-Benzylthioinosine Analogues as Ligands of Toxoplasma Gondii Adenosine Kinase. **2008**, 3934–3945.
- [242] Well, R. M. Van; Overkleeft, H. S.; Overhand, M.; Carstenen, E. V.; Marel, G. A. Van Der; Boom, J. H. Van. Parallel Synthesis of Cyclic Sugar Amino Acid / Amino Acid Hybrid Molecules. **2000**, 41, 9331–9335.
- [243] Well, R. M. Van; Marinelli, L.; Erkelens, K.; Marel, G. A. Van Der; Lavecchia, A.; Overkleeft, H. S.; Boom, J. H. Van; Kessler, H.; Overhand, M. Synthesis and Structural Analysis of Cyclic Oligomers Consisting of Furanoid and Pyranoid ϵ -Sugar Amino Acids. **2003**, 2303–2313.
- [244] Écija, P.; Uriarte, I.; Spada, L.; Davis, B. G.; Caminati, W.; Basterretxea, F. J.; Lesarri, A.; Cocinero, E. J. Furanosic Forms of Sugars: Conformational Equilibrium of Methyl β -D-Ribofuranoside. *Chem. Commun.* **2016**, 52 (37), 6241–6244.
- [245] Fleetwood, A.; Hughes, N. A. Convenient Synthesis of 2,3-O-Isopropylidene-5-Thio-D-Ribose and 5-Thio-D-Ribose; Synthesis of 1,4-Anhydro-2,3-O-Isopropylidene- α -D-Ribopyranose and 1,4-Anhydro-2,3-O-Isopropylidene-5-Thio- α -D-Ribopyranose. *Carbohydr. Res.* **1999**, 317, 204–209.
- [246] Bubb, W. A. NMR Spectroscopy in the Study of Carbohydrates: Characterizing the Structural Complexity. *Concepts Magn. Reson. Part A Bridg. Educ. Res.* **2003**, 19 (1), 1–19.
- [247] Finch, P. *Carbohydrates: Structures, Syntheses and Dynamics*; Finch, P., Ed.; Springer Science & Business Media, **1999**.
- [248] Podlasek, C. A.; Stripe, W. A.; Carmichael, I.; Shang, M.; Basu, B.; Serianni, A. S. ^{13}C - ^1H Spin-Coupling

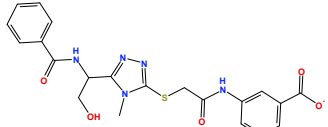
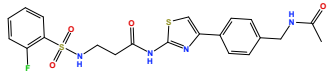
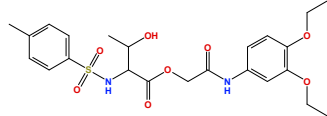
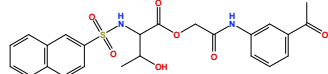
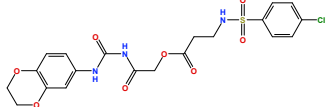
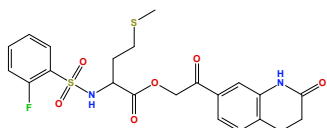
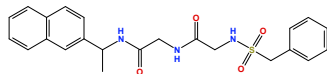
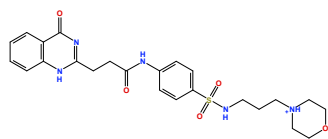
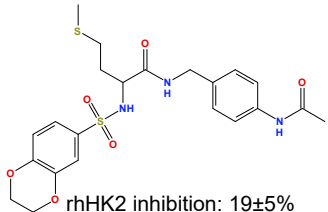
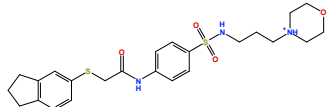
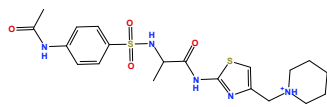
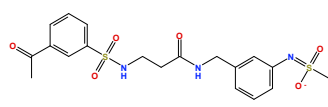
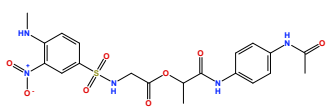
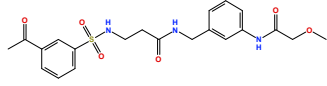
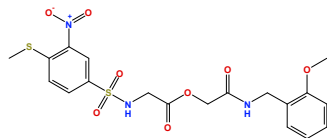
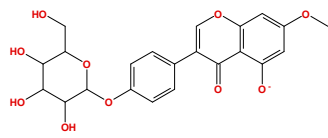
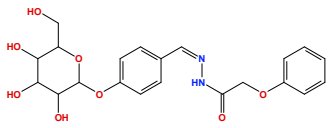
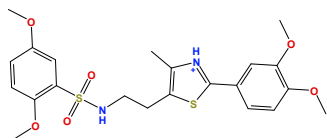
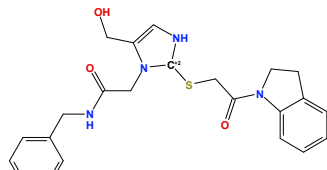
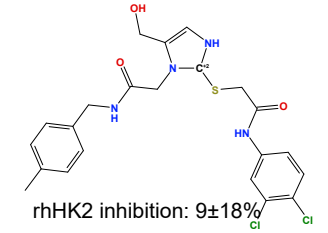
- Constants in the β -D-Ribofuranosyl Ring: Effect of Ring Conformation on Coupling Magnitudes. *J. Am. Chem. Soc.* **1996**, *118* (6), 1413–1425.
- [249] Raap, J.; Boom, J. H. van; Lieshout, H. C. van; Haasnoot, C. A. G. And Resonance Spectroscopy and Molecular Mechanics Study. **1988**, No. 13, 2736–2743.
- [250] Altona, C.; Sundaralingam, M. Conformational Analysis of the Sugar Ring in Nucleosides and Nucleotides. Improved Method for the Interpretation of Proton Magnetic Resonance Coupling Constants. *J. Am. Chem. Soc.* **1973**, *95* (7), 2333–2344.
- [251] Armarego, W. L. F. *Purification of Laboratory Chemicals*; EngineeringPro collection; Butterworth-Heinemann, **2017**.
- [252] Dong, M.; Horitani, M.; Dzikovski, B.; Pandelia, M.-E.; Krebs, C.; Freed, J. H.; Hoffman, B. M.; Lin, H. An Organometallic Complex Formed by an Unconventional Radical SAM Enzyme. *J Am Chem Soc* **2016**, *138* (31), 9755–9758.
- [253] McKinnon, K. M. Flow Cytometry: An Overview. *Curr. Protoc. Immunol.* **2018**, *2018* (February), 5.1.1-5.1.11.
- [254] Lewis, S. W. S. W.; Barni, F. Luminol. In *Wiley Encyclopedia of Forensic Science*; American Cancer Society, **2009**; pp 3–10.
- [255] Stott, D. I. Immunoblotting and Dot Blotting. *J. Immunol. Methods* **1989**, *119* (2), 153–187.
- [256] Agilent Technologies Inc. Agilent Seahorse XF Glycolysis Stress Test Kit - User Guide Kit 103020-100. Wilmington, DE, USA **2019**.
- [257] Agilent Technologies Inc. Agilent Seahorse XF Cell Mito Stress Test Kit - User Guide Kit 103015-100. Wilmington, DE, USA **2019**.
- [258] Lorkiewicz, P. K.; Gibb, A. A.; Rood, B. R.; He, L.; Zheng, Y.; Clem, B. F.; Zhang, X.; Hill, B. G. Integration of Flux Measurements and Pharmacological Controls to Optimize Stable Isotope-Resolved Metabolomics Workflows and Interpretation. *Sci. Rep.* **2019**, *9* (1), 1–17.
- [259] O'Rear, J. L.; Scocca, J. R.; Walker, B. K.; Kaiden, A.; Krag, S. S. Chinese Hamster Ovary Cells with Reduced Hexokinase Activity Maintain Normal GDP-Mannose Levels. *J. Cell. Biochem.* **1999**, *72* (1), 56–66.
- [260] Blaha, C.; Ramakrishnan, G. G.; Jeon, S.-M.; Nogueira, V.; Rho, H.; Kang, S.; Bhaskar, P.; Terry, A. R.; Aissa, A. F.; Frolov, M. V.; et al. A Novel Non-Catalytic Scaffolding Activity of Hexokinase 2 Contributes to EMT and Metastasis. *bioRxiv* **2021**, 2021.04.08.439049.
- [261] Wicks, R. T.; Azadi, J.; Mangraviti, A.; Zhang, I.; Hwang, L.; Joshi, A.; Bow, H.; Hutt-Cabezas, M.; Martin, K. L.; Rudek, M. A.; et al. Local Delivery of Cancer-Cell Glycolytic Inhibitors in High-Grade Glioma. *Neuro. Oncol.* **2015**, *17* (1), 70–80.
- [262] Marisi, G.; Cucchetti, A.; Ulivi, P.; Canale, M.; Cabibbo, G.; Solaini, L.; Foschi, F. G.; De Matteis, S.; Ercolani, G.; Valgiusti, M.; et al. Ten Years of Sorafenib in Hepatocellular Carcinoma: Are There Any Predictive and/or Prognostic Markers? *World J. Gastroenterol.* **2018**, *24* (36), 4152–4163.
- [263] Chu, D. B.; Troyer, C.; Mairinger, T.; Ortmayr, K.; Neubauer, S.; Koellensperger, G.; Hann, S. Isotopologue Analysis of Sugar Phosphates in Yeast Cell Extracts by Gas Chromatography Chemical Ionization Time-of-Flight Mass Spectrometry. *Anal. Bioanal. Chem.* **2015**, *407* (10), 2865–2875.
- [264] Rende, U.; Niittylä, T.; Moritz, T. Two-Step Derivatization for Determination of Sugar Phosphates in Plants by Combined Reversed Phase Chromatography/Tandem Mass Spectrometry. *Plant Methods* **2019**, *15* (1), 1–10.
- [265] Xiao, W.; Wang, R. S.; Handy, D. E.; Loscalzo, J. NAD(H) and NADP(H) Redox Couples and Cellular Energy Metabolism. *Antioxidants Redox Signal.* **2018**, *28* (3), 251–272.
- [266] Heinecke, A.; Eckhardt, W.; Horsch, M.; Bungartz, H.-J. Molecular Dynamics Simulation. In *Supercomputing for Molecular Dynamics Simulations*; Springer International Publishing, **2015**; pp 11–29.
- [267] Sliwoski, G.; Kothiwale, S.; Meiler, J.; Lowe Jr., E. W. Computational Methods in Drug Discovery. *Pharmacol Rev* **2014**, *66* (1), 334–395.
- [268] Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindah, E. Gromacs: High Performance Molecular Simulations through Multi-Level Parallelism from Laptops to Supercomputers. *SoftwareX* **2015**, *1–2*, 19–25.

