UNIVERSIDADE DE LISBOA
INSTITUTO SUPERIOR TÉCNICO

Molecular determinants of phosphatidylinositol 4,5-bisphosphate function and organization

Luís Pedro Borges Araújo

Supervisor:	**Doctor Fábio Monteiro Fernandes**
Co-Supervisors:	**Doctor Manuel Nuno de Sousa Pereira Simões de Melo**
	Doctor Manuel José Estevez Prieto

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

Jury final classification: Pass with Distinction and Honour

2022

TÉCNICO LISBOA

UNIVERSIDADE DE LISBOA
INSTITUTO SUPERIOR TÉCNICO

Molecular determinants of phosphatidylinositol 4,5-bisphosphate function and organization

Luís Pedro Borges Araújo

Supervisor:	Doctor Fábio Monteiro Fernandes
Co-Supervisors:	Doctor Manuel Nuno de Sousa Pereira Simões de Melo
	Doctor Manuel José Estevez Prieto

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

Jury final classification: Pass with Distinction and Honour

Jury

Chairperson:	Doctor Mário Nuno de Matos Sequeira Berberan e Santos, Instituto Superior Técnico, Universidade de Lisboa
Members of the Committee:	Doctor Arsénio do Carmo Sales Mendes Fialho, Instituto Superior Técnico, Universidade de Lisboa; Doctor Luis Miguel Santos Loura, Faculdade de Farmácia, Universidade de Coimbra; Doctor Fábio Monteiro Fernandes, Instituto Superior Técnico, Universidade de Lisboa; Doctor David Slochower, Vertex Pharmaceuticals, EUA.

Funding Institution

Fundação para a Ciência e a Tecnologia

2022

Acknowledgments

Science is a major collaborative endeavor. As such, none of the work developed during my PhD studies could have been possible without the support of my colleagues, friends and family. It has been an enormous pleasure to work alongside you. While many are mentioned throughout the text for their scientific contributions, I want to specifically acknowledge a few people.

My excellent supervisors, who have also become good friends:

Fábio Fernandes, who first gave me a chance to prove what I was worth. He took me in and taught me how to be a scientist. He gave me the freedom to follow my own path and make my own mistakes, without ever refusing me advice and guidance when I needed it.

Manuel Melo, who guided me into the world of computational biology with such joy and cheer. He constantly pushed me to challenge myself and leave my comfort zone, while giving me countless opportunities to establish myself and succeed in the field.

Prof. Manuel Prieto, who warmly welcomed me into his group and offered me nothing but support and motivation throughout these years.

All of my colleagues and friends from BSIRG at IBB (Former CQFM), who mentored, assisted and guided me in the stage of my career when I needed it most. Not only did we make excellent science but we also always found time for the craziest of adventures. My fellow colleagues from the Multiscale Modelling, Protein Modelling, Molecular Simulation, and Systems and Synthetic Biology labs who turned sitting down at a computer all day into a fun prospect.

My closest friends, who have always made sure I don't lose what little is left of my sanity and kept me close to my roots. My parents, sister, grandmother, and cat who always offered unconditional support and gave everything they had to make sure I could succeed. Especially the cat.

I must also thank Fundação para a Ciência e a Tecnologia – Ministério da Ciência, Tecnologia e Ensino Superior (FCT-MCTES, Portugal) as well as the Medical Biochemistry and Biophysics Doctoral Programme (M2B-PhD) for the PhD fellowship PD/BD/137492/2018. Also, I express my gratitute to the European Biophysical Societies' Association (EBSA) and the Portuguese biophysics society (SPBf) for sponsoring my participation in several scientific meetings.

Abstract

Phosphatidylinositol 4,5-bisphosphate ($PI(4,5)P_2$) is a minority membrane phospholipid that plays a crucial role in the regulation of several signaling processes at the inner leaflet of the plasma membrane. Lateral organization of $PI(4,5)P_2$ is critical to that end. It is thought that $PI(4,5)P_2$ is heterogeneously dispersed across the membrane, in localized $PI(4,5)P_2$-enriched nanodomains at particular sites and timings. Divalent cations were shown to have the ability to induce the formation of such nanodomains, even at physiological concentrations of both lipid and cation, and thus became key modulators of $PI(4,5)P_2$-lateral organization. Here, we use a combination of fluorescence spectroscopy and microscopy techniques, and coarse grained (CG) molecular dynamics (MD) simulations to study molecular determinants that regulate $PI(4,5)P_2$ lateral organization and their impact on its function, via the modulation of $PI(4,5)P_2$-protein interactions. Our results show that Ca^{2+} is able to directly impact $PI(4,5)P_2$-protein interactions through several mechanisms. Ca^{2+} was shown to induce charge screening of $PI(4,5)P_2$, significantly impacting $PI(4,5)P_2$-protein binding and affinity. $PI(4,5)P_2$-binding proteins were shown to not be able to fully sequester lipids from Ca^{2+}-induced $PI(4,5)P_2$ nanodomains, proving that these have the potential to compartmentalize $PI(4,5)P_2$-binding proteins, inhibiting, or possibly promoting interactions between them through changes in $PI(4,5)P_2$ nanodomain size. The effect of acyl-chain saturation on the calcium-dependent changes on the overall properties of $PI(4,5)P_2$-containing membranes was also reported. While in the absence of divalent cations all acyl-chain compositions behaved in the same manner, in the presence of Ca^{2+}, acyl-chain compositions yielded $PI(4,5)P_2$ nanodomains with significantly different biophysical properties. In particular, increasingly saturated acyl-chains yielded more ordered and structured nanodomains, which culminated in the formation of gel-like nanodomains for the fully saturated composition. The $PI(4,5)P_2$ biophysical properties explored in this work were then used as reference for the development of Martini 3 CG MD topologies of $PI(4,5)P_2$ and the remaining lipids in the phosphoinositide family. These topologies were developed with an expanded application scope to include, among others, accurate reproduction of cation-mediated phosphoinositide aggregation. Altogether, this work furthers the understanding of $PI(4,5)P_2$ lateral organization and hints at how some of its regulators might substantially impact $PI(4,5)P_2$ function.

Keywords: $PI(4,5)P_2$, Calcium, Lipid domains, Membrane lateral organization, lipid-protein interactions, Fluorescence spectroscopy and microscopy, Molecular dynamics simulations, Membrane models.

Resumo

O fosfatidilinositol 4,5-bisfosfato ($PI(4,5)P_2$) é um fosfolípido de membrana minoritário que desempenha um papel crucial na regulação de vários processos de sinalização no folheto interno da membrana plasmática. A organização lateral do $PI(4,5)P_2$ é fundamental para esse fim. Pensa-se que o $PI(4,5)P_2$ está heterogeneamente disperso pela membrana, em nanodomínios enriquecidos em $PI(4,5)P_2$ presentes em locais e tempos determinados. Foi demonstrado que os catiões divalentes têm a capacidade de induzir a formação de tais nanodomínios, mesmo a concentrações fisiológicas de lípido e catião, tornando-se assim em moduladores chave da organização lateral de $PI(4,5)P_2$. Neste trabalho utilizamos uma combinação de espectroscopia e microscopia de fluorescência, e simulações de dinâmica molecular (MD) *coarse grained* (CG) para estudar determinantes moleculares que regulam a organização lateral de $PI(4,5)P_2$, e o seu potencial impacto na sua função através da modulação das interacções $PI(4,5)P_2$-proteína. Os nossos resultados mostram que o Ca^{2+} deverá ter um impacto directo nas interacções $PI(4,5)P_2$-proteína através de vários mecanismos. Foi demonstrado que o Ca^{2+} induz a blindagem da carga do $PI(4,5)P_2$, impactando significativamente a afinidade da ligação $PI(4,5)P_2$-proteína. Foi também demonstrado que as proteínas ligantes de $PI(4,5)P_2$ não são capazes de sequestrar lípidos dos nanodomínios induzidos por Ca^{2+}, provando que estes têm o potencial de compartimentalizar as proteínas ligantes de $PI(4,5)P_2$ e modular interacções entre elas, através de alterações no tamanho dos nanodomínios. O efeito da saturação das cadeias de acilo sobre as alterações dependentes do cálcio nas propriedades globais de membranas contendo $PI(4,5)P_2$ foi também explorado. Enquanto na ausência de catiões divalentes, todas as composições de cadeias de acilo se comportaram da mesma maneira, na presença de Ca^{2+} as várias composições produziram nanodomínios de $PI(4,5)P_2$ com propriedades biofísicas distintas. Em particular, cadeias de acilo cada vez mais saturadas produziram nanodomínios mais ordenados e estruturados, o que culminou na formação de nanodomínios semelhantes a gel para a composição totalmente saturada. As propriedades biofísicas do $PI(4,5)P_2$ exploradas neste trabalho foram também utilizadas como referência para o desenvolvimento de topologias do $PI(4,5)P_2$, e dos restantes lípidos da família dos fosfatidilinositóis, para o campo de forças CG Martini 3. Estas topologias foram desenvolvidas com um âmbito de aplicação alargado para incluir, entre outras, a reprodução exacta da agregação de fosfatidilinositóis mediada por catiões. No seu conjunto, este trabalho expande o conhecimento sobre a organização lateral do $PI(4,5)P_2$ e dá pistas sobre como alguns dos seus reguladores podem ter um impacto substancial na sua função.

Palavras-chave: $PI(4,5)P_2$, Cálcio, Domínios Lipídicos, Organização lateral membranar, Interações lipido-proteína, Espectroscopia e microscopia de fluorescência, Simulações de dinâmica molecular, Modelos de membrana.

Resumo Alargado

O fosfatidilinositol 4,5-bisfosfato ($PI(4,5)P_2$) é um fosfolípido de membrana minoritário que desempenha um papel crucial na regulação de vários processos de sinalização no folheto interno da membrana plasmática. Embora o seu papel como fonte de mensageiros secundários seja conhecido há décadas, ao longo do tempo acumularam-se provas da sua importância como fosfolípido para a definição da identidade membranar em células eucarióticas. Devido à sua grande "cabeça" polar e elevada carga negativa, o $PI(4,5)P_2$ actua quase como um farol electrostático, interagindo com várias entidades moleculares, tais como proteínas de membrana, outros lípidos, catiões celulares, etc. Como resultado das suas propriedades biofísicas particulares, é um regulador importante de um largo espectro de eventos na membrana plasmática, incluindo a adesão e motilidade celular, transporte mediado por canais iónicos, endocitose e exocitose, entre outros.

A organização lateral do $PI(4,5)P_2$ é fundamental para a correcta regulação da sua função. Como está envolvido numa multiplicidade de funções celulares que ocorrem em paralelo, os níveis de $PI(4,5)P_2$ devem ser rigorosamente regulados para evitar flutuações significativas da sua concentração total. Isto implica que a regulação simultânea destas funções pelo $PI(4,5)P_2$ deve ocorrer através da presença de múltiplos nanodomínios enriquecidos em $PI(4,5)P_2$ presentes em locais e tempos determinados. Foi demonstrado que os catiões divalentes, como Ca^{2+} ou Mg^{2+}, têm a capacidade de induzir a formação de tais nanodomínios, mesmo a concentrações fisiológicas de lípido e catião, tornando-se assim em moduladores chave da organização lateral de $PI(4,5)P_2$.

Neste trabalho utilizamos uma combinação de espectroscopia e microscopia de fluorescência, e simulações de dinâmica molecular (MD) *coarse grained* (CG) para estudar determinantes moleculares que regulam a organização lateral de $PI(4,5)P_2$, e o seu potencial impacto na sua função através da modulação das interacções $PI(4,5)P_2$-proteína.

Os nossos resultados mostram que o Ca^{2+} deverá ter um impacto directo nas interacções $PI(4,5)P_2$-proteína através de vários mecanismos. Foi demonstrado que o Ca^{2+} induz a blindagem da carga do $PI(4,5)P_2$, impactando significativamente a afinidade da ligação $PI(4,5)P_2$-proteína. Este efeito sobre a afinidade da ligação mostrou ser dependente do tipo de especificidade e mecanismo da ligação $PI(4,5)P_2$-proteína. Foi também demonstrado que as proteínas ligantes de $PI(4,5)P_2$ não são capazes de sequestrar lípidos dos nanodomínios induzidos por Ca^{2+}, provando que estes têm o potencial de compartimentalizar as proteínas ligantes de $PI(4,5)P_2$ e modular interacções entre elas através de alterações no tamanho dos nanodomínios. Estes efeitos deverão actuar concertadamente, através de flutuações nos niveis de Ca^{2+}, para regular as dinâmicas de muitas proteínas ligantes de $PI(4,5)P_2$ cruciais para várias vias de sinalização.

O efeito da saturação das cadeias de acilo sobre as alterações dependentes do Ca^{2+} nas propriedades globais de membranas contendo $PI(4,5)P_2$ foi também explorado. A composição de cadeias acilo 1-

estearoyl-2-arachidonyl (18:0 20:4), representa até 70% do total de $PI(4,5)P_2$ em algumas linhas celulares, especialmente em tecido neuronal. No entanto, o perfil de distribuição das cadeias acilo de $PI(4,5)P_2$ ainda é amplo, desde cadeias altamente insaturadas até cadeias totalmente saturadas. Estas espécies menos abundantes tornam-se mais prevalentes em resposta a certos estímulos, stress, envelhecimento ou em situações de cancro. Nestes casos, a concentração das composições totalmente saturadas e mono-insaturadas aumenta significativamente, podendo ultrapassar a composição canónica de 18:0 20:4. Enquanto na ausência de catiões divalentes, todas as composições de cadeias de acilo se comportam da mesma maneira, na presença de Ca^{2+}, as várias composições de cadeias de acilo produziram nanodomínios de $PI(4,5)P_2$ com propriedades biofísicas distintas. Em particular, cadeias de acilo cada vez mais saturadas produziram nanodomínios mais ordenados e estruturados, culminando na formação de nanodomínios semelhantes a gel para a composição totalmente saturada. É provável que o aumento de ordem da membrana detectado nos nanodomínios de $PI(4,5)P_2$ tenha um impacto profundo na sua sinalização, bem como na função e organização das proteínas efectoras. Surpreendentemente, a formação destas fases ordenadas foi largamente abrogada na presença da composição canónica, 18:0 20:4. Os resultados fornecem assim uma explicação pela qual a evolução terá favorecido fortemente a composição de 18:0 20:4 para o $PI(4,5)P_2$ em células de mamíferos.

As propriedades biofísicas do $PI(4,5)P_2$ exploradas neste trabalho foram também utilizadas como referência para o desenvolvimento de topologias do $PI(4,5)P_2$, bem como dos restantes lípidos da família dos fosfatidilinositóis, para o campo de forças CG Martini 3. Mais do que uma simples actualização das versões existentes para Martini 2, estes novos modelos foram desenvolvidos independentemente, com maior precisão e com um âmbito de aplicação alargado para incluir, entre outros, a reprodução exacta da agregação de fosfatidilinositóis mediada por catiões. Os modelos desenvolvidos melhoraram substancialmente a dinâmica conformacional dos fosfoinositóis, mantendo-se estáveis a etapas de tempo típicas de Martini. Para além de reproduzirem a agregação de fosfatidilinositóis, são também capazes de reproduzir poses de ligação a proteínas determinadas experimentalmente. Os modelos desenvolvidos foram então utilizados para revisitar o trabalho sobre o impacto da composição de cadeias de acilo nas propriedades biofísicas dos nanodomínios $PI(4,5)P_2$, utilizando o campo de forças Martini 3. Aqui o foco foi dado às diferenças observadas entre os dois campos de forças.

No seu conjunto, este trabalho expande o conhecimento sobre a organização lateral do $PI(4,5)P_2$ e dá pistas sobre como alguns dos seus reguladores podem ter um impacto substancial na sua função.

Para além dos trabalhos relativos às propriedades biofísicas do $PI(4,5)P_2$, esta tese contém também dois estudos relacionados com o combate à Doença de coronavírus de 2019 (COVID-19). A COVID-19, causada pelo *severe acute respiratory syndrome coronavirus 2* (SARS-CoV-2), tornou-se uma pandemia global com mortalidade mais elevada do que a da gripe sazonal. A pandemia COVID-19 levou a uma dramática perda de vidas humanas em todo o mundo, e apresenta um desafio médico, social e económico sem precedentes na história moderna. Inicialmente, as nossas contribuições centraram-se na concepção e optimização de biofarmacêuticos contra o SARS-CoV-2, que foram desenvolvidos utilizando um protocolo em ROSETTA guiado por um algoritmo genético desenvolvido internamente.

O desenvolvimento desta abordagem, bem como a sua aplicação à concepção e optimização de uma biomolécula neutralizadora da SARS-CoV-2 baseada em ACE2 é descrita nesta tese.

Adicionalmente, com o aparecimento de várias variantes, estudámos também o impacto das variantes alfa, beta e delta na dinâmica conformacional e na acessibilidade à ACE2 do domínio de ligação ao receptor (RBD) da glicoproteína *Spike* do SARS-CoV-2. Simulações MD atomisticas longas mostram que em solução, o RBD apresenta duas conformações distintas em equilibrio: uma conformação "aberta", onde é livre de ligar à ACE2; e uma conformação "fechada", onde a superficie de ligação está bloqueada. As mutações presentes nas várias variantes desviaram este equilibrio na direção de conformações mais "abertas", e no caso da delta, causaram alterações substanciais da sua conformação que potencialmente poderão assistir na fuga a anticorpos.

Palavras-chave: PI(4,5)P_2, Cálcio, Domínios Lipídicos, Organização lateral membranar, Interações lipido-proteína, Espectroscopia e microscopia de fluorescência, Simulações de dinâmica molecular, Modelos de membrana.

Contents

List of abbreviations

In this thesis, acronyms will usually be expanded on first usage. Common acronyms, such as DNA, amino acid letter codes or IUPAC units, will not be expanded nor described.

- **AA** All-atom.
- **ACE2** Angiotensin-converting enzyme 2.
- **AFM** Atomic force microscopy.
- **ANTH** AP180 amino-terminal homology domain.
- **ApL** Area per lipid.
- **APP** Amyloid precursor protein.
- **BLI** Biolayer interferometry.
- **BODIPY-PC** 2-(4,4-difluoro-5-methyl-4-bora-3a,4a-diaza-s-indacene-3-dodecanoyl)-1-hexadecanoyl-*sn*-glycero-3-phosphocholine.
- **BSA** Bovine serum albumin.
- **C1** Protein kinase C conserved region 1 domain.
- **C2** Protein kinase C conserved region 2 domain.
- **CAP23** Brain abundant membrane attached signal protein 1.
- **CG** Coarse grained.
- **CHARMM** Chemistry at Harvard Macromolecular Mechanics. Atomistic forcefield for molecular dynamics.
- **COG or CoG** Center of geometry.
- **COM or CoM** Center of mass.
- **COVID-19** Coronavirus disease 2019. Diseased caused by the SARS-CoV-2 virus.
- **CV** Collective variable.
- **DAG** Diacylglycerol.
- **DOPC** 1,2-dioleoyl-*sn*-glycero-3-phosphocholine.
- **DOPE-Rho** 1,2-dioleoyl-*sn*-glycero-3-phosphoethanolamine-N-(lissamine rhodamine B sulfonyl).
- **DPH** 1,6-Diphenylhexatriene. Fluorescent probe for membrane fluidity measurements.
- **DPPC** 1,2-dipalmitoyl-*sn*-glycero-3-phosphocholine.
- **EDTA** Ethylenediamine tetraacetic acid.
- **EtOH** Ethanol.
- **FCS** Fluorescence correlation spectroscopy.
- **FFS** Fluorescence fluctuation spectroscopy
- **FLIM** Fluorescence-lifetime imaging microscopy.
- **FRAP** Fluorescence recovery after photobleaching.
- **FRET** Förster resonance energy transfer.
- **GA** Genetic algorithm.
- **Gag** Group-specific antigen protein.
- **GAP43** Growth associated protein of 43kDa.
- **GROMACS** Groningen Machine for Chemical Simulations. Molecular dynamics software package.

- **GUV** Giant unilamellar vesicle.
- **HBR** Highly basic region.
- **HEK** Human embryonic kidney cells.
- **HeLa** HeLa cell line.
- **HEPES** 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid. A common zwitterionic sulfonic acid buffer.
- **HIV-1** Type 1 human immunodeficiency virus.
- **I3P** Inositol trisphosphate.
- **LINCS** Linear constraint solver.
- **LUV** Large unilamellar vesicle.
- **MA** HIV Gag Matrix domain protein.
- **MARCKS** Myristoylated alanine-rich C kinase substrate.
- **MD** Molecular Dynamics.
- **MeOH** Methanol.
- **MERS-CoV** Middle East respiratory syndrome–related coronavirus.
- **MLV** Multilamellar vesicle.
- **MM-PBSA** Molecular mechanics – Poisson-Boltzmann surface area.
- **NBD** Nitrobenzoxadiazole.
- **NMR** Nuclear magnetic resonance.
- **NR5A1** Nuclear receptor steroidogenic factor-1.
- **PBS** Phosphate-buffered saline.
- **PC** Phosphatidylcholine.
- **PC12** PC12 Cell line
- **PCA** Principal component analysis.
- **PCH** Photon counting histogram.
- **PDB** Protein data bank.
- **PE** Phosphatidylethanolamine.
- **PH** Pleckstrin homology domain.
- **PH-YFP** Phospholipase C $\delta 1$ subtype – yellow fluorescent fusion protein.
- **PI** Phosphatidylinositol.
- **PI(3)P** Phosphatidylinositol 3-phosphate.
- **PI(3,4)P_2** Phosphatidylinositol 3,4-bisphosphate.
- **PI(3,4,5)P_3** Phosphatidylinositol 3,4,5-trisphosphate.
- **PI(3,5)P_2** Phosphatidylinositol 3,5-bisphosphate.
- **PI(4)P** Phosphatidylinositol 4-phosphate.
- **PI(4,5)P_2** Phosphatidylinositol 4,5-bisphosphate.
- **PI(5)P** Phosphatidylinositol 5-phosphate.
- **PIP4K** Phosphatidylinositol 5-phosphate 4-kinase.
- **PIP5K** Phosphatidylinositol 4-phosphate 5-kinase.
- **PLC** Phospholipase C.
- **POPC** 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphocholine.
- **POPS** 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phospho-L-serine.
- **PS** Phosphatidylserine.
- **PTEN** Phosphatase and tensin homologue on chromosome 10.
- **PVA** Polyvinyl alcohol.
- **PX** Phox domain.

- **QM** Quantum Mechanics.
- **QM/MM** Hybrid quantum mechanics/molecular mechanics approach.
- **RAT-1** Immortalized rat fibroblasts.
- **RBD** Receptor binding domain from the coronavirus spike glycoprotein.
- **RBM** Receptor binding motif.
- **RDF** Radial distribution function.
- **REF15** Rosetta energy function 2015.
- **RIN** Residue interaction networks.
- **S Protein** Coronavirus spike glycoprotein.
- **SARS-CoV** Severe acute respiratory syndrome coronavirus.
- **SARS-CoV-2** Severe acute respiratory syndrome coronavirus 2.
- **SASA** Solvent accessible surface area.
- **SD** Standard deviation.
- **SEM** Standard error of the mean.
- **SLB** Supported lipid bilayer.
- **SNARE** Soluble NSF attachment protein receptor.
- **SPR** Surface plasmon resonance.
- **STED** Stimulated emission depletion microscopy.
- **STORM** Stochastic optical reconstruction microscopy.
- **SUV** Small unilamellar vesicle.
- **TIP3P** Transferable intermolecular potential with 3 points.
- **T_m** Main transition temperature.
- **TMA-DPH** Trimethylammonium-diphenylhexatriene. Fluorescent probe for membrane fluidity measurements.
- **TopFluor-PI(4,5)P_2** Phosphatidylinositol 4,5-bisphosphate fluorescently labeled with a TopFluor fluorophore (1-oleoyl-2-6-[4-(dipyrromethene boron difluoride)butanoyl]aminohexanoyl-*sn*-glycero-3-phosphoinositol-4,5-bisphosphate).
- **tPnA** Trans-parinaric acid. Fluorescent probe typically used to detect the presence of membrane gel domains.
- **VMD** Visual molecular dynamics.
- **VOC** Variant of concern.
- **YFP** Yellow fluorescent protein.

List of symbols

In this thesis, symbols will usually be defined on first usage.

- β PCH instrument and probe-dependent factor.
- $\mathbf{B_{22}}$ Osmotic second virial coefficient.
- **C** Particle concentration.
- **D** Diffusion coefficient.
- **E** FRET efficiency.
- $\mathbf{F_i}$ Fraction weighed by brightness.
- **G** Calibration factor.
- **G(τ)** Autocorrelation of fluorescence fluctuations.
- ***H*** Mean curvature.
- **I** Fluorescence intensity.
- $\mathbf{I_{sc}}$ Interface energy score.
- $\mathbf{k_B}$ Boltzmann's constant.
- $\mathbf{K_c}$ Bending rigidity.
- $\mathbf{K_d}$ Dissociation constant.
- $\mathbf{k_{off}}$ Dissociation rate constant.
- $\mathbf{k_{on}}$ Association rate constant.
- $\mathbf{K_p}$ Partition coefficient.
- λ Wavelength.
- **N** Number of particles.
- $\overline{\mathbf{N}}$ Average number of particles.
- $\boldsymbol{< r >}$ Fluorescence anisotropy.
- $\mathbf{R_0}$ Foster Radius.
- **S** Axial to lateral dimension ratio.
- **T** Fraction of fluorophores in the triplet state.
- $\mathbf{T_m}$ Main transition temperature.
- $\mathbf{T_s}$ Number of detected photons per sampling time.
- **V** Observation volume.
- $\mathbf{V_0}$ Reference volume.
- ε Molecular brightness.
- γ Incomplete gamma function.
- **k** Photon counts.
- $\tau_\mathbf{d}$ Translational diffusion time.
- $\tau_\mathbf{t}$ Triplet lifetime.
- $\omega_\mathbf{xyz}$ Focal volume dimensions.

Outline

Biological membranes, or biomembranes, are complex structures, made up of a vast number of lipids, proteins and other components, which are tasked with enclosing the cell and separating its components from the surrounding environment. In eukaryotes, they also segregate specific processes and components into spatially localized compartments within the cell, leading to increased efficiency and restricted dissemination of metabolites. Due to their complex compositions, biomembranes are very heterogeneous, with transversal and lateral asymmetry of lipid distribution, segregation of certain lipids and proteins into membrane domains, interactions with the cytoskeleton, etc., leading us further away from the simplistic approach of the original fluid mosaic model of Singer and Nicolson. However, this almost chaotic heterogeneity is also at the foundation of some of the most important functions of biomembranes, such as the organization of reaction sequences, the regulation of traffic and communication to and from the cytosol and the regulation of cellular signaling processes. As such, the biochemical and biophysical characterisation of biomembranes and their complex organization, will ultimately contribute to a better understanding of cell biology as a whole.

This work is focused on a minority component of biomembranes, phosphatidylinositol 4,5-bisphosphate ($PI(4,5)P_2$). $PI(4,5)P_2$ is a membrane lipid, mostly present in the inner leaflet of the plasma membrane of mammalian cells. Due to its peculiar biochemical and biophysical characteristics — which will be further developed in this work — it is one of the most actively involved lipids in the regulation of vital cellular signaling pathways. For $PI(4,5)P_2$ to be engaged in a multitude of simultaneous cellular processes, its levels must be tightly regulated at a spatiotemporal level to avoid significant fluctuations of its low plasma membrane concentration. This is thought to occur through the presence of multiple localized pools of $PI(4,5)P_2$ in the membrane, possibly established via interactions with other other membrane components, such as proteins or divalent cations.

This dissertation is organized into 5 chapters, each of which is extensively based on published work. Unpublished data that are already submitted for publication were also included as articles on their own. Both the manuscripts' and the dissertation's page-numbering were kept. Apart from the downscaling to fit the page size, the published manuscripts were not otherwise altered. Supplementary material for each article is available in the respective appendix. The experimental biophysical methodologies carried out in this work, in particular, fluorescence spectroscopy and microscopy techniques were performed at Centro de Química-Física Molecular (CQFM) at Instituto Superior Técnico (IST), which later became the Biospectroscopy and Interfaces Research Group (BSIRG) of the Institute for Bioengineering and Biosciences (iBB), also at IST. Theoretical computational approaches, namely molecular dynamics (MD) simulations, were performed at the Multiscale Modeling Lab at Instituto de Tecnologia Química e Biológica António Xavier (ITQB NOVA).

Chapter I serves as an introduction to this work, and provides an in depth overview of some of the most important $PI(4,5)P_2$ molecular and biophysical properties, as well as their impact on $PI(4,5)P_2$

membrane dynamics, lateral organization, and interactions with other molecular partners. This overview is given through a biophysics oriented perspective, instead of the more typical biology oriented view, complementing many of the excellent cell biology focused reviews which already exist. The bulk of the introduction is based on a published review which was revised and adapted[1].

Temporal and spatial regulation of $PI(4,5)P_2$ concentration can achieve localized increases in the lipid, which are crucial for the activation and recruitment of peripheral proteins to the plasma membrane. Chapter II focuses on the impact of cation-induced $PI(4,5)P_2$ clusters on $PI(4,5)P_2$-protein interactions. In this chapter, I suggest that, given that $PI(4,5)P_2$ is likely to exist constitutively clustered *in vivo*, protein anchoring to the plasma membrane through $PI(4,5)P_2$, is likely not defined solely by a simple (monomeric $PI(4,5)P_2$)/(protein bound $PI(4,5)P_2$) equilibrium, but instead depends on complex protein interactions with $PI(4,5)P_2$ clusters. More so, I also investigate how these $PI(4,5)P_2$ nanodomains might also impact and modulate protein-protein interactions.

Chapter III focuses on an often overlooked regulator of $PI(4,5)P_2$ lateral organisation, its acyl-chain composition. While most phospholipids show considerable acyl-chain composition diversity, $PI(4,5)P_2$ is highly enriched in specific acyl chains. In fact, the most frequent fatty acyl-chain pair for $PI(4,5)P_2$ in mammalian cells is 1-stearoyl-2-arachidonyl (18:0 20:4), which accounts for up to 70 % of the total $PI(4,5)P_2$ pool in specific tissues. $PI(4,5)P_2$ molecular species exhibiting no polyunsaturation or even fully saturated acyl chains are less abundant, but have been reported to become more prevalent in response to certain stimuli, stress, ageing, or in cancer. In this chapter, I will explore how different acyl-chain compositions impact $PI(4,5)P_2$ and $PI(4,5)P_2$ nanodomain biophysical properties.

The experimental biophysical work shown in the previous chapters was complemented by molecular dynamics work carried out using the Martini 2 coarse-grain force field. The Martini CG model is the most widely used CG force field for biomolecular simulations and it has been successfully applied in the modeling of a variety of molecular processes — especially, in the modeling of biomembrane systems. With the recent release of the new and improved version 3 of the Martini force field, I had the opportunity to develop the parameters for inositol and the 8 phosphoinositol lipids. More than a simple version update from existing Martini 2 models, these are models developed independently from their Martini 2 counterparts, with greater accuracy and expanded application scope to include, among others, accurate reproduction of cation-mediated phosphoinositide aggregation. In chapter IV, I describe the development and validation of the phosphoinositide models, and revisit some of the previous molecular dynamics work on the impact of acyl-chain composition on $PI(4,5)P_2$ biophysical properties using the newly developed models.

During the course of my PhD work, a pandemic outbreak caused one of the biggest challenges in modern medicine history. Coronavirus disease 2019 (COVID-19), caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), has become a global pandemic with higher mortality than that of seasonal influenza. As of November 2021, over 5 million lives had been claimed by this disease. Like many other researchers, I dedicated a substantial amount of research time to the fight against SARS-

CoV-2. In chapter V, I describe the development process of a potential anti SARS-CoV-2 biomolecule, as well as the study of the impact of SARS-CoV-2 variants on the receptor binding domain conformational dynamics and ACE2 accessibility.

At the end of this dissertation, final remarks on the developed work are given, along with future perspectives on membrane biophysics, PI(4,5)P_2 lateral organisation, and its potential impact on cell biology.

Most of the work and data analysis presented in this dissertation was performed by me with the following exceptions:

- Article I: PH-YFP steady-state fluorescence spectroscopy experiments, which were performed by Marina E. Monteiro and Maria J. Sarmento;
- Article II: AFM measurements of PI(4,5)P_2 containing supported lipid bilayers, which were performed by Marco M. Domingues in collaboration with Instituto de Medicina Molecular (iMM);
- Article V: Atomistic simulations of RBD-ACE2 complexes, which were performed by Diana Lousa and Cláudio M. Soares in collaboration with the Protein Modeling lab at ITQB NOVA.
- Article VI: Atomistic simulations of RBD, which were performed by Diana Lousa and Cláudio M. Soares in collaboration with the Protein Modeling lab at ITQB NOVA.

Author contributions are also clearly stated at the end of each article presented in this dissertation.

1 | Introduction: Structure and lateral organization of PI(4,5)P_2

The plasma membrane is a complex structure tasked with enclosing the cell and separating it from the surrounding environment. While biomembranes provide structure and define the boundaries of the cell, their dynamic biochemical and biophysical characteristics also allow them to regulate traffic and communication to and from the cytosol, organize reaction sequences, and promote cellular processes. These biophysical properties are defined not only by the intrinsic physical and chemical properties of the lipids, proteins, and other components but also by their complex set of interactions. This complexity holds the key to many key cellular processes. A lipid that stands out in the landscape of the eukaryotic plasma membrane is phosphatidylinositol 4,5-bisphosphate (PI(4,5)P_2). PI(4,5)P_2 is the most abundant phosphoinositide in mammalian cells and is found primarily in the inner leaflet of the plasma membrane. It has also been found in endosomes, in the endoplasmic reticulum, and in the nucleus[2]. While its role as a source of secondary messengers during signaling events is known for decades[3], evidence has accumulated through time of its importance as an intact phospholipid for defining plasma membrane identity in eukaryotic cells. Due to its very large headgroup and multivalent negative charge, PI(4,5)P_2 acts almost like an electrostatic beacon, interacting specifically or non-specifically with several molecular entities such as membrane proteins, other lipids, cellular cations, etc. As a result of its particular biophysical properties, it is a major regulator of a wide spectrum of plasma membrane events, including cell adhesion and motility[4,5], ion channel transport[6,7], vesicle endocytosis[8,9,10,11,12], and exocytosis[9,13,14,15] (Figure 1). This wide-ranging reach of PI(4,5)P_2 as a critical functional lipid has made it an important research focus over the last decades in cell biology and more specifically in areas such as neuroscience, virology, and biophysics.

1.1 | The Phosphoinositide family

PI(4,5)P_2 is a member of the phosphoinositide (PI) lipid family. PIs are a small group of glycerophospholipids derived from phosphatidylinositol. These lipids consist of a characteristic inositol headgroup, which can undergo reversible phosphorylation and dephosphorylation leading to the formation of seven distinct

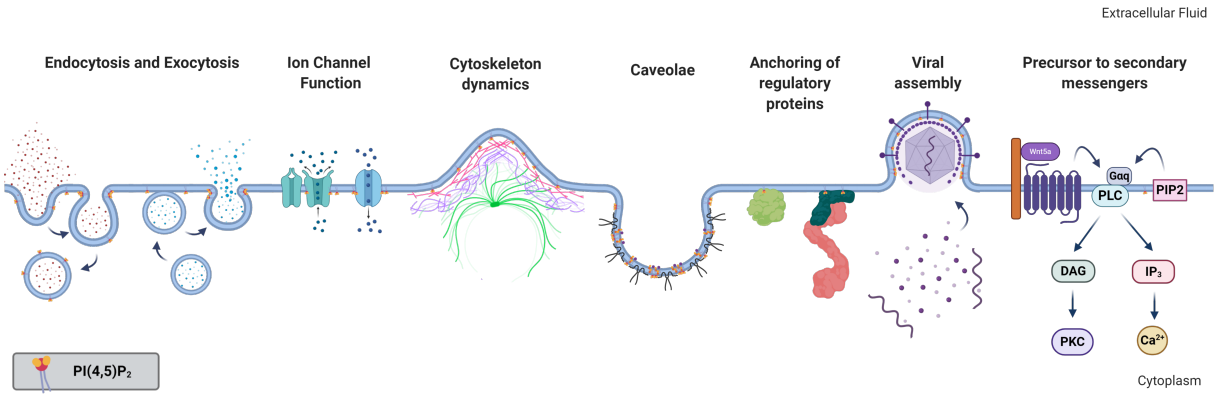


Fig. 1. Membrane processes associated with or dependent on PI(4,5)P_2. Figure created with BioRender.com.

phosphorylated species. While the parent lipid phosphatidylinositol represents roughly 10% of total membrane phospholipids in the eukaryotic cell, the phosphorylated derivatives account only for around 2-3%[2], with PI(4)P and PI(4,5)P_2 representing the bulk of these lipids[16]. Each of these seven species has a distinct subcellular distribution with a predominant localization in subsets of membranes. Additionally, within a given membrane the localization of a specific PI can be heterogeneous. Many PIs are overall in low abundance in the membrane but they can be found at high local concentrations in membrane domains not readily detected by conventional techniques[4,16]. Over the last couple of decades, PIs have been found to be one of the most ubiquitous signalling entities in eukaryotic metabolism. Their reach extends from controlling organelle biology to regulating cellular growth. Due to this all-encompassing reach they have also been linked to a number of human diseases. In fact, the inositide signalling pathway is considered a promising pharmaceutical target. For an excellent review on the major developments on PI cellular biology and their impact on disease see Balla (2013)[4].

1.2 | PI(4,5)P_2 Structure

1.2.1 | Headgroup conformation

The core PI(4,5)P_2 structure descends from its “parent” lipid, phosphatidylinositol. At the core, it consists of a myo-inositol headgroup. There are 9 existing isomers of inositol but the myo-inositol isomer is the most commonly found in eukaryotic cells. In its most stable conformation, it assumes the chair conformation where every hydroxyl substituent is at the equatorial position except for the hydroxyl in the position 2 of the ring which is in an axial position. The myo-inositol moiety is typically linked to the diacylglycerol (DAG) backbone via a phosphodiester bond in position 1. This leaves the hydroxyl groups in positions 2 to 6 open. However, only positions 3, 4 and 5 can be enzymatically phosphorylated to yield the 7 phosphorylated PI species. PI(4,5)P_2 is the result of phosphorylation in positions 4 and 5 of the headgroup by specific kinases and phosphatases (Figure 2, A).

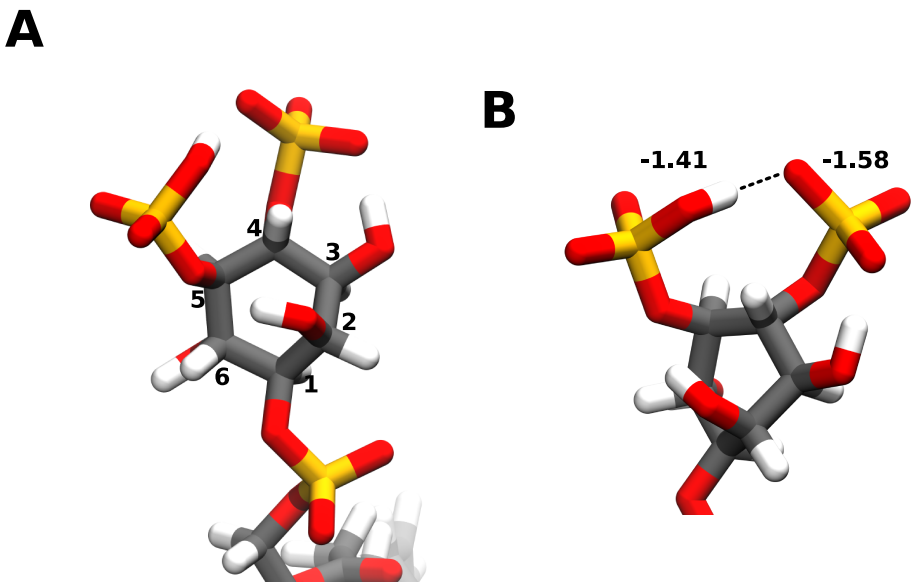


Fig. 2. PI(4,5)P_2 headgroup features. The PI(4,5)P_2 headgroup consists of myo-inositol ring where every hydroxyl substituent is at the equatorial position except for the hydroxyl in the position 2 of the ring which is in an axial position. In the case of PIP2 the hydroxyls in positions 4 and 5 are enzymatically phosphorylated. It is linked to the DAG backbone via a phosphodiester bond in position 1 (A). At pH 7.0, one of the phosphodiester proton dissociates, and the one remaining is shared between the two vicinal phosphomonoester groups. In terms of potential charge this means that, at pH 7.0, the charges would be -1.58 and -1.41 for the phosphomonoester groups at positions 4 and 5 respectively[17]. The lower charge of the 5-phosphomonoester group is attributed to a network of intramolecular hydrogen bonds that it is engaged in, which stabilize the proton (B). Carbon atoms are coloured in grey, hydrogen in white, oxygen in red and phosphorus in orange. Snapshots obtained from a simulation of a bilayer consisting of 95:5 mol ratio POPC: PI(4,5)P_2, using the CHARMM36m forcefield run in GROMACS2019. Images were modelled using VMD.

In mammals, the majority of PI(4,5)P_2 in the plasma membrane is synthesized from PI(4)P by type I PIP5Ks[18]. Type II PIP4K phosphorylate PI(5)P to synthesize a quantitatively minor pool of PI(4,5)P_2 localized in the Golgi[2,19,20]. It can also be produced by the dephosphorylation of PI(3,4,5)P_3 catalysed by PTEN and TPIP[21]. PI(4,5)P_2 hydrolysis is controlled by specific 4'- or 5'-phosphatases or by phospholipase breakdown in response to various stimuli. Dephosphorylation by specific phosphatases (primarily 5'-phosphatases) controls PI(4,5)P_2 steady-state levels and controls the extent of its signalling. Additionally, cleavage by phospholipases, such as phospholipase C (PLC), control PI(4,5)P_2 levels and originate metabolites that propagate and amplify cellular signalling. PI(4,5)P_2 levels, in general, are the result of a complex interplay of many cellular enzymes. While PI(4,5)P_2 metabolism falls out of the scope of this review, more information can be found elsewhere[4,22].

1.2.2 | Membrane conformation dynamics

In terms of molecular structure when inserted into the membrane, there are surprisingly very few experimental studies probing PI(4,5)P_2 dynamics. Since the dynamics of phosphatidylinositol or of the mono-phosphorylated PI(4)P has been addressed to some extent, we can estimate some of PI(4,5)P_2

properties from the behaviour of these closely related precursors. At the insertion depth of phosphatidylinositol and PI(4)P, when inserted into the membrane and in the absence of any interactions with other chemical partners, it is believed that the phosphodiester bond is located roughly at the same depth as the phosphodiester of phosphatidylcholine. Additionally, evidence points that the phosphodiester bond remains roughly parallel when compared to the membrane normal[23,24,25,26].

PI(4,5)P_2 headgroup tilt seems to be significantly impacted by phosphorylation. For the case of phsophatidylinositol, the headgroup is roughly perpendicular to the membrane plane, with the C4 hydroxyl as the most exposed to the water layer despite a slight tilt being observed arising from an intramolecular hydrogen bond established between the C2 hydroxyl and the *pro*-R-oxygen of the phosphodiester phosphate[23,24,25,27]. This is the result of the glyceryl-phosphate-inositol link being always very close to a trans, trans, trans, gauche- conformation, which brings the two hydrogen bond partners together. Interestingly, the formation of this hydrogen bond appears to be crucial for the recognition by PLC. However, it is not yet clear if it is formed when phosphatidylinositol is aggregated[24]. In the case of PI(4)P the headgroup tilt is more significant[23,25,26], and authors suggest it might be also due to the establishment of electrostatic interactions between the negatively charged phosphate and the positively charged choline headgroups from the neighbouring lipids. For PI(4)P, due to this more significant tilt, the C5 hydroxyl is the most accessible to the water layer.

In a Variable Angle Sample Spinning NMR study[26], PI(4,5)P_2 membrane conformation was studied in a membrane like environment consisting of neutral alkyl-poly(ethylene)glycol and long-chain alcohols. All possible conformations obtained showed a much more pronounced headgroup tilt for PI(4,5)P_2 than for PI(4)P, where the PI(4,5)P_2 headgroup would be laying almost parallel to the membrane surface. As this cannot arise from specific electrostatic interactions in this membrane model, it is likely that this is the result of more subtle effects such as water- or alcohol-mediated hydrogen bonding. As the analysis of NMR measurements of complex systems (such as PIs) is error prone and the “membrane matrix” used is far from being biologically relevant, the authors of this study could not be definitive in their conclusions regarding PI(4,5)P_2 orientation. Nevertheless, they were able to replicate what had been previously observed for PI(4)P in other membrane mimetics. If these observations are replicated in more relevant conditions, they will challenge the more established “concept” of how PI(4,5)P_2 is structurally displayed in the membrane and how it interacts with protein partners. All-atom (AA) molecular dynamics (MD) simulations of PI(4,5)P_2 in lipid membranes show a well-defined average head-tail angle of roughly 100° indicating that the headgroup would lie mostly flat along the membrane in agreement with the previous studies[28] (Figure 3). Poisson Boltzmann MD simulations, however, show a more conservative tilt of roughly 40°.

Overall, there are strong hints that, in the absence of interactions with other biochemical partners, PI(4,5)P_2 when inserted into a membrane has its headgroup lying tilted over the membrane. The extent of this tilt is still yet to be fully understood and likely depends not only on PI(4,5)P_2 intrinsic properties but also on the interactions with the neighboring lipid molecules.

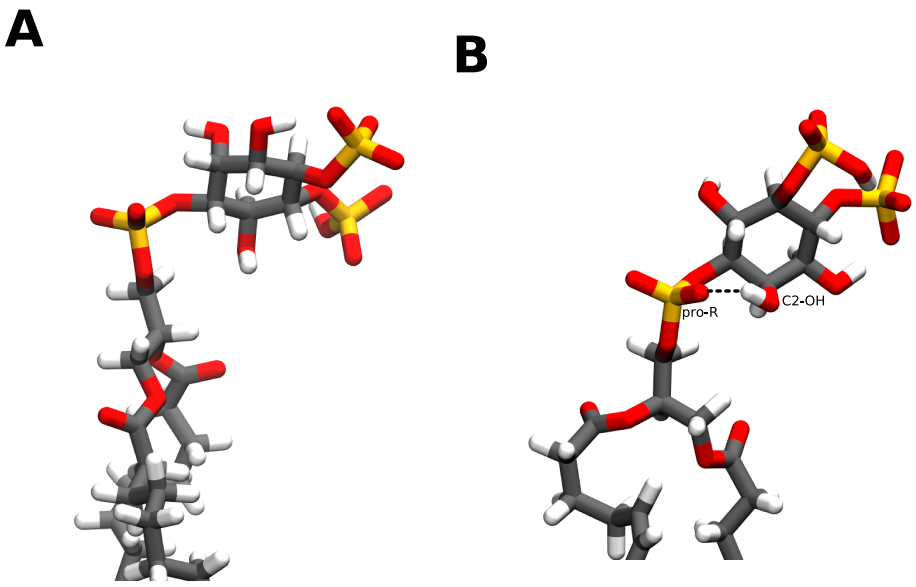


Fig. 3. Examples of PI(4,5)P_2 headgroup tilt when inserted into a phospholipid membrane. PI(4,5)P_2 presents a significant headgroup tilt, when inserted into a bilayer, ranging from almost parallel to the membrane plane (A) to a more conservative 40° tilt (B). Whilst the more dramatic headgroup tilt appears to be favoured from interactions between its negatively charged phosphate groups and the positively charged membrane surface, the more moderate tilt surges from the establishment of intramolecular hydrogen bonds between the C2 hydroxyl and the *pro*-R-oxygen of the phosphodiester phosphate. Carbon atoms are coloured in grey, hydrogen in white, oxygen in red and phosphorus in orange. Snapshots obtained from a simulation of a bilayer consisting of 95:5 mol ratio POPC: PI(4,5)P_2, using the CHARMM36m forcefield run in GROMACS2019. Images were modelled using VMD.

1.2.3 | Headgroup charge

An important aspect that is closely related to the conformation of PI(4,5)P_2 is the charge state of each of its headgroup phosphate groups. Whilst the charge state of the headgroup linking phosphate group is well defined at physiological pH (pKa between 1 to 3)[29], the headgroup phosphate charges are much more volatile. The charge state of these groups has mostly been studied experimentally for PI(4,5)P_2 in micelles and small unilamellar vesicles (SUVs), via 31P – NMR and the dependency of the chemical shift on the pH. Typically, PI(4,5)P_2 is considered to have approximately four negative charges at cytosolic pH. This result was inferred from the determination of the pKa of each headgroup phosphate in either pure or mixed vesicles of PI(4,5)P_2 using 31P-NMR[30]. These experiments determined a pKa value for the first protonation of roughly 6.7 and 7.7 for the phosphomonoester groups at position 4 and 5 respectively. In terms of potential charge, this would mean that at pH 7.2, the 5-phosphomonoester would be partially protonated at charge -1, the 4-phosphomonoester would be fully deprotonated with charge -2 and the phosphodiester would have charge -1.

However, studies have since shown that the PI(4,5)P_2 headgroup ionization behaviour (as well as for other phosphoinositides) can't be accurately described by a Henderson–Hasselbalch mechanism[17]. In the more detailed mechanism that was proposed[17], at pH values close to 4-5, both phosphomonoester groups are mono-protonated as previously described. Upon increasing the pH, one proton dissociates,

whilst the remaining one is shared between the two vicinal phosphomonoester groups (Figure 2, B). In terms of potential charge this means that, at pH 7.0, the charges would be -1.58 and -1.41 for the phosphomonoester groups at positions 4 and 5 respectively. The lower charge of the 5-phosphomonoester group is attributed to the fact that it establishes an hydrogen bond interaction with the hydroxyl group in the position 6, which in turn, is also forming a long range hydrogen bond with the phosphodiester group in the position 1. This weakens the first hydrogen bond slightly and thus the proton binds to the 5-phosphomonoester more tightly. These results gave a much more detailed look at the charge distribution of PI(4,5)P_2 and at the complex network of intra- and intermolecular hydrogen bonds that lead to the dissipation of the headgroup charge and are, very likely, part of the reason for why repulsion between phosphoinositides is much weaker than expected.

A fact that is often overlooked, is how the interaction of PI(4,5)P_2 with neighbouring molecules influences its charge distribution. *In vivo*, PI(4,5)P_2 is constantly in interaction not only with its neighbouring lipids, but also with proteins and cationic ions. The complex network of interactions formed by PI(4,5)P_2 with these partners leads to a greater distribution of its charge, effectively altering its headgroup charge. A study has shown that, in the absence of divalent cations, lipids with hydrogen bond donor capabilities could influence PI(4,5)P_2 headgroup charge[31]. Phosphatidylethanolamine (PE) was shown to influence the first step of deprotonation of the PI(4,5)P_2 headgroup, most likely by interacting preferably with the 5-phosphate. In the presence of PI, the first step of deprotonation was not affected, but a lower degree of ionization was observed for both phosphomonoester groups for the removal of the last shared proton. Curiously, phosphatidylserine (PS) was not shown to affect PI(4,5)P_2 headgroup ionization significantly. This study clearly showcases how the PI(4,5)P_2 local environment can affect PI(4,5)P_2 headgroup charge, a crucial feature responsible for a lot of its biological function.

1.2.4 | Acyl-chain composition

The acyl chain composition of lipids often plays an important functional role. These roles can be defined by specific interactions with proteins or by simply changing the overall biophysical properties of the surrounding membrane. In general, fatty acid profiles vary between phospholipid classes, tissues and species, and can also vary over time with dietary habits, stimuli or disease. Like many other lipids, PI(4,5)P_2 is also subject to these effects. In mammals, the phosphorylated myo-inositol headgroup is typically bound to a DAG moiety, which consists of two fatty acid chains bonded to a glycerol molecule via ester bonds at positions sn1 and sn2. The major fatty acid profile observed for PI(4,5)P_2 in mammals consists of 1-stearoyl-2-arachidonyl (18:0 20:4)[21]. Curiously, inositol-phosphate headgroups coupled to ceramide have also been identified in fungi, plants and protozoa[32]. However, I will focus only on the mammalian-relevant DAG-bound species in this review.

The 18:0 20:4 acyl chain profile consists of up to 70% of the total PI(4,5)P_2 lipid pool in some cell lines, especially in brain tissue. This enrichment is likely the combined outcome of substrate specificity for

1-stearoyl-2-arachidonyl-glycerol in several enzymes in the phosphatidylinositol cycle and the remodeling of phosphoinositide acyl-chains via the Land's cycle[33]. A more detailed look at how the cell might maintain this enrichment can be seen in this review by D'Souza et al (2014)[33]. However, PI(4,5)P_2 still has a broad distribution profile ranging from highly unsaturated chains to fully saturated ones[34]. These less abundant species become more prevalent in response to certain stimuli[35], stress[34,35], aging[34] or in cancer[36]. In these cases, fully saturated and mono-unsaturated compositions increase significantly in concentration, in some cases even surpassing the canonical 18:0 20:4 composition[35].

But why do cells spend so many resources in maintaining this particular acyl-chain composition? And why does the acyl-chain profile shift, sometimes dramatically, in specific conditions? The biological functions behind the enrichment in 18:0 20:4 are still not very clear. The enrichment in this configuration seems to be particularly prevalent in brain cells, where PI(4,5)P_2 has been associated with several stages of both endocytosis and exocytosis and has been considered an important mediator of synaptic vesicle trafficking[37]. It has been shown that arachidonate, and other polyunsaturated fatty acids such as docosahexaenoate (22:6), at the sn2 position, facilitate membrane shaping and fission activities. Additionally, asymmetric sn1-saturated-sn2-polyunsaturated phospholipids have been shown to provide efficient membrane vesiculation whilst maintaining lower membrane permeability[38]. These properties might provide significant mechanical benefits in these particular tissues. Of course, arachidonic acid in particular has biological activity of its own in addition to serving as the precursor for the synthesis of eicosanoids, such as prostaglandins and leukotrienes[39]. Overall, this particular theme has been given little attention thus far; nonetheless, it could be the key for some of PI(4,5)P_2 multifunctionality.

1.3 | Lateral organization of PI(4,5)P_2

Having looked at the core structural properties of PI(4,5)P_2 we now turn to its organization in the plasma membrane. As PI(4,5)P_2 is engaged in a multitude of cellular functions occurring in parallel, its levels must be tightly regulated to avoid significant fluctuations of its total plasma membrane concentration. This implies that the simultaneous regulation of these cellular functions by PI(4,5)P_2 must occur through the presence of multiple localized pools of this phospholipid in the plasma membrane[40]. PI(4,5)P_2 lateral organization in cells has been studied through a variety of techniques from Fluorescence correlation spectroscopy (FCS) and Fluorescence recovery after photobleaching (FRAP) to Atomic force microscopy (AFM). In FCS experiments carried out in Rat1 fibroblasts and HEK cells, researchers microinjected micelles of fluorescent labelled-PI(4,5)P_2 into cells and showed that the diffusion coefficient of PI(4,5)P_2 in these cells is significantly lower than expected for free phospholipids. The simplest interpretation of this result is that approximately two thirds of PI(4,5)P_2 in the inner leaflet of the plasma membrane is somehow sequestered[41]. Studies in PC12 cells have also shown, using stimulated emission depletion (STED) microscopy[42] and stochastic optical reconstruction microscopy (STORM) imaging techniques[43], that PI(4,5)P_2 is highly enriched in nanometre sized membrane domains, specific to this cellular model.

In fact, while the presence of segregated PI(4,5)P_2 pools can be partly explained by localized PI(4,5)P_2 synthesis and degradation through several kinases and phosphatases[44], it is also evident that membrane diffusion rates, in the absence of significant obstacles for diffusion, will always be higher than concentration changes due to enzymatic activity causing PI(4,5)P_2 to diffuse away faster than it can be produced. This means that it is unlikely that local synthesis can result in significant changes in the submicroscopic organization of PI(4,5)P_2 in the membrane[40]. PI(4,5)P_2 interactions with other cellular binding partners could alternatively explain the observed lateral organization of this phosphoinositide. Interactions with proteins, divalent cations, cholesterol and the cytoskeleton are the ones most likely to have such an impact. In this review I will give particular attention to the often neglected effect of divalent cations on the lateral organization of PI(4,5)P_2.

1.3.1 | Sequestration by proteins

One way to explain PI(4,5)P_2 lateral organization in the plasma membrane of cells, is that proteins can act as reversible buffers, binding much of the PI(4,5)P_2 present and then releasing it locally in response to specific signals[45]. Theoretical simulations predict that such sequestration can be achieved not only through specific interactions with PI(4,5)P_2, but also through nonspecific electrostatic interactions. In fact, polybasic proteins are able to sequester a lipid with a valence of 4 (such as PI(4,5)P_2) 1000-fold more effectively than a lipid with a valence of 1 (such as PS)[46,47]. Due to its highly negatively charged headgroup, PI(4,5)P_2 was confirmed to interact strongly with polybasic stretches of aminoacid residues[40,48]. Through these polybasic stretches, several proteins were found to laterally sequester PI(4,5)P_2 molecules in a reversible manner[49,50]. For an efficient buffering of PI(4,5)P_2 levels, these proteins would have to be present at a concentration comparable to PI(4,5)P_2, localize to the plasma membrane and be able to bind PI(4,5)P_2 with high affinity while being able to release it in response to stimuli. Proteins such as MARCKS[47,50,51], GAP43[45,52], CAP23[45], among many others, have been shown to be able to sequester PI(4,5)P_2 in such a manner. In the case described above of PI(4,5)P_2 domains detected in PC12 cells, these were found to be associated with the sequestration of PI(4,5)P_2 to clusters of the SNARE protein syntaxin-1[42,53,54]. This sequestration by syntaxin-1 is critical for the regulation of SNARE dependent membrane fusion[55,56].

Employing fluorescence and electron paramagnetic resonance spectroscopic tools, McLaughlin, Cafiso and co-workers[47,49] showed that a 24 amino acid peptide corresponding to the effector domain of MARCKS was able to efficiently sequester an average of 3 PI(4,5)P_2 molecules through non-specific electrostatic interactions. Importantly, this sequestration occurred even in the presence of physiological concentrations of the monovalent acidic phospholipid PS, confirming theoretical predictions. MARCKS sequestration of PI(4,5)P_2 has been shown to be important in the PI(4,5)P_2 mediated activation of TRPC-family Ca^{2+} channels[57], in the endocytosis of the amyloid precursor protein (APP)[58] and in the synaptic clustering of PI(4,5)P_2[59].

1.3.2 | PI(4,5)P_2 interactions with divalent cations

Several studies have shown that PIs and PI(4,5)P_2, in particular, are able to establish strong electrostatic interactions between their negatively charged headgroups and divalent cations. In the cellular PI(4,5)P_2 context, Ca^{2+} and Mg^{2+} stand out. Ca^{2+} is a common player in signal transduction and a second messenger in cells. Its levels are strictly controlled and maintained at low levels in the cytosol, with normal intracellular levels at around 100 nM (20 000 fold lower than extracellular levels)[60]. Upon stimulation, however, several signal transduction pathways can lead to transient increases of intracellular Ca^{2+} concentration up to around 1 µM, with local concentrations in the vicinity of open Ca^{2+} channels reaching hundreds of µM, before being regulated back to normal levels[61]. In fact, PI(4,5)P_2 has been reported to be associated with a variety of Ca^{2+} channels and a great number of these require PI(4,5)P_2 for proper function[4]. Mg^{2+}, on the other hand, is a less studied modulator of cell function. Mg^{2+} levels are well buffered in a narrow millimolar range between 0.25 mM and 1 mM[62,63] and are thus kept at a much higher concentration than those of Ca^{2+}. Both divalent cations have been shown to bind strongly to PI(4,5)P_2 and influence its lateral organization dramatically as discussed below.

Through QM/MM experiments we can get an insight on the molecular basis for cation binding to PI(4,5)P_2[28]. From a molecular point of view, when binding to a single PI(4,5)P_2 lipid, both Ca^{2+} and Mg^{2+} bind to PI(4,5)P_2 either in between the phosphomonoester groups (Figure 4, B) or solely near the 4-phophomonoester (Figure 4, A). However, simultaneous binding between the two phosphomonoester groups is approximately 10 kcal/mol more unfavourable[28]. Divalent cation binding to the phosphodiester group has also been observed[64].

When analysing the free energy associated with the removal of each divalent cation from its binding position, significantly more energy is required to remove Ca^{2+} into the bulk water than it is for Mg^{2+}. The difference in free energy could come from the fact that, in contrast to Ca^{2+}, Mg^{2+} appears to retain its first hydration shell in its equilibrium binding position. This causes its equilibrium binding position to be further away from the headgroup and leads to the formation of less hydrogen bonds, on average, between the headgroup and the surrounding water molecules. Interestingly, it was also shown that upon binding to Ca^{2+}, the remaining PI(4,5)P_2 headgroup proton at physiological pH could be favourably displaced and that the effective size of the PI(4,5)P_2 headgroup would significantly decrease[28]. In the presence of Mg^{2+}, the dissociation of the remaining proton was not favourable. However, the decrease in effective headgroup surface area was also observed albeit at a lesser extent. All of these cation induced changes can and will affect PI(4,5)P_2 dynamics, thus influencing local membrane dynamics as well as its interactions with protein binding partners.

Apart from simply binding to PI(4,5)P_2, both divalent cations also have the ability to crosslink PI(4,5)P_2 lipids. This induces the formation of very stable cation induced PI(4,5)P_2 nanodomains. It has been shown through different experimental techniques that divalent cations, and especially Ca^{2+}, are able to

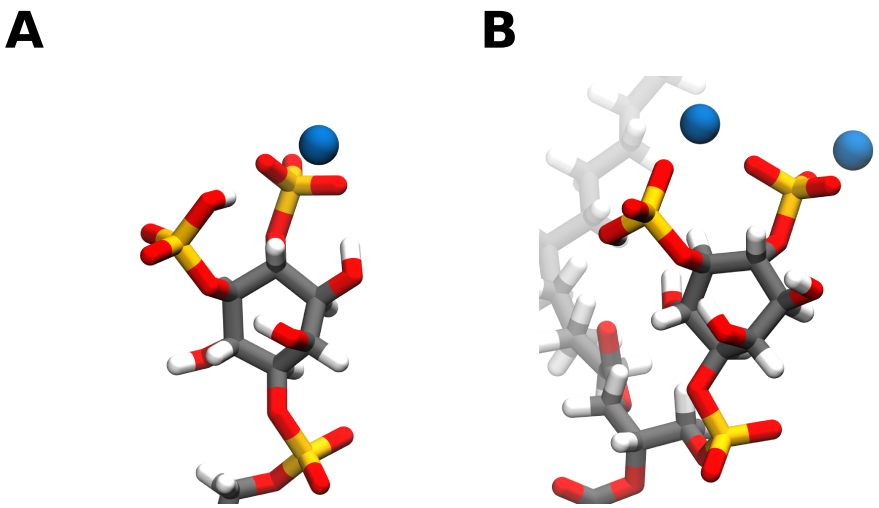


Fig. 4. Snapshots of Ca^{2+} ions interacting with the PI(4,5)P_2 headgroup phosphates. Ca^{2+} can bind to PI(4,5)P_2 either solely near the 4-phophomonoester (A) or in between the phosphomonoester groups (B). However, simultaneous binding between the two phosphomonoester groups is approximately 10 kcal/mol more unfavourable[28]. Carbon atoms are coloured in grey, hydrogen in white, oxygen in red, phosphorus in orange and Ca^{2+} in blue. Snapshots obtained from a simulation of a bilayer consisting of 95:5 mol ratio POPC: PI(4,5)P_2 in the presence of Ca^{2+} in a 5:1 Ca^{2+} to PI(4,5)P_2 ratio, using the CHARMM36m forcefield run in GROMACS2019. Images were modelled using VMD.

induce the formation of PI(4,5)P_2 nanodomains[a], even at physiological concentrations of cation and lipid. In lipid monolayers, these clusters can be detected through AFM.[65,66] (Figure 5). Through the use of fluorescent analogues of PI(4,5)P_2 , Ca^{2+}-induced clusters were shown to occur in model membranes at physiologically relevant concentrations of both Ca^{2+} and PI(4,5)P_2[67]. Other phosphoinositides have also shown some propensity to form cation-induced clusters. PI(3,5)P_2 has been found to form nanodomains by itself in the presence of physiological concentrations of Ca^{2+} cations.In the presence of Mg^{2+} PI(3,5)P_2 clustering was negligible[68]. On top of that, the clusters formed by PI(3,5)P_2 were much smaller and likely less stable than those formed by PI(4,5)P_2[68]. On the other hand, when the monophosphorylated PI(4)P was tested in the same type of experiments, no Ca^{2+}-induced clusters where observed[68].

From a molecular point of view, a single divalent cation can likely crosslink up to 2 PI(4,5)P_2 lipids by simultaneously binding each lipid phosphodiester and/or headgroup phosphomonoester groups via strong electrostatic interactions[69]. A single PI(4,5)P_2 lipid, however, can simultaneously bind up to 3 divalent cations, and thus be complexed with 3 other PI(4,5)P_2 lipids (Figure 6). This net of PI(4,5)P_2 – cation interactions can thus induce the formation of a grid of tightly crosslinked lipids. Whilst the main driving force behind the clustering appears to be cation crosslinking, the formation of a complex network of intermolecular hydrogen bonds, between the headgroup hydroxyl and phosphomonoester groups, very likely plays a role in thermodynamically favouring clustering. Due to the electrostatic nature of the cation interactions, the propensity to crosslink PI(4,5)P_2 lipids appears to be highly correlated with the affinity towards the divalent cation. Thus, Ca^{2+} shows a much greater clustering propensity than Mg^{2+}. In fact,

[a] PI(4,5)P_2 nanodomains and PI(4,5)P_2 clusters will often be used interchangeably in this work.

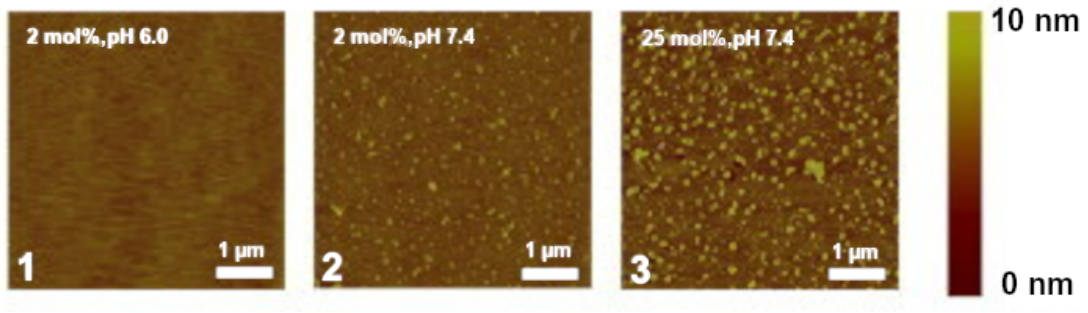


Fig. 5. Snapshots of experiments on mixed lipid monolayers, containing different mol % of SOPC and PI(4,5)P_2, while exposed to Ca^{2+}. Adapted from Biophysical Journal, 101, Ellenbroek, W.G.; Wang, Y.H.; Christian, D.A.; Discher, D.E.; Janmey, P.A.; Liu, A.J. Divalent cation-dependent formation of electrostatic PIP2 clusters in lipid monolayers. 2178–2184[66].

Ca^{2+} induced clusters have been shown to be significantly larger than those induced by Mg^{2+} at the same experimental conditions[65]. However, although Mg^{2+} has a much weaker affinity for PI(4,5)P_2 when compared to Ca^{2+}, its steady state levels are several orders of magnitude higher than those of Ca^{2+}, and at these mM concentrations it is also able to induce comparable PI(4,5)P_2 clustering[68].

As the formation of these clusters is driven mainly by the crosslinking of the phosphate groups, the nanodomains formed are composed of almost only PI(4,5)P_2. Studies have shown that other phosphorylated PI species can co-cluster with PI(4,5)P_2, albeit to a lesser degree, but that the parent lipid phosphatidylinositol cannot[70]. Incorporation into clusters also seems to be independent from acyl-chain composition[70], still, it is very likely that different acyl-chain compositions induce the formation of nanodomains with different biophysical properties.

These cation induced PI(4,5)P_2 nanodomains are much more than simply the sum of their elements. While Ca^{2+} is known to directly regulate the interaction of different protein domains to PI(4,5)P_2[71], the structure and dynamics of the phospholipid within the divalent cation-induced cluster presents markedly distinct biophysical characteristics than the monodisperse lipid. As mentioned previously, binding of divalent cations, and in particular Ca^{2+}, can alter PI(4,5)P_2 headgroup exposure leading to a decrease in solvent accessible area[28]. Additionally, as divalent cations accumulate, significant screening of the headgroup charges occurs, essentially shielding the large negatively charged headgroup from potential binding partners[72]. As PI(4,5)P_2 lipids are forced to accumulate in an enclosed area, further reorganization of the headgroups occurs, promoted by the molecular interactions of the divalent cations with the 3 phosphate groups[73], effectively altering the headgroup conformation. This local accumulation likely influences PI(4,5)P_2 acyl chain dynamics, and therefore, local membrane order. Studies have shown that PI(4,5)P_2 nanodomains have significantly reduced lateral dynamics[67] and that PI(4,5)P_2, which as a single lipid has a strong preference for disordered domains, displays significantly less affinity for disordered domains upon clustering[68]. All of these altered biophysical properties can, and most likely will, influence downstream PI(4,5)P_2 signalling by modulating its interactions with protein and lipid partners.

Altogether, these findings show that divalent cation mediated clustering can lead to the formation of specific sites in the membrane highly enriched in PI(4,5)P_2, while depleting the rest of the membrane[67].

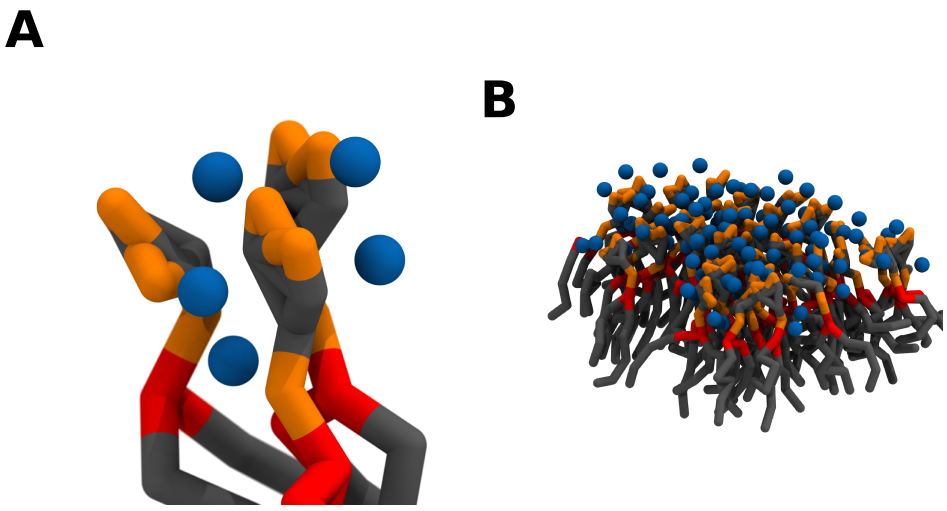


Fig. 6. Crosslinking of PI(4,5)P_2 lipids induces the formation of PI(4,5)P_2 nanodomains. As a single divalent cation can bind up to 2 PI(4,5)P_2 lipids and each lipid can potentially bind 3 divalent cations, a network of electrostatic interactions can crosslink PI(4,5)P_2 lipids together (A). As the number of clustered lipids increases, PI(4,5)P_2 nanodomains are formed (B). Coarse grain beads representing the inositol headgroup and acyl-chains are coloured in grey, the glycerol component in red, the phosphate groups in orange and Ca^{2+} in blue. Snapshots obtained from a simulation of a bilayer consisting of 95:5 mol ratio POPC: PI(4,5)P_2 in the presence of Ca^{2+} in a 5:1 Ca^{2+} to PI(4,5)P_2 ratio, using the martini 2.2 coarse-grained forcefield run in GROMACS2019. Images were modelled using VMD.

PI(4,5)P_2 is likely constitutively clustered in the membrane, crosslinked by Mg^{2+} ions alongside other minor phosphoinositide lipids. In the vicinity of active Ca^{2+} channels, where Ca^{2+} concentrations increase significantly upon opening of a channel, both ions will simultaneously contribute to the formation of the nanodomains, to form larger and more stable PI(4,5)P_2 nanodomains. These cation-induced nanodomains can influence not only PI(4,5)P_2 lateral organization and biophysical properties but also the way proteins interact with PI(4,5)P_2, by modulating their localization in the plasma membrane, their target recognition and binding affinity to PI(4,5)P_2, and even further downstream interactions with other proteins. Therefore, beyond the impact of Ca^{2+} on PI(4,5)P_2 levels in the membrane through activation of phospholipase activity, the direct interaction of divalent cations with PI(4,5)P_2 is expected to play a crucial role in the regulation of the biological activity of this phospholipid.

1.3.3 | Effect of cholesterol on PI(4,5)P_2 properties and distribution

Cholesterol is a crucial membrane component, implicated in a myriad of membrane processes. However, its most noted role is in the regulation of plasma membrane biophysical properties as a “fluidity buffer”. Whilst its effects can vary with different cholesterol contents, cholesterol, in general, decreases membrane fluidity by increasing lipid packing even leading to the cholesterol-dependent formation of coexisting liquid phases[74]. Like all the other phospholipids in the plasma membrane, PI(4,5)P_2 is also subject to these cholesterol-dependent effects.

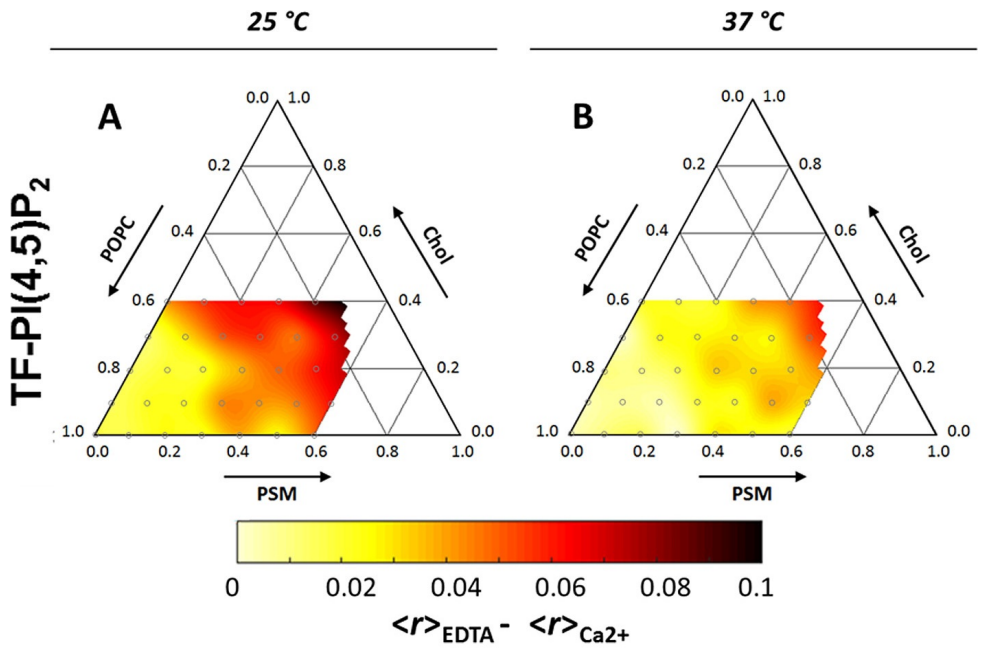


Fig. 7. Ternary diagram for the POPC:PSM:Chol lipid mixture at 25 (A) and 37 ° C (B). Color-code depicts decrease in measured fluorescence anisotropies of a PI(4,5)P_2 fluorescent analogue (TF-PI(4,5)P_2) upon inclusion of 100 µM Ca^{2+}. Since the decrease reflects homo-FRET between analogues incorporated within the same clusters, darker areas correspond to more efficient PI(4,5)P_2 clustering. Adapted from Sarmento, M.J.; Coutinho, A.; Fedorov, A.; Prieto, M.; Fernandes, F. Membrane order is a key regulator of divalent cation-induced clustering of PI(3,5)P_2 and PI(4,5)P_2. Langmuir 2017, 33, 12463–12477.[68]

Unsurprisingly, given its large negatively charged headgroup and highly unsaturated acyl chain, PI(4,5)P_2 was shown to preferentially partition into the less ordered cholesterol-poor phases of biphasic monolayers containing PI(4,5)P_2:SOPC:Chol[75]. However, after the addition of Ca^{2+}, the subsequent cation-induced PI(4,5)P_2 nanodomains were shown to have increased the miscibility of the coexisting domains in the cholesterol-containing monolayers[75]. Related results have been observed, in a study with fluorescent derivatives of PI(4,5)P_2 incorporated in ternary mixtures of POPC:SM:Chol. In this study, monodisperse PI(4,5)P_2 presented low miscibility in more ordered lipid phases, yet, after cation-induced clustering, the preference for disordered domains decreased by more than two-fold[75]. Importantly, the lipid composition of this ternary mixture was shown to have a marked influence both on the extent of PI(4,5)P_2 Ca^{2+}-induced clustering and on the size of clusters formed (Figure 7)[75] . Since the dimensions of PI(4,5)P_2 clusters were heavily dependent on temperature, it was concluded that the major factor regulating PI(4,5)P_2 clustering was membrane order and not the presence of a specific molecular partner in the membrane. This suggests that the insertion of PI(4,5)P_2 in more ordered domains is stabilized by the formation of cation-induced nanodomains.

In a cellular context the effect of cholesterol on PI(4,5)P_2 appears to be heavily dependent on cell type. In fibroblasts[76] and cultured pancreatic ß-cells[77] cholesterol depletion leads to decreased levels of free PI(4,5)P_2, whilst in HEK293[78] cholesterol enrichment was shown to promote PI(4,5)P_2 depletion.

Altogether, discussed findings appear to be in agreement with theories that associate PI(4,5)P_2 with the controversial cholesterol-rich microdomains, such as “lipid rafts” and caveolae, which are said to

be involved in regulating a variety of membrane functions. PI(4,5)P_2 has been found to be enriched in detergent-resistant membranes[79]. And while detergent extraction has been put into question on whether it induces an artefactual enrichment in PI(4,5)P_2[80], studies have shown PI(4,5)P_2 to colocalize with more ordered membrane domains[81,82,83,84] and to be sensitive to membrane curvature[84,85], both associated with these types of structures. Interestingly, in a study where a PI(4,5)P_2 specific phosphatase was targeted to either the "raft" or the "nonraft" membrane fractions of T cells, the authors were able to show clear evidence of compartmentalization of PI(4,5)P_2 dependent signalling in each of the fractions. When depleting the "nonraft" fraction of PI(4,5)P_2, T cells showed an increase in cell filopodia and cell spreading, whilst in contrast, when depleting the "raft" fraction of PI(4,5)P_2, T cells showed smooth membranes free of ruffling and filopodia among other effects. Findings also appear to suggest that roughly half of the PI(4,5)P_2 content is synthesized preferentially in these cholesterol-enriched domains[79]. Interestingly, in HeLa cells, cholesterol was shown to not significantly impact the lateral organization of PI(4,5)P_2, hinting that it is not associated with raft-like membrane patches[86], unlike what was observed for other systems[79,87]. As such, this is still an area of great controversy amongst researchers, and there is a lot yet to uncover before significant conclusions can be drawn on the importance of these microdomains for PI(4,5)P_2 lateral organization.

1.3.4 | Effect of the cytoskeleton and curvature on PI(4,5)P_2 lateral organization

PI(4,5)P_2 has been shown to be a major player in cytoskeleton dynamics, by interacting and regulating the activity of numerous enzymes and cytoskeletal proteins[88,89]. However, the cytoskeleton can also regulate PI(4,5)P_2, and in particular its lateral organization, via corralling by the cortical actin network. Cortical actin networks have been shown to be able to induce spatio-temporal confinement of phospholipids in the plasma membrane of living cells[90]. PI(4,5)P_2 should be no exception to this effect and, in fact, due to its close proximity with a variety of actin binding proteins[89] one can suspect it could be even more susceptible to these effects. Studies have shown that the cytoskeleton is responsible for some of the low mobility of PI(4,5)P_2 in atrial myocytes[91]. Additionally, the cytoskeleton has been shown to be critical for the formation of PI(4,5)P_2 enriched compartments in HeLa cells, as disruption of actin polymerization resulted in an almost homogeneous distribution of PI(4,5)P_2[86].

Curvature can also greatly influence PI(4,5)P_2 lateral organization. PI(4,5)P_2 has been found to undergo a transient increase at the phagocytic cup during the initiation of phagocytosis[92,93]. More recently, it was found that the curvature induced by the engagement of non-biological solid particles with the plasma membrane was enough to increase PI(4,5)P_2 concentrations at the site of contact. Additionally, as previously discussed, PI(4,5)P_2 has been associated with several stages of endocytosis and exocytosis, where curvature effects are paramount[37]. As a monodisperse lipid, PI(4,5)P_2 has an inverted cone-shaped structure[94] due to its very large inositol headgroup. As such, it is associated

with positive membrane curvature. After interacting with divalent cations, however, PI(4,5)P_2 presents a cone-shaped structure[94], likely due to the decrease in headgroup area as well as the aggregation of the headgroups after complexation with the cations. In this case, it would be associated with negative membrane curvature. Whether local curvature at the plasma membrane plays a major role in dictating PI(4,5)P_2 lateral organization or PI(4,5)P_2 lateral organization contributes to local curvature is not entirely clear. In a cellular context, it is likely dependent on the process in question and the overall result of both effects.

1.4 | Main perspectives and goals

Despite having been discussed separately in this work, all the complex biophysical properties discussed previously are tightly interlinked processes. It is the combined effect of these properties that allows PI(4,5)P_2 to be a major regulator of membrane dynamics despite being present at very low overall concentrations. While a lot of research has been conducted on these effects, many are still to be fully characterized, especially those associated with cation induced nanodomains, such as the lipid conformation in these structures, the extent of charge dissipation and the effect of these nanodomains on the local bilayer properties. A good molecular understanding of these effects is fundamental in order to better understand how PI(4,5)P_2 carries out its role as a major plasma membrane regulator.

Over the last decades, extensive research efforts have uncovered a multitude of different cellular roles of PI(4,5)P_2. However, the current view on the majority of the mechanisms associated with these functions neglect the almost certain presence of a highly significant, if not dominant, pool of this phospholipid that is not monodispersed. Special consideration should be given to the fact that in the plasma membrane, PI(4,5)P_2 must be either protein-bound or constitutively complexed with divalent cations within small clusters. PI(4,5)P_2 within these structures is bound to have significantly different properties from the monodisperse lipid. These properties can influence PI(4,5)P_2 interactions with its binding partners (such as proteins) as well as downstream protein-protein interactions. It is conceivable that many of PI(4,5)P_2 cellular functions are also regulated by the extent of this effect.

As such, crucial aspects of PI(4,5)P_2 function are yet to be properly unfolded. As was mentioned, temporal and spatial regulation of PI(4,5)P_2 can achieve localized increases in the concentration of the lipid, partly due to the effect of cation-induced clustering. The dramatic impact of physiological divalent cation concentrations on PI(4,5)P_2 clustering, suggests that protein anchoring to the plasma membrane through PI(4,5)P_2, is likely not defined solely by a simple (monomeric PI(4,5)P_2)/(protein bound PI(4,5)P_2) equilibrium, but instead depends on complex protein interactions with PI(4,5)P_2 clusters. However, the impact of the organization of PI(4,5)P_2 at the plasma membrane on its biomolecular interactions with PI(4,5)P_2 binding proteins is largely unknown.

The role of the highly conserved 1-stearoyl-2-arachidonyl PI(4,5)P_2 acyl-chain composition remains

unknown. A significant amount of cellular resources is spent on maintaining this dominant acyl-chain composition, while the less abundant and more saturated species have been shown to become more prevalent in response to certain stimuli, stress or aging. Nevertheless, the effect of these acyl-chain compositions on the biophysical properties of PI(4,5)P_2 and PI(4,5)P_2 nanodomains have not yet been described.

PI(4,5)P_2 is known to directly regulate several membrane remodeling events. While PI(4,5)P_2 is essential for membrane fusion processes such as SNARE-dependent membrane fusion, it has an intrinsic positive curvature which is expected to restraint the formation of the negatively curved fusion intermediates necessary for some of these mechanisms. How PI(4,5)P_2 can overcome this intrinsic curvature barrier is still not fully understood. Additionally, these intrinsic curvature properties may influence PI(4,5)P_2 regulation by inducing domain coupling effects across the bilayer or providing mechanical benefits in membrane shaping and fission events.

Molecular biophysics techniques, such as steady-state and time-resolved fluorescence spectroscopy, can be used accurately to follow and quantify the formation of PI(4,5)P_2 nanodomains as well as probe their micro and macro environment with the combined use of specific membrane probes in model membranes, such as large unilamellar vesciles (LUVs). Fluorescence spectroscopy can also be used to quantitatively follow the interaction of PI(4,5)P_2 with protein partners, and, with the assistance of more advanced tools, such as fluorescence fluctuation spectroscopy techniques, obtain binding and oligomerization rates.

The simultaneous deployment of theoretical approaches such as quantum mechanics or molecular mechanics simulations allow us to obtain an insight onto PI(4,5)P_2 properties that are hard or currently impossible to probe experimentally. Coarse-grained (CG) molecular dynamics simulations stand out from these techniques, as they allow us to test large system sizes at millisecond timescales. CG simulations can probe PI(4,5)P_2 aggregation dynamics, biophysical properties, large-scale curvature effects, as well as PI(4,5)P_2-Protein interactions at almost atomistic detail. The combined use of these experimental and theoretical techniques can allow us to further characterise and understand the biophysical properties and biological functions of PI(4,5)P_2.

2 | Impact of Ca^{2+}-dependent clustering on the properties of $PI(4,5)P_2$ binding proteins

Article I: Impact of Ca^{2+}-dependent $PI(4,5)P_2$ clustering on the properties of $PI(4,5)P_2$ binding proteins

Impact of Ca^{2+}-dependent $PI(4,5)P_2$ clustering on the properties of $PI(4,5)P_2$ binding proteins

Luís Borges-Araújo[†,‖,⊥], Marina E. Monteiro[†], Dalila Mil-Homens[†], Nuno Bernardes[†], Maria J. Sarmento[†], Ana Coutinho[†], Manuel Prieto[†], Manuel N. Melo[‖], Fábio Fernandes [†,‡,⊥,*]

† iBB-Institute for Bioengineering and Biosciences, Instituto Superior Técnico, Universidade de Lisboa, Lisbon, Portugal

⊥ Associate Laboratory i4HB—Institute for Health and Bioeconomy at Instituto Superior Técnico, Universidade de Lisboa, Lisbon, Portugal

‡ Department of Bioengineering; Instituto Superior Técnico, Universidade de Lisboa, 1049-001 Lisbon, Portugal;

‖ Instituto de Tecnologia Química e Biológica António Xavier, Universidade Nova de Lisboa, Av. da República, 2780-157 Oeiras, Portugal.

ABSTRACT: Despite its low abundance, phosphatidylinositol 4,5-bisphosphate ($PI(4,5)P_2$) is a key modulator of membrane associated signalling events in eukaryotic cells. Temporal and spatial regulation of $PI(4,5)P_2$ concentration can achieve localized increases in the levels of the lipid, which are crucial for the activation or recruitment of peripheral proteins to the plasma membrane. The recent observation of the dramatic impact of physiological divalent cation concentrations on $PI(4,5)P_2$ clustering, suggests that protein anchoring to the plasma membrane through $PI(4,5)P_2$, is likely not defined solely by a simple (monomeric $PI(4,5)P_2$)/(protein bound $PI(4,5)P_2$) equilibrium, but instead depends on complex protein interactions with $PI(4,5)P_2$ clusters. Nevertheless, the impact of the organization of $PI(4,5)P_2$ at the plasma membrane on its biomolecular interactions with $PI(4,5)P_2$ binding proteins is largely unknown. Using different advanced fluorescence spectroscopy and molecular dynamics methodologies, we characterized the impact of Ca^{2+} on $PI(4,5)P_2$ – protein interactions and membrane organization. We show that in giant unilamellar vesicles presenting $PI(4,5)P_2$, the membrane diffusion properties of pleckstrin homology (PH) domains tagged with a yellow fluorescent protein (YFP) are affected by the presence of Ca^{2+}, suggesting interaction of the protein with $PI(4,5)P_2$ clusters. Importantly, PH-YFP is found to dimerize in the membrane in the absence of Ca^{2+} and this oligomerization is inhibited in the presence of physiological concentrations of the divalent cation. Furthermore, it was also found that Ca^{2+} induced clustering of $PI(4,5)P_2$ could modulate the affinity of $PI(4,5)P_2$ binding proteins and peptides via a complex mixture of mechanisms. These results confirm that Ca^{2+}-dependent $PI(4,5)P_2$ clustering has the potential to dramatically influence affinity, oligomerization and organization of $PI(4,5)P_2$ binding proteins in the plasma membrane.

INTRODUCTION

The binding of peripheral proteins to biological membranes is known to be a fundamental process in cell function and homeostasis, not just by modulating local membrane dynamics, but also by organizing and regulating protein signalling complexes. In this context, lipid-protein interactions play a crucial role in the recruitment of peripheral proteins by establishing interactions through their headgroups and hydrocarbon chains with protein membrane binding domains. These membrane binding domains have distinct structures and can bind by a combination of stereochemical recognition of the lipid headgroup, and/or nonspecific long range electrostatic or hydrophobic interactions, leading to different binding affinities and specificities[1]. Simultaneously, biological membranes present a variable and complex composition due to the large repertoire of distinct lipids, that vary between different cell types and organelles. This gives rise to membranes with distinct biophysical properties that also influence protein-lipid interactions. Overall, peripheral protein membrane binding is a complex kinetic process that is spatially and temporally regulated, which is still not fully understood.