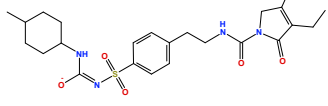
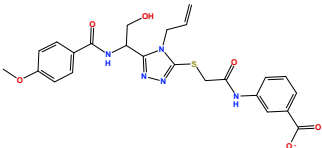
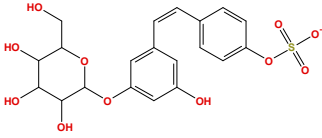
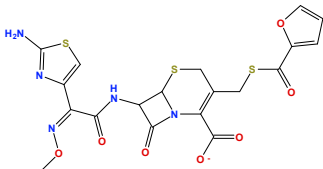
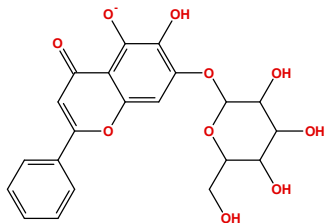
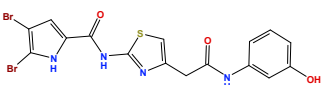
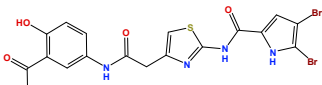
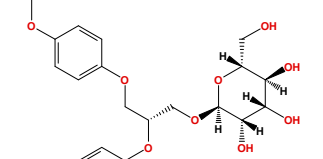
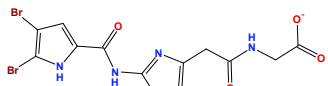
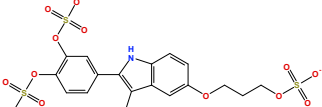
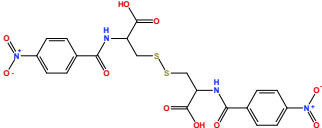
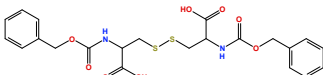
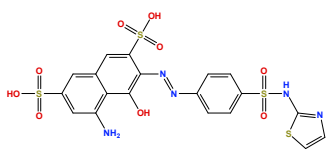
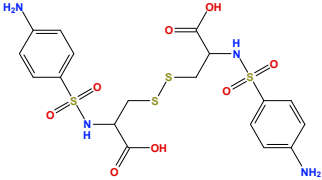
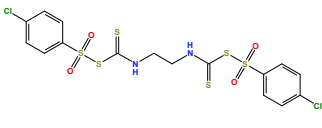
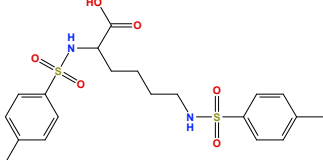
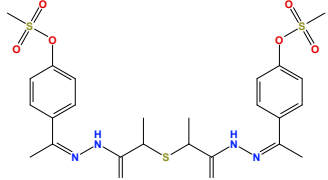
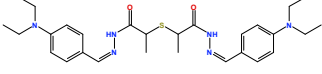
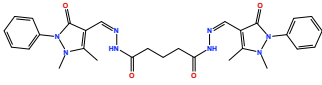
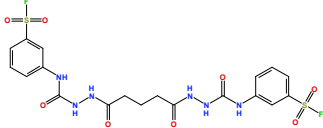
-
- [269] Makarewicz, T.; Kaźmierkiewicz, R. Improvements in GROMACS Plugin for PyMOL Including Implicit Solvent Simulations and Displaying Results of PCA Analysis. *J. Mol. Model.* **2016**, *22* (5), 109.
- [270] Van Der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. C. GROMACS: Fast, Flexible, and Free. *J. Comput. Chem.* **2005**, *26* (16), 1701–1718.
- [271] Ferreira, J. C.; Khrtli, A.; Shetler, C. L.; Mansoor, S.; Sensoy, O.; Rabeh, W. M.; Ali, L.; Sensoy, O.; Rabeh, W. M. Linker Residues Regulate the Activity and Stability of Hexokinase 2, a Promising Anticancer Target. *J. Biol. Chem.* **2020**, *jbc.RA120*.
- [272] Salentin, S.; Schreiber, S.; Haupt, V. J.; Adasme, M. F.; Schroeder, M. PLIP: Fully Automated Protein-Ligand Interaction Profiler. *Nucleic Acids Res.* **2015**, *43* (W1), W443–W447.
- [273] Stroet, M.; Caron, B.; Visscher, K. M.; Geerke, D. P.; Malde, A. K.; Mark, A. E. Automated Topology Builder Version 3.0: Prediction of Solvation Free Enthalpies in Water and Hexane. *J. Chem. Theory Comput.* **2018**, *14* (11), 5834–5845.
- [274] Schmid, N.; Eichenberger, A. P.; Choutko, A.; Riniker, S.; Winger, M.; Mark, A. E.; Van Gunsteren, W. F. Definition and Testing of the GROMOS Force-Field Versions 54A7 and 54B7. *Eur. Biophys. J.* **2011**, *40* (7), 843–856.
- [275] Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A Smooth Particle Mesh Ewald Method. *J. Chem. Phys.* **1995**, *103* (19), 8577–8593.
- [276] Darden, T.; York, D.; Pedersen, L. Particle Mesh Ewald: An N-log(N) Method for Ewald Sums in Large Systems. *J. Chem. Phys.* **1993**, *98* (12), 10089–10092.
- [277] Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual Molecular Dynamics. *J. Mol. Graph.* **1996**, *14* (1), 27–28,33–38.
- [278] Tannock, I. F.; Hickman, J. A. Limits to Personalized Cancer Medicine. *N. Engl. J. Med.* **2016**, *375* (13), 1289–1294.
- [279] Lee, Y. T.; Tan, Y. J.; Oon, C. E. Molecular Targeted Therapy: Treating Cancer with Specificity. *Eur. J. Pharmacol.* **2018**, *834* (January), 188–196.
- [280] Marquart, J.; Chen, E. Y.; Prasad, V. Estimation of the Percentage of US Patients with Cancer Who Benefit from Genome-Driven Oncology. *JAMA Oncol.* **2018**, *4* (8), 1093–1098.
- [281] Kazmierkiewicz, R. *Introduction to Molecular Modeling*; Gdansk, **2011**.
- [282] Wassermann, A. M.; Lounkine, E.; Hoepfner, D.; Le Goff, G.; King, F. J.; Studer, C.; Peltier, J. M.; Grippo, M. L.; Prindle, V.; Tao, J.; et al. Dark Chemical Matter as a Promising Starting Point for Drug Lead Discovery. *Nat. Chem. Biol.* **2015**, *11* (12), 958–966.
- [283] J. Richardson, S.; Bai, A.; A. Kulkarni, A.; F. Moghaddam, M. Efficiency in Drug Discovery: Liver S9 Fraction Assay As a Screen for Metabolic Stability. *Drug Metab. Lett.* **2016**, *10* (2), 83–90.

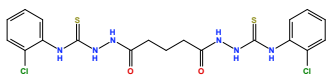
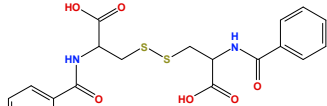
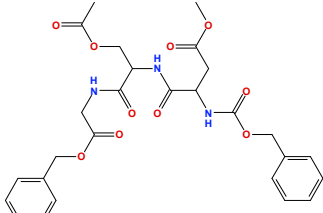
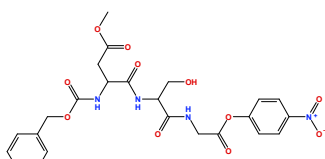
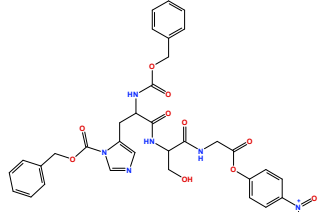
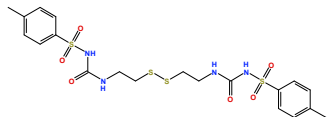
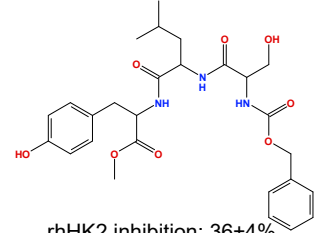
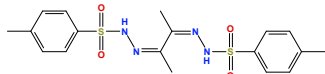
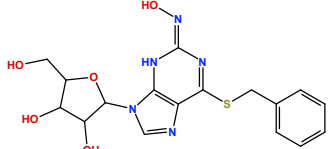
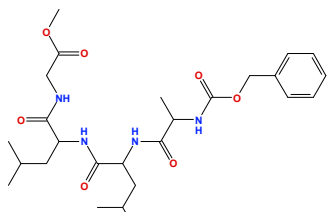
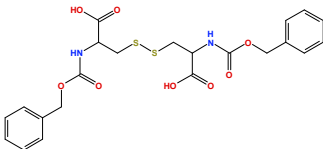
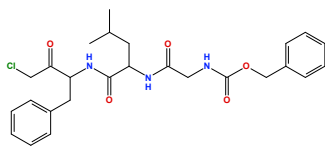
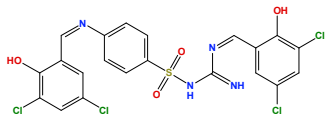
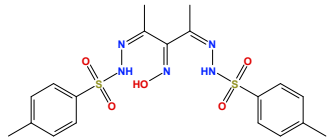
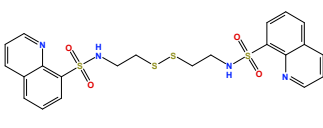
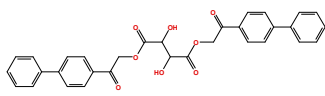
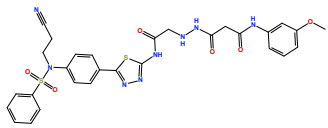
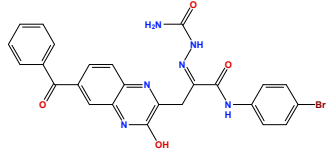
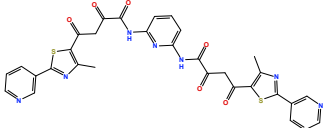
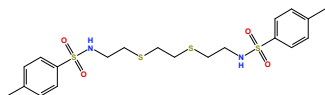
Annexes

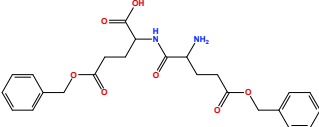
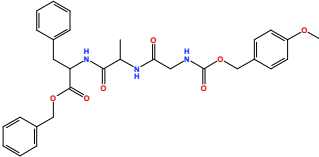
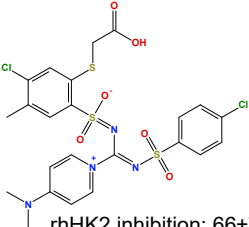
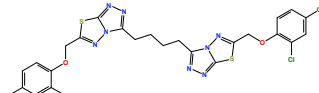
Annex 1. Structures of several glycolytic inhibitors (*cf* chapter II, section 1.2).



<p>CODE: 1-mol-SG</p>  <p>rhHK2 inhibition: 26\pm9%</p> <p>Score: 92 MW: 454.4870</p>	<p>CODE: 2-mol-SG</p>  <p>rhHK2 inhibition: 28\pm14%</p> <p>Score: 94 MW: 476.5530</p>	<p>CODE: 3-mol</p>  <p>rhHK2 inhibition: 4\pm4%</p> <p>Score: 90 MW: 494.5650</p>	<p>CODE: 4-mol</p>  <p>rhHK2 inhibition: 31\pm19%</p> <p>Score: 91 MW: 484.5290</p>
<p>CODE: 5-mol-SG</p>  <p>rhHK2 inhibition: 28\pm7%</p> <p>Score: 91 MW: 497.9120</p>	<p>CODE: 6-mol</p>  <p>rhHK2 inhibition: 30\pm8%</p> <p>Score: 92 MW: 496.5360</p>	<p>CODE: 7-mol</p>  <p>rhHK2 inhibition: 25\pm12%</p> <p>Score: 90 MW: 439.5360</p>	<p>CODE: 8-mol</p>  <p>rhHK2 inhibition: 6\pm4%</p> <p>Score: 93 MW: 500.6000</p>
<p>CODE: 9-mol</p>  <p>rhHK2 inhibition: 19\pm5%</p> <p>Score: 91 MW: 493.6050</p>	<p>CODE: 10-mol</p>  <p>rhHK2 inhibition: 35\pm4%</p> <p>Score: 97 MW: 490.6690</p>	<p>CODE: 11-mol</p>  <p>rhHK2 inhibition: 17\pm10%</p> <p>Score: 95 MW: 466.6070</p>	<p>CODE: 12-mol</p>  <p>rhHK2 inhibition: 24\pm8%</p> <p>Score: 91 MW: 452.5320</p>
<p>CODE: 13-mol-SG</p>  <p>rhHK2 inhibition: 16\pm8%</p> <p>Score: 90 MW: 493.4970</p>	<p>CODE: 14-mol-SG</p>  <p>rhHK2 inhibition: 34\pm6%</p> <p>Score: 90 MW: 447.5120</p>	<p>CODE: 15-mol-SG</p>  <p>rhHK2 inhibition: 31\pm8%</p> <p>Score: 93 MW: 483.5220</p>	<p>CODE: 16-mol-SG</p>  <p>rhHK2 inhibition: 23\pm13%</p> <p>Score: 96 MW: 446.4080</p>
<p>CODE: 17-mol</p>  <p>rhHK2 inhibition: 29\pm15%</p> <p>Score: 91 MW: 432.4290</p>	<p>CODE: 18-mol</p>  <p>rhHK2 inhibition: 27\pm7%</p> <p>Score: 92 MW: 479.5980</p>	<p>CODE: 19-mol</p>  <p>rhHK2 inhibition: 23\pm8%</p> <p>Score: 98 MW: 450.5630</p>	<p>CODE: 20-mol</p>  <p>rhHK2 inhibition: 9\pm18%</p> <p>Score: 101 MW: 493.4150</p>

<p>CODE: 21-mol</p>  <p>rhHK2 inhibition: 31\pm12% Score: 98 MW: 489.6170</p>	<p>CODE: 22-mol</p>  <p>rhHK2 inhibition: 31\pm7% Score: 102 MW: 510.5510</p>	<p>CODE: 23-mol</p>  <p>rhHK2 inhibition: 58\pm4% Score: 95 MW: 469.4430</p>	<p>CODE: 25-mol</p>  <p>rhHK2 inhibition: 40\pm4% Score: 87 MW: 522.5630</p>
<p>CODE: 28-mol</p>  <p>Score: 95 MW: 431.3730 rhHK2 inhibition: 49\pm12%</p>	<p>CODE: 1-Chem</p>  <p>rhHK2 inhibition: 43\pm23% Score: 85 MW: 500.1710</p>	<p>CODE: 2-Chem</p>  <p>rhHK2 inhibition: 47\pm9% Score: 86 MW: 543.1720</p>	<p>CODE: 3-Chem</p>  <p>rhHK2 inhibition: 53\pm11% Score: 80 MW: 466.4830</p>
<p>CODE: 4-Chem</p>  <p>rhHK2 inhibition: 52\pm3% Score: 80 MW: 465.1020</p>	<p>CODE: 5-Chem</p>  <p>rhHK2 inhibition: 56\pm5% Score: 98 MW: 550.5180</p>	<p>CODE: 1-NCI</p>  <p>rhHK2 inhibition: 40\pm10% Score: 89 MW: 538.5140</p>	<p>CODE: 2-NCI</p>  <p>rhHK2 inhibition: 44\pm8% Score: 93 MW: 508.5720</p>
<p>CODE: 3-NCI</p>  <p>rhHK2 inhibition: 60\pm7% Score: 91 MW: 585.6190</p>	<p>CODE: 4-NCI</p>  <p>rhHK2 inhibition: 33\pm8% Score: 90 MW: 550.6580</p>	<p>CODE: 5-NCI</p>  <p>rhHK2 inhibition: 37\pm13% Score: 94 MW: 561.6000</p>	<p>CODE: 6-NCI</p>  <p>rhHK2 inhibition: 28\pm8% Score: 89 MW: 454.5680</p>
<p>CODE: 8-NCI</p>  <p>rhHK2 inhibition: 47\pm4% Score: 93 MW: 598.7220</p>	<p>CODE: 9-NCI</p>  <p>rhHK2 inhibition: 23\pm11% Score: 94 MW: 524.7340</p>	<p>CODE: 10-NCI</p>  <p>rhHK2 inhibition: 44\pm12% Score: 88 MW: 556.6270</p>	<p>CODE: 11-NCI</p>  <p>rhHK2 inhibition: 43\pm13% Score: 92 MW: 562.5310</p>