Phosphatidylinositol 4,5-bisphosphate ($PI(4,5)P_2$) is the most abundant polyphosphoinositide found in the inner leaflet of the plasma membrane of mammalian cells, comprising around 1 mol% of the total membrane phospholipids[2,3]. Due to the two phosphorylations at positions 4 and 5 of the inositol headgroup it has a larger than average headgroup, with a highly negative charge density that stands out in the inner leaflet of the plasma membrane[4]. $PI(4,5)P_2$ is thought to have a distinct lateral organization, through the presence of different lipid pools across the inner plasma membrane. This is most likely not only the consequence of a non-homogeneous distribution of $PI(4,5)P_2$ but also of the establishment of $PI(4,5)P_2$–protein[5] and $PI(4,5)P_2$–cation interactions. In fact, studies indicate that as much as two thirds of $PI(4,5)P_2$ in the inner leaflet of the plasma membrane are somehow sequestered[6]. As a result of these distinct biophysical characteristics, $PI(4,5)P_2$ plays a critical role in membrane lipid-protein interactions, being responsible for the targeting of several protein domains to the

plasma membrane, as well as spatially and temporally regulating their activity[7]. In fact, $PI(4,5)P_2$ has been reported to be associated with numerous vital cell functions such as cell adhesion and motility[8], vesicle endocytosis[9–11], exocytosis[10,12,13] and ion channel transport[14]. There have been several $PI(4,5)P_2$-binding domains identified, among which are included the PX domains[15–17], C2 domains[18], ENTH domains[19], Tubby domains[20] and most notably the pleckstrin homology (PH) domains[21,22].

The pleckstrin homology (PH) domain consists of a 100-120 residue long amino acid sequence found in numerous proteins involved in a multitude of cellular signalling processes. Initially, it was thought that most PH domains directed membrane targeting of their host proteins by binding to phosphoinositides. However, it is now known that of the total number of PH domains detected in protein sequences, only a small minority group bind to phospholipids. In fact, from the known phospholipid binding PH domains, few do so specifically to PIs. Most bind weakly and non-specifically[23]. Despite this, the isolated PLCδ1 PH domain, in particular, was found to bind with high affinity and specificity to $PI(4,5)P_2$ and its soluble headgroup, inositol trisphosphate molecule (I3P)[24]. Studies of $PI(4,5)P_2$ binding by the PLCδ1 PH domain provided the first demonstrations of specific PI recognition by a PH domain and provided the foundations for how binding domains recognize specific phosphoinositides in cellular membranes[25]. Indeed, PLCδ1 PH domains are still to this day used as excellent protein models to study $PI(4,5)P_2$-protein interactions and $PI(4,5)P_2$ organization. Each characterized PH domain has essentially the same structure, consisting of a β sandwich closed off at one end by a C-terminal α helix [26], with the splayed corner of the β sandwich containing 3 variable loops, that have been suggested, by analogy with immunoglobulin-like domains, to be the phosphorylated inositol binding site[23]. It has also been found that phospholipid binding PH domains are electrostatically polarized with the positively charged end coinciding with the three variable loops. It is through this region that these domains bind to the polyphosphorylated inositol ring[27]. Through direct hydrogen bonds between the bound headgroup and seven amino acid residues, the 4- and 5- phosphate groups of the phosphorylated inositol headgroup are clamped and buried in the binding pocket. These interactions suggest that stereochemical cooperativity enhances specificity[28]. It is worth noting that while PLCδ1PH binds to $PI(4,5)P_2$ with high affinity, FRAP results suggest that its steady-state membrane localization is the result of its very rapid cycling between a membrane bound and a cytosolic state, with on-off rates in the range of seconds[29]. Also microdissociation and rebinding events have been detected for PH domains in single molecule studies, confirming a transient binding mechanism[30].

Recently, there has been a plethora of research focused on the interactions between $PI(4,5)P_2$ and divalent cations. $PI(4,5)P_2$ has been shown to establish strong electrostatic interactions between its negatively charged headgroup and divalent cations, especially with Ca^{2+}. In mammalian cells, Ca^{2+} is a common secondary messenger with an important role in signal transduction, typically binding and regulating proteins directly[31]. However, when Ca^{2+} interacts with $PI(4,5)P_2$, it alters its lateral organization and induces its segregation into microdomains[32–35]. This has been demonstrated at fully physiological conditions, through the incorporation of $PI(4,5)P_2$ fluorescent analogues on $PI(4,5)P_2$ clusters in free standing lipid bilayers using fluorescence spectroscopy methodologies[32]. Furthermore, Ca^{2+} has also been shown to influence $PI(4,5)P_2$ electrostatics via counterion accumulation, causing an electrostatic charge shielding effect that inhibits protein-lipid interactions[36]. Another important effect, is the influence of Ca^{2+} on the $PI(4,5)P_2$ headgroup conformation, via molecular interactions with the three phosphate groups[37]. These changes in headgroup conformation are expected to heavily influence $PI(4,5)P_2$ interactions with binding partners. In all, these studies imply that Ca^{2+}, a common signal transduction element with a very buffered low physiological concentration, has a strong role in the regulation of $PI(4,5)P_2$.

Many $PI(4,5)P_2$ binding proteins are also known to be sensitive to variations in Ca^{2+} concentration[7]. In most of these cases, Ca^{2+} causes changes in protein folding/electrostatics that affect the affinity of $PI(4,5)P_2$ binding proteins leading to an increase of binding to the membrane. In these proteins, $PI(4,5)P_2$ and Ca^{2+} binding typically occur at different binding sites[7]. However, the impact of the direct interaction of Ca^{2+} with $PI(4,5)P_2$ on binding proteins has been given little attention. It is worth noting that the formation of $PI(4,5)P_2$ clusters due to divalent cation cross-linking or the occurrence of the charge shielding effect, as described previously, are likely to have a dramatic effect not only on protein targeting or activity but also on the promotion/inhibition of protein-protein interactions. Ca^{2+} could then be a potential additional regulatory mechanism by interacting with $PI(4,5)P_2$ directly, and modulating protein affinity and binding protein organization. Such regulatory effects would be of special importance in the vicinity of Ca^{2+} channels where steep fluctuations of local Ca^{2+} concentrations occur, before buffering by Ca^{2+} binding proteins.

In this work, we aim to study how Ca^{2+}-induced $PI(4,5)P_2$ effects may influence $PI(4,5)P_2$ interactions with binding proteins and binding protein organization, focusing on the effects on two protein models, PH-YFP and PBP-10. With these two models we can probe the influence on both specific and nonspecific binding to $PI(4,5)P_2$, respectively. PH-YFP is a fusion protein that contains an isolated PLCδ1 PH domain linked to a YFP variant that is prone to dimerization. As a result, the PH-YFP construct is an excellent protein model to use as a case study of $PI(4,5)P_2$ modulation of protein-protein and protein-lipid interactions. PBP-10 consists of a 10 residue long, polycationic peptide derived from the $PI(4,5)P_2$ binding region in segment-2 of gelsolin (gelsolin residues 160-169), which was linked to rhodamine B[38,39]. It's binding to PIs appears to be governed mostly by electrostatics; however, it has been suggested that there may be a steric contribution, since gelsolin shows much higher affinity than similar peptides with higher positive charges[38]. These characteristics make PBP-10 an excellent model for studying non-specific protein- $PI(4,5)P_2$ interactions.

The results presented in this article show, through a complementary set of fluorescence techniques and coarse grained molecular simulations, that calcium within its physiological range of intracellular concentrations is able to significantly modulate directly $PI(4,5)P_2$-protein interactions, through a variety of mechanisms. Additionally, we also show that these effects can extend beyond lipid-protein interactions and influence protein-protein dynamics.

MATERIALS AND METHODS

Materials. 1-palmitoyl-2-oleoyl-*sn*-glicero-3-phosphocholine (POPC); 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phospho-*L*-serine (POPS), 1,2-dioleoyl-*sn*-glycero-3-phospho-(1'-myo-inositol-4',5'-bisphosphate) ($PI(4,5)P_2$), 1-oleoyl-2-{6-[4-(dipyrromethene boron difluoride)butanoyl]amino}hexanoyl-*sn*-glycero-3-phosphoinositol-4,5-bisphosphate (TopFluor $PI(4,5)P_2$), 2-(4,4-difluoro-5-Methyl-4-Bora-3a,4a-Diaza-s-Indacene-3-Dodecanoyl)-1-hexadecanoyl-*sn*-glycero-3-phosphocholine (BODIPY-PC), 1,2-dioleoyl-sn-glycero-3-phosphoethanolamine-N-(lissamine rhodamineB sulfonyl) (DOPE-Rho), and 1,2-dioleoyl-*sn*-glycero-3-phosphoethanolamine-N-(cap biotinyl) (biotinylated DOPE) were purchased from Avanti Polar Lipids (Alabaster, AL, USA). Lipid stock solutions were prepared in chloroform with the exception of the PIs that were prepared in chloroform:methanol (MeOH) (2:1 vol/vol). Both solvents were obtained from Merck (Darmstadt, Germany) and were of spectroscopic grade. Phosphate buffer saline (PBS), 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES), ethanol (EtOH), NaCl, Sucrose, Glucose, $CaCl_2$, Imidazole, Glycerol, bovine serum albumin (BSA), BSA-biotin and avidin were purchased from Sigma-Aldrich. Rhodamine 110 and Fluo-5N were obtained from Molecular Probes, Invitrogen (Eugene, Or).

PH-YFP expression and purification. PH-YFPxpET28a was made from YFP(d)-PHxpCDNA3[29], which was a gift from Dr. Kees Jalink, (Division of Cell Biology, Netherlands Cancer). Briefly, the PH-YFP sequence flanked by BamHI and NotI restriction sites was inserted into a pET28a vector. PH-YFPxpet28a was expressed in *Escherichia coli* BL21 (DE). The cells were transformed by electroporation and incubated in LB agar plates with kanamycin overnight at 37ºC. A pre-inocula was made with a single colony from the plate and incubated overnight. The appropriate pre-inocula volume was added for an initial OD600nm of 0.1 in 100mL LB medium containing kanamycin. The culture was incubated at 24ºC, 250rpm, until the OD600nm reached approximately 0.6-0.8. Expression was then induced with 0.2mM IPTG at 24ºC and 250rpm for 3 hours. Cells were harvested from the culture by centrifugation (6000*g* for 15 min) at 4ºC. The pellet was suspended in lysis buffer (50 mM PBS, 300 mM NaCl, pH 8.0) with an added protease inhibitor cocktail and sonicated until complete DNA fragmentation was observed. 10 x 15s bursts with a 15s cooling period in-between were performed. The lysate was centrifuged (17600*g* for 5 min) at 4ºC to remove cellular debris. The supernatant was transferred to a clean tube without disturbing the pellet and centrifuged again at 17600*g* for 1 hour at 4ºC. PH-YFP was then purified making use of its histidine tag using Protino Ni-TED 2000 packed columns (Macherey-Nagel). Cleared lysates were applied to the column and washed with 8 column volumes of lysis buffer. The polyhistidine tagged protein eluted with 5 column volumes of elution buffer (50 mM PBS, 300 mM NaCl, 250 mM imidazole, pH 8.0). 3mL fractions were collected by gravity flow and monitored for protein presence by UV 280nm absorption. Most of the eluted protein could be found in the first fractions. SDS-PAGE analysis was used to determine protein purity. The purified protein was stored in a 10mM PBS buffer containing 140mM NaCl and 10% glycerol at pH 7.4.

Synthesis of PBP-10. PBP-10 was synthesized and purified by TOCRIS bioscience (United Kingdom). Peptide was analysed by HPLC and Mass spectrometry showing 97% purity. It was delivered as a white lyophilised solid. Upon arrival, the peptide was solubilized in a 20% (vol/vol) acetonitrile / water solution to a 2mg/mL concentration. Stock solution was divided into aliquots and stored at -20ºC protected from sources of light. Before use, the protein stock solution was diluted in experiment buffer to a working stock concentration and any portion of the unused aliquot was discarded.

Liposome preparation. Large unilamellar vesicles (LUVs) were prepared by extrusion of multilamellar vesicles[40]. Briefly, the lipid mixtures were prepared, from phospholipid stock solutions, and dried under a nitrogen flux and left in vacuum for 3 hours to remove traces of solvent. Multilamellar vesicles were then obtained through the solubilisation of the lipid films in the appropriate experimental buffer. Freeze-thaw cycles were performed, using liquid nitrogen and a water bath typically set to 60 ºC. The thawing temperature used was always higher than the melting temperature of the lipid with the highest melting temperature in the mixture, to re-equilibrate and homogenize the samples. LUVs were obtained by extrusion at room temperature, using an Avanti Mini-Extruder (Alabaster, AL) and 100nm pore size polycarbonate membranes (Whatman, USA). Typically, at least 21 passages through the extruder were performed.

Giant unilamellar vesicles (GUVs) were obtained by gel-assisted formation, based on a method previously described[41]. The lipid mixtures were prepared, from stock solutions, in chloroform to a final concentration of 1.5mM. For the PH-YFP experiments, the mixtures were composed of 95% POPC and 5% $PI(4,5)P_2$. DOPE-Cap-biotin was included in the mixture at a biotinylated lipid/total lipid ratio of 1/750 000 [42]. A solution of 5% (w/w) polyvinyl alcohol (PVA) MW ~145 000 and 280 mM of sucrose was spread in a µ-slide chamber from Ibidi (Munich, Germany) and left to dry for 15 minutes at 50 ºC. The desired lipid mixture was then spread on the PVA surface. The solvent was evaporated for 15 minutes under vacuum. After evaporation of the solvent, the appropriate buffer solution was added, allowing for GUV formation for 60 minutes at room temperature. After the formation, GUVs were transferred to a µ-slide chamber with the appropriate coating and left to rest for 10 minutes to allow for GUV deposition and immobilization. In order to immobilize the GUVs through the interaction with the biotinylated lipid and minimize nonspecific protein adsorption to the surface, a mixed coating of BSA+BSA-biotin (9:1 molar ratio) and avidin was used. Failure to perform this passivation of the coverslip surface resulted always in significant PH-YFP adsorption. Ibidi µ-slide chambers were coated with 300µL of a 0.9mg/ml BSA and 0.1mg/mL BSA-biotin mixture for 1 hour. Afterwards the chambers were washed with filtered milliQ water and covered with a second layer of 300µL of 0.01mg/mL Avidin for 1 hour. BSA, BSA-biotin and avidin solutions were prepared with milliQ water. Before adding the GUV solution, the chambers were washed with filtered milliQ water.

Steady-state fluorescence spectroscopy. Fluorescence measurements were carried out with a SLM-Aminco 8100 Series 2 spectrofluorimeter (Rochester, NY) with double excitation and emission monochromators (MC-400), in right angle geometry. The light source was a 450 W Xe arc lamp and the reference a Rhodamine B quantum counter solution. Quartz cuvettes (0.5 x 0.5 cm) from Hellma Analytics were used. Temperature was controlled to 25 °C. Polarization of excitation and emission light was obtained by specific rotation of Glan-Thompson polarizers. Blank subtraction was taken into account in all fluorescence intensity and anisotropy measurements.

Steady-state fluorescence anisotropy, $< r >$ is defined as[43]:

$$< r > = \frac{I_{VV} - G * I_{VH}}{I_{VV} + 2 * G * I_{VH}} \; ; \; G = \frac{I_{HV}}{I_{HH}}$$

where I_{Vj} represents the steady-state vertical (parallel, I_{VV}) and horizontal (perpendicular, I_{VH}) components of the fluorescence emission with vertically polarized excitation. The G factor is measured using the vertical (I_{HV}) and horizontal (I_{HH}) components of the fluorescence emission with horizontaly polarized excitation.

PH-YFP partition to LUVs was followed by steady-state fluorescence intensity and anisotropy measurements, by exploiting the fluorescence emission of the attached YFP protein. PH-YFP membrane interactions with LUVs were studied in the presence of different lipid concentrations (up to 100 μM) and in the presence and absence of Ca^{2+}. LUVs used in these experiments were composed of POPC and varying molar ratios of (18:1)PI(4,5)P_2 (99:1, 97:3 and 95:5 molar ratios). 100 nM of PH-YFP was incubated for 5 minutes at 37ºC with the different lipid concentrations. The experimental buffer used for the PH-YFP experiments consisted of 10 mM Na2HPO4, 140 mM NaCl at pH 7.4. 5mM of EDTA was included in the experiment buffer to study the interactions in the absence of Ca^{2+}. To study the influence of Ca^{2+}, 100 μM of CaCl2 was included in the experimental buffer.

PBP-10 partition to PI(4,5)P_2 or PS containing LUVs, in the presence and absence of Ca^{2+}, was followed by steady-state fluorescence measurements by exploiting the strong reduction of fluorescence quantum yield that occurs whilst it binds to membranes containing acidic lipids[38,39].

PBP-10 partition to LUVs was studied in the presence of different lipid concentrations (up to 400 μM) and in the presence and absence of Ca^{2+}. LUVs used in these experiments were composed of POPC: (18:1) PI(4,5)P_2 (95:5 molar ration) or POPC:PS (80:20 molar ratio). 500 nM of PBP-10 was incubated for 5 minutes at 37 ºC with the different lipid concentrations. The experimental buffer used for these experiments consisted of 10 mM Hepes, 140 mM NaCl at pH 7.4. 5 mM of EDTA was included in the experiment buffer to study the interactions in the absence of Ca^{2+}. To study the influence of Ca^{2+}, 100 μM of $CaCl_2$ was included in the experiment buffer. All experiments were performed in triplicate using independent samples. PBP-10 was excited at 545 nm and fluorescence emission was collected from 565 to 650 nm. These measurements were carried out at 37ºC. The decrease in fluorescence intensity with increasing lipid concentrations can be used to calculate the partition coefficients of PBP-10 to LUVs, using the following equation[44]:

$$I = \frac{I_w + K_p \gamma_L [L] I_L}{1 + K_p \gamma_L [L]}$$

Where, I is the Integrated normalized fluorescence intensity, K_p is the partition coefficient between the lipid and aqueous phases, γ_L is the lipid molar volume, and [L] is the total lipid concentration. The subscripts W and L stand for the aqueous and the lipid phase, respectively. This equation can be fitted to the experimental data in order to recover the partition coefficients.

Förster resonance energy transfer measurements. All FRET measurements were carried out using the same optical setup as the previous fluorescence intensity measurements. In a first set of samples, no FRET acceptor was included, and the fluorescence emission spectra was traced, to obtain the fluorescence intensity of the donor in the absence of FRET. In a second set, the FRET acceptor was included in varying concentrations, and the fluorescence intensity of the donor in presence of FRET was determined. FRET efficiencies were calculated by evaluating the decrease of donor quantum yield through the changes in steady-state fluorescence intensity, as shown by the following equation:

$$E = 1 - \frac{I_{DA}}{I_D}$$

where the indexes DA and D refer to the spectral parameters of the donor in the presence or in the absence of acceptor, respectively.

FRET studies of PH-YFP interaction with PI(4,5)P_2 containing LUVs were carried out in similar experimental conditions as the corresponding fluorescence intensity measurements. 100 nM of PH-YFP was incubated with 100 nM of LUVs, for 5 minutes at 37ºC. The LUVs were composed of POPC and increasing concentrations of PI(4,5)P_2, ranging from 1 to 2.5% molar ratio. Rho-DOPE was added at 1% molar ratio to act as the FRET acceptor. Increasing Ca^{2+} concentrations were then tested.

FRET studies of PBP-10 interaction with TF-PI(4,5)P_2 and Bodipy-PC containing LUVs were carried out in similar experimental conditions as the corresponding fluorescence intensity measurements. Varying concentrations of PBP-10 were incubated with LUVs prepared with different lipidic compositions for 5 minutes at 37ºC. The LUVs used in these experiments were composed of either POPC:PI(4,5)P_2:TF-PI(4,5)P_2 (95:4.5:0.5 molar ratio) or POPC:Bodipy-PC:PI(4,5)P_2 (94.5:0.5:5 molar ratio). It is worth emphasizing that the mol% total of PI(4,5)P_2 (labelled and unlabelled) is the same for both lipid compositions (5 mol%). The FRET donors for this experiment were either TopFluor-PI(4,5)P_2 or Bodipy-PC, and PBP10 acted as the FRET acceptor for both experiments. The experiments were carried out at 250 μM total lipid concentration and at several PBP-10 concentrations, up to 20 μM. Both TF-PI(4,5)P_2 and Bodipy-PC were excited at 480 nm and fluorescence emission was collected from 490 to 600 nm.

Photobleaching Assay. Photobleaching of protein and liposome samples was performed using a Xe (450 W) light source, focused onto the sample via a magnifying glass. The samples were exposed to the light source for varying durations, in order to obtain different relative photobleaching percentages. Fluorescence anisotropy was then measured as previously described. We could then fit a model that

predicts the formation of oligomers to the obtained rate at which fluorescence anisotropy recovered with photobleaching percentage and attempt to determine the PH-YFP oligomerization number. This model is based on a binomial distribution, where it predicts the different fractions of still fluorescent oligomers for a given oligomerization number and photobleaching percentage.

Confocal fluorescence microscopy. Confocal laser scanning fluorescence microscopy measurements were performed on a Leica TCS SP5 (Leica Microsystems CMS GmbH, Manheim, Germany) inverted confocal microscope (DMI600). Excitation lines provided by an argon laser were focused into the sample through an apochromatic water immersion objective (63x, NA 1.2; Zeiss, Jena Germany). A 111.4 μm diameter pinhole in front of the image plane blocked out of focus signal. Images were acquired at 100 Hz, exciting PH-YFP at 488nm and collecting emission between 500 and 600 nm. For each GUV, an image was taken roughly at the equatorial plane. Membrane fluorescence and GUV radius were automatically quantified making use of a homemade MATLAB script. No correlation between GUV radius and membrane fluorescence was found.

Fluorescence fluctuation spectroscopy. FFS measurements were carried out using the same optical path described for the confocal imaging, except that fluorescence emission was detected using avalanche photodiodes (APDs) after passing through a 500-550 band-pass filter. Excitation of PH-YFP was performed with the 488 nm Ar laser line. For samples in solution, the focal volume was focused ~100μm above the top surface of the cover slide and five autocorrelation curves were sequentially obtained with an acquisition time of 60 seconds at a 500kHz sampling frequency. Fluorescence fluctuations from GUVs with bound fluorescent proteins were recorded from the top of the vesicle, with the focal volume centered in the focal plane with maximum fluorescence intensity. Five autocorrelation curves were sequentially obtained for each sample with an acquisition time of 20 seconds at a 100kHz sampling frequency.

Fluorescence correlation spectroscopy. For FCS measurements performed in simple samples in solution, a model that takes into account three-dimensional translational diffusion and triplet state formation for *n* diffusing species was used[45–47]. Where $G(\tau)$ is modelled by

$$G(\tau) = 1 + \frac{1}{N}\cdot\left(\frac{1 - T + T.e^{\frac{-\tau}{\tau_T}}}{1-T}\right)\cdot\left\{\sum_{i=1}^{n}\frac{f_i}{\left(1+\frac{\tau}{\tau_D}\right)\sqrt{1+\tau(S^{-2}\tau_D)}}\right\}$$

Where the axial (ω_z) to lateral (ω_{xy}) dimension ratio, S, is given by $S = \frac{\omega_z}{\omega_{xy}}$, and the translational diffusion time, τ_D, is given by $\tau_D = \frac{\omega_{xy}^2}{4D}$. N corresponds to the average number of fluorescent particles in the observation volume, f_i is the fractional population of *n* different diffusion species, τ_T is the triplet lifetime and T corresponds to $\frac{f_t}{1-f_t}$, where f_t is the fraction of fluorophores in the triplet state. For the modelling of data obtained from GUV membrane fluorescence, a model that considers two-dimensional translational diffusion through a Gaussian observation volume was used[48], described by

$$G(\tau) = \frac{1}{C\pi\omega_{xy}^2}\left(1+\frac{4D\tau}{\omega_{xy}^2}\right)^{-1} = \frac{1}{N}\left(1+\frac{\tau}{\tau_D}\right)^{-1}$$

where, C corresponds to the particle concentration. The dimensions of the focal volume were determined by calibration with rhodamine 110 in milliQ water using the same optical setup as the samples. A rhodamine 110 diffusion coefficient (D) of 440 $\mu m^2 s^{-1}$ was considered[49]. Analysis of all the FCS experimental data was carried out using the ISS Vista software. This program uses a Marquardt–Levenberg nonlinear least-squares fitting routine and the goodness of the fittings can be judged by the recovered χ^2 value and the random distribution of the weighted residuals.

Photon counting histogram. Recorded fluorescence fluctuation data can be used to generate photon counting histograms that access information on the molecular brightness, ε, and average number of particles. To describe the PCH of a freely diffusing species, we must first consider the case of a single, diffusing particle enclosed in a small volume V_0, where over a long enough observation time, the probability of the particle appearing at any position within V_0 will be the same. Combining the appropriate probability distribution of particle diffusion and PSF with Mandel's Formula, that describes the photon counting statistics for the semi-classical model of photon detection, we obtain the probability of detecting k > 0 photons for one fluorescent particle in V_0[50,51].

$$p^{(1)}(k;V_0,\varepsilon) = \frac{1}{V_0}\int Poisson\,[k,\varepsilon\cdot W(\overrightarrow{r})]d\vec{r}\;;\;\; Poisson(k,\lambda) = \frac{\lambda^k e^{-\lambda}}{k!}$$

This equation describes the PCH of a single diffusing fluorescent particle. It is the weighted average of Poissonian distributions for k counts. $W(\overrightarrow{r})$ is the function that describes the observation volume profile, considers excitation laser intensity and the detection efficiency as a function of the particle position. ε is the molecular brightness where:

$$\varepsilon = I_0^n \beta \eta_W T_s$$

The molecular brightness characterizes the average number of detected photons per sampling time (T_s) per molecule. Through the parameter β, the molecular brightness takes into account the particle fluorescence quantum yield, extinction coefficient, and all of the instrument-dependent factors. The ratio $\varepsilon_{sec} = \varepsilon/T_s$ is used as it allows a more convenient comparison between experiments. ε_{sec} is expressed in counts per second per molecule (cpsm)[51]. After defining the PCH for a single diffusing particle, we then consider the possibility of finding N identical but independent particles that diffuse within V_0. The distribution of the photon counts (k), will then be given by the Nth convolution of the distribution defined for a single particle, p, with itself:

$$p^{(N)}(k;V_0,\varepsilon) = (p^{(1)}\otimes p^{(1)}\otimes \ldots \otimes p^{(1)})(k;V_0,\varepsilon);$$

The case where no particles are present in the reference volume and no photons are detected is also considered.

$$p^{(0)}(k;V_0,\varepsilon)=\begin{cases}1, k=0\\0, k\neq 0\end{cases}$$

If we account for the fluctuation of the number of particles N inside V_0, given a certain concentration, c, then the probability of finding N particles in V_0 follows a Poisson distribution with the mean value cV_0. However, while c is adequate to characterize the concentration of particles, it is more straightforward to use the average of number of particles, $\overline{N}$, in a certain volume V, as a fitting parameter. In this way, the PCH for a single species is given by:

$$P(k;\overline{N},\varepsilon)=\sum_{N=0}^{\infty}p^{(N)}(k;Q,\varepsilon)\cdot Poisson(N,Q\overline{N});$$

$$Q=\frac{V_0}{V}$$

The volume ratio Q serves as a numerical simplification, cancelling the units of volume in the PCH integration, as $p^{(N)}(k;V_0,\varepsilon)$ is independent from the choice of V_0 as long as it is large enough to obtain a positive probability of zero photon counts[50,51]. In all our data analysis we arbitrarily chose Q = 6. To account for deviations of the experimental observation volume from the 3D Gaussian approximation, Huang et al.[50] implemented a correction that takes into consideration the contribution to the PCH integral from the out-of-focus region. This correction introduced a new parameter, F, that accounts for the fraction of detected photons from the non-Gaussian part of the observation volume. In this way, after the correction of the 3D Gaussian approximation, the PCH model for a confocal microscope with one-photon excitation is described by the following equations:

$$P(1;Q,\varepsilon)=\frac{1}{(1+F)^2}\left[p_G(1;Q,\varepsilon)+\frac{\varepsilon F}{2\sqrt{2Q}}\right]$$

$$P(k;Q,\varepsilon)=\frac{1}{(1+F)^2}p_G(k;Q,\varepsilon)\text{ for }k>1$$

and,

$$P(0;Q,\varepsilon)=1-\sum_{k}^{\infty}p\quad(k;Q,\varepsilon)$$

where,

$$p_G(k;Q,\varepsilon)=\frac{1}{Q\pi^{1/2}k!}\int_0^{\infty}\gamma\left(k,\varepsilon e^{-2x^2}\right)dx$$

p_G is the analytical expression for the PCH integration using a 3D Gaussian approximation for one-photon excitation and confocal detection and γ is the incomplete Gamma function. For more than two species, differences in photophysical properties, like quantum yield, are considered and included in the coefficient ε that will be different for each species. This allows us to detect and quantify the presence of different species with different brightness, such as monomers and oligomers in a sample. In the case of multiple independent species with different brightness, the photon counting histogram can be obtained by convoluting the photon counting histograms of the individual species:

$$P(k;\overline{N}_1,\varepsilon_1,\overline{N}_2,\varepsilon_2\ldots\overline{N}_n,\varepsilon_n)$$
$$=P(k;\overline{N}_1,\varepsilon_1)\otimes P\left(k;\overline{N}_2,\varepsilon_2\right)\otimes\ldots\otimes P(k;\overline{N}_n,\varepsilon_n)$$

For PH-YFP samples, these equations were globally fitted to the experimental data using a homemade MATLAB implementation by considering the presence of two populations with different brightness, ε_1 and ε_2, that correspond to the monomeric and oligomerized form of the protein, respectively. In this analysis, we consider that the monomer brightness does not change from sample to sample by linking ε_1 across all the curves obtained. The correction factor, F, can be determined by a calibration procedure using 5 nM of rhodamine 110 and subsequently fixed for the other samples measured in the same chamber slide.

Using the brightness and number of particles, obtained through PCH analysis of each data set, we calculated the fraction weighed by brightness (F_i), given by the following equation:

$$F_i=\frac{N_i\varepsilon_i}{\sum_1^i N_i\varepsilon_i},$$

where, N_i and ε_i refer to the number of particles and brightness, respectively, of population i. The fraction weighed by brightness can be used to quantify the total contribution of each population towards the total fluorescence collected. Here, we will use it to quantify each population of fluorescent proteins more accurately.

Molecular dynamics simulations. All coarse-grain simulations were modelled using the Martini2.2 forcefield[52] and run with GROMACS version 2018. The $PI(4,5)P_2$ topologies used were made by combining an in-house improved version of the existing $PI(4,5)P_2$ headgroup topology[53,54] with the standard lipid acyl-chain topologies for the composition studied (18:1). All other topologies were obtained directly from the Martini website. All simulations were run making use of the standard martini parameters. After initial energy minimization and equilibrium runs, the simulations were run at a 20 fs time step. Nonbonded interactions were cutoff at 1.1 nm and long-range interactions treated using the Verlet-neighborlist scheme. The V-rescale thermostat was used with a tau-t of 4.0, to maintain the temperature at 300 K. Constant pressure was maintained at 1.0 bar by semiisotropic pressure coupling to a parrinello-rahman barostat with a relaxation time of 16.0 ps.

All protein and peptide systems were run in membrane systems prepared separately. The membrane systems were built using the INSANE CG building tool[55]. The systems consisted of approximately 500 lipids at a 95:5 POPC:$PI(4,5)P_2$ molar ratio, in small 15 X 10 X 13 nm boxes. The systems were neutralized with counterions and then a further 140 mM of NaCl was added. These were simulated for up to 10 μs, either in the absence or in the presence of Ca^{2+}, at Ca^{2+}:$PI(4,5)P_2$ ratios of 0.5, 1, 3, 5 and 10.

PBP-10 simulations. For PBP-10 parameterization, a well resolved NMR structure of the $PI(4,5)P_2$ binding site of gelsolin was selected as the template (PDB:1SOL)[56]. The last 10 residues of this structure were used, as they are the ones that compose the peptide. The structure was converted to the Martini2.2 coarse-grain force-field using the martinize tool[57,58]. Bonded potentials between consecutive residues sufficed to restrain the protein secondary structure; the use of an elastic network was not necessary. When compared to the peptide used in the experimental setups, the modelled peptide doesn't contain the covalently bound rhodamine B molecule. The peptide is known to interact with $PI(4,5)P_2$ regardless of the fluorescent molecule attached; however, without it, it loses its ability to cross the bilayer[38]. For the PBP-10 simulation setup, 20 peptides were added to each leaflet of a previously prepared membrane bilayer. These

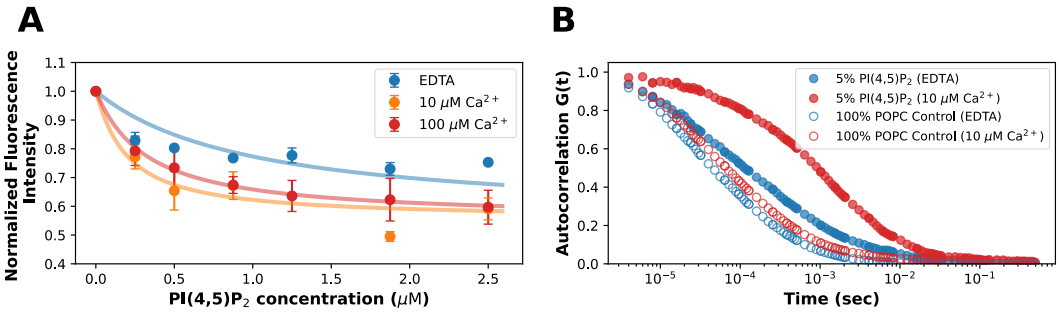


Figure 1. Low concentrations of calcium increase PH-YFP affinity towards LUVs containing 5% $PI(4,5)P_2$. (A) PH-YFP affinity towards POPC:Rho-DOPE:$PI(4,5)P_2$ (94:1:5 molar ratio) 100 nm LUVs was determined through the FRET efficiency between PH-YFP and Rho-DOPE in the presence of various calcium concentrations: 5 mM EDTA (blue), 10 μM Ca^{2+}(orange) and 100 μM Ca^{2+} (red). Dissociation constants were obtained by fitting a model assuming steady-state bimolecular binding and constant FRET efficiency in the membrane to the FRET curves. The lines represent the fit with this model. All samples were incubated with 100 nM of PH-YFP for 5 minutes. Error bars represent the standard deviation from at least three independent experiments. (B) Partition of 100 nM PH-YFP to 400 uM of POPC:$PI(4,5)P_2$ (95:5 molar ratio) LUVs in the absence (blue) and presence (red) of 10 μM Ca^{2+}. Partition was followed by FCS through the differences in PH-YFP diffusion properties caused by increasing lipid concentrations. The controls for Liposome fusion are also shown in empty markers.

conditions mimic the experimental setup of the FRET experiments, which aimed for saturation of the membrane with peptides. Additional counterions were added to account for peptide charge. Initially, the peptides were disposed as a grid on top of the membrane, equally distant from each other and restrained in the XY axis to allow for their incorporation into the membrane before aggregation outside of the membrane could occur[59]. Once all peptides were fully incorporated in the membrane, the restraints were removed, and the system was simulated normally.

PH domain simulations. For PH domain parameterization, a well resolved X-Ray diffraction structure of the pleckstrin homology domain from phospholipase C delta1 was used as the template (PDB: 1MAI at 1.9A resolution)[28]. The PH domain in this structure is in complex with an inositol trisphosphate molecule (I3P), which gives us a good starting conformation for studying $PI(4,5)P_2$ binding. The structure was converted to the martini coarse-grain force-field using the martinize[57,58] tool where the ElNeDyn elastic-network was used to maintain the protein secondary structure. For the free PH domain simulations, a PH domain was added to previously prepared membranes. The protein was added in a corner of the PBC box, and the protein restrained so that the $PI(4,5)P_2$ binding site would always be facing the membrane whilst probing its surface. Once a binding event had occurred, we lifted the restraints and the system was simulated normally. Facilitated binding simulations were also ran when the free simulations did not show any binding events (i.e. high Ca^{2+} to $PI(4,5)P_2$ ratios). In these, apart from the already in place restraint to keep the binding site facing the membrane, the PH domain was also restrained so that the centre of geometry of the binding site was always located on top of at least one $PI(4,5)P_2$ lipid. Additionally, a restraint was placed on the Z axis so that the protein centre of mass (COM) was never more than 4 nm away from the $PI(4,5)P_2$ phosphate layer.

Umbrella sampling simulations of the PH domain binding to $PI(4,5)P_2$ lipids were performed at various Ca^{2+} to $PI(4,5)P_2$ ratios, representing the experimental conditions used (0:1, 0.5:1 and 10:1 Ca^{2+} to $PI(4,5)P_2$). The collective variable (CV) was chosen to be the distance between the $PI(4,5)P_2$ PO4 bead and the COM of the PH domain. 12 umbrella sampling windows were setup with the CV spaced linearly between 3.75 and 1.00 nm, with a force constant of 300 kJ/mol being applied to the windows in the 3.75 to 2.25 nm range and a force constant of 600 kJ/mol being applied to the remaining. As with the facilitated binding simulations, to make sure that the PH domain interacts with the membrane via the $PI(4,5)P_2$ binding pocket, a restraint was applied so that the $PI(4,5)P_2$ binding site would always be facing the membrane surface. Similarly, a restraint in the XY plane was put in place to make sure that the PH domain would always be on top of at least one $PI(4,5)P_2$ lipid. To avoid membrane deformation or curvature issues, a potential was applied to all PO4 beads to restrain membrane travel in the ZZ axis to +/- 0.7 nm. The umbrella sampling runs in the presence of Ca^{2+} were performed both in membranes with pre-clustered $PI(4,5)P_2$ nanodomains and with simultaneous clustering of $PI(4,5)P_2$ nanodomains to account for hysteresis effects. The starting positions for each window were picked from an MD simulation where the PH domain started at 4.00 nm from the membrane and a restraint was applied in a step-wise fashion to approach it to the membrane. For the umbrellas ran in the presence of pre-clustered $PI(4,5)P_2$ nanodomains or in the absence of Ca^{2+}, membranes pre-simulated in the respective conditions were used. For the umbrellas ran with simultaneous clustering of $PI(4,5)P_2$, the starting positions were obtained from a simulation with a membrane pre-simulated in the absence of Ca^{2+}. Ca^{2+} was then added in the appropriate ratio replacing random water beads. After a quick minimization and relaxation step with the restraints in place, the window was run. For each window 3 replicas were ran for at least 1 μs each. Umbrella sampling unbiasing and error calculation was performed using Grossfield's implementation of WHAM[60]. All restraints and sampling methods were carried out using the open-source, community-developed PLUMED library[61], version 2.5[62].

RESULTS

Low physiological Ca^{2+} concentrations increase PH-YFP affinity towards $PI(4,5)P_2$ containing model membranes. The strong interactions between Ca^{2+} and $PI(4,5)P_2$ that have recently been reported, have been shown to influence $PI(4,5)P_2$ electrostatics as well as membrane organization, even at physiological concentrations. To track the impact of Ca^{2+} on $PI(4,5)P_2$-binding protein dynamics we made use of a fluorescent fusion protein, PH-YFP, as a reporter. PH-YFP is a fusion protein that contains an isolated PLCδ1 PH domain linked to the fluorescent protein variant, YFP. Apart from the $PI(4,5)P_2$ binding surface, that faces the membrane, the PH domain surface is fairly negatively charged and, so far, no tendency for oligomerization has been reported. On the other hand, the YFP variant used has been reported to be a weak dimer[63], which dimerizes when present at high concentrations. As a result, the PH-YFP construct, also showed a tendency to oligomerize, especially when bound to the membrane due to the local concentration effect. With the above mentioned in mind, PH-YFP turned out to be an excellent protein model to use as a case study of $PI(4,5)P_2$ modulation of protein-protein and protein-lipid interactions.

We started by characterizing the influence of low Ca^{2+} concentrations on PH-YFP affinity towards $PI(4,5)P_2$ containing LUVs through a FRET methodology (Figure 1, A). In this experiment, PH-YFP acted as the energy transfer donor, and Rho-DOPE acted as the energy transfer acceptor. This experiment allowed us to characterize the binding distribution of PH-YFP on the membrane, by analysing the degree of decrease in donor fluorescence emission, caused by the presence of acceptors in its proximity. Binding of PH-YFP to LUVs containing 5% $PI(4,5)P_2$ lead to a significant decrease in donor emission, in the presence and absence of Ca^{2+}. These quenching curves were then used to quantify PH-YFP binding, by fitting them to a model assuming steady-state bimolecular binding and constant PH-YFP FRET efficiency in the membrane. This allowed us to recover dissociation constants, K_d , for each experimental condition.

The presence of Ca^{2+} led to an overall increase in PH-YFP partition towards the LUVs, as shown by the decrease in dissociation constant observed when comparing the EDTA sample against the 10 μM Ca^{2+}sample. ($K_{d,EDTA} = 0.93 \pm 0.08$ μM , $K_{d,10\ \mu M\ Ca^{2+}} = 0.17 \pm 0.02$ μM). However, further increasing the Ca^{2+} concentration did not yield as significant changes ($K_{d,10\ \mu M\ Ca^{2+}} = 0.17 \pm 0.02$ μM, $K_{d,100\ \mu M\ Ca^{2+}} = 0.29 \pm 0.02$ μM). This is an interesting result, as Ca^{2+} does not interact directly with the PH domain to influence membrane binding. Therefore, the observed changes are most likely due to Ca^{2+} influencing $PI(4,5)P_2$ biophysical properties, favouring lipid-protein interactions. It is worth noting that the recovered K_d values obtained in the absence of Ca^{2+} are lower than those described in the literature for free PLCδ1 PH binding to POPC:$PI(4,5)P_2$ (98:2) LUVs ($K_d = 5$ μM determined from Hummel and Dreyer type measurements, $K_d = 2.5 \pm 1.3$ μM determined from Centrifugation experiments with Sucrose-loaded LUVs)[64]. These differences are likely associated with the different $PI(4,5)P_2$ levels.

We also followed PH-YFP partition towards POPC:$PI(4,5)P_2$ (95:5) LUVs, using FCS (Figure 1, B). FCS allowed us to analyse how PH-YFP diffusion properties varied with increasing lipid concentrations and compare these effects in the presence and absence of Ca^{2+}. As expected, whether in the presence or absence of Ca^{2+}, increasing concentrations of lipid led to an overall increase in PH-YFP autocorrelation at longer times, clearly showcasing the presence of a slower diffusing population of PH-YFP bound to the LUVs. However, when comparing both sets of results, we can determine that 10 uM of Ca^{2+} leads to an apparent significant increase in PH-YFP partition onto the membrane as it causes the most amount of autocorrelation at longer times. This data is in agreement with the previous FRET studies and confirms that low concentrations of Ca^{2+} cause an increase in PH-YFP binding. Liposome fusion controls were performed with a BODIPY-PC membrane probe, which showed that at this Ca^{2+} concentration, no substantial changes in liposome diffusion times are observed, as expected in the absence of aggregation. Higher concentrations of Ca^{2+} were not tested as these induce liposome fusion.

Overall, we can conclude that low Ca^{2+} concentrations lead to a significant increase in PH-YFP binding affinity for $PI(4,5)P_2$. Notably, this increase in binding affinity occurs even though Ca^{2+} is simultaneously causing significant charge shielding of the $PI(4,5)P_2$ headgroup which should reduce electrostatic interactions. These results appear to suggest that besides inducing cluster formation and influencing $PI(4,5)P_2$ acyl chain order, Ca^{2+} can also induce changes in the orientation of the headgroup, possibly altering lipid recognition and docking of proteins. This effect should be especially important, if the protein binds to $PI(4,5)P_2$ by interacting specifically with the headgroup, as is the case of the PH domain.

High Ca^{2+} concentrations reduce PH-YFP bound to $PI(4,5)P_2$ containing GUVs. Fluorescence confocal microscopy allowed for a simple method to analyse the effects of higher Ca^{2+} concentrations on model membranes without much of the noise caused by Ca^{2+}-induced vesicle aggregation and fusion. The fluorescence of GUVs containing 5% $PI(4,5)P_2$ (POPC:$PI(4,5)P_2$ at a 95:5 molar ratio) incubated with 250 nM of PH-YFP and increasing concentrations of Ca^{2+} and quantified the mean membrane fluorescence for each GUV. Analysing the obtained data (Figure 2, B), we can observe that Ca^{2+} induces a significant increase in membrane fluorescence from 20 to 100 μM Ca^{2+} which is in agreement with the previous FRET and FCS studies. However, from 150 μM Ca^{2+} to 300 μM Ca^{2+}, PH-YFP membrane fluorescence continuously decreases to values slightly lower than those obtained in the absence of Ca^{2+}, in a reproducible manner. Similar results that have showcased a decrease in PH domain affinity at high Ca^{2+} concentrations have been published recently for PLCδ1 PH at 1mM Ca^{2+} [37], while other authors have also shown that PLC activity is inhibited at non-physiological Ca^{2+} concentrations (>100 μM) by an unknown mechanism[36,65,66]. Here we suggest two mechanisms for this inhibition at higher Ca^{2+} concentrations.

The first mechanism is likely charge shielding of the $PI(4,5)P_2$ headgroup by Ca^{2+}, inhibiting interactions with binding proteins. As Ca^{2+} concentrations increase, this effect

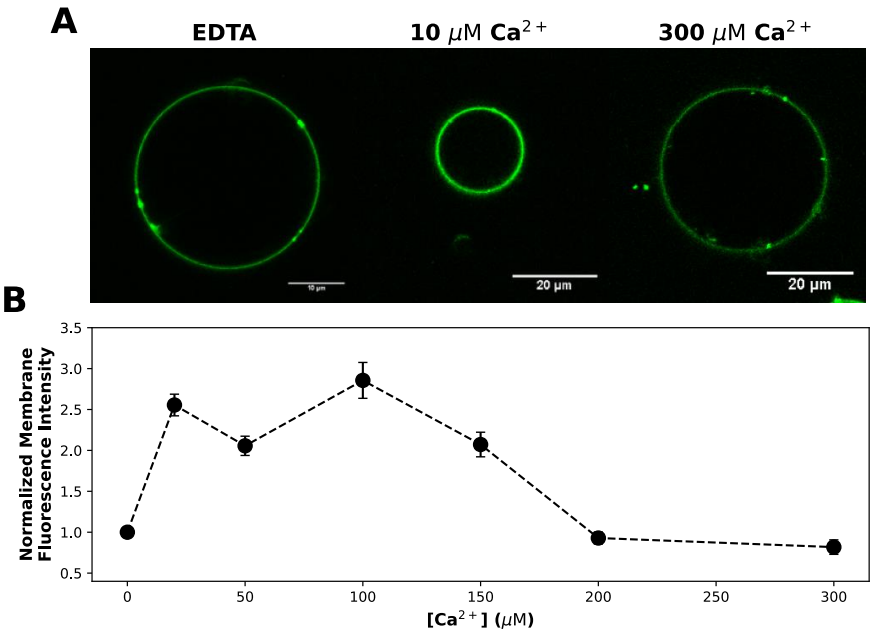


Figure 2. High calcium concentrations lead to a decrease of PH-YFP affinity towards $PI(4,5)P_2$ containing GUVs. Images taken from the equatorial plane of GUVs at increasing Ca^{2+} concentrations (A): (5 mM EDTA; 100 μM Ca^{2+} and 300 μM Ca^{2+}). (B) Overall analysis of the mean PH-YFP membrane fluorescence intensity with increasing concentrations of Ca^{2+}. Fluorescence intensity is normalized to the average intensity obtained for 5 mM EDTA to allow for easier comparison between experiments. At least 30 GUVs over 4 independent experiments were analysed for each composition, for a total of 338 GUVs. Dashed line is a guide to the eye.

becomes more prevalent and very likely starts competing with the binding of proteins with positively charged binding pockets, such as the PH domain, via screening of the electrostatic interactions. The second, more complex mechanism we propose is the increase in number of $PI(4,5)P_2$ lipids sequestered by the PH domain with increasing Ca^{2+} concentrations. On a fluid membrane containing 5% $PI(4,5)P_2$, 1 PH-YFP molecule can sequester in average 1 $PI(4,5)P_2$ lipid, however, if significant $PI(4,5)P_2$ clustering occurs, we can estimate, taking into account the PH domain membrane surface area, that 1 PH-YFP molecule can sequester up to 15 $PI(4,5)P_2$ lipids in average. This could cause a decrease in membrane fluorescence by limiting the amount of $PI(4,5)P_2$ available to bind PH-YFP, leading to a lower concentration of bound protein. However, this effect is only valid if PH-YFP is unable to remove the bound $PI(4,5)P_2$ from within the clusters. Protein sequestration is thought to be one of the causes for the observed non-homogeneous distribution of $PI(4,5)P_2$ in the plasma membrane[5]. Thus, proteins would act as reversible buffers, binding much of the $PI(4,5)P_2$ present and then releasing it locally in response to specific signals[67]. Proteins such as MARCKS[68,69,70], GAP43[67,71], CAP23[67] and syntaxin-1A[72] have been shown to be able to sequester $PI(4,5)P_2$ in such a manner. In fact, studies show that MARCKS alone could reversibly sequester much of the $PI(4,5)P_2$ in the plasma membrane[68]. MARCKS sequestration of $PI(4,5)P_2$ has been shown to be important in the $PI(4,5)P_2$ mediated activation of TRPC-family Ca^{2+} channels[73], in the endocytosis of the amyloid precursor protein (APP)[74] and in the synaptic clustering of $PI(4,5)P_2$[75]. However, the influence of Ca^{2+} on protein sequestration of $PI(4,5)P_2$ has not been given any attention thus far.

MD simulations showcase the effect of Ca^{2+} on PH domain $PI(4,5)P_2$ binding affinity. To better understand the origin of the complex effects of Ca^{2+} on PH domain affinity, we performed coarse grained molecular dynamics simulations using the Martini2.2 model. These were performed at varying Ca^{2+} to $PI(4,5)P_2$ ratios (0.5, 1, 3, 5, 10) to represent the range of experimental Ca^{2+} concentrations used. We started by running simulations with systems containing a PH domain and a previously relaxed and simulated membrane composed of $PI(4,5)P_2$ and POPC (5:95 mol ratio) in the presence of varying Ca^{2+} to $PI(4,5)P_2$ ratios (0, 0.5, 3, 5, 10). In the absence of Ca^{2+} and in the presence of low Ca^{2+} ratios, the PH domain easily bound $PI(4,5)P_2$ lipids and did not release the lipid over the course of the simulation (at least 10 μs.). At these ratios, the $PI(4,5)P_2$ binding pocket forms a tight cage around the $PI(4,5)P_2$ headgroup (Figure 3, B), where the positively charged side chains wrap around the negatively charged $PI(4,5)P_2$ headgroup. In this tight binding 3 LYS and 1 ARG sidechains interact with the headgroup phosphates whilst being locked inside the protein pocket via the interaction of a TRP and 2 ARG side chains with the phosphodiester phosphate. This tight cage wrapped around the headgroup is in agreement with previous atomistic studies[76] and similar to the original crystal structure bound to I3P[28] (Figure 3, A). Besides sequestering $PI(4,5)P_2$ in the specific binding pocket, 2 protein loops are also inserted into the membrane at headgroup depth, interacting with the lipid headgroups. This is also in agreement with previous atomistic simulation studies[76] that report the penetration of hydrophobic side chains into the core of the

bilayer and the formation of hydrogen bonds between some PH side chains and PC headgroups. Interestingly, it appears that, even in the absence of Ca^{2+}, these inserted regions promote a significant increase in local concentration of $PI(4,5)P_2$. This is likely due to polar interactions with the SER, MET and VAL sidechains present in the membrane inserted loops. In the presence of higher Ca^{2+} ratios (>3 Ca^{2+}:$PI(4,5)P_2$ ratio), however, the PH domain struggled to bind to $PI(4,5)P_2$, as it was unable to displace the Ca^{2+} in place and only able to bind to the edges of the nanodomains (Figure 3, C). The inability of the positively charged sidechains to displace the bound Ca^{2+} ion weaken protein binding significantly and effectively shield the $PI(4,5)P_2$ lipid from interacting with protein partners. An important observation however, is that the PH domain was not able to remove a $PI(4,5)P_2$ lipid from within the cluster, even at low Ca^{2+} ratios. This is a necessary condition for Ca^{2+} to be able to modulate $PI(4,5)P_2$-Protein interactions via the formation of cation-induced nanodomains.

To get a better picture and to extract the free energy for PH domain binding to $PI(4,5)P_2$ we resorted to umbrella sampling simulations where the reaction coordinate collective variable (CV) was chosen to be the distance between the $PI(4,5)P_2$ phosphodiester (PO4) bead and the COM of the PH domain (Figure 4). To account for hysteresis, the umbrella systems in the presence of Ca^{2+} were run both in membranes with pre-clustered $PI(4,5)P_2$ nanodomains and with simultaneous clustering of $PI(4,5)P_2$ nanodomains. This lets us work around the stickiness (and very long lifetime of exchange) of Ca^{2+} bound to $PI(4,5)P_2$ observed even at low Ca^{2+} ratios, which may reflect some of the limitations of the Martini 2.2 forcefield[77].

In the absence of Ca^{2+}, the PH domain binds to $PI(4,5)P_2$ with a binding free energy of -60 kJ/mol and an optimal distance between the PO4 and the protein COM of roughly 1.7 nm. In the absence of Ca^{2+}, the PH domain $PI(4,5)P_2$ binding pocket wraps around the headgroup tightly, as was seen in the previous MD simulations. At a 0.5 Ca^{2+} to $PI(4,5)P_2$ ratio ran with pre-clustered membranes, we see an increase in energy to -48 kJ/mol and in distance to roughly 2.25 nm. This is caused from the inability of PH domain sidechains to replace Ca^{2+} even at this low concentration. Despite being wrapped around the $PI(4,5)P_2$ headgroup, the PH domain couldn't displace the Ca^{2+} beads that were in place, so they remained bound to $PI(4,5)P_2$ within the binding pocket. This not only leads to electrostatic repulsion between the positively charged sidechains and the Ca^{2+} ion, but also to a less favourable packing of the headgroup in the binding pocket. Using the same Ca^{2+} ratio but with simultaneous clustering of $PI(4,5)P_2$, the landscape changes. The PMF minimum is located at 1.7 nm with an energy value of -62 kJ/mol, showing a marginal increase in affinity when compared to the runs in the absence of Ca^{2+}. In this case, the PH domain bound to the $PI(4,5)P_2$ headgroup first and the Ca^{2+} beads could not displace the positively charged residues of the binding pocket. Ca^{2+} did however lead to the clustering of $PI(4,5)P_2$, which, combined with the intrinsic clustering property of the PH domain, formed a small nanodomain around the protein. It is possible that the increase in local concentration of $PI(4,5)P_2$ promoted by Ca^{2+} could help stabilise PH domain binding, in comparison to the system without Ca^{2+}. However, in this system the effect was very minor.

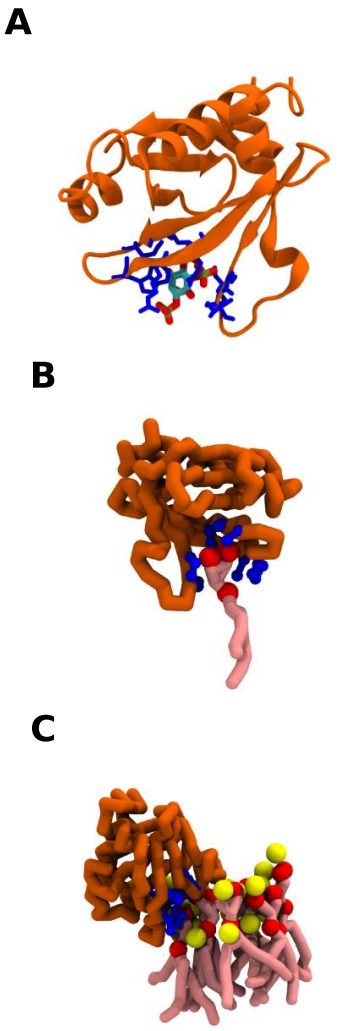


Figure 3. Closeup views of PH domain binding to $PI(4,5)P_2$. (A) PH domain bound to I3P (PDB: 1MAI [28]) (B) CG PH domain bound to a $PI(4,5)P_2$ lipid inserted in a POPC membrane. (C) CG PH domain bound to a $PI(4,5)P_2$ calcium induced nanodomain. The PH domains are represented in orange, with the $PI(4,5)P_2$ binding pocket region side chains being represented in blue. The $PI(4,5)P_2$ lipids are represented in pink, with the 3 phosphates discriminated in red. Ca^{2+} beads are represented in yellow.

At a 10:1 Ca^{2+} to $PI(4,5)P_2$ ratio ran with pre-clustered membranes, we don't observe a clear minimum in the PMF. Approaching the protein to the $PI(4,5)P_2$ containing membrane always results in a very unfavourable increase in energy. With simultaneous clustering, the landscape doesn't change much. Despite there being a minimum at roughly 1.7 nm of -12 kJ/mol, at this concentration Ca^{2+} was able to displace the PH domain binding pocket positive sidechains and restrict the protein to binding the edge of the nanodomains. This is in agreement with the effect that we observe in the experimental results for high Ca^{2+} concentrations and showcases the charge screening effect of Ca^{2+} that we suggest leads to an inhibition of protein binding.

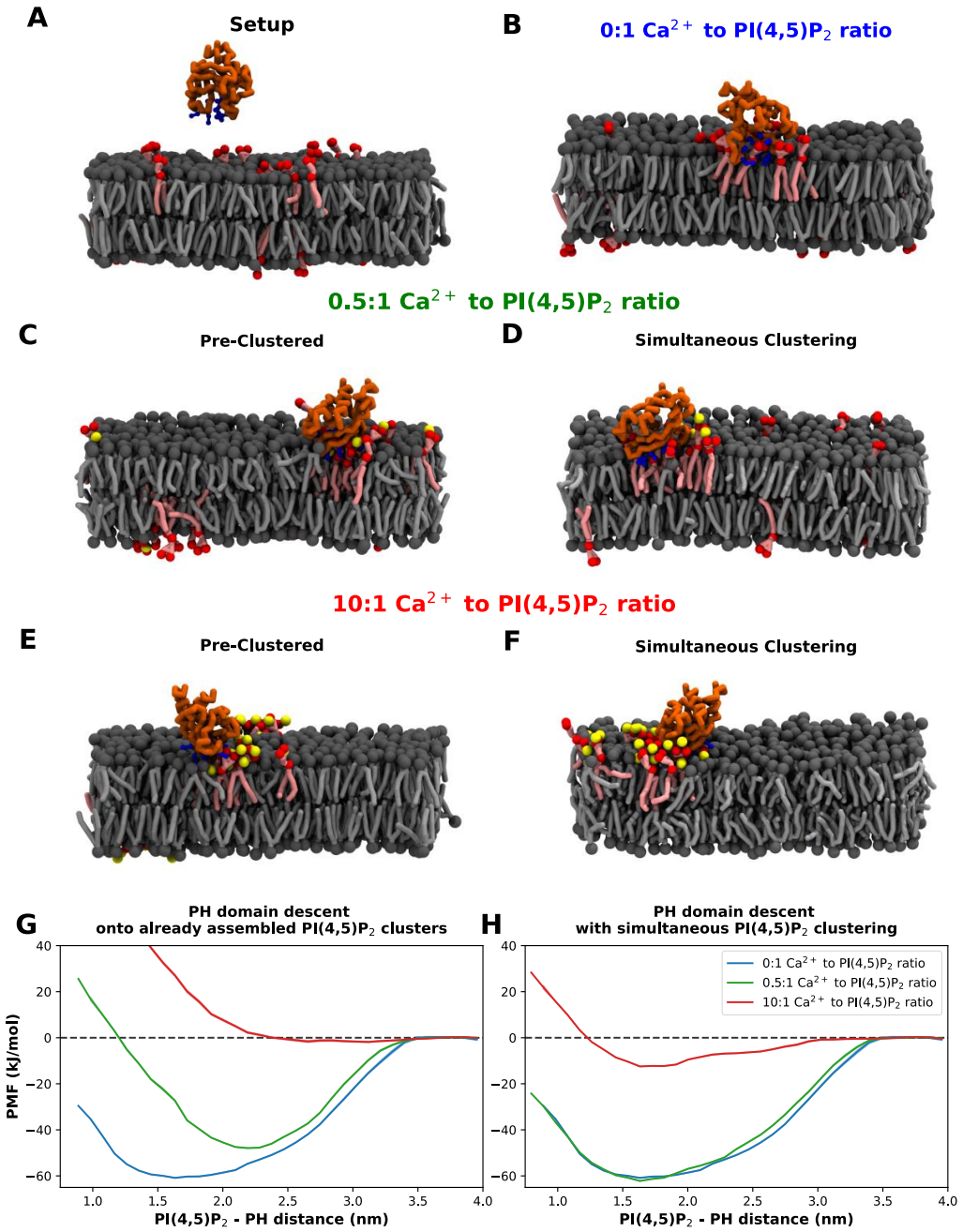


Figure 4. Effect of calcium concentration on PH domain binding affinity by umbrella sampling of CG MD simulations. Different umbrella snapshots showcasing the setup (A) and PH domain binding to $PI(4,5)P_2$ at 1.7 nm under different conditions (B-F). The PH domains are represented in orange, with the $PI(4,5)P_2$ binding pocket region side chains being represented in blue. The $PI(4,5)P_2$ lipids are represented in pink, with the 3 phosphates discriminated in red. POPC lipids are represented in grey. Ca^{2+} beads are represented in yellow. PMF landscapes for PH domain descent onto already assembled $PI(4,5)P_2$ clusters (G) and with simultaneous clustering (H) for the 3 Ca^{2+} ratios tested. The CV chosen is the distance between the $PI(4,5)P_2$ PO4 bead and the COM of the PH domain.

Ca^{2+} decreases PBP-10 partition towards $PI(4,5)P_2$ containing model membranes even at low concentrations. After studying the effect of Ca^{2+} on PH-YFP interactions with $PI(4,5)P_2$-containing membranes, we further investigated a distinct model of $PI(4,5)P_2$ binding protein, PBP-10. Polyphosphoinositide-binding peptide 10 (PBP-10) is a 10 residue long, membrane-permeant polycationic peptide derived from the $PI(4,5)P_2$ binding region

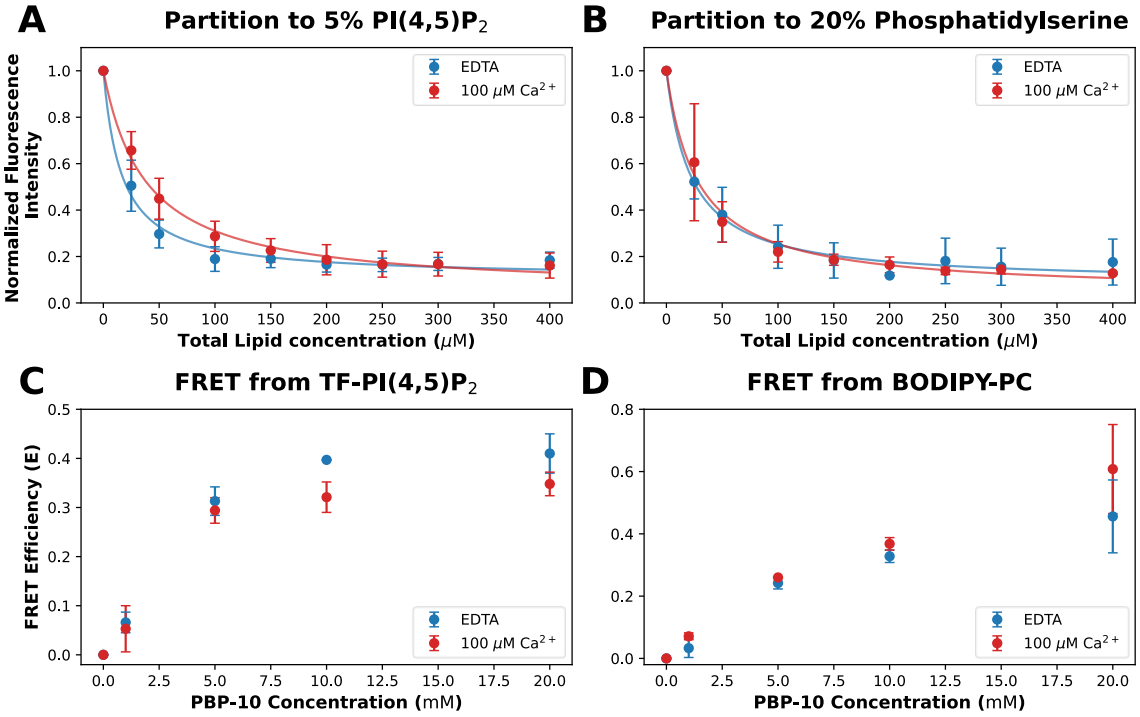


Figure 5. Low concentrations of calcium decrease PBP-10 partition towards $PI(4,5)P_2$ containing LUVs and modulate peptide interactions with membrane components. PBP-10 partition to (A) POPC:$PI(4,5)P_2$ (95:5 molar ratio) LUVs or (B) POPC:POPS (80:20 molar ratio), in the presence of 100μM Ca^{2+} (blue) and in the absence of Ca^{2+} (red). Partition was followed by the decrease in PBP-10 fluorescence intensity when bound to the membrane. The lines represent the fit from where the partition coefficient was recovered. Both lipid compositions were incubated with 500 nM of PBP-10 for 5 minutes. Error bars represent the standard deviation from three independent experiments. FRET efficiencies for the FRET pairs (C) TopFluor-$PI(4,5)P_2$/PBP-10 and (D) Bodipy-PC/PBP10 (donor/acceptor) in the presence and absence of Ca^{2+}. LUVs used in these experiments were composed of either POPC:$PI(4,5)P_2$:TopFluor-$PI(4,5)P_2$ (95:4.5:0.5 molar ratio) or POPC:Bodipy-PC:$PI(4,5)P_2$ (94.5:0.5:5 molar ratio). Total Lipid concentration for both experiments was 250 μM.

in segment-2 of gelsolin (Gelsolin residues 160-169) which was linked to rhodamine B[38,39]. While PH-YFP binds specifically to $PI(4,5)P_2$ by stereochemical recognition of the headgroup, PBP-10 binding to $PI(4,5)P_2$ is much less specific and is governed mostly by electrostatic interactions between its positively charged residues and the negatively charged lipid headgroup. Studying how Ca^{2+} might influence a less specific $PI(4,5)P_2$ binding protein model, will provide further insight on how modulation of $PI(4,5)P_2$ electrostatics and lateral organization in the membrane can influence lipid-protein and protein-protein interactions.

PBP-10 partition to negatively charged membranes was studied using steady-state fluorescence measurements, taking advantage of the linked rhodamine B fluorophore. Upon binding to membranes containing acidic lipids, PBP-10 undergoes a strong reduction in fluorescence intensity, likely caused by the self-quenching of rhodamine B groups as the peptide oligomerizes at the vesicle surface after membrane association[38,39]. Thus, PBP-10 was incubated with POPC:$PI(4,5)P_2$ (95:5 molar ratio) LUVs for 5 minutes, in the presence or absence of Ca^{2+}, and partition was followed through the decrease in fluorescent intensity experienced by PBP-10 when bound to the membrane (Figure 5.). The obtained data was used to calculate the partition coefficient. As already described, the mechanism responsible for the membrane-driven fluorescence quenching of PBP-10 is not fully understood and could be the result of oligomerization of the peptide. If this is the case, then the fluorescence quenching curves would not solely depend on membrane partition but also on the extent of its oligomerization. However, the curves are clearly hyperbolic, indicative of a simpler equilibrium. Also, previous unpublished results from our group show that the peptide membrane bound intensity is constant and not dependent on the density of peptide in the membrane. This suggests that in case some oligomerization takes place, the dissociation constant is extremely small, and a highly homogeneous population of oligomers is present. In this case, although a putative PBP-10 membrane-dependent oligomerization could interfere with membrane partition of the peptide, quenching curves can be analysed with a simple partition model, since all the membrane associated peptide would be in the same state, and only two species need to be considered.

Analysing the results, for LUVs containing 5% $PI(4,5)P_2$ (Figure 5. A), 100 μM Ca^{2+} has a significant effect on the partition of PBP-10 to the membrane. Comparing the data, we observe a 2.4 fold decrease in the partition coefficient in the presence of Ca^{2+}, when compared to the data obtained in the absence of it ($K_{p,EDTA} = (8.1 \pm 1.3) \times 10^4$ to $K_{p,Ca^{2+}} = (3.5 \pm 0.3) \times 10^4$). These results contrast with what was

obtained for PH-YFP, where the same concentration of Ca^{2+} induced an increase in PH-YFP affinity. This experiment was then repeated with LUVs containing 20% POPS (Figure 5. B), to simulate the membrane charge density of 5% $PI(4,5)P_2$ (POPS has a net charge of -1, whilst $PI(4,5)P_2$ has a net charge roughly equal to -4 at the experimental pH of 7.4 [78]). In this case, Ca^{2+} has no significant effect on the partition of PBP-10 to the membrane ($K_{p,EDTA} = (4.7 \pm 0.6) \times 10^4$ to $K_{p,Ca^{2+}} = (6.0 \pm 0.8) \times 10^4$). This result not only confirms that PBP-10 interaction with the membrane is ruled mostly by electrostatics, since we recover similar partition coefficients for membranes prepared with a similar charged density, but also that the effect of Ca^{2+} on PBP-10 partition originates from specific interaction between Ca^{2+} and $PI(4,5)P_2$, as it only occurs in LUVs containing 5 mol% $PI(4,5)P_2$. It also rules out an interaction between PBP-10 and Ca^{2+} in solution as the causer of the lower membrane affinity of the peptide in the presence of the divalent cation. Our proposed explanation for the observed results is based on electrostatic charge shielding caused by the interaction of Ca^{2+} with $PI(4,5)P_2$, which is not so prevalent with PS, given the significantly lower affinity between PS and Ca^{2+}. This effect was only observed for PH-YFP at higher Ca^{2+} concentrations, as it binds to $PI(4,5)P_2$ via specific interactions between the protein and the lipid headgroup which are stronger than the electrostatic interactions in play for PBP-10.

In order to gather a more detailed insight on how Ca^{2+} influences PBP-10 interactions with the membrane, we performed FRET studies between the fluorescent analogues, BODIPY- PC and TopFluor-$PI(4,5)P_2$, and PBP-10. On both experiments, we tested the effect of the presence and absence of Ca^{2+} on the FRET efficiency of each pair. LUVs used in these experiments were composed of either POPC:$PI(4,5)P_2$:TF-$PI(4,5)P_2$ (95:4.5:0.5 molar ratio) or POPC:Bodipy-PC:$PI(4,5)P_2$ (94.5:0.5:5 molar ratio). It is worth emphasizing that the mol% total of $PI(4,5)P_2$ (labelled and unlabelled) is the same for both lipid compositions (5 mol%). Experiments were always carried out at concentrations for which ~ 100% membrane association is observed for the peptide (250 μM total lipid concentration) even in the presence of Ca^{2+}, so that any change in FRET efficiencies can be ascribed to changes in the lateral organization of the peptide and phospholipid analogues. The two FRET pairs here used, present similar spectral overlaps, allowing for comparison between both sets of data (Figure 3.8 (A and B)).

Looking at the results for the TopFluor-$PI(4,5)P_2$/PBP-10 (donor/acceptor) FRET pair (Figure 5, C) we observed a clear decrease in FRET efficiency in response to 100μM Ca^{2+}, allowing us to conclude that Ca^{2+} is in fact leading to a decrease in interaction between PBP-10 and $PI(4,5)P_2$. This data is in agreement with the partition data, and the electrostatic charge shielding of $PI(4,5)P_2$ by Ca^{2+}. Interestingly, when analysing the results for the Bodipy-PC/PBP-10 FRET pair (Figure 5, D) we observe a slight opposite effect in response to the presence of Ca^{2+}. In the presence of 100 μM Ca^{2+}, FRET efficiency increases very slightly which could indicate that it leads to an increase in the interaction between the peptide and PC. In this way, Ca^{2+} levels are shown to not only influence the membrane affinity of PBP-10 but also modulate the number of $PI(4,5)P_2$ molecules sequestered by PBP-10 within the membrane. In vivo, the highly negatively charged headgroup from $PI(4,5)P_2$ acts like a "beacon" for cationic membrane binding proteins. Electrostatic charge shielding of $PI(4,5)P_2$ by divalent cations can neutralize this negative charge, influencing $PI(4,5)P_2$-Protein interactions and consequently the localization of cationic membrane binding proteins with lower specifity. The formation of Ca^{2+} induced clusters could then contribute to this effect by excluding small cationic peptides from within the nanodomains, decreasing the interactions between the peptides and $PI(4,5)P_2$, and releasing them back to the bulk membrane. This regulation by Ca^{2+} could be fundamental to the regulation of some signalling pathways associated with $PI(4,5)P_2$.

MD simulations show PBP-10 exclusion from within Ca^{2+} induced $PI(4,5)P_2$ nanodomains. To better understand how the Ca^{2+} induced $PI(4,5)P_2$ nanodomains were affecting the interactions of PBP-10 with $PI(4,5)P_2$, we performed coarse grained molecular dynamics simulations using the Martini2.2 model. These simulations were carried out using the same protein to lipid ratio as the FRET experiments, to better replicate the experimental setup. A 5:1 Ca^{2+} to $PI(4,5)P_2$ ratio was used, which is also close to the Ca^{2+} concentrations used experimentally. Here, PBP-10 was added to membrane systems that were previously prepared, equilibrated and ran with or without Ca^{2+} for at least 10 μs. After adding the peptides to both sides of the membrane, the protein-lipid system ran for at least 20 μs. After running for at least 10 μs, 2 new simulations were started from snapshots obtained from the previous systems. In Ca^{2+} containing systems, the ion was replaced by water beads, whilst for the system without Ca^{2+}, it was added by replacing random water beads. Ionic force was then adjusted accordingly to account for the changes in system charge. This was done to check for hysteresis. No hysteresis effects were found, when comparing the 4 simulations, and thus the results shown are from the first two systems.

Analysing the PBP-10 - $PI(4,5)P_2$ dynamics (Figure 6.) we see that, in the absence of Ca^{2+}, PBP-10 appears to induce some aggregation of $PI(4,5)P_2$ (Figure 6. D and F). This occurs mainly via interaction between its positively charged residues and the phosphodiester bond phosphate group. These $PI(4,5)P_2$ aggregates, however, were not as ordered (Supmat. Figure 1. A and B), stable or as large (Supmat. Figure 1. B (red) and C (blue)) as Ca^{2+} induced $PI(4,5)P_2$ nanodomains. In the presence of Ca^{2+}, PBP-10 did not seem to disrupt the already existing $PI(4,5)P_2$ nanodomains very substantially (Supmat. Figure 1. B and C (red lines)), however some disruption was observed. The fact that PBP-10 did not significantly disrupt $PI(4,5)P_2$ clusters, means $PI(4,5)P_2$ interactions with nonspecific peptides can be modulated via interaction with clusters. PBP-10 did, however, significantly deplete the number of Ca^{2+} ions from the PO4 region (Figure 6. G). The ions at the headgroup phosphates were less impacted (Figure 6. H and I). PBP-10 is located at the phosphodiester depth, and its positively charged residues replace the stabilizing action of the Ca^{2+} ions.

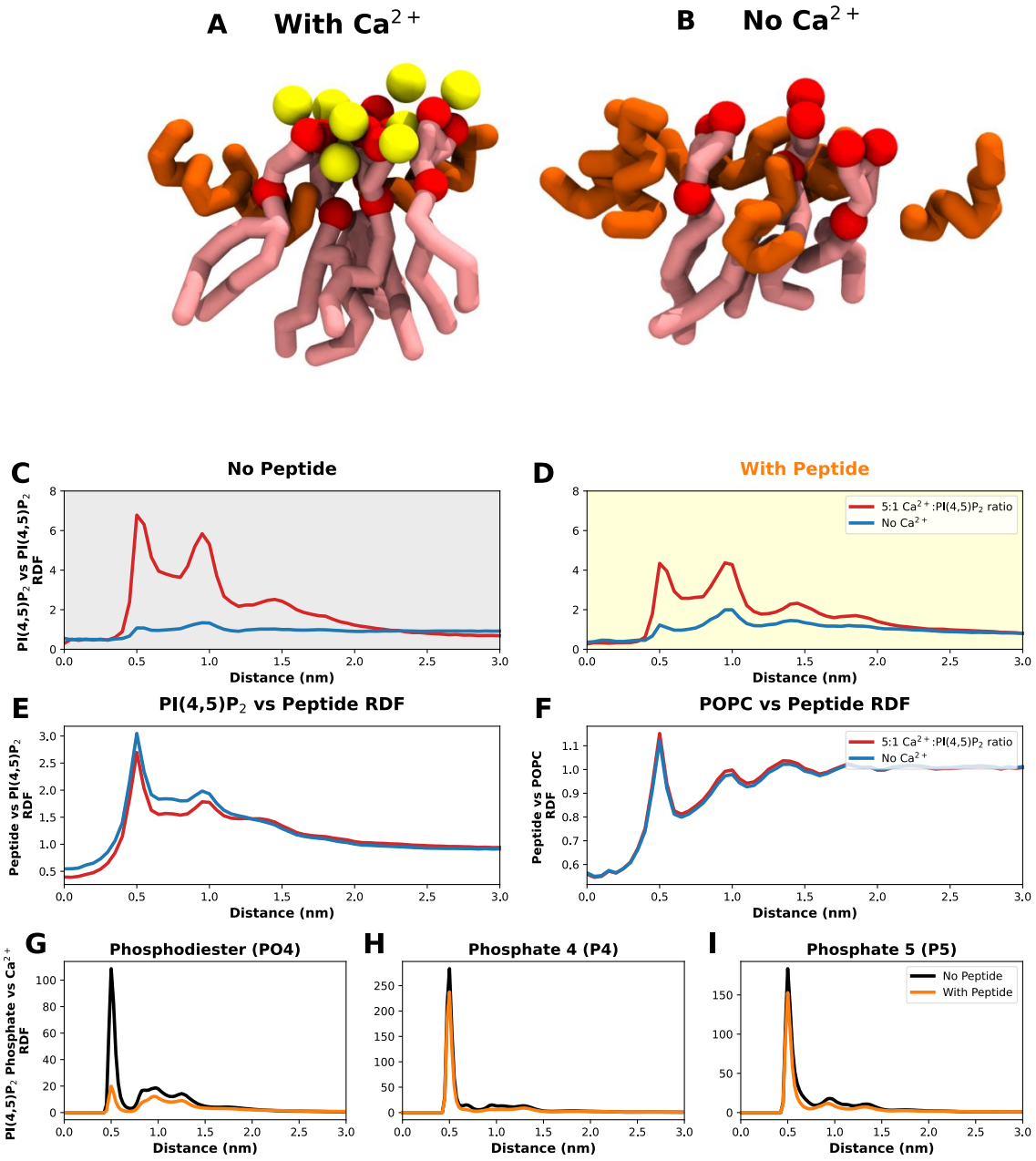


Figure 6. MD simulations show PBP-10 exclusion from within Ca^{2+} induced $PI(4,5)P_2$ nanodomains. Snapshots of PBP-10 interactions with $PI(4,5)P_2$ calcium induced clusters (A) and $PI(4,5)P_2$ disperse lipids (B). The PBP-10 peptides are represented in orange. The $PI(4,5)P_2$ lipids are represented in pink, with the 3 phosphates discriminated in red. Calcium beads are represented in yellow. $PI(4,5)P_2$ vs $PI(4,5)P_2$ RDF (PO4 vs PO4) analysis both in the absence (C) and presence of peptide (D) and either in the presence (red) or absence (blue) of Ca^{2+}. $PI(4,5)P_2$ vs Peptide (E) and POPC vs peptide RDF analysis (F) either in the presence (red) or absence (blue) of Ca^{2+}. $PI(4,5)P_2$ vs Ca^{2+} RDF analysis discriminated by phosphate bead (Phosphodiester: G, P4: H and P5: I) either in the presence (orange) or absence (black) of peptide.

It is likely that the positively charged residues are not as strong as Ca^{2+} in crosslinking the $PI(4,5)P_2$ phosphodiester phosphates, explaining the slight depletion observed in the size of the nanodomains. The formation of $PI(4,5)P_2$ nanodomains also appears to lead to a decrease in PBP-10 interactions with $PI(4,5)P_2$ (Figure 6., E), with a slight concomitant increase in interactions with POPC (Figure 6., F). These results, here seen through the RDFs, are in agreement with the FRET results shown previously. This once again appears to showcase that the formation of $PI(4,5)P_2$ nanodomains does not allow the incorporation of PBP-10 within them, restricting the peptide to interacting with the borders of the domain. This decreases interactions with $PI(4,5)P_2$ and releases the peptide to the bulk membrane, thus increasing interactions with POPC when compared to the system without Ca^{2+}.

Ca^{2+}-induced $PI(4,5)P_2$ effects influence PH-YFP organization and protein-protein interactions. After observing that Ca^{2+} induced $PI(4,5)P_2$ effects can influence PH-YFP and PBP-10 dynamics through complex competing mechanisms, we were interested in understanding whether these effects, and especially whether Ca^{2+} induced $PI(4,5)P_2$ clustering, could also influence protein-protein interactions, in this case specifically, PH-YFP membrane organization and oligomerization. PH-YFP is a fusion protein that consists of an isolated PLCδ1 PH domain linked to the fluorescent protein variant, YFP. The YFP variant used, as mentioned previously, is prone to dimerization and thus this system is an excellent model for membrane protein oligomerization.

Initially we aimed to determine whether or not PH-YFP was able to remove bound $PI(4,5)P_2$ from within the clusters and diffuse freely in the membrane through FCS studies. We started by studying the free diffusion of PH-YFP in solution (Supmat. Figure 2), in the presence and absence of Ca^{2+}. Both sets of autocorrelation data were properly fit to a single component 3D diffusion model and did not require a second component to be introduced to the analysis, suggesting a single population of diffusing species. We report that Ca^{2+} does not alter the diffusion of the protein in solution, which presented a diffusion coefficient in the absence and presence of Ca^{2+} of D = 70.09 ± 4.72 $\mu m^2 s^{-1}$ and D = 68.87 ± 1.74 $\mu m^2 s^{-1}$, respectively, confirming that Ca^{2+} does not impact the PH-YFP oligomerization state before membrane binding.

We continued by studying PH-YFP diffusion in GUVs in response to the same conditions (Figure 7., A and B). Every set of autocorrelation data was now properly fit to a single component 2D diffusion model, not requiring a second component be introduced to the analysis, suggesting a single population of diffusing species (membrane-bound protein). Representative curves are presented in figure 7. (A). Analysing the diffusion data (Figure 7. (B)), one can conclude that PH-YFP membrane diffusion appears to be dominated by $PI(4,5)P_2$ diffusion at this protein to lipid ratio, as the values observed (D = 6.12 ± 0.32 $\mu m^2 s^{-1}$ (D + SEM) for the samples in the absence of Ca^{2+}) are very similar to those reported by FCS for fluorescent analogues of $PI(4,5)P_2$ in POPC GUVs (D = 8.01 ± 0.13 $\mu m^2 s^{-1}$ (D + SEM) for the samples in the absence of Ca^{2+})[32]. This suggests that the diffusion of the protein-lipid complex is dominated by the viscosity experienced by the $PI(4,5)P_2$ inserted in the membrane, rather than the water exposed protein, as has been reported for other PH domains[79]. Looking at the effect of Ca^{2+}, the diffusion coefficient of PH-YFP decreased to some extent, going from D = 6.12 ± 0.32 $\mu m^2 s^{-1}$ (D + SEM) in the absence of Ca^{2+} to D = 4.27 ± 0.65 $\mu m^2 s^{-1}$ at 100μM Ca^{2+}. This effect was observed even at 20μM Ca^{2+} (D = 4.00 ± 0.42 $\mu m^2 s^{-1}$). This 30% decrease in PH-YFP diffusion in response to an increase in Ca^{2+} concentration is similar to the one previously reported, for a $PI(4,5)P_2$ fluorescent analogue in GUVs with Ca^{2+} [32]. Altogether, these results clearly confirm that PH-YFP domains after interaction with Ca^{2+}-induced $PI(4,5)P_2$ clusters, are not able to sequester the phospholipid from these $PI(4,5)P_2$ enriched nanodomains and remain associated with these structures, with restricted lateral diffusion. This helps explain the increased affinity of PH domains for $PI(4,5)P_2$ in the presence of Ca^{2+} and suggest that Ca^{2+} could potentially influence PH-YFP oligomerization.

Thus, we used PCH to follow PH-YFP oligomerization in solution and when bound to GUVs. If oligomerization occurs, at least 2 populations of fluorescent particles should be detected with distinct brightnesses. In the case of proteins, one population will correspond to the monomeric form of the protein and the other population, with a higher molecular brightness, will correspond to the oligomerized form. Analysing the PCH for PH-YFP in solution, both sets of data (in the presence and absence of Ca^{2+}) were properly fitted by a single brightness population model (Supmat. Figure 3, A). Furthermore, the brightness values confirm that Ca^{2+} does not have a direct effect on PH-YFP oligomerization in solution, as there was no significant difference in the brightness recovered for the single population, whether in the absence (ε = 0.068 ± 0.002 CPSM (Brightness ± SEM)) or presence of Ca^{2+} (ε = 0.070 ± 0.001 CPSM).

However, the photon counting histograms obtained for PH-YFP bound to the membrane of GUVs could no longer be properly fitted by a single population model and were instead fitted with a two-brightness population model. Representative photon counting histograms of PH-YFP bound to GUVs are shown in Figure 7., (C). Analysing the brightness data recovered (Supmat. Figure 3, B), one of the populations presented a brightness value with negligible differences between the samples in the absence and presence of the several Ca^{2+} concentrations (ε_1). This value was consistent with the value recovered for the protein in solution (Supmat. Figure 3, A) and therefore considered to represent the brightness of the monomeric membrane-bound PH-YFP population (5mM EDTA: ε_1 = 0.065 ± 0.003 CPSM; 20μM Ca^{2+}: ε_1 = 0.071 ± 0.003 CPSM; 50μM Ca^{2+}: ε_1 = 0.075 ± 0.010 CPSM; 100μM Ca^{2+}: ε_1 = 0.088 ± 0.003 CPSM). Importantly, in the presence of Ca^{2+}, this population was fully dominant, with the second population corresponding to less than 5 % of particles and < 12 % of total fluorescence in average. This result confirms that the presence of Ca^{2+} inhibits PH-YFP oligomerization in the membrane. Upon total chelation of Ca^{2+} by EDTA, the second population, which we considered to be the oligomeric membrane-bound PH-YFP population, presented a higher brightness value (5mM EDTA: ε2 = 0.172 ± 0.014 CPSM), consistent with a dimer (assuming no changes in quantum yield), as it is 2.3 ± 0.3 times higher than the

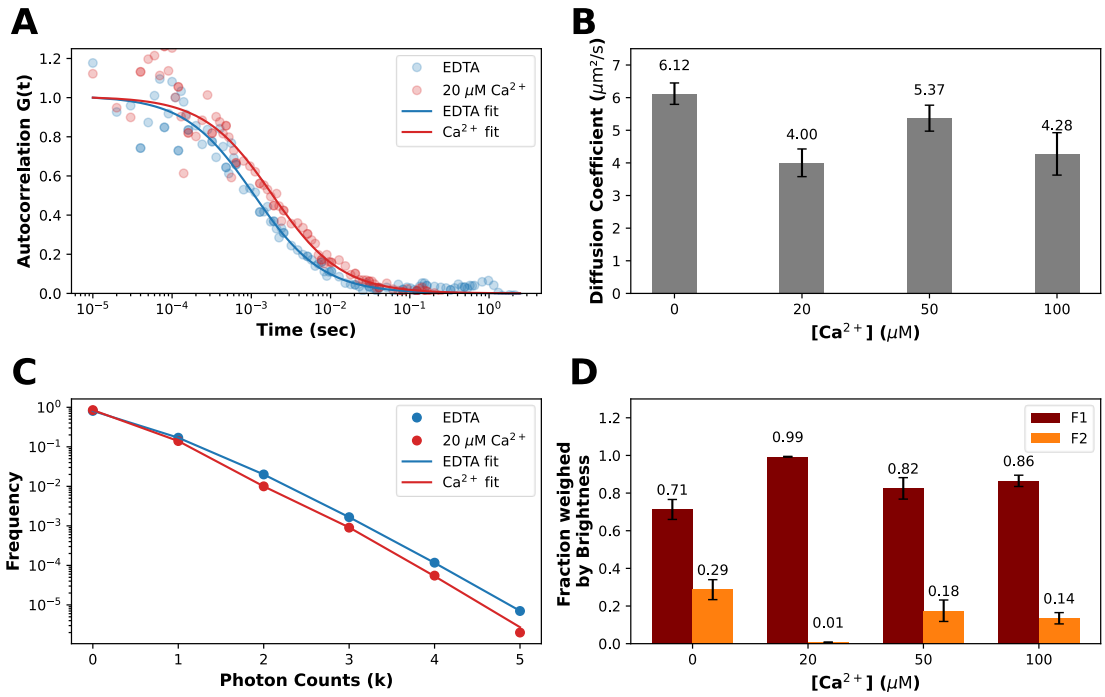


Figure 7. FFS studies of the influence of calcium on PH-YFP organization and protein-protein interactions. (A) Representative FCS autocorrelation curves of PH-YFP in POPC:$PI(4,5)P_2$ (95:5 molar ratio) GUVs, in the presence (red) and absence of Ca^{2+} (blue). (B) Diffusion coefficients obtained for PH-YFP in POPC:$PI(4,5)P_2$ (95:5 molar ratio) GUVs, in the presence and absence of Ca^{2+}. (C) Representative photon counting histograms of PH-YFP in POPC:$PI(4,5)P_2$ (95:5 molar ratio) GUVs in the presence (red) and absence (blue) of Ca^{2+}. Homogenous distribution of residuals shows appropriate fitting of the data. (data not shown) (D) Fraction weighed by brightness recovered for the two different populations detected for PH-YFP in POPC:$PI(4,5)P_2$ (95:5 molar ratio) GUVs, in the presence and absence of Ca^{2+}.

average of the values recovered for the lower brightness component. Formation of the PH-YFP dimer in the membrane was expected given the relatively low dissociation constant of YFP [63]. The fraction of dimers observed in the absence of Ca^{2+} corresponded to 28.7 ± 5.3% of protein (Fig 7D). In the presence of Ca^{2+}, as already discussed, the contribution of particles brighter than the monomer fell considerably, and the recovered brightness for the second component presented much higher uncertainty (20 μM Ca^{2+}: ε2 = 0.407 ± 0.083 CPSM; 50 μM Ca^{2+}: ε2 = 0.303 ± 0.066 CPSM; 100 μM Ca^{2+}: ε2 = 0.408 ± 0.111 CPSM). These values are also higher than the brightness value obtained for the dimer in the presence of EDTA, possibly reflecting the presence of a very small fraction of highly fluorescent, stable higher order oligomers of PH-YFP, which are not affected by the presence of Ca^{2+}, unlike dimers.