<p>CODE: 12-NCI</p>  <p>rhHK2 inhibition: 27±5% Score: 88 MW: 499.4470</p>	<p>CODE: 13-NCI</p>  <p>rhHK2 inhibition: 30±12% Score: 88 MW: 448.5200</p>	<p>CODE: 15-NCI</p>  <p>rhHK2 inhibition: 34±7% Score: 95 MW: 557.5560</p>	<p>CODE: 16-NCI</p>  <p>rhHK2 inhibition: 40±10% Score: 95 MW: 546.4890</p>
<p>CODE: 17-NCI</p>  <p>rhHK2 inhibition: 42±8% Score: 95 MW: 688.6500</p>	<p>CODE: 18-NCI</p>  <p>rhHK2 inhibition: 35±1% Score: 95 MW: 546.7140</p>	<p>CODE: 19-NCI</p>  <p>rhHK2 inhibition: 36±4% Score: 88 MW: 529.5900</p>	<p>CODE: 20-NCI</p>  <p>rhHK2 inhibition: 43±10% Score: 88 MW: 422.5300</p>
<p>CODE: 21-NCI</p>  <p>rhHK2 inhibition: 36±4% Score: 88 MW: 405.4350</p>	<p>CODE: 22-NCI</p>  <p>rhHK2 inhibition: 50±6% Score: 91 MW: 520.6270</p>	<p>CODE: 23-NCI</p>  <p>rhHK2 inhibition: 38±5% Score: 90 MW: 508.5720</p>	<p>CODE: 24-NCI</p>  <p>rhHK2 inhibition: 44±6% Score: 89 MW: 502.0110</p>
<p>CODE: 26-NCI</p>  <p>rhHK2 inhibition: 30±8% Score: 94 MW: 560.2450</p>	<p>CODE: 27-NCI</p>  <p>rhHK2 inhibition: 33±7% Score: 91 MW: 465.5550</p>	<p>CODE: 28-NCI</p>  <p>rhHK2 inhibition: 51±1% Score: 89 MW: 534.7060</p>	<p>CODE: 29-NCI</p>  <p>rhHK2 inhibition: 42±11% Score: 94 MW: 538.5520</p>
<p>CODE: 30-NCI</p>  <p>rhHK2 inhibition: 47±9% Score: 95 MW: 648.7250</p>	<p>CODE: 31-nci</p>  <p>rhHK2 inhibition: 56±9% Score: 93 MW: 547.3690</p>	<p>CODE: 32-NCI</p>  <p>rhHK2 inhibition: 45±5% Score: 97 MW: 653.7000</p>	<p>CODE: 33-NCI</p>  <p>rhHK2 inhibition: 39±5% Score: 90 MW: 488.7180</p>

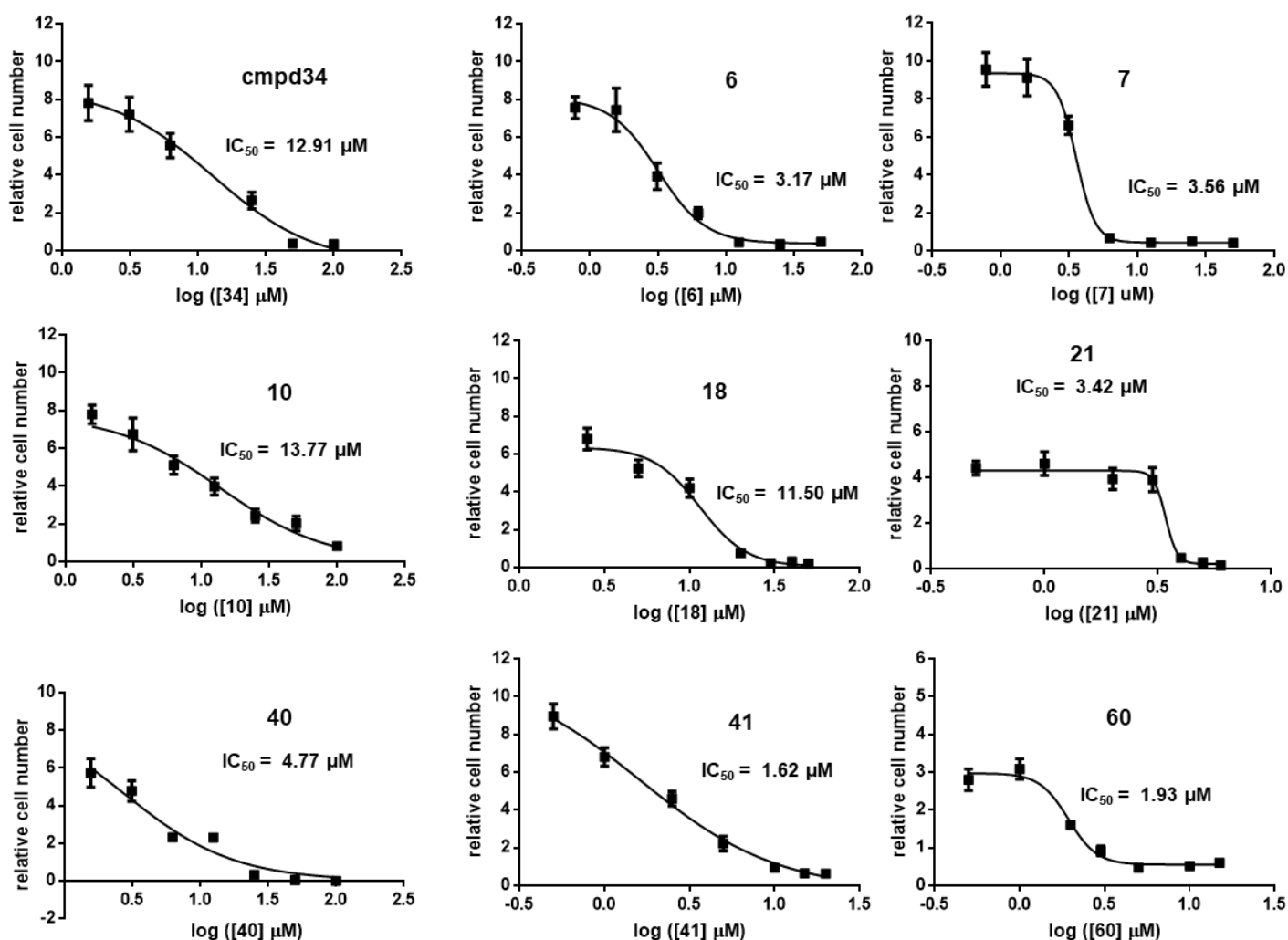
<p>CODE: 34-NCI</p>  <p>rhHK2 inhibition: 52±5% Score: 93 MW: 456.4950</p>	<p>CODE: 35-NCI</p>  <p>rhHK2 inhibition: 55±2% Score: 93 MW: 547.6080</p>	<p>CODE: 36-NCI</p>  <p>rhHK2 inhibition: 66±7% Score: 101 MW: 617.5550</p>	<p>CODE: 37-NCI</p>  <p>rhHK2 inhibition: 50±3% Score: 95 MW: 656.4060</p>
---	---	---	---

Annex 3. Correspondence between code number implemented in Chapter V and the molecule code defined previously (**annex 2, Chapter III, Chapter IV**).

Chapter V number	Molecule code ¹	Chapter V number	Molecule code	Chapter V number	Molecule code
1	19 ²	22	11-NCI	43	8 ²
2	23 ²	23	27-NCI	44	14 ²
3	"cmpd 25"	24	10-NCI	45	21 ²
4	4-chem	25	12-mol	46	9 ²
5	36-NCI	26	1-mol	47	22 ²
6	RuGly21 ³	27	3-mol	48	10 ²
7	RuGly22 ³	28	25-mol	49	16 ¹
8	3-chem	29	6-mol	50	6 ²
9	5-chem	30	2-mol	51	12 ²
10	3-NCI	31	8-mol	52	18 ²
11	22-NCI	32	5-mol	53	20 ²
12	23-NCI	33	21-mol	54	24 ²
13	35-NCI	34	32-NCI	55	Gly21 ⁴
14	16-mol	35	6-NCI	56	Gly22 ⁴
15	17-mol	36	7 ²	57	1-chem
16	20-mol	37	11 ²	58	2-chem
17	21-NCI	38	13 ²	59	2-NCI
18	28-NCI	39	17 ²	60	24-NCI
19	19-NCI	40	37-NCI	61	18-NCI
20	1-NCI	41	5-NCI	62	15-NCI
21	9-NCI	42	13-mol	63	19-NCI

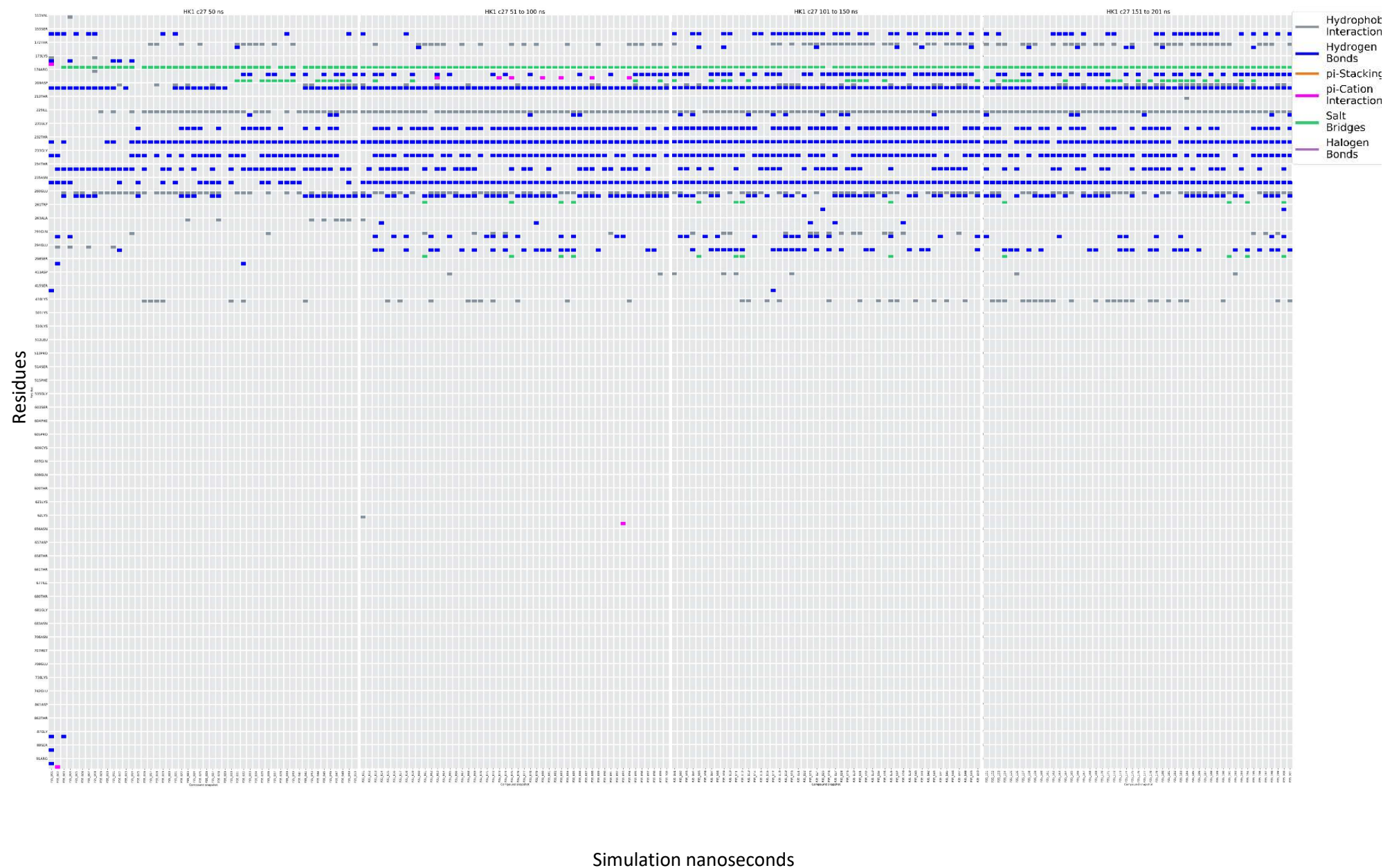
¹ Molecule code was defined after molecules acquisition or synthesis; Structures of these molecules are found in annex 3. ²These molecules were synthesized, structures are found in **chapter IV, section 7.2, schemes 12-32**; Structure of "cmpd 25" is found on **chapter IV, section 7.3, scheme 40**. ³ These molecules are described on **chapter III**, where structures are found (**section 2.6.3, figure 32**). ⁴ These molecules are described on chapter III as the carbohydrate ligands of molecules RuGly21 and RuGly22 (**section 2.6.3, figure 32**).

Annex 4. Dose-response curves used for the calculation of IC_{50} for the best inhibitors (“**cmpd 34**”, **6**, **7**, **10**, **18**, **21**, **40**, **41**, **60**) of Huh7 cells proliferation after 6 days of treatment.



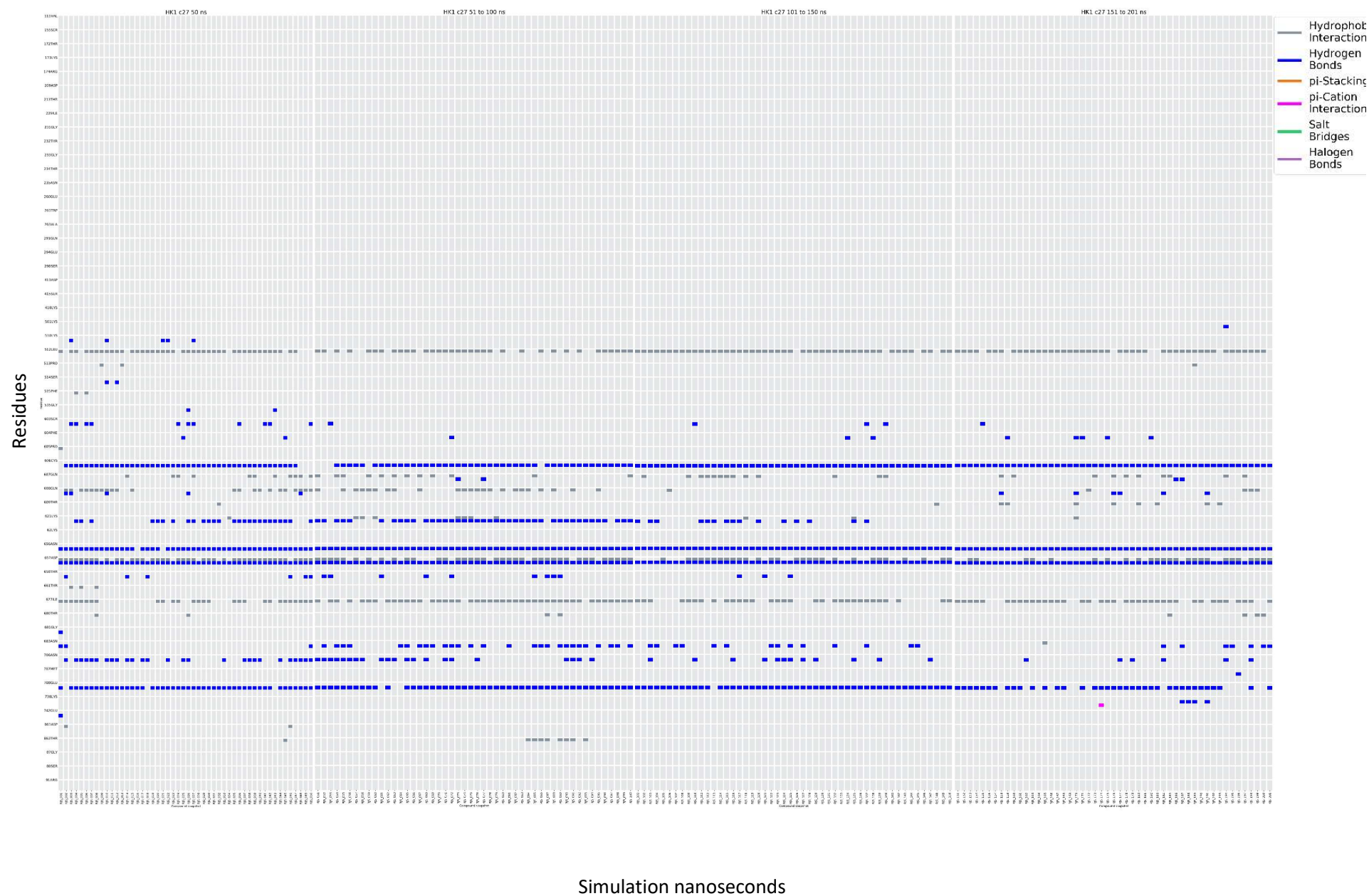
Annex 5. Protein-ligand interactions fingerprint (PLIF) plots for HK1 and HK2 with “cmpd 27”, “cmpd 30” and “cmpd 33” in both *N*-terminal and C-terminal domain.