The fraction weighed by brightness can be used to quantify the total contribution of each population towards the total fluorescence collected. Here, we will use it to more accurately quantify each population present on the surface of the GUVs (Figure 7., (D)). In the presence of Ca^{2+}, the contribution of the oligomerized population decreases to about half at 50 μM Ca^{2+} and 100 μM Ca^{2+} (F_2 = 17.5 ± 5.7% and F_2 = 13.5 ± 3.0%, respectively) and almost entirely disappears at 20 μM Ca^{2+} (F_2 = 0.7 ± 0.2%).

One possible model to explain this variation in oligomerized population, is based on the hypothesis that it could be correlated with an increase in $PI(4,5)P_2$ cluster size caused by increasing concentrations of Ca^{2+}. According to this model, in the absence of Ca^{2+}-induced clustering of $PI(4,5)P_2$, each PH domain after binding to the membrane is free to diffuse laterally within the membrane, and in this two-dimensional environment where protein concentration is dramatically higher than in solution, protein-protein interactions are favoured. On the other hand, in case of formation of relatively small $PI(4,5)P_2$ clusters (20 μM Ca^{2+}), the membrane is no longer presenting a homogeneous environment for protein diffusion, since $PI(4,5)P_2$ is being segregated from the bulk lipids. In this case, proteins after binding to $PI(4,5)P_2$ clusters smaller than the necessary to accommodate PH-YFP dimers (well under 30 lipid molecules), are not totally free to diffuse laterally and become trapped or compartmentalized within these nanodomains. Additionally, if $PI(4,5)P_2$ clusters exhibit electrostatic repulsion due to incomplete chelation, interaction between individual $PI(4,5)P_2$ clusters could be rare, further reducing the probability of protein-protein interactions. These effects are likely to inhibit the protein-protein interactions seen in the absence of Ca^{2+}. With increasing Ca^{2+} concentrations, however, cluster sizes increase and could start to allow for the accommodation of a dimer within a single nanodomain. This results in a slight recovery of the oligomerized population (50 μM and 100 μM Ca^{2+}). This proposed mechanism, offers a very interesting insight on how Ca^{2+} levels, through the formation of Ca^{2+}-induced $PI(4,5)P_2$ clusters, could

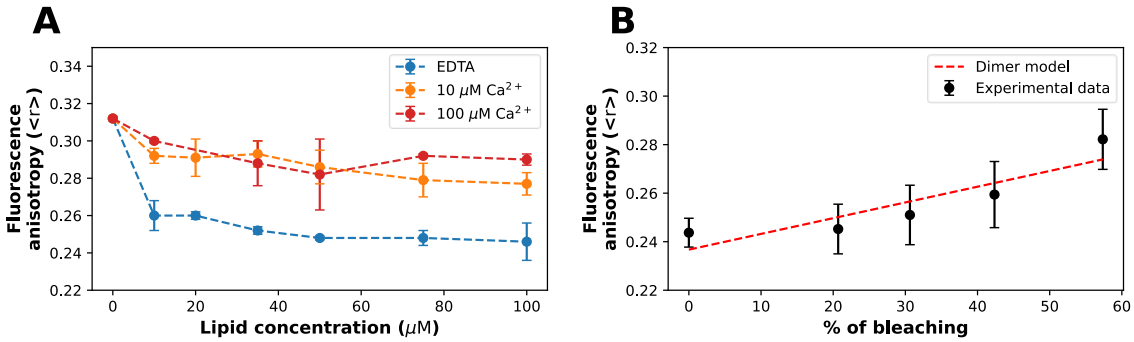


Figure 8. PH-YFP partitions to $PI(4,5)P_2$ containing LUVs and undergoes significant oligomerization in the absence of Ca^{2+}. (A) PH-YFP partition to POPC:$PI(4,5)P_2$ (95:5 molar ratio) LUVs, followed by the changes in steady-state fluorescence anisotropy, in the presence of 10 μM Ca^{2+} (orange), 100 μM Ca^{2+} (red) and in the presence of 5 mM EDTA (blue). Dashed lines are just a guide to the eye. (B) Photobleaching assay of 100 μM PH-YFP incubated with POPC:$PI(4,5)P_2$ (95:5 molar ratio) LUVs, in the presence of 5 mM EDTA.

possibly regulate $PI(4,5)P_2$ binding protein oligomerization and act as a regulatory step in signalling pathways.

To gather further evidence and confirm the FFS results, PH-YFP partition to POPC:$PI(4,5)P_2$ (95:5 molar ratio) LUVs was analysed through steady-state fluorescence depolarization in the presence of different Ca^{2+} concentrations. (Figure 8. A). As total lipid concentrations increase, we observe a general decrease of YFP fluorescence anisotropy. This decrease in anisotropy in response to an increase in lipid concentration, reflects PH-YFP dimerization in the membrane. PH-YFP dimers undergo Förster resonance energy transfer (homoFRET), increasing fluorescence depolarization and decreasing fluorescence anisotropies. However, analysing the different conditions we can clearly observe that, in the absence of Ca^{2+}, PH-YFP presents a significantly lower anisotropy when compared to the results from the samples in 10 μM and 100 μM Ca^{2+}. Since PH-YFP affinity is higher in the presence of these concentrations of Ca^{2+}, the decreased fluorescence depolarization obtained with EDTA can only be justified by a reduced oligomerization of the protein in the absence of $Ca2^{+}$-induced clustering. The results suggest that the formation of Ca^{2+}-induced $PI(4,5)P_2$ clusters could disrupt YFP oligomerization in the membrane to some extent. Furthermore, when 5 mM EDTA was added to PH-YFP previously incubated with 100 μM LUVs containing POPC:$PI(4,5)P_2$ (95:5 molar ratio) in 100 μM Ca^{2+} buffer (Supmat. Figure 4), we observed that fluorescence anisotropies decreased from the values observed in the presence of Ca^{2+} to, roughly, those observed in the absence of the cation. This result hints that the inhibition of PH-YFP oligomerization through the formation of Ca^{2+}-induced $PI(4,5)P_2$ clusters is reversible. In case this is a general property of $PI(4,5)P_2$ clusters, Ca^{2+}-induced clustering could act as a toggle switch to modulate interactions between $PI(4,5)P_2$-binding proteins in the plasma membrane.

To confirm that these variations in fluorescence anisotropy are due to PH-YFP oligomerization, as is suggested by the PCH results, we performed a photobleaching assay (Figure 8, B). In this assay, samples were exposed to intense light for varying amounts of time to obtain different photobleaching percentages. Whilst, photobleaching of PH-YFP in solution didn't cause any significant change in fluorescence anisotropy, photobleaching of PH-YFP samples incubated with 100 μM of $PI(4,5)P_2$ containing LUVs, in the absence of Ca^{2+}, led to the progressive recovery of fluorescence anisotropy values, reaching monomer anisotropy values at high photobleaching rates. This confirms that the decrease in fluorescence anisotropy observed for the samples in the absence of Ca^{2+}, occurs due to homoFRET induced by PH-YFP oligomerization. The rate at which the fluorescence anisotropy recovered with photobleaching percentage could be satisfactorily fit with a model that predicts the formation of dimers. However, this does not exclude the possibility of formation of higher order oligomers, as the quenching data could also be fit with a model that predicts the formation of trimers (although with a slightly lower accuracy). In fact, as the PCH brightness results also showed, PH-YFP is most likely found in an equilibrium with a minor fraction of high order oligomers.

DISCUSSION

The interaction of divalent cations with $PI(4,5)P_2$, especially Ca^{2+}, has been the focus of several studies. Divalent cations have been found to interact strongly with $PI(4,5)P_2$ and modulate its biophysical properties by influencing headgroup electrostatics via charge shielding[36], headgroup conformation[37], and $PI(4,5)P_2$ lateral organization by promoting the formation of clusters[32,80]. In all, this implies that divalent cations and especially Ca^{2+}, a common signal transduction element with a very buffered low physiological concentration, can have a strong role in the regulation of $PI(4,5)P_2$. These findings have led us to study not just how Ca^{2+} directly influences $PI(4,5)P_2$ but how it can also influence the interactions with $PI(4,5)P_2$ binding proteins. This is crucial, as it is through binding proteins that $PI(4,5)P_2$ carries out several of its important regulatory effects.

In an attempt to study how Ca^{2+} influences $PI(4,5)P_2$ binding proteins, we studied two different proteins that have the ability to bind to $PI(4,5)P_2$ through different mechanisms. We started by studying PH-YFP, a fusion protein consisting of a YFP domain with the ability to dimerize and a PH domain. The PH domain binds to $PI(4,5)P_2$ with high affinity and specificity via the stereochemical recognition of the phosphorylated inositol headgroup. Further on, we studied PBP-10, a 10 residue long polycationic peptide, which

interacts with $PI(4,5)P_2$ mainly through non-specific electrostatic interactions.

We started by determining how Ca^{2+}-induced $PI(4,5)P_2$ clustering could influence PH-YFP membrane affinity. Curiously, we observed a biphasic response to Ca^{2+} concentrations. At Ca^{2+} concentrations under 100 μM, we detected an increase in PH-YFP affinity towards the membrane. However, over 100 μM Ca^{2+}, membrane affinity seems to decrease once again to values slightly lower than those observed in the absence of Ca^{2+}. These results were consistent across several techniques and also observed in molecular dynamics simulations. This biphasic response is most likely the result of a complex mixture of competing effects. At low Ca^{2+} concentrations (under 100 μM), $PI(4,5)P_2$ headgroup conformation changes, promoted by the formation of $PI(4,5)P_2$ clusters, might facilitate its interaction with the PH domain. Simultaneously, the aggregation of $PI(4,5)P_2$ at low Ca^{2+} coordination ($PI(4,5)P_2$>>Ca^{2+}) might also provide a larger electrostatic charge gradient or induce headgroup conformation changes which could in turn increase PH-YFP binding affinity. On the other hand, for higher Ca^{2+} concentrations (over 100 μM Ca^{2+}), the increased charge screening of $PI(4,5)P_2$ by coordination with Ca^{2+} ($PI(4,5)P_2$<<Ca^{2+}) could reach a threshold where electrostatic interactions between $PI(4,5)P_2$ and PH domains are effectively cancelled out. Another effect which could take place at higher Ca^{2+} concentrations is the increase in number of sequestered $PI(4,5)P_2$ lipids by the PH domains, which would lead to an exponential decrease in lipids available for protein binding. The simultaneous contribution of all these processes is likely to modulate the response of $PI(4,5)P_2$ binding proteins to Ca^{2+}, and these effects will influence different proteins in different ways. Each stereospecific $PI(4,5)P_2$ binding protein will have its own most favourable headgroup conformation, and the extent by which Ca^{2+}-induced charge shielding of $PI(4,5)P_2$ affects protein binding affinity will depend on the electrostatic properties of each protein.

Ca^{2+}-induced $PI(4,5)P_2$ clustering had also a significant impact on PBP-10 partition to the membrane, where 100 μM Ca^{2+} led to a two-fold decrease in partition coefficient. These results contrast with the effect Ca^{2+} had on PH-YFP affinity. While at 100 μM, Ca^{2+} induced an increase in PH-YFP affinity, in the case of PBP-10 it leads to the opposite effect. Since PBP-10 is a small peptide that interacts with the membrane mostly via electrostatics and presents no specific $PI(4,5)P_2$ binding site, it stands to reason that it could be more affected by Ca^{2+}-induced charge shielding of the negatively charged headgroup. By the same logic, any influence on headgroup conformation should not contribute too much towards influencing PBP-10 interactions. Furthermore, we directly analysed how Ca^{2+} influenced PBP-10 interactions with the two individual membrane components, in the presence and absence of Ca^{2+} induced $PI(4,5)P_2$ clustering. We observed a decrease in interaction between PBP-10 and $PI(4,5)P_2$, in the presence of Ca^{2+}, which is in agreement with the electrostatic charge shielding of $PI(4,5)P_2$ by Ca^{2+}. On the other hand, the experiments also showed a slight increase in interaction with POPC, and thus these results can only be rationalized as reporting a decrease in the number of $PI(4,5)P_2$ lipids sequestered by the peptide. In the presence of Ca^{2+}, PBP-10 is now more in contact with the bulk lipids. The highly negatively charged inositol headgroup acts like a membrane beacon for cationic membrane binding proteins. Neutralizing this negative charge through an increase in Ca^{2+}, is expected to not only reduce the affinity of cationic membrane binding proteins for the plasma membrane but to also limit the extent of $PI(4,5)P_2$ interactions with the same proteins. This much was clear from our results, and can significantly influence protein localization and activity, which, in turn, will have regulatory effects on several signalling pathways.

Besides influencing membrane binding of proteins, it makes sense that Ca^{2+} could also influence membrane protein dynamics. We confirmed through FCS that PH-YFP membrane diffusion rates are slower in the presence of Ca^{2+}, asserting that PH-YFP remains associated to $PI(4,5)P_2$ clusters and is unable to sequester a lipid from these structures. This is also in agreement with what is seen in our MD simulations. This information is a strong indicator that cation induced clusters might, indeed, be able to influence protein dynamics.

Since PH-YFP undergoes oligomerization in the membrane, we could also quantify how cation induced $PI(4,5)P_2$ clustering influenced protein-protein interactions for this specific protein model. We observed that Ca^{2+}, in general, disrupted PH-YFP oligomerization in a reversible manner. Furthermore, this disruption was seen to be dependent on the concentration of Ca^{2+}, which could be the result of increasing $PI(4,5)P_2$ cluster sizes. In fact, an increase in cluster size is expected to increase the probability of incorporating multiple PH-YFP within the same $PI(4,5)P_2$ nanodomain. In case these changes in $PI(4,5)P_2$ organization occur within plasma membranes in response to Ca^{2+} stimulus, it is possible that a multitude of different protein-protein interactions are dramatically impacted, as proteins are included or excluded from interaction by an on – off Ca^{2+} switch. Such a regulatory mechanism could act as an important step in signalling pathways, especially in the vicinity of Ca^{2+} channels, where more steep fluctuations of local Ca^{2+} concentrations occur before buffering of the divalent cation by Ca^{2+} binding proteins.

An important fact, which is not directly addressed in this article, is the influence of Mg^{2+} on $PI(4,5)P_2$-protein interactions. While in vivo, intracellular Ca^{2+} is maintained at very low levels (around 100 nM), with transient spikes in concentration (up to hundreds of μM), Mg^{2+} levels are kept at much higher and stable concentrations. Free intracellular Mg^{2+} levels are typically around the 0.25-1 mM range. Although Mg^{2+} has a much weaker affinity for $PI(4,5)P_2$ when compared to Ca^{2+}, at mM concentrations it is also able to induce comparable $PI(4,5)P_2$ clustering as observed in the Ca^{2+} range of concentrations studied here[81]. This could mean that $PI(4,5)P_2$ is constitutively present in clusters and crosslinked with divalent cations. One also needs to consider, that Ca^{2+} and Mg^{2+} will have a combined influence over $PI(4,5)P_2$, which could buffer the effects caused by the fluctuation of Ca^{2+} levels. As the effect of Mg^{2+} has been often neglected, here included, it would be of interest for further studies to also emphasize its contribution.

Given that our results unequivocally associate the presence of divalent cations with considerable changes in $PI(4,5)P_2$ – protein affinity and $PI(4,5)P_2$-binding protein oligomerization, critical re-evaluations of $PI(4,5)P_2$ - protein dynamics must be carried out for $PI(4,5)P_2$ – binding proteins that take these effects into account.

CONCLUSION

In summary, our work clearly shows that Ca^{2+} within its physiological range of intracellular concentrations is able to directly modulate $PI(4,5)P_2$-protein interactions, through a complex combination of mechanisms. We found that after interaction of PLCδ1PH-YFP domains (here used as models of high affinity $PI(4,5)P_2$-binding proteins) with $PI(4,5)P_2$ presenting membranes in the presence of Ca^{2+}, the diffusion of the protein in the membrane is considerably lower than the diffusion rates observed in the absence of Ca^{2+}. That suggests that the protein interacts with Ca^{2+}-induced $PI(4,5)P_2$ clusters and is not able to fully sequester the bound phospholipid from these structures. This is of crucial relevance, since it proves that Ca^{2+}-induced $PI(4,5)P_2$ clusters, which only recently have been characterized in physiological concentrations of both $PI(4,5)P_2$ and Ca^{2+} [32], have the potential to compartmentalize $PI(4,5)P_2$ proteins, inhibiting, or possibly promoting interactions between them through changes in $PI(4,5)P_2$ cluster size. Ca^{2+} also led to significant charge screening of $PI(4,5)P_2$, altering local membrane electrostatics, and influencing protein binding and affinity. These cation-induced effects modulated different protein models in different ways, where the binding mechanism appears to play an important role. All of these factors, most likely, act in coordination to regulate the dynamics of many $PI(4,5)P_2$ binding proteins, crucial to various signalling pathways. In this way, the results of this study shed light on a new and largely uncharacterized mechanism, through which fluctuations in Ca^{2+} levels can have an impact on the membrane affinity and organization of $PI(4,5)P_2$ binding proteins.

ASSOCIATED CONTENT

Extended Methods and Materials providing additional analysis CG MD Analysis, FCS diffusion coefficients, Molecular brightness analysis and fluorescence anisotropy response to EDTA. Other data are available from the corresponding author upon request.

AUTHOR INFORMATION

Corresponding Author

Fábio Fernandes — iBB-Institute for Bioengineering and Biosciences, Instituto Superior Técnico, Universidade de Lisboa, Lisbon, Portugal; orcid.org/ 0000-0001-9112-265X; fernandesf@tecnico.ulisboa.pt

Author Contributions

All authors designed the experimental systems and L.B.A., M.E.M and M.J.S. performed them. L.B.A., F.F. and M.N.M. designed the simulation systems and L.B.A. performed them. L.B.A. performed all analysis. All authors contributed to manuscript writing and revision and have given approval to the final version of the manuscript.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENT

L.B.A. thanks the Medical Biochemistry and Biophysics Doctoral Program (M2B-PhD) and Fundação para a Ciência e a Tecnologia – Ministério da Ciência, Tecnologia e Ensino Superior (FCT-MCTES, Portugal) for PhD fellowship PD/BD/137492/2018. M.N.M. thanks FCT-MCTES for the "scientific employment stimulus" program CEECIND/04124/2017. M.N.M. also acknowledges FCT for funding project MOSTMICRO-ITQB, with references UIDB/04612/2020 and UIDP/04612/2020.

REFERENCES

(1) Whited, A. M.; Johs, A. The Interactions of Peripheral Membrane Proteins with Biological Membranes. *Chem. Phys. Lipids* **2015**, *192*, 51–59.
(2) Kwiatkowska, K. One Lipid, Multiple Functions: How Various Pools of PI(4,5)P2 Are Created in the Plasma Membrane. *Cell. Mol. Life Sci.* **2010**, *67* (23), 3927–3946.
(3) Ferrell, J. E.; Huests, W. H. Phosphoinositide Metabolism and the Morphology of Human Erythrocytes. *J. Cell Biol.* **1984**, *98* (6), 1992–1998.
(4) Balla, T. Phosphoinositides: Tiny Lipids with Giant Impact on Cell Regulation. *Physiol. Rev.* **2013**, *93* (3), 1019–1137.
(5) McLaughlin, S.; Wang, J.; Gambhir, A.; Murray, D. PIP2 and Proteins: Interactions, Organization, and Information Flow. *Annu. Rev. Biophys. Biomol. Struct.* **2002**, *31*, 151–175.
(6) Golebiewska, U.; Nyako, M.; Woturski, W.; Zaitseva, I.; McLaughlin, S. Diffusion Coefficient of Fluorescent Phosphatidylinositol 4,5-Bisphosphate in the Plasma Membrane of Cells. *Mol. Biol. Cell* **2008**, *19* (4), 1663–1669.
(7) Monteiro, M. E.; Sarmento, M. J.; Fernandes, F. Role of Calcium in Membrane Interactions by PI(4,5)P2-Binding Proteins. *Biochem. Soc. Trans.* **2014**, *42* (5), 1441–1446.
(8) Di Paolo, G.; De Camilli, P. Phosphoinositides in Cell Regulation and Membrane Dynamics. *Nature* **2006**, *443* (7112), 651–657.
(9) Höning, S.; Ricotta, D.; Krauss, M.; Späte, K.; Spolaore, B.; Motley, A.; Robinson, M.; Robinson, C.; Haucke, V.; Owen, D. J. Phosphatidylinositol-(4,5)-Bisphosphate Regulates Sorting Signal Recognition by the Clathrin-Associated Adaptor Complex AP2. *Mol. Cell* **2005**, *18* (5), 519–531.
(10) Wenk, M. R.; De Camilli, P. Protein-Lipid Interactions and Phosphoinositide Metabolism in Membrane Traffic: Insights from Vesicle Recycling in Nerve Terminals. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101* (22), 8262–8269.
(11) Simonsen, A.; Wurmser, A. E.; Emr, S. D.; Stenmark, H. The Role of Phosphoinositides in Membrane Transport. *Curr. Opin. Cell Biol.* **2001**, *13* (4), 485–492.
(12) Holz, R. W.; Hlubek, M. D.; Sorensen, S. D.; Fisher, S. K.; Balla, T.; Ozaki, S.; Prestwich, G. D.; Stuenkel, E. L.; Bittner, M. A. A Pleckstrin Homology Domain Specific for Phosphatidylinositol 4,5- Bisphosphate (PtdIns-4,5-P2) and Fused to Green Fluorescent Protein Identifies Plasma Membrane PtdIns-4,5-P2 as Being Important in Exocytosis. *J. Biol. Chem.* **2000**, *275* (23), 17878–17885.
(13) Martin, T. F. J. PI(4,5)P2 Regulation of Surface Membrane Traffic. *Curr. Opin. Cell Biol.* **2001**, *13* (4), 493–499.
(14) Hilgemann, D. W.; Feng, S.; Nasuhoglu, C. The Complex and Intriguing Lives of PIP2 with Ion Channels and Transporters. *Sci. STKE* **2001**, *2001* (111), re19.
(15) Yarar, D.; Surka, M. C.; Leonard, M. C.; Schmid, S. L. SNX9 Activities Are Regulated by Multiple Phosphoinositides through Both PX and BAR Domains. *Traffic* **2008**, *9* (1), 133–146.
(16) Song, X.; Xu, W.; Zhang, A.; Huang, G.; Liang, X.; Virbasius, J. V.; Czech, M. P.; Wayne Zhou, G. Phox Homology Domains Specifically Bind Phosphatidylinositol Phosphates. *Biochemistry* **2001**, *40* (30), 8940–8944.
(17) Chiang, S. H.; Hwang, J.; Legendre, M.; Zhang, M.; Kimura, A.; Saltiel, A. R. TCGAP, a Multidomain Rho GTpase-Activating Protein Involved in Insulin-Stimulated Glucose Transport. *EMBO J.* **2003**, *22* (11), 2679–2691.

(18) Bai, J.; Tucker, W. C.; Chapman, E. R. PIP2 Increases the Speed of Response of Synaptotagmin and Steers Its Membrane-Penetration Activity toward the Plasma Membrane. *Nat. Struct. Mol. Biol.* **2004**, *11* (1), 36–44.

(19) Itoh, T.; Koshiba, S.; Kigawa, T.; Kikuchi, A.; Yokoyama, S.; Takenawa, T. Role of the ENTH Domain in Phosphatidylinositol-4,5-Bisphosphate Binding and Endocytosis. *Science (80-.).* **2001**, *291* (5506), 1047–1051.

(20) Santagata, S.; Boggon, T. J.; Baird, C. L.; Gomez, C. A.; Zhao, J.; Wei Song Shan; Myszka, D. G.; Shapiro, L. G-Protein Signaling through Tubby Proteins. *Science (80-.).* **2001**, *292* (5524), 2041–2050.

(21) Lemmon, M. A.; Ferguson, K. M.; O'Brien, R.; Sigler, P. B.; Schlessinger, J. Specific and High-Affinity Binding of Inositol Phosphates to an Isolated Pleckstrin Homology Domain. *Proc. Natl. Acad. Sci. U. S. A.* **1995**, *92* (23), 10472–10476.

(22) Achiriloaie, M.; Barylko, B.; Albanesi, J. P. Essential Role of the Dynamin Pleckstrin Homology Domain in Receptor-Mediated Endocytosis. *Mol. Cell. Biol.* **1999**, *19* (2), 1410–1415.

(23) Lemmon, M. A.; Ferguson, K. M. Signal-Dependent Membrane Targeting by Pleckstrin Homology (PH) Domains. *Biochem. J.* **2000**, *350* (1), 1–18.

(24) Garcia, P.; Gupta, R.; Shah, S.; Morris, A. J.; Rudge, S. A.; Scarlata, S.; Petrova, V.; McLaughlin, S.; Rebecchi, M. J. The Pleckstrin Homology Domain of Phospholipase C-Δ1 Binds with High Affinity to Phosphatidylinositol 4,5-Bisphosphate in Bilayer Membranes. *Biochemistry* **1995**, *34* (49), 16228–16234.

(25) Lemmon, M. A. Pleckstrin Homology (PH) Domains and Phosphoinositides. *Biochem. Soc. Symp.* **2007**, *74* (1), 81–93.

(26) Schmoldt, A.; Benthe, H. F.; Haberland, G. Digitoxin Metabolism by Rat Liver Microsomes. *Biochem. Pharmacol.* **1975**, *24* (17), 1639–1641.

(27) Ferguson, K. M.; Lemmon, M. A.; Schlessinger, J.; Sigler, P. B. Crystal Structure at 2.2 Å Resolution of the Pleckstrin Homology Domain from Human Dynamin. *Cell* **1994**, *79* (2), 199–209.

(28) Ferguson, K. M.; Lemmon, M. A.; Schlessinger, J.; Sigler, P. B. Structure of the High Affinity Complex of Inositol Trisphosphate with a Phospholipase C Pleckstrin Homology Domain. *Cell* **1995**, *83* (6), 1037–1046.

(29) van der Wal, J.; Habets, R.; Várnai, P.; Balla, T.; Jalink, K. Monitoring Agonist-Induced Phospholipase C Activation in Live Cells by Fluorescence Resonance Energy Transfer. *J. Biol. Chem.* **2001**, *276* (18), 15337–15344.

(30) Knight, J. D.; Falke, J. J. Single-Molecule Fluorescence Studies of a PH Domain: New Insights into the Membrane Docking Reaction. *Biophys. J.* **2009**, *96* (2), 566–582.

(31) Clapham, D. E. Calcium Signaling. *Cell* **2007**, *131* (6), 1047–1058.

(32) Sarmento, M. J.; Coutinho, A.; Fedorov, A.; Prieto, M.; Fernandes, F. Ca2+ Induces PI(4,5)P2 Clusters on Lipid Bilayers at Physiological PI(4,5)P2 and Ca2+ Concentrations. *Biochim. Biophys. Acta* **2014**, *1838* (3), 822–830.

(33) Wang, Yu-hsiu; Collins, Agnieszka; Guo, Lin; Smith-Dupont, Kathryn B.; Gai, Feng; Svitkina, Tatyana; Janmey, P. A. Divalent Cation-Induced Cluster Formation by Polyphosphoinositol in Model Membranes. *J Am Chem Soc.* **2012**, *134* (7), 3387–3395.

(34) Wang, Y. H.; Slochower, D. R.; Janmey, P. A. Counterion-Mediated Cluster Formation by Polyphosphoinositides. *Chem. Phys. Lipids* **2014**, *182* (January), 38–51.

(35) Ellenbroek, W. G.; Wang, Y. H.; Christian, D. A.; Discher, D. E.; Janmey, P. A.; Liu, A. J. Divalent Cation-Dependent Formation of Electrostatic PIP2 Clusters in Lipid Monolayers. *Biophys. J.* **2011**, *101* (9), 2178–2184.

(36) Seo, J. B.; Jung, S. R.; Huang, W.; Zhang, Q.; Koh, D. S. Charge Shielding of PIP2 by Cations Regulates Enzyme Activity of Phospholipase C. *PLoS One* **2015**, *10* (12), 1–22.

(37) Bilkova, E.; Pleskot, R.; Rissanen, S.; Sun, S.; Czogalla, A.; Cwiklik, L.; Rog, T.; Vattulainen, I.; Cremer, P. S.; Jungwirth, P.; Coskun, Ü. Calcium Directly Regulates Phosphatidylinositol 4,5-Bisphosphate Headgroup Conformation and Recognition. *J. Am. Chem. Soc.* **2017**, jacs.6b11760.

(38) Cunningham, C. C.; Vegners, R.; Bucki, R.; Funaki, M.; Korde, N.; Hartwig, J. H.; Stossel, T. P.; Janmey, P. A. Cell Permeant Polyphosphoinositide-Binding Peptides That Block Cell Motility and Actin Assembly. *J. Biol. Chem.* **2001**, *276* (46), 43390–43399.

(39) Bucki, R.; Janmey, P. A.; Vegners, R.; Giraud, F.; Sulpice, J. C. Involvement of Phosphatidylinositol 4,5-Bisphosphate in Phosphatidylserine Exposure in Platelets: Use of a Permeant Phosphoinositide-Binding Peptide. *Biochemistry* **2001**, *40* (51), 15752–15761.

(40) Mayer, L. D.; Hope, M. J.; Cullis, P. R. Vesicles of Variable Sizes Produced by a Rapid Extrusion Procedure. *BBA - Biomembr.* **1986**, *858* (1), 161–168.

(41) Weinberger, A.; Tsai, F. C.; Koenderink, G. H.; Schmidt, T. F.; Itri, R.; Meier, W.; Schmatko, T.; Schröder, A.; Marques, C. Gel-Assisted Formation of Giant Unilamellar Vesicles. *Biophys. J.* **2013**, *105* (1), 154–164.

(42) Sarmento, M. J.; Prieto, M.; Fernandes, F. Reorganization of Lipid Domain Distribution in Giant Unilamellar Vesicles upon Immobilization with Different Membrane Tethers. *Biochim. Biophys. Acta - Biomembr.* **2012**, *1818* (11), 2605–2615.

(43) Lakowicz, J. R. *Principles of Fluorescence Spectroscopy*; Springer, 2006.

(44) Santos, N. C.; Prieto, M.; Castanho, M. A. R. B. Quantifying Molecular Partition into Model Systems of Biomembranes: An Emphasis on Optical Spectroscopic Methods. *Biochim. Biophys. Acta - Biomembr.* **2003**, *1612* (2), 123–135.

(45) Aragón, S. R.; Pecora, R. Fluorescence Correlation Spectroscopy as a Probe of Molecular Dynamics. *J. Chem. Phys.* **1976**, *64* (4), 1791–1803.

(46) Widengren, J.; Mets, Ü.; Rigler, R. Fluorescence Correlation Spectroscopy of Triplet States in Solution: A Theoretical and Experimental Study. *J. Phys. Chem.* **1995**, *99* (36), 13368–13379.

(47) Majer, G.; Melchior, J. P. Characterization of the Fluorescence Correlation Spectroscopy (FCS) Standard Rhodamine 6G and Calibration of Its Diffusion Coefficient in Aqueous Solutions. *J. Chem. Phys.* **2014**, *140* (9), 094201.

(48) Ries, J.; Schwille, P. New Concepts for Fluorescence Correlation Spectroscopy on Membranes. *Phys. Chem. Chem. Phys.* **2008**, *10* (24), 3487.

(49) Gendron, P. O.; Avaltroni, F.; Wilkinson, K. J. Diffusion Coefficients of Several Rhodamine Derivatives as Determined by Pulsed Field Gradient-Nuclear Magnetic Resonance and Fluorescence Correlation Spectroscopy. *J. Fluoresc.* **2008**, *18* (6), 1093–1101.

(50) Huang, B.; Perroud, T. D.; Zare, R. N. Photon Counting Histogram: One-Photon Excitation. *ChemPhysChem* **2004**, *5* (10), 1523–1531.

(51) Chen, Y.; Müller, J. D.; So, P. T. C.; Gratton, E. The Photon Counting Histogram in Fluorescence Fluctuation Spectroscopy. *Biophys. J.* **1999**, *77* (1), 553–567.

(52) Marrink, S. J.; Risselada, H. J.; Yefimov, S.; Tieleman, D. P.; De Vries, A. H. The MARTINI Force Field: Coarse Grained Model for Biomolecular Simulations. *J. Phys. Chem. B* **2007**, *111* (27), 7812–7824.

(53) López, C. A.; Sovova, Z.; van Eerden, F. J.; de Vries, A. H.; Marrink, S. J. Martini Force Field Parameters for Glycolipids. *J. Chem. Theory Comput.* **2013**, *9* (3), 1694–1708.

(54) Borges-Araújo, L.; Domingues, M. M.; Fedorov, A.; Santos, N. C.; Melo, M. N.; Fernandes, F. Acyl-Chain Saturation Regulates the Order of Phosphatidylinositol 4,5-Bisphosphate Nanodomains. *Commun. Chem.* **2021**, *4* (1), 1–13.

(55) Wassenaar, T. A.; Ingólfsson, H. I.; Böckmann, R. A.; Tieleman, D. P.; Marrink, S. J. Computational Lipidomics with Insane: A Versatile Tool for Generating Custom Membranes for Molecular Simulations. *J. Chem. Theory Comput.* **2015**, *11* (5), 2144–2155.

(56) Xian, W.; Vegners, R.; Janmey, P. A.; Braunlin, W. H. Spectroscopic Studies of a Phosphoinositide-Binding Peptide from Gelsolin: Behavior in Solutions of Mixed Solvent and Anionic Micelles. *Biophys. J.* **1995**, *69* (6), 2695–2702.

(57) Monticelli, L.; Kandasamy, S. K.; Periole, X.; Larson, R. G.; Tieleman, D. P.; Marrink, S. J. The MARTINI Coarse-Grained Force Field: Extension to Proteins. *J. Chem. Theory Comput.* **2008**, *4* (5), 819–834.

(58) De Jong, D. H.; Singh, G.; Bennett, W. F. D.; Arnarez, C.;

Wassenaar, T. A.; Schäfer, L. V.; Periole, X.; Tieleman, D. P.; Marrink, S. J. Improved Parameters for the Martini Coarse-Grained Protein Force Field. *J. Chem. Theory Comput.* **2013**, *9* (1), 687–697.

(59) Melo, M. N. Coarse-Grain Simulations of Membrane-Adsorbed Helical Peptides. In *Computational Peptide Science*; 2022.

(60) Grossfield, A. WHAM: The Weighted Histogram Analysis Method.

(61) Bonomi, M.; Bussi, G.; Camilloni, C.; Tribello, G. A.; Banáš, P.; Barducci, A.; Bernetti, M.; Bolhuis, P. G.; Bottaro, S.; Branduardi, D.; Capelli, R.; Carloni, P.; Ceriotti, M.; Cesari, A.; Chen, H.; Chen, W.; Colizzi, F.; De, S.; De La Pierre, M.; Donadio, D.; Drobot, V.; Ensing, B.; Ferguson, A. L.; Filizola, M.; Fraser, J. S.; Fu, H.; Gasparotto, P.; Gervasio, F. L.; Giberti, F.; Gil-Ley, A.; Giorgino, T.; Heller, G. T.; Hocky, G. M.; Iannuzzi, M.; Invernizzi, M.; Jelfs, K. E.; Jussupow, A.; Kirilin, E.; Laio, A.; Limongelli, V.; Lindorff-Larsen, K.; Löhr, T.; Marinelli, F.; Martin-Samos, L.; Masetti, M.; Meyer, R.; Michaelides, A.; Molteni, C.; Morishita, T.; Nava, M.; Paissoni, C.; Papaleo, E.; Parrinello, M.; Pfaendtner, J.; Piaggi, P.; Piccini, G. M.; Pietropaolo, A.; Pietrucci, F.; Pipolo, S.; Provasi, D.; Quigley, D.; Raiteri, P.; Raniolo, S.; Rydzewski, J.; Salvalaglio, M.; Sosso, G. C.; Spiwok, V.; Šponer, J.; Swenson, D. W. H.; Tiwary, P.; Valsson, O.; Vendruscolo, M.; Voth, G. A.; White, A. Promoting Transparency and Reproducibility in Enhanced Molecular Simulations. *Nature Methods*. Nature Publishing Group August 1, 2019, pp 670–673.

(62) Tribello, G. A.; Bonomi, M.; Branduardi, D.; Camilloni, C.; Bussi, G. PLUMED 2: New Feathers for an Old Bird. *Comput. Phys. Commun.* **2014**, *185* (2), 604–613.

(63) Shaner, N. C.; Steinbach, P. A.; Tsien, R. Y. A Guide to Choosing Fluorescent Proteins. *Nat. Methods* **2005**, *2* (12), 905–909.

(64) Rebecchi, M.; Peterson, A.; McLaughlin, S. Phosphoinositide-Specific Phospholipase C-Δ1 Binds with High Affinity to Phospholipid Vesicles Containing Phosphatidylinositol 4, 5-Bisphosphate. *Biochemistry* **1992**, *31* (51), 12742–12747.

(65) Ryu, S. H.; Suh, P. G.; Cho, K. S.; Lee, K. Y.; Rhee, S. G. Bovine Brain Cytosol Contains Three Immunologically Distinct Forms of Inositolphospholipid-Specific Phospholipase C. *Proc. Natl. Acad. Sci. U. S. A.* **1987**, *84* (October), 6649–6653.

(66) Ryu, S. H.; Cho, K. S.; Lee, K. Y.; Suh, P. G.; Rhee, S. G. Purification and Characterization of Two Immunologically Distinct Phosphoinositide-Specific Phospholipases C from Bovine Brain. *J. Biol. Chem.* **1987**, *262* (26), 12511–12518.

(67) Laux, T.; Fukami, K.; Thelen, M.; Golub, T.; Frey, D.; Caroni, P. GAP43, MARCKS, and CAP23 Modulate PI(4,5)P2 at Plasmalemmal Rafts, and Regulate Cell Cortex Actin Dynamics through a Common Mechanism. *J. Cell Biol.* **2000**, *149* (7), 1455–1471.

(68) Wang, J.; Gambhir, A.; Hangyás-Mihályné, G.; Murray, D.; Golebiewska, U.; McLaughlin, S. Lateral Sequestration of Phosphatidylinositol 4,5-Bisphosphate by the Basic Effector Domain of Myristoylated Alanine-Rich C Kinase Substrate Is Due to Nonspecific Electrostatic Interactions. *J. Biol. Chem.* **2002**, *277* (37), 34401–34412.

(69) Dietrich, U.; Krüger, P.; Gutberlet, T.; Käs, J. A. Interaction of the MARCKS Peptide with PIP2 in Phospholipid Monolayers. *Biochim. Biophys. Acta - Biomembr.* **2009**, *1788* (7), 1474–1481.

(70) Gambhir, A.; Hangyas-Mihalyne, G.; Zaitseva, I.; Cafiso, D. S.; Wang, J.; Murray, D.; Pentyala, S. N.; Smith, S. O.; McLaughlin, S. Electrostatic Sequestration of PIP2 on Phospholipid Membranes by Basic/Aromatic Regions of Proteins. *Biophys. J.* **2004**, *86* (4), 2188–2207.

(71) Tong, J.; Nguyen, L.; Vidal, A.; Simon, S. A.; Skene, J. H. P.; McIntosh, T. J. Role of GAP-43 in Sequestering Phosphatidylinositol 4,5-Bisphosphate to Raft Bilayers. *Biophys. J.* **2008**, *94* (1), 125–133.

(72) van den Bogaart, G.; Meyenberg, K.; Risselada, H. J.; Amin, H.; Willig, K. I.; Hubrich, B. E.; Dier, M.; Hell, S. W.; Grubmüller, H.; Diederichsen, U.; Jahn, R. Membrane Protein Sequestering by Ionic Protein–Lipid Interactions. *Nature* **2011**, *479* (7374), 552–555.

(73) Shi, J.; Birnbaumer, L.; Large, W. A.; Albert, A. P. Myristoylated Alanine-Rich C Kinase Substrate Coordinates Native TRPC1 Channel Activation by Phosphatidylinositol 4,5-Bisphosphate and Protein Kinase C in Vascular Smooth Muscle. *FASEB J.* **2014**, *28* (1), 244–255.

(74) Su, R.; Han, Z. Y.; Fan, J. P.; Zhang, Y. L. A Possible Role of Myristoylated Alanine-Rich C Kinase Substrate in Endocytic Pathway of Alzheimer's Disease. *Neurosci. Bull.* **2010**, *26* (4), 338–344.

(75) Trovò, L.; Ahmed, T.; Callaerts-Vegh, Z.; Buzzi, A.; Bagni, C.; Chuah, M.; VandenDriessche, T.; Balschun, D.; Dotti, C. G. Low Hippocampal PI(4,5)P2 Contributes to Reduced Cognition in Old Mice as a Result of Loss of MARCKS. *Nat. Publ. Gr.* **2013**, *16* (4), 4–4.

(76) Psachoulia, E.; Sansom, M. S. P. Interactions of the Pleckstrin Homology Domain with Phosphatidylinositol Phosphate and Membranes: Characterization via Molecular Dynamics Simulations. *Biochemistry* **2008**, *47* (14), 4211–4220.

(77) Alessandri, R.; Souza, P. C. T.; Thallmair, S.; Melo, M. N.; de Vries, A. H.; Marrink, S. J. Pitfalls of the Martini Model. *J. Chem. Theory Comput.* **2019**, *15* (10), 5448–5460.

(78) Levental, I.; Janmey, P. A.; Cēbers, a; Cebers, A. Electrostatic Contribution to the Surface Pressure of Charged Monolayers Containing Polyphosphoinositides. *Biophys.\ J.* **2008**, *95* (3), 1199–1205.

(79) Knight, J. D.; Lerner, M. G.; Marcano-Velázquez, J. G.; Pastor, R. W.; Falke, J. J. Single Molecule Diffusion of Membrane-Bound Proteins: Window into Lipid Contacts and Bilayer Dynamics. *Biophys. J.* **2010**, *99* (9), 2879–2887.

(80) Wang, Y.-H.; Collins, A.; Guo, L.; Smith-Dupont, K. B.; Gai, F.; Svitkina, T.; Janmey, P. A. Divalent Cation-Induced Cluster Formation by Polyphosphoinositides in Model Membranes. *J. Am. Chem. Soc.* **2012**, *134* (7), 3387–3395.

(81) Sarmento, M. J.; Coutinho, A.; Fedorov, A.; Prieto, M.; Fernandes, F. Membrane Order Is a Key Regulator of Divalent Cation-Induced Clustering of PI(3,5)P2 and PI(4,5)P2. *Langmuir* **2017**, *33* (43), 12463–12477.

3 | Acyl-chain saturation regulates the order of phosphatidylinositol 4,5-bisphosphate nanodomains

Article II: Acyl-chain saturation regulates the order of phosphatidylinositol 4,5-bisphosphate nanodomains

communications chemistry

ARTICLE

https://doi.org/10.1038/s42004-021-00603-1 OPEN

Acyl-chain saturation regulates the order of phosphatidylinositol 4,5-bisphosphate nanodomains

Luís Borges-Araújo[1,2], Marco M. Domingues[3], Alexander Fedorov[1], Nuno C. Santos[3], Manuel N. Melo[2] & Fábio Fernandes[1,4✉]

Phosphatidylinositol 4,5-bisphosphate ($PI(4,5)P_2$) plays a critical role in the regulation of various plasma membrane processes and signaling pathways in eukaryotes. A significant amount of cellular resources are spent on maintaining the dominant 1-stearoyl-2-arachidonyl $PI(4,5)P_2$ acyl-chain composition, while less abundant and more saturated species become more prevalent in response to specific stimuli, stress or aging. Here, we report the impact of acyl-chain structure on the biophysical properties of cation-induced $PI(4,5)P_2$ nanodomains. $PI(4,5)P_2$ species with increasing levels of acyl-chain saturation cluster in progressively more ordered nanodomains, culminating in the formation of gel-like nanodomains for fully saturated species. The formation of these gel-like domains was largely abrogated in the presence of 1-stearoyl-2-arachidonyl $PI(4,5)P_2$. This is, to the best of our knowledge, the first report of the impact of $PI(4,5)P_2$ acyl-chain composition on cation-dependent nanodomain ordering, and provides important clues to the motives behind the enrichment of $PI(4,5)P_2$ with polyunsaturated acyl-chains. We also show how Ca^{2+}-induced $PI(4,5)P_2$ nanodomains are able to generate local negative curvature, a phenomenon likely to play a role in membrane remodeling events.

[1] Institute for Bioengineering and Biosciences (IBB) and Associate Laboratory i4HB-Institute for Health and Bioeconomy, Instituto Superior Técnico, Universidade de Lisboa, Lisbon, Portugal. [2] Instituto de Tecnologia Química e Biológica António Xavier, Universidade Nova de Lisboa, Av. da República, 2780-157 Oeiras, Portugal. [3] Instituto de Medicina Molecular, Faculdade de Medicina, Universidade de Lisboa, Av. Prof. Egas Moniz, 1649-028 Lisbon, Portugal. [4] Department of Bioengineering, Instituto Superior Técnico, Universidade de Lisboa, 1049-001 Lisbon, Portugal. ✉email: fernandesf@tecnico.ulisboa.pt

Phosphoinositides are an important class of glycerophospholipids, which play a variety of diverse and specific roles across several eukaryotic membrane structures. Phosphatidylinositol 4,5-bisphosphate ($PI(4,5)P_2$) is one of the most abundant phosphoinositides, comprising ~1% of the total membrane phospholipids in mammalian cells[1]. It is mostly localized in the inner leaflet of the plasma membrane, where it plays a crucial role in multiple pathways, especially those related with membrane dynamics[2]. $PI(4,5)P_2$ has been associated with several membrane processes, such as vesicle trafficking[3,4], cytoskeletal regulations[5], ion channel function[6], viral assembly[7], and budding[8].

It is remarkable that a single lipid class present at low and constant steady-state levels can act as an important regulator in so many different, yet simultaneous, signaling pathways[9]. Lateral organization of $PI(4,5)P_2$ is critical to that end. It is thought that $PI(4,5)P_2$ is heterogeneously dispersed across the membrane, in localized $PI(4,5)P_2$-enriched clusters at particular sites and timings[10]. These clusters are promoted not only through localized depletion and synthesis of $PI(4,5)P_2$ but also by the formation of $PI(4,5)P_2$ nanodomains, promoted by the interaction of its negatively charged phosphorylated headgroup with divalent cations (such as Ca^{2+} and Mg^{2+}) or positively charged proteins[11,12]. It has been shown that divalent cations have the ability to induce the dramatic segregation of $PI(4,5)P_2$ to form highly enriched nanodomains, even at physiological concentrations of both lipid and cation[10,13]. Interactions between the anionic headgroup phosphates and divalent cations not only screen the electrostatic repulsion between headgroups but can also crosslink adjacent lipids. These $PI(4,5)P_2$ clusters have been shown to have higher affinity for more ordered lipid phases than the monomeric species[14].

The higher affinity of clustered $PI(4,5)P_2$ towards ordered membrane domains is puzzling, given its bulky headgroup and acyl-chain unsaturation profile, which are expected to strongly favor incorporation into more disordered membrane phases. Although most phospholipids show considerable acyl-chain composition diversity, $PI(4,5)P_2$ molecules are highly enriched in specific acyl chains[15]. In fact, the most frequent fatty acyl-chain pair for $PI(4,5)P_2$ in mammalian cells is 1-stearoyl-2-arachidonyl (18:0 20:4)[16]. This combination consists of up to 70% of the total $PI(4,5)P_2$ pool in some cases, particularly in brain tissue[17–19]. This enrichment is likely the combined outcome of specific substrate specificity for 1-stearoyl-2-arachidonyl-glycerol in several enzymes in the phosphatidylinositol cycle and of phosphoinositide acyl-chain remodeling via the Land's cycle[17].

The biological functions that call for this specific enrichment are still not clear. It has been shown that arachidonate (20:4) and other polyunsaturated fatty acids such as docosahexaenoate (22:6), when at the *sn*-2 position, facilitate membrane shaping and fission activities[20]. In addition, asymmetric *sn*-1-saturated-*sn*-2-polyunsaturated phospholipids have been shown to promote efficient membrane vesiculation, while maintaining low membrane permeability[20]. $PI(4,5)P_2$ has been associated with several stages of both endocytosis and exocytosis, being considered an important mediator of synaptic vesicle trafficking, where this composition might provide mechanical benefits[3]. Despite the dominant presence of the canonical 18:0 20:4 species, $PI(4,5)P_2$ acyl-chain composition is not fully uniform and includes a larger range of acyl-chain lengths and unsaturation profiles[19]. $PI(4,5)P_2$ molecular species exhibiting no polyunsaturation or even fully saturated acyl chains are less abundant, but have been reported to become more prevalent in response to certain stimuli[21], stress[19,21], aging[19], or in cancer[22]. In one of these cases, the levels of $PI(4,5)P_2$ with this profile of acyl-chain composition even surpassed the canonical 18:0 20:4 composition[21].

However, there is still hardly any research done on the impact of these more saturated species in $PI(4,5)P_2$ organization. More saturated acyl chains are expected to lead to a more ordered and less fluid membrane landscape. $PI(4,5)P_2$ species exhibiting lower unsaturation or even full saturation are likely to present significantly different lateral distribution that that of the canonical 18:0 20:4 form. Considering the importance of $PI(4,5)P_2$ for sorting of membrane proteins, such difference is likely to have profound influence in many $PI(4,5)P_2$-dependent cellular processes. In addition, interactions of proteins with specific $PI(4,5)P_2$ acyl chains have already been reported. HIV-1 Gag polyprotein was found to sequester unsaturated acyl chains from $PI(4,5)P_2$ and store them in a hydrophobic pocket during the membrane-anchoring process[23]. Several enzymes, such as phosphatidylinositol-4-phosphate-5-kinase, also show preference for some acyl-chain configurations when using the lipid as either a substrate or as an activator[17]. Further insight on how the acyl-chain profile can influence $PI(4,5)P_2$ biophysical properties may help in shedding some light on how it can affect downstream $PI(4,5)P_2$-dependent processes.

In this study, we investigated three different $PI(4,5)P_2$ acyl-chain configurations, representative of the broad spectrum observed in vivo, in order to better understand the extent of the impact of acyl-chain saturation on $PI(4,5)P_2$ biophysical properties and on cation-induced $PI(4,5)P_2$ nanodomains. Through a combination of atomic force microscopy (AFM), coarse-grained (CG) molecular dynamics (MD) simulations and fluorescence spectroscopy techniques, we show that different $PI(4,5)P_2$ acyl-chain compositions lead to the formation of Ca^{2+}-induced nanodomains with distinct biophysical properties. Specifically, we observed that as saturation increases, $PI(4,5)P_2$ forms more ordered nanodomains, which can ultimately culminate in the formation of gel-like nanodomains with the fully saturated acyl-chain composition. These results provide an important biophysical insight on the motives behind the conserved pattern of enrichment of $PI(4,5)P_2$ with polyunsaturated acyl chains.

$PI(4,5)P_2$ is known to directly regulate several membrane-remodeling events[3,4,24,25]. Although $PI(4,5)P_2$ is essential for membrane fusion processes, such as soluble NSF attachment receptor (SNARE)-dependent membrane fusion, it has an intrinsic positive curvature, which is expected to restrain the formation of the negatively curved intermediates necessary for some of these mechanisms[3,4,24,25]. Using CG MD simulations, we also show here how $PI(4,5)P_2$ nanodomains associate with negatively curved membranes. Generation of negative curvature by clustered $PI(4,5)P_2$ is likely to play a major rule in membrane-remodeling events.

Results

Acyl-chain composition does not influence the formation of $PI(4,5)P_2$ nanodomains. In order to study the influence of acyl-chain composition on the properties of $PI(4,5)P_2$ nanodomains, three compositions were chosen as representative of the broad spectrum seen in vivo. (18:0 20:4) $PI(4,5)P_2$ was an obvious choice, as it is the dominant species detected in mammalian cells. To represent the less abundant, more saturated species, $(18:1)_2$ and $(16:0)_2$ $PI(4,5)P_2$ were chosen.

To characterize the lateral distribution of $PI(4,5)P_2$ in lipid membranes, both in the presence and in the absence of divalent cations, we employed the fluorescent analog TF-$PI(4,5)P_2$. This analog has already been successfully used in the detection of $PI(4,5)P_2$ nanodomains in large unilamellar vesicles (LUVs)[10,13]. TF-$PI(4,5)P_2$ is capable of undergoing homo Förster resonance energy transfer (FRET) and, as a result, fluorescence depolarization takes place upon enrichment of the analog within nanoclusters in the membrane. Thus, nanoclustering of PI(4,5)

COMMUNICATIONS CHEMISTRY | https://doi.org/10.1038/s42004-021-00603-1 ARTICLE

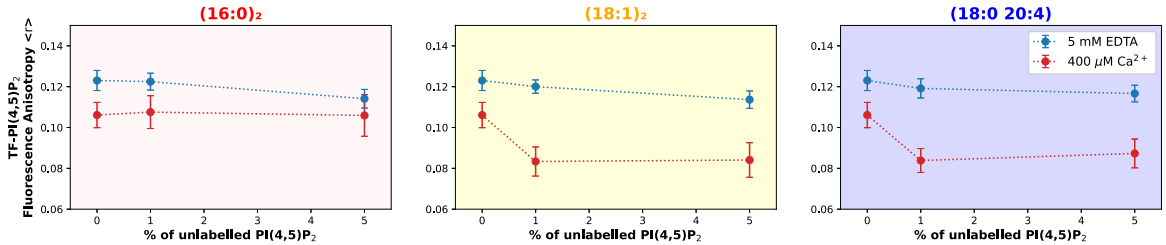


Fig. 1 Cation-induced clusters are formed independently of PI(4,5)P_2 acyl-chain composition as seen by homo-FRET of the TF-PI(4,5)P_2 analog. PI(4,5)P_2 cluster formation was determined through the incorporation of 0.1 mol% of TF-PI(4,5)P_2 in LUVs containing POPC and increasing concentrations of unlabeled PI(4,5)P_2. The experiments were carried out for the three unlabeled PI(4,5)P_2 species, with different acyl-chain composition. TF-PI(4,5)P_2 fluorescence anisotropy (<*r*>) values were measured in the presence (400 μM Ca^{2+}, red) and absence (5 mM EDTA, blue) of calcium. Error bars represent the SD from $N = 3$ independent experiments. A significant decrease in TF-PI(4,5)P_2 fluorescence anisotropy in the presence of calcium is observed for the three PI(4,5)P_2 species ((16:0)2 PI(4,5)P_2: $F(1,12) = 17.87$, $p = 0.0012$; $(18:1)_2$ PI(4,5)P_2: $F(1,12) = 96.84$; $p < 0.0001$; (18:0 20:4) PI(4,5)P_2: $F(1,12) = 107.7$, $p < 0.0001$).

P_2 can be followed through a decrease in fluorescence anisotropy (<*r*>) of TF-PI(4,5)P_2[10]. During Ca^{2+}-induced clustering, the presence of higher concentrations of unlabeled PI(4,5)P_2 contributes to the formation of larger clusters, promoting TF-PI(4,5)P_2 sequestration, which is detectable through a decrease of <*r*>[10].

TF-PI(4,5)P_2 was incorporated in 1-palmitoyl-2-oleoyl-glycero-3-phosphocholine (POPC):PI(4,5)P_2 LUVs at an analog to lipid ratio of 0.1 mol%. At such low concentrations, only residual FRET takes place in the absence of clustering[10]. As expected, we observed for all compositions that, in the absence of free calcium, the recovered fluorescence anisotropy of the analog, <*r*>, was 0.125 (Fig. 1), a value consistent with monodispersed TF-PI(4,5)P_2 with no aggregation[10].

A concentration of 400 μM Ca^{2+} was used for Ca^{2+}-induced PI(4,5)P_2 clustering experiments. This concentration is well within the expected levels of the divalent cation on local Ca^{2+} nanodomains after opening of channels[26]. Importantly, in the presence of 400 μM Ca^{2+}, the fluorescence anisotropy decreased for all compositions, which is indicative of higher FRET and sequestration of the analog into PI(4,5)P_2 clusters, confirming that aggregation occurs independently of acyl-chain composition. Increasing concentrations of unlabeled PI(4,5)P_2 were used to characterize the concentration dependence of clustering. For the $(18:1)_2$ and (18:0 20:4) PI(4,5)P_2 species in the presence of calcium, a similar decrease in anisotropy with increasing concentrations of unlabeled PI(4,5)P_2 is observed ($F(1,12) = 0.1445$, $p = 0.7105$). These results suggest that the extent of PI(4,5)P_2 clustering obtained for $(18:1)_2$ and (18:0 20:4) is comparable. Surprisingly, the decrease in <*r*> for $(16:0)_2$ PI(4,5)P_2 in the presence of calcium is independent from the concentration of unlabeled PI(4,5)P_2 ($F(2,12) = 1.023$, $p = 0.3887$).

Fluorescence lifetime ($\bar{\tau}$) of the analog is also decreased upon insertion into PI(4,5)P_2 clusters[10]. In the absence of calcium, $\bar{\tau}$ is identical for all PI(4,5)P_2 species (Supplemantary Fig. 1), reflecting the comparable local environment provided by the bulk membrane, composed mostly of POPC. When 400 μM Ca^{2+} was included, the fluorescence lifetime of the analog decreased similarly for all PI(4,5)P_2 species (Supplementary Fig. 1). The data for fully saturated PI(4,5)P_2 were once again puzzling, as the lifetime of the analog greatly increased when higher concentrations of $(16:0)_2$ PI(4,5)P_2 were used. Incorporation of the analog into PI(4,5)P_2 nanodomains is confirmed by fluorescence correlation spectroscopy (FCS) measurements (Supplementary Fig. 2), which show a decrease in the TF-PI(4,5)P_2 diffusion coefficient across the three acyl-chain compositions consistent with the sequestration into Ca^{2+}-induced clusters. Previous results from our laboratory have already shown that although Ca^{2+} is able to induce a minor decrease in diffusion coefficient of membranes enriched in other anionic lipids, the impact on PI(4,5)P_2-containing membranes is considerably larger, reflecting the specificity of the process[10]. It is noteworthy that the diffusion coefficient measured here reflects the diffusion of the cluster and the diffusion rates inside clusters fall quite below the resolution of the confocal microscope and are not resolved by FCS.

The interpretation of these results is simple for the mono- and polyunsaturated compositions, and clearly indicates the sequestration of the analog within PI(4,5)P_2 nanodomains of similar properties. However, the photophysical behavior of the analog in the presence of Ca^{2+}-induced $(16:0)_2$ PI(4,5)P_2 clusters suggests that the properties of these nanodomains differ dramatically from those obtained with unsaturated species. Given the saturated acyl chains of $(16:0)_2$ PI(4,5)P_2, it is possible that clustering of this lipid gives rise to highly ordered domains, whereas unsaturated PI(4,5)P_2 species give rise to more fluid nanodomains. This would be consistent with the fluorescence anisotropy and lifetime of the analog within $(16:0)_2$ PI(4,5)P_2 domains. In fact, fluorescence anisotropy values of membrane probes in ordered membrane phases are known to increase due to limited rotational diffusion of the fluorophore, while the fluorescence lifetime is often increased[27].

The nature of the different Ca^{2+}-induced PI(4,5)P_2 clusters was also analyzed by AFM on supported lipid bilayers (SLBs) of 1,2-dioleoyl-sn-glycero-3-phosphocholine (DOPC):PI(4,5)P_2 95:5 (Fig. 2a). DOPC was used instead of POPC, as was done for all other experiments, as it can undergo lipid interdigitation and enhance contrast between the bulk membrane and other phases, in this case, PI(4,5)P_2 nanodomains. All three PI(4,5)P_2 species are confirmed to produce nanodomains. The properties of $(16:0)_2$ PI(4,5)P_2 clusters were once more distinguishable from the other PI(4,5)P_2 species: although the height of nanodomains formed from unsaturated PI(4,5)P_2 is found to be only slightly above that of the bulk DOPC bilayer, the height of $(16:0)_2$ PI(4,5)P_2 domains was, on average, 2 nm thicker than the rest of the membrane. This difference gives rise to a clear second peak in the histograms of bilayer height (Fig. 2b). This large increase in bilayer thickness suggests a dramatic reorganization of lipid structure for $(16:0)_2$ PI(4,5)P_2. AFM studies of SLBs composed of DOPC and 1,2-diplamitoyl-sn-glycero-3-phosphocholine (DPPC) have shown DPPC gel domains to be roughly 2 nm thicker than the bulk membrane[28], whereas Martini CG MD simulations of fluid DOPC:$(16:0)_2$ PI(4,5)P_2 systems have shown differences in lipid height of roughly 1 nm over a single lipid bilayer (Supplementary Fig. 9). In the context of $(16:0)_2$ PI(4,5)P_2 gel-like nanodomains, it is likely that this difference in height would be slightly larger, in the 2 nm range seen experimentally.

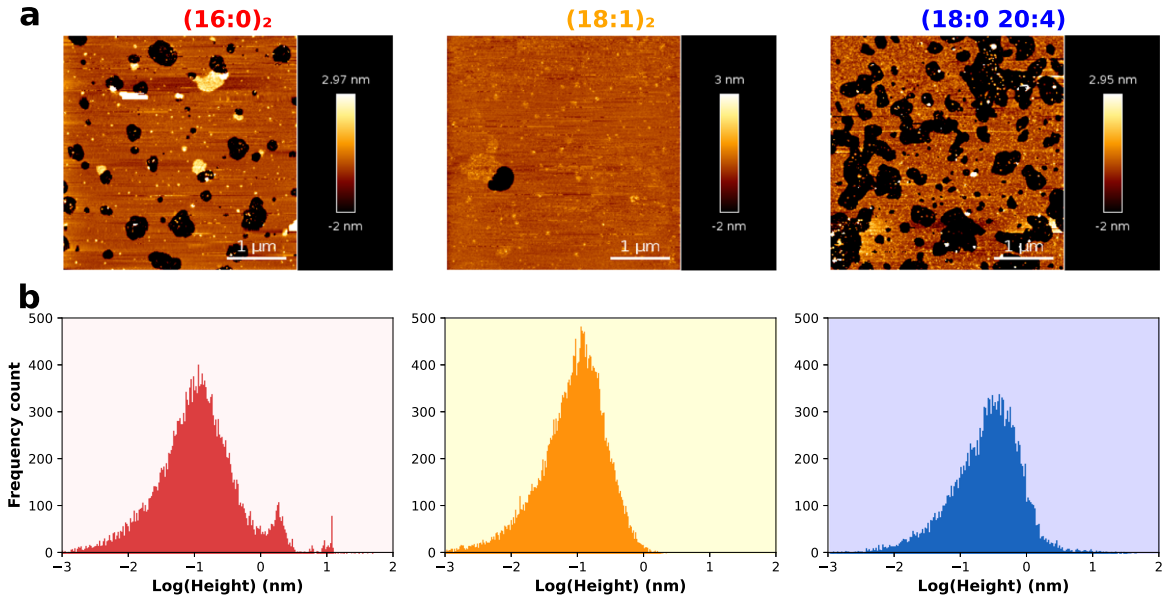


Fig. 2 Cation-induced clusters are formed independently of PI(4,5)P_2 acyl-chain composition, as seen by AFM measurements of supported lipid bilayers. PI(4,5)P_2 cluster formation was detected through the AFM measurement of SLBs containing DOPC and 5% PI(4,5)P_2. Experiments were carried out for the three PI(4,5)P_2 species in study and figures are labeled according to the acyl-chain composition of the PI(4,5)P_2 species employed. Topographical images were acquired (**a**) and analyzed with first- or second-level flattening, using the JPK data processing software, from which the membrane heigh profiles (**b**) were obtained. Dark patches correspond to defects in the supported membrane. The height values of each pixel (higher than 0 nm) were transformed into $\log_{10}$ value, avoiding the influence of the height values of defects on the histograms. The frequency count axis represents the number of data points grouped into each bin on the height histograms.

Although both $(16:0)_2$ PI(4,5)P_2 and $(18:1)_2$ PI(4,5)P_2 mostly gave rise to nanodomains of 30–50 nm diameter, the polyunsaturated PI(4,5)P_2 promoted the formation of smaller clusters, predominantly below these values (Supplementary Fig. 3). The nanodomains observed for all PI(4,5)P_2 species were often unstable in SLBs and substantial fractions were seen to cleave off from the mica surface during the washing steps where unfused vesicles are rinsed from the surface. This gave rise to the dark patches observed in Fig. 2a. Given the large headgroup of PI(4,5)P_2, it is likely that the curvature of these nanodomains differs considerably from the bulk DOPC, which can justify the observed instability.

Ca^{2+}-induced clustering leads to an increase in local membrane order, even for polyunsaturated acyl-chain compositions. The fluorescent analog and AFM results strongly indicate dramatic differences in nanodomain order for the different PI(4,5)P_2 species. To quantify these differences, we made use of 1-(4-Trimethylammoniumphenyl)-6-phenyl-1,3,5-hexatriene *p*-toluenesulfonate (TMA-DPH), a membrane probe commonly used to study membrane fluidity[29]. As in these experiments PI(4,5)P_2 only comprises a small percentage of membrane lipids, measuring the average membrane order is bound to underestimate any change observed within the phosphoinositide-rich domains. The cationic trimethylammonium moiety of TMA-DPH should be anchored near the phosphate and carbonyl/ester region of the membrane[29] to maximize favorable interactions with electronegative oxygen atoms. Electrostatic attraction is expected to cause TMA-DPH to prefer the vicinity of the anionic PI(4,5)P_2 molecules and their nanodomains, thus providing a strategy to probe the PI(4,5)P_2 local environment.

The fluorescence anisotropy of TMA-DPH was measured in multilamellar vesicles (MLVs) composed of POPC and increasing concentrations of the different species of PI(4,5)P_2 (Fig. 3a). Although in the absence of calcium, a minor increase in $\langle r \rangle_{TMA\text{-}DPH}$ is observed for higher concentrations of all PI(4,5)P_2 variants, the increase is more evident in the presence of the divalent cation. A similar behavior is observed for the fluorescence lifetime of the probe (Supplementary Fig. 4). Surprisingly, these results show that for all acyl-chain compositions analyzed here, an increase in local membrane order takes place upon Ca^{2+}-induced PI(4,5)P_2 nanodomain formation. The extent of the increase in membrane order correlates with the degree of saturation in the acyl chains, with much higher increases in $\langle r \rangle_{TMA\text{-}DPH}$ for the fully saturated species and more moderate changes observed for the polyunsaturated PI(4,5)P_2.

Once more, the behavior of $(16:0)_2$ PI(4,5)P_2 completely stands out from that of the other compositions, showing a much larger increase in $\langle r \rangle_{TMA\text{-}DPH}$ in the presence of calcium, whereas presenting a similar behavior in the absence of the divalent cation. As a comparison, a study on the impact of cholesterol on POPC membrane order found a similar increase in $\langle r \rangle_{TMA\text{-}DPH}$ only upon adding 40% of the sterol[30].

In order to evaluate the stability of these structures, we measured $\langle r \rangle_{TMA\text{-}DPH}$ in liposomes containing 5% PI(4,5)P_2 at increasing temperatures (Fig. 3b). In the absence of calcium, the behavior of all compositions was similar to that of pure POPC control membranes. The $\langle r \rangle_{TMA\text{-}DPH}$ decreases due to increased dynamics as the temperature increases. In the presence of calcium, $\langle r \rangle_{TMA\text{-}DPH}$ was noticeably higher throughout the range of tested temperatures and no clear inflection point is observed. These results confirm that, within this temperature range, there is no temperature dependence for the formation of PI(4,5)P_2 nanodomains. Remarkably, unlike the solid-like gel phase domain of saturated lipids such as $(16:0)_2$ PC (DPPC), which exhibit phase transition to fluid phase at a melting temperature of

COMMUNICATIONS CHEMISTRY | https://doi.org/10.1038/s42004-021-00603-1 ARTICLE

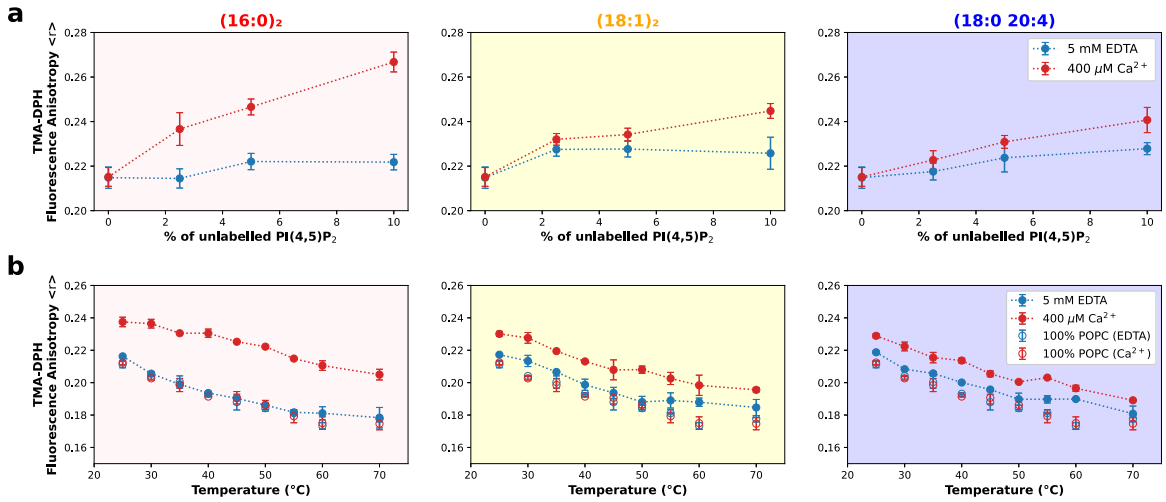


Fig. 3 TMA-DPH fluorescence spectroscopy shows that calcium-induced PI(4,5)P_2 nanodomains are more ordered than monodisperse PI(4,5)P_2, even for unsaturated acyl-chain compositions. PI(4,5)P_2 local membrane order was determined through the incorporation of TMA-DPH at a 1:300 lipid ratio in MLVs containing POPC and increasing concentrations of unlabeled PI(4,5)P_2. The experiments were done for the three acyl-chain compositions in study. For all acyl-chain compositions analyzed here, an increase in local membrane order takes place upon Ca^{2+}-induced PI(4,5)P2 nanodomain formation ($(16:0)_2$ PI(4,5)P_2: F(1,16) = 148.3, $p = 0.0001$; $(18:1)_2$ PI(4,5)P_2: F(1,16) = 19.71, $p = 0.0004$; (18:0 20:4) PI(4,5)P_2: F(1,16) = 12.23, $p < 0.0030$). TMA-DPH fluorescence anisotropy (<*r*>) (**a**) was measured in the presence (400 μM Ca^{2+}, red) and absence (5 mM EDTA, blue) of calcium. The TMA-DPH response to a temperature gradient (**b**) was measured in samples composed of POPC:PI(4,5)P_2 95:5. Pure POPC samples were measured as controls. All samples were measured both in the presence and absence of calcium. Error bars for all measurements represent the SD from $N = 3$ independent experiments.

41 °C[31], the highly ordered phase created by $(16:0)_2$ PI(4,5)P_2 is stable up to 70 °C.

Ca^{2+} induces the formation of $(16:0)_2$ PI(4,5)P_2 gel-like nanodomains. As saturated phospholipids with $(16:0)_2$ acyl chains, such as DPPC, are prone to the formation of solid-like gel phases at room temperature, we aimed to clarify whether that was the case for Ca^{2+}-induced $(16:0)_2$ PI(4,5)P_2 nanodomains, as this would explain both the dramatic increase in membrane order and the increase in membrane thickness detected by AFM. To that end, we performed experiments with the membrane probe *trans*-parinaric acid (tPnA). tPnA is the ideal membrane probe for detection of gel lipid domains, as it is one of a few known to exhibit preferential partition to these phases[32]. Large increases in both its anisotropy and fluorescence lifetime are typical hallmarks of gel-phase-like behavior. These parameters were measured in MLVs composed of POPC and increasing concentrations of the different species of PI(4,5)P_2 (Fig. 4a).

In the absence of calcium, tPnA anisotropy did not respond to increasing concentrations of PI(4,5)P_2, for any of the studied acyl-chain compositions. This was expected, as the previous experiments indicated that in the absence of calcium, PI(4,5)P_2 is homogenous and monodisperse. In the presence of calcium, we observed no effect for both $(18:1)_2$ PI(4,5)P_2 and (18:0 20:4) PI(4,5)P_2. Again, this was expected, as the previous experiments did not hint at these nanodomains having gel-like properties. However, for $(16:0)_2$ PI(4,5)P_2 in the presence of calcium, we observe a steep increase in both fluorescence anisotropy and fluorescence lifetime (Fig. 4a, b), confirming that Ca^{2+}-induced $(16:0)_2$ PI(4,5)P_2 nanodomains have a gel-phase-like behavior.

The formation of $(16:0)_2$ PI(4,5)P_2 gel-like phase in the presence of physiologically relevant concentrations of calcium is reminiscent of the previously observed formation of gel-like phosphatidylserine domains by Ca^{2+}-induced crosslinking of the anionic phosphoserine (PS) headgroups[33]. In the case of PS, gel formation only occurs at considerably higher calcium concentrations, which limits the biological relevance of this phenomena for that phospholipid. On the other hand, the higher affinity of PI(4,5)P_2 for the divalent cation promotes the occurrence of the phenomena within the range seen here. Nevertheless, the formation of a gel-like phase of PI(4,5)P_2 at low lipid concentrations is particularly surprising given the much larger headgroup, which poses greater steric hindrance in the organization of the gel crystal. In addition, the charge density is much higher for PI(4,5)P_2, leading to greater repulsion between the headgroups. The thermotropic behavior of $(16:0)_2$ PI(4,5)P_2 in the presence of calcium was also tested (Fig. 4c). Unlike the observation made with TMA-DPH, in this case, an inflexion point is clearly visible on the anisotropy data, reflecting a transition from gel-like to fluid phase. Interestingly, the transition temperature of $(16:0)_2$ PI(4,5)P_2 to the fluid phase (T_m) (obtained from the midpoint of the $<r>_{tPnA}$ transition in Fig. 4c) occurs at 45 °C, a temperature significantly higher than the one of the corresponding phosphatidylcholine, DPPC (41 °C).

The differences between the data obtained from tPnA and TMA-DPH reflect the absence of partition of TMA-DPH into gel phases. Hence, although tPnA data reflects gel domains specifically, TMA-DPH fails to incorporate into these and will only probe PI(4,5)P_2 nanodomain environments that failed to transition to the gel phase, possibly due to high local POPC content.

Interestingly, the presence of (18:0 20:4) PI(4,5)P_2 has a clear impact in the formation of $(16:0)_2$ PI(4,5)P_2 gel-like domains, as seen from Fig. 4d. When equal amounts of both PI(4,5)P_2 species were present, the fluorescence anisotropy value of tPnA dropped significantly, reflecting a decrease in the extent of gel-like phase formation. In addition, when the canonical polyunsaturated PI(4,5)P_2 is present at a higher concentration than $(16:0)_2$ PI(4,5)P_2, the formation of PI(4,5)P_2 gel-like nanodomains is completely abrogated. This result showcases the importance of (18:0 20:4)

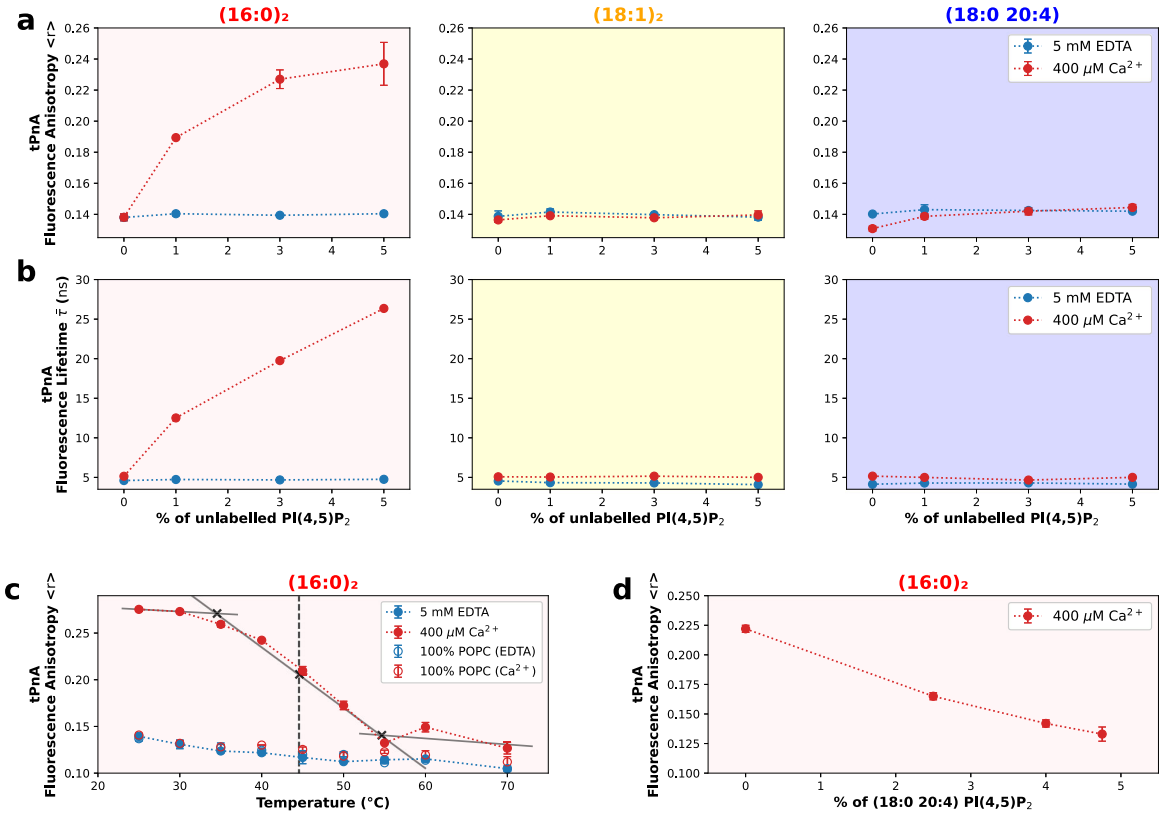


Fig. 4 $(16:0)_2$ PI(4,5)P_2 forms gel-like nanodomains upon undergoing calcium-induced clustering. PI(4,5)P_2 gel-like properties were detected through the incorporation of tPnA at a 1:300 lipid ratio in MLVs containing POPC and increasing concentrations of unlabeled PI(4,5)P_2. Experiments were done for the three acyl-chain compositions in study. tPnA fluorescence anisotropy (**a**) and fluorescence intensity weighed lifetime ($\bar{\tau}$) (**b**) were measured in the presence (400 μM Ca^{2+}, red) and absence (5 mM EDTA, blue) of calcium. In the presence of calcium, we observed no effect on tPnA fluorescence lifetime for both (18:1)2 PI(4,5)P_2 ($F(1,12) = 1.564$, $p = 0.2349$) and (18:0 20:4) PI(4,5)P_2 ($F(1,12) = 0.7426$, $p = 0.4057$). For (16:0)2 PI(4,5)P_2 in the presence of calcium, we observe a steep increase in both fluorescence anisotropy and fluorescence lifetime ($F(1,12) = 708.1$, $p \leq 0.0001$). The thermal profile of tPnA anisotropy (**c**) was measured in samples composed of POPC:PI(4,5)P_2 90:10. Controls were also carried out with samples of pure POPC. The impact of increasing concentrations of (18:0 20:4) PI(4,5)P_2 on the formation of $(16:0)_2$ PI(4,5)P_2 gel-like nanodomains (**d**) was detected in MLVs, containing POPC: PI(4,5)P_2 (95:5) at several $(16:0)_2$ to (18:0 20:4) PI(4,5)P_2 ratios, through the incorporation of tPnA at a 1:300 lipid ratio. Error bars for fluorescence anisotropy measurements represent the SD from $N = 3$ independent experiments.

PI(4,5)P_2 as an essential modulator of PIP2 nanodomain order, by decreasing the order within PI(4,5)P_2 nanodomains and thus maintaining their fluidity.

In summary, the results strongly point to two levels of ordering of the Ca^{2+}-induced PI(4,5)P_2 nanodomains above that of the bulk membrane. The formation of a gel-like phase reported by tPnA is unique to saturated PI(4,5)P_2 species and the corresponding T_m is higher than that of DPPC, suggesting that the PI(4,5)P_2 gel-like phase is even more stable than that of PC. On the other hand, a general local ordering reported by TMA-DPH is detected for all PI(4,5)P_2 species, including those polyunsaturated, and this ordering is not disrupted by temperature in the range evaluated here. The extent of this ordering is directly related to the level of acyl-chain unsaturation of PI(4,5)P_2.

These results are surprising given the large lipid headgroup in play. Biophysical properties such as membrane hydration, thickness, and flexibility within PI(4,5)P_2 nanodomains are expected to differ considerably from the rest of the membrane. Such changes are likely to have an impact in membrane events associated to PI(4,5)P_2. More ordered domains are expected to influence PI(4,5)P_2 headgroup accessibility, affecting lipid–protein interactions at the core of PI(4,5)P_2 role in the inner leaflet. In addition, PI(4,5)P_2 clusters have already been shown by us to prefer partitioning to more ordered domains[14]. It is likely that differences in acyl-chain saturation also regulate the distribution of PI(4,5)P_2 nanodomains within the plasma membrane.

CG MD simulations showcase the effects of the different PI(4,5)P_2 nanodomains on lipid bilayer structure and dynamics. To gather more structural information on the influence of the acyl-chain composition on the structure and dynamics of these nanodomains, we performed CG MD simulations using the Martini forcefield[34]. These CG MD simulations allow us to study not only PI(4,5)P_2 and PI(4,5)P_2 nanodomain dynamics, but also the effects of these nanodomains on the overall biophysical membrane properties. To this effect, large 50×50 nm^2 membrane patches containing 10% PI(4,5)P_2 were simulated, in the presence and absence of calcium, for at least 19 μs, for the three acyl-chain compositions. These membrane systems were built with 10% PI(4,5)P_2 to mimic the high concentrations of this lipid in the local environment around cation-induced PI(4,5)P_2 nanodomains. The last 2 μs of each simulation were considered for analysis (Fig. 5).

COMMUNICATIONS CHEMISTRY | https://doi.org/10.1038/s42004-021-00603-1 ARTICLE

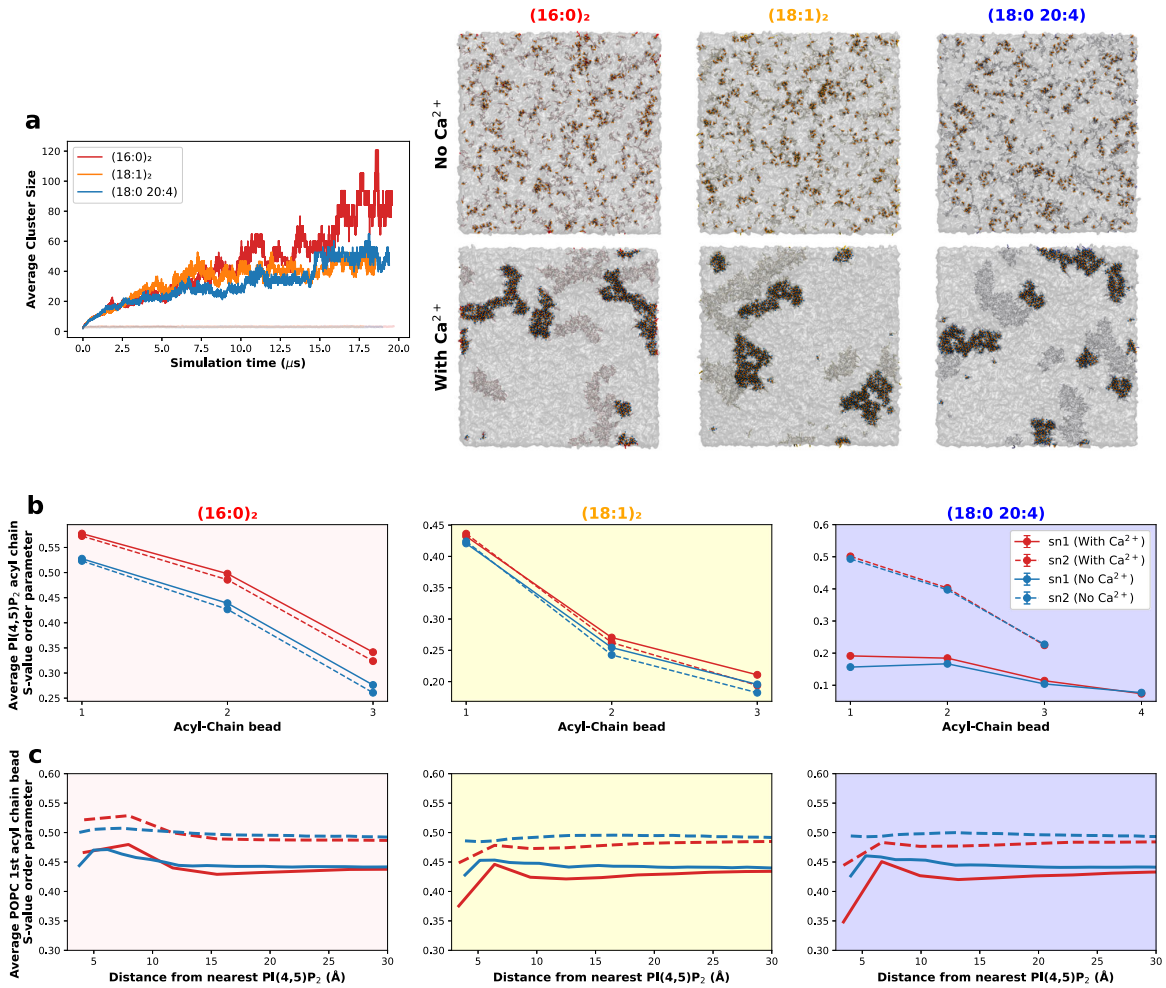


Fig. 5 CG MD simulations showcase the impact of acyl-chain composition on PI(4,5)P_2 and PI(4,5)P_2 nanodomain biophysical properties. a Average PI(4,5)P_2 cluster size over the course of the simulation for the three acyl-chain compositions studied, both in the presence and absence of calcium. Final simulation snapshots of the large membrane systems are also shown. PI(4,5)P_2 lipid headgroups are depicted in gray, with the phosphates discriminated in orange. PI(4,5)P_2 acyl chains are colored according to the corresponding color code. Ca^{2+} ions are represented in blue. The bulk POPC lipids are represented by the translucent gray surface. **b** PI(4,5)P_2 acyl-chain S-value order parameter for all acyl-chain beads of each composition, in the presence (red) and absence (blue) of calcium. **c** Dependency of the S-value order of POPC's first acyl-chain bonds on the distance from the nearest PI(4,5)P_2 molecule, in the presence and absence of calcium.

Analyzing the dynamics of Ca^{2+}-induced nanodomain formation for each composition, through the average PI(4,5)P_2 cluster size over the course of the simulation (Fig. 5a), we observed little difference between acyl-chain compositions. Nanodomain formation, at least at saturating divalent cation concentrations, seems to depend only on the probability of PI(4,5)P_2 molecules finding one another, while chelating a calcium ion. This is due to the slow PI(4,5)P_2-Ca^{2+} and PI(4,5)P_2-Ca^{2+}-PI(4,5)P_2 unbinding dynamics, caused by the extremely high affinity of the divalent cation for the negatively charged lipid headgroup[35]. Thus, the rate-limiting step of PI(4,5)P_2 aggregation, in saturating divalent cation concentrations, is lipid diffusion. As all compositions show similar diffusion coefficients (Supplementary Fig. 5D), they end up yielding similarly sized clusters with similar formation dynamics.

Looking at how nanodomain formation influences the biophysical properties of each PI(4,5)P_2 species, we observed an increase in lipid acyl-chain order (seen through the increase of the S-value parameter, which is described in detail in Supplementary Information) for all three compositions (Fig. 5b). This increase in order, however, is lower for the polyunsaturated and monounsaturated compositions, whereas the results for $(16:0)_2$ PI(4,5)P_2 show a considerable increase in the order parameter (~10%, at the first acyl-chain beads). For the $(16:0)_2$ and $(18:1)_2$ species, the increase in membrane order is transmitted up to the last acyl-chain bead, where an increase in order is still observed. For the polyunsaturated composition, however, nanodomain formation results in only minor increases in acyl-chain order. In fact, the stearoyl acyl-chain of this PI(4,5)P_2 species is barely affected by the formation of these nanodomains (1% increase) and the arachidoyl acyl chain only senses this increase in membrane order in the first two beads before dissipating. These results appear to suggest that the 18:0 20:4 composition can better disperse the increase in membrane order, caused by the organization of the PI(4,5)P_2 headgroups during nanodomain formation and, thus, can better maintain membrane fluidity. Overall, these findings are in excellent agreement with the previous fluorescence data that showed

an increase in PI(4,5)P_2 local membrane order across all the acyl-chain compositions studied, with a particularly high increase for the fully saturated composition, and more moderate changes for the polyunsaturated species.