HK1-“cmpd 27” (NTD)

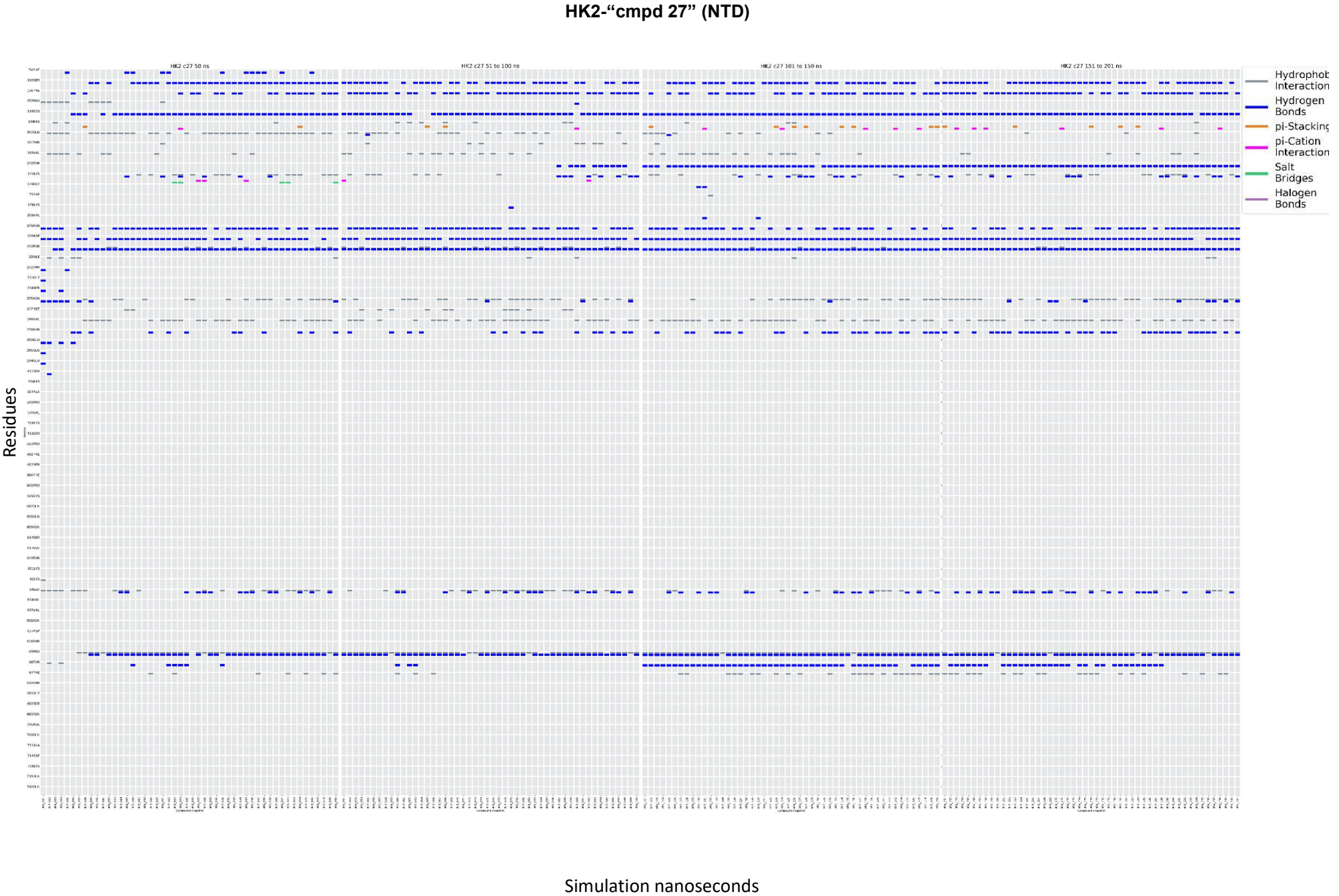


Annex 5. Protein-ligand interactions fingerprint (PLIF) plots for HK1 and HK2 with “cmpd 27”, “cmpd 30” and “cmpd 33” in both *N*-terminal and C-terminal domain.

HK1-“cmpd 27” (CTD)

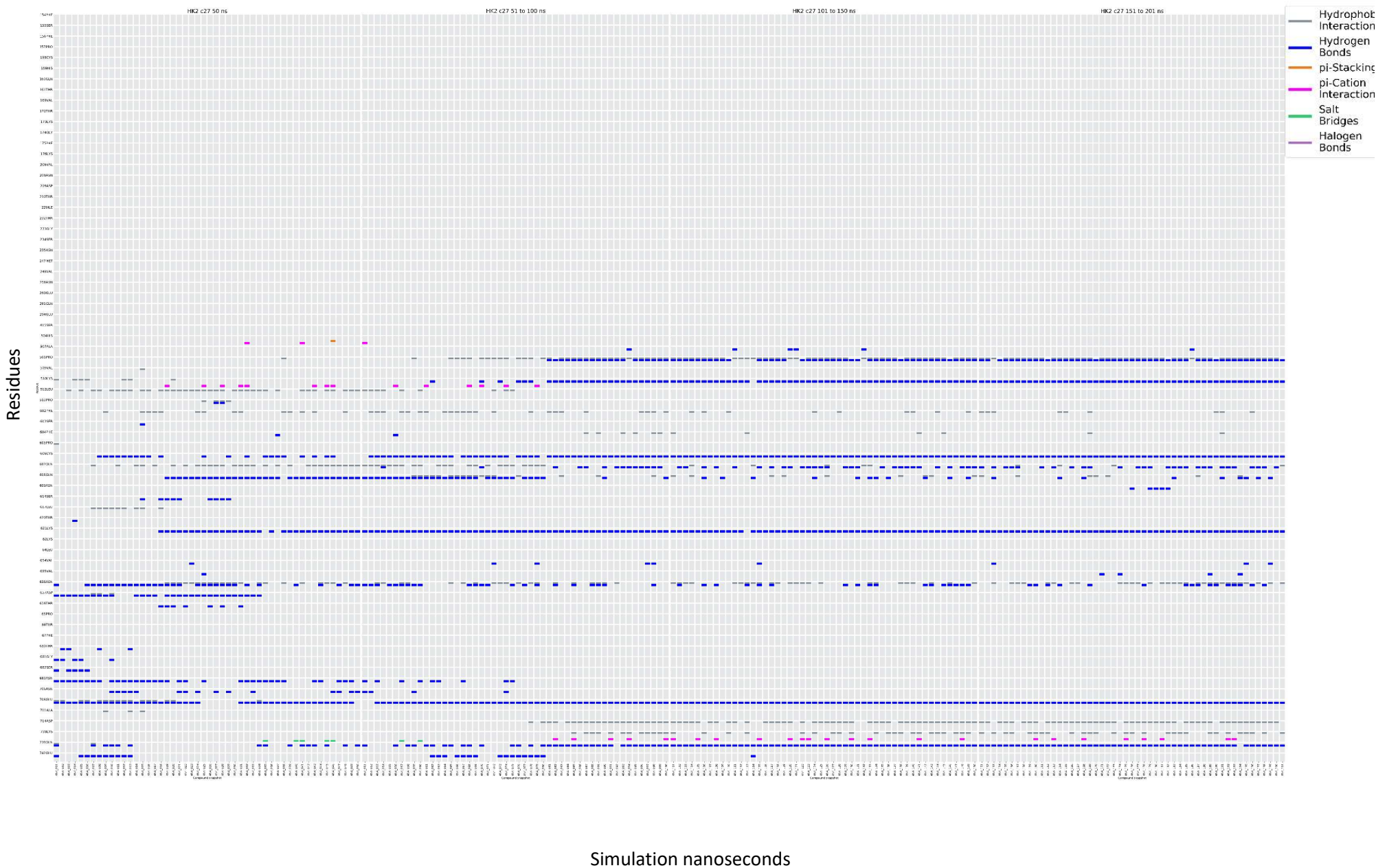


Annex 5. Protein-ligand interactions fingerprint (PLIF) plots for HK1 and HK2 with “cmpd 27”, “cmpd 30” and “cmpd 33” in both N-terminal and C-terminal domain.