In addition to the acyl-chain order, we also looked at the extent of their organization. As lipids become tightly packed, they become more structurally organized by extending and straightening their acyl chains in a geometric hexagonal disposition, creating a tightly packed unit with increased van der Waals interactions. These highly ordered hexagonal structures are hallmarks of the formation of gel phases[36,37]. To this purpose, we measured the average number of PI(4,5)P_2 acyl chains within hexagonal acyl-chain lattices (Supplementary Fig. 5B). As expected, in the absence of calcium, none of the PI(4,5)P_2 conformations form hexagonal acyl-chain lattices. However, in the presence of calcium, we observe an increase in acyl-chain hexagonality for the monounsaturated composition and a particularly drastic increase for the fully saturated composition. This effect is minimal for the polyunsaturated acyl-chain composition. These findings suggest that cation-induced PI(4,5)P_2 nanodomains formed from more saturated acyl-chain compositions exhibit not only higher order but also higher organization and packing. For the $(16:0)_2$ composition, one could expect that these structures could eventually act as nucleation points for the formation of gel or ripple phases.

The effects of these nanodomains are not merely confined to the PI(4,5)P_2 lipids within. We found that the nanodomains also affected the order of the surrounding lipids, especially those in the immediate vicinity of the nanodomains. For the poly- and monounsaturated compositions, the domains formed led to a decrease in acyl-chain order of the POPC lipids within 10 Å of PI(4,5)P_2 (Fig. 5c). This is likely the result of POPC having to accommodate the increased density of unsaturated acyl chains at the border of the PI(4,5)P_2 nanodomains. Interestingly, we observed the opposite effect for the fully saturated composition. The nanodomains formed with the fully saturated PI(4,5)P_2 led to an increase in acyl-chain order for POPC lipids within 10 Å of PI(4,5)P_2. For all compositions, this effect dissipates for POPC lipids further away from the nanodomains.

Curiously, we have found that the PI(4,5)P_2 nanodomains were strongly unregistered (decoupled) across the bilayer (Fig. 5a and Supplementary Fig. 5A). This may be caused by local membrane curvature/tension in the vicinity of Ca^{2+}-induced nanodomains, even though curvature was being restricted by an applied potential. As such, we prepared new systems designed to study membrane curvature during PI(4,5)P_2 nanodomain formation. To this effect, $50 \times 20\ nm^2$ asymmetric membrane patches, containing 10 mol% PI(4,5)P_2 in the bottom leaflet only, were simulated for up to 10 μs, for the three acyl-chain compositions (Fig. 6). In the absence of calcium, no PI(4,5)P_2 clustering occurs and no noticeable membrane undulations are observed.

When calcium was added to the simulations, consolidating PI(4,5)P_2 nanodomains quickly became associated with a markedly negative curvature for all PI(4,5)P_2 species (Fig. 6). There were no immediate apparent differences in the curvature generated by the three PI(4,5)P_2 species. These results are of particular importance, given the requirement of PI(4,5)P_2 for membrane-remodeling events requiring the formation of highly negatively curved intermediates[3,4,24,25]. As PI(4,5)P_2 itself is associated with positive curvature, as a result of the bulky headgroup, the presence of PI(4,5)P_2 within nanodomains could be responsible for alleviating tension on local curvature. Other Martini CG studies have found that inducing membrane curvature, by applying sufficient lateral pressure, could promote the enrichment of PI(4,5)P_2 at negatively curved membrane areas in the absence of Ca^{2+}[38]. However, in our systems without induced curvature, PI(4,5)P_2 alone was not sufficient to generate curvature undulations, and required Ca^{2+}-induced aggregation to associate with negative curvature.

To get further structural insight on the formation of gel-like nanodomains of $(16:0)_2$ PI(4,5)P_2 in the presence of calcium, we performed additional simulations in an attempt to detect the formation of gel-like domains in smaller 500 lipid systems. Gel phase-like nanodomains were obtained, albeit at higher PI(4,5)P_2 mol % (50 mol%) and at lower temperatures (280 K or 6.85 °C) (Fig. 7a) than the experimental conditions. These domains were PI(4,5)P_2-rich, but contained some POPC molecules incorporated in their structure. The underestimation of the gel-fluid lipid phase transition temperature is well known for Martini forcefield, as seen for the canonical case of DPPC[39]. In addition, it is known that small amounts of disorder-inducing membrane components (such as POPC) are sufficient to disrupt gel-like phases in Martini systems[40].

To overcome the kinetic barrier posed by the formation of a stable nucleation patch, which slows down the study of the formation of gel phases in Martini, we applied a restrained crystal seeding technique. Briefly, we placed a PI(4,5)P_2 gel crystal, obtained from one of the previous gel-forming simulations (Fig. 7b), and restrained it in the middle of a membrane system (Fig. 7c). This restrained crystal acts as a stable nucleation point, which we can use to probe the gel growth phase for lower PI(4,5)P_2 mol% and higher temperatures (Fig. 7c). Using this strategy, the nucleation is shown to induce a consolidation of $(16:0)_2$ PI(4,5)P_2 gel phase formation for the lipid concentration range employed in the experimental studies (10 mol%) (Fig. 7c). In these systems, we followed the average first acyl-chain bead S-value, which gave us an idea of the order of the gel, as well as the hexagonality of the system, which hints at the number of lipids in the gel lattice. The results were compared against a DPPC system built in the same manner.

Looking at the S-value order parameter over a temperature gradient, calculated exclusively from the gel-forming lipid (PI(4,5)P_2 or DPPC) (Fig. 7f), no clear inflexion point is observed for any of the systems. This is due to the restrained gel crystal in all systems, which prevents the gel phase from completely disappearing, by allowing for the transient incorporation of unrestrained lipids in the crystal. Regardless, PI(4,5)P_2 in the presence of calcium presents a higher S-value than that of DPPC in the same conditions. This confirms the experimental evidence from tPnA fluorescence, indicating that that the gel domains formed by PI(4,5)P_2 are more ordered than the canonical gel-forming lipid DPPC. From the slope of this curve, we can infer the thermal stability of these gel domains and, once again, we observe that the domains formed by PI(4,5)P_2 in the presence of calcium are more stable than those formed by DPPC in the same conditions.

Analyzing the system-wide hexagonality (Fig. 7d), we observe that both PI(4,5)P_2 systems lead to higher hexagonality values, regardless of the presence of calcium, when compared to DPPC. This result hints that PI(4,5)P_2 must have a higher propensity to organize itself and the surrounding lipids into a hexagonal lattice. This suggests that the interactions between PI(4,5)P_2 headgroups, further stabilized by Ca^{2+}, lead to a headgroup organization that enhances gel formation.

Within the plasma membrane, an increase of $(16:0)_2$ PI(4,5)P_2 levels in a context of elevated calcium is likely to give rise to gel-like domains in the inner leaflet, which would not only influence PI(4,5)P_2 lateral organization and availability for interaction with other molecular partners but also potentially impact local membrane properties, such as curvature, thickness, and permeability. Concentrations of Mg^{2+} within mM levels have been shown by us to induce PI(4,5)P_2 clustering as well[14], and although not explored in this work, it is likely that high free Mg^{2+} levels, such as those observed intracellularly, might generate

COMMUNICATIONS CHEMISTRY | https://doi.org/10.1038/s42004-021-00603-1 ARTICLE

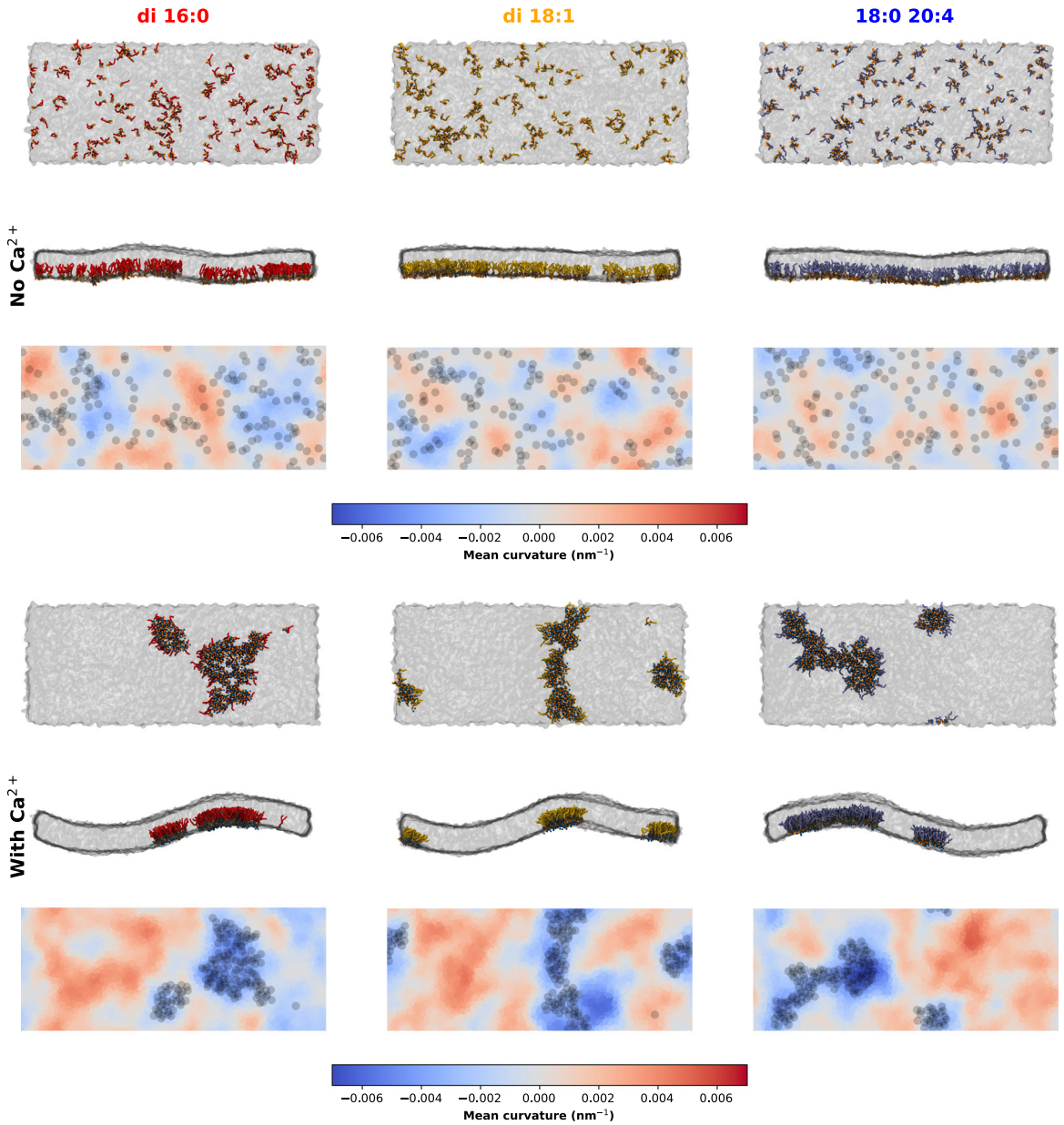


Fig. 6 Calcium-induced PI(4,5)P_2 clustering induces negative membrane curvature, as seen by CG MD simulations. Curvature analysis of snapshots from asymmetric membrane simulations containing 10 mol% PI(4,5)P_2 in the inner leaflet. Snapshots are shown for the three acyl-chain compositions studied, both in the presence and absence of calcium. For each system, a top and side view are presented, as well as a top view colored by local mean curvature value. PI(4,5)P_2 phosphodiester (PO4) beads are represented by the black circles. In the molecular representations, PI(4,5)P_2 headgroups are depicted in gray, with the phosphates discriminated in orange. PI(4,5)P_2 acyl chains are colored according to the corresponding color code. Ca^{2+} ions are represented in blue. The bulk POPC lipids are represented by the translucent gray surface.

similar saturated PI(4,5)P_2 gel-like domains. If that is the case, then the formation of saturated PI(4,5)P_2 gel-like nanodomains will be constitutive and independent of calcium levels, with unforeseen consequences.

Discussion

By investigating representative acyl-chain compositions of the spectrum observed in vivo, we report the effect of acyl-chain saturation on the calcium-dependent changes on the overall properties of PI(4,5)P_2-containing membranes. In the absence of divalent cations, all acyl-chain compositions behave in the same manner. However, for cation-induced PI(4,5)P_2 nanodomains, it is clear that the acyl-chain compositions yield structures with significantly different biophysical properties. The key observation is that increasing saturation yields more ordered and structured nanodomains, which can, in the case of the fully saturated composition, culminate in the formation of gel-like nanodomains. Although calcium-dependent gel-like $(16:0)_2$

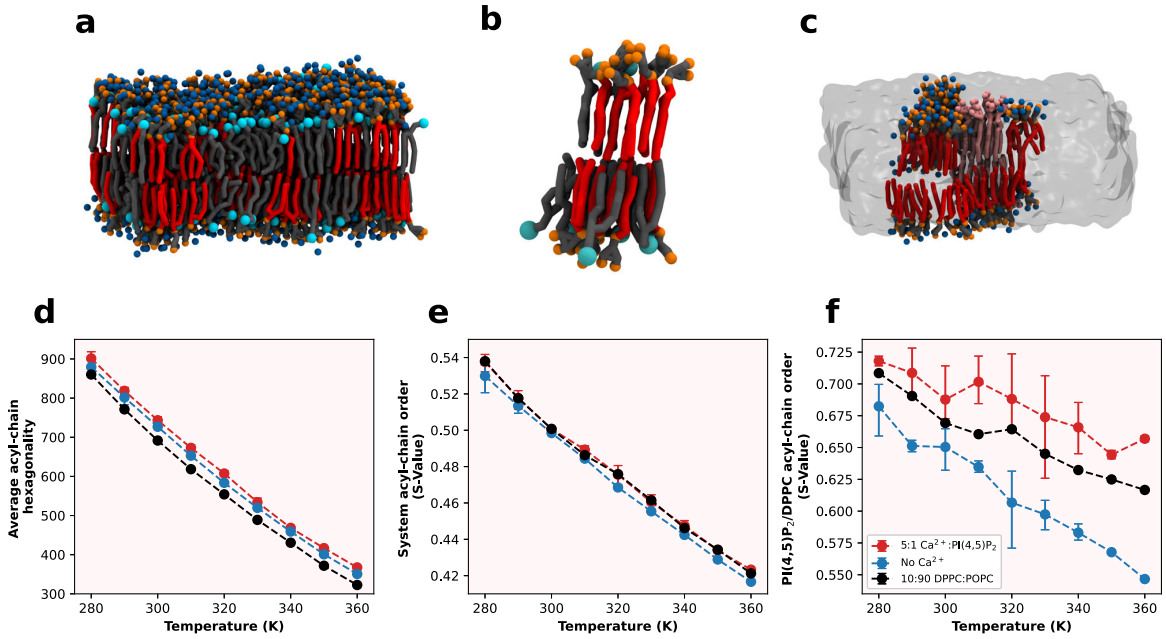


Fig. 7 CG MD study of (16:0)$_2$ PI(4,5)P$_2$ gel-phase nanodomains. a Final snapshot of a gel-forming simulation containing 50% mol PI(4,5)P_2 at 280 K. PI(4,5)P_2 gel crystals (**b**) were obtained from this simulation and used for the crystal scaffold simulations (**c**), which were used to characterize gel-phase properties. **d** Average system-wide acyl-chain hexagonality dependency with temperature for each of the gel-phase crystal scaffold systems studied. **e** Average system-wide S-value order parameter, calculated on the first acyl-chain bead of every lipid acyl chain, for each of the gel-phase crystal scaffold systems studied. **f** S-value order parameter for the first acyl-chain bead of every gel-forming lipid acyl chain for each of the systems studied. Error bars for all measurements represent the SD from $N = 3$ independent simulation experiments.

PI(4,5)P_2 nanodomains are shown to have a melting temperature of ~45 °C, the increase in general local ordering reported for all PI(4,5)P_2 acyl chains (in the presence of calcium) could not be disrupted by the temperatures tested here (up to 70 °C). The extent of this ordering was directly related to the level of acyl-chain unsaturation of PI(4,5)P_2.

The increased membrane order detected within calcium-dependent PI(4,5)P_2 nanodomains is likely to have a profound impact on PI(4,5)P_2 signaling and on the function and organization of effector proteins. Remarkably, the formation of this phase was largely abrogated in the presence of 1-stearoyl-2-arachidonyl PI(4,5)P_2. Our findings thus provide an explanation of why evolution strongly favored the 18:0 20:4 acyl-chain composition for PI(4,5)P_2 in mammalian cells. The presence of the polyunsaturated acyl-chain is expected to guarantee not only low bending rigidity, but also that the ordering within PI(4,5)P_2 nanodomains remains low. On the other hand, the saturated acyl-chain in the *sn*-1 position could be important to reduce membrane permeability[20].

Another possible outcome of the formation of ordered calcium-induced PI(4,5)P_2 clusters is that interactions with proteins exhibiting lower binding affinity for the phospholipid could be switched-off, whereas proteins forming a tighter complex with PI(4,5)P_2 remain bound. In fact, both the presence of calcium as the modified membrane structure are expected to have an impact on PI(4,5)P_2–protein interactions. In this scenario, the triggering of PI(4,5)P_2 clustering could offer an additional level of regulation for PI(4,5)P_2 signaling.

The molecular simulations carried out here were in clear agreement with the experimental data. Moreover, the simulations identified that the more unsaturated PI(4,5)P_2 species were more effective in dissipating the increase in order upon clustering over the length of their acyl chains. In addition, simulations also showed that the increase in membrane order was not limited to PI(4,5)P_2 domains, but extended to the surrounding lipids as well.

We also report an association of calcium-dependent PI(4,5)P_2 nanodomains with negative curvature, a phenomenon that is likely to play a role in membrane-remodeling events, directly regulated by the presence of PI(4,5)P_2. These results suggest that the monodisperse PI(4,5)P_2–PI(4,5)P_2 nanodomain relation might play a key role in controlling the formation of fusion intermediates required for some of the membrane-remodeling events associated with this phospholipid.

Methods

Materials. POPC, DOPC, 1,2-dioleoyl-*sn*-glycero-3-phospho-(1'-*myo*-inositol-4',5'-bisphosphate) (di18:1 PI(4,5)P_2), 1-stearoyl-2-arachidonoyl-*sn*-glycero-3-phospho-(1'-*myo*-inositol-4',5'-bisphosphate) (18:0 20:4 PI(4,5)P_2), 1-oleoyl-2-{6-[4-(dipyrrometheneboron difluoride)butanoyl]amino}hexanoyl-*sn*-glycero-3-phosphoinositol-4,5-bisphosphate (TopFluor PI(4,5)P_2), and 1,2-dioleoyl-*sn*-glycero-3-phosphoethanolamine-N-(cap biotinyl) (biotinylated DOPE) were purchased from Avanti Polar Lipids (Alabaster, AL, USA). 1,2-Dipalmitoyl-*sn*-glycero-3-phospho-(1'-*myo*-inositol-4',5'-bisphosphate) (di16:0 PI(4,5)P2) was from Echelon Biosciences (Salt Lake City, UT, USA). Lipid stock solutions were prepared in chloroform, with the exception of the phosphoinositides, which were prepared in chloroform:methanol (MeOH) 2:1 (v/v). Both solvents were obtained from Merck (Darmstadt, Germany) and were of spectroscopic grade. 4-(2-Hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES), ethanol (EtOH), NaCl, Sucrose, EDTA, glucose, and $CaCl_2$ were from Sigma-Aldrich (St. Louis, MO, USA). TMA-DPH, tPnA, Rhodamine 110, and Fluo-5N were from Molecular Probes, Invitrogen (Eugene, OR, USA).

Liposome preparation. MLVs were prepared by solubilization of lipid films. Briefly, the lipid mixtures were prepared from phospholipid stock solutions, dried under a nitrogen flux, and left in vacuum for 3 h to remove traces of solvent. MLVs were then obtained through the solubilization of the lipid films in the appropriate experimental buffer. Freeze–thaw cycles were performed, using liquid nitrogen and a water bath typically set to 60 °C. The thawing temperature used was always above the melting temperature of the lipid with the highest melting temperature in the mixture, to re-equilibrate and homogenize the samples. LUVs were prepared by extrusion of MLVs[41],

at room temperature, using an Avanti Mini-Extruder (Alabaster, AL, USA) and 100 nm pore-size polycarbonate membranes (Whatman, Little Chalfont, UK). Typically, at least 21 passages through the extruder were performed.

Giant unilamellar vesicles (GUVs) were obtained by gel-assisted formation, based on a method previously described[42]. The lipid mixtures were prepared, from stock solutions, in chloroform to a final concentration of 1.5 mM. DOPE-Cap-biotin was included in the mixture[43] at a biotinylated lipid/total lipid ratio of 1:750,000. A solution of 5% (w/w) polyvinyl alcohol (PVA) ($M_W \sim 145{,}000$ Da) and 280 mM of sucrose was spread in a µ-slide chamber from Ibidi (Munich, Germany) and left to dry for 15 min at 50 °C. The desired lipid mixture was then spread on the PVA surface. The solvent was evaporated for 15 min under vacuum. Afterwards, the appropriate buffer solution was added, allowing for GUV formation for 60 min at room temperature. Then, GUVs were transferred to a µ-slide chamber with the appropriate coating and left to rest for 10 min, to allow for GUV deposition and immobilization. GUVs were immobilized in the µ-slide chamber through binding of the biotinylated lipid to surface avidin, which was previously used to coat the chamber slide. Coating was carried out using a mixture of bovine serum albumin (BSA), BSA-biotin (9:1 mol/mol), and avidin. Ibidi µ-slide chambers were coated with 300 µL of a 0.9 mg ml^{-1} BSA and 0.1 mg mL^{-1} BSA-biotin mixture for 1 h. Afterwards, the chambers were washed with Milli-Q water and covered with a second layer of 300 µL of 0.01 mg mL^{-1} avidin for 1 h. BSA, BSA-biotin, and avidin solutions were prepared with Milli-Q water. Before adding the GUV suspension, the chambers were also washed with Milli-Q water.

Unless indicated otherwise, all liposome samples were suspended in a buffer solution containing 10 mM HEPES, 140 mM NaCl, and either 400 µM $CaCl_2$ or 5 mM EDTA. Buffer Ca^{2+} concentrations in the micromolar range were determined using the fluorescent calcium indicator Fluo-5N pentapotassium salt, following the instructions of the manufacturer.

For samples containing TF-PI(4,5)P_2, the fluorescent lipid probe was added to the lipid film mixtures at a 0.1 mol% ratio. For TMA-DPH and tPnA samples, the fluorescent lipid probe was added to the lipid film mixtures at a 1:300 lipid ratio (0.33 mol% ratio). MLVs were used for fluorescence studies using TMA-DPH and tPnA, where the probe's anisotropy and fluorescence lifetime were followed. These properties are not influenced by multilamellarity. For studies using TF-PI(4,5)P_2, however, LUVs were used in order to avoid possible energy transfer events between bilayers, which would complicate the interpretation of homo-FRET results.

Steady-state fluorescence spectroscopy. Fluorescence measurements were carried out with a SLM-Aminco 8100 Series 2 fluorescence spectrophotometer (Rochester, NY, USA) with double excitation and emission monochromators (MC-400), in right angle geometry. The light source was a 450 W Xe arc lamp and the reference a Rhodamine B quantum counter solution. Quartz cuvettes (0.5 × 0.5 cm^2) from Hellma (Müllheim, Germany) were used and temperature set to 25 °C. Polarization of excitation and emission light was obtained with Glan-Thompson polarizers. Blank subtraction was taken into account in all measurements.

Steady-state fluorescence anisotropy, $<r>$ is defined as[27]:

$$<r> = \frac{I_{VV} - G^*I_{VH}}{I_{VV} + 2^*G^*I_{VH}};\ G = \frac{I_{HV}}{I_{HH}}$$

where I_{Vj} represents the steady-state vertical (parallel, I_{VV}) and horizontal (perpendicular, I_{VH}) components of the fluorescence emission with vertically polarized excitation. The G factor is measured using the vertical (I_{HV}) and horizontal (I_{HH}) components of the fluorescence emission with horizontaly polarized excitation.

Time-resolved fluorescence spectroscopy. Fluorescence decay measurements were carried out using the time-correlated single-photon timing technique, as described elsewhere[44]. The emission wavelength was selected by a Jobin Yvon HR320 monochromator (Horiba Jobin Yvon, Kyoto, Japan). Then, 0.5 × 0.5 cm^2 quartz cuvettes from Hellma were used. Blank decays were acquired, and photon counts were negligible. The fluorescence intensity decays were described by a sum of exponentials:

$$i(t) = \sum_i \alpha_i \exp\left(-\frac{t}{\tau_i}\right)$$

where α_i is the normalized amplitude and τ_i is the ith lifetime component. The amplitude-weighted average lifetime is defined as:

$$\overline{\tau} = \sum_i \alpha_i \tau_i$$

Data analysis was performed with the TRFA software (Scientific Software Technologies Center, Minsk, Belarus), based on Levenberg–Marquardt nonlinear least-squares fitting. The goodness of the fit was judged from the experimental χ^2-weighted residuals and autocorrelation plot. In every analysis, χ^2 was below 1.3, and both the residuals and the autocorrelation were randomly distributed around zero.

Statistical analysis. Statistical analysis was performed for steady-state fluorescence anisotropy results using regular two-way analysis of variance tests. The two factors accounted for were PI(4,5)P_2 concentration and whether the sample was in the presence or absence of calcium. F-statistics, degrees of freedom, and p-values are reported in the manuscript as called for. No post hoc comparisons were performed. Statistical analysis was performed using GraphPad Prism, GraphPad Software, La Jolla California USA.

SLB preparation. SLBs were formed by the vesicle fusion rupture method[45,46]. Lipids were mixed at the appropriate molar proportions in a round-bottom flask, dried under nitrogen stream, and left overnight in vacuum to remove any traces of chloroform. Lipid mixtures were hydrated in citrate buffer, to neutralize PI(4,5)P_2 negative charge and promote vesicle adsorption in the negatively charged mica surface. The multilamellar lipid suspension was power sonicated using a Vibra-Cell ultrasonicator (Sonics & Materials, Newtown, CT, USA) for three times, in cycles of 3 min, with pulsed sonication and 3 min resting in ice. The clear lipid suspension was centrifuged in a microcentrifuge Z 233 M-2 (HERMLE Labortechnik, Wehingen, Germany), for 5 min at 16,500 × g to remove titanium particles, large vesicles, and debris. After this, 500 µL of a 10× diluted lipid suspension was pipetted onto freshly cleaved mica along with a 3 mM final concentration of $CaCl_2$, in a custom-built well. The sample was incubated in a humidity chamber at 60 °C, above all lipids main transition temperature, for a maximum of 60 min. This procedure allows small unilamellar vesicles to adsorb and rupture on the surface of the mica, forming a flat continuous bilayer[45,46]. After incubation, the bilayer was washed 10–25 times with 100 µL of 60 °C HEPES buffer, using a pipette. The washing procedure was performed parallel to the bilayer surface. This ensures that unfused vesicles, either in suspension or deposited on the bilayer surface, are removed. In all samples, the hydrated bilayers were let to cool down at room temperature, enabling phase separation to occur.

Atomic force microscopy. AFM was performed using a JPK Nanowizard IV (JPK Instruments, Berlin, Germany). Bilayers were imaged in quantitative imaging (QI) mode, a recent innovation in which the apparatus modulates the z-piezo to perform a fast force curve on each pixel of the image[47,48]. This avoids lateral friction and allows for better control of the tip force during measurements. QI mode allows several mechanical properties to be calculated from the force applied and the tip-sample separation. Throughout the imaging, the maximum applied force was 200 pN, in order not to affect the sample structure[49]. Images were obtained with a resolution of 256 × 256 pixels, at a scan rate of 1 Hz. AFM measurements were performed at room temperature, which varied from 22 °C to 25 °C.

Before measurements, cantilever spring constants were quantified by the thermal noise method[50] and cantilever sensitivity was measured by performing a force curve on a clean freshly cleaved mica surface, in HEPES buffer. For the used qpBioAC CB2 probes (NanoWorld AG, Neuchâtel, Switzerland), the spring constants obtained were on the 0.06–0.18 N m^{-1} range and the sensitivity 7.6 ± 1.2 nm V^{-1}. Approximately three to five separated areas, of 10 × 10 µm^2 and 20 × 20 µm^2, from at least two bilayers, prepared in different days, from different lipid stocks, were imaged to obtain representative data and assure the reproducibility of the measurements.

Topographical images were analyzed with first or second-level flattening, using the JPK data processing software. The sizes of the domains were evaluated by several cross-sections on the SLB images. The number of cross-sections varied from 70 to ~300, to obtain representative data for the domains formed in each SLB.

MD simulations. The Martini 2.2 CG model for biomolecular simulations was employed throughout this study[34]. The used PI(4,5)P_2 topologies were constructed combining an in-house improved version of the existing PI(4,5)P_2 headgroup parameters[51] with the standard Martini lipid acyl-chain topologies for each of the compositions studied. The rationale behind the improved PI(4,5)P_2 headgroup parameters is described in detail in the Supplementary MMethods (Supplementary Figs. 7–11). These topologies are provided, in the standard GROMACS format, in Supporting Information (PIP2.itp) and are also available from the authors. All other topologies were obtained directly from the Martini developers' website[52]. All simulations were run using the GROMACS simulation package version 2018[53]. The membrane systems were built and solvated using the *insane.py* CG building tool[54]. Counterions were added to neutralize the system as necessary and 140 mM NaCl was added on top of that to all systems. When required, Ca^{2+} was added at the appropriate Ca^{2+}:PI(4,5)P_2 ratio by replacing water particles and maintaining the system charge neutral by adding Cl^- counterions.

Nonbonded interactions were cutoff at 1.1 nm and Coulombic interactions were treated using reaction-field electrostatics[55]. The particle neighbor list was updated using the Verlet list scheme. The v-rescale thermostat[56] was used with a tau-t of 4.0 ps to maintain the temperature at 300 K. Constant pressure was semi-isotropically coupled to 1.0 bar using a Parrinello–Rahman barostat[57] with a relaxation time of 16.0 ps. After initial energy minimization and pressure/temperature equilibration runs, simulations were run at a 20 fs time step.

Three types of systems were simulated as follows:

— *large* systems consisting of ~8500 lipids at a 90:10 POPC:PI(4,5)P_2 molar ratio in a 50 × 50 × 13 nm^3 box were simulated for at least 19 µs. Simulations were ran for each composition both in the presence and absence of Ca^{2+}. For these systems only, membrane undulations were limited by applying a weak (200 kJ mol^{-1} nm^{-2}) flat-bottomed potential in z, restraining the glycerol beads of all lipids to remain within a 2.0 nm vertical distance of the simulation box center.

— *asymmetric* membrane systems, consisting of ~3500 lipids, in a 50 × 20 × 13 nm^3 box, which were simulated for at least 10 µs. Their top membrane

layer consists solely of POPC, whereas the bottom layer consists of a 90:10 POPC:PI(4,5)P_2 mixture. The total number of lipids in each lipid layer was adjusted, matching the top and bottom leaflet areas, thereby avoiding the introduction of tension. To this end, the area per lipid of each layer's composition was calculated from smaller symmetric simulations ran for each of the compositions, both in the presence and absence of Ca^{2+}.

— *crystal seed* systems, where gel-phase formation was probed by simulations of a PI(4,5)P_2 gel crystal seed restrained in the middle of an initially fluid membrane. Gel-phase crystals were obtained through regular simulations of POPC: PI(4,5)P_2 mixtures (~500 lipids; $15 \times 10 \times 13$ nm^3 boxes) with PI(4,5)P_2 molar ratios ranging from 30% to 50 %, simulated for up to 10 µs. These simulations were run with Ca^{2+}, at a 5:1 Ca^{2+}:PI(4,5)P_2, using the same general MD parameters, except for the temperature, which was maintained at 280 K. From these initial simulations, PI(4,5)P_2-rich gel phases were formed, from which small ~20 lipid crystals were excised. These crystals were inserted in membranes of the same size at a 90:10 POPC:PI(4,5)P_2 molar ratio, simulated for at least 5 µs. These were run either in the absence or in the presence of calcium, at a 5:1 Ca^{2+}:PI(4,5)P_2 ratio, and at temperatures ranging from 280 to 360 K. Throughout all the simulation steps, the acyl chain and glycerol beads of the crystal seeds were position-restrained in all dimensions with a force constant of 10,000 kJ mol^{-1} nm^{-2}. The headgroups were left unrestrained to better accommodate new lipids in the growing crystal. Three replicates, with independent gel-phase crystals, were run for each condition.

All simulations were analyzed making use of in-house developed Python3 programs using the MDAnalysis package[58]. We also used the IPython[59], numpy[60], SciPy[61], scikit-learn[62], and matplotlib[63] packages for scientific computing in Python. The PyCurv[64] package was used for curvature estimation. Visualization and renderization of the simulations was performed with the molecular graphics viewer VMD[65]. The last 2 µs of each simulation were used for analysis. See the "Methods" section of the Supporting Information for details on the analysis methods.

Data availability

Extended Methods and Materials providing details on the rationale behind the improvement of the Martini 2.2 PI(4,5)P2 model, details behind the simulation analysis, additional time-resolved fluorescence spectroscopy results, fluorescence correlation spectroscopy results on TF-PI(4,5)P2 incorporation in nanodomains, AFM analysis of nanodomain sizes, and further simulation analysis of PI(4,5)P2 biophysical properties are available within Supplementary Information. The molecule description file containing the CG parameters for the Martini 2.2 PI(4,5)P_2 models used in this study in text format compatible with GROMACS software, as well as initial and final configurations for the molecular dynamics simulation systems are available from https://github.com/MeloLab/PhosphoinositideParameterization. Other data are available from the corresponding author upon reasonable request.

Received: 16 May 2021; Accepted: 10 November 2021;
Published online: 29 November 2021

References

1. Monteiro, M. E., Sarmento, M. J. & Fernandes, F. Role of calcium in membrane interactions by PI(4,5)P2-binding proteins. *Biochem. Soc. Trans.* **42**, 1441–6 (2014).
2. Di Paolo, G. & De Camilli, P. Phosphoinositides in cell regulation and membrane dynamics. *Nature* **443**, 651–657 (2006).
3. Koch, M. & Holt, M. Coupling exo- and endocytosis: an essential role for PIP2 at the synapse. *Biochim. Biophys. Acta* **1821**, 1114–1132 (2012).
4. Martin, T. F. J. in *Subcellular Biochemistry* Vol. 59, 111–130 (Springer New York, 2012).
5. dos Remedios, C. G. & Nosworthy, N. J. in *Actin-Binding Proteins and Disease* 290–297, https://doi.org/10.1007/978-0-387-71749-4_12 (Springer New York, 2008).
6. Suh, B.-C. & Hille, B. PIP 2 is a necessary cofactor for ion channel function: how and why? *Annu. Rev. Biophys.* **37**, 175–195 (2008).
7. Mücksch, F. et al. Quantification of phosphoinositides reveals strong enrichment of PIP2 in HIV-1 compared to producer cell membranes. *Sci. Rep.* **9**, 17661 (2019).
8. Bertin, A. et al. Phosphatidylinositol-4,5-bisphosphate promotes budding yeast septin filament assembly and organization. *J. Mol. Biol.* **404**, 711–731 (2010).
9. Mandal, K. Review of PIP2 in cellular signaling, functions and diseases. *Int. J. Mol. Sci.* **21**, 8342 (2020).
10. Sarmento, M. J., Coutinho, A., Fedorov, A., Prieto, M. & Fernandes, F. Ca2+ induces PI(4,5)P2 clusters on lipid bilayers at physiological PI(4,5)P2 and Ca2+ concentrations. *Biochim. Biophys. Acta* **1838**, 822–830 (2014).
11. Rauch, M. E., Ferguson, C. G., Prestwich, G. D. & Cafiso, D. S. Myristoylated alanine-rich C kinase substrate (MARCKS) sequesters spin-labeled phosphatidylinositol 4,5-bisphosphate in lipid bilayers. *J. Biol. Chem.* **277**, 14068–14076 (2002).
12. Wang, J. et al. Lateral sequestration of phosphatidylinositol 4,5-bisphosphate by the basic effector domain of myristoylated alanine-rich C kinase substrate is due to nonspecific electrostatic interactions. *J. Biol. Chem.* **277**, 34401–34412 (2002).
13. Wen, Y., Vogt, V. M. & Feigenson, G. W. Multivalent cation-bridged PI(4,5)P2 clusters form at very low concentrations. *Biophys. J.* **114**, 2630–2639 (2018).
14. Sarmento, M. J., Coutinho, A., Fedorov, A., Prieto, M. & Fernandes, F. Membrane order is a key regulator of divalent cation-induced clustering of PI(3,5)P2 and PI(4,5)P2. *Langmuir* **33**, 12463–12477 (2017).
15. Bozelli, J. C. & Epand, R. M. Specificity of acyl chain composition of phosphatidylinositols. *Proteomics* **19**, 1900138 (2019).
16. De Craene, J.-O., Bertazzi, D., Bär, S. & Friant, S. Phosphoinositides, major actors in membrane trafficking and lipid signaling pathways. *Int. J. Mol. Sci.* **18**, 634 (2017).
17. D'Souza, K. & Epand, R. M. Enrichment of phosphatidylinositols with specific acyl chains. *Biochim. Biophys. Acta* **1838**, 1501–1508 (2014).
18. Hicks, A. M., DeLong, C. J., Thomas, M. J., Samuel, M. & Cui, Z. Unique molecular signatures of glycerophospholipid species in different rat tissues analyzed by tandem mass spectrometry. *Biochim. Biophys. Acta* **1761**, 1022–1029 (2006).
19. Traynor-Kaplan, A. et al. Fatty-acyl chain profiles of cellular phosphoinositides. *Biochim. Biophys. Acta* **1862**, 513–522 (2017).
20. Manni, M. M. et al. Acyl chain asymmetry and polyunsaturation of brain phospholipids facilitate membrane vesiculation without leakage. *Elife* **7**, e34394 (2018).
21. Mujalli, A. et al. Profiling of phosphoinositide molecular species in human and mouse platelets identifies new species increasing following stimulation. *Biochim. Biophys. Acta* **1863**, 1121–1131 (2018).
22. Koizumi, A. et al. Increased fatty acyl saturation of phosphatidylinositol phosphates in prostate cancer progression. *Sci. Rep.* **9**, 13257 (2019).
23. Saad, J. S. et al. Structural basis for targeting HIV-1 Gag proteins to the plasma membrane for virus assembly. *Proc. Natl Acad. Sci. USA* **103**, 11364–9 (2006).
24. James, D. J., Khodthong, C., Kowalchyk, J. A. & Martin, T. F. J. Phosphatidylinositol 4,5-bisphosphate regulates SNARE-dependent membrane fusion. *J. Cell Biol.* **182**, 355–366 (2008).
25. Mu, L. et al. A phosphatidylinositol 4,5-bisphosphate redistribution-based sensing mechanism initiates a phagocytosis programing. *Nat. Commun.* **9**, 4259 (2018).
26. Tadross, M. R., Tsien, R. W. & Yue, D. T. Ca2+ channel nanodomains boost local Ca^{2+} amplitude. *Proc. Natl Acad. Sci. USA* **110**, 15794–15799 (2013).
27. Lakowicz, J. R. *Principles of Fluorescence Spectroscopy* (Springer, 2006).
28. Attwood, S., Choi, Y. & Leonenko, Z. Preparation of DOPC and DPPC supported planar lipid bilayers for atomic force microscopy and atomic force spectroscopy. *Int. J. Mol. Sci.* **14**, 3514–3539 (2013).
29. do Canto, A. M. T. M. et al. Diphenylhexatriene membrane probes DPH and TMA-DPH: a comparative molecular dynamics simulation study. *Biochim. Biophys. Acta* **1858**, 2647–2661 (2016).
30. Shrivastava, S., Paila, Y. D., Dutta, A. & Chattopadhyay, A. Differential effects of cholesterol and its immediate biosynthetic precursors on membrane organization. *Biochemistry* **47**, 5668–5677 (2008).
31. Marsh, D. *Handbook of Lipid Bilayers* (CRC, 2013).
32. Sklar, L. A. The partition of cis-parinaric acid and *trans*-parinaric acid among aqueous, fluid lipid, and solid lipid phases. *Mol. Cell. Biochem.* **32**, 169–177 (1980).
33. Florine, K. I. & Feigenson, G. W. Influence of the calcium-induced gel phase on the behavior of small molecules in phosphatidylserine and phosphatidylserine–phosphatidylcholine multilamellar vesicles. *Biochemistry* **26**, 1757–1768 (1987).
34. Marrink, S. J., Risselada, H. J., Yefimov, S. & Tieleman, D. P. & De Vries, A. H. The MARTINI force field: coarse grained model for biomolecular simulations. *J. Phys. Chem. B* **111**, 7812–7824 (2007).
35. Alessandri, R. et al. Pitfalls of the Martini Model. *J. Chem. Theory Comput.* **15**, 5448–5460 (2019).
36. Koynova, R. & Caffrey, M. Phases and phase transitions of the phosphatidylcholines. *Biochim. Biophys. Acta* **1376**, 91–145 (1998).
37. Nagle, J. F. & Tristram-Nagle, S. Structure of lipid bilayers. *Biochim. Biophys. Acta* **1469**, 159–95 (2000).
38. Lin, X. et al. Roles of PIP2 in the membrane binding of MIM I-BAR: insights from molecular dynamics simulations. *FEBS Lett.* **592**, 2533–2542 (2018).
39. Marrink, S. J., Risselada, J. & Mark, A. E. Simulation of gel phase formation and melting in lipid bilayers using a coarse grained model. *Chem. Phys. Lipids* **135**, 223–244 (2005).
40. Carpenter, T. S. et al. Capturing phase behavior of ternary lipid mixtures with a refined Martini coarse-grained force field. *J. Chem. Theory Comput.* **14**, 6050–6062 (2018).
41. Mayer, L. D., Hope, M. J. & Cullis, P. R. Vesicles of variable sizes produced by a rapid extrusion procedure. *Biochim. Biophys. Acta* **858**, 161–168 (1986).
42. Weinberger, A. et al. Gel-assisted formation of giant unilamellar vesicles. *Biophys. J.* **105**, 154–164 (2013).

43. Sarmento, M. J., Prieto, M. & Fernandes, F. Reorganization of lipid domain distribution in giant unilamellar vesicles upon immobilization with different membrane tethers. *Biochim. Biophys. Acta* **1818**, 2605–2615 (2012).
44. Loura, L. M. S., Fedorov, A. & Prieto, M. Partition of membrane probes in a gel/fluid two-component lipid system: a fluorescence resonance energy transfer study. *Biochim. Biophys. Acta* **1467**, 101–112 (2000).
45. El Kirat, K., Morandat, S. & Dufrêne, Y. F. Nanoscale analysis of supported lipid bilayers using atomic force microscopy. *Biochim. Biophys. Acta* **1798**, 750–765 (2010).
46. Franquelim, H. G., Gaspar, D., Veiga, A. S., Santos, N. C. & Castanho, M. A. R. B. Decoding distinct membrane interactions of HIV-1 fusion inhibitors using a combined atomic force and fluorescence microscopy approach. *Biochim. Biophys. Acta* **1828**, 1777–1785 (2013).
47. Chopinet, L., Formosa, C., Rols, M. P., Duval, R. E. & Dague, E. Imaging living cells surface and quantifying its properties at high resolution using AFM in QI^{TM} mode. *Micron* **48**, 26–33 (2013).
48. Smolyakov, G., Formosa-Dague, C., Severac, C., Duval, R. E. & Dague, E. High speed indentation measures by FV, QI and QNM introduce a new understanding of bionanomechanical experiments. *Micron* **85**, 8–14 (2016).
49. Aufderhorst-Roberts, A., Chandra, U. & Connell, S. D. Three-phase coexistence in lipid membranes. *Biophys. J.* **112**, 313–324 (2017).
50. Hutter, J. L. & Bechhoefer, J. Calibration of atomic-force microscope tips. *Rev. Sci. Instrum.* **64**, 1868–1873 (1993).
51. López, C. A., Sovova, Z., van Eerden, F. J., de Vries, A. H. & Marrink, S. J. Martini force field parameters for glycolipids. *J. Chem. Theory Comput.* **9**, 1694–1708 (2013).
52. Martini Coarse Grain Forcefield for Biomolecules. http://cgmartini.nl/. (2021).
53. Abraham, M. J. et al. Gromacs: high performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1–2**, 19–25 (2015).
54. Wassenaar, T. A., Ingólfsson, H. I., Böckmann, R. A., Tieleman, D. P. & Marrink, S. J. Computational lipidomics with insane: a versatile tool for generating custom membranes for molecular simulations. *J. Chem. Theory Comput.* **11**, 2144–2155 (2015).
55. De Jong, D. H., Baoukina, S., Ingólfsson, H. I. & Marrink, S. J. Martini straight: boosting performance using a shorter cutoff and GPUs. *Comput. Phys. Commun.* **199**, 1–7 (2016).
56. Bussi, G., Donadio, D. & Parrinello, M. Canonical sampling through velocity rescaling. *J. Chem. Phys.* **126**, 014101 (2007).
57. Parrinello, M. & Rahman, A. Polymorphic transitions in single crystals: a new molecular dynamics method. *J. Appl. Phys.* **52**, 7182–7190 (1981).
58. Michaud-Agrawal, N., Denning, E. J., Woolf, T. B. & Beckstein, O. MDAnalysis: a toolkit for the analysis of molecular dynamics simulations. *J. Comput. Chem.* **32**, 2319–2327 (2011).
59. Pérez, F. & Granger, B. E. *IPython: A System for Interactive Scientific Computing*. www.python.org. (2007).
60. Harris, C. R. et al. Array programming with NumPy. *Nature* **585**, 357–362 (2020).
61. Virtanen, P. et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat. Methods* **17**, 261–272 (2020).
62. Fabian, P. et al. Scikit-learn: machine learning in Python. *J. Machine Learn. Res.* **12** http://scikit-learn.sourceforge.net. (2011).
63. Hunter, J. D. Matplotlib: a 2D graphics environment. *Comput. Sci. Eng.* **9**, 90–95 (2007).
64. Salfer, M., Collado, J. F., Baumeister, W., FernándezBusnadiego, R. & Martínez-Sánchez, A. Reliable estimation of membrane curvature for cryo-electron tomography. *PLoS Comput. Biol.* **16**, e1007962 (2020).
65. Humphrey, W., Dalke, A. & Schulten, K. VMD: visual molecular dynamics. *J. Mol. Graph.* **14**, 33–38 (1996).

Acknowledgements

We thank the group of SJ Marrink's, at the University of Groningen, for making available atomistic simulation trajectories for PI(4,5)P_2 re-parameterization. L.B.A. thanks Fundação para a Ciência e a Tecnologia – Ministério da Ciência, Tecnologia e Ensino Superior (FCT-MCTES, Portugal) for the PhD fellowship PD/BD/137492/2018. M.M.D and M.N.M. thank FCT-MCTES for the "scientific employment stimulus" programs CEECIND/02961/2017 and CEECIND/04124/2017, respectively. This work is financed by national funds from FCT - Fundação para a Ciência e a Tecnologia, I.P., in the scope of the project UIDB/04565/2020 and UIDP/04565/2020 of the Research Unit Institute for Bioengineering and Biosciences - iBB and the project LA/P/0140/2020 of the Associate Laboratory Institute for Health and Bioeconomy - i4HB. We also acknowledge funding by FCT-MCTES projects IF/00386/2015/CP1272/CT0010 and by the European Regional Development Fund (FEDER), through the Regional Operational Programme of Lisbon (PORLISBOA 2020) and the Competitiveness and Internationalization Operational Programme (COMPETE 2020) of the Portugal 2020 framework (LISBOA-01-0145-FEDER-031057 and PTDC/BTM-SAL/31057/2017).

Author contributions

L.B.A. and M.N.M. designed the simulation systems and L.B.A. performed them. L.B.A., F.F., M.M.D., and N.C.S. designed the experimental measurements. L.B.A., M.M.D., and A.F. performed them. All authors contributed to manuscript writing and revision and have given approval to the final version of the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s42004-021-00603-1.

Correspondence and requests for materials should be addressed to Fábio Fernandes.

Peer review information *Communications Chemistry* thanks Urzula Golebiewska and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2021

4 | Development, validation and implementation of a PI(4,5)P_2 model for the Martini 3 coarse grain force field

Article III: Improved parameterization of phosphatidylinositide lipid headgroups for the Martini 3 coarse grain force field

JCTC Journal of Chemical Theory and Computation

pubs.acs.org/JCTC Article

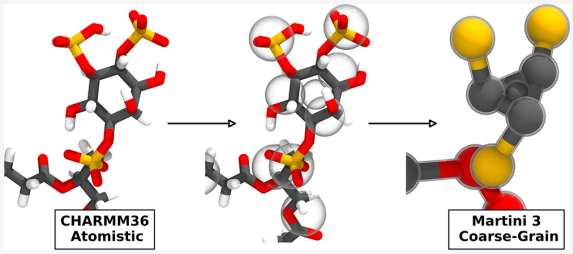
Improved Parameterization of Phosphatidylinositide Lipid Headgroups for the Martini 3 Coarse-Grain Force Field

Luís Borges-Araújo, Paulo C. T. Souza, Fábio Fernandes, and Manuel N. Melo*

Cite This: https://doi.org/10.1021/acs.jctc.1c00615 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Phosphoinositides are a family of membrane phospholipids that play crucial roles in membrane regulatory events. As such, these lipids are often a key part of molecular dynamics simulation studies of biological membranes, in particular of those employing coarse-grain models because of the potential long times and sizes of the involved membrane processes. Version 3 of the widely used Martini coarse-grain force field has been recently published, greatly refining many aspects of biomolecular interactions. In order to properly use it for lipid membrane simulations with phosphoinositides, we put forth the Martini 3-specific parameterization of inositol, phosphatidylinositol, and seven physiologically relevant phosphorylated derivatives of phosphatidylinositol. Compared to parameterizations for earlier Martini versions, focus was put on a more accurate reproduction of the behavior seen in both atomistic simulations and experimental studies, including the signaling-relevant phosphoinositide interaction with divalent cations. The models that we develop improve upon the conformational dynamics of phosphoinositides in the Martini force field and provide stable topologies at typical Martini time steps. They are able to reproduce experimentally known protein-binding poses as well as phosphoinositide aggregation tendencies. The latter was tested both in the presence and absence of calcium and included correct behavior of PI(4,5)P_2 calcium-induced clusters, which can be of relevance for regulation.



Downloaded via Lu&#65533;&#65533;s Borges-Araujo on December 30, 2021 at 14:12:11 (UTC).
See https://pubs.acs.org/sharingguidelines for options on how to legitimately share published articles.

INTRODUCTION

Phosphoinositides are a small group of glycerophospholipids, derived from the reversible phosphorylation of phosphatidylinositol (PI), that account for around 10 to 15% of the total membrane phospholipid content of eukaryotic cells.[1] While these lipids are minor components of eukaryotic biomembranes, research over the last couple of decades has established their role as major regulators of cell dynamics and signaling in eukaryotic metabolism. Their impact extends far beyond the cell membrane, where they take part in regulating several processes such as endocytosis/exocytosis,[1−5] ion channel regulation,[6−8] cellular signaling,[1,2] or cytoskeleton reorganization,[1,2,9] impacting downstream processes such as cellular proliferation, migration, survival, and differentiation.[10] Additionally, as it often happens with ubiquitous signaling entities, disorders affecting their metabolism are responsible for several human conditions,[7] which can range from neurological pathologies[7,11] (such as Charcot−Marie−Tooth neuropathy) to cancer.[12,13]

The core structure of every phosphoinositide lipid is inherited from PI. PI is composed of a diacylglycerol connected to a distinctive *myo*-inositol ring by a phosphodiester bond. While there are 9 possible isomers of inositol, *myo*-inositol is the isomer most commonly found in eukaryotic cells.[14] It assumes a chair conformation where every hydroxyl substituent is at the equatorial position, except for the hydroxyl group in the position 2 of the ring (Figure 1). In forming PI, the hydroxyl group in position 1 of *myo*-inositol takes part in the phosphodiester bond with the diacylglycerol backbone. The hydroxyl groups in positions 3, 4, and 5 of the PI ring can be enzymatically phosphorylated in every possible combination, yielding the seven phosphorylated phosphoinositide species found in mammalian cells (Figure 1): phosphatidylinositol 3-phosphate (PI(3)P), phosphatidylinositol 4-phosphate (PI(4)P), phosphatidylinositol 5-phosphate (PI(5)P), phosphatidylinositol 3,4-bisphosphate (PI(3,4)P_2), phosphatidylinositol 3,5-bisphosphate (PI(3,5)P_2), phosphatidylinositol 4,5-bisphosphate (PI(4,5)P_2), and phosphatidylinositol 3,4,5-trisphosphate (PI(3,4,5)P_3).

While PI is synthesized in the endoplasmic reticulum and then distributed throughout the cell, each of the phosphorylated species has a distinct subcellular distribution within subsets of membranes as a result of the localization of specific PI kinases and phosphatases that catalyze their local formation

Received: June 21, 2021

ACS Publications © XXXX American Chemical Society

Journal of Chemical Theory and Computation pubs.acs.org/JCTC Article

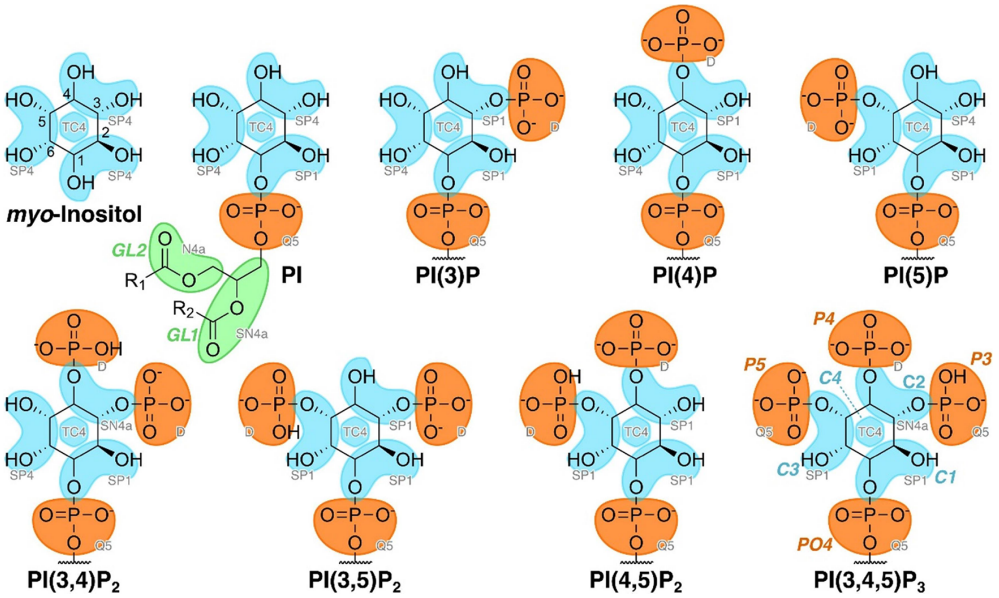


Figure 1. Chemical structures and coarse-grained mapping of the inositol/phosphoinositides parameterized in this work. The inositol ring carbon numbering is overlaid on the *myo*-inositol structure. Atom-to-bead mappings are indicated by the colored shapes, with an extra bead assigned to the center of the inositol ring (ring beads in blue, phosphate beads in orange). For the PI structure, the glycerol and acyl esters are also depicted (R_1 and R_2 abbreviating the sn1 and sn2 chains, respectively), as well as their mapping (in green). Assigned Martini 3 bead types are indicated for each bead as overlaid gray text. For the structure of PI(3,4,5)P_3, in which all possible phosphoinositide beads are present, and for the glycerol/acyl ester groups in PI, bead names are indicated as bold-italic colored text (following the convention in this text that bead names are typed in italics, to distinguish from bead types).

from PI. Apart from their distinct distributions, each phosphoinositide species also establishes distinct interactions with specific binding partners. Phosphoinositides achieve most of their direct signaling effects through interactions between their headgroup and protein partners. While many of them are transient, low-affinity interactions with basic residues within unstructured protein domains, high-affinity specific interactions with specialized domains also occur, some of which are able to distinguish between phosphoinositide isomers.[2] Their distinct subcellular localization, combined with the distinct set of interactions established with cellular binding partners, grants each isomer specific roles and regulatory functions within the cell.

Additionally, some phosphoinositides and PI(4,5)P_2, in particular, are able to establish strong electrostatic interactions with divalent cations (Ca^{2+} and Mg^{2+}), culminating in the cation-dependent aggregation into stable phosphoinositide nanodomains.[15−17] They are formed at physiological concentrations of both lipids and divalent cations, and it is likely that some are found constitutively clustered in vivo. While these nanodomains are often overlooked, they potentially impact phosphoinositide−protein interactions.

The wide-range influence of the phosphoinositide family has piqued research interest, leading to numerous experimental and theoretical research studies.[7] Theoretical studies, such as molecular dynamics (MD) simulations, have the potential to provide molecular-level structural and dynamic insight into complex systems, such as biological membranes, that is otherwise inaccessible through experiments. This gap in knowledge is particularly evident in the case of phosphoinositides, where there is still a considerable lack of understanding about many of the molecular mechanisms behind their function, such as lateral organization, interactions with binding partners, cation-induced nanodomain formation, or induction of curvature.

While MD studies at an all-atom (AA) level of resolution of lipid membrane systems can be performed, they are still limited by current computational power in terms of attainable system size and time scales. Coarse-grained (CG) MD models, however, have allowed access to scales of tens to hundreds of nm and into the millisecond—termed the mesoscale—at which many of the phosphoinositide processes become of relevance.[18] The Martini CG model[19] is the most widely used CG force field for biomolecular simulations,[20,21] and it has been successfully applied in the modeling of a variety of molecular processes, especially in the modeling of biomembrane systems. The large-scale Martini simulation of an average mammalian plasma membrane model[18] stands out as one of the most complex simulation to date. Neuronal plasma membrane systems, where phosphoinositides are thought to play an important role, have also been simulated.[22] Martini applications extend to phosphoinositide−protein interactions, having been able to predict the location of PI(4,5)P_2 binding sites on Kir channels,[23] later confirmed by PI(4,5)P_2-bound protein crystal structures.[24]

Martini is based on a four-to-one mapping scheme, meaning that on average, four heavy (nonhydrogen) atoms, and associated hydrogens, are represented as a single CG particle, or a "bead". Each bead is assigned a type dependent on its underlying chemical nature in terms of polarity and charge. Different interparticle nonbonded potentials are associated to the interactions between each type, effectively encoding a bead's interaction preferences. Martini molecules are then assembled from beads in a building-block fashion, employing

https://doi.org/10.1021/acs.jctc.1c00615
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

2-, 3-, and 4-body bonded potentials (for bonds, angles, and torsions, respectively) to define preferred configurations—matching known behavior from either experiments or from finer simulation sources—and delimiting the overall explorable molecular configuration space.

Within the Martini philosophy, bead type nonbonded potentials have been parameterized to reproduce thermodynamic properties of the represented chemical moieties, in particular the partitioning between polar and nonpolar phases. It is an underlying assumption that, when bonded together, the additive properties of the individual Martini beads can overall represent the properties of the entire modeled molecule. However, in Martini 2, this was found to not hold in specific cases where mappings finer than 4-to-1 were employed, or when intermolecular bead−bead distances were very short. This led to a number of nonobvious pitfalls,[20] the most important being the observations that some molecules become too hydrophobic or self-interact too strongly, in particular proteins[20,25−27] and carbohydrates,[20,28] such as phosphoinositides. Indeed, phosphoinositides included in the pioneering Martini plasma membrane study were found to form dimers and trimers more frequently than expected.[18]

A new version of the Martini force field, Martini 3, has been recently released.[29] This new implementation of the force field addresses the major limitations of Martini 2 by including bead types specific for mappings finer than 4-to-1. The release of the new force field version also greatly increases the discrimination of chemical space by defining a larger number of bead types, allowing for more flexibility and accuracy when creating new models. In addition to expanding the nonbonded landscape, the Martini 3 release also establishes new guidelines for the bonded parameterization of molecules in tandem with the new bead types. In this sense, the new Martini release provides a good opportunity to revisit, update, and improve existing models.

The Martini 2 parameters for phosphatidylinositol and two of its phosphorylated forms were first parameterized by López et al.[30] using the GROMOS53A6 atomistic force field as a reference. In this initial effort, headgroup motion was left mostly unrestrained, with only one torsion potential controlling it, and even this potential ended up being routinely ignored in production simulations[31] as it led to numerical instability at typical Martini time steps.[32] Different improvements on these Martini 2 parameters have been proposed,[18,22] mostly focused on stabilizing bonded behavior. These approaches are still limited by the Martini 2 nonbonded insufficiencies and by an overall excessive freedom of the inositol ring.

In this work, we parameterize inositol and 8 phosphoinositide lipids for Martini 3, using the CHARMM36 atomistic force field as a structural atomistic reference. In addition to compatibility with the rest of the Martini 3 landscape, Martini 3 phosphoinositide models are expected to have more accurate interaction propensities than their Martini 2 counterparts. This parameterization effort is also an opportunity to expand the covered species to include all relevant phosphorylations. Finally, a new take on the parameterization of phosphoinositides can aim at better modeling (i) their complex structural dynamics[33]—a result of large headgroups, networks of intra-/intermolecular hydrogen bonds, and variable phosphate protonations—and (ii) the often overlooked interaction of phosphoinositides with divalent cations, which can drive nanodomain formation of physiological relevance, especially involving the PI(4,5)P_2 headgroup.[15,17]

METHODS

All simulations were run using the GROMACS simulation package version 2019[34] and analyzed making use of in-house developed Python3 programs using the MDAnalysis package.[35] We also used the IPython,[36] numpy,[37] SciPy,[38] scikit-learn,[39] Voro++,[40] and matplotlib[41] packages for scientific computing in Python. Visualization and rendering of the simulations were performed with the molecular graphics viewer VMD.[42] Octanol−water partition free energies were calculated from the individual CG solvation free energies into each solvent, as described elsewhere.[43] See the Methods section of the Supporting Information for details on specific analysis methods.

Atomistic Models. All atomistic models used as the parameterization targets were simulated using the CHARMM36[44,45] force field, with the TIP3P water model. All topologies used are readily available in CHARMM-GUI.[46,47] The topologies used for each of the parameterization targets can be seen in Supporting Information Table S1.

To validate the parameterization of the inositol molecule, simulations were performed of *myo*-inositol in water, either singly or as a solution of eight molecules. For single *myo*-inositol molecule systems, the sugar molecule was inserted in a 5 × 5 × 5 nm simulation box, which was then fully solvated. As for the multiple *myo*-inositol system, eight sugars were placed in a cubic box and solvated at an 8.0 water/sugar (weight/weight) concentration.

For the lipid membrane systems, initial structures were all generated, using the CHARMM-GUI membrane builder module,[48−51] by arranging the lipids on a regular array in the bilayer (*xy*) plane. Membranes were built with roughly 240 lipids per leaflet, which were then solvated by ∼11,000 water molecules. Every system was neutralized and an additional 140 mM NaCl was added to better represent physiological conditions.

After initial energy minimization and equilibration runs, all atomistic simulations were run at a 2 fs time step. The LINCS[52] algorithm was applied to constrain all bonded hydrogens. Van der Waals forces were switched off smoothly from 1.0 to 1.2 nm, and electrostatics were computed using particle-mesh Ewald summation.[53] The particle neighbor list was updated using the Verlet list scheme. The system temperature was maintained at 310 K by coupling to a Nosé−Hoover thermostat[54] with a 1 ps coupling constant, while pressure was coupled to 1.0 bar using a Parrinello−Rahman barostat[55] (isotropically for aqueous systems, semi-isotropically for membrane systems) with a 5 ps coupling time.

All atomistic simulations were run for at least 2 μs, and unless otherwise stated, the last 1 μs was used for analysis.

Coarse-Grained Models. All coarse-grained simulations were modeled using the Martini 3 CG model for biomolecular simulations.[29] All topologies, other than the ones parameterized here, are available alongside the force field.[29] Along this text, bead names will be typed in italics, to distinguish from bead types.

To validate the inositol headgroup, CG simulations were performed of a single *myo*-inositol sugar molecule in water as well as of multiple sugars, as done with the atomistic models.

For the lipid membrane systems, initial structures were built and solvated using the insane.py CG building tool[56] by

https://doi.org/10.1021/acs.jctc.1c00615
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

Journal of Chemical Theory and Computation pubs.acs.org/JCTC Article

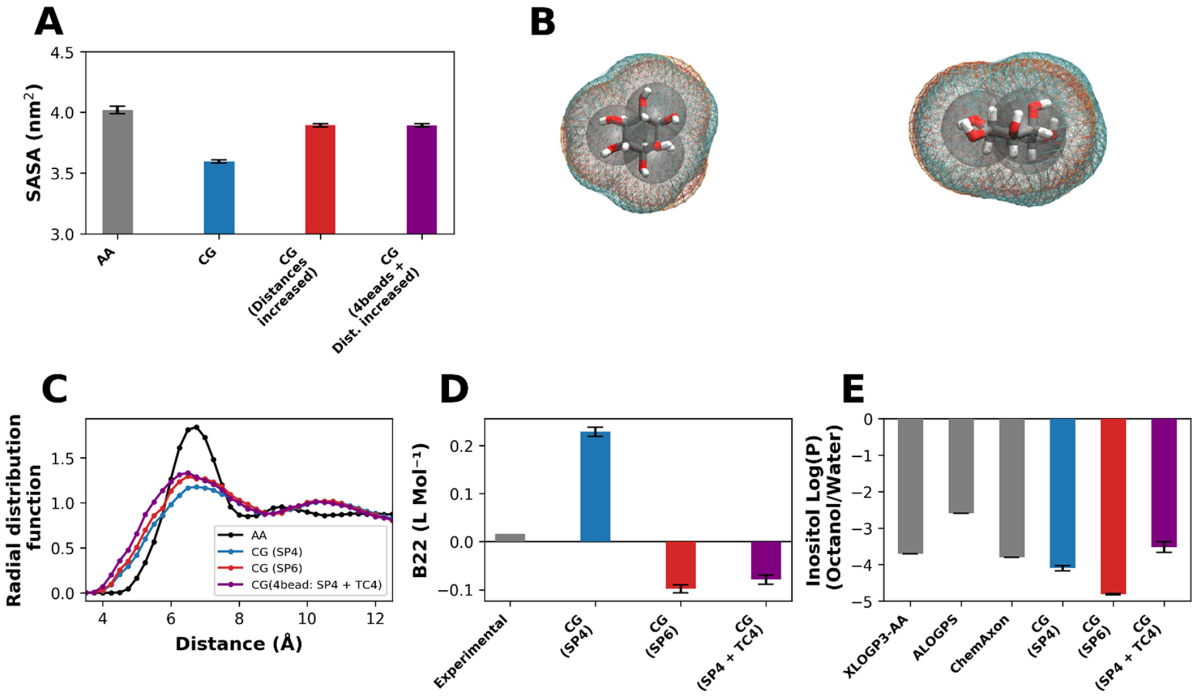


Figure 2. *myo*-Inositol Martini 3 CG model properties. SASA (A), inositol c.o.m. vs inositol c.o.m. RDF (C), B_{22} osmotic coefficient (D), and inositol log(*P*) values (E) for several iterations of the inositol CG model as well as the appropriate references. Error bars for the B_{22} measurements (D) represent the standard error of the mean determined from 16,200 ns measurements over the course of 3.5 μs (Supporting Information Figure S1). Otherwise, when applicable, error bars represent the standard deviation from 3 independent simulation experiments. Connolly surfaces (B) for 2 iterations of the CG inositol model (c.o.g.-mapped bead distance: red wireframe; 20% increased bead distance: orange wireframe) as well as the atomistic reference (blue wireframe).

arranging the lipids on a regular array in the bilayer (*xy*) plane to obtain roughly 180 lipids per leaflet, solvated by roughly 8000 water beads. All CG systems were solvated with regular Martini water beads; while smaller water beads are available, they are still not fully balanced to be used as solvents. Larger systems containing up to 700 lipids per leaflet were also built for characterization of inositide lipid properties and Ca^{2+}-induced inositide clustering. Counterions were added to neutralize the systems as necessary, plus 140 mM NaCl ionic strength. When testing divalent cation aggregation, Ca^{2+} was added to the systems at a 5:1 Ca^{2+}-to-phosphoinositide ratio, by replacing water particles and maintaining the system charge neutral by adding Cl^- counterions.

Two different proteins were also studied. Protein structures were obtained from the Protein Data Bank[57] (PDB) (PLCδ1 PH domain PDB: 1MAI[58] and the Kir2.2 channel PDB: 6M84[59]). All CG protein models were constructed using Martinize2,[60] with an applied elastic network with a bond force constant of 700 kJ/mol and a cutoff distance of 0.8 nm. Side chain corrections were also applied.[61] Both proteins were placed in lipid membrane systems. The CG model of the PH domain was placed in the corner of the simulation box with a random initial orientation. The CG model of the Kir2.2 channel was positioned roughly as described in previous simulation studies.[23,62]

Nonbonded interactions were cut off at 1.1 nm, and Coulombic interactions were treated using reaction-field electrostatics[63] with a dielectric constant of 15 and an infinite reaction-field dielectric constant. The particle neighbor list was updated using the Verlet list scheme. A *v*-rescale thermostat[64] was used with a coupling time of 4.0 ps to maintain the temperature at 310 K (unless other temperatures are specified). Constant pressure was isotropically or semi-isotropically (in the case of membranes) coupled to 1.0 bar using the Parrinello–Rahman barostat[55] with a relaxation time of 16.0 ps. After initial energy minimization and pressure/temperature equilibration runs, simulations were run at a 20 fs time step.

All CG systems were run for at least 10 μs, and unless otherwise stated, the last 2 μs of each simulation was used for subsequent analysis.

All developed models are provided as GROMACS-compatible topology files, as Supporting Information associated to this manuscript. Current and future parameter iterations can be tracked in the associated GitHub repository at https://github.com/MeloLab/PhosphoinositideParameters.

RESULTS AND DISCUSSION

Our aim was to develop phosphoinositide parameters up to date with the new Martini 3 framework, also taking into account important phosphoinositide biophysical properties not previously accounted for, such as protonation and interaction with divalent cations. An overview of the mappings used in this work is given in Figure 1. While previous work on Martini 2 phosphoinositides served as inspiration for some of our parameterization strategies, the Martini 3 models in this work were developed from independently tuned parameters, following the parameterization guidelines for the new Martini 3

https://doi.org/10.1021/acs.jctc.1c00615
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

force field.[29] These guidelines call for using center of geometry (c.o.g.)-based mapping of atomistic structures, taking into account the respective hydrogen atoms. CG parameters were fit to corresponding atomistic simulations ran with the CHARMM36 force field. In addition, CHARMM36 was proven to be an excellent model for the simulation of phospholipid biological membranes, outperforming competing force fields in terms of experimental data reproducibility.[65] Additionally, the parameterization of the phosphoinositide models for CHARMM36 has been thoroughly described,[66,67] and these parameters have been extensively used for phosphoinositide atomistic simulation studies, both in the presence and absence of divalent cations.[68−71]

With this in mind, we first revisited the parameterization of the core molecule in all phosphoinositide headgroups, the *myo*-inositol sugar.

***myo*-Inositol Sugar Core Parameterization.** The bead mapping behind the CG *myo*-inositol model followed the Martini 3 rules for 4−1 cyclic/branched fragments, assigning an S-bead for each 2 consecutive ring carbon atoms and their respective hydroxyl groups (Figure 1, *myo*-inositol). In our model, carbons 1 and 2, 6 and 5, and 4 and 3 correspond to the CG beads *C1*, *C2*, and *C3*, respectively. This contrasts with the inositol model developed for Martini 2,[30] in which the ring was rotated by 1 inositol carbon, meaning that carbons 6 and 1, 4 and 5, and 3 and 2 represented CG beads *C1*, *C2* and *C3*, respectively. While both mappings are valid within the Martini guidelines, our mapping should perform better in PI(4,5)P_2 calcium-induced aggregation experiments, as discussed ahead.

To model the bonded parameters for our CG representation, AA simulations of an inositol sugar molecule in water were used as a reference (Supporting Information Figure S2), with initial bonded parameters obtained from the c.o.g.-based mapping of atomistic structures, as mentioned above. To confirm that the c.o.g.-based mapping accurately reflects the atomistic configurations, we compared CG and AA solvent accessible surface areas (SASA). This indicates whether the molecular volume and shape of the CG inositol molecule is representative of the corresponding AA structure. The average SASA values recovered for the single inositol in water systems show a discrepancy of roughly 13% between the initial CG model and the atomistic parameterization target (Figure 2A). While the current version of Martini 3 systematically underestimates the SASA,[29] a 13% discrepancy falls beyond a target 10% error margin. Analyzing the Connolly surfaces (Figure 2A), we can see that the CG model (red) is not capturing the behavior of the AA model (blue) accurately enough. This is in part because the hexagon-like inositol ring is now modeled by a triangular arrangement of CG beads, creating a molecular shape mismatch. Additionally, as the hydroxyl groups on the *myo*-inositol ring are mostly equatorial, they increase the surface area significantly. To improve the model, the CG interparticle distances in the triangular arrangement of inositol were lengthened by 20%. This reduced the discrepancy between the atomistic and CG model to an acceptable 5%, as seen by the recovered SASA values as well as the corresponding Connolly surface (Figure 2A), and was the bonded arrangement used in subsequent inositol parameterization.

***myo*-Inositol Bead Type Assignment.** To assist in our selection of particle types for the *myo*-inositol model, we set up its octanol/water partition as a target. As a starting point for inositol, beads types ranging from SP4 to SP6 (from less to more polar) were chosen from existing Martini 3 examples representing ethylene glycol moieties. These beads types also roughly correspond to those used by the original Martini 2 inositol model.[30] We ran several octanol/water partition simulations of inositols composed of each bead type and compared the results against theoretical predictions by XLOGP3-AA,[72] ALOGPS,[73−75] and the ChemAxon Consensus method[76,77] (Figure 1E). Experimental values could not be found for *myo*-inositol, possibly because of its strong hydrophilicity, which hampers the use of standard log(P) determination methods. It is worth noting that while these computational predictions can be quite accurate, at extreme hydrophilicities/hydrophobicities, there can often be large disagreements between methods and between predictions and experiments. The theoretical log(P) predictors, XLOGP3-AA and ChemAxon, yielded log(P) values of −3.7 and −3.8, respectively, reflecting the strongly hydrophilic character of the *myo*-inositol ring. ALOGPS yielded a somewhat less hydrophilic log(P) value of −2.6. The initially tested CG inositol compositions (all SP4, all SP5, or all SP6 beads) yielded log(P) values between −4 and −5, significantly more hydrophilic than any of the predictions. The bead choice closest to the predictions was SP4, the least polar tested type, at a log(P) of −4.15. Use of Martini 3 types less polar than SP4, which do not typically represent small diols, was avoided since it may compromise the resulting transferability within the Martini 3 framework.

As a further characterization of the suitability of assigned bead types, we probed inositol self-interaction in solution, comparing CG systems to their atomistic CHARMM36 counterpart in terms of center-of-mass (c.o.m.) inositol−inositol radial distribution functions (RDFs; Figure 2C). In these RDFs, the position of the first CG neighbor matches that of the AA reference. The CG behavior differs, however, in having offset second-neighbor peaks, which was already a feature of the Martini 2 version of the model.[30] The Martini 3 models also differ from the AA reference in having a less pronounced first-neighbor binding. In this aspect, the more polar type (SP6) performs somewhat better.

Given concerns that atomistic models may overestimate self-interaction in sugar molecules,[78] the second osmotic virial coefficient (B_{22}) was also determined and compared to experimental values. The B_{22} coefficient can be used to estimate the self-interaction of sugar molecules, where $B_{22} < 0$ indicates net attraction and $B_{22} > 0$ indicates repulsion between molecules. A B_{22} value of 0.016 mol·L^{-1} has been determined for *myo*-inositol[79] (Figure 2D). CG inositol molecules (Figure 2D) composed of SP4 beads yielded a B_{22} value of 0.228 mol·L^{-1}, indicating a more repulsive behavior than expected. The more polar SP6 beads, however, yielded a value of −0.098 mol·L^{-1}, indicating a slightly more attractive behavior than the experimental one. Compared to previous Martini 2 models, all Martini 3 models tested here yielded B_{22} values much closer to the experimental range. Martini 2 models of similar sugars have typically overestimated their aggregation propensity, recovering values that differ from the experimental values by over one order of magnitude (−1.2 to −2.0 mol·L^{-1}).[28]

Ring-Center Apolar Interaction Site. To improve our model, we added an additional Martini 3 tiny-type (T) particle to the center of the inositol ring—bead *C4*—representing the hydrophobic sugar ring core. A TC4 particle was added as a virtual interaction site, constructed in the plane as a linear combination of the ring particle positions. With this strategy,

https://doi.org/10.1021/acs.jctc.1c00615
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

the inositol model consisting of the SP4 ring particles and the hydrophobic TC4 virtual site yielded the best matching of the atomistic and experimental behavior in all our tests. While it did not influence the SASA (as the additional particle is buried in the middle of the ring), it led to log(P) values in much better agreement with theoretical predictors (log(P) = −3.74), drastically improved inositol self-interaction to near atomistic levels (B_{22} = −0.079 mol·L^{-1}), and better approximated the atomistic first-neighbor behavior in Figure 2C. This *myo*-inositol model was chosen as the basis for subsequent phosphoinositide parameterization.

In addition to improving the performance of our model, the added T particle helps in avoiding underestimation of the molecular interaction density. Although S particles are called for when mapping 4−1 branched or cyclic fragments, they must not be overused. As described in the Martini 3 parameterization guidelines,[29] the maximum mismatch should be 1 nonhydrogen atom for each 10 nonhydrogen atoms mapped by CG beads. *myo*-Inositol has a total of 12 nonhydrogen atoms. Given that a single S particle represents 3 CG-mapped nonhydrogen atoms, the 3 S particles used to map the inositol ring only account for 9 nonhydrogen atoms. The addition of a T bead in the center of the ring reduces this mismatch and helps avoid interaction density issues.

Phosphoinositide Parameterization. Having finalized the *myo*-inositol sugar parameterization, we moved on to the phosphoinositide family. Phosphoinositides were parameterized against reference AA counterparts modeled with CHARMM36. The atomistic parameterization target systems for our lipid models were bilayers containing roughly 250 lipids and composed of 95% POPC and 5% of the phosphoinositide of interest. This system was chosen over the more common bilayer system containing 100% of the lipid of interest as it better replicates the phosphoinositide physiological and experimental environments: all phosphoinositides are minority membrane lipids, and most cannot even form stable bilayers at high mole fractions. As a control for lipid overparameterization, a test system composed by a single lipid of interest in a water box is also used. The inclusion of this system reduces parameterization bias for specific membrane-only behavior.

Phosphatidylinositol. The parameterization strategy for PI (Supporting Information Figure S20) was (i) to bind the inositol model to the phosphodiester phosphate bead from the standard Martini 3 phosphoacylglycerol topology (using 1-palmitoyl-2-oleoyl lipid tail parameters) and then (ii) to control the headgroup tilt and rotation by means of appropriate angle and dihedral potentials. All of the inositol headgroup particle types were inherited from the standalone topology, except for the *C1* bead, which was changed from SP4 to SP1 to reflect the loss of polarity because one of its hydroxyls now takes part in the phosphodiester bond with the phosphate in the *PO4* bead.

It should be noted that the used phosphoacylglycerol topological basis may be subject to optimization in the future, as the full potential offered by Martini 3 is realized. Our parameterization takes that into account and, for added robustness, only involves headgroup bonded terms with the phosphate and glycerol moieties. This also ensures optimal transferability across Martini 3 phosphoglycerides, which differ only in acyl chain composition.

Regarding the bonded parameters, distances between the headgroup inositol ring beads were adjusted as necessary to account for slight ring torsions originating from intramolecular hydrogen bonds when inserted onto the lipid diacyl glycerol moiety. To control the headgroup tilt, an angle potential and 2 bonds were used. The *C2−C1−PO4* angle was used to control headgroup tilt with respect to the *PO4* bead, while *PO4-GL1* and *PO4-GL2* bonds were used to control the *PO4* tilt with respect to the glycerol backbone. This pair of bonds was used in place of the more common *PO4−GL1−GL2* angle potential as it allowed us to avoid repulsion between the *PO4* and *GL2* particles, thus better representing the atomistic description. With these parameters in place, we were able to control the headgroup tilt and, importantly, the angle between the inositol headgroup and the membrane surface. To control headgroup rotation, the *C3−C2−C1−PO4* and *C2−C1−PO4−GL1* dihedral potentials were put in place. To avoid numerical instabilities and thus allow for longer time steps, restricted bending potentials were used in place of the typically used cosine-harmonic angle potentials[32] when angles that take part in dihedral potentials had distributions close to 0 or 180 degrees. The remaining lipid parameters are acyl chain-dependent and strictly follow the current Martini 3 lipid parameters. Our approach closely represented the distributions obtained from CG-mapped AA lipid bilayers and lipid-in-solution systems (Supporting Information Figure S3).

Phosphatidylinositol Phosphates. With the core PI lipid fully parameterized, we were able to use it as a building block for the phosphorylated species. Our strategy to add phosphorylations to the core PI lipid was to use a single bond to the inositol ring, one angle potential and one dihedral potential to control the phosphate position relative to the ring plane. When necessary, a weak secondary bond to the ring was employed, to overcome intramolecular repulsions. This was especially useful when parameterizing the polyphosphorylated inositides, where the phosphate beads lie closer to the inositol rings than in the monophosphorylated inositides (this can be seen for PI(3,5)P_2 through the smaller *P3−C3−C1* angle compared to the same angle obtained for PI(3)P, Figures S4 and S8). For each phosphate, the main bond was set matching the respective CG-mapped AA distance. When used, the weak secondary bond could be adjusted slightly to allow for the correct angle between ring−phosphate and ring−ring bonds.

In the case of polyphosphorylated inositides, there were instances where two phosphates lay within the interaction range of each other. This was often the result of transient interactions between adjacent phosphate groups, mediated by the last phosphate proton. To avoid repulsion and accurately represent the target behavior, a weak bond was put in place between them. This was relevant in lipids with adjacent phosphorylations, such as PI(3,4)P_2, PI(4,5)P_2, and PI(3,4,5)-P_3.

As was the case with the phosphodiester, the beads that took part in the phosphomonoester bonds had their polarity reduced from SP4 to SP1 (or from SP1 to SN4a in the case of two phosphomonoester groups bonded to the same bead) to reflect the loss of polarity from the hydroxyl group that is now taking part in the bond.

The addition of phosphate groups induced significant ring torsions and changes to the headgroup tilt angle due to the formation of intramolecular hydrogen bonds with the ring hydroxyls and phosphodiester oxygens.[33] The bonds between the inositol headgroup beads, as well as the angle and dihedral potentials in place, were slightly adjusted for each lipid to reflect these dynamics. Our approach correctly represented the distributions obtained from the CG-mapped AA lipid bilayers

https://doi.org/10.1021/acs.jctc.1c00615
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

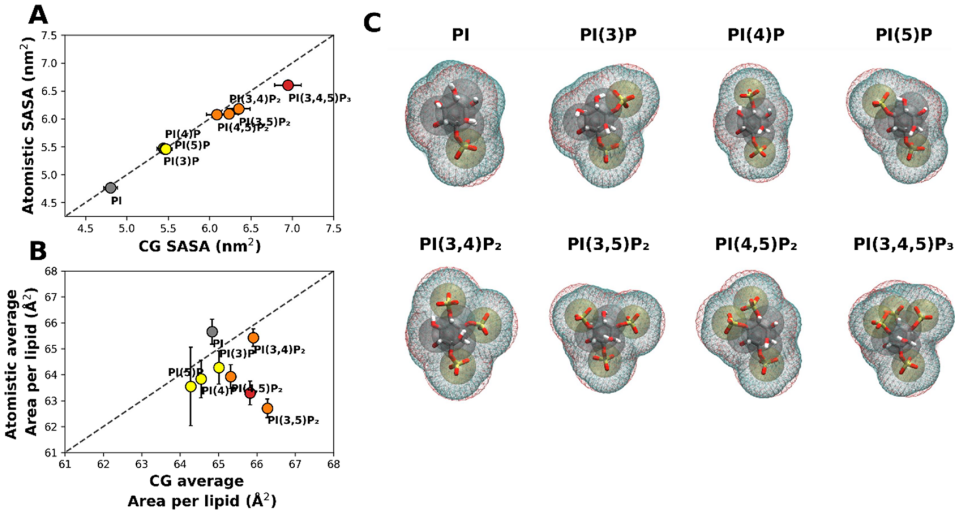


Figure 3. Phosphoinositide lipid headgroup SASA, Connolly surfaces, and ApL. Comparison of phosphoinositide headgroup SASA values (A) and ApL values (B) recovered for the Martini 3.0 and AA models. Colors indicate lipid phosphorylation (gray: no headgroup phosphates; yellow: monophosphorylated; orange: bis-phosphorylated; red: tris-phosphorylated). The $y = x$ relation was drawn as a guide. Connolly surfaces (C) were determined from AA and CG-mapped structures, which were then overlayed on top of each other. The AA surface is represented in blue and the CG surface in red.

and lipid-in-solution systems (Supporting Information Figures S4−S10).

Headgroup Phosphate Protonation State. The protonation state of the phosphodiester group is well-defined, being fully deprotonated at physiological pH (pK_a between 1 and 3),[80] and a −1 charge was assigned to it. For phosphatidylinositide monophosphates, such as PI(4)P ($pK_a(P4)$ = 6.2), the monoprotonated form exists at physiological pH, but the fully deprotonated form is still the dominant population.[81] Headgroup phosphate beads of the monophosphorylated lipids were therefore assigned −2 charges.

The protonation behavior of phosphatidylinositol bisphosphates, however, is much more complex. It has been shown that the ionization behavior of phosphoinositides with adjacent phosphate groups cannot be accurately described by a Henderson−Hasselbalch mechanism.[82] Instead, at physiological pH, one proton from the adjacent phosphate group dissociates, while the remaining one is stabilized by being shared between the two vicinal phosphomonoester groups.[82] To add to this complexity, it is likely that a small, fully deprotonated population also exists.[82] With this in mind, the headgroup phosphate beads of the double-phosphorylated lipids with adjacent phosphates (PI(4,5)P_2 and PI(3,4)P_2) were set at a charge of −1.5 each, reflecting the most common charge state (1 phosphate group fully deprotonated with the last proton group being mostly shared between the 2 groups) when monodisperse and not interacting with other partners. As for the only double-phosphorylated lipid that does not have adjacent phosphates, PI(3,5)P_2, each of its phosphates presents a mostly well-behaved Henderson−Hasselbach mechanism.[82] At physiological pH, each of its phosphates is split almost evenly into populations of monoprotonated and fully deprotonated species. For this reason, each phosphate bead was set at a charge of −1.5 as well, reflecting the most common overall charge state.

PI(3,4,5)P_3, where the three phosphorylations are adjacent, is also a complex case. Its behavior is similar to that of the phosphatidylinositol bisphosphate lipids, with protons being shared between adjacent phosphates.[82,83] As such, the most common charge state at physiological pH is to have 2 phosphate groups fully deprotonated, with the third phosphate group monoprotonated and sharing its last proton with the adjacent phosphate groups. With this in mind, the *P3* and *P5* phosphate beads of our CG model of PI(3,4,5)P_3 were set at a charge of −1.3, and the *P4* bead was set at −1.4 to reflect this behavior. The slightly higher charge for the *P4* phosphate is in agreement with experimental observations.[82,83]

Phosphate Particle Types. The −1-charged phosphodiester phosphate was assigned the bead type Q5, in line with its charge and with the rest of Martini 3 phospholipid headgroups. For the −2-charged phosphate particles, regular-size D particles—added in Martini 3[29] for divalent hard ions—were used. The −1.5-charged phosphate particles, in PI(3,4)P_2, PI(3,5)P_2, and PI(4,5)P_2, were modeled as D particles as well, as the last proton is very weakly bonded to the phosphates, conferring significant divalent character. For PI(3,4,5)P_3, however, the *P3* and *P5* phosphates were modeled as Q5 beads, while the *P4* phosphate was modeled as a D bead. It is worth noting that beads with noninteger charges more anionic than −1 are a parameterization edge case that was not explicitly included in the Martini 3 model. However, Martini 3 was designed to allow the assignment of partial charges when representing charge delocalization over several beads—a situation somewhat analogous to proton sharing between adjacent phosphates.