Annex 5. Protein-ligand interactions fingerprint (PLIF) plots for HK1 and HK2 with “cmpd 27”, “cmpd 30” and “cmpd 33” in both N-terminal and C-terminal domain.

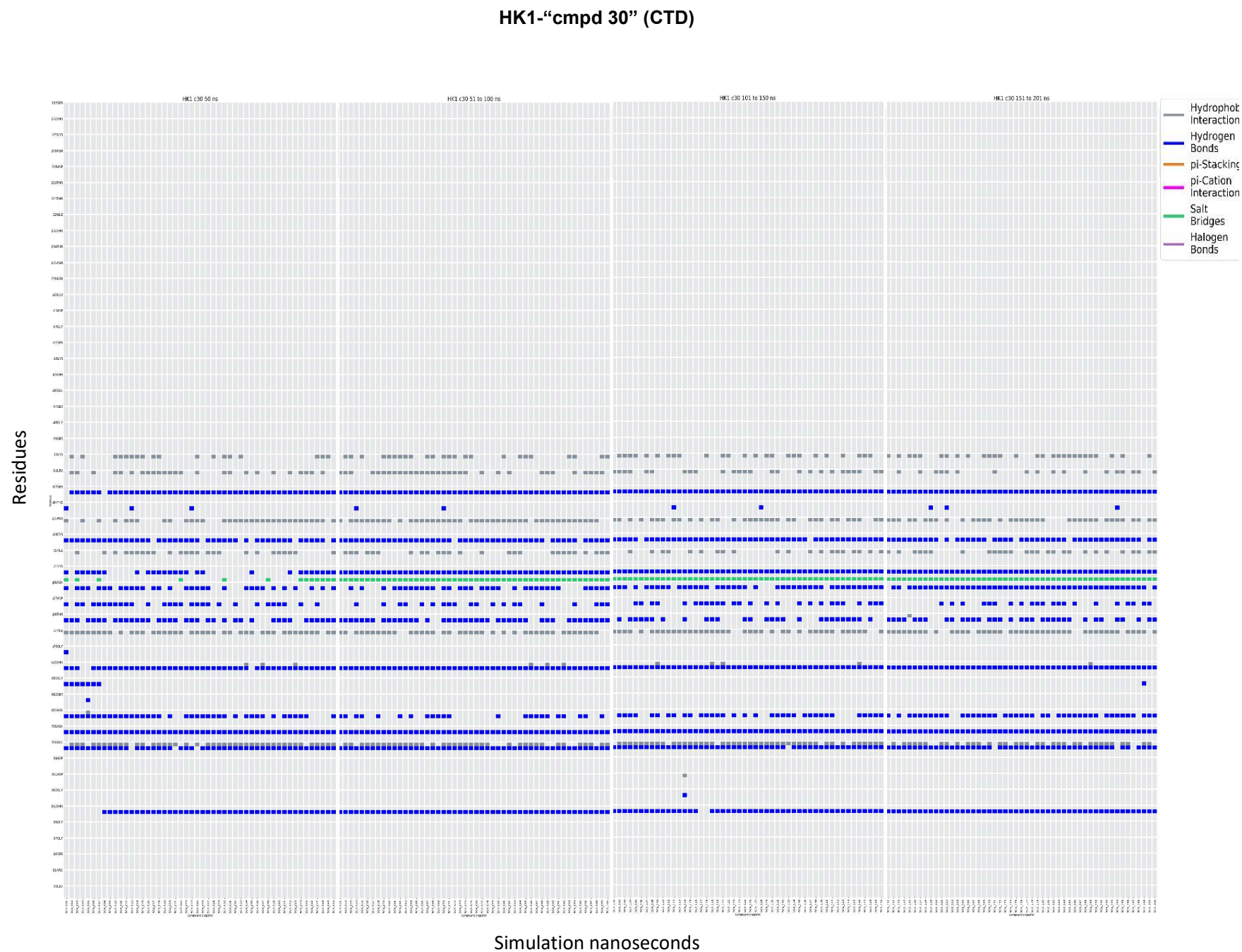
HK2-“cmpd 27” (CTD)



Annex 5. Protein-ligand interactions fingerprint (PLIF) plots for HK1 and HK2 with “cmpd 27”, “cmpd 30” and “cmpd 33” in both *N*-terminal and C-terminal domain.



Annex 5. Protein-ligand interactions fingerprint (PLIF) plots for HK1 and HK2 with “cmpd 27”, “cmpd 30” and “cmpd 33” in both N-terminal and C-terminal domain.

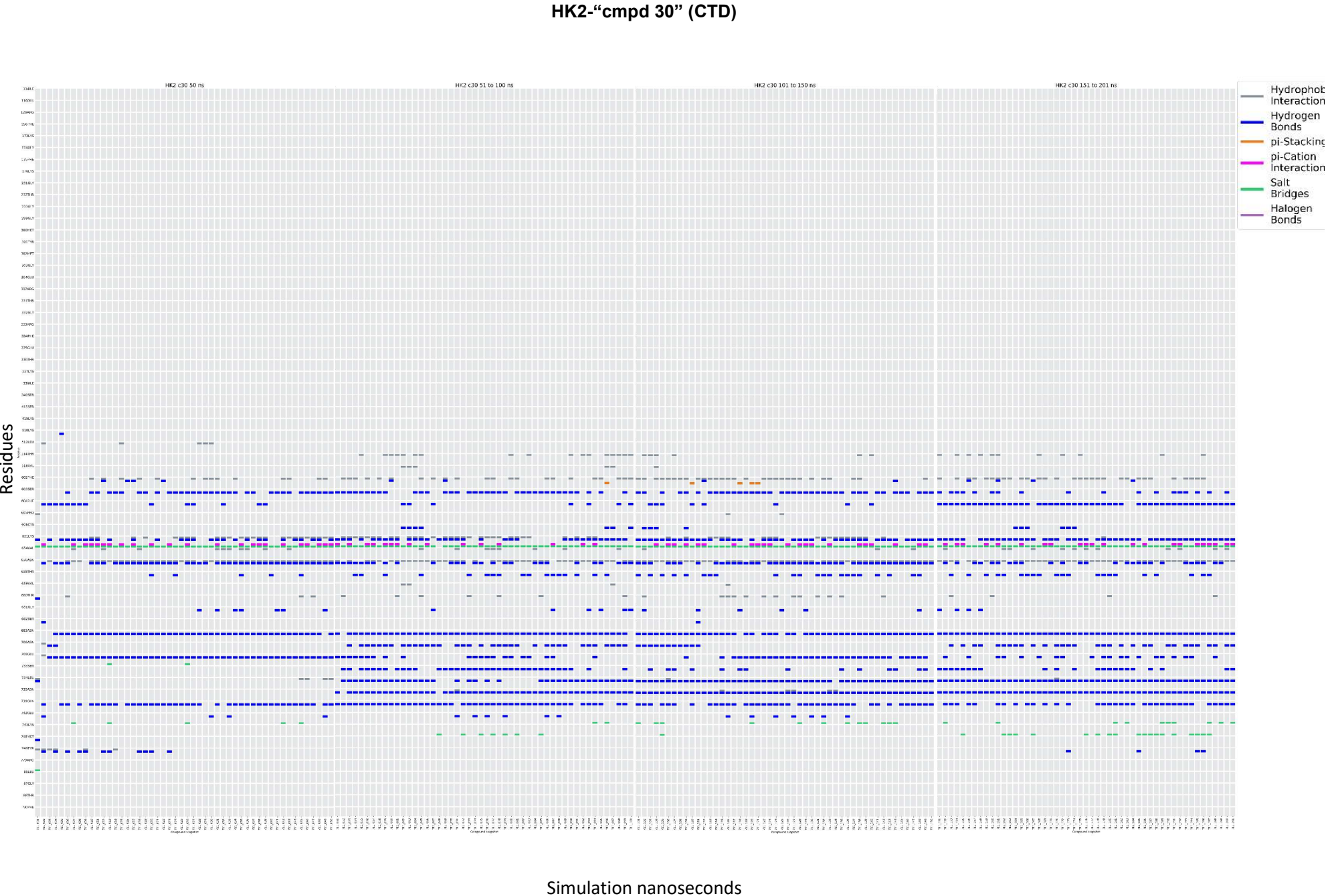


Annex 5. Protein-ligand interactions fingerprint (PLIF) plots for HK1 and HK2 with “compd 27”, “compd 30” and “compd 33” in both *N*-terminal and *C*-terminal domain.