Apart from the particle types and charges in our CG model, protonation was also taken into consideration when building atomistic references. For the monophosphorylated lipids, we opted to use fully deprotonated lipids as our target structural reference. In the case of the bis- and tris-phosphorylated species, however, atomistic systems contained semiprotonated lipids, where one of the headgroup phosphates is monoprotonated and the other headgroup phosphates are fully deprotonated, in all possible combinations. Our parameter-

https://doi.org/10.1021/acs.jctc.1c00615
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

Journal of Chemical Theory and Computation pubs.acs.org/JCTC Article

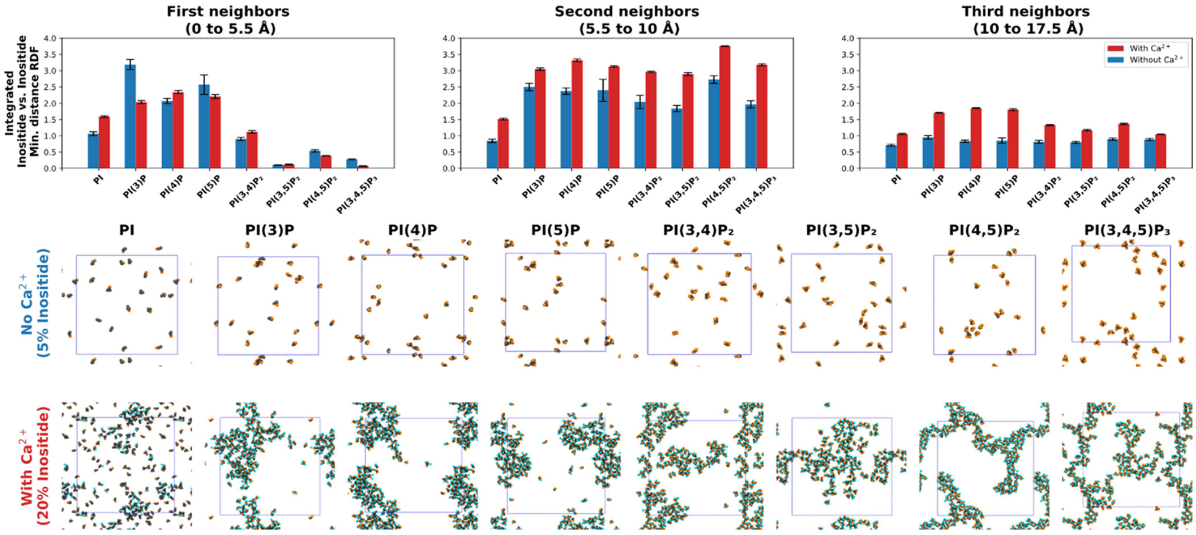


Figure 4. Phosphoinositide aggregation in response to calcium. Integrated inositide vs inositide minimum distance RDFs corresponding to the first, second, and third local neighbors, for the 8 inositide lipids. They were obtained from systems in the presence (red) and absence (blue) of calcium at 300 K. Systems in the presence of calcium were built at a 20:80 inositide:POPC ratio, mimicking high inositide concentrations in the local vicinity of phosphoinositide domains, and Ca^{2+} was added at a 2.5:1 Ca^{2+}:inositide ratio. Systems in the absence of calcium were built at a 5:95 inositide ratio, mimicking monodisperse inositide conditions. The full RDFs are included in Supporting Information Figure S14. Final simulation snapshots of the systems are also shown, viewing the top leaflet along the membrane normal, with phosphoinositide lipid headgroups depicted in gray, phosphates in orange, and Ca^{2+} ions in blue.

izations were then left as flexible as possible to allow the lipid to visit all possible conformations, allowing for a better representation of the flexibility and variety of conformations that occur in vivo at physiological pH. While this falls outside the scope of this work, a possible future enhancement would be to leverage the multiple phosphate protonation states, allowing for simulation of phosphoinositides under constant-pH conditions, with methods such as those of Baptista et al.[84] or Grünewald et al.[85] The phosphoinositide models developed in this work would serve as a great starting point; however, significantly more adaptation would be required before they could be used with these methods.

VALIDATION

Membrane Behavior. Having fully parameterized the phosphatidylinositol lipid family, we then aimed at confirming that the CG lipids accurately represent expected biophysical properties. We started by assessing whether the c.o.g-mapping correctly reflected the atomistic headgroup volumetry, SASA-wise, as was done for *myo*-inositol. SASA comparisons for each lipid headgroup (from, and including, the phosphodiester phosphate) and their respective atomistic descriptions are shown in Figure 3A. As phosphorylation increases, and with it the headgroup size, there is a clear increase in SASA. This is well-recovered by our CG models. Less phosphorylated species, like PI, slightly underestimate the SASA in the inositol ring region (Martini 3 tends to underestimate atomistic SASA values[29]). This is compensated for by each additional phosphorylation, in that the added individual SASA is slightly larger than that of atomistic phosphates (Figure 3C). As such, the bis- and tris-phosphorylated inositides end up having a slightly higher SASA than the atomistic model, well within the 10% accepted margin and all but PI(3,4,5)P_3 within 5% (Figure 3A and Supporting Information Figure S11). For PI, but not for the other phosphoinositides, ring distances were increased to recover the correct SASA behavior, as had been the case for the analogous *myo*-inositol.

To confirm lipid insertion depth and headgroup surface protrusion when inserted into POPC bilayers, we compared the membrane z axis densities between CG systems and AA references (Supporting Information Figure S13). It can be seen that all lipids' phosphodiester group inserts at the same depth as the POPC phosphodiester group. This not only matches the atomistic reference but also experimental evidence that indicates that this is the preferred depth for PI and PI(4)P.[86−89] *myo*-Inositol rings are slightly more protruded (up to 1 Å) than in the AA reference, whereas individual headgroup phosphates can protrude up to 2 Å more. Since phosphate−ring distances, as well as most of the bonded terms influencing the inositol tilt relative to the glycerol backbone, are well-reproduced by our models (Supporting Information Figures S3−S10), the excessive protrusions may instead result from an insufficient headgroup-glycerol tilt relative to the lipids' acyl tails (Supporting Information Figure S22). Because we aim for our inositol models to be widely compatible with different tails, in the general scope of Martini 3 lipids, we refrained from introducing specific corrections in this regard.

To quantify the membrane density of each lipid, the area per lipid (ApL) was determined (Figure 3B). Interestingly, unlike the SASA, the ApL measured in atomistic references does not show a clear dependency on headgroup size or phosphorylation. This showcases how phosphoinositides, in general, are able to maintain an ApL similar to that of POPC despite a much larger headgroup. Our CG models were able to closely reproduce the atomistic ApL values within a 5% discrepancy (Supporting Information Figure S12). PI(3,5)P_2 shows the highest discrepancy out of all the lipids. We hypothesize that this is again due to insufficient headgroup tilt (Supporting

https://doi.org/10.1021/acs.jctc.1c00615
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

Information Figure S22), which plays an important role in the inositide ApL. Although our work was done to maintain most of the characteristic headgroup tilt through the bonded parameters, we could not fully replicate this behavior. This is likely caused by a decreased range of motion inherent to the coarseness of the Martini force field, not only impacting the inositide headgroup tilt but also stiffening the bulk POPC headgroup environment, which hinders inositide headgroup mobility. While this should impact all inositide lipids, it appears to be particularly relevant for PI(3,5)P_2 and PI(3,4,5)-P_3, which show a clear bimodal headgroup tilt in the AA reference (Supporting Information Figure S22). Nonetheless, in the broader context of membrane lipids, the ∼6% discrepancy for PI(3,5)P_2 is still well within an acceptable range. It is also worth noting that it is not fully clear how accurate the ApL of the CHARMM inositide models actually is. In this sense, the atomistic reference ApL provides guidance and quality control but is not a parameterization target as strict as the reproduction of SASA values or bonded distributions.

From a dynamic perspective, our phosphoinositide models diffuse quite faster than the experimentally measured (a PI(4,5)P_2 fluorescent analogue (TopFluor PI(4,5)P_2) diffusion coefficient[15] of $\sim 8 \times 10^{-8}$ $cm^2 \cdot s^{-1}$ and a Martini 3 PI(4,5)P_2 diffusion coefficient of 2.7×10^{-5} $cm^2 \cdot s^{-1}$, Supporting Information Figure S25). Such a speedup is expected for the Martini model, where the discrepancy has been seen to be larger for charged molecules,[90] which presumably face reduced friction due to the simplified Martini electrostatic representation.

Altogether, our analyses show that the membrane properties of the phosphoinositides parameterized in this work are represented within the fidelity expected from the Martini 3 force field.

Phosphoinositide Aggregation in Response to Divalent Cations. A significant portion of phosphoinositide lateral organization is dependent on the interaction of these lipids with divalent cations. Several phosphoinositides are known to undergo cation-dependent aggregation, forming stable lipid aggregates. They are formed at physiological concentrations of lipids and cations for some of the lipids in the phosphoinositide family, and it is likely that some of them are found constitutively clustered in vivo.[15−17] As such, it is of importance that our models accurately represent this aggregation behavior.

We tested the response of each phosphoinositide to the presence of calcium at a 2.5:1 Ca^{2+}:phosphoinositide ratio (Figure 4). These systems were built at a 20:80 inositide:-POPC ratio to mimic the exceptionally high inositide concentrations in the local vicinity of phosphoinositide domains. In the systems where calcium is included, phosphoinositide headgroup phosphates are assigned a charge of −2, and their bead type was changed to D, when applicable. This mimics the displacement of the last phosphate proton when calcium binds phosphoinositide headgroups. To test the phosphoinositide aggregation in the absence of calcium, we employed membranes at a 5:95 inositide:POPC ratio, mimicking monodisperse phosphoinositide conditions.

The analysis of inositide vs inositide RDFs obtained from the minimum distance between headgroups, combined with the visual analysis of the structures obtained, proved useful in characterizing and comparing the domains formed by each inositide species. In general, the first-neighbor peak in these RDFs represents direct inositide−inositide contacts. The second-neighbor peaks represent inositide−inositide−inositide contacts, but, in the case of systems with calcium, they mostly represent calcium-linked inositide dimers (i.e., inositide−Ca^{2+}−inositide contacts). Further peaks represent larger contact chains.

In the absence of divalent cations, none of the parameterized PI lipids self-aggregate (Figure 4). This is in agreement with most of the experimental findings that point at PIs being monodisperse in fluid bilayers.[15,17,91] The monophosphorylated inositides showed the most first-neighbor contacts compared to the bis- and tris-phosphorylated lipids. This hints that the additional bulky phosphates and associated negative charge in polyphosphorylated inositides either increase the repulsion between the monodisperse lipids or lead to more inefficient stacking of the headgroups, or both. Interestingly, the monophosphorylated inositides also showed increased first-neighbor contacts when compared to PI, showing that the added repulsion of the single headgroup phosphate can be compensated for by the establishment of additional interactions. Additionally, no considerable concentration-dependent phosphoinositide aggregation was observed, even at a 20% lipid concentration (Supporting Information Figure S21), where no substantial structured lipid aggregation takes place, despite some localized accumulation.

In the presence of calcium, however, all inositol lipids except for PI quickly undergo very significant cation-induced aggregation. This can be clearly seen, both through the final simulation snapshots as well as through the integrated RDFs (Figure 4). It is worth noting that running these systems with PME electrostatic treatment did not impact the formation of calcium-induced clusters (Supporting Information Figure S24A,B). While PI did not form calcium-induced clusters, it induced some adsorption of calcium at the phosphodiester depth, which led to the formation of PI dense areas, with no defined structure. This can be clearly seen in the simulation snapshots and integrated RDF peaks (Figure 4).

We see no significant calcium-induced increase in the density of the first neighbors among the phosphorylated lipids, but we see a large increase in density of the second and third neighbors, for all inositides. PI(4,5)P_2 formed the densest calcium-induced nanodomains, with barely any POPC mixed in with the inositide. This can be clearly seen in the increased density of the second neighbors. These domains are irregularly shaped, with a lobed morphology and a fairly organized stacking of the inositol headgroups (Supporting Information Figure S17).

The monophosphorylated inositides formed less dense, irregular domains that contained a significant amount of POPC lipids within. These domains also showed much more irregular stacking of the headgroups (Supporting Information Figure S17). The rough circular shape of these domains led to a significant increase in density of the third neighbors when compared to the other inositides. The reduced headgroup charge, compared to the poly-phosphoryated inositides, likely impacted the cross-linking of monophosphorylated inositides, leading to less dense domains. For PI(3,4)P_2, PI(3,5)P_2, and PI(3,4,5)P_3, domains were formed with a density similar to the monophosphorylated cases, although apparently less localized. While PI(3,5)P_2 and PI(3,4,5)P_3 domains show very little headgroup organization, PI(3,4)P_2 domains are slightly more organized, as was the case with PI(4,5)P_2 (Supporting Information Figure S17). PI(4,5)P2 stands out from the other poly-phosphorylated inositides, by presenting a larger

https://doi.org/10.1021/acs.jctc.1c00615
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

Journal of Chemical Theory and Computation pubs.acs.org/JCTC Article

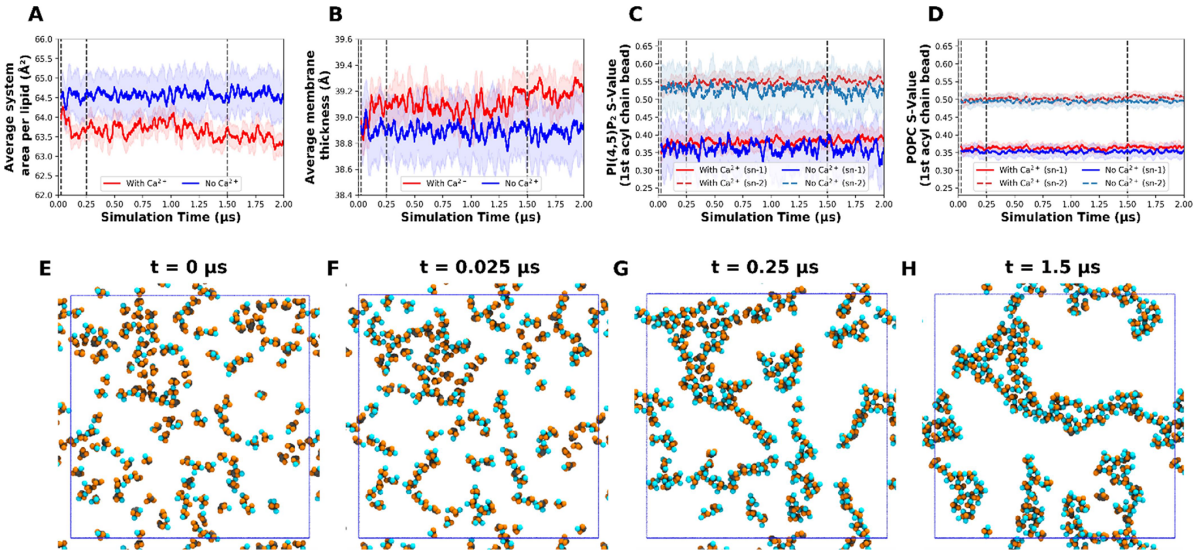


Figure 5. PI(4,5)P_2 calcium-induced cluster growth. System area per lipid (A), membrane thickness (B), and PI(4,5)P_2 (C) and POPC (D) order parameters during the initial calcium-induced cluster growth stages. They were obtained from 20:80 PI(4,5)P_2:POPC systems containing Ca^{2+} at a 2.5:1 Ca^{2+}:PI(4,5)P_2 ratio at 300 K. The values shown are the running average over the course of 50 frames, with the shaded areas representing the running standard deviation. (E−H) Simulation snapshots at four time points are also shown, with the same viewpoint and coloring as in Figure 4.

favored contact area (Supporting Information Figure S23), which likely assists in the assembly of larger, denser aggregates.

The observed CG-simulated behavior is roughly in agreement with available experimental results. Calcium-induced clusters of PI(4,5)P_2 and PI(3,5)P_2 have been observed in model membranes at physiological concentrations of both calcium and lipids, with PI(4,5)P_2 inducing larger and likely more stable clusters than PI(3,5)P_2[16] (the behavior of PI(4,5)P_2 is discussed and explored in greater detail below). This agrees with the behavior of our Martini 3 bis-phosphorylated inositides. On the other hand, when the monophosphorylated PI(4)P was tested in the same experimental conditions, no calcium-induced clusters were observed.[16] However, monophosphorylated PI species can cocluster with bis-phosphorylated inositides such as PI(4,5)P_2, showing that they have some clustering propensity.[17] We could not find any experimental evidence on cation-induced aggregation of PI(3,4,5)P_3; however, it is believed that it likely does.[92] As for PI, there is no experimental evidence for its cation-induced clustering in model membranes, and it has been shown that it does not cocluster with other inositide lipids.[17] This is in agreement with our observation of no significant PI aggregation in the presence of calcium.

Phosphoinositide Aggregation and Choice of Mapping. In our models, the inositol sugar mapping is rotated by 1 atomistic carbon when compared to the Martini 2 inositol mapping, i.e., carbons 1 and 2, 6 and 5, and 4 and 3 correspond to the CG beads *C1*, *C2*, and *C3* respectively. This was done to simplify the modeling of PI(4,5)P_2, keeping it symmetric with the headgroup phosphate beads bonded to separate inositol beads. While both mappings are valid, this should slightly improve PI(4,5)P_2 self-interactions, as well as headgroup packing in cation-induced nanodomains.

In comparison to our Martini 3 models, the Martini 2 models lead to much more dramatic aggregation (Supporting Information Figure S15), with every phosphoinositide undergoing significant calcium-induced clustering. This more intense aggregation is likely to be the combined result of several effects: overly loose phosphoinositide conformational dynamics, less precise cation−phosphate interactions, and overestimated phosphoinositide self-interactions. The drastic increase in first-neighbor density and the highly organized headgroup structure of the domains formed with the Martini 2 models hint that overestimated self-interactions are likely the main driver. Such an overestimation is one of the main aspects that Martini 3 mitigates.[20,29]

Characterization of PI(4,5)P_2 Calcium-Induced Clusters. PI(4,5)P_2 has been shown to tightly bind both calcium and magnesium through strong electrostatic interactions between the negatively charged headgroup phosphates and the divalent ions. These divalent cations can influence PI(4,5)P_2 lateral organization dramatically, inducing the formation of nanodomains. Cation-induced nanodomains have been observed in model membranes at physiological concentrations of both lipids and cations using fluorescence spectroscopy methods[15−17] and AFM.[93] PI(4,5)P_2 nanodomains have also been observed in PC12 cells using single-molecule imaging techniques.[9495] Additionally, it has been shown in fluorescence correlation spectroscopy experiments carried out in Rat1 fibroblasts and HEK cells that the diffusion of PI(4,5)P_2 is significantly slower than expected for free phospholipids, concluding that approximately two-thirds of PI(4,5)P_2 in the inner leaflet of the plasma membrane is somehow sequestered.[96]

Our Martini 3 PI(4,5)P_2 CG model can undergo the formation of these cation-induced clusters in a qualitatively accurate manner and, in contrast to the other lipids, form fairly organized structures. There are no experimental studies on the structural organization of PI(4,5)P_2 within these nanodomains, and little available structural evidence comes from atomistic MD studies that do not actually form fully assembled

https://doi.org/10.1021/acs.jctc.1c00615
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

Journal of Chemical Theory and Computation pubs.acs.org/JCTC Article

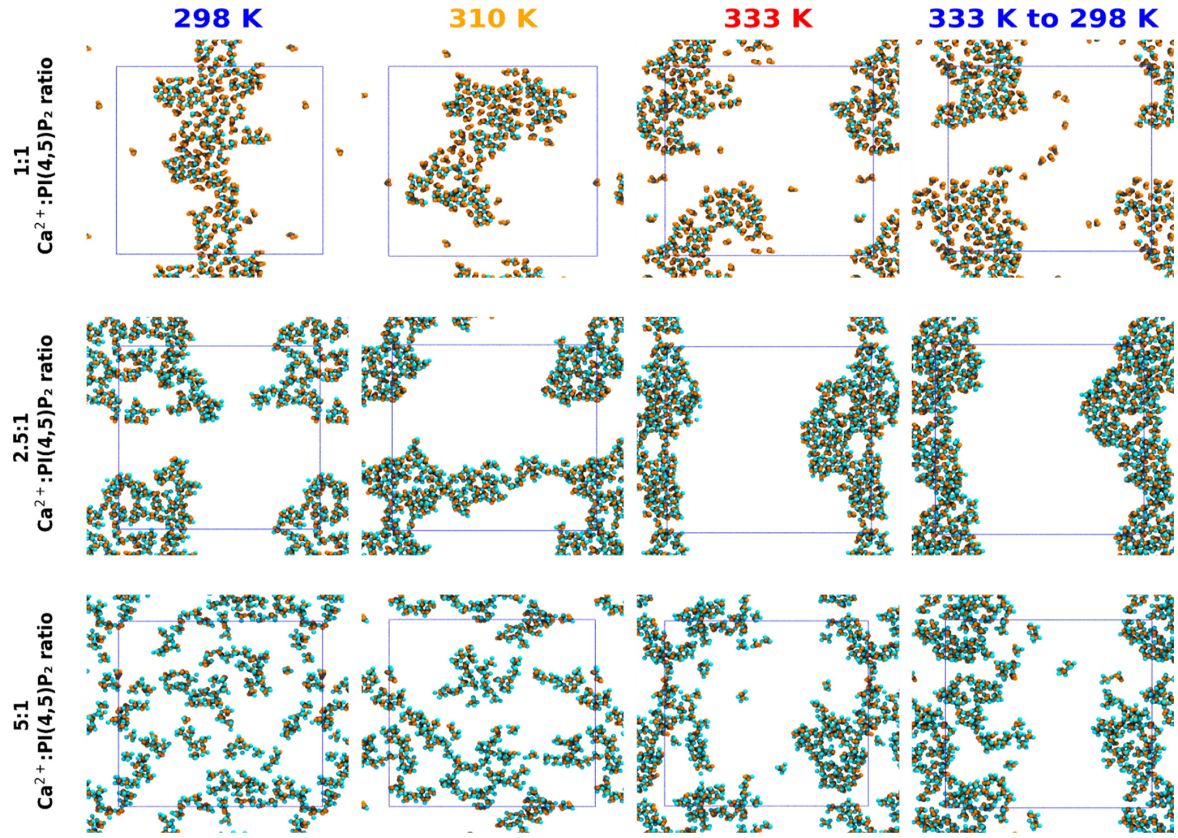


Figure 6. Effects of temperature and Ca^{2+} concentration on calcium-induced PI(4,5)P_2 clusters. Final simulation snapshots obtained from 20:80 PI(4,5)P_2:POPC systems containing Ca^{2+} at 1:1, 2.5:1, or 5:1 Ca^{2+}:PI(4,5)P_2 ratios and ran at 298, 310, or 333 K. Snapshots from control runs, where the final conformations from the simulations running at 333 K were used to restart the simulation at 298 K, are also shown. The same viewpoint and component colors as in Figure 4 were employed.

PI(4,5)P_2 nanodomains but rather smaller aggregates due to either the system size, timescale, or force field limitations.[71,97]

Figure 5 depicts the formation dynamics of PI(4,5)P_2 clusters. At t = 0, after minimization and relaxation, a fair amount of calcium has already adsorbed onto the membrane surface. A continued decrease in the ApL with increasing calcium adsorption is observed over the simulation time. Also at t = 0, we can already see the formation of the first aggregates, with PI(4,5)P_2 assembling into small filaments.

After 25 ns, the filament-like clusters continue to grow, and at 250 ns, large 10 lipid filament structures are observed alongside a significant decrease in the ApL and an increase in membrane thickness. At this point, these structures are in qualitative agreement with those found by Han et al.[98] who, despite a significantly different setup (monolayers composed of 100% PI(4,5)P_2), observed the formation of stringlike calcium-induced clusters. Han et al. also detected a decrease in the ApL, in agreement with our findings. Other authors have also reported similar cluster structures and ApL responses, using atomistic bilayer systems with low PI(4,5)P_2 concentrations.[71]

After 1.5 μs, the PI(4,5)P_2 filament-like structures begin to coalesce and form the denser, lobed structures. With the formation of these structures, we also begin to see a slight increase in the acyl chain order of both PI(4,5)P_2 and POPC. The ordering of PI(4,5)P_2 acyl chains is also in agreement with experimental data.[99] At longer timescales, the domains continue to coalesce (Figure 4, PI(4,5)P_2), and these effects on the membrane properties are either maintained or further magnified by the increase in cluster density (Supporting Information Figure S18).

Finally, we tested how the formation of PI(4,5)P_2 domains responds to varying temperature and Ca^{2+}:PI(4,5)P_2 ratios (Figure 6). The temperature, up to the tested 333 K, does not appear to significantly impact lipid organization. While we observe larger clusters at higher temperatures, it looks as if it is simply the result of the speedup kinetics (i.e., a higher lipid diffusion coefficient). To confirm this, the final conformations from the systems ran at 333 K were used to restart the simulations at 298 K. No significant changes to the clusters were observed from lowering the temperature. Both PI(4,5)-P_2–PI(4,5)P_2 and PI(4,5)P_2–POPC RDFs confirm these observations (Supporting Information Figure S19), which are also in agreement with experimental results.[99]

The calcium concentration, however, influences the formation of PI(4,5)P_2 domains in a more complex manner. At a 1:1 Ca^{2+}:PI(4,5)P_2 ratio, the clusters formed are not very dense, with some room for other lipids to be incorporated. When the Ca^{2+}:PI(4,5)P_2 ratio is increased to 2.5, clusters become much denser and tighter packed. However, a further increase in the Ca^{2+}:PI(4,5)P_2 ratio to 5 causes clusters to

https://doi.org/10.1021/acs.jctc.1c00615
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

Journal of Chemical Theory and Computation pubs.acs.org/JCTC Article

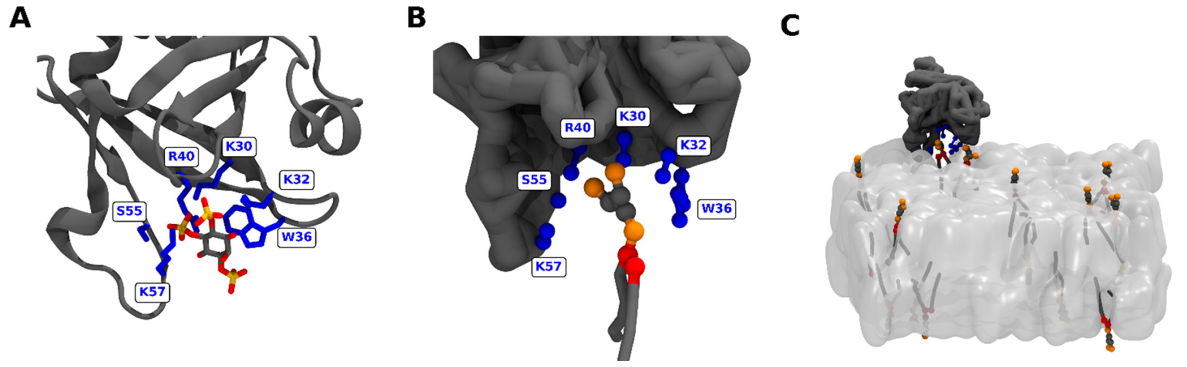


Figure 7. PI(4,5)P_2 recognition and binding by PLCδ1 PH. PLCδ1 PH domain bound to inositol 1,4,5-trisphosphate as determined by X-ray diffraction (PDB: 1MAI[58]) (A). Martini 3 CG PLCδ1 PH domain bound to PI(4,5)P_2 (B). Snapshot of the CG membrane system showcasing PH domain membrane recognition and PI(4,5)P_2 binding (C). PH domains are depicted in gray, with the PI(4,5)P_2 binding pocket amino acid residues explicitly shown in blue. POPC is shown as the translucent surface.

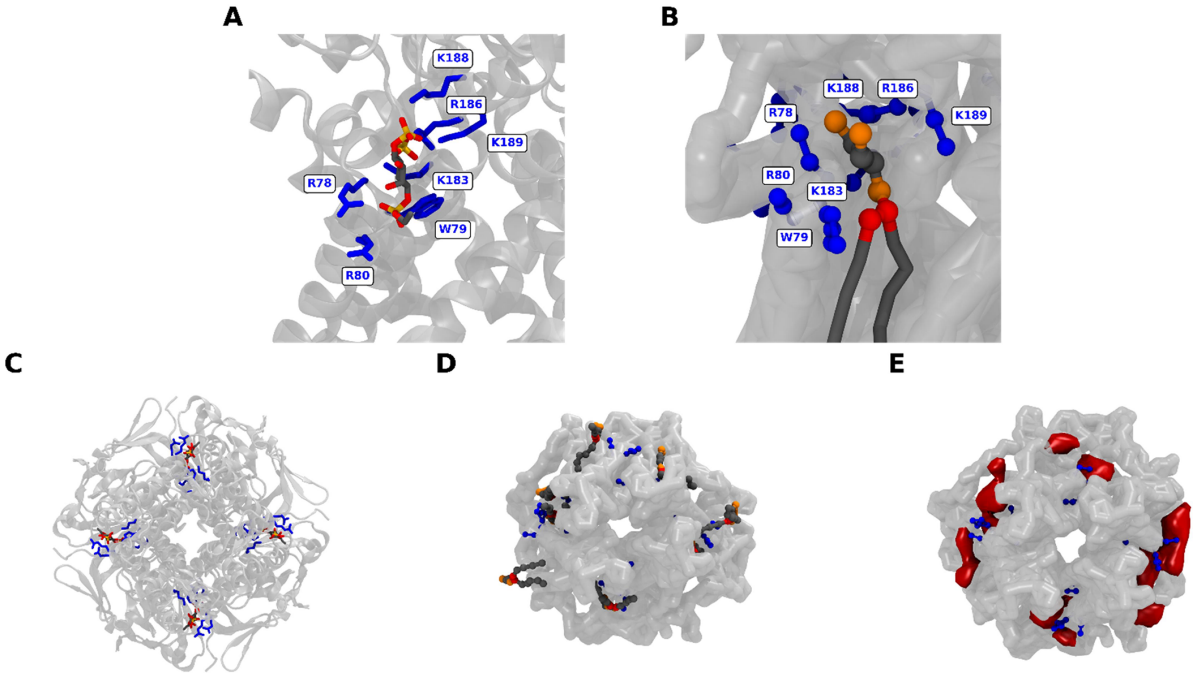


Figure 8. PI(4,5)P_2 recognition and binding to the Kir2.2 channel. PI(4,5)P_2 binding site (A) and top view (C) of Kir2.2 bound to PI(4,5)P_2 as determined by X-ray diffraction (PDB: 6M84[59]). PI(4,5)P_2 binding site (B), top view (D), and PI(4,5)P_2 density map (E) of Martini 3 Kir2.2 bound to PI(4,5)P_2. Kir2.2 channels are depicted in gray, with the PI(4,5)P_2 binding pocket amino acid residues explicitly shown in blue. PI(4,5)P_2 densities are shown in red volume maps, at an isovalue corresponding to at least a 3% occupancy.

become disrupted; while some PI(4,5)P_2 filament-like structures are still formed, they are sparsely interconnected and not very condensed, especially at temperatures below 333 K. At high calcium concentrations, PI(4,5)P_2 headgroups no longer gain from sharing cations, as there are enough ions to fully screen the lipids without bridging lipids together.

PI(4,5)P_2 Recognition and Binding by Proteins. A crucial area of phosphoinositide research is the interaction with proteins. Phosphoinositide−protein interactions are at the base of their downstream cellular signaling, and as such, our models must correctly replicate the experimental behavior. To test this, we probed the recognition and binding of our Martini 3 model of PI(4,5)P_2 by 2 canonical PI(4,5)P_2-binding proteins, the pleckstrin homology (PH) domain from phospholipase Cδ (PLCδ1) and an inward rectifier potassium (Kir2.2) channel.

The isolated PLCδ1 PH domain was found to bind to PI(4,5)P_2 with high affinity and specificity.[100,101] In fact, these studies provided the first demonstrations of specific PI recognition by a PH domain and showed how binding domains recognize specific phosphoinositides in membranes.[102] PLCδ1 PH domains are still to this day used as excellent protein models to study PI(4,5)P_2−protein interactions and PI(4,5)P_2 organization.

https://doi.org/10.1021/acs.jctc.1c00615
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

To test whether PI(4,5)P_2 recognition by the PH domain is in agreement with previous findings, we built membrane systems at a 5:95 PI(4,5)P_2:POPC ratio and placed a PH domain in the corner of the simulation box. Over the course of the simulation, the PH domain correctly captured its binding orientation and probed the surface of the membrane, eventually recognizing and binding a PI(4,5)P_2 lipid in its binding pocket (Figure 7C). The interactions established by the Martini 3 model (Figure 7B) are in agreement with those established by the X-ray structure of the PH domain bound to the soluble headgroup of PI(4,5)P_2[58] (Figure 7A), as well as those observed in previous Martini 2 simulation studies.[61,103] Most of the amino acid contacts that were observed in the X-ray structure were also observed in our Martini 3 model (K30, K32, W36, R40, S55, and K57), as well as a very similar binding pose (Figure 7A,B). These findings point at a successful PH domain recognition and binding of PI(4,5)P_2.

Kir channels are tetrameric transmembrane potassium channels composed of identical subunits. These channels were found to be regulated by several lipid species, especially PI(4,5)P_2, which was found to activate mammalian Kir channels.[104,105] Resolved crystal structures of Kir channels with bound PI(4,5)P_2,[59] as well as previous Martini 2 simulation studies,[23,62] show a binding pocket from which the PI(4,5)P_2 headgroup may interact with both the transmembrane and cytoplasmic domain, favoring the open conformation.

To test the PI(4,5)P_2 binding to Kir channels, an asymmetric membrane system was built, composed of full POPC on the upper leaflet and of a 5:95 POPC:PI(4,5)P_2 ratio on the lower leaflet. A Kir2.2 channel was positioned roughly as described in previous simulation studies.[23,62] Over the course of the simulation, PI(4,5)P_2 lipids diffused freely and converged upon a site common to all subunits (Figure 8D,E), which overlaps with the PI(4,5)P_2 binding site detected by X-ray crystallography[59] (Figure 8C). Indeed, PI(4,5)P_2 lipids were found bound to the described Kir binding site (Figure 8B) and established amino acid contacts (K189, K186, K188, K183, R78, R80, and W79), which are in agreement with those observed in the resolved X-ray structure[59] as well as in previous Martini 2 studies.[23,62]

The recognition and binding behavior, observed with both the PLCδ1 PH domain and the Kir2.2 channel, shows that the Martini 3 PI(4,5)P_2 model is correctly interacting with proteins and replicating experimental findings. While other inositide species were not tested, by following the Martini chemical building block approach, we expect a degree of transferability that should lead to a representative behavior also for those cases.

CONCLUSIONS

In this work, we successfully develop and validate Martini 3 CG topologies for inositol and 8 phosphoinositides. More than a simple version update from existing Martini 2 models, these are models developed independently of their Martini 2 counterparts, with greater accuracy and expanded application scope to include, among others, accurate reproduction of cation-mediated phosphoinositide aggregation. This enabled a thorough characterization of calcium-induced PI(4,5)P_2 clusters, which showed biophysical and structural properties in agreement with available AA and experimental evidence.

The models that we developed here were built and tested atop the phosphoacylglycerol backbone initially put forth with Martini 3. Should that model be refined, phosphoinositide behavior could be affected. The parameterization strategy that we lay out here can then be employed again to adjust the phosphoinositide models. In fact, force field development is a constantly ongoing effort, as attested to by the recent release of the significantly improved Martini 3, even as Martini 2 enjoys wide application as successfully used. Our phosphoinositide models should be no exception, and we look forward to revisiting their parameterization as CG methodologies evolve, as more accurate atomistic models are developed, and as new experimental data become available.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jctc.1c00615.

Extended methods and materials providing details behind the simulation analysis, as well as additional analysis results on phosphoinositide parameterization, calcium-induced aggregation, and characterization of PI(4,5)P_2 calcium-induced clusters (PDF)

AUTHOR INFORMATION

Corresponding Author

Manuel N. Melo − *Instituto de Tecnologia Química e Biológica António Xavier, Universidade Nova de Lisboa, Oeiras 2780-157, Portugal;* orcid.org/0000-0001-6567-0513; Email: m.n.melo@itqb.unl.pt

Authors

Luís Borges-Araújo − *iBB-Institute for Bioengineering and Biosciences, Instituto Superior Técnico, Universidade de Lisboa, Lisbon 1049-001, Portugal; Associate Laboratory i4HB—Institute for Health and Bioeconomy, at Instituto Superior Técnico, Universidade de Lisboa, Lisbon 1049-001, Portugal; Instituto de Tecnologia Química e Biológica António Xavier, Universidade Nova de Lisboa, Oeiras 2780-157, Portugal;* orcid.org/0000-0003-3004-7750

Paulo C. T. Souza − *Molecular Microbiology and Structural Biochemistry, UMR 5086 CNRS & University of Lyon, Lyon F-69367, France;* orcid.org/0000-0003-0660-1301

Fábio Fernandes − *iBB-Institute for Bioengineering and Biosciences and Department of Bioengineering, Instituto Superior Técnico, Universidade de Lisboa, Lisbon 1049-001, Portugal; Associate Laboratory i4HB—Institute for Health and Bioeconomy, at Instituto Superior Técnico, Universidade de Lisboa, Lisbon 1049-001, Portugal*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jctc.1c00615

Author Contributions

L.B.-A., P.C.T.S., and M.N.M. designed the models. All authors designed the simulation/validation systems. L.B.-A. and M.N.M. designed the analyses. L.B.-A. performed simulations and analyses. All authors contributed to manuscript writing and revision and have given approval to the final version of the manuscript.

Notes

The authors declare no competing financial interest.

martini_v3.0_phosphoinositides_v1.0.itp is a molecule topology description file, in a text format compatible with the GROMACS software, containing the CG parameters for the

https://doi.org/10.1021/acs.jctc.1c00615
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

Martini 3 phosphoinositide lipid models developed in this study, linked to several acyl chain profiles and the soluble inositol and phosphoinositide headgroup models derived from the lipid topologies.

■ ACKNOWLEDGMENTS

The authors thank S. Thallmair for critical discussion during parameter development. L.B.-A. thanks the Medical Biochemistry and Biophysics Doctoral Program (M2B-PhD) and Fundação para a Ciência e a Tecnologia − Ministério da Ciência, Tecnologia e Ensino Superior (FCT-MCTES, Portugal) for PhD fellowship PD/BD/137492/2018. M.N.M. thanks FCT-MCTES for the "scientific employment stimulus" program CEECIND/04124/2017. M.N.M. also acknowledges FCT for the funding project MOSTMICRO-ITQB, with references UIDB/04612/2020 and UIDP/04612/2020.

■ REFERENCES

(1) Viaud, J.; Mansour, R.; Antkowiak, A.; Mujalli, A.; Valet, C.; Chicanne, G.; Xuereb, J. M.; Terrisse, A. D.; Séverin, S.; Gratacap, M. P.; Gaits-Iacovoni, F.; Payrastre, B. Phosphoinositides: Important Lipids in the Coordination of Cell Dynamics. *Biochimie* **2016**, *125*, 250−258.

(2) Balla, T. Phosphoinositides: Tiny Lipids with Giant Impact on Cell Regulation. *Physiol. Rev.* **2013**, *93*, 1019−1137.

(3) Höning, S.; Ricotta, D.; Krauss, M.; Späte, K.; Spolaore, B.; Motley, A.; Robinson, M.; Robinson, C.; Haucke, V.; Owen, D. J. Phosphatidylinositol-(4,5)-Bisphosphate Regulates Sorting Signal Recognition by the Clathrin-Associated Adaptor Complex AP2. *Mol. Cell* **2005**, *18*, 519−531.

(4) Posor, Y.; Eichhorn-Grünig, M.; Haucke, V. Phosphoinositides in Endocytosis. *Biochimica et Biophysica Acta - Molecular and Cell Biology of Lipids*; Elsevier B.V. June 1, 2015, pp. 794−804.

(5) Levin, R.; Grinstein, S.; Schlam, D. Phosphoinositides in Phagocytosis and Macropinocytosis. *Biochimica et Biophysica Acta - Molecular and Cell Biology of Lipids*;. Elsevier B.V. June 1, 2015, pp. 805−823.

(6) Hilgemann, D. W.; Feng, S.; Nasuhoglu, C. The Complex and Intriguing Lives of PIP2 with Ion Channels and Transporters. *Sci. STKE* **2001**, *2001*, re19.

(7) Dickson, E. J.; Hille, B. Understanding Phosphoinositides: Rare, Dynamic, and Essential Membrane Phospholipids. *Biochem. J.* **2019**, *476*, 1−23.

(8) Hille, B.; Dickson, E. J.; Kruse, M.; Vivas, O.; Suh, B. C. Phosphoinositides Regulate Ion Channels. *Biochimica et Biophysica Acta - Molecular and Cell Biology of Lipids*; Elsevier B.V. June 1, 2015, pp. 844−856.

(9) Tsujita, K.; Itoh, T. Phosphoinositides in the Regulation of Actin Cortex and Cell Migration. *Biochimica et Biophysica Acta - Molecular and Cell Biology of Lipids*; Elsevier B.V. June 1, 2015, pp. 824−831.

(10) Fiume, R.; Faenza, I.; Sheth, B.; Poli, A.; Vidalle, M. C.; Mazzetti, C.; Abdul, S. H.; Campagnoli, F.; Fabbrini, M.; Kimber, S. T.; Mariani, G. A.; Xian, J.; Marvi, M. V.; Mongiorgi, S.; Shah, Z.; Divecha, N. Nuclear Phosphoinositides: Their Regulation and Roles in Nuclear Functions. *Int. J. Mol. Sci.* **2019**, *20*, 2991.

(11) Raghu, P.; Joseph, A.; Krishnan, H.; Singh, P.; Saha, S. Phosphoinositides: Regulators of Nervous System Function in Health and Disease. *Frontiers in Molecular Neuroscience*; Frontiers Media S.A. August 23, 2019, p 208.

(12) Bunney, T. D.; Katan, M. Phosphoinositide Signalling in Cancer: Beyond PI3K and PTEN. *Nature Reviews Cancer*; Nature Publishing Group May 2010, pp. 342−352.

(13) Thapa, N.; Tan, X.; Choi, S.; Lambert, P. F.; Rapraeger, A. C.; Anderson, R. A. The Hidden Conundrum of Phosphoinositide Signaling in Cancer. *Trends in Cancer;* Cell Press July *1*, 2016, pp. 378−390.

(14) Irvine, R. F. Thematic Review Series: Living History of Lipids: A Short History of Inositol Lipids. *Journal of Lipid Research;* American Society for Biochemistry and Molecular Biology Inc. November *1*, 2016, pp. 1987−1994.

(15) Sarmento, M. J.; Coutinho, A.; Fedorov, A.; Prieto, M.; Fernandes, F. Ca2+ Induces PI(4,5)P2 Clusters on Lipid Bilayers at Physiological PI(4,5)P2 and Ca2+ Concentrations. *Biochim. Biophys. Acta* **2014**, *1838*, 822−830.

(16) Sarmento, M. J.; Coutinho, A.; Fedorov, A.; Prieto, M.; Fernandes, F. Membrane Order Is a Key Regulator of Divalent Cation-Induced Clustering of PI(3,5)P2 and PI(4,5)P2. *Langmuir* **2017**, *33*, 12463−12477.

(17) Wen, Y.; Vogt, V. M.; Feigenson, G. W. Multivalent Cation-Bridged PI(4,5)P2 Clusters Form at Very Low Concentrations. *Biophys. J.* **2018**, *114*, 2630−2639.

(18) Ingólfsson, H. I.; Melo, M. N.; Van Eerden, F. J.; Arnarez, C.; Lopez, C. A.; Wassenaar, T. A.; Periole, X.; De Vries, A. H.; Tieleman, D. P.; Marrink, S. J. Lipid Organization of the Plasma Membrane. *J. Am. Chem. Soc.* **2014**, *136*, 14554−14559.

(19) Marrink, S. J.; Risselada, H. J.; Yefimov, S.; Tieleman, D. P.; De Vries, A. H. The MARTINI Force Field: Coarse Grained Model for Biomolecular Simulations. *J. Phys. Chem. B* **2007**, *111*, 7812−7824.

(20) Alessandri, R.; Souza, P. C. T.; Thallmair, S.; Melo, M. N.; de Vries, A. H.; Marrink, S. J. Pitfalls of the Martini Model. *J. Chem. Theory Comput.* **2019**, *15*, 5448−5460.

(21) Marrink, S. J.; Corradi, V.; Souza, P. C. T.; Ingólfsson, H. I.; Tieleman, D. P.; Sansom, M. S. P. Computational Modeling of Realistic Cell Membranes. *Chemical Reviews;* American Chemical Society May *8*, 2019, pp. 6184−6226.

(22) Ingólfsson, H. I.; Carpenter, T. S.; Bhatia, H.; Bremer, P. T.; Marrink, S. J.; Lightstone, F. C. Computational Lipidomics of the Neuronal Plasma Membrane. *Biophys. J.* **2017**, *113*, 2271−2280.

(23) Stansfeld, P. J.; Hopkinson, R.; Ashcroft, F. M.; Sansom, M. S. P. PIP 2 -Binding Site in Kir Channels: Definition by Multiscale Biomolecular Simulations. *Biochemistry* **2009**, *48*, 10926−10933.

(24) Hansen, S. B.; Tao, X.; MacKinnon, R. Structural Basis of PIP2 Activation of the Classical Inward Rectifier K+ Channel Kir2.2. *Nature* **2011**, *477*, 495−498.

(25) Stark, A. C.; Andrews, C. T.; Elcock, A. H. Toward Optimized Potential Functions for Protein-Protein Interactions in Aqueous Solutions: Osmotic Second Virial Coefficient Calculations Using the MARTINI Coarse-Grained Force Field. *J. Chem. Theory Comput.* **2013**, *9*, 4176−4185.

(26) Javanainen, M.; Martinez-Seara, H.; Vattulainen, I. Excessive Aggregation of Membrane Proteins in the Martini Model. *PLoS One* **2017**, *12*, No. e0187936.

(27) Majumder, A.; Straub, J. E. Addressing the Excessive Aggregation of Membrane Proteins in the MARTINI Model. *J. Chem. Theory Comput.* **2021**, *17*, 2513−2521.

(28) Schmalhorst, P. S.; Deluweit, F.; Scherrers, R.; Heisenberg, C.-P.; Sikora, M. Overcoming the Limitations of the MARTINI Force Field in Simulations of Polysaccharides. *J. Chem. Theory Comput.* **2017**, *13*, 5039−5053.

(29) Souza, P. C. T.; Alessandri, R.; Barnoud, J.; Thallmair, S.; Faustino, I.; Grünewald, F.; Patmanidis, I.; Abdizadeh, H.; Bruininks, B. M. H.; Wassenaar, T. A.; Kroon, P. C.; Melcr, J.; Nieto, V.; Corradi, V.; Khan, H. M.; Domański, J.; Javanainen, M.; Martinez-Seara, H.; Reuter, N.; Best, R. B.; Vattulainen, I.; Monticelli, L.; Periole, X.; Tieleman, D. P.; de Vries, A. H.; Marrink, S. J. Martini 3: A General Purpose Force Field for Coarse-Grained Molecular Dynamics. *Nat. Methods* **2021**, *18*, 382−388.

(30) López, C. A.; Sovova, Z.; van Eerden, F. J.; de Vries, A. H.; Marrink, S. J. Martini Force Field Parameters for Glycolipids. *J. Chem. Theory Comput.* **2013**, *9*, 1694−1708.

(31) *Martini 2 lipid topology for C16:0/18:1 PIP2 (POP2)* http://cgmartini.nl/images/parameters/lipids/Phosphatidylinositols/POP2/martini_v2.0_POP2_01.itp (accessed Apr 13, 2021).

(32) Bulacu, M.; Goga, N.; Zhao, W.; Rossi, G.; Monticelli, L.; Periole, X.; Tieleman, D. P.; Marrink, S. J. Improved Angle Potentials

https://doi.org/10.1021/acs.jctc.1c00615
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

for Coarse-Grained Molecular Dynamics Simulations. *J. Chem. Theory Comput.* **2013**, *9*, 3282−3292.
(33) Borges-Araújo, L.; Fernandes, F. Structure and Lateral Organization of Phosphatidylinositol 4,5-Bisphosphate. *Molecules* **2020**, *25*, 3885.
(34) 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.
(35) Michaud-Agrawal, N.; Denning, E. J.; Woolf, T. B.; Beckstein, O. MDAnalysis: A Toolkit for the Analysis of Molecular Dynamics Simulations. *J. Comput. Chem.* **2011**, *32*, 2319−2327.
(36) Pérez, F.; Granger, B. E. IPython: A System for Interactive Scientific Computing. *Comp. Sci. Eng.* **2007**, 21−29.
(37) Harris, C. R.; Millman, K. J.; van der Walt, S. J.; Gommers, R.; Virtanen, P.; Cournapeau, D.; Wieser, E.; Taylor, J.; Berg, S.; Smith, N. J.; Kern, R.; Picus, M.; Hoyer, S.; van Kerkwijk, M. H.; Brett, M.; Haldane, A.; del Río, J. F.; Wiebe, M.; Peterson, P.; Gérard-Marchant, P.; Sheppard, K.; Reddy, T.; Weckesser, W.; Abbasi, H.; Gohlke, C.; Oliphant, T. E. Array Programming with NumPy. *Nature* **2020**, *585*, 357−362.
(38) Virtanen, P.; Gommers, R.; Oliphant, T. E.; Haberland, M.; Reddy, T.; Cournapeau, D.; Burovski, E.; Peterson, P.; Weckesser, W.; Bright, J.; van der Walt, S. J.; Brett, M.; Wilson, J.; Millman, K. J.; Mayorov, N.; Nelson, A. R. J.; Jones, E.; Kern, R.; Larson, E.; Carey, C. J.; Polat, İ.; Feng, Y.; Moore, E. W.; VanderPlas, J.; Laxalde, D.; Perktold, J.; Cimrman, R.; Henriksen, I.; Quintero, E. A.; Harris, C. R.; Archibald, A. M.; Ribeiro, A. H.; Pedregosa, F.; van Mulbregt, P.; SciPy 1.0 Contributors; Vijaykumar, A.; Bardelli, A. P.; Rothberg, A.; Hilboll, A.; Kloeckner, A.; Scopatz, A.; Lee, A.; Rokem, A.; Woods, C. N.; Fulton, C.; Masson, C.; Häggström, C.; Fitzgerald, C.; Nicholson, D. A.; Hagen, D. R.; Pasechnik, D. V.; Olivetti, E.; Martin, E.; Wieser, E.; Silva, F.; Lenders, F.; Wilhelm, F.; Young, G.; Price, G. A.; Ingold, G. L.; Allen, G. E.; Lee, G. R.; Audren, H.; Probst, I.; Dietrich, J. P.; Silterra, J.; Webber, J. T.; Slavič, J.; Nothman, J.; Buchner, J.; Kulick, J.; Schönberger, J. L.; de Miranda Cardoso, J. V.; Reimer, J.; Harrington, J.; Rodríguez, J. L. C.; Nunez-Iglesias, J.; Kuczynski, J.; Tritz, K.; Thoma, M.; Newville, M.; Kümmerer, M.; Bolingbroke, M.; Tartre, M.; Pak, M.; Smith, N. J.; Nowaczyk, N.; Shebanov, N.; Pavlyk, O.; Brodtkorb, P. A.; Lee, P.; McGibbon, R. T.; Feldbauer, R.; Lewis, S.; Tygier, S.; Sievert, S.; Vigna, S.; Peterson, S.; More, S.; Pudlik, T.; Oshima, T.; Pingel, T. J.; Robitaille, T. P.; Spura, T.; Jones, T. R.; Cera, T.; Leslie, T.; Zito, T.; Krauss, T.; Upadhyay, U.; Halchenko, Y. O.; Vázquez-Baeza, Y. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nat. Methods* **2020**, *17*, 261−272.
(39) Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; Vanderplas, J.; Passos, A.; Cournapeau, D.; Brucher, M.; Perrot, M.; Duchesnay, É. Scikit-Learn: Machine Learning in Python. *J. Mach. Learn Re.s* **2011**, *12*, 2825.
(40) Rycroft, C. H. VORO++: A Three-Dimensional Voronoi Cell Library in C++. *Chaos;* American Institute of Physics Inc. December *19*, 2009, p 041111, DOI: 10.1063/1.3215722.
(41) Hunter, J. D. Matplotlib: A 2D Graphics Environment. *Comput. Sci. Eng.* **2007**, *9*, 90−95.
(42) Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual Molecular Dynamics. *J. Mol. Graph.* **1996**, *14*, 33−38.
(43) Sousa, F. M.; Lima, L. M. P.; Arnarez, C.; Pereira, M. M.; Melo, M. N. Coarse-Grained Parameterization of Nucleotide Cofactors and Metabolites: Protonation Constants, Partition Coefficients, and Model Topologies. *J. Chem. Inf. Model.* **2021**, 335.
(44) Huang, J.; Mackerell, A. D. CHARMM36 All-Atom Additive Protein Force Field: Validation Based on Comparison to NMR Data. *J. Comput. Chem.* **2013**, *34*, 2135−2145.
(45) Huang, J.; Rauscher, S.; Nawrocki, G.; Ran, T.; Feig, M.; De Groot, B. L.; Grubmüller, H.; MacKerell, A. D. CHARMM36m: An Improved Force Field for Folded and Intrinsically Disordered Proteins. *Nat. Methods* **2017**, *14*, 71−73.
(46) Jo, S.; Kim, T.; Iyer, V. G.; Im, W. CHARMM-GUI: A Web-Based Graphical User Interface for CHARMM. *J. Comput. Chem.* **2008**, *29*, 1859−1865.
(47) Lee, J.; Cheng, X.; Swails, J. M.; Yeom, M. S.; Eastman, P. K.; Lemkul, J. A.; Wei, S.; Buckner, J.; Jeong, J. C.; Qi, Y.; Jo, S.; Pande, V. S.; Case, D. A.; Brooks, C. L., III; MacKerell, A. D., Jr.; Klauda, J. B.; Im, W. CHARMM-GUI Input Generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM Simulations Using the CHARMM36 Additive Force Field. *J. Chem. Theory Comput.* **2016**, *12*, 405−413.
(48) Wu, E. L.; Cheng, X.; Jo, S.; Rui, H.; Song, K. C.; Dávila-Contreras, E. M.; Qi, Y.; Lee, J.; Monje-Galvan, V.; Venable, R. M.; Klauda, J. B.; Im, W. CHARMM-GUI Membrane Builder toward Realistic Biological Membrane Simulations. *Journal of Computational Chemistry;* John Wiley and Sons Inc. October *15*, 2014, pp. 1997−2004.
(49) Jo, S.; Lim, J. B.; Klauda, J. B.; Im, W. CHARMM-GUI Membrane Builder for Mixed Bilayers and Its Application to Yeast Membranes. *Biophys. J.* **2009**, *97*, 50−58.
(50) Lee, J.; Patel, D. S.; Ståhle, J.; Park, S. J.; Kern, N. R.; Kim, S.; Lee, J.; Cheng, X.; Valvano, M. A.; Holst, O.; Knirel, Y. A.; Qi, Y.; Jo, S.; Klauda, J. B.; Widmalm, G.; Im, W. CHARMM-GUI Membrane Builder for Complex Biological Membrane Simulations with Glycolipids and Lipoglycans. *J. Chem. Theory Comput.* **2019**, *15*, 775−786.
(51) Jo, S.; Kim, T.; Im, W. Automated Builder and Database of Protein/Membrane Complexes for Molecular Dynamics Simulations. *PLoS One* **2007**, *2*, No. e880.
(52) Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M. LINCS: A Linear Constraint Solver for Molecular Simulations. *J. Comput. Chem.* **1997**, *18*, 1463−1472.
(53) 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*, 10089−10092.
(54) Evans, D. J.; Holian, B. L. The Nose-Hoover Thermostat. *J. Chem. Phys.* **1985**, *83*, 4069−4074.
(55) Parrinello, M.; Rahman, A. Polymorphic Transitions in Single Crystals: A New Molecular Dynamics Method. *J. Appl. Phys.* **1981**, *52*, 7182−7190.
(56) Wassenaar, T. A.; Ingólfsson, H. I.; Böckmann, R. A.; Tieleman, D. P.; Marrink, S. J. Computational Lipidomics with Insane: A Versatile Tool for Generating Custom Membranes for Molecular Simulations. *J. Chem. Theory Comput.* **2015**, *11*, 2144−2155.
(57) 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 Research;* Oxford University Press January *1*, 2000, pp. 235−242.
(58) Ferguson, K. M.; Lemmon, M. A.; Schlessinger, J.; Sigler, P. B. Structure of the High Affinity Complex of Inositol Trisphosphate with a Phospholipase C Pleckstrin Homology Domain. *Cell* **1995**, *83*, 1037−1046.
(59) Zangerl-Plessl, E. M.; Lee, S. J.; Maksaev, G.; Bernsteiner, H.; Ren, F.; Yuan, P.; Stary-Weinzinger, A.; Nichols, C. G. Atomistic Basis of Opening and Conduction in Mammalian Inward Rectifier Potassium (Kir2.2) Channels. *J. Gen. Physiol.* **2020**, *152*, No. e201912422.
(60) *GitHub - marrink-lab/vermouth-martinize: Describe and apply transformation on molecular structures and topologies* https://github.com/marrink-lab/vermouth-martinize (accessed Jun 7, 2021).
(61) Herzog, F. A.; Braun, L.; Schoen, I.; Vogel, V. Improved Side Chain Dynamics in MARTINI Simulations of Protein-Lipid Interfaces. *J. Chem. Theory Comput.* **2016**, *12*, 2446−2458.
(62) Duncan, A. L.; Corey, R. A.; Sansom, M. S. P. Defining How Multiple Lipid Species Interact with Inward Rectifier Potassium (Kir2) Channels. *Proc. Natl. Acad. Sci. U. S. A.* **2020**, *117*, 7803−7813.
(63) De Jong, D. H.; Baoukina, S.; Ingólfsson, H. I.; Marrink, S. J. Martini Straight: Boosting Performance Using a Shorter Cutoff and GPUs. *Comput. Phys. Commun.* **2016**, *199*, 1−7.

https://doi.org/10.1021/acs.jctc.1c00615
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

(64) Bussi, G.; Donadio, D.; Parrinello, M. Canonical Sampling through Velocity Rescaling. *J. Chem. Phys.* **2007**, *126*, No. 014101.
(65) Botan, A.; Favela-Rosales, F.; Fuchs, P. F. J.; Javanainen, M.; Kanduč, M.; Kulig, W.; Lamberg, A.; Loison, C.; Lyubartsev, A.; Miettinen, M. S.; Monticelli, L.; Määttä, J.; Ollila, O. H. S.; Retegan, M.; Róg, T.; Santuz, H.; Tynkkynen, J. Toward Atomistic Resolution Structure of Phosphatidylcholine Headgroup and Glycerol Backbone at Different Ambient Conditions. *J. Phys. Chem. B* **2015**, *119*, 15075−15088.
(66) Wang, Y. H.; Slochower, D. R.; Janmey, P. A. Counterion-Mediated Cluster Formation by Polyphosphoinositides. *Chem. Phys. Lipids* **2014**, *182*, 38−51.
(67) Hatcher, E. R.; Guvench, O.; MacKerell, A. D., Jr. CHARMM Additive All-Atom Force Field for Acyclic Polyalcohols, Acyclic Carbohydrates, and Inositol. *J. Chem. Theory Comput.* **2009**, *5*, 1315−1327.
(68) Lupyan, D.; Mezei, M.; Logothetis, D. E.; Osman, R. A Molecular Dynamics Investigation of Lipid Bilayer Perturbation by PIP2. *Biophys. J.* **2010**, *98*, 240−247.
(69) Slochower, D. R.; Huwe, P. J.; Radhakrishnan, R.; Janmey, P. A. Quantum and All-Atom Molecular Dynamics Simulations of Protonation and Divalent Ion Binding to Phosphatidylinositol 4,5-Bisphosphate (PIP 2). *J. Phys. Chem. B* **2013**, *117*, 8322−8329.
(70) Slochower, D. R.; Wang, Y.-H.; Radhakrishnan, R.; Janmey, P. A. Physical Chemistry and Membrane Properties of Two Phosphatidylinositol Bisphosphate Isomers. *Phys. Chem. Chem. Phys.* **2015**, *17*, 12608−12615.
(71) Bradley, R. P.; Slochower, D. R.; Janmey, P. A.; Radhakrishnan, R. Divalent Cations Bind to Phosphoinositides to Induce Ion and Isomer Specific Propensities for Nano-Cluster Initiation in Bilayer Membranes. *R. Soc. Open Sci.* **2020**, *7*, 192208.
(72) Cheng, T.; Zhao, Y.; Li, X.; Lin, F.; Xu, Y.; Zhang, X.; Li, Y.; Wang, R.; Lai, L. Computation of Octanol-Water Partition Coefficients by Guiding an Additive Model with Knowledge. *J. Chem. Inf. Model.* **2007**, *47*, 2140−2148.
(73) Tetko, I. V.; Tanchuk, V. Y.; Kasheva, T. N.; Villa, A. E. P. Estimation of Aqueous Solubility of Chemical Compounds Using E-State Indices. *J. Chem. Inf. Comput. Sci.* **2001**, *41*, 1488−1493.
(74) Tetko, I. V.; Bruneau, P. Application of ALOGPS to Predict 1-Octanol/Water Distribution Coefficients, LogP, and LogD, of AstraZeneca in-House Database. *J. Pharm. Sci.* **2004**, *93*, 3103−3110.
(75) Tetko, I. V.; Poda, G. I. Application of ALOGPS 2.1 to Predict Log D Distribution Coefficient for Pfizer Proprietary Compounds. *J. Med. Chem.* **2004**, *47*, 5601−5604.
(76) Klopman, G.; Li, J. Y.; Wang, S.; Dimayuga, M.; Wang, S.; Dimayuga, M. Computer Automated Log P Calculations Based on an Extended Group Contribution Approach. *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 752−781.
(77) *ChemAxon Chemicalize* http://www.chemicalize.org/.
(78) Lay, W. K.; Miller, M. S.; Elcock, A. H. Optimizing Solute-Solute Interactions in the GLYCAM06 and CHARMM36 Carbohydrate Force Fields Using Osmotic Pressure Measurements. *J. Chem. Theory Comput.* **2016**, *12*, 1401−1407.
(79) Davis, D. J.; Burlak, C.; Money, N. P. Osmotic Pressure of Fungal Compatible Osmolytes. *Mycol. Res.* **2000**, *104*, 800−804.
(80) Brown, W. H. *Organic Chemistry*; Brooks/Cole Cengage Learning, 2009.
(81) van Paridon, P. A.; de Kruijff, B.; Ouwerkerk, R.; Wirtz, K. W. A. Polyphosphoinositides Undergo Charge Neutralization in the Physiological PH Range: A 31P-NMR Study. *Biochim. Biophys. Acta Lipids Lipid Metab.* **1986**, *877*, 216−219.
(82) Kooijman, E. E.; King, K. E.; Gangoda, M.; Gericke, A. Ionization Properties of Phosphatidylinositol Polyphosphates in Mixed Model Membranes. *Biochemistry* **2009**, *48*, 9360−9371.
(83) Graber, Z. T.; Thomas, J.; Johnson, E.; Gericke, A.; Kooijman, E. E. Effect of H-Bond Donor Lipids on Phosphatidylinositol-3,4,5-Trisphosphate Ionization and Clustering. *Biophys. J.* **2018**, *114*, 126−136.
(84) Baptista, A. M.; Teixeira, V. H.; Soares, C. M. Constant-PH Molecular Dynamics Using Stochastic Titration. *J. Chem. Phys.* **2002**, *117*, 4184.
(85) Grünewald, F.; Souza, P. C. T.; Abdizadeh, H.; Barnoud, J.; Vries, A. H.; de Marrink, S. J. Titratable Martini Model for Constant PH Simulations. *J. Chem. Phys.* **2020**, *153*, No. 024118.
(86) Bradshaw, J. P.; Bushby, R. J.; Giles, C. C. D.; Saunders, M. R.; Saxena, A. The Headgroup Orientation of Dimyristoylphosphatidylinositol-4-Phosphate in Mixed Lipid Bilayers: A Neutron Diffraction Study. *Biochim. Biophys. Acta, Biomembr.* **1997**, *1329*, 124−138.
(87) Zhou, C.; Garigapati, V.; Roberts, M. F. Short-Chain Phosphatidylinositol Conformation and Its Relevance to Phosphatidylinositol-Specific Phospholipase C. *Biochemistry* **1997**, *36*, 15925−15931.
(88) Bradshaw, J. P.; Bushby, R. J.; Giles, C. C. D.; Saunders, M. R. Orientation of the Headgroup of Phosphatidylinositol in a Model Biomembrane as Determined by Neutron Diffraction. *Biochemistry* **1999**, *38*, 8393−8401.
(89) Kishore, A. I.; Prestegard, J. H. Molecular Orientation and Conformation of Phosphatidylinositides in Membrane Mimetics Using Variable Angle Sample Spinning (VASS) NMR. *Biophys. J.* **2003**, *85*, 3848−3857.
(90) Vazquez-Salazar, L. I.; Selle, M.; de Vries, A. H.; Marrink, S. J.; Souza, P. C. T. Martini Coarse-Grained Models of Imidazolium-Based Ionic Liquids: From Nanostructural Organization to Liquid−Liquid Extraction. *Green Chem.* **2020**, *22*, 7376−7386.
(91) Fernandes, F.; Loura, L. M. S.; Fedorov, A.; Prieto, M. Absence of Clustering of Phosphatidylinositol-(4,5)-Bisphosphate in Fluid Phosphatidylcholine. *J. Lipid Res.* **2006**, *47*, 1521−1525.
(92) Gericke, A. Is Calcium Fine-Tuning Phosphoinositide-Mediated Signaling Events Through Clustering? *Biophys. J.* **2018**, *114*, 2483−2484.
(93) Ellenbroek, W. G.; Wang, Y. H.; Christian, D. A.; Discher, D. E.; Janmey, P. A.; Liu, A. J. Divalent Cation-Dependent Formation of Electrostatic PIP2 Clusters in Lipid Monolayers. *Biophys. J.* **2011**, *101*, 2178−2184.
(94) van den Bogaart, G.; Meyenberg, K.; Risselada, H. J.; Amin, H.; Willig, K. I.; Hubrich, B. E.; Dier, M.; Hell, S. W.; Grubmüller, H.; Diederichsen, U.; Jahn, R. Membrane Protein Sequestering by Ionic Protein−Lipid Interactions. *Nature* **2011**, *479*, 552−555.
(95) Wang, J.; Richards, D. A. Segregation of PIP2 and PIP3 into Distinct Nanoscale Regions within the Plasma Membrane. *Biol. Open* **2012**, *1*, 857−862.
(96) Golebiewska, U.; Nyako, M.; Woturski, W.; Zaitseva, I.; McLaughlin, S. Diffusion Coefficient of Fluorescent Phosphatidylinositol 4,5-Bisphosphate in the Plasma Membrane of Cells. *Mol. Biol. Cell* **2008**, *19*, 1663−1669.
(97) Bilkova, E.; Pleskot, R.; Rissanen, S.; Sun, S.; Czogalla, A.; Cwiklik, L.; Rog, T.; Vattulainen, I.; Cremer, P. S.; Jungwirth, P.; Coskun, Ü. Calcium Directly Regulates Phosphatidylinositol 4,5-Bisphosphate Headgroup Conformation and Recognition. *J. Am. Chem. Soc.* **2017**, 4019.
(98) Han, K.; Gericke, A.; Pastor, R. W. Characterization of Specific Ion Effects on PI(4,5)P2 Clustering: Molecular Dynamics Simulations and Graph-Theoretic Analysis. *J. Phys. Chem. B* **2020**, *124*, 1183−1196.
(99) Borges-Araújo, L.; Domingues, M. M.; Fedorov, A.; Santos, N. C.; Melo, M. N.; Fernandes, F. Acyl-Chain Saturation Regulates the Order of Phosphatidylinositol 4,5-Bisphosphate Nanodomains. *Commun. Chem.* **2021**, *4*, 1−13.
(100) Lemmon, M. A.; Ferguson, K. M. Signal-Dependent Membrane Targeting by Pleckstrin Homology (PH) Domains. *Biochem. J.* **2000**, *350*, 1−18.
(101) Garcia, P.; Gupta, R.; Shah, S.; Morris, A. J.; Rudge, S. A.; Scarlata, S.; Petrova, V.; McLaughlin, S.; Rebecchi, M. J. The Pleckstrin Homology Domain of Phospholipase C-Delta 1 Binds with High Affinity to Phosphatidylinositol 4,5-Bisphosphate in Bilayer Membranes. *Biochemistry* **1995**, *34*, 16228−16234.

https://doi.org/10.1021/acs.jctc.1c00615
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

(102) Lemmon, M. A. Pleckstrin Homology (PH) Domains and Phosphoinositides. *Biochem. Soc. Symp.* **2007**, *74*, 81.
(103) Khan, H. M.; Souza, P. C. T.; Thallmair, S.; Barnoud, J.; De Vries, A. H.; Marrink, S. J.; Reuter, N. Capturing Choline-Aromatics Cation-Π Interactions in the MARTINI Force Field. *J. Chem. Theory Comput.* **2020**, *16*, 2550−2560.
(104) D'Avanzo, N.; Cheng, W. W. L.; Doyle, D. A.; Nichols, C. G. Direct and Specific Activation of Human Inward Rectifier K+ Channels by Membrane Phosphatidylinositol 4,5-Bisphosphate. *J. Biol. Chem.* **2010**, *285*, 37129−37132.
(105) Lacin, E.; Aryal, P.; Glaaser, I. W.; Bodhinathan, K.; Tsai, E.; Marsh, N.; Tucker, S. J.; Sansom, M. S. P.; Slesinger, P. A. Dynamic Role of the Tether Helix in PIP2-Dependent Gating of a G Protein-Gated Potassium Channel. *J. Gen. Physiol.* **2017**, *149*, 799−811.

https://doi.org/10.1021/acs.jctc.1c00615
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

Article IV: Revisiting the impact of acyl-chain composition on phosphatidylinositol 4,5-bisphosphate nanodomains using the Martini 3 force field

Revisiting the impact of acyl-chain composition on phosphatidylinositol 4,5-bisphosphate nanodomains using the Martini 3 force field

Luís Borges-Araújo [†,§,⊥], Fábio Fernandes [†,‡,⊥], Manuel N. Melo[§,*]

† iBB-Institute for Bioengineering and Biosciences, Instituto Superior Técnico, Universidade de Lisboa, Lisbon, Portugal

⊥ Associate Laboratory i4HB—Institute for Health and Bioeconomy at Instituto Superior Técnico, Universidade de Lisboa, Lisbon, Portugal

‡ Department of Bioengineering; Instituto Superior Técnico, Universidade de Lisboa, 1049-001 Lisbon, Portugal;

§ Instituto de Tecnologia Química e Biológica António Xavier, Universidade Nova de Lisboa, Av. da República, 2780-157 Oeiras, Portugal.

ABSTRACT: Phosphatidylinositol 4,5-bisphosphate (PI(4,5)P_2) is a minority membrane lipid that plays a complex, critical role in a large spectrum of membrane regulatory processes. The Martini coarse-grained (CG) model is the most widely used CG force field for biomolecular molecular dynamics simulations, especially in the modelling of biomembrane systems where PI(4,5)P_2 has been extensively included. Recently, version 3 of the Martini forcefield was released, greatly refining molecular interactions and addressing many of the limitations of the previous model. We have since developed a new set of phosphoinositide headgroup parameters specific to Martini 3, focused on more accurately reproducing phosphoinositide behavior seen in both experimental studies and atomistic simulations. Here, we use the improved Martini 3 force field, as well as our newly developed phosphoinositide parameters, to revisit our previous Martini CG MD work regarding the impact of acyl-chain composition on the biophysical properties of PI(4,5)P_2 and PI(4,5)P_2 nanodomains. The key observation that increasing saturation yields more ordered and structured nanodomains is well recovered in Martini 3, showcasing the importance of the canonical (18:0 20:4) PI(4,5)P_2 acyl-chain composition for the maintenance of membrane fluidity within PI(4,5)P_2 nanodomains. Representativity is extended to the formation of gel-like PI(4,5)P_2 nanodomains when using saturated acyl chains, with which the Martini 2 model had struggled. This behavior was compared to that of Martini 3 phosphatidylcholine lipids in binary mixtures of mixed acyl chain saturation. Finally, the Martini 3 models were shown to have almost no specificity for negatively curved membrane regions, in contrast with Martini 2 results, which we discuss in terms of the newer models' accuracy and impact on mechanistic views of PI(4,5)P_2.

INTRODUCTION

Phosphatidylinositol 4,5-bisphosphate (PI(4,5)P_2) is a minority membrane lipid in the inner leaflet of the eukaryotic plasma membrane[1], that plays a crucial role in the regulation of a multitude of plasma membrane processes and signalling pathways[2]. Due to its impact in a large spectrum of processes, ranging from endocytosis/exocytosis[3–6], to cytoskeleton reorganization[7] and even cellular proliferation[8], PI(4,5)P_2 has been the focus of numerous studies.

Lateral organization of PI(4,5)P_2 has been the subject of particular attention. PI(4,5)P_2 nanodomains, can be promoted by the interaction of its negatively charged phosphorylated headgroup with divalent cations (such as Ca^{2+} and Mg^{2+}) or positively-charged proteins[9,10]. It has been shown that divalent cations have the ability to induce the dramatic segregation of PI(4,5)P_2 to form highly enriched nanodomains, even at physiological concentrations of both lipid and cation [11,12]. Ca^{2+}-induced PI(4,5)P_2 have been shown to substantially impact membrane biophysical properties, the partition of PI(4,5)P_2 and are likely to also impact PI(4,5)P_2 – protein interactions.

The Martini 2 coarse-grained (CG) model[13] is the most widely used CG force field for biomolecular molecular dynamics (MD) simulations[14,15], having been successfully applied in the modelling of a variety of molecular processes — especially, in the modelling of biomembrane systems[16]. Martini 2 CG MD simulations have been used to successfully study the properties of binary and ternary lipid mixtures[17–19], elastic properties of membranes[20–23], lipid-protein interactions[24] and even realistic models of mammalian plasma membranes[25,26]. PI(4,5)P_2 has been extensively included in these CG MD studies with the Martini forcefield quickly becoming a cornerstone of PI(4,5)P_2 biophysical research.

An example of this is our recent work on the impact of acyl-chain saturation on PI(4,5)P_2 biophysical properties, which combined experimental biophysical techniques with Martini 2 CG MD simulations[27]. We observed that PI(4,5)P_2 species with increasing levels of acyl-chain saturation clustered in progressively more ordered nanodomains, culminating in the formation of gel-like nanodomains for fully saturated species. This points at the importance of the canonical 1-stearoyl-2-arachidonyl (18:0 20:4) acyl-composition to guarantee, not only low bending rigidity and membrane permeability[28], but also low membrane order within PI(4,5)P_2 nanodomains.

Recently Martini 3, a new and improved version of the Martini force field, was released[29]. Martini 3 addresses the major limitations of Martini 2 by not only including bead types specific for mappings finer than 4-to-1, but also by expanding the number of bead types, allowing for a larger discrimination of the chemical space. These changes allow for more accurate and flexible models. The new Martini release also provided us with the opportunity to improve the existing phosphoinositide models, such as PI(4,5)P_2[30]. These improved models were developed independently from the previous martini 2 topologies, with a focus on more accurately reproducing phosphoinositide behaviour seen in both atomistic simulations and experimental studies, including, among others, accurate reproduction of cation-mediated phosphoinositide aggregation.

In this work, we are revisiting our previous Martini CG MD work regarding the impact of acyl-chain composition on the biophysical properties of PI(4,5)P_2 and PI(4,5)P_2 nanodomains, using the improved Martini 3 force field as well as the newly developed phosphoinositide topologies. Focus was given to the differences between the two forcefields. Overall, the major findings observed with the previous Martini 2 simulations were well preserved in the new Martini 3 model, both in good agreement with experimental findings.

The key observation that increasing saturation yields more ordered and structured nanodomains is well recovered in the Martini 3 model, showcasing the importance of the canonical (18:0 20:4) PI(4,5)P_2 composition for the maintenance of membrane fluidity within PI(4,5)P_2 nanodomains. This increase in order with acyl-chain saturation, has been shown to culminate in the formation of gel-like PI(4,5)P_2 nanodomains for fully saturated compositions, which Martini 2 has struggled in reproducing. With Martini 3 this was improved, inducing the formation of stable gel-like domains at substantially higher temperatures, in agreement with experimental results. Slight differences are observed on the extent of association of PIP2 nanodomains with negatively curved membrane regions, but they are still well within the experimental observations. This work not also helps validate the new and improved Martini 3 model for biophysical studies of PI(4,5)P_2 and PI(4,5)P_2 nanodomains, but also shows that Martini 2 continues to be a very solid and accurate model within the CG resolution.

METHODS

The Martini 3 CG model for biomolecular simulations was employed throughout this study[29]. The parameterization and development of the PI(4,5)P_2 has been detailed by us elsewhere[30], and the parameters are freely available from https://github.com/MeloLab/PhosphoinositideParameters. All other topologies were obtained directly from the http://cgmartini.nl website[31]. All simulations were run using the GROMACS simulation package version 2020[32]. The membrane systems were built and solvated using the *insane.py* CG building tool[33]. Counterions were added to neutralize the system as necessary and 140 mM NaCl was added on top of that to all systems. When required, Ca^{2+} was added at a 2.5 Ca^{2+}:PI(4,5)P_2 ratio by replacing water particles and maintaining the system charge neutral by adding Cl^- counterions.

Nonbonded interactions were cutoff at 1.1 nm and Coulombic interactions were treated using reaction-field electrostatics[34] with a dielectric constant of 15 and an infinite reaction-field dielectric constant. The particle neighbor list was updated using the Verlet list scheme. The v-rescale thermostat[35] was used with a coupling time of 4.0 ps to maintain the temperature at 310 K (unless other temperature is specified). Constant pressure was semi-isotropically coupled to 1.0 bar using the Parrinello-Rahman barostat[36] with a relaxation time of 16.0 ps. After initial energy minimization and pressure/temperature equilibration runs, simulations were run at a 20 fs time step.

Three types of systems were simulated:

— *large* systems consisting of approximately 7000 lipids at a 90:10 POPC:PI(4,5)P_2 molar ratio in a 45 × 45 × 13 nm^3 box were simulated for at least 20 µs. Simulations were ran for each composition both in the presence and absence of Ca^{2+}. For these systems only, membrane undulations were limited by applying a weak (200 kJ mol^{-1} nm^{-2}) flat-bottomed potential in *z*, restraining the glycerol beads of all lipids to remain within a 2.0 nm vertical distance of the simulation box center.

— *asymmetric* membrane systems, consisting of either approximately 7000 lipids at a 90:10 POPC:PI(4,5)P_2 molar ratio in a 45 × 45 × 13 nm^3 box, or approximately 3500 lipids, in a 50 × 20 × 13 nm^3 box, were simulated for at least 10 µs. Their top membrane layer consists solely of POPC whilst the bottom layer consists of a 90:10 POPC:PI(4,5)P_2 mixture. The total number of lipids in each lipid layer was adjusted, matching the top and bottom leaflet areas, thereby avoiding the introduction of tension. To this end, the area per lipid of each layer's composition was calculated from smaller symmetric simulations ran for each of the compositions, both in the presence and absence of Ca^{2+}.

— *gel-phase probing* systems, consisting of approximately 750 lipids in a 15 × 15 × 13 nm^3. These systems were run at multiple temperatures (280 K, 290 K, 300 K, 310 K) and at multiple DPPC/PI(4,5)P_2 molar ratios (10:90, 50:50, 90:10). PI(4,5)P_2 systems were run both in the presence and absence of calcium. The systems ran at the highest temperature were built using *insane.py*, like previously described. The remaining lower temperature systems were started from the last frame of the highest temperature simulation, re-equilibrated and ran at the appropriate temperature.

All simulations were analyzed making use of in-house developed Python3 programs using the MDAnalysis package[37]. We also used the IPython[38], NumPy[39], SciPy[40], scikit-learn[41] and matplotlib[42] packages for scientific computing in Python. Visualization and rendering of the simulations was

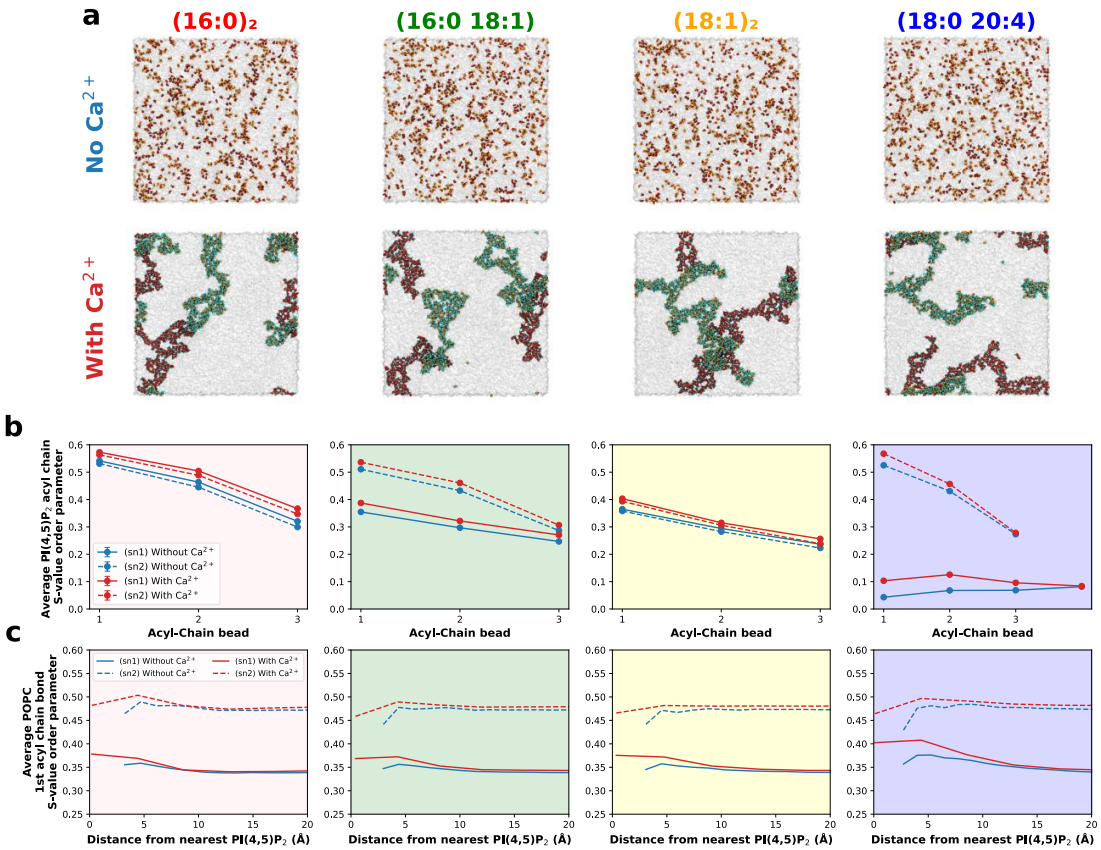


Figure 1. Impact of acyl-chain composition on PI(4,5)P_2 nanodomain biophysical properties. (a) Final snapshots of the large membrane patches composed of 90:10 POPC:PI(4,5)P_2, for each of the four acyl-chain compositions studied in the presence and absence of Ca^{2+}. Inositol headgroup beads are depicted in gray, headgroup phosphate beads in orange, phosphodiester beads in red and Ca^{2+} ions in blue. Bulk POPC lipids are represented by the gray surface. (b) PI(4,5)P_2 acyl-chain S-value order parameter recovered for all acyl-chain bead pairs, for each studied composition in the presence (red) and absence (blue) of Ca^{2+}. (c) POPC's first acyl-chain bond S-Value dependence on the distance from the nearest PI(4,5)P_2 lipid, in the presence and absence of Ca^{2+}. Distance between lipids is calculated from the phosphodiester (PO4) bead positions.

performed with the molecular graphics viewer VMD[43]. Unless otherwise stated, the last 2 µs of each simulation were used for analysis. See the Supplementary Methods section of the Supporting Information for details on the analysis methods.

RESULTS AND DISCUSSION

Three compositions were chosen as representative of the acyl chain spectrum seen *in vivo*: 1-stearoyl-2-arachidonyl (18:0 20:4); 1,2-dioleoyl (18:1)$_2$; and 1,2-dipalmitoyl (16:1)$_2$, which also match those employed in our previous Martini 2 study[27]. In the same work, experimental biophysical methods showed that in the absence of Ca^{2+}, no substantial differences were observed between PI(4,5)P_2 lipids with these acyl chain compositions. However, once nanodomain formation was induced by Ca^{2+} a significant increase in membrane order was detected within them. This was observed to impact all acyl chain compositions, with increasing levels of acyl chain saturation yielding progressively more ordered nanodomains, and where the fully saturated species stood out by assembling into gel-like nanodomains[27]. Our Martini 2 CG MD simulations in that work aimed at recovering greater structural information on these phenomena, focusing on large 50 × 50 nm^2 membrane patches containing 10% PI(4,5)P_2, in the presence and absence of Ca^{2+}. Analysis results largely agreed with experimental biophysical measurements.

In this work, we revisit these simulations, by performing them using the recently released Martini 3 forcefield[29], alongside phosphoinositide headgroup models we recently adapted and improved specifically for Martini 3[30]. We also expand on the previous work by including an additional acyl-chain composition: 1-palmitoyl-2-oleoyl (16:0 18:1). As before, large membrane patches containing 10% PI(4,5)P_2 were simulated for each of the four acyl chain compositions, in the presence and absence of Ca^{2+}, for at least 20 µs.

Along this section, the parallels and differences between the results obtained with the two Martini versions are described. The last two sub-sections then puts these comparisons in the context of each model's features, and of our understanding of PI(4,5)P_2 domain formation and impact.

Acyl-chain impact on the morphology and order of Ca^{2+}-induced PI(4,5)P_2 nanodomains. We observed no

substantial differences in size and shape of the Ca^{2+}-induced domains between the different compositions (Figure 1a). All compositions formed similarly sized nanodomains with an irregular lobed morphology. In the absence of Ca^{2+}, no composition aggregated in any significant manner. Both size and shape are well in agreement with those recovered from Martini 2 simulations. With Ca^{2+}-induced clustering, we also observe a decrease in area per lipid (ApL), and an increase in membrane thickness (Supplementary Figure S1 A and B), in agreement with Martini 2[27] and atomistic simulation results[44,45].

Regarding the impact of the acyl chain composition on the biophysical properties of PI(4,5)P_2 nanodomains, we observe a consistent increase in acyl chain order for the four compositions tested, as seen by the increase in the S-value parameter (Figure 1B). For the more saturated species ($(16:0)_2$ and (16:0 18:1)), this increase in order is consistently transmitted along both acyl-chains, while for the more unsaturated species ($(18:1)_2$ and (18:0 20:4)) ordering is more significant for the beads closest to the headgroup. These effects are in agreement with the Martini 2 results, which were already in good agreement with the experimental biophysical measurements[27].

It is worth noting that the Ca^{2+}-induced increase in membrane order for the mixed saturated-polyunsaturated PI(4,5)P_2 (18:0 20:4) is slightly larger with the Martini 3 model than with Martini 2. With Martini 2, the stearoyl chain of this species was barely affected by aggregation, but with Martini 3 the first beads of the chain become significantly more ordered. The increase in order of the arachidonoyl chain is also slightly larger with Martini 3. As previously mentioned, this increase is still well dispersed along the acyl chains, showcasing the role of the (18:0 20:4) composition as a fluidity buffer.

The effect of these nanodomains on the order of the surrounding lipids was also probed. With the Martini 3 model, all compositions yielded nanodomains that increased the order of the surrounding POPC within 10 Å (Figure 1C). This result contrasts with that of Martini 2, where the poly- and monounsaturated compositions led to a decrease in the order of the surrounding POPC, and only the fully saturated composition increased the order of the microenvironment surrounding the nanodomains.

Impact of Ca^{2+}-induced PI(4,5)P_2 nanodomains on membrane curvature and flexibility. With our previous Martini 2 model simulations, we were surprised to observe that the PI(4,5)P_2 nanodomains were strongly unregistered (decoupled) across the bilayer[27]. This led us to hypothesize that the decoupling could be the result of local membrane curvature tension caused by Ca^{2+}-induced nanodomains. This was confirmed once we ran simulations without curvature restrictions and observed the association of PI(4,5)P_2 nanodomains with markedly negative curvature undulations[27].

In the Martini 3 simulations performed in this work, we do not observe any significant interleaflet decoupling between PI(4,5)P_2 nanodomains in the presence of Ca^{2+}(Supplementary Figure S2). In fact, most of the acyl chain compositions ((18:0 20:4), $(18:1)_2$ and $(16:1)_2$) show very slight registration across leaflets. This observation hinted that PI(4,5)P_2 curvature effects in Martini 3 may be substantially weaker than those observed with Martini 2.

To characterize curvature effects, 45 × 45 nm^2 asymmetric membrane patches, containing 10 mol% PI(4,5)P_2 in the bottom leaflet, were simulated for up to 10 μs, for all acyl-chain compositions, both in the presence and absence of Ca^{2+}. Regardless of the presence of Ca^{2+}, and contrary to what had been observed with Martini 2, only minimal stabilization of membrane buckling was observed, for all the simulated systems (Figure 2). In addition, Martini 3 PI(4,5)P_2 associated only very weakly with regions of negative curvature, independently of acyl-chain composition, with little difference in behavior between Ca^{2+}-induced nanodomains and monodisperse PI(4,5)P_2 in the absence of Ca^{2+}. The observation of little to no curvature preference is relevant because PI(4,5)P_2, at least when monodisperse, is thought to promote *positive* curvature, owing to its large, phosphorylated inositol headgroup. Other Martini 2 CG studies have also observed that, when membrane curvature is induced, monodisperse PI(4,5)P_2 (in the absence of Ca^{2+}) becomes enriched in negatively curved regions[46]. Analyses of curvature preference were repeated with 20 × 50 nm^2 asymmetric membrane patches, containing 10 mol% PI(4,5)P_2 in the bottom leaflet, to parallel simulations carried out with Martini 2[27]. Also no PI(4,5)P_2 preference for negative curvature was observed in these simulations (Supplementary Figure S3), confirming that Martini 3 PI(4,5)P_2 models are less susceptible to lipid sorting by curvature than Martini 2 ones.

Since a different curvature preference/sorting of Martini 3 PI(4,5)P_2 could be a result of an increase in membrane stiffness, we also determined the bending rigidity, K_c, of our square membrane patch systems, for all acyl chain compositions, as well as for a system containing 100 % POPC, in the presence and absence of Ca^{2+} (Supplementary Figure S4). Comparing the K_c values recovered for our 100 % POPC systems (17.5 ± 2.4 k_BT, Supplementary Figure S4) with those obtained with Martini 2 (29.8 ± 5.5 k_BT[23]), we observe a decrease in bending rigidity, hinting at a more flexible membrane. However, both values are still well within the range of experimentally determined values for pure POPC (5.8 – 49 k_BT[23]). The addition of 10% PI(4,5)P_2 to the inner leaflet, regardless of the acyl chain composition, did not significantly impact the bending rigidity. Addition of Ca^{2+}, and the concomitant formation of PI(4,5)P_2 nanodomains, led only to minor increases in K_c.

Characterization of $(16:0)_2$ PI(4,5)P_2 gel-phase nanodomains. As PI(4,5)P_2 acyl chain saturation increases, it clusters in progressively more ordered nanodomains, culminating in the formation of gel-like PI(4,5)P_2 nanodomains when using the $(16:0)_2$ acyl chain composition[27]. Given the necessary proximity of the large, negatively charged, phosphorylated inositol headgroups, the formation of these structures is quite surprising and can greatly impact the biophysical properties within nanodomains. These changes are likely to also affect membrane events associated with PI(4,5)P_2 as well as the PI(4,5)P_2-protein interactions, especially in cases where more saturated PI(4,5)P_2 acyl chain compositions are more prevalent[47–49].

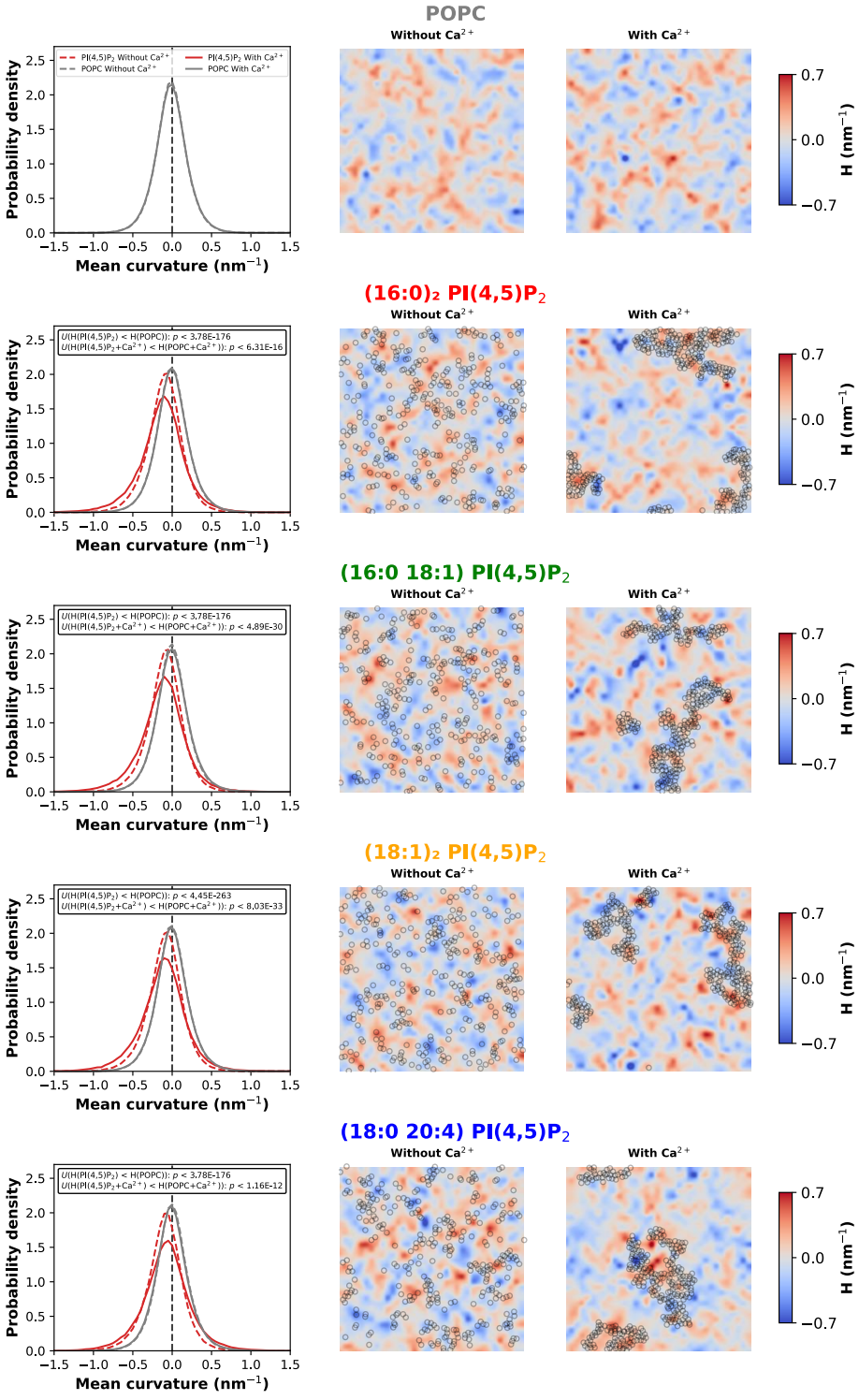


Figure 2. Association of Ca^{2+}-induced PI(4,5)P_2 nanodomains with membrane curvature fluctuations. Histogram of mean curvature associated with either PI(4,5)P_2 (red) or POPC (grey), from asymmetric membrane simulations containing 10 mol% PI(4,5)P_2 in the inner leaflet, for each of the four studied acyl chain compositions in the presence (solid line) and absence (dashed line) of Ca^{2+}. For each system, top view snapshots are shown, colored by the local curvature value. PI(4,5)P_2 phosphodiester beads are represented by the black circles. Mann-Whitney U tests were used to determine the statistical significance between PI(4,5)P_2 and POPC associated mean curvature.

5

Using Martini 2, we struggled to observe the formation of these gel-like nanodomains. Martini 2 gel-like $(16:0)_2$ PI(4,5)P_2 nanodomains did assemble, albeit at higher PI(4,5)P_2 concentrations (50 mol%) and at lower temperatures (280 K) than what we observed experimentally (T_m = 45 °C or 318 K for 1-10 mol% PI(4,5)P_2[27]). Indeed, lipid gel phases in Martini 2 are known not only to underestimate gel-fluid lipid phase transition temperatures — as reported for 1,2-dipalmitoyl-sn-glycero-3-phosphocholine (DPPC)[18] — but also to be especially sensitive to small amounts of disorder-inducing membrane components, such as POPC[50].

Here, we revisit the formation of $(16:0)_2$ PI(4,5)P_2 gel-phase nanodomains using Martini 3. We simulated 15 × 15 symmetric membrane patches containing multiple $(16:0)_2$ PI(4,5)P_2:POPC molar ratios, at temperatures ranging from 280 to 310 K. These systems were run both in the presence and absence of Ca^{2+}. For comparison, these systems were also run with the canonical gel forming lipid, DPPC, instead of PI(4,5)P_2. The formation of gel phases was followed via the system hexagonality, ApL, as well as the average acyl-chain order (Figure 3).

We observe that DPPC gel-phase formation only occurs for systems containing 90:10 DPPC:POPC molar ratio, at temperatures under 290 K (Figure 3 and Supplementary Figure S6). The Martini 3 DPPC gel-fluid phase transition temperature at this molar ratio occurs between 290 and 300 K, which is in agreement with the experimental phase transition temperature for 90:10 DPPC:POPC[51] and already a substantial improvement over the 270 – 280 K transition temperature in Martini 2 simulations of pure DPPC[18]. The DPPC:POPC experimental phase diagram expects the coexistence of gel and fluid phases at 50:50 DPPC:POPC mol ratio at temperatures under 300 K[51], which is still not correctly recovered by Martini 3. Altogether, this quick characterization of the DPPC gel forming behavior shows that the Martini 3 forcefield has substantially improved the underestimation of the gel-fluid transition temperature but is still sensitive to disorder inducing membrane components.

When using $(16:0)_2$ PI(4,5)P_2 as the gel-forming lipid we do not observe the formation of a gel phase in the absence of Ca^{2+} (Figure 3 and Supplementary Figure S7). At 90 mol% of PI(4,5)P_2 we observe signs of bilayer instability, such as abnormally low values of acyl chain order and high values of ApL. This is in agreement with experimental findings that show that, in the absence of other lipids, phosphoinositides such as PI(4,5)P_2 are prone to form micellar structures[52].

In the presence of Ca^{2+} (Figure 3 and Supplementary Figure S8), we observe the aggregation of PI(4,5)P_2 into nanodomains, and at 90:10 PI(4,5)P_2:POPC mol% ratio, gel domains are formed across the temperature range tested, covering almost the entire simulated membrane patch. At 50:50 PI(4,5)P_2:POPC ratio, gel domains are observed across the temperature range tested in coexistence with fluid POPC-rich phases. $(16:0)_2$ PI(4,5)P_2 shows substantial gel character, with most POPC lipids retaining a fair amount of fluidity. At a 10:90 $(16:0)_2$ PI(4,5)P_2:POPC mol% ratio no gel-phase behavior is observed.

Overall, while Martini 3 still does not reproduce the formation of $(16:0)_2$ PI(4,5)P_2 gel phases at low PI(4,5)P_2 concentrations (<10 mol%), the behavior at higher mol ratios is much improved. Gel domains are stable in the range of temperatures tested, at least up to 310 K, in agreement with experimental results[27]. At lower mol ratios, $(16:0)_2$ PI(4,5)P_2 gel-like nanodomains are observed in coexistence with more fluid POPC-rich phases, demonstrating an increased tolerance for disorder-inducing membrane components when compared to DPPC. The Martini 3 model thus shows promising indications that it can be successfully used to study lipid bilayer phase change behavior with greater accuracy than its predecessor.

Model properties behind Martini 2 and 3 differences. One of the main aspects that may account for the change we observed between Martini 2 and Martini 3 behavior is the new PI(4,5)P_2 headgroup parameterization: in the Martini 3 models we developed, the PI(4,5)P_2 headgroup has a more restricted conformational freedom, that better reproduces behavior observed in atomistic references. This may result in a lower ability of the inositol headgroup to undergo the reorientation required for being more compacted in negative curvatures. This is likely more relevant in PI(4,5)P_2 aggregates, where PI(4,5)P_2 headgroups must pack against one another. At the same time, the Martini 3 PI(4,5)P_2 reparameterization has made use of the more detailed Martini 3 particle type landscape to better describe headgroup–headgroup interactions[30], to match the aggregation propensities of different PI(4,5)P_2 species in the presence of Ca^{2+}. These improved interactions result in larger headgroup packing distances (Supplementary figure S9), which are likely less conducive to negative curvatures.

Another factor likely relevant to the observed differences is the new set of interactions affecting all glycerolipids — not just the phosphoinositides — under Martini 3[29]: one of the two beads of glycerol is now mapped to an S-type particle (with a smaller Lennard-Jones sigma parameter), several alkyl interbead distances and force constants have been adjusted and the bead types underlying the phosphodiester (PO4) and phosphatidylcholine (NC3) beads have been adjusted. It is reasonable to expect the lateral pressure profile of lipids to have changed, which could both account for different curvature behavior, as well as for the differences in ordering of the acyl chains — also in line with the better phase agreement of the binary PC mixtures. It will be interesting to see how the continuous development of the standard Martini 3 lipid acyl chains, as well as of the lipid headgroups refines this behavior moving forward.

Impact on the understanding of PI(4,5)P_2 behavior. CG MD has previously been employed to characterize the mechanisms of PI(4,5)P_2 curvature sensing and/or aggregation. It is therefore important to assess how the changes brought about by new CG models affect our view of these processes. Some of the differences we observe in this work, using Martini 3, reflect PI(4,5)P_2 features more in line with experimentally known behavior — for instance, the better phase representation of $(16:0)_2$ PI(4,5)P_2 aggregates. Such improvements are very much welcome to have in a computational model, but are not very informative on their own, since they merely match aspects that are already known. After all, a model only becomes truly useful when predicting behavior for which there's no other authoritative source of information. In this sense, the main behavior change we report here is that PI(4,5)P_2 seems to not have any strong

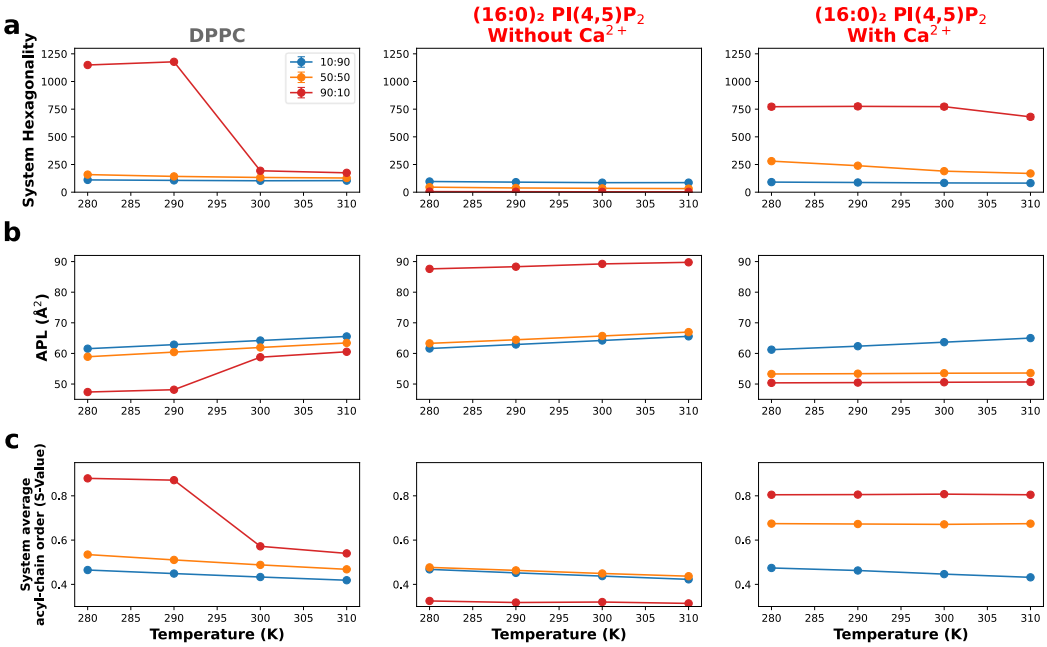


Figure 3. (16:0)$_2$ PI(4,5)P_2 / DPPC gel-phase formation study. (a) System wide hexagonality, (b) ApL and (c) system average acyl-chain order recovered from gel-phase probing systems. These systems consisted of either DPPC or PI(4,5)P_2 in POPC at multiple mol% ratios (10:90, 50:50, 90:10), which were simulated at multiple temperatures (280 K, 290 K, 300 K, 310 K). PI(4,5)P_2 systems were run both in the presence and absence of Ca^{2+}.

preference for negative curvature, when we and others had observed otherwise with Martini 2[27,46]. On the one hand, this is closer to an expectation of positive curvature preference by PI(4,5)P_2 due to the bulkiness of its headgroup; on the other hand, curvature neutrality explains less readily the purported participation of PI(4,5)P_2 in endocytic mechanisms — for which negative curvature is a necessity[53–55].

Alongside curvature preference, our observations also indicate a much lower membrane bending rigidity using Martini 3. This is not specific to PI(4,5)P_2-containing membranes and falls, together with Martini 2 K_c values, within a large range of reported experimental measurements. Another difference was also observed in how the increased order of the PI(4,5)P_2 nanodomains impacted their surrounding microenvironment. With Martini 3, PI(4,5)P_2 nanodomains of any acyl-chain composition led to an increase in order of the surrounding POPC, contrasting with Martini 2 where the unsaturated compositions led to a decrease in order and only the fully saturated composition ordered its microenvironment. In this case, the Martini 3 observations are more consistent, in that an ordered domain seems likely to also order its surroundings, rather than to disorder them.

From the considerations over each individual discrepancy above, and given the overall increase in accuracy from Martini 2 to Martini 3, we are inclined to better trust the results obtained with the latter. However, force field development is a continuous endeavor, and our views of PI(4,5)P_2 behavior may again be updated as new models are put forth. Likewise, our hope is that our results spur the undertaking of specific experiments that test our conclusions by producing data that better describes PI(4,5)P_2 behavior, which, in turn, can be used to further refine our CG models.

CONCLUSION

We revisited the impact of acyl-chain composition[27] on the biophysical properties of PI(4,5)P_2 and PI(4,5)P_2 nanodomains, using the improved Martini 3 CG MD force field[29] as well as our recently developed Martini 3 phosphoinositide topologies[30]. We compared our results with those of our previous Martini 2 work on the same topic. The finding that increasing saturation results in more ordered and structured nanodomains is well recovered in the Martini 3 model, highlighting the importance of the polyunsaturated (18:0 20:4) PI(4,5)P_2 composition for the maintenance of membrane fluidity within PI(4,5)P_2 nanodomains. Representativity was extended to the formation of gel-like (16:0)$_2$ PI(4,5)P_2 nanodomains with which the Martini 2 model had struggled. Moreso, the Martini 3 models were shown to have almost no specificity for negatively curved membrane regions, unlike the Martini 2 models. The key differences between the Martini 2 and Martini 3 models were discussed in depth, considering both the newer models' accuracy and parameterization strategy, as well as the known experimental PI(4,5)P_2 mechanistic behavior.

Overall, the major findings observed with the previous Martini 2 simulations were well preserved in the new Martini 3 model, and both in agreement with experimental findings. Not only does this help validate the new and improved Martini 3 model for biophysical studies of PI(4,5)P_2 and PI(4,5)P_2 nanodomains, but it also shows that Martini 2 continues to be a very solid and accurate model within the CG

resolution. This serves as a reminder that one should not immediately dismiss the thousands of existing Martini 2 simulation results based on the release of the new model. Nonetheless, it will be interesting to see how the continuous development of the Martini 3 model in the near future further extends accuracy and applicability, especially with the refinement of the lipid acyl-chains and glycerol backbone, as well as the development of other membrane components, such as cholesterol.

ASSOCIATED CONTENT

Extended Methods and Materials providing details behind the simulation analysis, as well as additional analysis results on the biophysical properties of PI(4,5)P_2 and PI(4,5)P_2 nanodomains, curvature association of PI(4,5)P_2, membrane rigidity, and gel-phase formation. Other data are available from the corresponding author upon request.

AUTHOR INFORMATION

Corresponding Author

Manuel N. Melo — Instituto de Tecnologia Química e Biológica António Xavier, Universidade Nova de Lisboa, Oeiras 2780-157, Portugal; orcid.org/0000-0001-6567-0513; m.n.melo@itqb.unl.pt

Author Contributions

All authors designed the simulation systems. L.B.A. and M.N.M. designed the analyses. L.B.A. performed simulations and analyses. All authors contributed to manuscript writing and revision and have given approval to the final version of the manuscript.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENT

L.B.A. thanks the Medical Biochemistry and Biophysics Doctoral Program (M2B-PhD) and Fundação para a Ciência e a Tecnologia – Ministério da Ciência, Tecnologia e Ensino Superior (FCT-MCTES, Portugal) for PhD fellowship PD/BD/137492/2018. M.N.M. thanks FCT-MCTES for the "scientific employment stimulus" program CEECIND/04124/2017. M.N.M. also acknowledges FCT for funding project MOSTMICRO-ITQB, with references UIDB/04612/2020 and UIDP/04612/2020.

REFERENCES

(1) Di Paolo, G.; De Camilli, P. Phosphoinositides in Cell Regulation and Membrane Dynamics. *Nature* **2006**, *443* (7112), 651–657.

(2) Mandal, K. Review of PIP2 in Cellular Signaling, Functions and Diseases. *Int. J. Mol. Sci.* **2020**, *21* (21), 1–20.

(3) Wenk, M. R.; De Camilli, P. Protein-Lipid Interactions and Phosphoinositide Metabolism in Membrane Traffic: Insights from Vesicle Recycling in Nerve Terminals. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101* (22), 8262–8269.

(4) Posor, Y.; Eichhorn-Grünig, M.; Haucke, V. Phosphoinositides in Endocytosis. *Biochimica et Biophysica Acta - Molecular and Cell Biology of Lipids*. Elsevier B.V. June 1, 2015, pp 794–804.

(5) Levin, R.; Grinstein, S.; Schlam, D. Phosphoinositides in Phagocytosis and Macropinocytosis. *Biochimica et Biophysica Acta - Molecular and Cell Biology of Lipids*. Elsevier B.V. June 1, 2015, pp 805–823.

(6) Martin, T. F. J. PI(4,5)P2-Binding Effector Proteins for Vesicle Exocytosis. *Biochimica et Biophysica Acta - Molecular and Cell Biology of Lipids*. Elsevier B.V. June 1, 2015, pp 785–793.

(7) dos Remedios, C. G.; Nosworthy, N. J. The Role of PIP2 in Actin, Actin-Binding Proteins and Disease. In *Actin-Binding Proteins and Disease*; Springer New York, 2008; pp 290–297.

(8) Berridge, M. J. Inositol Trisphosphate and Diacylglycerol: Two Interacting Second Messengers. *Annu. Rev. Biochem.* **1987**, *56* (1), 159–193.

(9) Rauch, M. E.; Ferguson, C. G.; Prestwich, G. D.; Cafiso, D. S. Myristoylated Alanine-Rich C Kinase Substrate (MARCKS) Sequesters Spin-Labeled Phosphatidylinositol 4,5-Bisphosphate in Lipid Bilayers. *J. Biol. Chem.* **2002**, *277* (16), 14068–14076.

(10) Wang, J.; Gambhir, A.; Hangyás-Mihályné, G.; Murray, D.; Golebiewska, U.; McLaughlin, S. Lateral Sequestration of Phosphatidylinositol 4,5-Bisphosphate by the Basic Effector Domain of Myristoylated Alanine-Rich C Kinase Substrate Is Due to Nonspecific Electrostatic Interactions. *J. Biol. Chem.* **2002**, *277* (37), 34401–34412.

(11) Sarmento, M. J.; Coutinho, A.; Fedorov, A.; Prieto, M.; Fernandes, F. Ca2+ Induces PI(4,5)P2 Clusters on Lipid Bilayers at Physiological PI(4,5)P2 and Ca2+ Concentrations. *Biochim. Biophys. Acta* **2014**, *1838* (3), 822–830.

(12) Wen, Y.; Vogt, V. M.; Feigenson, G. W. Multivalent Cation-Bridged PI(4,5)P2 Clusters Form at Very Low Concentrations. *Biophys. J.* **2018**, *114* (11), 2630–2639.

(13) Marrink, S. J.; Risselada, H. J.; Yefimov, S.; Tieleman, D. P.; De Vries, A. H. The MARTINI Force Field: Coarse Grained Model for Biomolecular Simulations. *J. Phys. Chem. B* **2007**, *111* (27), 7812–7824.

(14) Alessandri, R.; Souza, P. C. T.; Thallmair, S.; Melo, M. N.; de Vries, A. H.; Marrink, S. J. Pitfalls of the Martini Model. *J. Chem. Theory Comput.* **2019**, *15* (10), 5448–5460.

(15) Marrink, S. J.; Corradi, V.; Souza, P. C. T.; Ingólfsson, H. I.; Tieleman, D. P.; Sansom, M. S. P. Computational Modeling of Realistic Cell Membranes. *Chemical Reviews*. American Chemical Society May 8, 2019, pp 6184–6226.

(16) Cherniavskyi, Y.; Tieleman, D. P. 19. Simulations of Biological Membranes with the Martini Model. In *Characterization of Biological Membranes*; De Gruyter, 2019; pp 551–568.

(17) Baoukina, S.; Rozmanov, D.; Tieleman, D. P. Composition Fluctuations in Lipid Bilayers. *Biophys. J.* **2017**, *113* (12), 2750–2761.

(18) Marrink, S. J.; Risselada, J.; Mark, A. E. Simulation of Gel Phase Formation and Melting in Lipid Bilayers Using a Coarse Grained Model. *Chem. Phys. Lipids* **2005**, *135* (2), 223–244.

(19) Arnarez, C.; Webb, A.; Rouvière, E.; Lyman, E. Hysteresis and the Cholesterol Dependent Phase Transition in Binary Lipid Mixtures with the Martini Model. *J. Phys. Chem. B* **2016**, *120* (51), 13086–13093.

(20) Hu, M.; Briguglio, J. J.; Deserno, M. Determining the Gaussian Curvature Modulus of Lipid Membranes in Simulations. *Biophys. J.* **2012**, *102* (6), 1403–1410.

(21) Hoiles, W.; Gupta, R.; Cornell, B.; Cranfield, C.; Krishnamurthy, V. The Effect of Tethers on Artificial Cell Membranes: A Coarse-Grained Molecular Dynamics Study. *PLoS One* **2016**, *11* (10), e0162790.

(22) Baoukina, S.; Marrink, S. J.; Tieleman, D. P. Molecular Structure of Membrane Tethers. *Biophys. J.* **2012**, *102* (8), 1866–1871.

(23) Fowler, P. W.; Hélie, J.; Duncan, A.; Chavent, M.; Koldsø, H.; Sansom, M. S. P. Membrane Stiffness Is Modified by Integral Membrane Proteins. *Soft Matter* **2016**, *12* (37), 7792–7803.

(24) Stansfeld, P. J.; Hopkinson, R.; Ashcroft, F. M.; Sansom, M. S. P. PIP 2 -Binding Site in Kir Channels: Definition by Multiscale Biomolecular Simulations. *Biochemistry* **2009**, *48* (46), 10926–10933.

(25) Ingólfsson, H. I.; Carpenter, T. S.; Bhatia, H.; Bremer, P. T.; Marrink, S. J.; Lightstone, F. C. Computational Lipidomics of the Neuronal Plasma Membrane. *Biophys. J.* **2017**, *113* (10), 2271–2280.

(26) Ingólfsson, H. I.; Melo, M. N.; Van Eerden, F. J.; Arnarez, C.; Lopez, C. A.; Wassenaar, T. A.; Periole, X.; De Vries, A. H.; Tieleman, D. P.; Marrink, S. J. Lipid Organization of the Plasma Membrane. *J. Am. Chem. Soc.* **2014**, *136* (41), 14554–14559.

(27) Borges-Araújo, L.; Domingues, M. M.; Fedorov, A.; Santos, N. C.; Melo, M. N.; Fernandes, F. Acyl-Chain Saturation Regulates the Order of Phosphatidylinositol 4,5-Bisphosphate Nanodomains. *Commun. Chem.* **2021**, *4* (1), 1–13.

(28) Manni, M. M.; Tiberti, M. L.; Pagnotta, S.; Barelli, H.; Gautier, R.; Antonny, B. Acyl Chain Asymmetry and Polyunsaturation of Brain Phospholipids Facilitate Membrane Vesiculation without Leakage. *Elife* **2018**, *7*.

(29) Souza, P. C. T. T.; Alessandri, R.; Barnoud, J.; Thallmair, S.; Faustino, I.; Grünewald, F.; Patmanidis, I.; Abdizadeh, H.; Bruininks, B. M. H. H.; Wassenaar, T. A.; Kroon, P. C.; Melcr, J.; Nieto, V.; Corradi, V.; Khan, H. M.; Domański, J.; Javanainen, M.; Martinez-Seara, H.; Reuter, N.; Best, R. B.; Vattulainen, I.; Monticelli, L.; Periole, X.; Tieleman, D. P.; de Vries, A. H.; Marrink, S. J. Martini 3: A General Purpose Force Field for Coarse-Grained Molecular Dynamics. *Nat. Methods* **2021**, *18* (4), 382–388.

(30) Borges-Araújo, L.; Souza, P. C. T.; Fernandes, F.; Melo, M. N. Improved Parameterization of Phosphatidylinositide Lipid Headgroups for the Martini 3 Coarse-Grain Force Field. *J. Chem. Theory Comput.* **2021**, acs.jctc.1c00615.

(31) Martini Coarse Grain Forcefield for Biomolecules http://cgmartini.nl/.

(32) 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.

(33) Wassenaar, T. A.; Ingólfsson, H. I.; Böckmann, R. A.; Tieleman, D. P.; Marrink, S. J. Computational Lipidomics with Insane: A Versatile Tool for Generating Custom Membranes for Molecular Simulations. *J. Chem. Theory Comput.* **2015**, *11* (5), 2144–2155.

(34) De Jong, D. H.; Baoukina, S.; Ingólfsson, H. I.; Marrink, S. J. Martini Straight: Boosting Performance Using a Shorter Cutoff and GPUs. *Comput. Phys. Commun.* **2016**, *199*, 1–7.

(35) Bussi, G.; Donadio, D.; Parrinello, M. Canonical Sampling through Velocity Rescaling. *J. Chem. Phys.* **2007**, *126* (1), 014101.

(36) Parrinello, M.; Rahman, A. Polymorphic Transitions in Single Crystals: A New Molecular Dynamics Method. *J. Appl. Phys.* **1981**, *52* (12), 7182–7190.

(37) Michaud-Agrawal, N.; Denning, E. J.; Woolf, T. B.; Beckstein, O. MDAnalysis: A Toolkit for the Analysis of Molecular Dynamics Simulations. *J. Comput. Chem.* **2011**, *32* (10), 2319–2327.

(38) Pérez, F.; Granger, B. E. *IPython: A System for Interactive Scientific Computing*; 2007; Vol. 9.

(39) Harris, C. R.; Millman, K. J.; van der Walt, S. J.; Gommers, R.; Virtanen, P.; Cournapeau, D.; Wieser, E.; Taylor, J.; Berg, S.; Smith, N. J.; Kern, R.; Picus, M.; Hoyer, S.; van Kerkwijk, M. H.; Brett, M.; Haldane, A.; del Río, J. F.; Wiebe, M.; Peterson, P.; Gérard-Marchant, P.; Sheppard, K.; Reddy, T.; Weckesser, W.; Abbasi, H.; Gohlke, C.; Oliphant, T. E. Array Programming with NumPy. *Nature* **2020**, *585* (7825), 357–362.

(40) Virtanen, P.; Gommers, R.; Oliphant, T. E.; Haberland, M.; Reddy, T.; Cournapeau, D.; Burovski, E.; Peterson, P.; Weckesser, W.; Bright, J.; van der Walt, S. J.; Brett, M.; Wilson, J.; Millman, K. J.; Mayorov, N.; Nelson, A. R. J.; Jones, E.; Kern, R.; Larson, E.; Carey, C. J.; Polat, İ.; Feng, Y.; Moore, E. W.; VanderPlas, J.; Laxalde, D.; Perktold, J.; Cimrman, R.; Henriksen, I.; Quintero, E. A.; Harris, C. R.; Archibald, A. M.; Ribeiro, A. H.; Pedregosa, F.; van Mulbregt, P. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nat. Methods* **2020**, *17* (3), 261–272.

(41) Fabian, P.; Michel, V.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Vanderplas, J.; Cournapeau, D.; Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Thirion, B.; Grisel, O.; Dubourg, V.; Passos, A.; Brucher, M.; Perrot, M.; Duchesnay, É. *Scikit-Learn: Machine Learning in Python*; 2011; Vol. 12.

(42) Hunter, J. D. Matplotlib: A 2D Graphics Environment. *Comput. Sci. Eng.* **2007**, *9* (3), 90–95.

(43) Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual Molecular Dynamics. *J. Mol. Graph.* **1996**, *14* (1), 33–38.

(44) Han, K.; Gericke, A.; Pastor, R. W. Characterization of Specific Ion Effects on PI(4,5)P 2 Clustering: Molecular Dynamics Simulations and Graph-Theoretic Analysis. *J. Phys. Chem. B* **2020**, *124* (7), 1183–1196.

(45) Bradley, R. P.; Slochower, D. R.; Janmey, P. A.; Radhakrishnan, R. Divalent Cations Bind to Phosphoinositides to Induce Ion and Isomer Specific Propensities for Nano-Cluster Initiation in Bilayer Membranes. *R. Soc. Open Sci.* **2020**, *7* (5).

(46) Lin, X.; Wang, H.; Lou, Z.; Cao, M.; Zhang, Z.; Gu, N. Roles of PIP2 in the Membrane Binding of MIM I-BAR: Insights from Molecular Dynamics Simulations. *FEBS Lett.* **2018**, *592* (15), 2533–2542.

(47) Traynor-Kaplan, A.; Kruse, M.; Dickson, E. J.; Dai, G.; Vivas, O.; Yu, H.; Whittington, D.; Hille, B. Fatty-Acyl Chain Profiles of Cellular Phosphoinositides. *Biochim. Biophys. Acta* **2017**, *1862* (5), 513–522.

(48) Mujalli, A.; Chicanne, G.; Bertrand-Michel, J.; Viars, F.; Stephens, L.; Hawkins, P.; Viaud, J.; Gaits-Iacovoni, F.; Severin, S.; Gratacap, M. P.; Terrisse, A. D.; Payrastre, B. Profiling of Phosphoinositide Molecular Species in Human and Mouse Platelets Identifies New Species Increasing Following Stimulation. *Biochim. Biophys. Acta - Mol. Cell Biol. Lipids* **2018**, *1863* (9), 1121–1131.

(49) Koizumi, A.; Narita, S.; Nakanishi, H.; Ishikawa, M.; Eguchi, S.; Kimura, H.; Takasuga, S.; Huang, M.; Inoue, T.; Sasaki, J.; Yoshioka, T.; Habuchi, T.; Sasaki, T. Increased Fatty Acyl Saturation of Phosphatidylinositol Phosphates in Prostate Cancer Progression. *Sci. Rep.* **2019**, *9* (1).

(50) Carpenter, T. S.; López, C. A.; Neale, C.; Montour, C.; Ingólfsson, H. I.; Di Natale, F.; Lightstone, F. C.; Gnanakaran, S. Capturing Phase Behavior of Ternary Lipid Mixtures with a Refined Martini Coarse-Grained Force Field. *J. Chem. Theory Comput.* **2018**, *14* (11), 6050–6062.

(51) Marsh, D. *Handbook of Lipid Bilayers*, 2nd Editio.; CRC Press, 2013.

(52) Sugiura, Y. Structure of Molecular Aggregates of 1-(3-Sn-Phosphatidyl)-l-Myo-Inositol 3,4-Bis(Phosphate) in Water. *BBA - Biomembr.* **1981**, *641* (1), 148–159.

(53) Koch, M.; Holt, M. Coupling Exo- and Endocytosis: An Essential Role for PIP2 at the Synapse. *Biochim. Biophys. Acta* **2012**, *1821* (8), 1114–1132.

(54) James, D. J.; Khodthong, C.; Kowalchyk, J. A.; Martin, T. F. J. Phosphatidylinositol 4,5-Bisphosphate Regulates SNARE-Dependent Membrane Fusion. *J. Cell Biol.* **2008**, *182* (2), 355–366.

(55) Mu, L.; Tu, Z.; Miao, L.; Ruan, H.; Kang, N.; Hei, Y.; Chen, J.; Wei, W.; Gong, F.; Wang, B.; Du, Y.; Ma, G.; Amerein, M. W.; Xia, T.; Shi, Y. A Phosphatidylinositol 4,5-Bisphosphate Redistribution-Based Sensing Mechanism Initiates a Phagocytosis Programing. *Nat. Commun.* **2018**, *9* (1), 4259.

5 | Tackling the SARS-CoV-2 pandemic

In early 2020, half way through my PhD studies, the coronavirus disease 2019 (COVID-19) pandemic began. This pandemic, caused by the novel severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), began with an outbreak in Wuhan, China in December of 2019 and quickly spread across the globe. By March 2020, the world health organization (WHO) had declared it a global pandemic. The COVID-19 pandemic has led to a dramatic loss of human life worldwide, and presents a medical, social and economic challenge unprecedented in modern history. As of mid December 2021, more than 277 million cases and 5.4 million deaths had been reported, making it one of the deadliest pandemics in history.

While policymakers tried to contain and mitigate the pandemic and medical professionals fought to prevent further loss of life, the scientific community worldwide united in an attempt to find solutions to the virus. The response was exceptional, substantially mitigating the spread and impact of COVID-19: within weeks the SARS-CoV-2 genome had been fully sequenced[95]; within a few months multiple COVID-19 diagnostic methods had been developed and deployed worldwide[96]; and, within a year, several vaccines had been developed[97]. It is worth noting that this effort took place despite the enormous burden that the pandemic placed on academia. Lockdown measures limited access to research institutes and laboratories impacting the work of researchers and post-grad students, cooperation between groups was hindered and access to learning was compromised.

At the start of the pandemic, a COVID-19 taskforce was created at ITQB NOVA. Many labs redirected part of their research efforts to the fight against SARS-CoV-2 and worked, often in collaboration, on developing new diagnostics, vaccines, pharmaceuticals, and on unveiling molecular mechanisms of viral action. A sense of responsibility led me to volunteer for the cause and join in on our lab's contributions the fight. This chapter is composed by two articles which describe the work I developed for this undertaking.

Initially, our contributions focused on the design and optimization of biopharmaceuticals against SARS-CoV-2, which were performed using an in-house developed genetic algorithm (GA)-guided ROSETTA approach. Article V describes in detail the development of this approach, as well as its application to the design and optimization of an ACE2-based SARS-CoV-2 neutralizing biomolecule. With the appearance of several SARS-CoV-2 variants we also aimed to study the impact of the alpha, beta and delta variants on the conformational dynamics and ACE2 accessibility of the SARS-CoV-2 Spike (S) glycoprotein receptor

binding domain (RBD). This work, which provides important clues as to how the variants might present a substantial fitness advantage due to increased availability for ACE2 binding as well as antibody escape through epitope occlusion, is reported in Article VI.

Article V: Genetic algorithm-guided optimization of anti-SARS-CoV-2 biomolecules

Genetic algorithm-guided optimization of anti-SARS-CoV-2 biomolecules.

Luís Borges-Araújo †,⊥,‡, Diana Lousa‡, Cláudio M. Soares‡, Manuel N. Melo‡,*

† iBB-Institute for Bioengineering and Biosciences, Instituto Superior Técnico, Universidade de Lisboa, Lisbon, Portugal

⊥ Associate Laboratory i4HB—Institute for Health and Bioeconomy at Instituto Superior Técnico, Universidade de Lisboa, Lisbon, Portugal

‡ Instituto de Tecnologia Química e Biológica António Xavier, Universidade Nova de Lisboa, Av. da República, 2780-157 Oeiras, Portugal.

ABSTRACT: Coronavirus disease 2019 (COVID-19), which is caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), is a global pandemic which at the time of writing, has claimed approximately 3 million lives since its outbreak. Containment of the SARS-CoV-2 infection requires reducing viral transmission and replication, and as such the S protein, responsible for the interaction of SARS-CoV-2 with the host cells, via binding to the angiotensin-converting enzyme 2 (ACE2), becomes an obvious target for pharmaceutical action as a potential neutralizing decoy receptor. In this work, we showcase the design, optimization, and *in silico* characterization of an improved anti-SARS-CoV-2 ACE2 protein, using an in-house developed genetic-algorithm (GA) guided ROSETTA approach. The designed molecule proved to be a promising candidate for *in vitro* testing after extensive *in silico* characterization, showing significantly increased binding affinity, whilst maintaining promising indicators of fold stability. Additionally, the developed GA-guided approach proved to be a reliable and efficient method to obtain computational protein designs.

INTRODUCTION

Coronavirus disease 2019 (COVID-19), which is caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2)[1–3], is a global pandemic with mortality higher than that of seasonal influenza[4]. At the time of writing, more than 133 million cases have been reported and approximately 3 million lives have been claimed since its outbreak[5].

Infection by SARS-CoV-2 requires the fusion of viral and host cell membranes, at either the cell surface or the endosomal membrane. Like severe acute respiratory syndrome coronavirus (SARS-CoV) and the Middle East respiratory syndrome-related coronavirus (MERS-CoV), the fusion process is mediated by the viral envelope spike glycoprotein, or S protein. Upon viral attachment or uptake, host factors trigger large-scale conformational rearrangements in S protein, including a refolding step that leads directly to membrane fusion and viral entry [6–11].

Containment of the SARS-CoV-2 infection requires reducing viral transmission and replication, and as such the S protein becomes an obvious target for pharmaceutical action. Already a prime immune target, the S protein is targeted by neutralizing antibody responses and protective immunity, in contrast to most other viral proteins[12–15]. The S protein is responsible for the interaction of SARS-CoV-2 with the host cells via binding to the angiotensin-converting enzyme 2 (ACE2)[7,9,16]. It can be divided into two regions, S1 and S2, where the extra-viral S1 region contains the receptor-binding domain (RBD)[17]. The RBD is responsible for mediating the interactions between the S protein and ACE2.

ACE2 is an enzyme that is typically anchored to the cell membrane and plays a key role in the cardiovascular system by hydrolyzing angiotensin II into angiotensin (1–7)[18]. Due to its complex role as the receptor for SARS-CoV-2, ACE2 presents itself as a promising therapeutic target for COVID-19[19]. Apart from binding to ACE2, SARS-CoV-2 has also been found to favor ACE2 cleavage from the membrane, causing increased levels of soluble ACE2, reducing the cardioprotective effects of ACE2 and contributing to the increased inflammatory response[19].

Given that ACE2 binds to the S protein RBD, it is a prime template candidate for the computational design of a neutralizing anti-SARS-CoV-2 biomolecule. In this work, we showcase the design, optimization and *in silico* characterization of an improved anti-SARS-CoV-2 ACE2 protein, using an in-house developed genetic-algorithm (GA) guided ROSETTA approach. The designed molecule proved to be a promising candidate for *in vitro* testing after extensive *in silico* characterization, showing significantly increased binding affinity, whilst maintaining promising indicators of fold stability. Additionally, we present a detailed outlook on protein design methods and on the use of genetic algorithms in protein design, as well as a description of our developed algorithm. Our GA-guided design method allowed us to quickly and efficiently probe the optimization sequence

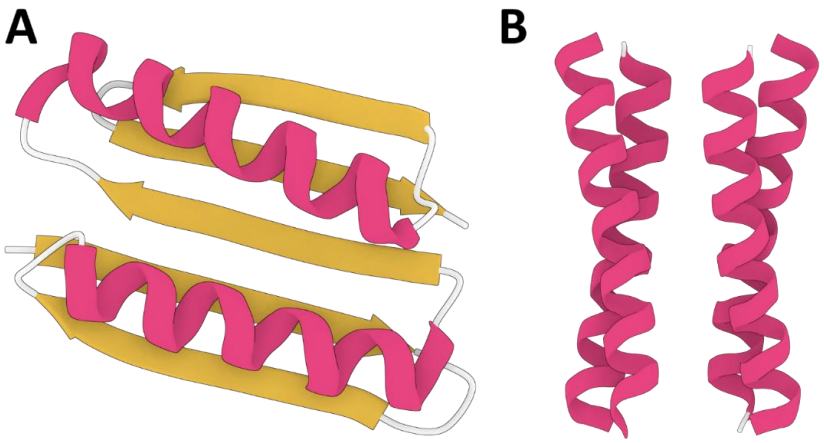


Figure 1. Examples of computational protein designs. (A) Crystal structure of the TOP7 design[27] obtained by X-ray diffraction (PDB ID: 1QYS) and (B) the structure of a transmembrane Zn^{2+} transporting four-helix bundle[30] obtained by solution NMR (PDB ID: 2MUZ).

space, whilst having full control over the process and obtaining results similar to that of the standard Rosetta design methods.

A short outlook on computational protein design. Computational protein design has been defined as *the computer-aided rational (or semi-rational) design of a protein (or part thereof) to fold to a requested structure or to facilitate a requested function or biophysical property*[20]. Rational protein design, as we know it today, dates back to the 1970's with Chris Anfinsen's Nobel prize for the "connection between the amino acid sequence and the biologically active conformation" of proteins[21], and thus, the start of the sequence-structure-function relationships that are at the core of protein science. However, the original idea of it goes back to 1902, when Emil Fischer emphasized, during his own Nobel lecture, that molecules (such as proteins) can be elucidated by the reverse method, namely, chemical design from decomposition products (i.e. aminoacids)[22].

The first rational design of a protein is attributed to Gutte, in 1975, who pioneered the field by designing a functioning, half-length version of bovine ribonuclease in which the aminoacid excursions away from the active site were truncated[23,24]. The first computer-aided rational design would come in 1985, when DeGrado designed, synthesized and characterized a 17-residue helical peptide that was, at the time, the tightest calmodulin-binding peptide produced[25,26]. Today, while the field is still constantly evolving, the plethora of successfully computationally designed proteins highlight the growth and potential of the field. Researchers have not only designed completely new folds (such as TOP7[27]) but also new functions (such as the design of new enzymatic functions[28]). These success stories are not only the result of increasing computational power, but also of the increasing quality of computational tools, models, and methods available, as well as a better understanding of protein structure and function.

Protein design concepts. From a practical standpoint, protein design is seen as "the inverse protein-folding problem", a term defined by Yue and Dill in 1992[29]. While the "protein folding problem" looks at searching for the lowest-energy conformation for a given amino acid sequence, protein design aims at identifying the sequence that stabilizes a requested and defined conformation. As such, protein design may also target aspects which are not a specific structure, per se, but rather a specific characteristic thereof, such as interactions or stability. Due to this parallelism between protein folding and protein design, the methods and tools used often intersect and have common development paths.

De novo vs template-based design. In general, protein design can be split into two broad categories: *de novo* design, and template-based design. In *de novo* design a totally new fold and/or function are pursued, such as the TOP7 design (Figure 1A) of Kuhlman and Baker[27], the transmembrane Zn^{2+} transporter of DeGrado[30] (Figure 1B) or the more recent enzyme designs of Baker[31]. Once the desired fold/structure is chosen, a model (or models) of the protein backbone adopting the target fold is constructed and used for subsequent steps. Much of the recent work in the field of de novo protein design has been focused on developing improved methods for constructing protein backbones that are physically realizable[32,33]. In template-based design, however, the sequence and structure of naturally evolved proteins are modified to achieve new functions or properties. Examples of template-based design are the solubilization of membrane proteins[34,35] or the improvement of protein-protein interactions by modifying their specificity[36,37] and/or affinity[38]. In this case, the backbone of the template protein is used for the subsequent design.

Having defined a starting model or set of models for the protein backbone, the next step in the design process is to identify an amino acid sequence that will stabilize the desired conformation or property. This process can be divided into two iterative steps: the evaluation of the favorability of

a particular sequence and the subsequent search for more favorable sequences.

Evaluating a particular sequence. When analyzing the favorability of a particular sequence, optimization software must first *search for lower-energy side chain conformations*, which are evaluated according to a specific *energy function*. In our work, this optimization was carried out using the Rosetta software for macromolecular modeling[39].

Rosetta originally started development in the mid-1990s, aiming at tackling the protein structure prediction and protein folding problems. Since then, it has added an impressive number of applications to tackle diverse modeling tasks from protein design, to incorporating nuclear magnetic resonance data (NMR) or even incorporating interactions with nucleic acids. Additionally, the core algorithms and energy functions have been continuously improved and refined over the last two decades. While over time, the user and developer community's main goals moved on to macromolecular design, performance in the critical assessment of protein structure prediction – Round XIII (CASP13) blind prediction challenge remains remarkably respectable[40]. Overall, Rosetta has enabled notable scientific progress in the fields of computational biology, including protein and enzyme design, ligand docking and protein structure prediction[41].

Energy Functions. Energy functions describe the total energy of a determined system as a functional of its state. These are typically parameterized using a variety of benchmarks that focus on reproducing the features of naturally occurring proteins, emphasizing residue-pair interactions, atomic packing, hydrogen bonding, solvation, protein torsion angles, and rotamer conformation preferences.

The Rosetta energy function 2015 (REF15)[42] was used in this work. REF15 approximates the energy of a biomolecule conformation, ΔE_{Total}, as a linear combination of 19 energy terms, E_i, that are functions of geometric degrees of freedom, Θ_i, and chemical identities, aa_i. These are scaled by a specific weight, w_i,on each term:

$$\Delta E_{Total} = \sum_i w_i \, E_i \, (\Theta_i, \mathrm{aa}_i) \qquad (1)$$

From the 19 energy terms, 7 terms describe the energies of interactions between nonbonded atom pairs, which are important for atomic packing, electrostatics, and solvation, 5 terms are used to model hydrogen and disulfide bonds, and the final 7 terms are statistical potentials used to describe the backbone and side-chain torsional preferences in proteins. The specific weights, w_i, have been adjusted to balance the contributions from the various terms. These are crucial, as van der Waals solvation, and electrostatics energies partially capture some torsional behavior, overlapping with the statistical torsional preference parameters. This overlap can cause errors as a result of double counting of atomic or residue-specific contributions[42,43]. All these parameters have evolved over the years by optimization of the performance of multiple scientific benchmark methods.

While REF15 can reproduce several thermodynamic observables, as well as structure-based tests, there are still several limitations[41]. The energy function still does not directly estimate entropy, although some terms implicitly model this component accounting for more than half of the entropy. Some knowledge-based terms are derived from high-resolution crystal structures, which represent a single state on the energy landscape, and thus do not account for protein flexibility. Additionally, these same knowledge-based terms are less interpretable and transferable than the physics-based terms. The solvation model is treated implicitly, which speeds up the scoring process but hinders the explicit modeling of interactions with ions and water molecules. These shortcomings of the energy function are important to take into consideration when analyzing optimization results.

Side chain conformation search. For the search for lower-energy side chain conformations, a variety of methods have been developed, which range from deterministic approaches, such as dead-end elimination, mean-field optimization or Markov random field solvers, to stochastic approaches, such as simulated annealing, Monte Carlo or genetic algorithms[33,44]. While deterministic methods have access to the complete set of data, and, if they converge, are bound to find the global minimum energy conformation, stochastic search methods have an intrinsic random component, which may give a slightly different solution each run. While this solution is likely to be a good solution, it is not guaranteed to be the best solution. However, this randomness, in general, can significantly accelerate the search progress and thus reduce computational cost. Additionally, it can aid the method to escape local minima, eventually approaching the global optimum. Given the astronomical size of the search space for a regularly sized protein, deterministic approaches are often unfeasible.

Regardless of the method used, most simplify the side-chain optimization problem by restricting side-chain motion to a set of commonly observed conformations (or rotamers) observed in high-resolution structures from the PDB[33,45]. As each amino acid sidechain can only realistically accept a discrete number of conformations (rotamers), rotamer libraries can greatly reduce the conformational search space, speeding up the process.

The Rosetta software's primary algorithm for optimizing sidechains is called the *packer*, which uses Monte Carlo sampling with simulated annealing to identify the low energy rotamers[46]. Monte Carlo methods in Rosetta use the Metropolis-Hastings algorithm[47]. In summary, when iterating steps where a random move is made, that alters the pose in some way, the resulting change in energy, ΔE, is determined and the move is then accepted or rejected based on the Metropolis criterion. The Metropolis criterion states that if a move results in a decrease in energy, then it is always accepted. However, if the move results in an increase in energy, then it will only be accepted with a probability, P, equal to:

$$P = e^{-\Delta E/(k_B T)} \qquad (2)$$

Where the temperature factor, $k_B T$, determines how the probability of accepting/rejecting a move scales with the change in energy, ΔE. This criterion means that the larger the increase in energy, ΔE, the lower the probability of the move being accepted. Simultaneously, larger values of $k_B T$ allow moves that result in larger increases in energy to be accepted more frequently, while smaller values only permit moves that result in smaller increases in energy.

While searches can be carried out with fixed values of $k_B T$, it is often advantageous to vary this value over the course of an optimization. For example, using a high $k_B T$ value during early optimization stages can allow the structure to explore conformations with higher energy barriers and escape local energy minima, while lowering the $k_B T$ value at later stages might help the conformation reach the bottom of the lowest energy well that it had previously found. Monte Carlo methods that cycle the value of $k_B T$ over the course of an optimization are broadly called simulated annealing methods.

In general, when *packer* is called, it lists all possible rotamers at each position and then pre-computes all possible pairs of interacting rotamers, calculating and storing their interaction energies. *Packer* then carries out the simulated annealing-based search of rotamer combinations, randomly selecting an amino acid and replacing the rotamer at that position with a randomly selected rotamer from the library of allowed rotamers for that position.

As the overall conformation is often very sensitive to the precise 3D structure of the protein, gradient-based minimization of backbone and sidechain torsion angles is also performed. Very small atomic displacements (<1 Å), often resultant from small changes in backbone conformations, can lead to very significant jumps in energy due to strong steric repulsions or loss of favorable interactions. In Rosetta, the *minimizer* algorithm takes on this task, finding the nearest local minimum in the energy function by means of a line search optimization.

Searching for more favorable sequences. Evaluating a particular sequence is only part of the process. The overall goal is to identify the sequence that stabilizes a requested and defined conformation and as such the software must search through sequence space. To optimize the sequence, Rosetta generally carries out the search with Monte Carlo sampling with simulated annealing, much like when searching for low energy rotamers. In our implementation, this part of the process was replaced by a genetic algorithm implemented in Python3, as described further ahead. This was done to gain more control over the optimization process as well as improve the computational performance of the sequence search.

Protein Docking. When designing an interaction between two proteins, a *docking* step is often put in place. *Docking* is a method which can search for the preferred orientation of one molecule (ligand) in reference to a second (target), binding each other to form a complex. Two types of docking protocols can be used, *local docking*, where we assume that we already have some information about the binding of the two molecules/proteins (i.e. known binding surface or known complex structure), or *global docking*, where no information is available on the binding sites in the proteins.

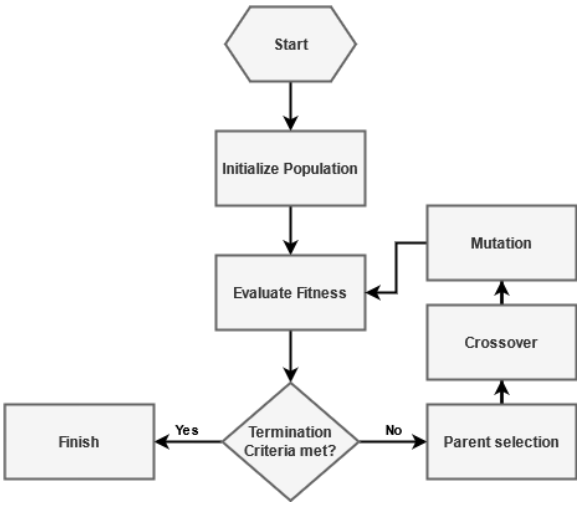


Figure 2. General GA workflow diagram.

In Rosetta, when running *local docking*, a structure with the two proteins within 10 Å and with the binding pockets roughly facing each other is used as the starting point. The chain defined as the ligand is then randomly perturbed (translation and rotation) before the start of every individual docking optimization. When running *global docking*, Rosetta assumes a general spherical structure for the proteins, and rotates the ligand randomly around the target. Before the start of every individual run the starting position of the unbound proteins are completely randomized, so their initial position is not that significant.

Regardless of the type of docking, in Rosetta it always undergoes a two-stage protocol. In the first stage, aggressive sampling is done in centroid mode, where the two proteins are pulled together and slid into contact. These translation and rotation moves are randomly selected following a Monte Carlo method. In the second stage, smaller movements take place in full atom mode. In both steps, the protocol always assumes a fixed backbone and can only undergo the translation and rotation moves as well as sidechain packing. However, the docking protocol can often originate small clashes, so the structures are often further refined, with sidechain re-packing or backbone and sidechain torsion angle minimization[46].

Combining protein design and genetic algorithms. *Genetic algorithm fundamentals.* Genetic algorithms (GAs) are search heuristics inspired by the theory of natural evolution, which are particularly well suited for multicriteria optimization over solution spaces that are either categorical or non-differentiable. The goal of a GA, like all metaheuristics, is to find or select a sufficiently good solution to an optimization problem. A GA operates on a population of *individuals,* where those that are evaluated as the *fittest* are

selected for reproduction, producing offspring for the next generation. Each individual is characterized by a set of properties, or *chromosomes*, which can be mutated and altered during the process. These chromosomes are typically sequences of variables and represent a solution to the problem in question. Individuals are judged based on their *fitness*, which is a function of their chromosomes and used as the measure of how good a solution it is to the problem. GAs have their origins in the 1950s and with increasing computational power they have been successfully applied to solve a wide range of real-world optimization problems in most fields of study, from biology and medicine to physics, mathematics, economics, and even social sciences.

General GA algorithm. The typical steps in a GA workflow can be seen in figure 2. In general, a genetic algorithm starts with the generation of an initial population. This initial population may be generated at random, seeded by other heuristics or through previous knowledge of the problem. The choice of population is an important consideration. It should not be too large, as it can significantly slow down the optimization, however it should also not be too small, as it might not provide a deep enough gene pool to obtain a reasonable accurate solution. However, there is no rule of thumb to determine population size. Optimal population size depends on the nature of the optimization problem, number of dimensions, mutation rates, number of generations and intended outcome. As such, this parameter should be tested and optimized according to the constraints of the problem.

At this point, the initial population is evaluated, and a fitness value attributed to each individual. This is done through a fitness function, which takes the individual solution to the problem and provides a numerical output of how good the solution is. Obviously, fitness functions are heavily dependent on the problem. However, they should always have a couple of characteristics: they should accurately and quantitatively measure how fit a given solution is, in a clear and concise manner, and it should be sufficiently fast to compute. Fitness functions are called several times per iteration and as such they should be implemented as efficiently as possible, otherwise they become a large bottleneck.

Once every individual has been evaluated, the mating process starts. Mating begins with parent selection, which is the process of selecting the individuals that, in turn, mate and recombine to create the offspring for the next generation. In steady-state or incremental GA, only one or two offspring are generated per iteration, while in generational GA, the entire population (or most of it) is replaced by the new one at the end of the iteration.

Parent selection is crucial to the convergence rate of the algorithm, as fitter parents drive better solutions. However, care must be taken to prevent premature convergence, and an effort must be made to maintain good population diversity. Parent selection is typically carried out through one of several strategies: Roulette wheel selection, rank selection or tournament selection. K-way tournament selection is one of the most popular selection methods. In this method, K individuals are selected from the population at random, and a tournament is run among them, selecting the fittest candidate. This is repeated several times until there are enough selected pairs to mate and produce the next generation. The same individual can be chosen in multiple tournaments. If the tournament size is larger, weaker individuals have a smaller chance of being selected for mating; as such, one can control the selection pressure and rate of convergence through the tournament size. Having selected a pair of parents, their genetic information is now combined to generate the offspring. In traditional genetic algorithms this is typically done through chromosome crossover. In k-point crossover, k crossover points are picked randomly from the parent chromosomes, originating k+1 segments. These segments are swapped between the parent chromosomes, originating two offspring individuals with genetic information distinct from each other and their parents.

At this point, mutations might also be applied with a, usually low, probability. Random mutations are essential to the exploration of the search space and the successful convergence of the GA to a reasonable solution. Mutations involve randomly altering a gene to another permitted value for the corresponding variable. When constructing the new population, elitist selection is often applied, where the best individual — the champion — from the current generation carries over to the next one unaltered. This guarantees that the best solution obtained by the GA will not decrease from one generation to the next. The new population is now, iteratively, re-evaluated, mated and mutated until termination criteria are met. These are picked such that the solution is close to optimal at the end of the optimization. Typically, termination criteria are either: No fitness improvement for X iterations; a certain absolute number of generations; or the fitness value reaching a pre-defined value.

Applying GAs to protein design. GAs have also been successfully applied to the field of protein design. Genetic algorithms have been applied to search the sequence space in multistate computational design of proteins, combining both positive and negative designs[48–50]. This overall method is particularly useful when optimizing a design for specificity and has since been implemented into Rosetta with the purpose of optimizing a single sequence to perform differently in distinct contexts[51]. They have also been successfully combined with *in vitro* assays to quickly identify 44 highly potent antimicrobial peptides[52]. This method used the GA to search the amino acid sequence space, starting with a 13-mer natural AMP, and using the in vitro bacterial assay as the fitness function. In this work, we implement a GA method to search the sequence space for the optimization of anti-SARS-CoV-2 biomolecules.

GA-guided ROSETTA protein design. In this work, we showcase the design of an improved anti-SARS-CoV-2 biomolecule, using a GA guided ROSETTA approach. In our method, a GA was used to search the sequence space for each of the selected amino acid positions. This GA was implemented by us in Python3 using the DEAP package. An overview of our method can be seen in figure 3.

Like all GAs, our method starts with the generation of the initial population. Each individual's chromosome is defined as a string of letters, each of which represents an amino acid at a particular position (i.e. AAA represents Ala – Ala - Ala). The initial population is randomly generated, from the 18 available amino acids. Glycine and Proline are not

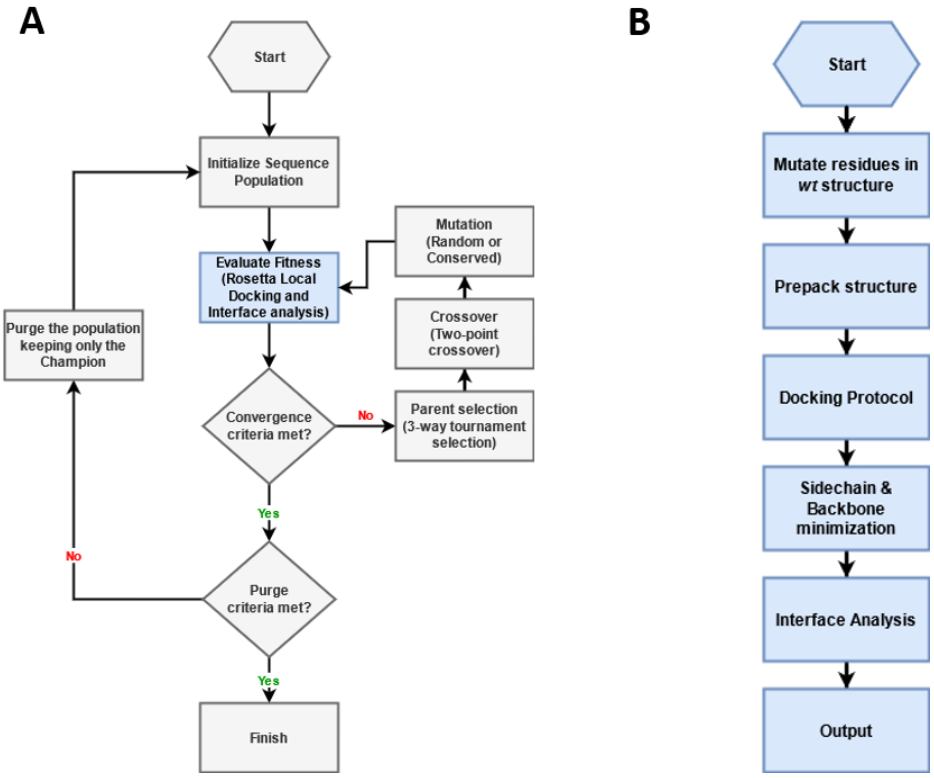


Figure 3. GA-Guided Rosetta Design algorithm workflow.

considered in our design, as mutations to either of these amino acids are often disruptive. To this initial population, the wild type sequence is added. The *wt* sequence is kept in every population during our optimization, guaranteeing that information about the *wt* sequence is propagated through the iterations and assuring that the best result always performs at least as well as the *wt* sequence.

At this point the first fitness evaluation takes place. Fitness of each individual is defined as the Interface energy score (I_{sc}) obtained for each sequence by Rosetta. Each individual undergoes a docking and relaxation routine. In summary, the mutations are applied to a previously relaxed *wt* structure of the biomolecule bound to the relevant target (i.e. ACE2 bound to the SARS-COV-2 RBD), and the sidechains are then re-packed in preparation for the local docking runs. Local docking takes place, slightly refining the relative orientation of each protein and then optimizing the sidechain rotamers. Structure optimization is finalized by minimizing the sidechain and backbone dihedrals. At this point protein-protein interface analysis can take place and a I_{sc} is obtained. For each individual 5 replicas are run and the best result is stored (with differences between replicas expected due to the non-deterministic nature of the packing procedure). If an individual has already been evaluated in a previous iteration, the result is compared to the stored I_{sc} and the best one is considered.

In our GA, parent selection is done through a 3-way tournament selection without reposition and mating through two-point crossover. At this point, individuals can be selected for mutation with a 50% probability. If selected, there is $1/length(chromosome)$ probability of mutation to each gene, half of which will be random mutations, and the other half conservative mutations. Conservative mutations are mutations where the amino acid is changed to one of the same type, i.e. a Glutamine to an Asparagine. When constructing the new population, the best individual, as well as the *wt* sequence, are carried over unaltered to the new generation. This not only guarantees that the best solution does not decrease over iterations, but also that it will never be worse than the *wt* sequence.

In our model, a population is deemed converged when there is not enough genetic diversity to assemble enough unique pairs of individuals to generate the next iteration of offspring. At this point, the population is "purged" by discarding the entire population, with the exception of the fittest individual. This individual will be carried over to a new, randomized population which will continue to be iterated over. After 3 "purges" without fitness improvement, the optimization is terminated.

This GA-guided Rosetta approach allowed us to quickly search the optimization sequence space, while having full control over the optimization process. The approach yielded similar results to the standard Rosetta design methods using less CPU time. Additionally, whilst we used the straightforward Interface energy score (I_{sc}) as our fitness, more complex functions considering other parameters can be easily implemented. This opens the door for multiple objective optimizations, such as optimizing for specificity, as

has previously been done, or even simultaneously optimizing protein-protein interactions and protein solubility.

COMPUTATIONAL METHODS

GA-guided protein design. The original ACE2 bound to RBD structure was obtained from PDB (PDB ID: 6M0J[53]). Glycosilations were removed and the structure relaxed using Rosetta relax[54–57], to refine the conformational space around the crystallographic structure.

Design target residues were determined from ACE2 surface residues, using the distance between the 2 proteins as the selection criterium. ACE2 surface residues were determined by calculating each residues' solvent accessible surface area (SASA), as determined by gmx SASA[58,59], and excluding those that had a SASA of less than 0.25 nm^2. The 14 ACE2 surface residues that were within the 3 Å cutoff were selected. Glycine and Proline residues were not considered in the selection. Hotspot residues (i.e. residues that contribute the most for *wt* ACE2 binding to SARS-CoV-2 RBD), were determined by an Alanine scan, using Rosetta[60]. Residues D30, D38 and Y41 were identified, as they contributed over 2 REU towards the binding energy. These residues were also not considered in the selection of our Design target residues. As such, the 11 design target residues are: Q24, T27, K31, H34, Q42, L79, M82, Y83, N330, N330, K353 and R393 (Final list of residues to mutate: QTKHQLMYNKR).

A GA was used to control the optimization algorithm and search the sequence space for each of the selected amino acid positions. This GA was implemented in Python3 using the DEAP package. An overview of our method can be seen in figure 3. A population size of 36 individuals was used. Parent selection and mating was carried out through a 3-way tournament selection and 2-point crossover of unique pairs of parent individuals. Once there is not enough genetic diversity to assemble enough unique pairs of individuals to generate the next iteration of offspring, the population is purged, with the exception of the fittest individual. This individual will be carried over to a new randomized population. After 3 purges without fitness improvement, the optimization is terminated.

Fitness is defined as the Interface energy score (I_{sc}) recovered for each sequence by a Rosetta docking routine implemented in RosettaScripts[61] which interfaces with the GA using bash. An overview of the routine can be seen in figure 2B. For each individual 5 replicas are run and the best result is stored. If an individual has already been evaluated in a previous iteration, the result is compared to the stored values and the best one is considered. Further refinement Rosetta runs, for specific sequences, used a similar docking routine. At least 100 designs were generated for each refinement run.

Importance of each mutation was determined using the rosetta revert_design_to_native application[62,63]. This application determines the contribution of each aminoacid substitution at the interface of a design, by singly reverting these residues back to their wildtype identities and computing the effects of the reversion on the I_{sc}.

Atomistic molecular dynamics simulation systems. All atomistic simulations were performed with the GROMACS 2020.3[58,59] package and modelled using the Amber14sb[64], forcefield alongside the TIP3P water model[65]. Simulations were performed of each ACE2 protein structure in water as well as in a complex with the SARS-CoV-2 S-protein RBD in water. Each structure was inserted in a truncated dodecahedron box filled with water molecules (minimum of 1.2 nm between protein and box walls). The total charge of the system was neutralized and additional Na^+ and Cl^- ions were added to the solution to reach an ionic strength of 0.1 M.

The system was energy-minimized using the steepest descent method for a maximum of 50000 steps with position restraints on the heteroatom positions by restraining them to the crystallographic coordinates using a force constant of 1000 kJ/mol in the X, Y and Z positions. Before performing the production runs, an initialization process was carried out in 5 stages of 100 ps each. Initially, all heavy-atoms were restrained using a force constant of 1000 kJ/mol/nm, and at the final stage only the only C-α atoms were position-restrained using the same force constant. In the first stage, the Berendsen temperature algorithm[66] was used to initialize and maintain the simulation at 300 K, using a temperature coupling constant of 0.01 ps, without pressure control. The second stage continued to use the Berendsen temperature algorithm but with a force constant of 0.1 ps. The third stage kept the same temperature control settings, but introduced pressure coupling with the Berendsen pressure algorithm[66] with a pressure coupling constant of 5.0 ps and applied isotropically. The fourth stage changed the temperature algorithm to V-rescale[67], with a temperature coupling constant of 0.1 ps, and the pressure algorithm to Parrinello-Rahman[68] with a temperature coupling constant of 5.0 ps. The fifth stage is equal to the fourth stage, but position restraints are only applied to Cα atoms.

During the simulation, the equations of motion were integrated using a timestep of 2 fs. The temperature was maintained at 300 K, using the V-rescale[67] algorithm with a time constant of 0.1 ps, and the pressure was maintained at 1 bar using the Parrinello–Rahman[11] pressure coupling algorithm, with a time constant of 5 ps; pressure coupling was applied isotropically.

Long-range electrostatic interactions were treated with the PME[69,70] method, using a grid spacing of 0.12 nm, with cubic interpolation. The neighbor list was updated every twenty steps and the cutoff scheme used was Verlet with 0.8 nm as the cut-off radius. All bonds were constrained using the LINCS algorithm[71]. All systems were simulated for at least 2 µs over 5 replicates. The last 1 µs of each replicate was used for analysis.

RESULTS AND DISCUSSION

The goal for this work was to optimize and design a set of proteins for the neutralization of the SARS-CoV-2 S-Protein RBD, for potential therapeutic or prophylactic use. Here, the design, refinement and in silico characterization of the Angiotensin-converting enzyme 2 (ACE2) protein is described.

ACE2 (Figure 4) is an enzyme that is typically attached to the outer cell membrane and plays a key role in the cardiovascular system by catalyzing the hydrolysis of angiotensin

II into angiotensin[18]. Studies have shown that the ACE2 enzyme, previously identified as the cellular receptor for SARS-CoV, also acts as the receptor for SARS-CoV-2 by mediating the interactions with the S glycoprotein on the surface of the virion. This process is part of the concerted action that culminates in viral entry to the target cells. Due to its importance in the mechanism, ACE2 has been labelled as a potential pharmaceutical target for the treatment of the disease, with promising use as a neutralizing decoy receptor for SARS-CoV-2. However, the human ACE2 sequence has been found to be suboptimal for binding the SARS-CoV-2 RBD[72] and, as such, efforts have been made to optimize it.

Design and refinement of the ACE2 protein. *Structure preparation and selection of design target residues.* The core structure from which we based our optimization on was a crystal structure of *wt* ACE2 bound with SARS-CoV-2 spike RBD, obtained by X-ray diffraction (PDB ID: 6M0J[53]). In preparation for the optimization, glycosylations were removed and the structure relaxed, to remove potential crystallographic clashes and refine the conformation space around the structure.

The target residues for optimization were selected from the ACE2 surface residues that were within 3 Å of the RBD. Care was taken not to include Gly and Pro residues which could destabilize the core structure. Hotspot residues were also removed from the selection. This process resulted in the selection of 11 potential design target residues: Q24, T27, K31, H34, Q42, L79, M82, Y83, N330, K353 and R393 (Final *wt* sequence: QTKHQLMYNKR).

Other than mutations to the design target residues, 2 other mutations were applied to every design: H374N and H378N. These mutations are designed to neutralize ACE2 enzymatic activity[73], by replacing the 2 Histidine residues which are responsible for complexing the Zn^{2+} ion in the active center. This is an important step in using ACE2 as a potential pharmaceutical, as it plays an important role in the cardiovascular system which must be nullified before potential administration.

Optimization and Refinement of potential designs. At this point, the ACE2 structure was optimized for RBD binding by GA guided mutation of the design target residues (figure 5A). After 140 generations and 4 purges the algorithm reached termination criteria and yielded the following potential design #4 VIFIYWLFAWN. This design showed a -11.8 REU (roughly 30 %) improvement in interface score in comparison to the *wt* sequence. It is worth noting however, that the Total Score of the potential design is significantly higher than that of the *wt*. The potential design showed an increase of 28.9 REU (1.15 %) in comparison to the *wt* sequence. The total score takes into account not only the binding energy but also the stability of each protein conformation and the solvation energies. As such this increase in energy hints at a decrease in protein stability or solubility of our design.

To confirm the performance of the GA, 4 sequences were chosen from the course of the optimization. For each sequence, at least 200 docking runs were performed to confirm their interface scores and measure whether they are suffering from undersampling (figure 5C). The interface scores recovered were similar to those obtained during the optimization and, most importantly, the order of improvement was kept the same, proving the efficacy of the algorithm.

Figure 4. ACE2 bound to SARS-CoV-2 RBD. Crystal structure of SARS-CoV-2 spike RBD (orange) bound with ACE2 (green), acquired by X-ray diffraction[53] (PDB ID: 6M0J).

To refine our design, and test each of our mutated residues, we individually reverted each mutation back to the wild type aminoacid and determined the impact on the interface score (figure 5B). From this reversion, two mutations offered no significant benefit over the *wt* residues: Q42Y and R393N. As such, these 2 mutations were reversed and the new sequence, #5 VIFIQWLFAWR, was further tested with at least 200 docking runs (figure 5C). This sequence yielded a -10.8 REU (27 %) improvement in interface score in comparison to the *wt* sequence, 1 REU less than the best sequence, #4. However, the new sequence performed much better in terms of total score, showing an increase of only 1 REU compared to 28.9 REU of sequence #4. This hints at sequence #5 being much more stable than #4.

Having concluded the optimization, out of the 11 potential design target residues, 9 mutations were applied: Q24V, T27I, K31F, H34I, L79W, M82L, Y83F, N330A and K353W. Procko et al. has created a library of ACE2 mutants that increase RBD binding and created an enrichment map of aminoacid substitutions for each ACE2 aminoacid position[72]. Out of the 9 mutations found by our model, 4 mutations (Q24V, H34I, N330A and K353W) were depleted in this library, 4 were neither significantly enriched or depleted (K31F, T27I, M82L, Y83F) and 1 was highly enriched (L79W).

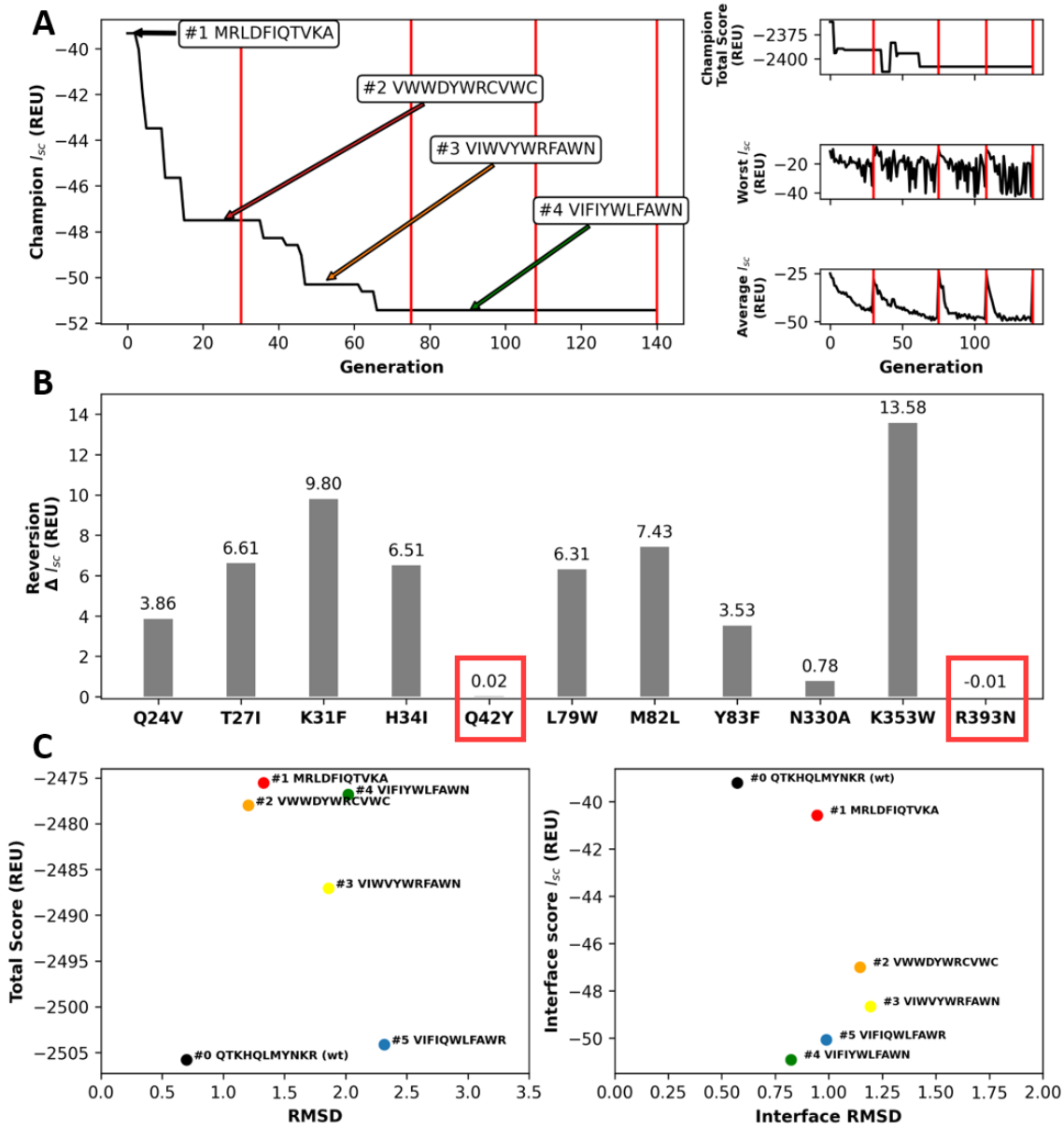


Figure 5. GA-Guided optimization and refinement of ACE2. (A) Optimization progress and best interface score (I_{sc}) over the course of the GA iterations. Champion total score, worst I_{sc} and average I_{sc} over the course of the optimization are also shown. The red bars refer to iterations where a purge occurred. The best sequence, as well as 3 others used to confirm the performance of the GA are shown. (B) Difference in interface score upon reversion of each design mutation back to the *wt* residue. The two mutations which were eventually reverted are marked in red. (C) Best performing structures for each of the selected sequences, as well as the final sequence obtained after the 2 reversions. At least 200 docking runs were performed for each sequence from which the best structure was selected. RMSD and interface RMSD are calculated using the crystallographic structure as reference.

From the 4 mutations that were found depleted in the author's model, we scanned all possible replacements for the 2 most aggressive mutations (N330A and K353W). K353W was the best possible aminoacid replacement at this position, yielding both the best Interface and total scores (figure S1). For position 330, N330F resulted in a very slightly improved interface score (>1 REU improvement), however, this replacement resulted in a significant worsening of the total energy, hinting at a more unstable structure (figure S2). N330A provided the best interface and total score combination. Sequence #5 VIFIQWLFAWR was thus chosen for further tests and characterization.

Characterization of the binding interface. Analyzing the binding interface of the designed sequence and comparing it to the *wt* sequence, one thing that stands out is the increased hydrophobicity (figure 6 A, B, C, D). Interface

hydrophobic solvent accessible surface area (SASA) increased from 885 to 1236 Å^2, while polar SASA decreased from 779 to 747 Å^2. Increased hydrophobicity can increase binding affinity but can also induce unwanted effects such as protein instability, misfolding or non-specific aggregation. It is worth noting, however, that most mutations towards more hydrophobic residues are mostly restricted to left sector (per the division in figure 6) of the 2 alpha helices that make up most of the binding interface (figures 6 C & D). The fact that this hydrophobic region is restricted reduces the probability of issues, however further in vitro and MD tests must be performed to be certain about the overall fold stability.

The overall interface area also increased significantly (from 1664 to 1984 Å^2), however, the number of predicted hydrogen bonds between the two structures decreased from 8 to 5. This hints at a slight loss in binding specificity. With the simultaneous increase in surface area and decrease in the number of predicted hydrogen bonds we also saw an increase in unsatisfied buried hydrogen bonds at the interface from 1 to 11. This was a parameter that at the time of optimization wasn't taken into consideration for the optimization process, however, it clearly hints at more available binding energy gains that could potentially be unlocked by taking into account some of these parameters into the fitness function.

Looking at specific amino acid interactions on the left section of the binding interface, the *wt* ACE2 (Figure 6E) establishes binding via hydrophobic interactions mediated by M82, L79 and the RBD F482. A pi-stacking interaction between ACE2 Y83 and RBD F482 was also in place. Y83 appeared to be supported by T27 and Q24 who could also establish some interactions with the backbone of RBD G476 and the sidechain of Y473. On the designed ACE2 (Figure 6H), both L79W and Y83F now establish a potential double pi-stacking interaction with RBD F486. These interactions are supported by M82L, Q24V and T27I who establish non-specific hydrophobic interactions with several RBD residues.

On the middle section of the binding interface, the *wt* ACE2 (Figure 6F) establishes a significant salt bridge between D30 and RBD K417. D30 is one of the hotspot residues which was kept unaltered in our optimization. K31 establishes interactions with RBD Q493, and H34 with RBD Y453. These 2 interactions appear to be quite non-specific. On the designed interface (Figure 6I), K31F now leads to the establishment of a pi-stacking interaction with RBD Y489, supported by H34I which establishes non-specific hydrophobic interactions with RBD L455.

On the right section of the binding interface, the *wt* ACE2 (Figure 6G) establishes mostly polar interactions, via Q24, D38, Y41, N330 and K353. On the designed interface (Figure 6J), 2 mutations are in place. N330A and K353W. N330A establishes a hydrophobic interaction with the methyl group of the RBD T500 sidechain. It is less clear what K353W interacts with, however. This residue is placed in a tight pocket lined by the protein backbone of several residues but surrounded by several polar sidechains. Nonetheless, scanning of this residue with Rosetta reveals that K353W is the best possible aminoacid replacement at this position.

It is worth noting that this analysis, is based on the best scoring function obtained by docking to both *wt* and designed sequences. As such, it provides limited information on the dynamics of these interactions and does not take into account explicit solvent and ion effects as well as protein conformational dynamics effects. To obtain more information on these interactions, atomistic molecular dynamics simulations were performed.

Atomistic simulation characterization of the optimized structures. To better probe the interaction and conformational dynamics of these proteins, as well as the effects of solvation on these properties, atomistic molecular dynamics simulations were carried out using the Amber14sb[64] forcefield. In this study 3 ACE2 protein sequences were tested in 2 different simulation systems. We simulated the *wt, wt* + H374N and H378N (HH) and our designed sequence. These sequences were then either simulated by themselves in water, to test protein stability and conformation dynamics, or complexed with RBD to test binding interface interactions.

ACE2 Conformational stability and dynamics. ACE2 conformational stability and dynamics were analyzed from systems containing the ACE2 protein in water. We started by analyzing the root-mean-square deviation (RMSD) of atomic positions of each of the systems from the X-ray structure. The three simulated variants yielded very similar, low RMSD values, pointing at quite stable structures. This result is quite promising for the overall stability of our design as its fluctuations are similar to that of the *wt* variant in the timescale here probed.

Nonetheless, some RMSD fluctuations were observed, especially on some replicas of the designed sequence. These fluctuations were originated from the opening (Figure S3A) and closing (Figure S3B) of the ACE2 active center pocket. ACE2, like other metallopeptidases, undergoes a large hinge-bending motion when binding a substrate or inhibitor, locking it in the active center[74]. To probe these dynamics over the course of the simulations we looked at both the volume of the active center (Figure S3D), as well was the angle made by the opening and closing of the active center pocket (Figure S3C).

Looking at the pocket dynamics of the free ACE2 variants, we see that the *wt* and HH variants show similar pocket open angles despite the *wt* variant showing a slightly larger active center volume. The designed variant shows a slightly wider distribution of opening angles with the peak skewed to lower angle values than the *wt* and HH variants. It also yields a much smaller active center pocket volume. The designed ACE2 shows much higher variance between the replicas in comparison to the *wt* and HH sequences.

When analyzing the simulations containing the ACE2 molecules complexed with the RBD, we observe a slight decrease in the population with larger interdomain angles for the *wt* variant, although no significant decrease in active center volume is observed. Similar behavior is observed for the HH variant. This contrasts with other ACE2 simulation studies, which have shown that the complexation with the RBD locks the structure in the open conformation[75]. The designed variant shows a similar behavior, yielding a smaller

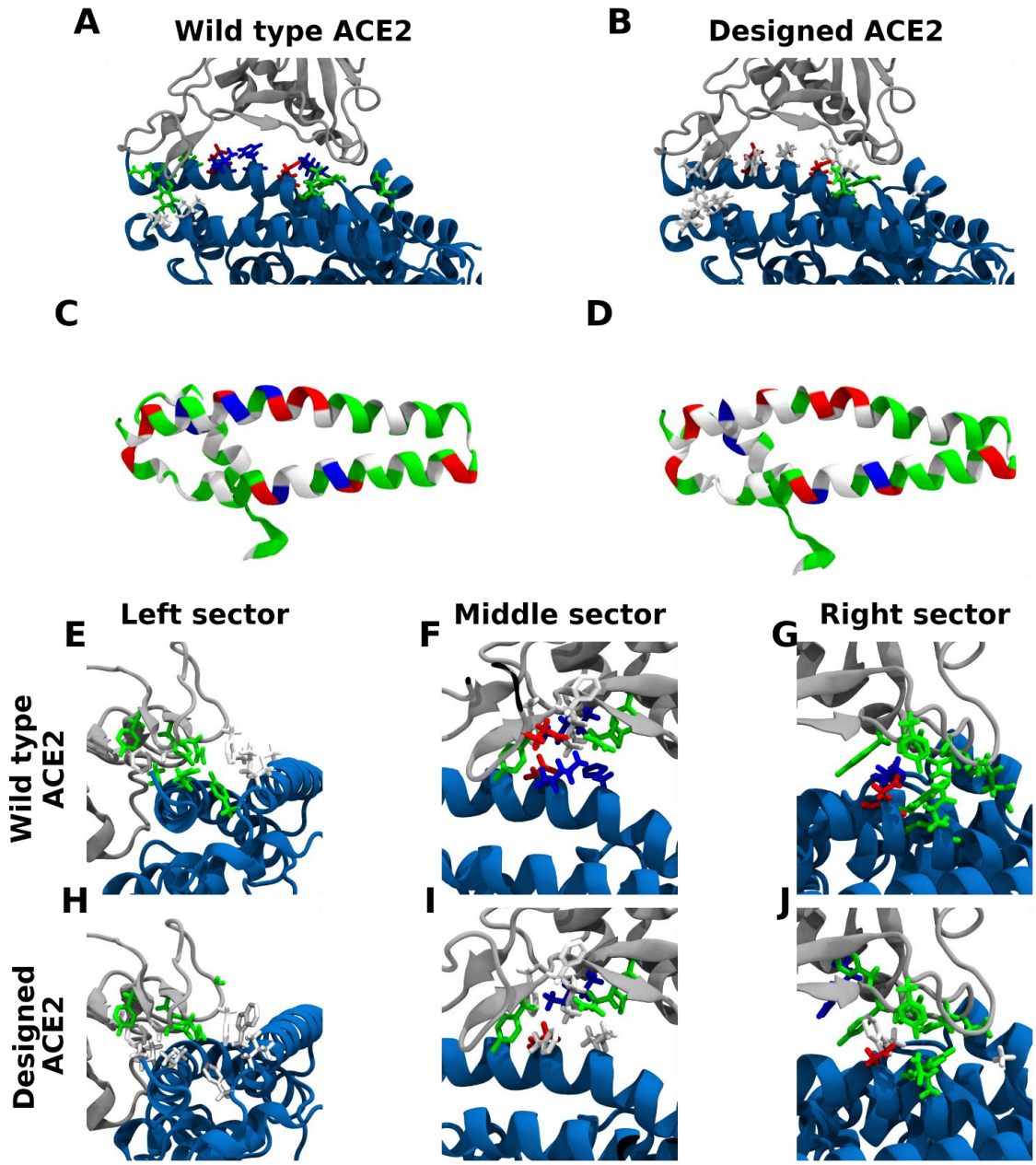


Figure 6. Interface analysis of both wild type and designed ACE2 proteins. (A and B) General look at both *wt* and designed ACE2 (blue) RBD (white) interfaces. (C and D) Hydrophobicity profile of the 2 interface helixes from both *wt* and designed sequences. (E to J) Close-up on the interactions established by both *wt* and designed ACE2 sequences. Aminoacid color code: non-polar: white, basic: blue, acidic: red, polar: green.

interdomain angle and active center volumes. These dynamics are likely to play a role in the ACE2 binding affinity. And while the extent of their influence is not yet clear it is unlikely that these would influence molecule stability.

ACE2 – RBD binding characterization. To characterize ACE2-RBD binding we deployed simulation systems consisting of ACE2 complexed with RBD for each of the studied sequences. We started by analyzing the contact matrix for each simulated system. Each system yielded a very similar contact matrix (figure S4 A, B & C), with a similar number of average total contacts per frame (145, 145 and 143 for the *wt*, HH and designed sequences, respectively) (table S1).