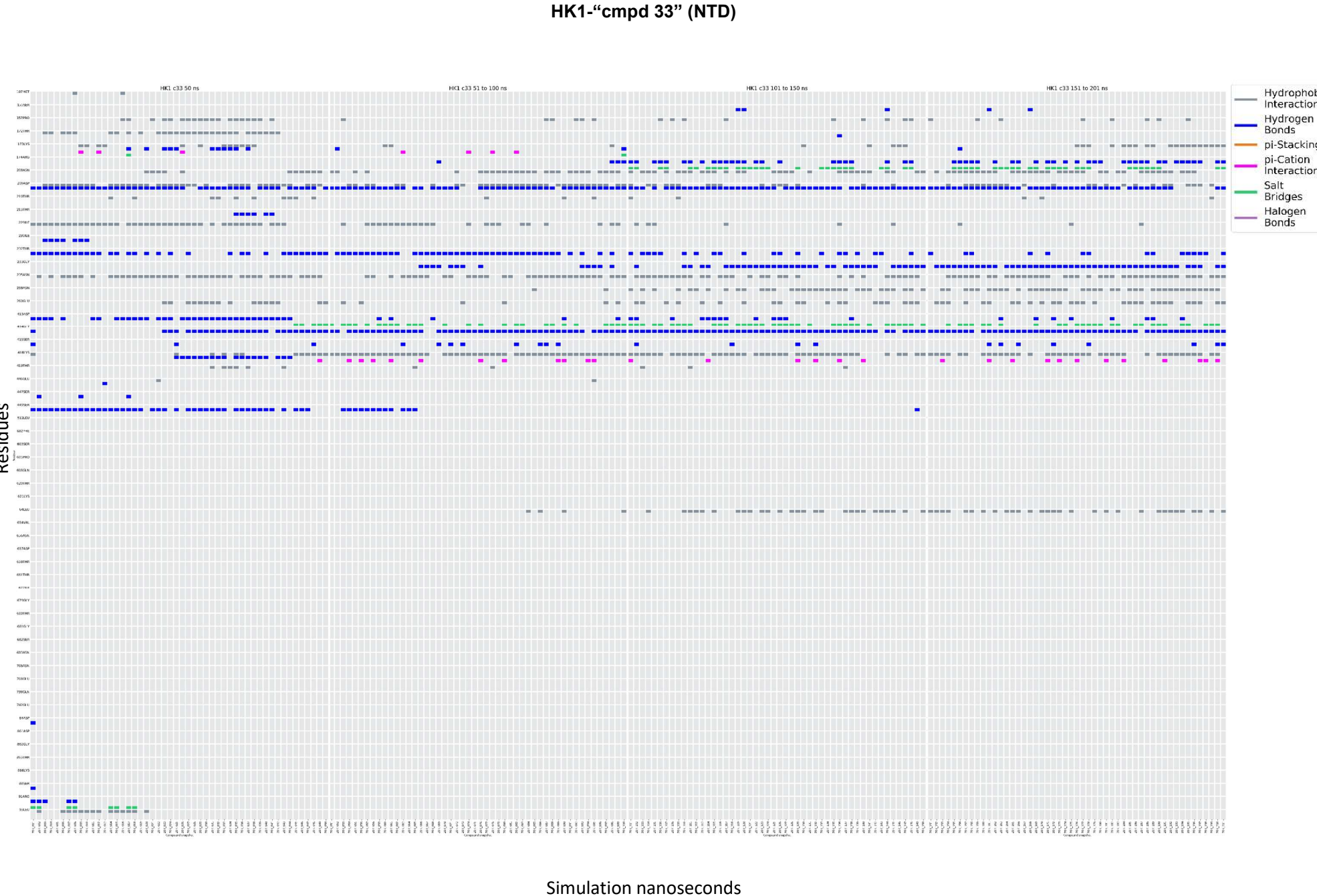
HK2-“cmpd 30” (NTD)



Annex 5. Protein-ligand interactions fingerprint (PLIF) plots for HK1 and HK2 with “cmpd 27”, “cmpd 30” and “cmpd 33” in both N-terminal and C-terminal domain.

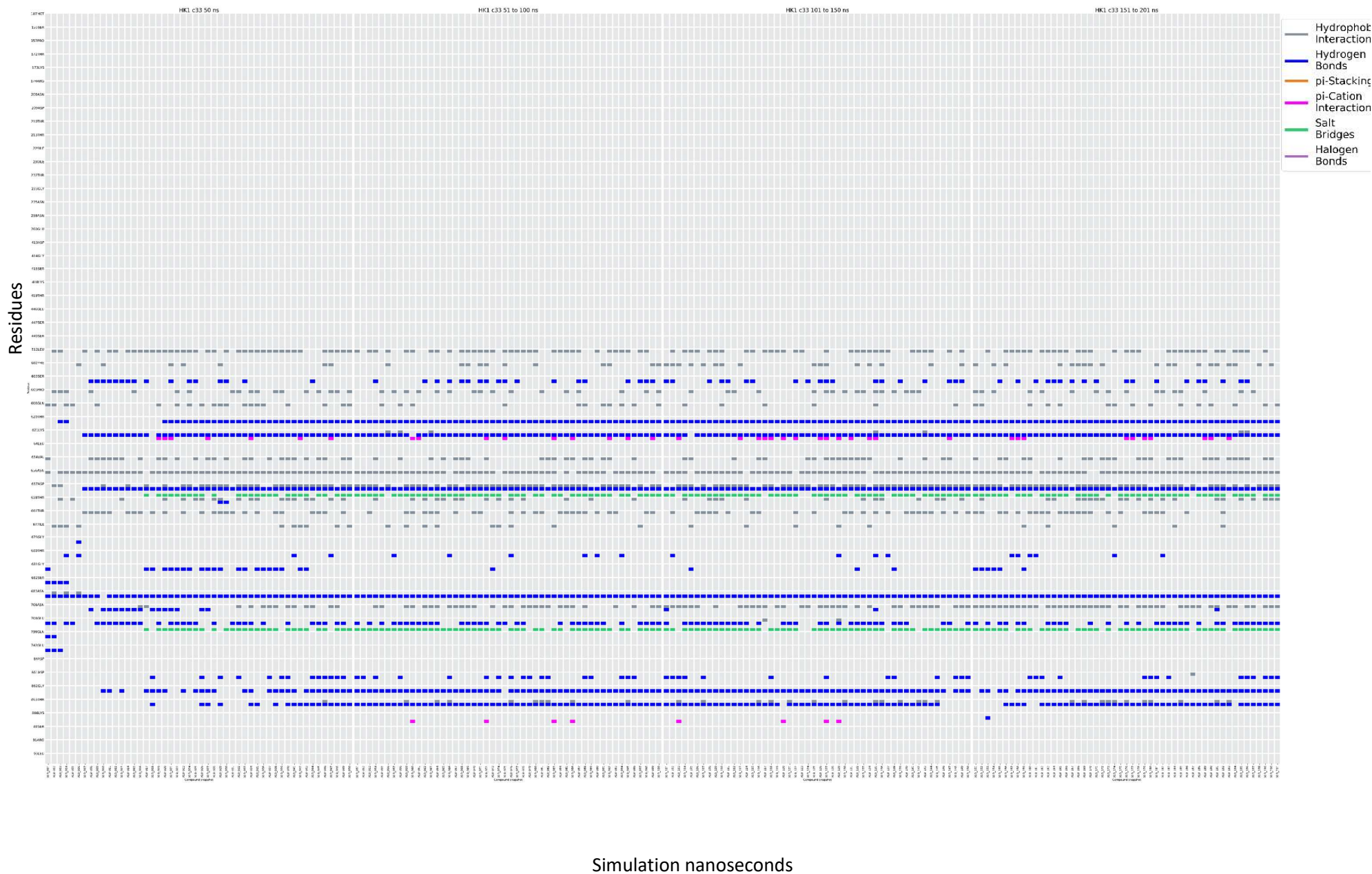


Annex 5. Protein-ligand interactions fingerprint (PLIF) plots for HK1 and HK2 with “cmpd 27”, “cmpd 30” and “cmpd 33” in both *N*-terminal and C-terminal domain.



Annex 5. Protein-ligand interactions fingerprint (PLIF) plots for HK1 and HK2 with “cmpd 27”, “cmpd 30” and “cmpd 33” in both *N*-terminal and C-terminal domain.

HK1-“cmpd 33” (CTD)



Annex 5. Protein-ligand interactions fingerprint (PLIF) plots for HK1 and HK2 with “cmpd 27”, “cmpd 30” and “cmpd 33” in both *N*-terminal and C-terminal domain.

HK2-“cmpd 33” (NTD)