Looking at the contact matrix for the designed sequence (figure S4 C), we can analyze whether the contacts present in the rosetta design are kept during the MD simulations

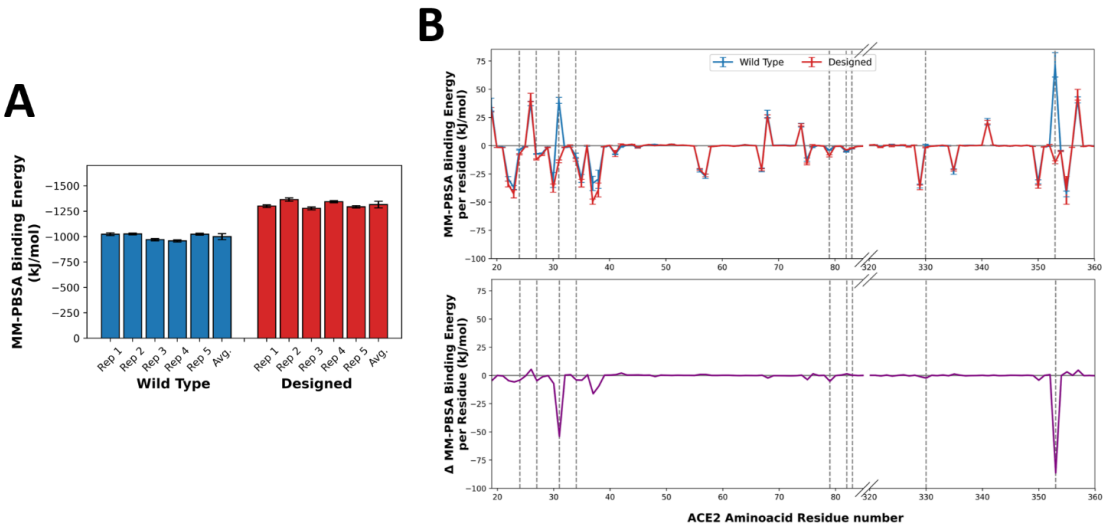


Figure 7. MM-PBSA analysis of the ACE2-RBD binding. (A) MM-PBSA binding energy for the *wt* and designed sequences. Results for each replica as well as the overall average are shown. (B) MM-PBSA binding energy contribution per residue for both *wt* and designed sequences as well as the difference between the two. Mutated residues are marked by dashed lines.

and if any new contacts appear. Looking at the left section of the binding interface, the interactions between L79W and Y83F with RBD F486 are also quite prevalent in the atomistic simulations. The support interactions established by M82L, Q24V and T27I are also present, with M82L also interacting with RBD F486 while Q24V and T27I interact non-specifically with RBD residues 473 to 477 and 486 to 489. On the middle section of the binding interface, K31F maintains its interaction with RBD Y489, however it seems to be quite promiscuous also interacting with neighboring RBD residues. H34I maintains its interaction with RBD L455 but also interacts quite significantly with RBD Q493, S494 and Y495. On the right section, the hydrophobic interaction between N330A and RBD T500 is maintained. K353W continues to establish quite complex interactions, being within interaction range of RBD residues 495 to 498, 501, 502, 504 and 505. In all, it seems as if most of these interactions are quite well maintained when simulating the Rosetta structure.

To compare the different binding interfaces, we can simply subtract the contact matrices from one another. We started by comparing the *wt* and HH sequences (Figure S5), where we found no significant differences in contact frequency between them. This result hints that the H374N and H378N mutations do not influence ACE2-RBD binding significantly. Experimental studies have also shown this to be the case[73]. Comparing the designed and *wt* sequences, however, we observe some significant differences (figure S6). Some of these differences are simple to interpret, such as the relative loss of contact between K31F with RBD P491 and subsequent increase in contact with RBD F486, in comparison with the *wt* contacts. Other more complex changes are also observed, such as the loss of contacts between N330A and RBD residues N501 and G502 with a possibly correlated increase in contacts by the non-mutated neighboring ACE2 residues T324, Q325, G236 and F327. In all, the number of contacts was mostly maintained, however one must consider that not all contacts have equal binding energies.

To attempt to quantify the binding energy of each simulated sequence and their respective residue contribution we employed the molecular mechanics Poisson-Boltzmann surface area (MM-PBSA) method using the g_mmpbsa tool[76,77]. MM-PBSA is an implicit solvent free energy method which can be used to determine the free energy of molecular systems and predict binding affinities. While the method has some weaknesses, the major being not predicting solute entropy, which leads to unrealistic energy estimate values for larger molecules, it has been successfully applied in drug and protein design by being able to determine relative binding affinity of similar structures with fairly high accuracy[78]. As such, we employed it in this work to probe the binding affinity of the *wt* and designed sequences and the specific contribution of each amino acid residue towards it.

Looking at the MM-PBSA binding energy (figure 7A), within the same tested sequence, all the 5 replicas behaved similarly, for both the *wt* and designed systems. This hints at quite a stable binding interaction during the timescale probed. It is worth noting that the total binding energy values recovered for either sequence are very unrealistic in scale, which is mostly a consequence of not accounting for entropic contributions in such large molecules. One must be careful to interpret the energy for these protein-protein complexes and rather look at their relative differences. This can be done in our case, as both proteins are overall quite similar, and both are analyzed binding the same partner.

Comparing the 2 systems (figure 7A), the designed sequence yielded a 31% increase in binding energy when compared to the *wt* (-1000.63 ± 29.81 and -1316.36 ± 32.50, for the wt and designed systems respectively). This is very similar to the relative increase of 27 % determined by

Rosetta, which points at our Rosetta models being in good agreement with the MD models.

Looking at the MM-PBSA binding energy per residue (figure 7B), one can compare the *wt* and designed binding interfaces. The binding energy per residue landscape is quite similar between the 2 sequences. Most mutations on the designed sequence yield contribute only with smaller improvements to the binding, but 2 mutations stand out significantly, K31F and K353W. Both these mutations also had the biggest impact when testing their reversion back to the native aminoacids with Rosetta (Figure 5B).

K31F is the substitution that now establishes a pi-stacking interaction with RBD Y489, and while our Rosetta scans showed that it greatly contributed towards binding, it was neither enriched or depleted in the enrichment map of aminoacid substitutions for increased RBD binding by Procko et al.[72]. K353W is the substitution that is placed in a tight pocket lined by the protein backbone of several residues but surrounded by several polar sidechains. This substitution was depleted on the enrichment map of substitutions, but it was found to be the best possible residue for that position when scanning with Rosetta.

Most of the gains from these 2 positions, come from the fact that the *wt* residues in place were quite unfavorable for binding. With this in mind, a few other *wt* residues which were not considered in this design stand out as they are slightly unfavorable for binding: K26, K68, K74, K94, K341 and R357. K26 is found to be under heavy selection pressure by Procko et al.[72], while the others were not. Some of these might be potential future targets for further optimization.

Summarizing these results, this designed sequence seems to be a stable and high affinity promising candidate for further *in vitro* testing. These tests may not only yield a promising therapeutic molecule for the neutralization of SARS-CoV-2 but will also be crucial to validate our current *in silico* protocol. Validation of our protocol and models will be crucial to further refine and improve our methods which will help in designing better molecules. Expression, purification, and characterization of this design will be carried out by the Protein Regulation in Plants laboratory at ITQB NOVA.

Future potential optimizations. Analyzing the outcome of the optimization often reveals potential future optimizations, not only for the ACE2 protein but also for the GA method. Obviously, some of these potential optimizations are dependent on establishing correlation between our computational and future experimental results.

Concerning the algorithm, further optimizations will likely be focused on improving the fitness function. In our current optimization the fitness function was merely the interface score as calculated by the Rosetta. However, it is likely that we can achieve better results with a more complex fitness function that considers not only the interface score, but also possibly the total complex score, the number of inter-protein hydrogen bonds, number of unsatisfied buried hydrogen bonds, etc. This can be done using a weighed linear combination of all these terms, calibrating the contribution of each specific term according to their importance.

Concerning further optimizations to the ACE2 protein, we found a few promising targets that were not explored in this work. Glycosilations, of either the ACE2 protein or the SARS-CoV-2 S protein RBD, were not considered in our optimizations. While taking glycosilations into account is not trivial, due to their immense conformational freedom, they can be considered in our Rosetta protocol.

Procko et al.[72] found that both ACE2 N90 and T92, which form a consensus N-glycosylation motif, are significant hot spots of enriched mutations that favor RBD binding. Any substitutions on either position, with the exception of T92S, prevents the N-glycosilation and as such it is likely that this glycosylation somehow hinders binding. Including point mutations to either of these residues in our designs is likely to be very favorable.

With the emergence of several SARS-CoV-2 variants, with several mutations to the spike protein and especially in the RBD, some of which with increased transmission ratios, there is now a need for therapeutical molecules which can universally target these variants. These different variants can be accounted for in our algorithm, by docking the design sequences not only with the *wt* ACE2 RBD but also dock them with the mutated variants and then use a weighed linear combination of the fitness value obtained for each variant as the overall fitness.

While some of these optimizations are specific to the ACE2 molecule, others are more general and will allow for a stronger algorithm which can be used for the fast and efficient screening of neutralizing molecules to fight not only SARS-CoV-2 but also other diseases.

CONCLUSIONS

By applying a GA-guided Rosetta docking protocol, we designed an optimized ACE2 protein for the neutralization of SARS-CoV-2 S-protein RBD. This approach allowed us to quickly probe the optimization sequence space, whilst having full control over the process. With this protocol we obtained similar results to that of the standard Rosetta design methods whilst being more computationally efficient.

The designed molecule proved to be a promising candidate for *in vitro* testing after extensive *in silico* characterization, showing significantly increased binding affinity ($\sim$ 30 %), whilst maintaining promising indicators of fold stability. This candidate molecule will now undergo in vitro testing and characterization by the Proteome Regulation in Plants lab at ITQB NOVA, which will provide definitive evidence on the success of the protocol.

ASSOCIATED CONTENT

Extended Methods and Materials providing additional analysis results on alternative mutations and ACE2-RBD contacts.

AUTHOR INFORMATION

Corresponding Author

Manuel N. Melo — Instituto de Tecnologia Química e Biológica António Xavier, Universidade Nova de Lisboa, Oeiras 2780-157, Portugal; orcid.org/0000-0001-6567-0513 ; m.n.melo@itqb.unl.pt

Author Contributions

L.B.A. and M.N.M. designed the GA-guided ROSETTA design method. All authors designed the simulation/validation systems. D.L. and C.M.S. performed atomistic simulations. L.B.A. performed all analysis.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENT

The authors thank Carlos Cruz, Isabel Abreu, Mariana Valério, Inês M. Luís and Raquel Nogueira for the helpful discussions and continuous input to this work.

REFERENCES

(1) Andersen, K. G.; Rambaut, A.; Lipkin, W. I.; Holmes, E. C.; Garry, R. F. The Proximal Origin of SARS-CoV-2. *Nature Medicine*. Nature Research April 1, 2020, pp 450–452.

(2) Wu, F.; Zhao, S.; Yu, B.; Chen, Y. M.; Wang, W.; Song, Z. G.; Hu, Y.; Tao, Z. W.; Tian, J. H.; Pei, Y. Y.; Yuan, M. L.; Zhang, Y. L.; Dai, F. H.; Liu, Y.; Wang, Q. M.; Zheng, J. J.; Xu, L.; Holmes, E. C.; Zhang, Y. Z. A New Coronavirus Associated with Human Respiratory Disease in China. *Nature* **2020**, *579* (7798), 265–269.

(3) Zhu, N.; Zhang, D.; Wang, W.; Li, X.; Yang, B.; Song, J.; Zhao, X.; Huang, B.; Shi, W.; Lu, R.; Niu, P.; Zhan, F.; Ma, X.; Wang, D.; Xu, W.; Wu, G.; Gao, G. F.; Tan, W. A Novel Coronavirus from Patients with Pneumonia in China, 2019. *N. Engl. J. Med.* **2020**, *382* (8), 727–733.

(4) Zhou, F.; Yu, T.; Du, R.; Fan, G.; Liu, Y.; Liu, Z.; Xiang, J.; Wang, Y.; Song, B.; Gu, X.; Guan, L.; Wei, Y.; Li, H.; Wu, X.; Xu, J.; Tu, S.; Zhang, Y.; Chen, H.; Cao, B. Clinical Course and Risk Factors for Mortality of Adult Inpatients with COVID-19 in Wuhan, China: A Retrospective Cohort Study. *Lancet* **2020**, *395* (10229), 1054–1062.

(5) Worldometer. Coronavirus Update (Live): Cases and Deaths from COVID-19 Virus Pandemic https://www.worldometers.info/coronavirus (accessed Apr 7, 2021).

(6) Li, F. Structure, Function, and Evolution of Coronavirus Spike Proteins. *Annual Review of Virology*. Annual Reviews Inc. September 29, 2016, pp 237–261.

(7) Hoffmann, M.; Kleine-Weber, H.; Schroeder, S.; Krüger, N.; Herrler, T.; Erichsen, S.; Schiergens, T. S.; Herrler, G.; Wu, N. H.; Nitsche, A.; Müller, M. A.; Drosten, C.; Pöhlmann, S. SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. *Cell* **2020**, *181* (2), 271-280.e8.

(8) Wan, Y.; Shang, J.; Graham, R.; Baric, R. S.; Li, F. Receptor Recognition by the Novel Coronavirus from Wuhan: An Analysis Based on Decade-Long Structural Studies of SARS Coronavirus. *J. Virol.* **2020**, *94* (7).

(9) Bosch, B. J.; van der Zee, R.; de Haan, C. A. M.; Rottier, P. J. M. The Coronavirus Spike Protein Is a Class I Virus Fusion Protein: Structural and Functional Characterization of the Fusion Core Complex. *J. Virol.* **2003**, *77* (16), 8801–8811.

(10) Walls, A. C.; Tortorici, M. A.; Snijder, J.; Xiong, X.; Bosch, B. J.; Rey, F. A.; Veesler, D. Tectonic Conformational Changes of a Coronavirus Spike Glycoprotein Promote Membrane Fusion. *Proc. Natl. Acad. Sci. U. S. A.* **2017**, *114* (42), 11157–11162.

(11) de Vries, R. D.; Schmitz, K. S.; Bovier, F. T.; Predella, C.; Khao, J.; Noack, D.; Haagmans, B. L.; Herfst, S.; Stearns, K. N.; Drew-Bear, J.; Biswas, S.; Rockx, B.; McGill, G.; Dorrello, N. V.; Gellman, S. H.; Alabi, C. A.; de Swart, R. L.; Moscona, A.; Porotto, M. Intranasal Fusion Inhibitory Lipopeptide Prevents Direct-Contact SARS-CoV-2 Transmission in Ferrets. *Science (80-.).* **2021**, *371* (6536), 1379–1382.

(12) Buchholz, U. J.; Bukreyev, A.; Yang, L.; Lamirande, E. W.; Murphy, B. R.; Subbarao, K.; Collins, P. L. Contributions of the Structural Proteins of Severe Respiratory Syndrome Coronavirus to Protective Immunity. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101* (26), 9804–9809.

(13) Deming, D.; Sheahan, T.; Heise, M.; Yount, B.; Davis, N.; Sims, A.; Suthar, M.; Harkema, J.; Whitmore, A.; Pickles, R.; West, A.; Donaldson, E.; Curtis, K.; Johnston, R.; Baric, R. Vaccine Efficacy in Senescent Mice Challenged with Recombinant SARS-CoV Bearing Epidemic and Zoonotic Spike Variants. *PLoS Med.* **2006**, *3* (12), 2359–2375.

(14) Suthar, M. S.; Zimmerman, M. G.; Kauffman, R. C.; Mantus, G.; Linderman, S. L.; Hudson, W. H.; Vanderheiden, A.; Nyhoff, L.; Davis, C. W.; Adekunle, O.; Affer, M.; Sherman, M.; Reynolds, S.; Verkerke, H. P.; Alter, D. N.; Guarner, J.; Bryksin, J.; Horwath, M. C.; Arthur, C. M.; Saakadze, N.; Smith, G. H.; Edupuganti, S.; Scherer, E. M.; Hellmeister, K.; Cheng, A.; Morales, J. A.; Neish, A. S.; Stowell, S. R.; Frank, F.; Ortlund, E.; Anderson, E. J.; Menachery, V. D.; Rouphael, N.; Mehta, A. K.; Stephens, D. S.; Ahmed, R.; Roback, J. D.; Wrammert, J. Rapid Generation of Neutralizing Antibody Responses in COVID-19 Patients. *Cell Reports Medicine*. medRxiv May 8, 2020.

(15) Robbiani, D. F.; Gaebler, C.; Muecksch, F.; Lorenzi, J. C. C.; Wang, Z.; Cho, A.; Agudelo, M.; Barnes, C. O.; Gazumyan, A.; Finkin, S.; Hägglöf, T.; Oliveira, T. Y.; Viant, C.; Hurley, A.; Hoffmann, H. H.; Millard, K. G.; Kost, R. G.; Cipolla, M.; Gordon, K.; Bianchini, F.; Chen, S. T.; Ramos, V.; Patel, R.; Dizon, J.; Shimeliovich, I.; Mendoza, P.; Hartweger, H.; Nogueira, L.; Pack, M.; Horowitz, J.; Schmidt, F.; Weisblum, Y.; Michailidis, E.; Ashbrook, A. W.; Waltari, E.; Pak, J. E.; Huey-Tubman, K. E.; Koranda, N.; Hoffman, P. R.; West, A. P.; Rice, C. M.; Hatziioannou, T.; Bjorkman, P. J.; Bieniasz, P. D.; Caskey, M.; Nussenzweig, M. C. Convergent Antibody Responses to SARS-CoV-2 in Convalescent Individuals. *Nature* **2020**, *584* (7821), 437–442.

(16) Wang, Q.; Zhang, Y.; Wu, L.; Niu, S.; Song, C.; Zhang, Z.; Lu, G.; Qiao, C.; Hu, Y.; Yuen, K. Y.; Wang, Q.; Zhou, H.; Yan, J.; Qi, J. Structural and Functional Basis of SARS-CoV-2 Entry by Using Human ACE2. *Cell* **2020**, *181* (4), 894-904.e9.

(17) Walls, A. C.; Park, Y. J.; Tortorici, M. A.; Wall, A.; McGuire, A. T.; Veesler, D. Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell* **2020**, *181* (2), 281-292.e6.

(18) Samavati, L.; Uhal, B. D. ACE2, Much More Than Just a Receptor for SARS-COV-2. *Front. Cell. Infect. Microbiol.* **2020**, *10*, 317.

(19) Michaud, V.; Deodhar, M.; Arwood, M.; Al Rihani, S. B.; Dow, P.; Turgeon, J. ACE2 as a Therapeutic Target for COVID-19; Its Role in Infectious Processes and Regulation by Modulators of the Raas System. *J. Clin. Med.* **2020**, *9* (7), 1–27.

(20) Samish, I. The Framework of Computational Protein Design. In *Methods in Molecular Biology*; Humana Press Inc., 2017; Vol. 1529, pp 3–19.

(21) Forsén, S. *Nobel Lectures in Chemistry 1971 – 1980*; WORLD SCIENTIFIC, 1993.

(22) Fischer, E. Nobel Lecture: Syntheses in the Purine and Sugar Group. Nobel lecture http://www.nobelprize.org/nobel_prizes/chemistry/laureates/1902/fischer-lecture.html (accessed Mar 4, 2021).

(23) Gutte, B. A Synthetic 70 Amino Acid Residue Analog of Ribonuclease S Protein with Enzymic Activity. *J. Biol. Chem.* **1975**, *250* (3), 889–904.

(24) Richardson, J. S.; Richardson, D. C. The de Novo Design of Protein Structures. *Trends Biochem. Sci.* **1989**, *14* (7), 304–309.

(25) Samish, I. Achievements and Challenges in Computational Protein Design. In *Methods in Molecular Biology*; Humana Press Inc., 2017; Vol. 1529, pp 21–94.

(26) DeGrado, W. F.; Prendergast, F. G.; Wolfe, H. R.; Cox, J. A. The Design, Synthesis, and Characterization of Tight-Binding Inhibitors of Calmodulin. *J. Cell. Biochem.* **1985**, *29* (2), 83–93.

(27) Kuhlman, B.; Dantas, G.; Ireton, G. C.; Varani, G.; Stoddard, B. L.; Baker, D. Design of a Novel Globular Protein Fold with Atomic-Level Accuracy. *Science (80-.).* **2003**, *302* (5649), 1364–1368.

(28) ˜widerek, K.; Tuñón, I.; Moliner, V.; Bertran, J. Computational Strategies for the Design of New Enzymatic Functions. *Archives of Biochemistry and Biophysics*. Academic Press Inc. September 15, 2015, pp 68–79.

(29) Yue, K.; Dill, K. A. Inverse Protein Folding Problem: Designing Polymer Sequences. *Proc. Natl. Acad. Sci. U. S. A.* **1992**, *89* (9), 4163–4167.

(30) Joh, N. H.; Wang, T.; Bhate, M. P.; Acharya, R.; Wu, Y.; Grabe,

M.; Hong, M.; Grigoryan, G.; DeGrado, W. F. De Novo Design of a Transmembrane Zn2+-Transporting Four-Helix Bundle. *Science (80-.).* **2014**, *346* (6216), 1520–1524.

(31) Huang, P.-S.; Love, J. J.; Mayo, S. L. A de Novo Designed Protein-Protein Interface. *Protein Sci.* **2007**, *16* (12), 2770–2774.

(32) Huang, P. S.; Boyken, S. E.; Baker, D. The Coming of Age of de Novo Protein Design. *Nature*. Nature Publishing Group September 14, 2016, pp 320–327.

(33) Kuhlman, B.; Bradley, P. Advances in Protein Structure Prediction and Design. *Nature Reviews Molecular Cell Biology*. Nature Publishing Group November 1, 2019, pp 681–697.

(34) Slovic, A. M.; Kono, H.; Lear, J. D.; Saven, J. G.; DeGrado, W. F. Computational Design of Water-Soluble Analogues of the Potassium Channel KcsA. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101* (7), 1828–1833.

(35) Slovic, A. M. Computational Design of a Water-Soluble Analog of Phospholamban. *Protein Sci.* **2003**, *12* (2), 337–348.

(36) Shifman, J. M.; Mayo, S. L. Exploring the Origins of Binding Specificity through the Computational Redesign of Calmodulin. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100* (23), 13274–13279.

(37) Kortemme, T.; Joachimiak, L. A.; Bullock, A. N.; Schuler, A. D.; Stoddard, B. L.; Baker, D. Computational Redesign of Protein-Protein Interaction Specificity. *Nat. Struct. Mol. Biol.* **2004**, *11* (4), 371–379.

(38) Potapov, V.; Reichmann, D.; Abramovich, R.; Filchtinski, D.; Zohar, N.; Ben Halevy, D.; Edelman, M.; Sobolev, V.; Schreiber, G. Computational Redesign of a Protein-Protein Interface for High Affinity and Binding Specificity Using Modular Architecture and Naturally Occurring Template Fragments. *J. Mol. Biol.* **2008**, *384* (1), 109–119.

(39) Leaver-Fay, A.; Tyka, M.; Lewis, S. M.; Lange, O. F.; Thompson, J.; Jacak, R.; Kaufman, K.; Renfrew, P. D.; Smith, C. A.; Sheffler, W.; Davis, I. W.; Cooper, S.; Treuille, A.; Mandell, D. J.; Richter, F.; Ban, Y. E. A.; Fleishman, S. J.; Corn, J. E.; Kim, D. E.; Lyskov, S.; Berrondo, M.; Mentzer, S.; Popović, Z.; Havranek, J. J.; Karanicolas, J.; Das, R.; Meiler, J.; Kortemme, T.; Gray, J. J.; Kuhlman, B.; Baker, D.; Bradley, P. Rosetta3: An Object-Oriented Software Suite for the Simulation and Design of Macromolecules. In *Methods in Enzymology*; Academic Press Inc., 2011; Vol. 487, pp 545–574.

(40) Kryshtafovych, A.; Schwede, T.; Topf, M.; Fidelis, K.; Moult, J. Critical Assessment of Methods of Protein Structure Prediction (CASP)—Round XIII. *Proteins Struct. Funct. Bioinforma.* **2019**, *87* (12), 1011–1020.

(41) Leman, J. K.; Weitzner, B. D.; Lewis, S. M.; Adolf-Bryfogle, J.; Alam, N.; Alford, R. F.; Aprahamian, M.; Baker, D.; Barlow, K. A.; Barth, P.; Basanta, B.; Bender, B. J.; Blacklock, K.; Bonet, J.; Boyken, S. E.; Bradley, P.; Bystroff, C.; Conway, P.; Cooper, S.; Correia, B. E.; Coventry, B.; Das, R.; De Jong, R. M.; DiMaio, F.; Dsilva, L.; Dunbrack, R.; Ford, A. S.; Frenz, B.; Fu, D. Y.; Geniesse, C.; Goldschmidt, L.; Gowthaman, R.; Gray, J. J.; Gront, D.; Guffy, S.; Horowitz, S.; Huang, P. S.; Huber, T.; Jacobs, T. M.; Jeliazkov, J. R.; Johnson, D. K.; Kappel, K.; Karanicolas, J.; Khakzad, H.; Khar, K. R.; Khare, S. D.; Khatib, F.; Khramushin, A.; King, I. C.; Kleffner, R.; Koepnick, B.; Kortemme, T.; Kuenze, G.; Kuhlman, B.; Kuroda, D.; Labonte, J. W.; Lai, J. K.; Lapidoth, G.; Leaver-Fay, A.; Lindert, S.; Linsky, T.; London, N.; Lubin, J. H.; Lyskov, S.; Maguire, J.; Malmström, L.; Marcos, E.; Marcu, O.; Marze, N. A.; Meiler, J.; Moretti, R.; Mulligan, V. K.; Nerli, S.; Norn, C.; Ó'Conchúir, S.; Ollikainen, N.; Ovchinnikov, S.; Pacella, M. S.; Pan, X.; Park, H.; Pavlovicz, R. E.; Pethe, M.; Pierce, B. G.; Pilla, K. B.; Raveh, B.; Renfrew, P. D.; Burman, S. S. R.; Rubenstein, A.; Sauer, M. F.; Scheck, A.; Schief, W.; Schueler-Furman, O.; Sedan, Y.; Sevy, A. M.; Sgourakis, N. G.; Shi, L.; Siegel, J. B.; Silva, D. A.; Smith, S.; Song, Y.; Stein, A.; Szegedy, M.; Teets, F. D.; Thyme, S. B.; Wang, R. Y. R.; Watkins, A.; Zimmerman, L.; Bonneau, R. Macromolecular Modeling and Design in Rosetta: Recent Methods and Frameworks. *Nature Methods*. Nature Research July 1, 2020, pp 665–680.

(42) Alford, R. F.; Leaver-Fay, A.; Jeliazkov, J. R.; O'Meara, M. J.; DiMaio, F. P.; Park, H.; Shapovalov, M. V.; Renfrew, P. D.; Mulligan, V. K.; Kappel, K.; Labonte, J. W.; Pacella, M. S.; Bonneau, R.; Bradley, P.; Dunbrack, R. L.; Das, R.; Baker, D.; Kuhlman, B.; Kortemme, T.; Gray, J. J. The Rosetta All-Atom Energy Function for Macromolecular Modeling and Design. *J. Chem. Theory Comput.* **2017**, *13* (6), 3031–3048.

(43) Renfrew, P. D.; Butterfoss, G. L.; Kuhlman, B. Using Quantum Mechanics to Improve Estimates of Amino Acid Side Chain Rotamer Energies. *Proteins Struct. Funct. Genet.* **2008**, *71* (4), 1637–1646.

(44) Gainza, P.; Nisonoff, H. M.; Donald, B. R. Algorithms for Protein Design. *Current Opinion in Structural Biology*. Elsevier Ltd August 1, 2016, pp 16–26.

(45) Dunbrack, R. L. Rotamer Libraries in the 21st Century. *Current Opinion in Structural Biology*. Elsevier Ltd August 1, 2002, pp 431–440.

(46) Vikram K. Mulligan. Optimizing Side-Chains: Introduction to the Packer https://www.rosettacommons.org/demos/latest/tutorials/Optimizing_Sidechains_The_Packer/Optimizing_Sidechains_The_Packer (accessed Mar 9, 2021).

(47) Robert, C. P. The M Etropolis– H Astings Algorithm . In *Wiley StatsRef: Statistics Reference Online*; John Wiley & Sons, Ltd: Chichester, UK, 2015; pp 1–15.

(48) Leaver-Fay, A.; Jacak, R.; Stranges, P. B.; Kuhlman, B. A Generic Program for Multistate Protein Design. *PLoS One* **2011**, *6* (7), 20937.

(49) Havranek, J. J.; Harbury, P. B. Automated Design of Specificity in Molecular Recognition. *Nat. Struct. Biol.* **2003**, *10* (1), 45–52.

(50) Ashworth, J.; Taylor, G. K.; Havranek, J. J.; Quadri, S. A.; Stoddard, B. L.; Baker, D. Computational Reprogramming of Homing Endonuclease Specificity at Multiple Adjacent Base Pairs. *Nucleic Acids Res.* **2010**, *38* (16), 5601–5608.

(51) Leaver-Fay, A. Multistate design and distributing the load with MPI, "mpi_msd" https://www.rosettacommons.org/docs/latest/application_documentation/design/mpi-msd (accessed Mar 12, 2021).

(52) Yoshida, M.; Hinkley, T.; Tsuda, S.; Abul-Haija, Y. M.; McBurney, R. T.; Kulikov, V.; Mathieson, J. S.; Galiñanes Reyes, S.; Castro, M. D.; Cronin, L. Using Evolutionary Algorithms and Machine Learning to Explore Sequence Space for the Discovery of Antimicrobial Peptides. *Chem* **2018**, *4* (3), 533–543.

(53) Lan, J.; Ge, J.; Yu, J.; Shan, S.; Zhou, H.; Fan, S.; Zhang, Q.; Shi, X.; Wang, Q.; Zhang, L.; Wang, X. Structure of the SARS-CoV-2 Spike Receptor-Binding Domain Bound to the ACE2 Receptor. *Nature* **2020**, *581* (7807), 215–220.

(54) Nivón, L. G.; Moretti, R.; Baker, D. A Pareto-Optimal Refinement Method for Protein Design Scaffolds. *PLoS One* **2013**, *8* (4), e59004.

(55) Conway, P.; Tyka, M. D.; DiMaio, F.; Konerding, D. E.; Baker, D. Relaxation of Backbone Bond Geometry Improves Protein Energy Landscape Modeling. *Protein Sci.* **2014**, *23* (1), 47–55.

(56) Khatib, F.; Cooper, S.; Tyka, M. D.; Xu, K.; Makedon, I.; Popović, Z.; Baker, D.; Players, F. Algorithm Discovery by Protein Folding Game Players. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108* (47), 18949–18953.

(57) Tyka, M. D.; Keedy, D. A.; André, I.; Dimaio, F.; Song, Y.; Richardson, D. C.; Richardson, J. S.; Baker, D. Alternate States of Proteins Revealed by Detailed Energy Landscape Mapping. *J. Mol. Biol.* **2011**, *405* (2), 607–618.

(58) 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.

(59) Lindahl; Abraham; Hess; Spoel, van der. GROMACS 2020.3 Source Code. **2020**.

(60) DdGScan https://new.rosettacommons.org/docs/latest/scripting_documentation/RosettaScripts/Filters/filter_pages/DdGScanFilter (accessed Apr 14, 2021).

(61) Fleishman, S. J.; Leaver-Fay, A.; Corn, J. E.; Strauch, E. M.; Khare, S. D.; Koga, N.; Ashworth, J.; Murphy, P.; Richter, F.; Lemmon, G.; Meiler, J.; Baker, D. Rosettascripts: A Scripting Language Interface to the Rosetta Macromolecular Modeling Suite. *PLoS One* **2011**, *6* (6), e20161.

(62) Revert design to native application https://new.rosettacommons.org/docs/latest/application_d

ocumentation/design/revert-design-to-native (accessed Apr 14, 2021).
(63) Fleishman, S. J.; Whitehead, T. A.; Ekiert, D. C.; Dreyfus, C.; Corn, J. E.; Strauch, E. M.; Wilson, I. A.; Baker, D. Computational Design of Proteins Targeting the Conserved Stem Region of Influenza Hemagglutinin. *Science (80-.).* **2011**, *332* (6031), 816–821.
(64) Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. Ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from Ff99SB. *J. Chem. Theory Comput.* **2015**, *11* (8), 3696–3713.
(65) Mark, P.; Nilsson, L. Structure and Dynamics of the TIP3P, SPC, and SPC/E Water Models at 298 K. *J. Phys. Chem. A* **2001**, *105* (43), 9954–9960.
(66) Berendsen, H. J. C.; Postma, J. P. M.; Van Gunsteren, W. F.; Dinola, A.; Haak, J. R. Molecular Dynamics with Coupling to an External Bath. *J. Chem. Phys.* **1984**, *81* (8), 3684–3690.
(67) Bussi, G.; Donadio, D.; Parrinello, M. Canonical Sampling through Velocity Rescaling. *J. Chem. Phys.* **2007**, *126* (1), 014101.
(68) Parrinello, M.; Rahman, A. Polymorphic Transitions in Single Crystals: A New Molecular Dynamics Method. *J. Appl. Phys.* **1981**, *52* (12), 7182–7190.
(69) 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.
(70) 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.
(71) Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M. LINCS: A Linear Constraint Solver for Molecular Simulations. *J. Comput. Chem.* **1997**, *18* (12), 1463–1472.
(72) Chan, K. K.; Dorosky, D.; Sharma, P.; Abbasi, S. A.; Dye, J. M.; Kranz, D. M.; Herbert, A. S.; Procko, E. Engineering Human ACE2 to Optimize Binding to the Spike Protein of SARS Coronavirus 2. *Science (80-.).* **2020**, *369* (6508), 1261–1265.
(73) Svilenov, H. L.; Sacherl, J.; Reiter, A.; Wolff, L.; Chen, C.-C.; Stern, M.; Wachs, F.-P.; Simonavicius, N.; Pippig, S.; Wolschin, F.; Buchner, J.; Brockmeyer, C.; Protzer, U. Efficient Inhibition of SARS-CoV-2 Strains by a Novel ACE2-IgG4-Fc Fusion Protein with a Stabilized Hinge Region. *bioRxiv*. bioRxiv December 7, 2020, p 2020.12.06.413443.
(74) Towler, P.; Staker, B.; Prasad, S. G.; Menon, S.; Tang, J.; Parsons, T.; Ryan, D.; Fisher, M.; Williams, D.; Dales, N. A.; Patane, M. A.; Pantoliano, M. W. ACE2 X-Ray Structures Reveal a Large Hinge-Bending Motion Important for Inhibitor Binding and Catalysis. *J. Biol. Chem.* **2004**, *279* (17), 17996–18007.
(75) Nami, B.; Ghanaeian, A.; Ghanaeian, K.; Nami, N. The Effect of ACE2 Inhibitor MLN-4760 on the Interaction of SARS-CoV-2 Spike Protein with Human ACE2: A Molecular Dynamics Study. *ChemRxiv*. ChemRxiv April 21, 2020.
(76) Kumari, R.; Kumar, R.; Lynn, A. G-Mmpbsa -A GROMACS Tool for High-Throughput MM-PBSA Calculations. *J. Chem. Inf. Model.* **2014**, *54* (7), 1951–1962.
(77) Baker, N. A.; Sept, D.; Joseph, S.; Holst, M. J.; McCammon, J. A. Electrostatics of Nanosystems: Application to Microtubules and the Ribosome. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98* (18), 10037–10041.
(78) Homeyer, N.; Gohlke, H. Free Energy Calculations by the Molecular Mechanics Poisson–Boltzmann Surface Area Method. *Mol. Inform.* **2012**, *31* (2), 114–122.

Article VI: SARS-CoV-2 variants impact RBD conformational dynamics and ACE2 accessibility

SARS-CoV-2 variants impact RBD conformational dynamics and ACE2 accessibility

Mariana Valério[†,§], Luís Borges-Araújo[†,⊥,‡,§], Manuel N. Melo[†,*], Diana Lousa[†,*], Cláudio M. Soares[†,*]

† Instituto de Tecnologia Química e Biológica António Xavier, Universidade Nova de Lisboa, Av. da República, 2780-157 Oeiras, Portugal.

⊥ iBB-Institute for Bioengineering and Biosciences, Instituto Superior Técnico, Universidade de Lisboa, Lisbon, Portugal

‡ Associate Laboratory i4HB—Institute for Health and Bioeconomy at Instituto Superior Técnico, Universidade de Lisboa, Lisbon, Portugal

§ M.V. and L.B.A. contributed equally to this work.

ABSTRACT: The coronavirus disease 2019 (COVID-19) pandemic, caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), has killed over 5 million people and is causing a devastating social and economic impact all over the world. The rise of new variants of concern (VOCs) represents a difficult challenge due to the loss vaccine and natural immunity, and increased transmissibility. All circulating VOCs contain mutations in the spike glycoprotein, which mediates fusion between the viral and host cell membranes, via its receptor binding domain (RBD) that binds to angiotensin-converting enzyme 2 (ACE2). In an attempt to understand the effect of RBD mutations in circulating VOCs, a lot of attention has been given to the RBD-ACE2 interaction. However, this type of analysis is limited, since it ignores more indirect effects, such as the conformational dynamics of the RBD itself. Observing that some VOCs mutations occur in residues that are not in direct contact with ACE2, we hypothesized that they could affect RBD conformational dynamics. To test this, we performed long atomistic (AA) molecular dynamics (MD) simulations to investigate the structural dynamics of *wt* RBD, and that of three circulating VOCs (alpha, beta, and delta). Our results show that in solution, wt RBD presents two distinct conformations: an "open" conformation where it is free to bind ACE2; and a "closed" conformation, where the RBM ridge blocks the binding surface. The alpha and beta variants significantly impact the open/closed equilibrium, shifting it towards the open conformation by roughly 20%. This shift likely increases ACE2 binding affinity. Simulations of the currently predominant delta variant RBD were extreme in this regard, in that a closed conformation was never observed. Instead, the system alternated between the before mentioned open conformation and an alternative "reversed" one, with a significantly changed orientation of the RBM ridge flanking the RBD. This alternate conformation could potentially provide a fitness advantage not only due to increased availability for ACE2 binding, but also by aiding antibody escape through epitope occlusion. These results support the hypothesis that VOCs, and particularly the delta variant, impact RBD conformational dynamics in a direction that simultaneously promotes efficient binding to ACE2 and antibody escape.

INTRODUCTION

Coronavirus disease 2019 (COVID-19), caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2)[1–3], is a global pandemic with higher mortality than that of seasonal influenza[4]. As of November 2021, 0ver 5 million lives had been claimed by this disease[5]. Infection by SARS-CoV-2 requires the fusion of viral and host cell membranes, at either the cell surface or the endosomal membrane[6]. As for the severe acute respiratory syndrome coronavirus (SARS-CoV) and the Middle East respiratory syndrome-related coronavirus (MERS-CoV), the SARS-CoV-2 fusion process is mediated by the viral envelope spike (S) glycoprotein[6]. Upon viral attachment or uptake, host factors trigger large-scale conformational rearrangements in the S protein, including a refolding step that leads directly to membrane fusion and viral entry [7–12].

The SARS-CoV-2 S protein is composed of a signal peptide located at the N-terminus (residues 1-13) and 2 subunits, S1 (residues 14-685) and S2 (residues 686-1273)[13]. The S1 and S2 subunits are responsible for receptor binding and membrane fusion, respectively[13]. The S1 subunit consists of a N-terminal domain (residues 14-305) and a receptor binding domain, or RBD (residues 319-541). The RBD is responsible for the interaction of SARS-CoV-2 with host cells via binding to the angiotensin-converting enzyme 2 (ACE2)[8,10,13,14], a regulator of the renin-angiotensin system. Binding to ACE2 is one of the first steps in what is considered to be the main mode of SARS-CoV-2 viral entry, and as such, a lot of attention has been given to the SARS-CoV-2

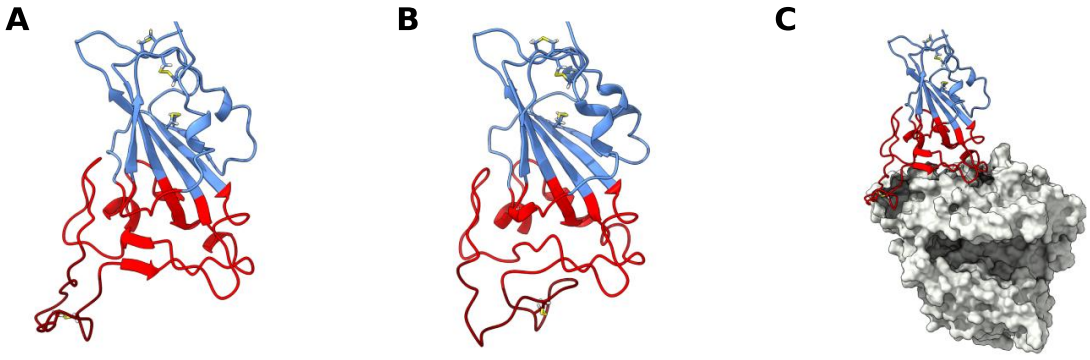


Figure 1. SARS-CoV-2 receptor binding domain (RBD) structure. Structure of *wt* RBD in the open (A) and closed (B) conformations. Snapshots obtained from the AA MD simulations. Disulfide bonds are represented in yellow sticks. Structure of *wt* RBD bound to ACE2 is also shown (C). The RBM region is colored red and the ridge in dark red, with the rest of the protein being colored in blue. ACE2 is in grey.

RBD – ACE2 complex due to both its mechanistic implications[15–20] and pharmaceutical potential[21–27]. However, not much attention has been given to the dynamics of the RBD by itself.

The RBD core structure when bound to ACE2 (Figure 1A) consists of a twisted five stranded antiparallel β sheet (β1, β2, β3, β4 and β7), with short connecting helices and loops[28]. This core β sheet structure is further stabilized by 3 disulfide bonds. Between the core β4 and β7 strands (residues 438-506), there is an extended region containing 2 short β strands (β5 and β6), the alpha 4 and alpha 5 helices and loops. This region is the receptor-binding motif (RBM), which contains most of the residues that are responsible for interacting with ACE2[28,29]. When complexed with ACE2, the RBM folds into a concave surface, that accommodates the N-terminal α-helix of ACE2, with a ridge (residues 471 to 491) on one side, formed by a disulfide-bridge-stabilized loop (Cys480–Cys488). It is in this surface that several RBM residues establish specific and non-specific interaction with ACE2 residues[28].

From the available experimental structural data the core β-sheet structure is very stable, but the RBM seems to be quite dynamic and not as structurally defined, unless bound to other proteins, like ACE2[14,30–33] or antibody fragments[34–40]. Molecular dynamics (MD) simulation studies have also mostly focused on RBD complexed with these proteins, and while there are MD simulation studies of free RBD, they have either been short simulations[41–43] or not focused on RBM dynamics[41,44]. As such, not much is known about the conformational dynamics of this motif when unbound. This is relevant because the conformational dynamics of the SARS-CoV-2 RBD and RBM might not only play an important role in receptor recognition and binding but also provide important information for the development of newer improved pharmaceuticals.

Recently, a significant number of naturally occurring mutations to the SARS-CoV-2 S protein have also been reported[45–48]. Many of these mutations have been identified in the RBD, some of which have rapidly become the dominant viral variant in certain regions due to their significant fitness advantage[45–48]. Many of these RBD mutations are thought to increase fitness by increasing ACE2 binding affinity or by escaping neutralization by anti-SARS-CoV-2 monoclonal antibodies[49]. Still, the impact of these mutations on the structural dynamics of RBD and the RBM have not yet been investigated.

In this work, we use atomistic (AA) molecular dynamics (MD) simulation methods to investigate the structural dynamics of SARS-CoV-2 RBD, and that of three naturally occurring variants of concern (VOCs): B.1.1.7, or alpha, variant[47] (N501Y); B.1.351, or beta, variant[45] (K417N E484K N501Y); and B.1.617.2, or delta, variant[46] (L452R T478K). Our results show that the RBM dynamics of *wt* RBD are such that it is not always in a conformation competent for ACE2 binding (Figure 1). Conversely, all variants, and delta in particular, stabilize binding-competent configurations. The conformational space visited by the variants thus putatively increases ACE2 binding efficiency and may further provide fitness advantage by aiding in antibody escape.

METHODS

Molecular dynamics simulations. All atomistic simulations were performed with the GROMACS 2020.3[50,51] package and modelled using the Amber14sb[52], forcefield alongside the TIP3P water model[53]. The initial *wt* RBD structure was obtained from PDB ID: 6M0J[30], which corresponds to an ACE2 bound conformation of RBD; ACE2 was excluded from this structure. The different RBD variants were generated by mutating the appropriate residues in the *wt* RBD using PyMOL[54]. Simulations were performed on each RBD protein structure in water. Each structure was inserted in a truncated dodecahedron box filled with water molecules (considering a minimum distance of 1.2 nm between protein and box walls). The total charge of the system was neutralized with the required number of Na^+ ions, with additional Na^+ and Cl^- ions added to the solution to reach an ionic strength of 0.1 M.

The system was energy-minimized using the steepest descent method for a maximum of 50000 steps with position restraints on the heteroatom positions by restraining them to the crystallographic coordinates using a force constant of 1000 kJ/mol in the X, Y and Z positions. Before performing

the production runs, an initialization process was carried out in 5 stages of 100 ps each. Initially, all heavy-atoms were restrained using a force constant of 1000 kJ/mol/nm, and at the final stage only the only Cα atoms were position-restrained using the same force constant. In the first stage, the Berendsen thermostat[55] was used to initialize and maintain the simulation at 300 K, using a temperature coupling constant of 0.01 ps, without pressure control. The second stage continued to use the Berendsen thermostat but now with a coupling constant of 0.1 ps. The third stage kept the same temperature control, but introduced isotropic pressure coupling with the Berendsen barostat[55], with a coupling constant of 5.0 ps. The fourth stage changed the thermostat to V-rescale[56], with a temperature coupling constant of 0.1 ps, and the barostat to Parrinello-Rahman[57] with a pressure coupling constant of 5.0 ps. The fifth stage is equal to the fourth stage, but position restraints are only applied on Cα atoms. For production simulations, conditions were the same as for the fifth stage, but without any restraints. In all cases, 2 fs integration steps were used. Long-range electrostatic interactions were treated with the PME[58,59] scheme, using a grid spacing of 0.12 nm, with cubic interpolation. The neighbor list was updated every twenty steps with a Verlet cutoff with a 0.8 nm radius. All bonds were constrained using the LINCS algorithm[60].

Simulations of each of the RBD proteins were performed for at least 7 μs over 5 replicates (the *wt* was simulated for 15 μs, and the alpha, beta and delta variants for 7 μs each). The first 3 μs of simulation were considered as equilibration time and the remaining frames were used for analysis. Visualization and rendering of simulation snapshots was performed with the molecular graphics viewers VMD[61], PyMOL[54] and UCSF Chimera[62].

Principal Component Analysis. PCA is a standard dimensionality reduction method that we apply here to the (3N-6)-dimensional space of possible RBD conformations (in our case, N being the number of RBD residues). PCA consists of a linear transformation that changes a set of possibly correlated dimensions into a set of linearly uncorrelated, mutually orthogonal ones, called principal components (PCs). The first PC can be defined as the direction that accounts for as much of the variance in the data as possible, with each successive PC accounting for as much of the remaining variance as possible. Reduction of data dimensionality is achieved by retaining only a few of the first PCs — which represent the strongest correlations in the data, in our case, the most important conformational motions —, thus sacrificing some information for simplicity. Discussions of the mathematical and computational backgrounds can be found elsewhere[63–66].

In this work, PCA was applied to sets of conformational coordinates obtained from MD simulations. Prior to PCA, each conformation was translationally and rotationally fitted to the RBD core Cα carbons of the *wt* crystal structure (hence the –6 in the dimensionality). PCs were determined using MDAnalysis[67], from the entire pool of simulation trajectories, considering only the coordinates of the RBD's Cα carbons. The dimensionality was reduced to the 2 most representative PCs, preserving a large part of the variance. RBD structures for each simulation frame, for each variant, could then be projected as points in this two-dimensional space, enabling a simplified visual representation of the conformation space explored by the RBD in each case.

The probability density function for each trajectory projection was estimated using a gaussian kernel estimator[65,68] implemented in LandscapeTools' *get_density* software as described elsewhere[65,69]. This procedure defines a probability density function P(r), with the values of P(r) being stored for the position of each data point and for the nodes of a two-dimensional uniform grid, with a mesh size of 0.5 Å. These values were used to define an energy surface, calculated as[65]:

$$E(r) = -RT \ln\left(\frac{P(r)}{P_{max}}\right)$$

Where P_{max} is the maximum of the probability density function, P(r). The energy surface landscapes were analyzed by determining the energy minima and respective basins. The basins were defined as the set of all conformations whose steepest descent path along the energy surface leads to a particular minimum[65,70,71]. Here, the steepest descent paths for each grid cell were computed, with each conformation inheriting the path of its corresponding grid cell. Landscape regions with $E > 6\ k_BT$ were discarded, resulting in the final set of basins for each data set.

Residue interaction network analysis. Residue interaction networks (RINs) are graph representations of protein structures, where the nodes represent amino acid residues and the edges represent interactions between residues. Pairwise residue interactions were analyzed for the 5000 lowest energy conformations obtained for the most populated open, closed and reversed conformation basins of the energy surface landscapes of each RBD variant, using RIP-MD[72]. Several types of interactions between AAs were probed: Cα contacts, hydrogen bonds, salt bridges, disulfide bonds, cation-π, π–π, Arg-Arg, Coulomb and van der Waals. The parameters defining each interaction, as well as their mathematical formulation can be found elsewhere[72]. Once the interactions were determined, the interaction networks were visualized using Cytoscape[73].

RESULTS AND DISCUSSION

Our aim was to study the conformational dynamics of the SARS-CoV-2 RBD, as well as that of several other SARS-CoV-2 variants in solution. To this effect, we simulated the *wt*, alpha, beta and delta variants of the SARS-CoV-2 RBD. The gamma variant was not studied due to its similarity to the beta variant (in the RBD region, both variants share the E484K and N501Y mutations; the beta variant also contains the K417N mutation while the gamma variant has K417T).

***wt* RBD presents two distinct RBM conformations in aqueous solution.** Visual inspection of the trajectories obtained in the simulation of *wt* RBD in water revealed that large dynamic conformational changes occur in the RBM region (Figure 2A, Supplementary Video S1). The dynamics observed appear to show an opening and closing of the ACE2 binding surface of the RBM. To better characterize these conformational dynamics, we performed principal component analysis (PCA) on the coordinates recovered

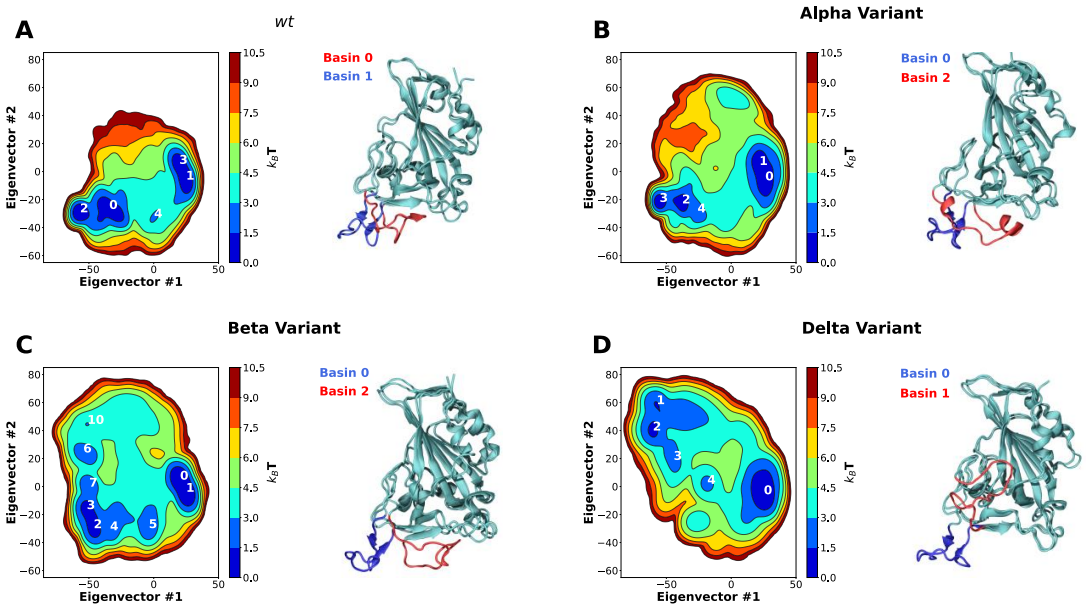


Figure 2. Two-dimension principal component analysis (PCA) of SARS-CoV-2 RBD conformational dynamics in water. Plots of the first two principal components determined from the Cα backbone of the *wt* RBD (A) as well as the alpha (B), beta (C) and delta (D) variants. Basins with $K_BT < 3$ are numbered in each figure. Snapshots of the lowest energy structures for selected open and closed basins are also shown. The ridge regions of the open and closed snapshots are colored in blue and red, respectively.

from these simulations, reducing them to 2 principal components; this 2D configuration space sampling was expressed as free energy landscapes (Figure 2).

For *wt* RBD, we observe two deep basin clusters (Figure 2A), as well as several other lesser basins. Closer analysis on the RBD conformations that make up each basin shows that *wt* basins 1 and 3 correspond to conformations close to the ACE2-bound one determined by X-ray crystallography[30] (Figure 1A and 2A). We named these "open" configurations. The second basin cluster (basins 0 and 2), however, was made up by quite distinct conformations. In these basins, the loop that makes up the RBM has twisted, and collapsed over the ACE2 binding surface, effectively hiding it from the solvent (Figure 1B and 2A). We named these conformations "closed". Further analysis of the PCA results reveals that the *wt* RBD is in a closed state for more than half of the simulation time (~55.5%, Supplementary Table S1). Given that in these conformations the RBM closes on itself, hiding the ACE2 binding surface, we can speculate that RBD would be unable to effectively bind to ACE2 and initiate the ACE2-dependent infection process. Moreover, the open and closed states were visited reversibly (Supplementary Figure S1), indicating that our simulations were not kinetically trapped in either basin.

Residue interaction network (RIN) analysis was performed for the 5000 lowest energy structures of basins 1 (open) and 0 (closed). From the identified interactions, we selected those that were present in over 50% of the simulation frames (Supplementary Figure S3). We also only considered interactions that are established by RBM residues, or those in their immediate vicinity. These RINs were then used to probe the different intramolecular interactions established in each of the conformations.

In the open conformation, the RBD ridge is stabilized by a triple π-stacking interaction between residues Y489-F456-Y473 and a hydrogen bond between Y489-Y473. Additionally, two hydrogen bonds are established between residues Y453 and E493, which help stabilize the formation of a small β-sheet (Figure 3A).

In the closed conformation, however, the π-stacking interactions are broken, and new interactions with RBD core residues are formed in their place. F456 forms a stable π-stacking with Y421, Y489 forms a transient π-stacking interaction with F486 and Y473 forms a hydrogen bond with the backbone of Y451. Moreover, E484 forms a salt bridge with R403, that is found in the RBD core, and a hydrogen bond with K417 (Figure 3E). This hydrogen bond does not show up in the RIN, as K417 can establish a bond with each of the two glutamate oxygens, each with ~40% prevalence (each thus below our 50% selection cutoff). These two interactions, together with the formation of three hydrogen bonds (C480–S494–G482–Q493) are responsible for the "closing" of the ridge and consequent shielding of the ACE2 binding surface. The importance of the E484-R403 and E484–K417 interactions for the closing of the loop was confirmed by simulating the E484K and K417N mutants. Either of these single mutations were enough to completely deplete the closed conformation (Supplementary Figure S2A and B for E484K and K417N, respectively). This shows that both these interactions are crucial for the stabilization of

the *wt* closed state. Still, several other transient hydrogen bonds, formed between residues L492, G493 and S494 of strand β6, and T478, C480, N481, G482 and E484 of the RBM ridge, assist in stabilizing the structure.

The closed conformation does not seem to substantially impact RBD secondary structure (Supplementary Figure S7). The largest impact appears to be limited to residues 473-474 and 488-489, that in the open state display a slight β-sheet character. However, upon closing, this β-sheet character disappears. This effect comes from residues 473 and 489 no longer participating in the triple π-stacking that was likely stabilizing this region.

Apart from impacting ACE2 accessibility, the closing of the RBM ridge also decreases the solvent accessible surface area (SASA) of the RBD by slightly over 3 % (Supplementary Table S2).

Although other studies have noted the high flexibility in the RBM region of the RBD[41–43], this is, as far as we know, the first report of this "hinge" mechanism which can effectively hide the ACE2 binding surface of the RBD from binding partners. While it is likely that induced fit interactions might assist in opening a closed conformation for binding to ACE2, it is safe to assume that the closed conformation will have its binding to ACE2 substantially hindered when compared to an open conformation.

SARS-CoV-2 alpha and beta variants impact RBM conformational dynamics and exposure. The first SARS-CoV-2 variant of concern to be identified was first detected in the UK. It is often referred to as B.1.1.7 or alpha variant and has only one mutation in the RBD region — N501Y. A second variant emerged soon after in South Africa, independently of B.1.1.7, referred to as B.1.351 or beta variant. In the RBD region, this variant shares the N501Y mutation with the alpha variant and includes two others: K417N and E484K[47].

Like the *wt* variant, MD simulations of the RBDs from the alpha and beta variants also showed the prevalence of two sets of RBM conformations, corresponding to open and closed conformations (Supplementary Videos S2 and S3). PCA analysis of the alpha variant trajectory shows two deep basin clusters (Figure 2B), basins 0 and 1, and basins 2 and 3, which correspond to open and closed conformations respectively. However, unlike the *wt* variant, the alpha variant remains most of the simulation time in an open conformation (~72.64 %, Supplementary Table S1). The beta variant (Figure 2C) also has two deep basin clusters (basins 0 and 1, and basins 2 and 3), corresponding to open and closed conformations, respectively. Like the alpha variant, beta remains in an open conformation for substantially longer time than the *wt* (~69 %, Supplementary Table S1). In both cases, and as for *wt*, our simulations were able to reversibly visit either state (Supplementary Figure S1).

Both alpha and beta variants shift the open/closed equilibrium towards more open conformations by roughly 20%. A closing ΔΔG was calculated from the ratio between time spent in the open and closed states, where the time spent in each individual open and closed basin was added together (Supplementary Table S1). The equilibrium shift led to a decrease in the closing ΔΔG from 0.55 ± 0.17 kJ/mol, in the case of *wt* RBD, to -2.44 ± 0.22 and -2.09 ± 0.14 kJ/mol, for the alpha and beta variants, respectively. As mentioned previously, it is likely that only the open conformations are fully available to bind to ACE2, meaning that these mutations substantially increase the accessibility of RBD to ACE2, and probably impact ACE2-RBD binding.

By analyzing the intramolecular residue interactions for both variants we observe that the interactions which stabilize the open conformation in the *wt* variant are conserved in both alpha and beta variants, namely the triple π-stacking between residues Y489–F456–Y473, as well as the hydrogen bond between Y489 and Y473. An additional hydrogen bond between Q493 and Y453 assists in stabilizing the β6 strand (Figure 3B and 3C).

Interestingly, in both the open and closed conformations of the alpha variant, the interactions established by residue Y501 (alpha's only mutation in the RBD) that were previously present in the *wt* variant are maintained in the alpha variant (two hydrogen bonds established through the residue backbones: Q458–Y501 and Y501–Q506). However, the main interactions that stabilize the closed conformations differ between the alpha variant and *wt* (although some transient hydrogen bonds between strand β6 and the RBM ridge do remain). Instead of the E484–R403 salt bridge seen for *wt*, in the alpha variant the closed conformation is promoted by the formation of hydrophobic interactions between the mutated Y501, V483 and F486 (Figure 3F). This arrangement hinders the establishment of the E484–R403 salt-bridge (as can be seen in Supplementary Video 2) while being itself less stable than the open conformations. This is the likely cause for the decrease in percentage of closed state observed for alpha. Progression to the E484–R403 salt-bridge may also be prevented in part by the establishment of a short α-helix, discussed ahead.

In the beta variant, the closed conformation is notably impacted by both the E484K and the N501Y mutations. The E484K mutation prevents the formation of the E484–R403 salt bridge that was crucial for the stability of the closed conformation in the *wt* variant. However, unlike the single E484K mutant (Supplementary Figure S2), the beta variant can still reach a closed conformation. This is because it can establish the same hydrophobic interaction between Y501 and V483 as the alpha variant (Figure 3G). This closed state is also stabilized by the same transient hydrogen bonds between strand β6 and the RBM ridge seen in the *wt* and alpha variants

Concerning the secondary structure, there are no substantial differences between the alpha or beta open states and the *wt* open state (Figure S7). However, upon closing, both alpha and delta form a small α-helix between residues 475 and 490, for roughly 30% of the simulation time. This helical character might be relevant for the alpha variant, as it assists in facing the E484 sidechain away from R403 (Figure 3F and 3G), hindering the formation of the salt-bridge. Additionally, the alpha variant also shows some helicity in residues 482 to 489, which likely arises from contacts between residues in this helix and the mutated N501Y.

Curiously, while the alpha variant also shows a considerable decrease in SASA upon closing (~5%), the beta variant shows no substantial change.

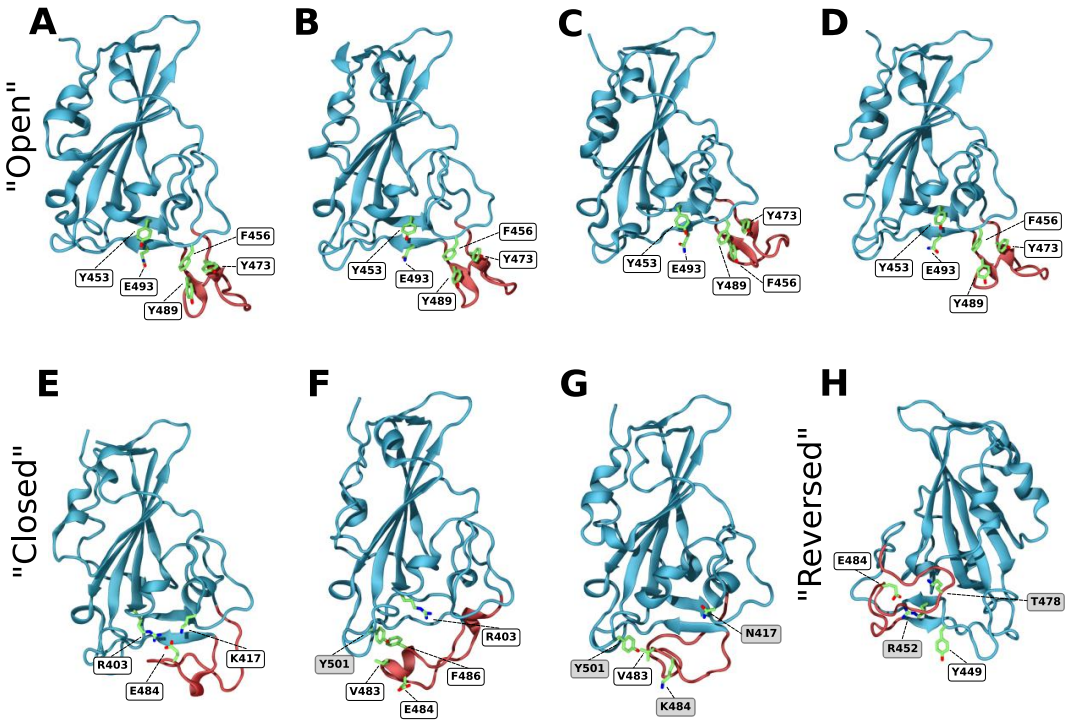


Figure 3. Closeup snapshots of SARS-CoV-2 RBD intramolecular interactions that stabilize the various conformations. Snapshots from AA MD simulations showcasing crucial intramolecular interactions responsible for stabilizing the open, closed, and reversed conformations for the *wt* (A, E), alpha (B, F), beta (C, G) and delta (D, H) RBD variants. The ridge region of the RBD is colored in red and residues of interest in green. Text labels indicate relevant residues, with shaded labels indicating mutations relative to *wt*. All figures are rotated 180° relative to Figures 1 and 2, apart from snapshot H.

Overall, these results showcase a possible alternative mechanism for how the alpha and beta variants might facilitate viral entry into the host cells. By shifting the open/closed equilibrium towards the ACE2-accessible open conformation, both of these variants are facilitating ACE2–RBD binding, which will inevitably lead to an increase in binding affinity and enhanced receptor-dependent infection.

SARS-CoV-2 delta variant shows conformational dynamics distinct from the other variants. As of November 2021 the global dominant SARS-CoV-2 variant is the B.1.617.2 (or delta)[46]. It contains two mutations in the RBD region: L452R and T478K. Like the *wt*, alpha and beta variants, MD simulations of the delta RBD show the prevalence of two sets of RBM conformations, one of which corresponds to the *wt* open conformation (Supplementary Video S4) and is stabilized by the same interactions observed for the three other variants (Figure 3D). However, unlike those variants, MD simulations of the delta RBD do not show the occurrence of a closed conformation at all. Instead, an alternative open conformation is present, which we refer to as reversed. PCA analysis of the delta variant trajectory, show two deep basins, 0 and 2 in Figure 2D, which correspond to the open and reversed conformations, respectively. As for the other variants, simulations were able to reversibly visit the two states (Supplementary Figure S1).

The reversed conformation showcases the incredible flexibility of the RBM region, which not only opens and closes over the ACE2 binding surface of the RBD but acts as a two-way hinge that reverse-folds to the side of the RBD. This alternative conformation might also prove significant advantages over the *wt* open state: RBD-targeting antibodies are known to bind via recognition of the RBM ridge region[17,74]; the reversed state putatively hides this region from antibody recognition, while still providing an open ACE2 binding surface for infection.

A hydrogen bond between the mutated R452 on strand β5 and Y449 appears to be one of the main driving forces folding the delta variant's ridge region backwards. This interaction destabilizes the β5 strand and enables the ridge to move up and interact with the core. Transient interactions between ridge residues G476, S477 as well as the mutated K478 with residues R346, F347 and N354 of strand β1 stabilize the contact between the ridge loop and the RBD core, keeping it locked in place (Figure 3H).

Regarding the secondary structure, much like the other variants, the delta open conformation is very similar to that of the *wt* (Supplementary Table S3). However, as expected,

the reversed conformation shows substantial differences. In this state, the two small beta strands formed by residues 473-474 and 488-489, present in the open conformation, are completely lost. Additionally, the beta-sheet formed by strands β5 and β6 becomes less prevalent, likely due to the L452R mutation (one of the β5 strand residues that destabilizes the β-sheet by establishing a new interaction with Y449). Curiously, like in the alpha and beta variants, there is also a significant alpha helical character between residues 490 and 475.

As for the closed conformations of the *wt*, alpha and beta variants, the delta reversed conformation also leads to a decrease in SASA (∼ 3%). Unlike the closed conformations, however, this alternative open conformation still presents a fully accessible ACE2 binding surface.

Impact of SARS-CoV-2 variants on ACE2 binding affinity. To find experimental basis for our results, we compiled ACE2-RBD binding kinetics data from recent studies[75–81] (Supplementary Table S3). These results were obtained by surface plasmon resonance (SPR) and biolayer interferometry (BLI) and encompass data regarding both the *wt* and studied variants. Additionally, we compiled results obtained for just the RBD as well as for the entire S protein. While the binding kinetics values recovered from these studies are not fully consistent with each other, likely due to differences in particular experimental setups, they are mostly in the same range, and appear to follow similar trends.

Regarding the equilibrium dissociation constant (K_d), all variants have an increased binding affinity when compared to the *wt*. With the currently available data, however, it is hard to distinguish between the efficiency of the several variants, with the alpha and beta variants showing a slightly better affinity than delta.

To get more information, we analyzed both the association (k_{on}) and dissociation rate constants (k_{off}). k_{off} reflects the lifetime of the protein-protein complex and as such, the strength of the interaction. We observe a consistent decrease in k_{off} for the variants in comparison to the *wt*. The alpha and beta variants stand out from delta in this regard, with substantially lower k_{off} values. These results hint at the variants interacting more strongly with ACE2 than the *wt*, with the alpha and beta complexes being substantially more stable than those of delta. Several other MD studies have studied the impact of these mutations on the contacts between RBD and ACE2 and have shown how the substantially altered ACE2-RBD interaction network of the alpha and beta variants might be outperforming that of the *wt* variant[82–86]. The delta variant does not contain mutations to the RBD ACE2 binding surface and, as such, the interactions established are not substantially different from those of *wt*. This is reflected in a closer, although still lower, k_{off} value.

The variants also substantially impact k_{on}. This rate constant reflects the efficiency with which protein–protein collisions lead to a bound state. While a couple of studies show no significant impact[75,79], most show that the variants lead to a substantial increase in k_{on} , reflecting an increase in ACE2 accessibility[76–78,80,81]. We propose that this can be explained by the significant changes in RBM conformational dynamics that we have here described, where mutations lead to a decrease in prevalence of the closed state, favoring binding. As such, our results point to an alternative mechanism for enhancing RBD-ACE2 binding, not by directly strengthening ACE2-RBD interactions, but rather by boosting, via modulation of ridge dynamics, the ACE2 binding competence.

Emergent VOCs share relevant RBM mutations with alpha beta and delta. Recently, a new VOC — B.1.1.529 or omicron — has emerged which is overtaking delta as the dominant variant in some world regions[87]. The omicron variant contains 15 mutations in the RBD region, 10 of which are concentrated in the RBM. Some of those mutations are also observed, or are similar to those, in the alpha, beta and delta variants: K417N, T478K, E484A and N501Y. From our work, we can expect this large number of mutations to heavily impact the open/closed equilibrium we observed for *wt* RBD. In particular, the presence of the T478K mutation — shared with the delta variant — points towards possible alternative conformations like delta's reversed state. Just as for delta, these conformations are likely to improve antibody escape, providing omicron with a substantial fitness advantage.

CONCLUSION

In this work we performed AA MD simulations of the SARS-CoV-2 RBD, as well as that of the alpha, beta and delta VOCs, to characterize the impact of the mutations on RBD conformational dynamics in solution.

Our results show that the *wt* RBD adopts two distinct conformations in equilibrium: an open conformation where the RBD is free to bind ACE2; and a closed conformation, where the RBM ridge blocks the ACE2 binding surface and likely hinders binding to ACE2. We characterized the two states and showed that they originate from specific intramolecular interactions between residues of the RBM ridge and those of the ACE2 binding surface. As far as we know, this is the first report of this "hinge-like" mechanism which can effectively shield the ACE2 binding surface from the solvent and binding partners. This mechanism is yet to be seen in experimentally solved RBD structures, which have thus far struggled to fully resolve the unbound RBM region[20,29,88]. The RBM is found unresolved in most structures due to the large flexibility of the region, and those that are fully resolved are often structures of RBD complexed with either ACE2[14,30–33], antibodies[34–40] or itself by dimerizing via the ACE2 binding surface[89,90].

The three variants tested in this work, significantly impacted the open/closed equilibrium we observed for *wt* RBD. Both alpha and beta variants shifted the equilibrium towards more open conformations by roughly 20%, while the delta variant did not show the presence of a closed conformation at all. This shift towards more open conformations likely enhances ACE2 binding affinity by increasing accessibility to the RBM and facilitating binding. Several experimental binding studies have shown that these variants lead to a substantial increase in ACE2-RBD binding association rate constant, reflecting an increased ACE2 accessibility, in agreement with our findings.

Additionally, the delta variant showed an alternative open conformation, distinct from that of the other variants.

This alternative conformation keeps the ACE2 binding surface open and accessible for binding, but significantly alters the conformation of the RBM ridge. This state presents a substantially altered ridge region, which bends backwards towards the RBD core, shielding some of it from exposure. We hypothesize that this may provide a fitness advantage by aiding in antibody escape: many RBD-targeting antibodies are known to target the RBM ridge region[35,74,91,92]. In the alternative open conformation, the ridge may be not as easily recognized, while the ACE2 binding surface remains unobstructed for infection.

These results show that the mutations found in the three VOCs (alpha, beta and delta) impact RBD conformational dynamics in a direction that promotes efficient binding to ACE2 and (in the case of the delta variant) antibody escape, an effect which has thus far been disregarded. In this context, our findings can also help explain some of the antibody-evading characteristics of the emergent omicron variant.

ASSOCIATED CONTENT

Supporting Information. Extended materials providing video description, additional analysis results on RMSD, SASA, RINs, additional detail on the PCA analysis and compilation of ACE2-RBD experimental binding results. Videos S1 through S4 show the RBD conformational dynamics of the *wt*, alpha, beta and delta variants.

AUTHOR INFORMATION

Corresponding Authors

* Cláudio M. Soares - Instituto de Tecnologia Química e Biológica António Xavier, Universidade Nova de Lisboa, Av. da República, 2780-157 Oeiras, Portugal; E-mail: claudio@itqb.unl.pt
* Diana Lousa - Instituto de Tecnologia Química e Biológica António Xavier, Universidade Nova de Lisboa, Av. da República, 2780-157 Oeiras, Portugal; E-mail: dlousa@itqb.unl.pt
* Manuel N. Melo - Instituto de Tecnologia Química e Biológica António Xavier, Universidade Nova de Lisboa, Av. da República, 2780-157 Oeiras, Portugal; E-mail: m.n.melo@itqb.unl.pt

Author Contributions

§ M.V. and L.B.A. contributed equally to this work. D.L. and C.M.S. designed the simulation setup. DL prepared the systems' topologies and C.M.S. performed the simulations. M.V., L.B.A., M.N.M., D.L. and C.M.S. designed the analysis and M.V. and L.B.A. performed them. All authors contributed to manuscript writing and revision and have given approval to the final version of the manuscript.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENT

The authors thank António M. Baptista and Sara R. R. Campos for their helpful discussion and input, and for making available their density and energy landscape analysis package, LandscapeTools. L.B.A. thanks the Medical Biochemistry and Biophysics Doctoral Program (M2B-PhD) and Fundação para a Ciência e a Tecnologia, I.P. (FCT) for PhD fellowship PD/BD/137492/2018. M.V. thanks FCT for the PhD fellowship SFRH/BD/148542/2019. M.N.M thanks FCT for fellowship CEECIND/04124/2017. D.L. acknowledges FCT project PTDC/CCI-BIO/28200/2017. C.M.S. and M.N.M further acknowledge FCT project MOSTMICRO-ITQB, with references UIDB/04612/2020 and UIDP/04612/2020.

REFERENCES

(1) Andersen, K. G.; Rambaut, A.; Lipkin, W. I.; Holmes, E. C.; Garry, R. F. The Proximal Origin of SARS-CoV-2. *Nature Medicine*. Nature Research April 1, 2020, pp 450–452.
(2) Wu, F.; Zhao, S.; Yu, B.; Chen, Y. M.; Wang, W.; Song, Z. G.; Hu, Y.; Tao, Z. W.; Tian, J. H.; Pei, Y. Y.; Yuan, M. L.; Zhang, Y. L.; Dai, F. H.; Liu, Y.; Wang, Q. M.; Zheng, J. J.; Xu, L.; Holmes, E. C.; Zhang, Y. Z. A New Coronavirus Associated with Human Respiratory Disease in China. *Nature* **2020**, *579* (7798), 265–269.
(3) Zhu, N.; Zhang, D.; Wang, W.; Li, X.; Yang, B.; Song, J.; Zhao, X.; Huang, B.; Shi, W.; Lu, R.; Niu, P.; Zhan, F.; Ma, X.; Wang, D.; Xu, W.; Wu, G.; Gao, G. F.; Tan, W. A Novel Coronavirus from Patients with Pneumonia in China, 2019. *N. Engl. J. Med.* **2020**, *382* (8), 727–733.
(4) Zhou, F.; Yu, T.; Du, R.; Fan, G.; Liu, Y.; Liu, Z.; Xiang, J.; Wang, Y.; Song, B.; Gu, X.; Guan, L.; Wei, Y.; Li, H.; Wu, X.; Xu, J.; Tu, S.; Zhang, Y.; Chen, H.; Cao, B. Clinical Course and Risk Factors for Mortality of Adult Inpatients with COVID-19 in Wuhan, China: A Retrospective Cohort Study. *Lancet* **2020**, *395* (10229), 1054–1062.
(5) World Health Organization. WHO Coronavirus (COVID-19) Dashboard https://covid19.who.int/ (accessed Nov 29, 2021).
(6) Jackson, C. B.; Farzan, M.; Chen, B.; Choe, H. Mechanisms of SARS-CoV-2 Entry into Cells. *Nat. Rev. Mol. Cell Biol.* **2021**, 1–18.
(7) Li, F. Structure, Function, and Evolution of Coronavirus Spike Proteins. *Annual Review of Virology*. Annual Reviews Inc. September 29, 2016, pp 237–261.
(8) Hoffmann, M.; Kleine-Weber, H.; Schroeder, S.; Krüger, N.; Herrler, T.; Erichsen, S.; Schiergens, T. S.; Herrler, G.; Wu, N. H.; Nitsche, A.; Müller, M. A.; Drosten, C.; Pöhlmann, S. SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. *Cell* **2020**, *181* (2), 271-280.e8.
(9) Wan, Y.; Shang, J.; Graham, R.; Baric, R. S.; Li, F. Receptor Recognition by the Novel Coronavirus from Wuhan: An Analysis Based on Decade-Long Structural Studies of SARS Coronavirus. *J. Virol.* **2020**, *94* (7).
(10) Bosch, B. J.; van der Zee, R.; de Haan, C. A. M.; Rottier, P. J. M. The Coronavirus Spike Protein Is a Class I Virus Fusion Protein: Structural and Functional Characterization of the Fusion Core Complex. *J. Virol.* **2003**, *77* (16), 8801–8811.
(11) Walls, A. C.; Tortorici, M. A.; Snijder, J.; Xiong, X.; Bosch, B. J.; Rey, F. A.; Veesler, D. Tectonic Conformational Changes of a Coronavirus Spike Glycoprotein Promote Membrane Fusion. *Proc. Natl. Acad. Sci. U. S. A.* **2017**, *114* (42), 11157–11162.
(12) de Vries, R. D.; Schmitz, K. S.; Bovier, F. T.; Predella, C.; Khao, J.; Noack, D.; Haagmans, B. L.; Herfst, S.; Stearns, K. N.; Drew-Bear, J.; Biswas, S.; Rockx, B.; McGill, G.; Dorrello, N. V.; Gellman, S. H.; Alabi, C. A.; de Swart, R. L.; Moscona, A.; Porotto, M. Intranasal Fusion Inhibitory Lipopeptide Prevents Direct-Contact SARS-CoV-2 Transmission in Ferrets. *Science (80-.).* **2021**, *371* (6536), eabf4896.
(13) Huang, Y.; Yang, C.; Xu, X. feng; Xu, W.; Liu, S. wen. Structural and Functional Properties of SARS-CoV-2 Spike Protein: Potential Antivirus Drug Development for COVID-19. *Acta Pharmacologica Sinica*. Springer Nature September 1, 2020, pp 1141–1149.
(14) Wang, Q.; Zhang, Y.; Wu, L.; Niu, S.; Song, C.; Zhang, Z.; Lu, G.; Qiao, C.; Hu, Y.; Yuen, K. Y.; Wang, Q.; Zhou, H.; Yan, J.; Qi, J. Structural and Functional Basis of SARS-CoV-2 Entry by Using Human ACE2. *Cell* **2020**, *181* (4), 894-904.e9.

(15) Hussain, M.; Jabeen, N.; Raza, F.; Shabbir, S.; Baig, A. A.; Amanullah, A.; Aziz, B. Structural Variations in Human ACE2 May Influence Its Binding with SARS-CoV-2 Spike Protein. *J. Med. Virol.* **2020**, *92* (9), 1580–1586.

(16) Ali, F.; Elserafy, M.; Alkordi, M. H.; Amin, M. ACE2 Coding Variants in Different Populations and Their Potential Impact on SARS-CoV-2 Binding Affinity. *Biochem. Biophys. Reports* **2020**, *24*, 100798.

(17) Alenquer, M.; Ferreira, F.; Lousa, D.; Valério, M.; Medina-Lopes, M.; Bergman, M.-L.; Gonçalves, J.; Demengeot, J.; Leite, R. B.; Lilue, J.; Ning, Z.; Penha-Gonçalves, C.; Soares, H.; Soares, C. M.; Amorim, M. J. Signatures in SARS-CoV-2 Spike Protein Conferring Escape to Neutralizing Antibodies. *PLOS Pathog.* **2021**, *17* (8), e1009772.

(18) Lupala, C. S.; Li, X.; Lei, J.; Chen, H.; Qi, J.; Liu, H.; Su, X.-D. Computational Simulations Reveal the Binding Dynamics between Human ACE2 and the Receptor Binding Domain of SARS-CoV-2 Spike Protein. *Quant. Biol.* **2021**, *9* (1), 61–72.

(19) Yan, F.-F.; Gao, F. Comparison of the Binding Characteristics of SARS-CoV and SARS-CoV-2 RBDs to ACE2 at Different Temperatures by MD Simulations. *Brief. Bioinform.* **2021**, *22* (2), 1122–1136.

(20) Xu, C.; Wang, Y.; Liu, C.; Zhang, C.; Han, W.; Hong, X.; Wang, Y.; Hong, Q.; Wang, S.; Zhao, Q.; Wang, Y.; Yang, Y.; Chen, K.; Zheng, W.; Kong, L.; Wang, F.; Zuo, Q.; Huang, Z.; Cong, Y. Conformational Dynamics of SARS-CoV-2 Trimeric Spike Glycoprotein in Complex with Receptor ACE2 Revealed by Cryo-EM. *Sci. Adv.* **2021**, *7* (1).

(21) Cao, L.; Goreshnik, I.; Coventry, B.; Case, J. B.; Miller, L.; Kozodoy, L.; Chen, R. E.; Carter, L.; Walls, A. C.; Park, Y. J.; Strauch, E. M.; Stewart, L.; Diamond, M. S.; Veesler, D.; Baker, D. De Novo Design of Picomolar SARS-CoV-2 Miniprotein Inhibitors. *Science (80-.).* **2020**, *370* (6515).

(22) Alexpandi, R.; De Mesquita, J. F.; Pandian, S. K.; Ravi, A. V. Quinolines-Based SARS-CoV-2 3CLpro and RdRp Inhibitors and Spike-RBD-ACE2 Inhibitor for Drug-Repurposing Against COVID-19: An in Silico Analysis. *Front. Microbiol.* **2020**, *0*, 1796.

(23) Awad, I. E.; Abu-Saleh, A. A.-A. A.; Sharma, S.; Yadav, A.; Poirier, R. A. High-Throughput Virtual Screening of Drug Databanks for Potential Inhibitors of SARS-CoV-2 Spike Glycoprotein. *https://doi.org/10.1080/07391102.2020.1835721* **2020**.

(24) Padhi, A. K.; Seal, A.; Khan, J. M.; Ahamed, M.; Tripathi, T. Unraveling the Mechanism of Arbidol Binding and Inhibition of SARS-CoV-2: Insights from Atomistic Simulations. *Eur. J. Pharmacol.* **2021**, *894*, 173836.

(25) Kumar, V.; Liu, H.; Wu, C. Drug Repurposing against SARS-CoV-2 Receptor Binding Domain Using Ensemble-Based Virtual Screening and Molecular Dynamics Simulations. *Comput. Biol. Med.* **2021**, *135*, 104634.

(26) Patel, C. N.; Goswami, D.; Jaiswal, D. G.; Parmar, R. M.; Solanki, H. A.; Pandya, H. A. Pinpointing the Potential Hits for Hindering Interaction of SARS-CoV-2 S-Protein with ACE2 from the Pool of Antiviral Phytochemicals Utilizing Molecular Docking and Molecular Dynamics (MD) Simulations. *J. Mol. Graph. Model.* **2021**, *105*, 107874.

(27) Muhseen, Z. T.; Hameed, A. R.; Al-Hasani, H. M. H.; Tahir ul Qamar, M.; Li, G. Promising Terpenes as SARS-CoV-2 Spike Receptor-Binding Domain (RBD) Attachment Inhibitors to the Human ACE2 Receptor: Integrated Computational Approach. *J. Mol. Liq.* **2020**, *320*, 114493.

(28) Lan, J.; Ge, J.; Yu, J.; Shan, S.; Zhou, H.; Fan, S.; Zhang, Q.; Shi, X.; Wang, Q.; Zhang, L.; Wang, X. Structure of the SARS-CoV-2 Spike Receptor-Binding Domain Bound to the ACE2 Receptor. *Nature* **2020**, *581* (7807), 215–220.

(29) Wrapp, D.; Wang, N.; Corbett, K. S.; Goldsmith, J. A.; Hsieh, C.-L.; Abiona, O.; Graham, B. S.; McLellan, J. S. Cryo-EM Structure of the 2019-NCoV Spike in the Prefusion Conformation. *Science (80-.).* **2020**, *367* (6483), 1260–1263.

(30) Lan, J.; Ge, J.; Yu, J.; Shan, S.; Zhou, H.; Fan, S.; Zhang, Q.; Shi, X.; Wang, Q.; Zhang, L.; Wang, X. Structure of the SARS-CoV-2 Spike Receptor-Binding Domain Bound to the ACE2 Receptor. *Nat. 2020 5817807* **2020**, *581* (7807), 215–220.

(31) Li, F. Structure of SARS Coronavirus Spike Receptor-Binding Domain Complexed with Receptor. *Science (80-.).* **2005**, *309* (5742), 1864–1868.

(32) Shang, J.; Ye, G.; Shi, K.; Wan, Y.; Luo, C.; Aihara, H.; Geng, Q.; Auerbach, A.; Li, F. Structural Basis of Receptor Recognition by SARS-CoV-2. *Nat. 2020 5817807* **2020**, *581* (7807), 221–224.

(33) Zhou, T.; Tsybovsky, Y.; Gorman, J.; Rapp, M.; Cerutti, G.; Chuang, G.-Y.; Katsamba, P. S.; Sampson, J. M.; Schön, A.; Bimela, J.; Boyington, J. C.; Nazzari, A.; Olia, A. S.; Shi, W.; Sastry, M.; Stephens, T.; Stuckey, J.; Teng, I.-T.; Wang, P.; Wang, S.; Zhang, B.; Friesner, R. A.; Ho, D. D.; Mascola, J. R.; Shapiro, L.; Kwong, P. D. Cryo-EM Structures of SARS-CoV-2 Spike without and with ACE2 Reveal a PH-Dependent Switch to Mediate Endosomal Positioning of Receptor-Binding Domains. *Cell Host Microbe* **2020**, *28* (6), 867-879.e5.

(34) Rapp, M.; Guo, Y.; Reddem, E. R.; Yu, J.; Liu, L.; Wang, P.; Cerutti, G.; Katsamba, P.; Bimela, J. S.; Bahna, F. A.; Mannepalli, S. M.; Zhang, B.; Kwong, P. D.; Huang, Y.; Ho, D. D.; Shapiro, L.; Sheng, Z. Modular Basis for Potent SARS-CoV-2 Neutralization by a Prevalent VH1-2-Derived Antibody Class. *Cell Rep.* **2021**, *35* (1).

(35) Tortorici, M. A.; Beltramello, M.; Lempp, F. A.; Pinto, D.; Dang, H. V.; Rosen, L. E.; McCallum, M.; Bowen, J.; Minola, A.; Jaconi, S.; Zatta, F.; De Marco, A.; Guarino, B.; Bianchi, S.; Lauron, E. J.; Tucker, H.; Zhou, J.; Peter, A.; Havenar-Daughton, C.; Wojcechowskyj, J. A.; Case, J. B.; Chen, R. E.; Kaiser, H.; Montiel-Ruiz, M.; Meury, M.; Czudnochowski, N.; Spreafico, R.; Dillen, J.; Ng, C.; Sprugasci, N.; Culap, K.; Benigni, F.; Abdelnabi, R.; Foo, S. Y. C.; Schmid, M. A.; Cameroni, E.; Riva, A.; Gabrieli, A.; Galli, M.; Pizzuto, M. S.; Neyts, J.; Diamond, M. S.; Virgin, H. W.; Snell, G.; Corti, D.; Fink, K.; Veesler, D. Ultrapotent Human Antibodies Protect against SARS-CoV-2 Challenge via Multiple Mechanisms. *Science (80-.).* **2020**, *370* (6519), 950–957.

(36) Wu, Y.; Wang, F.; Shen, C.; Peng, W.; Li, D.; Zhao, C.; Li, Z.; Li, S.; Bi, Y.; Yang, Y.; Gong, Y.; Xiao, H.; Fan, Z.; Tan, S.; Wu, G.; Tan, W.; Lu, X.; Fan, C.; Wang, Q.; Liu, Y.; Zhang, C.; Qi, J.; Gao, G. F.; Gao, F.; Liu, L. A Noncompeting Pair of Human Neutralizing Antibodies Block COVID-19 Virus Binding to Its Receptor ACE2. *Science (80-.).* **2020**, *368* (6496), 1274–1278.

(37) Bertoglio, F.; Fühner, V.; Ruschig, M.; Heine, P. A.; Abassi, L.; Klünemann, T.; Rand, U.; Meier, D.; Langreder, N.; Steinke, S.; Ballmann, R.; Schneider, K.-T.; Roth, K. D. R.; Kuhn, P.; Riese, P.; Schäckermann, D.; Korn, J.; Koch, A.; Chaudhry, M. Z.; Eschke, K.; Kim, Y.; Zock-Emmenthal, S.; Becker, M.; Scholz, M.; Moreira, G. M. S. G.; Wenzel, E. V.; Russo, G.; Garritsen, H. S. P.; Casu, S.; Gerstner, A.; Roth, G.; Adler, J.; Trimpert, J.; Hermann, A.; Schirrmann, T.; Dübel, S.; Frenzel, A.; Heuvel, J. Van den; Čičin-Šain, L.; Schubert, M.; Hust, M. A SARS-CoV-2 Neutralizing Antibody Selected from COVID-19 Patients Binds to the ACE2-RBD Interface and Is Tolerant to Most Known RBD Mutations. *Cell Rep.* **2021**, *36* (4).

(38) Kreye, J.; Reincke, S. M.; Kornau, H.-C.; Sánchez-Sendin, E.; Corman, V. M.; Liu, H.; Yuan, M.; Wu, N. C.; Zhu, X.; Lee, C.-C. D.; Trimpert, J.; Höltje, M.; Dietert, K.; Stöffler, L.; Wardenburg, N. von; Hoof, S. van; Homeyer, M. A.; Hoffmann, J.; Abdelgawad, A.; Gruber, A. D.; Bertzbach, L. D.; Vladimirova, D.; Li, L. Y.; Barthel, P. C.; Skriner, K.; Hocke, A. C.; Hippenstiel, S.; Witzenrath, M.; Suttorp, N.; Kurth, F.; Franke, C.; Endres, M.; Schmitz, D.; Jeworowski, L. M.; Richter, A.; Schmidt, M. L.; Schwarz, T.; Müller, M. A.; Drosten, C.; Wendisch, D.; Sander, L. E.; Osterrieder, N.; Wilson, I. A.; Prüss, H. A Therapeutic Non-Self-Reactive SARS-CoV-2 Antibody Protects from Lung Pathology in a COVID-19 Hamster Model. *Cell* **2020**, *183* (4), 1058-1069.e19.

(39) Hansen, J.; Baum, A.; Pascal, K. E.; Russo, V.; Giordano, S.; Wloga, E.; Fulton, B. O.; Yan, Y.; Koon, K.; Patel, K.; Chung, K. M.; Hermann, A.; Ullman, E.; Cruz, J.; Rafique, A.; Huang, T.; Fairhurst, J.; Libertiny, C.; Malbec, M.; Lee, W. Y.; Welsh, R.; Farr, G.; Pennington, S.; Deshpande, D.; Cheng, J.; Watty, A.; Bouffard, P.; Babb, R.; Levenkova, N.; Chen, C.; Zhang, B.; Hernandez, A. R.; Saotome, K.; Zhou, Y.; Franklin, M.; Sivapalasingam, S.; Lye, D. C.; Weston, S.; Logue, J.; Haupt, R.;

Frieman, M.; Chen, G.; Olson, W.; Murphy, A. J.; Stahl, N.; Yancopoulos, G. D.; Kyratsous, C. A. Studies in Humanized Mice and Convalescent Humans Yield a SARS-CoV-2 Antibody Cocktail. *Science (80-.).* **2020**, *369* (6506), 1010–1014.

(40) Yuan, M.; Liu, H.; Wu, N. C.; Lee, C. C. D.; Zhu, X.; Zhao, F.; Huang, D.; Yu, W.; Hua, Y.; Tien, H.; Rogers, T. F.; Landais, E.; Sok, D.; Jardine, J. G.; Burton, D. R.; Wilson, I. A. Structural Basis of a Shared Antibody Response to SARS-CoV-2. *Science (80-.).* **2020**, *369* (6507), 1119–1123.

(41) Baral, P.; Bhattarai, N.; Hossen, M. L.; Stebliankin, V.; Gerstman, B. S.; Narasimhan, G.; Chapagain, P. P. Mutation-Induced Changes in the Receptor-Binding Interface of the SARS-CoV-2 Delta Variant B.1.617.2 and Implications for Immune Evasion. *Biochem. Biophys. Res. Commun.* **2021**, *574*, 14–19.

(42) Bhattarai, N.; Baral, P.; Gerstman, B. S.; Chapagain, P. P. Structural and Dynamical Differences in the Spike Protein RBD in the SARS-CoV-2 Variants B.1.1.7 and B.1.351. *J. Phys. Chem. B* **2021**, *125* (26), 7101–7107.

(43) Williams, J. K.; Wang, B.; Sam, A.; Hoop, C. L.; Case, D. A.; Baum, J. Molecular Dynamics Analysis of a Flexible Loop at the Binding Interface of the SARS-CoV-2 Spike Protein Receptor-Binding Domain. *Proteins Struct. Funct. Bioinforma.* **2021**.

(44) Nelson, G.; Buzko, O.; Bassett, A.; Spilman, P.; Niazi, K.; Rabizadeh, S.; Soon-Shiong, P. Millisecond-Scale Molecular Dynamics Simulation of Spike RBD Structure Reveals Evolutionary Adaption of SARS-CoV-2 to Stably Bind ACE2. *bioRxiv* **2020**, 2020.12.11.422055.

(45) Tegally, H.; Wilkinson, E.; Giovanetti, M.; Iranzadeh, A.; Fonseca, V.; Giandhari, J.; Doolabh, D.; Pillay, S.; San, E. J.; Msomi, N.; Mlisana, K.; Gottberg, A. von; Walaza, S.; Allam, M.; Ismail, A.; Mohale, T.; Glass, A. J.; Engelbrecht, S.; Zyl, G. Van; Preiser, W.; Petruccione, F.; Sigal, A.; Hardie, D.; Marais, G.; Hsiao, M.; Korsman, S.; Davies, M.-A.; Tyers, L.; Mudau, I.; York, D.; Maslo, C.; Goedhals, D.; Abrahams, S.; Laguda-Akingba, O.; Alisoltani-Dehkordi, A.; Godzik, A.; Wibmer, C. K.; Sewell, B. T.; Lourenço, J.; Alcantara, L. C. J.; Pond, S. L. K.; Weaver, S.; Martin, D.; Lessells, R. J.; Bhiman, J. N.; Williamson, C.; Oliveira, T. de. Emergence and Rapid Spread of a New Severe Acute Respiratory Syndrome-Related Coronavirus 2 (SARS-CoV-2) Lineage with Multiple Spike Mutations in South Africa. *medRxiv* **2020**, *10*, 2020.12.21.20248640.

(46) Cherian, S.; Potdar, V.; Jadhav, S.; Yadav, P.; Gupta, N.; Das, M.; Rakshit, P.; Singh, S.; Abraham, P.; Panda, S.; Team, N. SARS-CoV-2 Spike Mutations, L452R, T478K, E484Q and P681R, in the Second Wave of COVID-19 in Maharashtra, India. *Microorganisms* **2021**, *9* (7), 1542.

(47) Science Brief: Emerging SARS-CoV-2 Variants | CDC https://www.cdc.gov/coronavirus/2019-ncov/science/science-briefs/scientific-brief-emerging-variants.html (accessed Sep 30, 2021).

(48) Abdool Karim, S. S.; de Oliveira, T. New SARS-CoV-2 Variants — Clinical, Public Health, and Vaccine Implications. *N. Engl. J. Med.* **2021**, *384* (19), 1866–1868.

(49) Shah, M.; Ahmad, B.; Choi, S.; Woo, H. G. Mutations in the SARS-CoV-2 Spike RBD Are Responsible for Stronger ACE2 Binding and Poor Anti-SARS-CoV MAbs Cross-Neutralization. *Comput. Struct. Biotechnol. J.* **2020**, *18*, 3402–3414.

(50) 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.

(51) Lindahl; Abraham; Hess; Spoel, van der. GROMACS 2020.3 Source Code. **2020**.

(52) Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. Ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from Ff99SB. *J. Chem. Theory Comput.* **2015**, *11* (8), 3696–3713.

(53) Mark, P.; Nilsson, L. Structure and Dynamics of the TIP3P, SPC, and SPC/E Water Models at 298 K. *J. Phys. Chem. A* **2001**, *105* (43), 9954–9960.

(54) Schrödinger, LLC. *The {PyMOL} Molecular Graphics System, Version~1.8*; 2015.

(55) Berendsen, H. J. C.; Postma, J. P. M.; Van Gunsteren, W. F.; Dinola, A.; Haak, J. R. Molecular Dynamics with Coupling to an External Bath. *J. Chem. Phys.* **1984**, *81* (8), 3684–3690.

(56) Bussi, G.; Donadio, D.; Parrinello, M. Canonical Sampling through Velocity Rescaling. *J. Chem. Phys.* **2007**, *126* (1), 014101.

(57) Parrinello, M.; Rahman, A. Polymorphic Transitions in Single Crystals: A New Molecular Dynamics Method. *J. Appl. Phys.* **1981**, *52* (12), 7182–7190.

(58) 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.

(59) 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.

(60) Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M. LINCS: A Linear Constraint Solver for Molecular Simulations. *J. Comput. Chem.* **1997**, *18* (12), 1463–1472.

(61) Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual Molecular Dynamics. *J. Mol. Graph.* **1996**, *14* (1), 33–38.

(62) Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. UCSF Chimera?A Visualization System for Exploratory Research and Analysis. *J. Comput. Chem.* **2004**, *25* (13), 1605–1612.

(63) Jolliffe, I. T. *Principal Component Analysis*; Springer Series in Statistics; Springer-Verlag: New York, 2002.

(64) Jolliffe, I. T.; Cadima, J. Principal Component Analysis: A Review and Recent Developments. *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.* **2016**, *374* (2065).

(65) Campos, S. R. R.; Baptista, A. M. Conformational Analysis in a Multidimensional Energy Landscape: Study of an Arginylglutamate Repeat. *J. Phys. Chem. B* **2009**, *113* (49), 15989–16001.

(66) Rencher, A. C. Methods of Multivariate Analysis. **2002**, 708.

(67) Michaud-Agrawal, N.; Denning, E. J.; Woolf, T. B.; Beckstein, O. MDAnalysis: A Toolkit for the Analysis of Molecular Dynamics Simulations. *J. Comput. Chem.* **2011**, *32* (10), 2319–2327.

(68) Silverman, B. W. *Density Estimation for Statistics and Data Analysis*; Routledge, 2018.

(69) Campos, S. R. R.; Baptista, A. M. Molecular Simulation Lab — In-House Software https://www.itqb.unl.pt/labs/molecular-simulation/in-house-software (accessed Nov 29, 2021).

(70) Becker, O. M.; Karplus, M. The Topology of Multidimensional Potential Energy Surfaces: Theory and Application to Peptide Structure and Kinetics. *J. Chem. Phys.* **1998**, *106* (4), 1495.

(71) Stillinger, F. H.; Weber, T. A. Packing Structures and Transitions in Liquids and Solids. *Science (80-.).* **1984**, *225* (4666), 983–989.

(72) Contreras-Riquelme, S.; Garate, J.-A.; Perez-Acle, T.; Martin, A. J. M. RIP-MD: A Tool to Study Residue Interaction Networks in Protein Molecular Dynamics. *PeerJ* **2018**, *6* (12), e5998.

(73) Shannon, P.; Markiel, A.; Ozier, O.; Baliga, N. S.; Wang, J. T.; Ramage, D.; Amin, N.; Schwikowski, B.; Ideker, T. Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Res.* **2003**, *13* (11), 2498.

(74) Wu, N. C.; Yuan, M.; Liu, H.; Lee, C.-C. D.; Zhu, X.; Bangaru, S.; Torres, J. L.; Caniels, T. G.; Brouwer, P. J. M.; Gils, M. J. van; Sanders, R. W.; Ward, A. B.; Wilson, I. A. An Alternative Binding Mode of IGHV3-53 Antibodies to the SARS-CoV-2 Receptor Binding Domain. *Cell Rep.* **2020**, *33* (3).

(75) McCallum, M.; Walls, A. C.; Sprouse, K. R.; Bowen, J. E.; Rosen, L.; Dang, H. V.; deMarco, A.; Franko, N.; Tilles, S. W.; Logue, J.; Miranda, M. C.; Ahlrichs, M.; Carter, L.; Snell, G.; Pizzuto, M. S.; Chu, H. Y.; Voorhis, W. C. Van; Corti, D.; Veesler, D. Molecular Basis of Immune Evasion by the Delta and Kappa SARS-CoV-2 Variants. *bioRxiv* **2021**, 2021.08.11.455956.

(76) Tian, F.; Tong, B.; Sun, L.; Shi, S.; Zheng, B.; Wang, Z.; Dong, X.; Zheng, P. N501Y Mutation of Spike Protein in SARS-CoV-2 Strengthens Its Binding to Receptor ACE2. *Elife* **2021**, *10*.

(77) Laffeber, C.; de Koning, K.; Kanaar, R.; Lebbink, J. H. G. Experimental Evidence for Enhanced Receptor Binding by Rapidly Spreading SARS-CoV-2 Variants. *J. Mol. Biol.* **2021**, *433* (15), 167058.

(78) Supasa, P.; Zhou, D.; Dejnirattisai, W.; Liu, C.; Mentzer, A. J.;

Ginn, H. M.; Zhao, Y.; Duyvesteyn, H. M. E.; Nutalai, R.; Tuekprakhon, A.; Wang, B.; Paesen, G. C.; Slon-Campos, J.; López-Camacho, C.; Hallis, B.; Coombes, N.; Bewley, K. R.; Charlton, S.; Walter, T. S.; Barnes, E.; Dunachie, S. J.; Skelly, D.; Lumley, S. F.; Baker, N.; Shaik, I.; Humphries, H. E.; Godwin, K.; Gent, N.; Sienkiewicz, A.; Dold, C.; Levin, R.; Dong, T.; Pollard, A. J.; Knight, J. C.; Klenerman, P.; Crook, D.; Lambe, T.; Clutterbuck, E.; Bibi, S.; Flaxman, A.; Bittaye, M.; Belij-Rammerstorfer, S.; Gilbert, S.; Hall, D. R.; Williams, M. A.; Paterson, N. G.; James, W.; Carroll, M. W.; Fry, E. E.; Mongkolsapaya, J.; Ren, J.; Stuart, D. I.; Screaton, G. R. Reduced Neutralization of SARS-CoV-2 B.1.1.7 Variant by Convalescent and Vaccine Sera. *Cell* **2021**, *184* (8), 2201-2211.e7.

(79) Wirnsberger, G.; Monteil, V.; Eaton, B.; Postnikova, E.; Murphy, M.; Braunsfeld, B.; Crozier, I.; Kricek, F.; Niederhöfer, J.; Schwarzböck, A.; Breid, H.; Jimenez, A. S.; Bugajska-Schretter, A.; Dohnal, A.; Ruf, C.; Gugenberger, R.; Hagelkruys, A.; Montserrat, N.; Holbrook, M. R.; Oostenbrink, C.; Shoemaker, R. H.; Mirazimi, A.; Penninger, J. M. Clinical Grade ACE2 as a Universal Agent to Block SARS-CoV-2 Variants. *bioRxiv* **2021**, 2021.09.10.459744.

(80) Souza, A. S. de; Amorim, V. M. de F.; Guardia, G. D. A.; Santos, F. R. C. dos; Santos, F. F. dos; Souza, R. F. de; Juvenal, G. de A.; Huang, Y.; Ge, P.; Jiang, Y.; Paudel, P.; Ulrich, H.; Galante, P. A. F.; Guzzo, C. R. Molecular Dynamics Analysis of Fast-Spreading Severe Acute Respiratory Syndrome Coronavirus 2 Variants and Their Effects in the Interaction with Human Angiotensin-Converting Enzyme 2. *bioRxiv* **2021**, 2021.06.14.448436.

(81) Saville, J. W.; Mannar, D.; Zhu, X.; Srivastava, S. S.; Berezuk, A. M.; Demers, J.-P.; Zhou, S.; Tuttle, K. S.; Sekirov, I.; Kim, A.; Li, W.; Dimitrov, D. S.; Subramaniam, S. Structural and Biochemical Rationale for Enhanced Spike Protein Fitness in Delta and Kappa SARS-CoV-2 Variants. *bioRxiv* **2021**, 2021.09.02.458774.

(82) Socher, E.; Conrad, M.; Heger, L.; Paulsen, F.; Sticht, H.; Zunke, F.; Arnold, P. Computational Decomposition Reveals Reshaping of the SARS-CoV-2–ACE2 Interface among Viral Variants Expressing the N501Y Mutation. *J. Cell. Biochem.* **2021**.

(83) Ali, F.; Kasry, A.; Amin, M. The New SARS-CoV-2 Strain Shows a Stronger Binding Affinity to ACE2 Due to N501Y Mutant. *Med. Drug Discov.* **2021**, *10*, 100086.

(84) Luan, B.; Wang, H.; Huynh, T. Enhanced Binding of the N501Y-Mutated SARS-CoV-2 Spike Protein to the Human ACE2 Receptor: Insights from Molecular Dynamics Simulations. *FEBS Lett.* **2021**, *595* (10), 1454–1461.

(85) Ahmed, W.; Philip, A. M.; Biswas, K. H. Stable Interaction Of The UK B.1.1.7 Lineage SARS-CoV-2 S1 Spike N501Y Mutant With ACE2 Revealed By Molecular Dynamics Simulation. *bioRxiv* **2021**, 2021.01.07.425307.

(86) Nelson, G.; Buzko, O.; Spilman, P.; Niazi, K.; Rabizadeh, S.; Soon-Shiong, P. Molecular Dynamic Simulation Reveals E484K Mutation Enhances Spike RBD-ACE2 Affinity and the Combination of E484K, K417N and N501Y Mutations (501Y.V2 Variant) Induces Conformational Change Greater than N501Y Mutant Alone, Potentially Resulting in an Escape Mutant. *bioRxiv* **2021**, 2021.01.13.426558.

(87) CoVariants https://covariants.org/variants/21K.Omicron (accessed Nov 29, 2021).

(88) Walls, A. C.; Park, Y. J.; Tortorici, M. A.; Wall, A.; McGuire, A. T.; Veesler, D. Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell* **2020**, *181* (2), 281-292.e6.

(89) Norman, A.; Franck, C.; Christie, M.; Hawkins, P. M. E.; Patel, K.; Ashhurst, A. S.; Aggarwal, A.; Low, J. K. K.; Siddiquee, R.; Ashley, C. L.; Stein, M.; Triccas, J. A.; Turville, S.; Mackay, J. P.; Passioura, T.; Payne, R. J. Discovery of Cyclic Peptide Ligands to the SARS-CoV-2 Spike Protein Using MRNA Display. *ACS Cent. Sci.* **2021**, *7* (6), 1001–1008.

(90) Jiang, W.; Wang, J.; Jiao, S.; Gu, C.; Xu, W.; Chen, B.; Wang, R.; Chen, H.; Xie, Y.; Wang, A.; Li, G.; Zeng, D.; Zhang, J.; Zhang, M.; Wang, S.; Wang, M.; Gui, X. Characterization of MW06, a Human Monoclonal Antibody with Cross-Neutralization Activity against Both SARS-CoV-2 and SARS-CoV. *MAbs* **2021**, *13* (1), 1953683.

(91) Yao, H.; Sun, Y.; Deng, Y.-Q.; Wang, N.; Tan, Y.; Zhang, N.-N.; Li, X.-F.; Kong, C.; Xu, Y.-P.; Chen, Q.; Cao, T.-S.; Zhao, H.; Yan, X.; Cao, L.; Lv, Z.; Zhu, D.; Feng, R.; Wu, N.; Zhang, W.; Hu, Y.; Chen, K.; Zhang, R.-R.; Lv, Q.; Sun, S.; Zhou, Y.; Yan, R.; Yang, G.; Sun, X.; Liu, C.; Lu, X.; Cheng, L.; Qiu, H.; Huang, X.-Y.; Weng, T.; Shi, D.; Jiang, W.; Shao, J.; Wang, L.; Zhang, J.; Jiang, T.; Lang, G.; Qin, C.-F.; Li, L.; Wang, X. Rational Development of a Human Antibody Cocktail That Deploys Multiple Functions to Confer Pan-SARS-CoVs Protection. *Cell Res. 2020 311* **2020**, *31* (1), 25–36.

(92) Fu, D.; Zhang, G.; Wang, Y.; Zhang, Z.; Hu, H.; Shen, S.; Wu, J.; Li, B.; Li, X.; Fang, Y.; Liu, J.; Wang, Q.; Zhou, Y.; Wang, W.; Li, Y.; Lu, Z.; Wang, X.; Nie, C.; Tian, Y.; Chen, D.; Wang, Y.; Zhou, X.; Wang, Q.; Yu, F.; Zhang, C.; Deng, C.; Zhou, L.; Guan, G.; Shao, N.; Lou, Z.; Deng, F.; Zhang, H.; Chen, X.; Wang, M.; Liu, L.; Rao, Z.; Guo, Y. Structural Basis for SARS-CoV-2 Neutralizing Antibodies with Novel Binding Epitopes. *PLOS Biol.* **2021**, *19* (5), e3001209.

6 | Conclusions and future perspectives

From the early descriptions of Gorter and Grendel[98], to the more current updated version of the Singer-Nicolson fluid-mosaic model[99], our view on the structure of biomembranes has been constantly evolving. Recently, a lot of focus has been given to the lateral organization of membrane components, and in particular, to the formation of spatiotemporally regulated membrane domains which can potentially segregate and modulate many cellular processes.

As described in Chapter 1, PI(4,5)P_2 is a prime example of this complex lateral organization. It is a ubiquitous signaling entity, present in multiple localized pools in the plasma membrane where it recruits and modulates several regulatory proteins. Due to its importance, PI(4,5)P_2 lateral organization has been the focus of extensive work. While localized synthesis and degradation of PI(4,5)P_2 is not sufficient to explain the submicroscopic organization observed, PI(4,5)P_2 interactions with other cellular entities have helped elucidate the observed lateral organization of this lipid. The interaction of PI(4,5)P_2 with divalent cation stands out as a crucial element of its lateral organization. Divalent cations, and in particular Ca^{2+}, have been shown to not only impact PI(4,5)P_2 electrostatics and headgroup conformation, but also promote the formation of PI(4,5)P_2-rich nanodomains by crosslinking of the headgroups, even at physiological concentrations of cation and lipid[67]. Many of the effects associated with these nanodomains are yet to be fully understood, especially the extent to which they can impact PI(4,5)P_2-dependent function and signalling. In this work, our main goal was to expand on some of these unknowns. Specifically, how cation-induced effects can potentially impact PI(4,5)P_2-protein interactions, and how they can also amplify the effect of other lateral organization modulators such as acyl-chain composition.

Using two model proteins, our results clearly show that Ca^{2+}, within its physiological range of intracellular concentrations, is able to modulate PI(4,5)P_2-protein interactions via a complex set of competing mechanisms. Not only can divalent cations impact the affinity of proteins towards PI(4,5)P_2, but they can also compartmentalize PI(4,5)P_2 binding proteins through the formation of cation-induced nanodomains. By compartmentalizing proteins, PI(4,5)P_2 nanodomains were found to substantially impact protein-protein interactions, with the mechanism likely being coupled to PI(4,5)P_2 nanodomain size. Our results clearly note the importance of accounting for divalent cations when studying PI(4,5)P_2 - protein

interactions. Surprisingly, the impact of divalent cations on PI(4,5)P_2 - protein interactions is very rarely considered, constituting a serious oversight of this research field.

We also investigated representative PI(4,5)P_2 acyl-chain compositions of the spectrum observed *in vivo*, and reported the effect of acyl-chain saturation on the Ca^{2+}-dependent changes of the biophysical properties of PI(4,5)P_2-containing membranes. While in the absence of divalent cations all acyl-chain compositions behave in the same manner, for cation-induced PI(4,5)P_2 nanodomains, it is clear that the acyl-chain compositions yield structures with significantly different biophysical properties. The key conclusion is that increasing saturation yields more ordered and structured nanodomains, which can, in the case of the fully saturated composition, culminate in the formation of gel-like nanodomains. These results provide an explanation of why evolution favored the 18:0 20:4 PI(4,5)P_2 acyl-chain composition in mammalian cells. The presence of the polyunsaturated acyl-chain is expected to guarantee not only low bending rigidity, but also that the ordering within PI(4,5)P_2 nanodomains remains low. More so, this work once again showcases how divalent cations can dramatically impact PI(4,5)P_2 biophysical properties, specifically membrane order, permeability, PI(4,5)P_2 partition, and curvature association.

Overall, the results presented in the first two chapters elucidate mechanisms by which divalent cations can modulate, not only, PI(4,5)P_2 compartmentalization and lateral organization but also PI(4,5)P_2 function and interaction with cellular binding partners. The results also showcase the potential of combining fluorescence microscopy and spectroscopy approaches with computational molecular dynamics techniques, such as Martini CG MD simulations, to study both membrane biophysics and lipid-protein interactions. Fluorescence spectroscopy and microscopy techniques, especially when coupled with carefully selected fluorescent probes, allow for the accurate measurement of physical quantities. These include membrane biophysical properties such as membrane order, fusion, hydration, permeability and lipid diffusion, but also protein-protein and protein-lipid interactions such as protein-lipid binding and affinity or protein-protein oligomerization parameters. These features are well complemented by information obtained with MD simulations such as Martini CG MD, which can provide molecular level detail on the underlying interactions and structures responsible for the features observed. Additionally, MD simulations allow researchers to further the boundaries of experimental work by quickly testing new systems that would otherwise be very costly or outright impossible to perform experimentally.

The accuracy of the information obtained by these MD models relies heavily on the quality of the topologies used to describe the different molecular entities. As such, we applied our knowledge of PI(4,5)P_2 chemical and biophysical properties to develop and validate Martini 3 CG topologies of PI(4,5)P_2 and the remaining lipids in the phosphoinositide family. In this work, we developed these models independently of their Martini 2 counterparts, with greater accuracy and expanded application scope to include, among others, accurate reproduction of cation-mediated phosphoinositide aggregation. These models will hopefully be used by the Martini community in the future and thus indirectly contribute to thousands of studies on PI(4,5)P_2, and general phosphoinositide, biophysics.

Moving forward, an important aspect of PI(4,5)P_2 biophysics which is yet to be fully understood is its

association with membrane curvature. In this work we attempted to explore the association of PI(4,5)P_2 and PI(4,5)P_2 nanodomains with regions of specific membrane curvature. As mentioned previously, while PI(4,5)P_2 is essential for membrane fusion processes such as SNARE-dependent membrane fusion, it has an intrinsic positive curvature which is expected to restrain the formation of the negatively curved fusion intermediates necessary for some of these mechanisms.

In CG MD simulations, we observed that in the presence of divalent cations, PI(4,5)P_2 nanodomains would overcome this intrinsic curvature barrier and associate themselves with negatively curved membrane regions. Still, the extent of this association was hard to gauge and depended on the Martini force field used. Late in my PhD studies, we also attempted more complex simulation systems, such as the pulling of membrane tethers, in an attempt to understand how mechanically induced membrane curvature might influence the lateral organization of PI(4,5)P_2. However, due to a mixture of lack of time, high complexity and computational cost of the simulation systems involved, these studies are not yet completed and were not included in this thesis. Experimentally, studying PI(4,5)P_2 association with curvature is no easy task either. Micropipette aspiration is a powerful technique that could be used to study this. Hopefully, with time, these efforts into researching the impact of curvature on PI(4,5)P_2 biophysics will be taken over and resumed.

Apart from the work on PI(4,5)P_2 biophysics, part of my research efforts were also spent tackling the SARS-CoV-2 pandemic. In this thesis, two of these works are described. Our first contributions were based on the design and optimization of biopharmaceuticals against SARS-CoV-2, which were performed using an in-house developed genetic algorithm (GA)-guided ROSETTA approach. In this work, we applied this method to design and optimize an ACE2-based SARS-CoV-2 neutralizing molecule which could have potential use as pharmaceutical. The molecule designed in this work is, at the time of writing, undergoing *in vitro* testing.

We also worked on the impact of the alpha, beta and delta variants on the conformational dynamics and ACE2 accessibility of the SARS-CoV-2 Spike (S) glycoprotein receptor binding domain (RBD). In this work, long all-atom (AA) MD simulations show that in solution, *wt* RBD presents two distinct conformations: an “open” conformation where it is free to bind ACE2; and a “closed” conformation, where the RBM ridge blocks the binding surface. The alpha and beta variants shifted the open/closed equilibrium towards the open conformation by roughly 20%, while the delta variant did not show the presence of a closed conformation at all. This shift towards more open conformations likely enhances ACE2 binding affinity by increasing accessibility to the RBM and facilitating binding. Curiously, delta also showed an alternative open conformation, distinct from that of the other variants, where the ACE2 binding surface is kept open and accessible for binding but it significantly alters the conformation of the RBM ridge. We hypothesize that this may provide a fitness advantage by aiding in antibody escape given that many RBD-targeting antibodies are known to target the RBM ridge region.

Bibliography

[1] L. Borges-Araújo and F. Fernandes. Structure and Lateral Organization of Phosphatidylinositol 4,5-bisphosphate. *Molecules*, 25(17):3885, aug 2020. ISSN 1420-3049. doi: 10.3390/molecules25173885. URL https://www.mdpi.com/1420-3049/25/17/3885.

[2] J. Viaud, R. Mansour, A. Antkowiak, A. Mujalli, C. Valet, G. Chicanne, J. M. Xuereb, A. D. Terrisse, S. Séverin, M. P. Gratacap, F. Gaits-Iacovoni, and B. Payrastre. Phosphoinositides: Important lipids in the coordination of cell dynamics. *Biochimie*, 125:250–258, 2016. ISSN 61831638. doi: 10.1016/j.biochi.2015.09.005.

[3] M. J. Berridge. Inositol trisphosphate and diacylglycerol: two interacting second messengers. *Annual review of biochemistry*, 56(1):159–193, jun 1987. ISSN 00664154. doi: 10.1146/annurev.bi.56.070187.001111.

[4] T. Balla. Phosphoinositides: tiny lipids with giant impact on cell regulation. *Physiological reviews*, 93 (3):1019–1137, 2013. ISSN 1522-1210. doi: 10.1152/physrev.00028.2012. URL http://www.ncbi.nlm.nih.gov/pubmed/23899561.

[5] K. Tsujita and T. Itoh. Phosphoinositides in the regulation of actin cortex and cell migration, jun 2015. ISSN 18792618.

[6] D. W. Hilgemann, S. Feng, and C. Nasuhoglu. The complex and intriguing lives of PIP2 with ion channels and transporters. *Science's STKE : signal transduction knowledge environment*, 2001 (111):re19, 2001. ISSN 1525-8882. doi: 10.1126/stke.2001.111.re19.

[7] B. Hille, E. J. Dickson, M. Kruse, O. Vivas, and B. C. Suh. Phosphoinositides regulate ion channels, jun 2015. ISSN 18792618.

[8] S. Höning, D. Ricotta, M. Krauss, K. Späte, B. Spolaore, A. Motley, M. Robinson, C. Robinson, V. Haucke, and D. J. Owen. Phosphatidylinositol-(4,5)-bisphosphate regulates sorting signal recognition by the clathrin-associated adaptor complex AP2. *Molecular Cell*, 18(5):519–531, may 2005. ISSN 10972765. doi: 10.1016/j.molcel.2005.04.019. URL http://www.ncbi.nlm.nih.gov/pubmed/15916959http://linkinghub.elsevier.com/retrieve/pii/S1097276505012827.

[9] M. R. Wenk and P. De Camilli. Protein-lipid interactions and phosphoinositide metabolism in membrane traffic: Insights from vesicle recycling in nerve terminals. *Proceedings of the National*

Academy of Sciences of the United States of America, 101(22):8262–8269, jun 2004. ISSN 00278424. doi: 10.1073/pnas.0401874101. URL http://www.ncbi.nlm.nih.gov/pubmed/15146067http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC420382http://www.pnas.org/cgi/doi/10.1073/pnas.0401874101.

[10] A. Simonsen, A. E. Wurmser, S. D. Emr, and H. Stenmark. The role of phosphoinositides in membrane transport. *Current Opinion in Cell Biology*, 13(4):485–492, aug 2001. ISSN 09550674. doi: 10.1016/S0955-0674(00)00240-4. URL http://www.ncbi.nlm.nih.gov/pubmed/11454456.

[11] Y. Posor, M. Eichhorn-Grünig, and V. Haucke. Phosphoinositides in endocytosis, jun 2015. ISSN 18792618.

[12] R. Levin, S. Grinstein, and D. Schlam. Phosphoinositides in phagocytosis and macropinocytosis, jun 2015. ISSN 18792618.

[13] R. W. Holz, M. D. Hlubek, S. D. Sorensen, S. K. Fisher, T. Balla, S. Ozaki, G. D. Prestwich, E. L. Stuenkel, and M. A. Bittner. A pleckstrin homology domain specific for phosphatidylinositol 4,5- bisphosphate (PtdIns-4,5-P2) and fused to green fluorescent protein identifies plasma membrane PtdIns-4,5-P2 as being important in exocytosis. *Journal of Biological Chemistry*, 275(23):17878–17885, jun 2000. ISSN 00219258. doi: 10.1074/jbc.M000925200. URL http://www.ncbi.nlm.nih.gov/pubmed/10747966http://www.jbc.org/cgi/doi/10.1074/jbc.M000925200http://www.jbc.org/lookup/doi/10.1074/jbc.M000925200.

[14] T. F. Martin. PI(4,5)P2 regulation of surface membrane traffic. *Current Opinion in Cell Biology*, 13(4):493–499, aug 2001. ISSN 09550674. doi: 10.1016/S0955-0674(00)00241-6. URL http://www.ncbi.nlm.nih.gov/pubmed/11454457.

[15] T. F. Martin. PI(4,5)P2-binding effector proteins for vesicle exocytosis, jun 2015. ISSN 18792618.

[16] G. Di Paolo and P. De Camilli. Phosphoinositides in cell regulation and membrane dynamics. *Nature*, 443(7112):651–657, 2006. ISSN 0028-0836. doi: 10.1038/nature05185.

[17] E. E. E. Kooijman, K. E. E. King, M. Gangoda, and A. Gericke. Ionization properties of phosphatidylinositol polyphosphates in mixed model membranes. *Biochemistry*, 48(40):9360–9371, 2009. ISSN 00062960. doi: 10.1021/bi9008616.

[18] H. Ishihara, Y. Shibasaki, N. Kizuki, T. Wada, Y. Yazaki, T. Asano, and Y. Oka. Type I phosphatidylinositol-4-phosphate 5-kinases. Cloning of the third isoform and deletion/substitution analysis of members of this novel lipid kinase family. *Journal of Biological Chemistry*, 273(15):8741–8748, 1998. ISSN 00219258. doi: 10.1074/jbc.273.15.8741.

[19] J. H. Clarke, P. C. Emson, and R. F. Irvine. Localization of phosphatidylinositol phosphate kinase IIgamma in kidney to a membrane trafficking compartment within specialized cells of the nephron. *American journal of physiology. Renal physiology*, 295(5):F1422–F1430, 2008. ISSN 0363-6127. doi: 10.1152/ajprenal.90310.2008.

[20] M. W. Bunce, I. V. Boronenkov, and R. A. Anderson. Coordinated activation of the nuclear ubiquitin ligase Cul3-SPOP by the generation of phosphatidylinositol 5-phosphate. *Journal of Biological Chemistry*, 283(13):8678–8686, 2008. ISSN 00219258. doi: 10.1074/jbc.M710222200.

[21] J.-O. De Craene, D. Bertazzi, S. Bär, and S. Friant. Phosphoinositides, Major Actors in Membrane Trafficking and Lipid Signaling Pathways. *International Journal of Molecular Sciences*, 18(3):634, mar 2017. ISSN 1422-0067. doi: 10.3390/ijms18030634. URL http://www.mdpi.com/1422-0067/18/3/634.

[22] E. J. Dickson and B. Hille. Understanding phosphoinositides: rare, dynamic, and essential membrane phospholipids. *Biochemical Journal*, 476(1):1–23, jan 2019. ISSN 0264-6021. doi: 10.1042/BCJ20180022. URL /biochemj/article/476/1/1/84/Understanding-phosphoinositides-rare-dynamic-andhttps://portlandpress.com/biochemj/article/476/1/1/84/Understanding-phosphoinositides-rare-dynamic-and.

[23] J. P. Bradshaw, R. J. Bushby, C. C. Giles, M. R. Saunders, and A. Saxena. The headgroup orientation of dimyristoylphosphatidylinositol-4-phosphate in mixed lipid bilayers: a neutron diffraction study. *Biochimica et Biophysica Acta - Biomembranes*, 1329(1):124–138, oct 1997. URL https://www.sciencedirect.com/science/article/pii/S000527369700093Xhttp://www.ncbi.nlm.nih.gov/pubmed/9370250.

[24] C. Zhou, V. Garigapati, and M. F. Roberts. Short-chain phosphatidylinositol conformation and its relevance to phosphatidylinositol-specific phospholipase C. *Biochemistry*, 36(50):15925–15931, dec 1997. ISSN 00062960. doi: 10.1021/bi9716175. URL http://www.ncbi.nlm.nih.gov/pubmed/9398326.

[25] J. P. Bradshaw, R. J. Bushby, C. C. Giles, and M. R. Saunders. Orientation of the headgroup of phosphatidylinositol in a model biomembrane as determined by neutron diffraction. *Biochemistry*, 38 (26):8393–401, jun 1999. ISSN 0006-2960. doi: 10.1021/bi990338+. URL http://www.ncbi.nlm.nih.gov/pubmed/10387085.

[26] A. I. Kishore and J. H. Prestegard. Molecular Orientation and Conformation of Phosphatidylinositides in Membrane Mimetics Using Variable Angle Sample Spinning (VASS) NMR. *Biophysical Journal*, 85(6):3848–3857, dec 2003. ISSN 00063495. doi: 10.1016/S0006-3495(03)74799-7.

[27] P. M. Hansbro, S. J. Byard, R. J. Bushby, P. J. Turnbull, N. Boden, M. R. Saunders, R. Novelli, and D. G. Reid. The conformational behaviour of phosphatidylinositol in model membranes: 2H-NMR studies. *Biochimica et biophysica acta*, 1112(2):187–96, dec 1992. ISSN 0006-3002. doi: 10.1016/0005-2736(92)90391-x. URL http://www.ncbi.nlm.nih.gov/pubmed/1457451.

[28] D. R. Slochower, P. J. Huwe, R. Radhakrishnan, and P. A. Janmey. Quantum and All-Atom Molecular Dynamics Simulations of Protonation and Divalent Ion Binding to Phosphatidylinositol 4,5-Bisphosphate (PIP 2). *The Journal of Physical Chemistry B*, 117(28):8322–8329, jul

2013. ISSN 1520-6106. doi: 10.1021/jp401414y. URL http://pubs.acs.org/doi/abs/10.1021/jp401414yhttps://pubs.acs.org/doi/10.1021/jp401414y.

[29] W. H. Brown. *Organic chemistry.* Brooks/Cole Cengage Learning, 2009. ISBN 0495388572.

[30] P. A. van Paridon, B. de Kruijff, R. Ouwerkerk, and K. W. Wirtz. Polyphosphoinositides undergo charge neutralization in the physiological pH range: a 31P-NMR study. *Biochimica et Biophysica Acta (BBA) - Lipids and Lipid Metabolism*, 877(1):216–219, jun 1986. ISSN 00052760. doi: 10.1016/0005-2760(86)90137-2. URL http://www.ncbi.nlm.nih.gov/pubmed/3013316https://pubmed.ncbi.nlm.nih.gov/3013316/https://linkinghub.elsevier.com/retrieve/pii/0005276086901372.

[31] Z. T. T. Graber, Z. Jiang, A. Gericke, and E. E. E. Kooijman. Phosphatidylinositol-4,5-bisphosphate ionization and domain formation in the presence of lipids with hydrogen bond donor capabilities. *Chemistry and Physics of Lipids*, 165(6):696–704, sep 2012. ISSN 00093084. doi: 10.1016/j.chemphyslip.2012.07.003. URL http://www.ncbi.nlm.nih.gov/pubmed/22820347.

[32] A. Björkbom, H. Ohvo-Rekilä, P. Kankaanpää, T. K. Nyholm, B. Westerlund, and J. P. Slotte. Characterization of membrane properties of inositol phosphorylceramide. *Biochimica et Biophysica Acta - Biomembranes*, 1798(3):453–460, mar 2010. ISSN 00052736. doi: 10.1016/j.bbamem.2009.11.003.

[33] K. D'Souza and R. M. Epand. Enrichment of phosphatidylinositols with specific acyl chains. *Biochimica et Biophysica Acta (BBA)*, 1838(6):1501–1508, jun 2014. ISSN 00052736. doi: 10.1016/j.bbamem.2013.10.003. URL https://linkinghub.elsevier.com/retrieve/pii/S0005273613003568.

[34] A. Traynor-Kaplan, M. Kruse, E. J. Dickson, G. Dai, O. Vivas, H. Yu, D. Whittington, and B. Hille. Fatty-acyl chain profiles of cellular phosphoinositides. *Biochimica et Biophysica Acta (BBA)*, 1862(5):513–522, may 2017. ISSN 13881981. doi: 10.1016/j.bbalip.2017.02.002. URL http://linkinghub.elsevier.com/retrieve/pii/S1388198117300264https://linkinghub.elsevier.com/retrieve/pii/S1388198117300264.

[35] A. Mujalli, G. Chicanne, J. Bertrand-Michel, F. Viars, L. Stephens, P. Hawkins, J. Viaud, F. Gaits-Iacovoni, S. Severin, M. P. Gratacap, A. D. Terrisse, and B. Payrastre. Profiling of phosphoinositide molecular species in human and mouse platelets identifies new species increasing following stimulation. *Biochimica et Biophysica Acta - Molecular and Cell Biology of Lipids*, 1863(9):1121–1131, sep 2018. ISSN 18792618. doi: 10.1016/j.bbalip.2018.06.009. URL https://linkinghub.elsevier.com/retrieve/pii/S138819811830132X.

[36] A. Koizumi, S. Narita, H. Nakanishi, M. Ishikawa, S. Eguchi, H. Kimura, S. Takasuga, M. Huang, T. Inoue, J. Sasaki, T. Yoshioka, T. Habuchi, and T. Sasaki. Increased fatty acyl saturation of phosphatidylinositol phosphates in prostate cancer progression. *Scientific Reports*, 9(1), dec 2019. ISSN 20452322. doi: 10.1038/s41598-019-49744-3.

[37] M. Koch and M. Holt. Coupling exo- and endocytosis: An essential role for PIP2 at the synapse. *Biochimica et Biophysica Acta (BBA)*, 1821(8):1114–1132, aug 2012. ISSN 13881981. doi: 10.1016/j.bbalip.2012.02.008. URL https://linkinghub.elsevier.com/retrieve/pii/S1388198112000406.

[38] M. M. Manni, M. L. Tiberti, S. Pagnotta, H. Barelli, R. Gautier, and B. Antonny. Acyl chain asymmetry and polyunsaturation of brain phospholipids facilitate membrane vesiculation without leakage. *eLife*, 7, mar 2018. ISSN 2050084X. doi: 10.7554/eLife.34394.

[39] D. L. Nelson and M. M. Cox. *Lehninger Principles of Biochemistry Sixth Edition*. macmillan, 6th edition, 2013. URL http://www.macmillanlearning.com/catalog/Product/lehningerprinciplesofbiochemistry-sixthedition-nelson.

[40] S. McLaughlin, J. Wang, A. Gambhir, and D. Murray. PIP 2 and Proteins: Interactions, Organization, and Information Flow. *Annual Review of Biophysics and Biomolecular Structure*, 31(1):151–175, jun 2002. ISSN 1056-8700. doi: 10.1146/annurev.biophys.31.082901.134259. URL http://www.annualreviews.org/doi/10.1146/annurev.biophys.31.082901.134259.

[41] U. Golebiewska, M. Nyako, W. Woturski, I. Zaitseva, and S. McLaughlin. Diffusion coefficient of fluorescent phosphatidylinositol 4,5-bisphosphate in the plasma membrane of cells. *Molecular Biology of the Cell*, 19(4):1663–1669, apr 2008. ISSN 10591524. doi: 10.1091/mbc.E07-12-1208. URL http://www.ncbi.nlm.nih.gov/pubmed/18256277http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2291420http://www.molbiolcell.org/cgi/doi/10.1091/mbc.E07-12-1208.

[42] G. van den Bogaart, K. Meyenberg, H. J. Risselada, H. Amin, K. I. Willig, B. E. Hubrich, M. Dier, S. W. Hell, H. Grubmüller, U. Diederichsen, and R. Jahn. Membrane protein sequestering by ionic protein–lipid interactions. *Nature*, 479(7374):552–555, oct 2011. ISSN 0028-0836. doi: 10.1038/nature10545. URL http://www.nature.com/doifinder/10.1038/nature10545.

[43] J. Wang and D. A. Richards. Segregation of PIP2 and PIP3 into distinct nanoscale regions within the plasma membrane. *Biology Open*, 1(9):857–862, sep 2012. ISSN 20466390. doi: 10.1242/bio.20122071. URL http://www.ncbi.nlm.nih.gov/pubmed/23213479http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3507238.

[44] B. Payrastre, K. Missy, S. Giuriato, S. Bodin, M. Plantavid, and M. P. Gratacap. Phosphoinositides: Key players in cell signalling, in time and space. *Cellular Signalling*, 13(6):377–387, jun 2001. ISSN 08986568. doi: 10.1016/S0898-6568(01)00158-9. URL http://www.ncbi.nlm.nih.gov/pubmed/11384836.

[45] T. Laux, K. Fukami, M. Thelen, T. Golub, D. Frey, and P. Caroni. GAP43, MARCKS, and CAP23 modulate PI(4,5)P2 at plasmalemmal rafts, and regulate cell cortex actin dynamics through a common mechanism. *Journal of Cell Biology*, 149(7):1455–1471, 2000. ISSN

00219525. doi: 10.1083/jcb.149.7.1455. URL http://jcb.rupress.org/content/149/7/1455?ijkey=b6cad8574126ba9e7d3c33dc15e36341f84d0fb3{&}keytype2=tf{_}ipsecsha.

[46] J. Wang, A. Gambhir, S. McLaughlin, and D. Murray. A Computational Model for the Electrostatic Sequestration of PI(4,5)P2 by Membrane-Adsorbed Basic Peptides. *Biophysical Journal*, 86(4): 1969–1986, apr 2004. ISSN 00063495. doi: 10.1016/S0006-3495(04)74260-5. URL https://linkinghub.elsevier.com/retrieve/pii/S0006349504742605.

[47] A. Gambhir, G. Hangyas-Mihalyne, I. Zaitseva, D. S. Cafiso, J. Wang, D. Murray, S. N. Pentyala, S. O. Smith, and S. McLaughlin. Electrostatic Sequestration of PIP2 on Phospholipid Membranes by Basic/Aromatic Regions of Proteins. *Biophysical Journal*, 86(4):2188–2207, apr 2004. ISSN 00063495. doi: 10.1016/S0006-3495(04)74278-2. URL http://www.ncbi.nlm.nih.gov/pubmed/15041659http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC1304070.

[48] S. McLaughlin and D. Murray. Plasma membrane phosphoinositide organization by protein electrostatics. *Nature*, 438(December):605–611, 2005. ISSN 0028-0836. doi: 10.1038/nature04398.

[49] M. E. Rauch, C. G. Ferguson, G. D. Prestwich, and D. S. Cafiso. Myristoylated alanine-rich C kinase substrate (MARCKS) sequesters spin-labeled phosphatidylinositol 4,5-bisphosphate in lipid bilayers. *Journal of Biological Chemistry*, 277(16):14068–14076, apr 2002. ISSN 00219258. doi: 10.1074/jbc.M109572200. URL http://www.ncbi.nlm.nih.gov/pubmed/11825894http://www.jbc.org/lookup/doi/10.1074/jbc.M109572200.

[50] J. Wang, A. Gambhir, G. Hangyás-Mihályné, D. Murray, U. Golebiewska, and S. McLaughlin. Lateral sequestration of phosphatidylinositol 4,5-bisphosphate by the basic effector domain of myristoylated alanine-rich C kinase substrate is due to nonspecific electrostatic interactions. *Journal of Biological Chemistry*, 277(37):34401–34412, sep 2002. ISSN 00219258. doi: 10.1074/jbc.M203954200. URL http://www.ncbi.nlm.nih.gov/pubmed/12097325http://www.jbc.org/cgi/doi/10.1074/jbc.M203954200http://www.jbc.org/lookup/doi/10.1074/jbc.M203954200.

[51] U. Dietrich, P. Krüger, T. Gutberlet, and J. A. Käs. Interaction of the MARCKS peptide with PIP2 in phospholipid monolayers. *Biochimica et Biophysica Acta - Biomembranes*, 1788(7):1474–1481, 2009. ISSN 00052736. doi: 10.1016/j.bbamem.2009.04.001. URL http://www.sciencedirect.com/science/article/pii/S0005273609001126.

[52] J. Tong, L. Nguyen, A. Vidal, S. A. Simon, J. H. Skene, and T. J. McIntosh. Role of GAP-43 in sequestering phosphatidylinositol 4,5-bisphosphate to raft bilayers. *Biophysical Journal*, 94(1):125–133, jan 2008. ISSN 00063495. doi: 10.1529/biophysj.107.110536. URL http://www.ncbi.nlm.nih.gov/pubmed/17827240http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2134862.

[53] K. Aoyagi, T. Sugaya, M. Umeda, S. Yamamoto, S. Terakawa, and M. Takahashi. The activation of exocytotic sites by the formation of phosphatidylinositol 4,5-bisphosphate microdomains at syntaxin

clusters. *Journal of Biological Chemistry*, 280(17):17346–17352, apr 2005. ISSN 00219258. doi: 10.1074/jbc.M413307200.

[54] D. H. Murray and L. K. Tamm. Clustering of syntaxin-1A in model membranes is modulated by phosphatidylinositol 4,5-bisphosphate and cholesterol. *Biochemistry*, 48(21):4617–4625, jun 2009. ISSN 00062960. doi: 10.1021/bi9003217.

[55] D. J. James, C. Khodthong, J. A. Kowalchyk, and T. F. Martin. Phosphatidylinositol 4,5-bisphosphate regulates SNARE-dependent membrane fusion. *Journal of Cell Biology*, 182(2):355–366, jul 2008. ISSN 1540-8140. doi: 10.1083/jcb.200801056. URL https://rupress.org/jcb/article/182/2/355/34989/Phosphatidylinositol-45bisphosphate-regulates.

[56] I. Milosevic, J. B. Sørensen, T. Lang, M. Krauss, G. Nagy, V. Haucke, R. Jahn, and E. Neher. Plasmalemmal phosphatidylinositol-4,5-bisphosphate level regulates the releasable vesicle pool size in chromaffin cells. *Journal of Neuroscience*, 25(10):2557–2565, mar 2005. ISSN 02706474. doi: 10.1523/JNEUROSCI.3761-04.2005.

[57] J. Shi, L. Birnbaumer, W. A. Large, and A. P. Albert. Myristoylated alanine-rich C kinase substrate coordinates native TRPC1 channel activation by phosphatidylinositol 4,5-bisphosphate and protein kinase C in vascular smooth muscle. *FASEB Journal*, 28(1):244–255, jan 2014. ISSN 15306860. doi: 10.1096/fj.13-238022. URL http://www.ncbi.nlm.nih.gov/pubmed/24022404http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3868830.

[58] R. Su, Z. Y. Han, J. P. Fan, and Y. L. Zhang. A possible role of myristoylated alanine-rich C kinase substrate in endocytic pathway of Alzheimer's disease. *Neuroscience Bulletin*, 26(4):338–344, aug 2010. ISSN 16737067. doi: 10.1007/s12264-010-0131-0. URL http://link.springer.com/10.1007/s12264-010-0131-0.

[59] L. Trovò, T. Ahmed, Z. Callaerts-Vegh, A. Buzzi, C. Bagni, M. Chuah, T. VandenDriessche, D. Balschun, and C. G. Dotti. Low hippocampal PI(4,5)P2 contributes to reduced cognition in old mice as a result of loss of MARCKS. *Nature Publishing Group*, 16(4):4–4, 2013. doi: 10.1038/nn.3342. URL https://www.nature.com/neuro/journal/v16/n4/pdf/nn.3342.pdf.

[60] D. E. Clapham. Calcium signaling. *Cell*, 80(2):259–268, 1995. ISSN 00928674. doi: 10.1016/0092-8674(95)90408-5.

[61] M. J. Berridge. Calcium microdomains: Organization and function. *Cell Calcium*, 40(5-6):405–412, 2006. ISSN 01434160. doi: 10.1016/j.ceca.2006.09.002. URL http://www.sciencedirect.com/science/article/pii/S0143416006001709.

[62] R. D. Grubbs. Intracellular magnesium and magnesium buffering, sep 2002. ISSN 09660844. URL http://www.ncbi.nlm.nih.gov/pubmed/12206391.

[63] V. Trapani, G. Farruggia, C. Marraccini, S. Iotti, A. Cittadini, and F. I. Wolf. Intracellular magnesium detection: imaging a brighter future. *The Analyst*, 135:1855–1866, 2010. ISSN 0003-2654. doi: 10.1039/c0an00087f.

[64] D. R. Slochower, Y.-H. Wang, R. Radhakrishnan, and P. A. Janmey. Physical chemistry and membrane properties of two phosphatidylinositol bisphosphate isomers. *Phys. Chem. Chem. Phys.*, 17(19):12608–12615, 2015. ISSN 1463-9076. doi: 10.1039/C5CP00862J. URL http://dx.doi.org/10.1039/C5CP00862J{%}5Cnhttp://xlink.rsc.org/?DOI=C5CP00862J.

[65] Y.-H. Wang, A. Collins, L. Guo, K. B. Smith-Dupont, F. Gai, T. Svitkina, and P. A. Janmey. Divalent Cation-Induced Cluster Formation by Polyphosphoinositides in Model Membranes. *Journal of the American Chemical Society*, 134(7):3387–3395, feb 2012. ISSN 0002-7863. doi: 10.1021/ja208640t. URL http://www.ncbi.nlm.nih.gov/pubmed/22280226http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3445022https://pubs.acs.org/doi/10.1021/ja208640t.

[66] W. G. Ellenbroek, Y. H. Wang, D. A. Christian, D. E. Discher, P. A. Janmey, and A. J. Liu. Divalent cation-dependent formation of electrostatic PIP2 clusters in lipid monolayers. *Biophysical Journal*, 101(9):2178–2184, 2011. ISSN 00063495. doi: 10.1016/j.bpj.2011.09.039. URL /pmc/articles/PMC3207175/?report=abstracthttps://www.ncbi.nlm.nih.gov/pmc/articles/PMC3207175/http://dx.doi.org/10.1016/j.bpj.2011.09.039.

[67] M. J. Sarmento, A. Coutinho, A. Fedorov, M. Prieto, and F. Fernandes. Ca2+ induces PI(4,5)P2 clusters on lipid bilayers at physiological PI(4,5)P2 and Ca2+ concentrations. *Biochimica et Biophysica Acta (BBA)*, 1838(3):822–830, mar 2014. ISSN 00052736. doi: 10.1016/j.bbamem.2013.11.020. URL http://dx.doi.org/10.1016/j.bbamem.2013.11.020https://linkinghub.elsevier.com/retrieve/pii/S0005273613004288.

[68] M. J. Sarmento, A. Coutinho, A. Fedorov, M. Prieto, and F. Fernandes. Membrane Order Is a Key Regulator of Divalent Cation-Induced Clustering of PI(3,5)P2 and PI(4,5)P2. *Langmuir*, 33(43):12463–12477, oct 2017. ISSN 15205827. doi: 10.1021/acs.langmuir.7b00666.

[69] J. A. Cohen and M. Cohen. Adsorption of monovalent and divalent cations by phospholipid membranes. The monomer-dimer problem. *Biophysical Journal*, 36(3):623–651, 1981. ISSN 00063495. doi: 10.1016/S0006-3495(81)84756-X.

[70] Y. Wen, V. M. Vogt, and G. W. Feigenson. Multivalent Cation-Bridged PI(4,5)P2 Clusters Form at Very Low Concentrations. *Biophysical Journal*, 114(11):2630–2639, jun 2018. ISSN 00063495. doi: 10.1016/j.bpj.2018.04.048. URL http://www.ncbi.nlm.nih.gov/pubmed/29874613http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC6129474https://linkinghub.elsevier.com/retrieve/pii/S0006349518305733.

[71] M. E. Monteiro, M. J. Sarmento, and F. Fernandes. Role of calcium in membrane interactions by PI(4,5)P2-binding proteins. *Biochemical Society transactions*, 42(5):1441–6, 2014. ISSN 1470-8752. doi: 10.1042/BST20140149. URL http://www.ncbi.nlm.nih.gov/pubmed/25233429.

[72] J. B. Seo, S. R. Jung, W. Huang, Q. Zhang, and D. S. Koh. Charge shielding of PIP2 by cations regulates enzyme activity of phospholipase C. *PLoS ONE*, 10(12):1–22, 2015. ISSN 19326203. doi: 10.1371/journal.pone.0144432.

[73] E. Bilkova, R. Pleskot, S. Rissanen, S. Sun, A. Czogalla, L. Cwiklik, T. Rog, I. Vattulainen, P. S. Cremer, P. Jungwirth, and Ü. Coskun. Calcium directly regulates phosphatidylinositol 4,5-bisphosphate headgroup conformation and recognition. *Journal of the American Chemical Society*, page jacs.6b11760, 2017. ISSN 0002-7863. doi: 10.1021/jacs.6b11760. URL http://pubs.acs.org/doi/abs/10.1021/jacs.6b11760.

[74] W. K. Subczynski, M. Pasenkiewicz-Gierula, J. Widomska, L. Mainali, and M. Raguz. High Cholesterol/Low Cholesterol: Effects in Biological Membranes: A Review. *Cell Biochemistry and Biophysics*, 75(3-4):369–385, dec 2017. ISSN 10859195. doi: 10.1007/s12013-017-0792-7.

[75] I. Levental, D. A. Christian, Y.-H. H. Wang, J. J. Madara, D. E. Discher, and P. A. Janmey. Calcium-dependent lateral organization in phosphatidylinositol 4,5-bisphosphate (PIP2)- and cholesterol-containing monolayers. *Biochemistry*, 48(34):8241–8248, sep 2009. ISSN 00062960. doi: 10.1021/bi9007879. URL https://pubs.acs.org/doi/10.1021/bi9007879.

[76] J. Kwik, S. Boyle, D. Fooksman, L. Margolis, M. P. Sheetz, and M. Edidin. Membrane cholesterol, lateral mobility, and the phosphatidylinositol 4,5-bisphosphate-dependent organization of cell actin. *Proceedings of the National Academy of Sciences of the United States of America*, 100(SUPPL. 2):13964–13969, nov 2003. ISSN 00278424. doi: 10.1073/pnas.2336102100. URL http://www.ncbi.nlm.nih.gov/pubmed/14612561http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC283529.

[77] M. Hao and J. S. Bogan. Cholesterol regulates glucose-stimulated insulin secretion through phosphatidylinositol 4,5-bisphosphate. *Journal of Biological Chemistry*, 284(43):29489–29498, oct 2009. ISSN 00219258. doi: 10.1074/jbc.M109.038034. URL http://www.ncbi.nlm.nih.gov/pubmed/19729450http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2785582.

[78] R. M. Epand. Proteins and cholesterol-rich domains, jul 2008. ISSN 00052736. URL http://www.ncbi.nlm.nih.gov/pubmed/18423371.

[79] L. J. Pike and J. M. Miller. Cholesterol depletion delocalizes phosphatidylinositol bisphosphate and inhibits hormone-stimulated phosphatidylinositol turnover. *Journal of Biological Chemistry*, 273(35):22298–22304, aug 1998. ISSN 00219258. doi: 10.1074/jbc.273.35.22298. URL http://www.ncbi.nlm.nih.gov/pubmed/9712847.

[80] J. Van Rheenen, E. M. Achame, H. Janssen, J. Calafat, and K. Jalink. PIP2 signaling in lipid domains: A critical re-evaluation. *EMBO Journal*, 24(9):1664–1673, may 2005. ISSN 02614189. doi: 10.1038/sj.emboj.7600655.

[81] L. Mu, Z. Tu, L. Miao, H. Ruan, N. Kang, Y. Hei, J. Chen, W. Wei, F. Gong, B. Wang, Y. Du, G. Ma, M. W. Amerein, T. Xia, and Y. Shi. A phosphatidylinositol 4,5-bisphosphate redistribution-based sensing mechanism initiates a phagocytosis programing. *Nature Communications*, 9(1):4259, dec 2018. ISSN 2041-1723. doi: 10.1038/s41467-018-06744-7. URL http://www.nature.com/articles/s41467-018-06744-7.

[82] G. A. Ponuwei. A glimpse of the ERM proteins. *Journal of Biomedical Science*, 23(1):35, mar 2016. ISSN 14230127. doi: 10.1186/s12929-016-0246-3. URL http://www.ncbi.nlm.nih.gov/pubmed/26983550http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC4794931.

[83] D. M. Taglieri, D. A. Delfín, and M. M. Monasky. Cholesterol regulation of PIP2: Why cell type is so important. *Frontiers in Physiology*, 3 JAN:492, jan 2013. ISSN 1664042X. doi: 10.3389/fphys.2012.00492. URL http://journal.frontiersin.org/article/10.3389/fphys.2012.00492/abstract.

[84] L. Chierico, A. S. Joseph, A. L. Lewis, and G. Battaglia. Live cell imaging of membrane/cytoskeleton interactions and membrane topology. *Scientific Reports*, 3(1), sep 2014. ISSN 20452322. doi: 10.1038/srep06056. URL http://www.ncbi.nlm.nih.gov/pubmed/25205456http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC4159633.

[85] C. Johnson and W. Rodgers. Spatial Segregation of Phosphatidylinositol 4,5-Bisphosphate (PIP2) Signaling in Immune Cell Functions. *Immunology, Endocrine & Metabolic Agents in Medicinal Chemistry*, 8(4):349–357, jan 2009. ISSN 18715222. doi: 10.2174/187152208787169233. URL http://www.ncbi.nlm.nih.gov/pubmed/19956793http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2771939.

[86] M. J. Sarmento, L. Borges-Araújo, S. N. Pinto, N. Bernardes, J. C. Ricardo, A. Coutinho, M. Prieto, and F. Fernandes. Quantitative FRET Microscopy Reveals a Crucial Role of Cytoskeleton in Promoting PI(4,5)P2 Confinement. *International Journal of Molecular Sciences*, 22(21):11727, oct 2021. ISSN 1422-0067. doi: 10.3390/ijms222111727. URL https://www.mdpi.com/1422-0067/22/21/11727.

[87] A. Fujita, J. Cheng, K. Tauchi-Sato, T. Takenawa, and T. Fujimoto. A distinct pool of phosphatidylinositol 4,5-bisphosphate in caveolae revealed by a nanoscale labeling technique. *Proceedings of the National Academy of Sciences*, 106(23):9256–9261, jun 2009. ISSN 0027-8424. doi: 10.1073/pnas.0900216106. URL https://pubmed.ncbi.nlm.nih.gov/19470488/http://www.pnas.org/cgi/doi/10.1073/pnas.0900216106.

[88] T. Nebl, S. W. Oh, and E. J. Luna. Membrane cytoskeleton: PIP2 pulls the strings, may 2000. ISSN 09609822.

[89] L. Zhang, Y. S. Mao, P. A. Janmey, and H. L. Yin. Phosphatidylinositol 4, 5 bisphosphate and the actin cytoskeleton. *Sub-Cellular Biochemistry*, 59:177–215, 2012. ISSN 03060225. doi: 10.1007/978-94-007-3015-1_6. URL http://www.ncbi.nlm.nih.gov/pubmed/22374091.

[90] D. M. Andrade, M. P. Clausen, J. Keller, V. Mueller, C. Wu, J. E. Bear, S. W. Hell, B. C. Lagerholm, and C. Eggeling. Cortical actin networks induce spatio-temporal confinement of phospholipids in the plasma membrane -A minimally invasive investigation by STED-FCS. *Scientific Reports*, 5(1):1–12, jun 2015. ISSN 20452322. doi: 10.1038/srep11454.

[91] H. Cho, Y. A. Kim, J. Y. Yoon, D. Lee, J. H. Kim, S. H. Lee, and W. K. Ho. Low mobility of phosphatidylinositol 4,5-bisphosphate underlies receptor specificity of Gq-mediated ion channel regulation in atrial myocytes. *Proceedings of the National Academy of Sciences of the United States of America*, 102(42):15241–15246, oct 2005. ISSN 00278424. doi: 10.1073/pnas.0408851102.

[92] E. F. Corbett-Nelson, D. Mason, J. G. Marshall, Y. Collette, and S. Grinstein. Signaling-dependent immobilization of acylated proteins in the inner monolayer of the plasma membrane. *Journal of Cell Biology*, 174(2):255–265, jul 2006. ISSN 00219525. doi: 10.1083/jcb.200605044.

[93] R. J. Botelho, M. Teruel, R. Dierckman, R. Anderson, A. Wells, J. D. York, T. Meyer, and S. Grinstein. Localized biphasic changes in phosphatidylinositol-4,5-bisphosphate at sites of phagocytosis. *Journal of Cell Biology*, 151(7):1353–1367, dec 2000. ISSN 00219525. doi: 10.1083/jcb.151.7.1353.

[94] J. Zimmerberg and L. V. Chernomordik. Membrane fusion. *Advanced Drug Delivery Reviews*, 38(3): 197–205, aug 1999. ISSN 0169409X. doi: 10.1016/S0169-409X(99)00029-0.

[95] F. Wu, S. Zhao, B. Yu, Y. M. Chen, W. Wang, Z. G. Song, Y. Hu, Z. W. Tao, J. H. Tian, Y. Y. Pei, M. L. Yuan, Y. L. Zhang, F. H. Dai, Y. Liu, Q. M. Wang, J. J. Zheng, L. Xu, E. C. Holmes, and Y. Z. Zhang. A new coronavirus associated with human respiratory disease in China. *Nature*, 579(7798):265–269, mar 2020. ISSN 14764687. doi: 10.1038/s41586-020-2008-3. URL `https://pubmed.ncbi.nlm.nih.gov/32015508/https://pubmed.ncbi.nlm.nih.gov/32015508/?dopt=Abstract`.

[96] D. K. W. Chu, Y. Pan, S. M. S. Cheng, K. P. Y. Hui, P. Krishnan, Y. Liu, D. Y. M. Ng, C. K. C. Wan, P. Yang, Q. Wang, M. Peiris, and L. L. M. Poon. Molecular Diagnosis of a Novel Coronavirus (2019-nCoV) Causing an Outbreak of Pneumonia. *Clinical Chemistry*, 66(4):549–555, apr 2020. ISSN 0009-9147. doi: 10.1093/clinchem/hvaa029. URL `https://academic.oup.com/clinchem/article/66/4/549/5719336`.

[97] F. P. Polack, S. J. Thomas, N. Kitchin, J. Absalon, A. Gurtman, S. Lockhart, J. L. Perez, G. Pérez Marc, E. D. Moreira, C. Zerbini, R. Bailey, K. A. Swanson, S. Roychoudhury, K. Koury, P. Li, W. V. Kalina, D. Cooper, R. W. Frenck, L. L. Hammitt, Ö. Türeci, H. Nell, A. Schaefer, S. Ünal, D. B. Tresnan, S. Mather, P. R. Dormitzer, U. Şahin, K. U. Jansen, and W. C. Gruber. Safety and Efficacy of the BNT162b2 mRNA Covid-19 Vaccine. *New England Journal of Medicine*, 383(27):2603–2615, dec 2020. ISSN 0028-4793. doi: 10.1056/NEJMoa2034577. URL `http://www.nejm.org/doi/10.1056/NEJMoa2034577`.

[98] E. Gorter and F. Grendel. On bimolecular layers of lipoids on the chromocytes of the blood. *Journal of Experimental Medicine*, 41(4):439–444, mar 1925. ISSN 15409538. doi: 10.1084/jem.41.4.439. URL `http://www.ncbi.nlm.nih.gov/pubmed/19868999http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2130960`.

[99] G. L. Nicolson. The Fluid - Mosaic Model of Membrane Structure: Still relevant to understanding the structure, function and dynamics of biological membranes after more than 40 years. *Biochimica*

et Biophysica Acta - Biomembranes, 1838(6):1451–1466, 2014. ISSN 18792642. doi: 10.1016/j.bbamem.2013.10.019. URL http://dx.doi.org/10.1016/j.bbamem.2013.10.019.

A | Supplementary Material for Chapter 2

Supplementary Material

Impact of Ca^{2+}-dependent $PI(4,5)P_2$ clustering on the properties of $PI(4,5)P_2$ binding proteins

Luís Borges-Araújo[†,‖,⊥], Marina E. Monteiro[†], Dalila Mil-Homens[†], Nuno Bernardes[†], Maria J. Sarmento[†], Ana Coutinho[†], Manuel Prieto[†], Manuel N. Melo[‖], Fábio Fernandes [†,‡,⊥,*]

† iBB-Institute for Bioengineering and Biosciences, Instituto Superior Técnico, Universidade de Lisboa, Lisbon, Portugal

⊥ Associate Laboratory i4HB—Institute for Health and Bioeconomy at Instituto Superior Técnico, Universidade de Lisboa, Lisbon, Portugal

‡ Department of Bioengineering; Instituto Superior Técnico, Universidade de Lisboa, 1049-001 Lisbon, Portugal;

‖ Instituto de Tecnologia Química e Biológica António Xavier, Universidade Nova de Lisboa, Av. da República, 2780-157 Oeiras, Portugal.

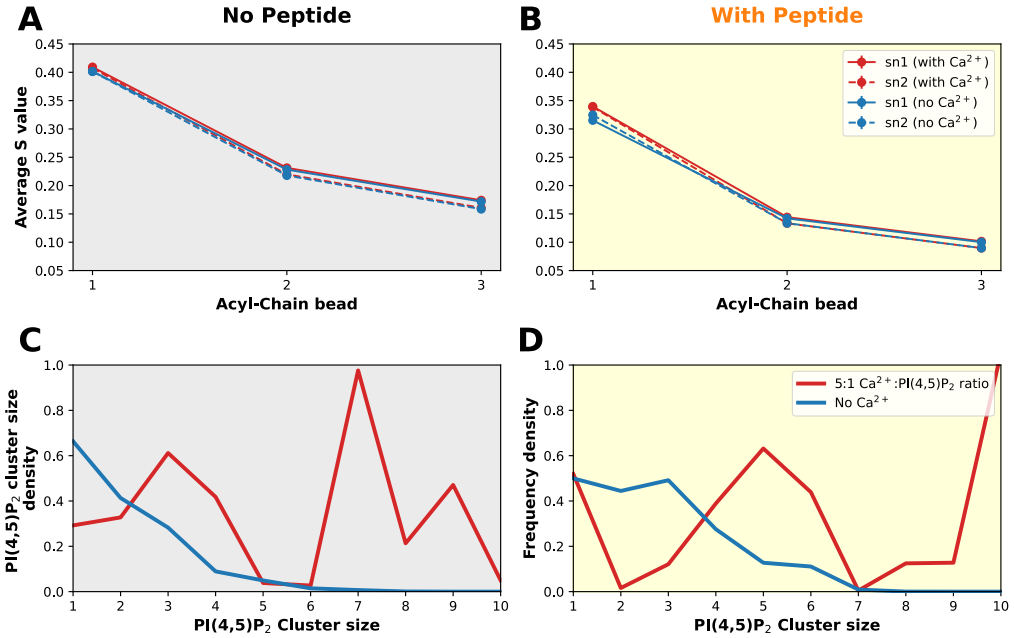


Figure. S1. Impact of PBP-10 on the biophysical properties of PI(4,5)P_2 and PI(4,5)P_2 clusters. The S-value order parameter was calculated for each acyl-chain bead, both in the presence (red) and absence (blue) of calcium, in systems with (B) and without (A) PBP-10. PI(4,5)P_2 cluster size histogram analysis both in the absence (C) and presence of peptide (D) and either in the presence (red) or absence (blue) of Ca^{2+}.

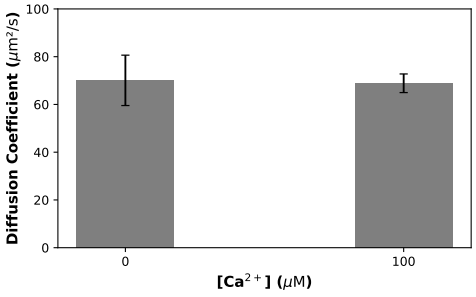


Figure. S2. Diffusion coefficients recovered for PH-YFP dispersed in buffer solution. Diffusion coefficients obtained by FCS for a 250 nM solution of PH-YFP, in the presence and absence of calcium. Mean and Standard deviation values shown.

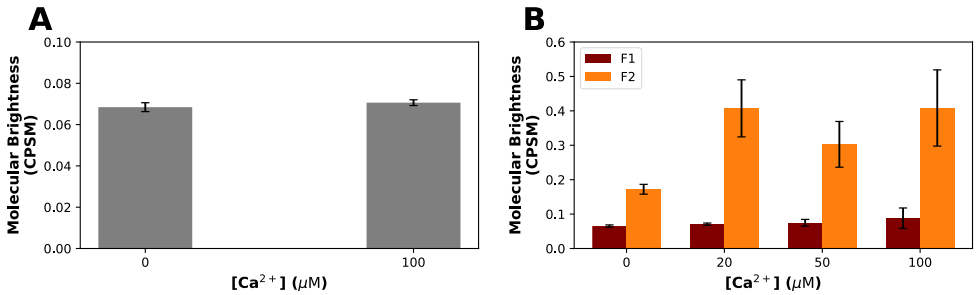


Figure. S3. Molecular brightness values recovered for PH-YFP samples. Molecular brightness values recovered by PCH of both a 250 nM solution of PH-YFP (A) and of PH-YFP in POPC:PI(4,5)P_2 (95:5 molar ratio) GUVs, in the presence and absence of calcium. For the PH-YFP samples in GUVs, 2 dominant brightness populations were recovered, a monomeric (red) and an oligomeric (orange) population.

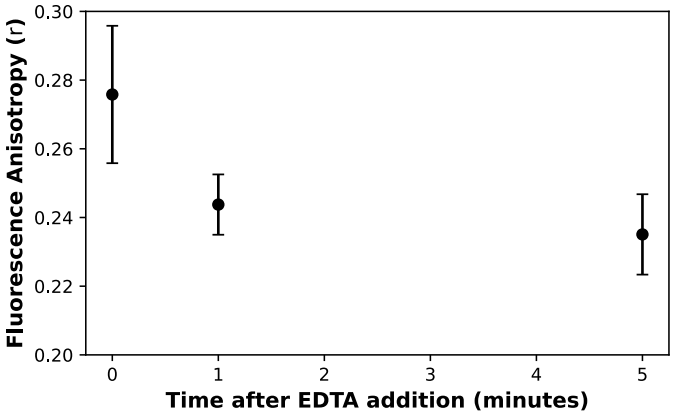


Figure. S4. Kinetics of fluorescence anisotropy response to the addition of 5mM EDTA to PH-YFP in 100 μM of PI(4,5)P_2 containing LUVs in the presence of 100 μM Ca^{2+}. Error bars refer to standard deviation values.

B | Supplementary Material for Chapter 3

Supporting Information

Acyl-chain saturation regulates the order of phosphatidylinositol 4,5-bisphosphate nanodomains

Luís Borges-Araújo [†,‡], Marco M. Domingues [§], Alexander Fedorov [†], Nuno C. Santos [§], Manuel N. Melo[‡], Fábio Fernandes [†,‖] *

† Institute for Bioengineering and Biosciences (IBB) and Associate Laboratory i4HB-Institute for Health and Bioeconomy, Instituto Superior Técnico, Universidade de Lisboa, Lisbon, Portugal

‡ Instituto de Tecnologia Química e Biológica António Xavier, Universidade Nova de Lisboa, Av. da República, 2780-157 Oeiras, Portugal.

§ Instituto de Medicina Molecular, Faculdade de Medicina, Universidade de Lisboa, Av. Prof. Egas Moniz, 1649-028 Lisbon, Portugal;

‖ Department of Bioengineering; Instituto Superior Técnico, Universidade de Lisboa, 1049-001 Lisbon, Portugal;

* Author to whom correspondence should be addressed; email: fernandesf@tecnico.ulisboa.pt

SUPPLEMENTARY METHODS

Improved Martini 2.2 coarse-grain parameterization of PI(4,5)P_2. With the goal of using the Martini 2.2 coarse-grained (CG) molecular dynamics (MD) forcefield for studies of PI(4,5)P_2 in lipid membranes, the parameters of the phosphatidylinositol (PI) headgroup and of PI(4,5)P_2 were improved from existing parameters[1]. A focus was made on more accurately describing the behavior seen in atomistic simulations and in improving the numerical stability of the topology. The existing PI/ PI(4,5)P_2 parameters had a dihedral potential in place that led to instability at typical Martini timesteps[2]; they also describe the conformational dynamics of the lipid headgroup with insufficient accuracy. The topology further showed an unintended propensity for aggregation, even in the absence of divalent cations. This is not observed experimentally, where only divalent cations have the affinity to induce this kind of lateral reorganization. Part of this behavior could be explained by an overly loose headgroup conformation.

PI was reparameterized in an aqueous solution and in a pure PI membrane as described in López et al. 2013[1]. PI(4,5)P_2 was only parameterized in aqueous solution. The same all-atom (AA) simulations that were used in López et al. 2013[1] were used as reference for the improved parameterization as well. The AA trajectories were converted into CG trajectories using the center of mass of the appropriate AA particles[1]. These backmapped trajectories were then used to obtain the several dihedral, angle and distance distributions, that were here considered as reference. Both the single lipid in water and the membrane systems were built and solvated as described in López et al. 2013[1]. All CG simulations were modelled using the Martini 2.2 CG model for biomolecular simulations[3] and run with GROMACS[4] version 2018. For the CG simulations, nonbonded interactions were cutoff at 1.1 nm and Coulombic interactions were treated using reaction-field electrostatics[5]. The particle neighbor list was updated using the Verlet list scheme. The equations of motion were integrated

using a timestep of 20 fs. The Stochastic Dynamics (SD) integrator was used as the thermostat, with an inverse friction constant of 2.0, maintaining the temperature at 310 K. Constant pressure was maintained at 1.0 bar by coupling to a Berendsen barostat with a relaxation time of 3.0 ps. For the simulations used to test the stability of the topology, the V-rescale thermostat was used, with a tau-t of 1.0, to maintain the temperature at 310 K. As was done in López et al. 2013[1], anisotropic scaling was performed for all bilayer systems, whilst for the aqueous systems isotropic scaling was used. The base topology for PI (DPPI) was obtained from the Martini coarse grain forcefield website[6].

To better match the PI conformations observed in the AA simulations and improve the stability of the topology at longer timesteps, the dihedral, angle and bond potentials were tuned. No bead types were changed from the original topology. The strategy employed was to replace the dihedral that was in place to control the headgroup orientation relative to the glycerol backbone (C3-C1-PO4-GL1) by more robust ones (i.e., less likely to operate over particle angles that can become collinear), and to adjust particle angles to improve overall stability. This was done by defining three dihedrals (C1-PO4-GL1-GL2, C2-C3-GL1-GL2 and C2-C3-C1-PO4), adding 3 new angles (GL2-GL1-C3, C2-C3-GL1 and PO4-GL1-GL2) and adjusting the angles that were already defined (C3-C1-PO4, C2-C1-PO4 and C1-PO4-GL1). For angles that were involved in the dihedrals, and that could reach 180°, a restricted bending potential was used, in place of the regular harmonic angle potential, to avoid numerical instabilities. This new approach better represented the distributions obtained from the backmapped AA simulations (figure S6) and yielded a topology stable in membrane simulations for at least 10μs at both 20 and 30 fs timesteps. The recovered values for the average area per lipid and average membrane thickness, at 310 K, were 0.598±0.004 nm^2 and 4.08±0.02 nm, respectively. These are in agreement with the values obtained by AA simulations, in López et al. 2013, of 0.65nm^2 and 3.9 nm, and with those obtained with the previous CG topology of 0.62±0.004 nm^2 and 4.0 nm.

Since PI(4,5)P_2 is a phosphorylated derivative of PI, the PI CG topology was used as a building block for the parameterization, to which the P4 and P5 phosphates were added. To connect these phosphates to the headgroup, we maintained the strategy that was employed by López et al. 2013[1] of using 2 bonds to inositol headgroup beads and 1 dihedral potential, to control the phosphate orientation relative to the ring plane. For each phosphate, the main bond was set at the appropriate AA distance and the secondary bond was increased slightly in order to control the phosphate angle relative to the inositol ring. This artificial increase in distance for the secondary bond is required due to the also increased distance between the inositol headgroup beads, which was put in place to improve area per lipid values. Contrary to what was done in López et al. 2013[1], harmonic bonds were used instead of constraints to obtain better conformation flexibility (these can eventually be set as constraints if necessary). Additionally, as the two phosphates are placed very close together, an exclusion between them was put in place. Only a slight adjustment to one of the dihedral potentials (C2-C3-C1-PO4) that controls the inositol headgroup (derived from the PI CG topology) was necessary. We were able to accurately represent the phosphate dynamics obtained from the backmapped AA simulations (figure S7). This topology was stable in membrane simulations for at least 50 μs using both 20 and 30 fs timesteps. The recovered values for the average area per lipid and average membrane thickness, at 310 K, were 0.715±0.007 nm^2 and 3.64±0.02 nm, respectively.

As mentioned above, with the original PI(4,5)P_2 topology we observed a propensity for unintended aggregation, especially through the interaction with sodium, which is not observed experimentally. With the improved PI(4,5)P_2 topology we can observe a significantly better propensity for aggregation (figure S8, A and B). This is clearly reflected in the PI(4,5)P_2 vs. PI(4,5)P_2 radial distribution function, where we observe a significant decrease in PI(4,5)P_2 neighbors with the improved topology (figure S8, C). It must be noted that the propensity for PI(4,5)P_2 to undergo

calcium induced clustering, was already present in the original topology and was not introduced by our modifications (figure S10).

Molecular dynamics simulation analysis. All simulations were analyzed making use of in-house developed Python3 programs using the MDAnalysis package. We also used the IPython[7], numpy[8], SciPy[9], scikit-learn[10] and matplotlib[11] packages for scientific computing in Python. Visualization and renderization of the simulations was performed with the molecular graphics viewer VMD[12]. The last 2 μs of each simulation were used for analysis.

— The *S-value order parameter* is a geometric parameter which can be used to compare the different lipid tails' overall order/flexibility. Lipid tail S-value order parameters were calculated for each lipid tail bond, of each independent lipid as described previously[13]. The S-value used here is the extension to coarse-grain of the concept of lipid C-H bond order parameter, and is defined as: $S = \frac{1}{2}(3\langle(\cos\theta)^2\rangle - 1)$, where θ is the angle between the vector along a particular lipid tail bead pair and the normal of the bilayer, which we approximate as the system's *z*-axis. When calculating the S-Value vs. the distance to a reference target, each S-Value obtained was then binned according to their distance to the reference in the *xy* plane. Bins with less than analyzed 200000 elements over the simulation time (2 μs) were removed from analysis.

—The *hexagonality index* was used in comparing the ability of the different lipid tails to organize into a hexagonal matrix. Hexagonality was calculated for the first lipid tail bead of each individual lipid. A lipid tail bead was considered to be within an hexagonal matrix if the following criteria were met: i) the tail bead is within 6 Å of the 6 closest neighbors; and ii) each of the 6 closest neighbors to the target bead also has at least 2 neighbors within that cutoff. To eliminate some thermal fluctuation noise the hexagonality index was calculated from a trajectory with coordinates averaged over a 5-frame running window.

—*Average area per lipid (APL)* was calculated from the average box area divided by the number of lipids in each leaflet.

—*Lipid lateral diffusion* was calculated from the mean square displacement (MSD) of the molecules in the membrane plane over a range of window sizes: 1, 3, 5, 10, 25, 50, 100, 200, 300, 400, 500 ns. Diffusion constants were then estimated from non-linear least squares fitting to the two parameter equation described previously[14,15]: $MSD = ND_\alpha t^\alpha$, where N represents the degrees of freedom coefficient, which in the case of planar phospholipid bilayers is 2, t represents time and D_α is measured in units of length^2/time$^\alpha$. Standard deviation for both parameters was taken from the square root of the diagonal of the covariance matrix obtained from the non-linear squares fit. Lateral diffusion was calculated for each lipid independently considering the position of its PO4 bead.

—*Radial distribution functions (RDF)* are obtained by processing the distance matrix from one reference set of positions to one or more sets of positions. The distance arrays are histogrammed and averaged over the simulation time. The RDF is then obtained by normalizing the histogram by the number of reference positions, volume of the bin and average particle density of the positions. In our cases the RDFs are calculated over the *xy* membrane plane and thus the *z* component of the positions is not taken into consideration. Additionally, normalization employs the area of the bin instead of the volume.

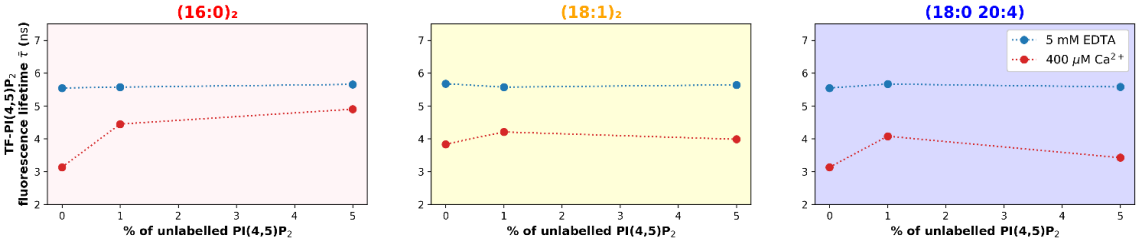
—*Membrane curvature analysis* was performed using the PyCurv module and adapting the methods described elsewhere[16]. In summary, the PO4 beads from one of the bilayer leaflets were selected and

used to generate a triangulated surface. To correctly calculate the curvature near the periodic boundaries, a neighboring periodic copy of the membrane was added in each of the *xy* directions before surface triangulation. This surface was then processed, and a surface graph generated from which the local surface normal, principal directions and curvatures were estimated. Principal directions and curvatures were estimated using the vector voting (VV) method using a neighborhood of 5 nm. This neighborhood radius was calibrated to better show the features of interest. In general, the neighborhood radius should be set to the radius of the smallest feature of interest on the input surface.

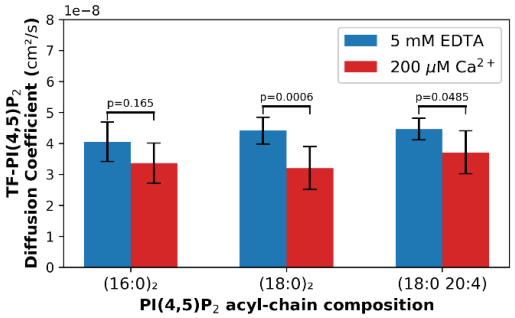
Fluorescence correlation spectroscopy. FFS measurements were performed on a Leica TCS SP5 (Leica Microsystems CMS GmbH, Manheim, Germany) inverted confocal microscope (DMI600). Excitation lines provided by an argon laser were focused into the sample through an apochromatic water immersion objective (63×, NA 1.2; Zeiss, Jena Germany). A 111.4 μm diameter pinhole in front of the image plane blocked out of focus signal. Fluorescence emission was detected using avalanche photodiodes (APDs) after passing through a 500–550 nm band-pass filter. Excitation of TF-PI(4,5)P2 was performed with the 488 nm Ar laser line. Fluorescence fluctuations from GUVs with incorporated fluorescent probes were recorded from the top of the vesicle, with the focal volume centered in the focal plane with maximum fluorescence intensity. Five autocorrelation (AC) curves, G(τ), were sequentially obtained for each sample, with an acquisition time of 20 s at a 100 kHz sampling frequency. Assuming a Gaussian measurement volume, AC curves were fitted using a model accounting for 2D translational diffusion, and [17]:

$$G(\tau) = \frac{1}{C\pi\omega_{xy}^2}\left(1+\frac{4D\tau}{\omega_{xy}^2}\right)^{-1} = \frac{1}{N}\left(1+\frac{\tau}{\tau_D}\right)^{-1}$$

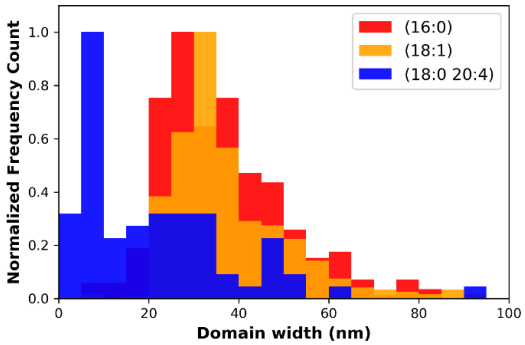
The dimensions of the focal volume were determined by calibration with rhodamine 110 in milliQ water using the same optical setup as the samples. A rhodamine 110 diffusion coefficient (D) of 440 $\mu m^2 s^{-1}$ was considered[18]. Analysis of all the FCS experimental data was carried out using the ISS Vista software. This program employs a Levenberg–Marquardt nonlinear least-squares fitting routine and the goodness of the fittings can be judged by the recovered χ^2 value and the random distribution of the weighted residuals.



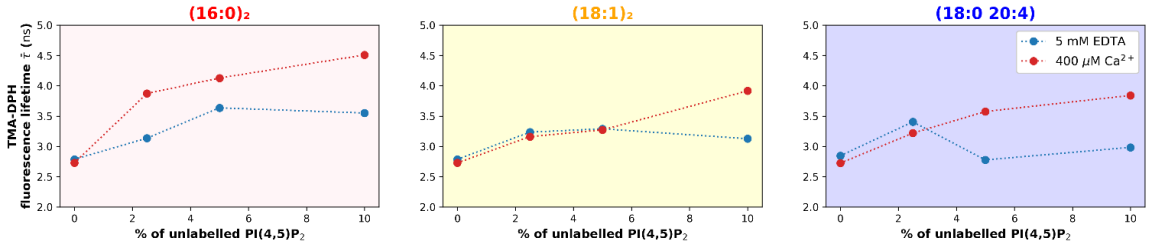
Supplementary Figure 1. Cation-induced clusters are formed independently of PI(4,5)P_2 acyl-chain composition as seen by homoFRET of the TF- PI(4,5)P_2 analogue. PI(4,5)P_2 cluster formation was determined through the incorporation of 0.1% of TF- PI(4,5)P_2 in 100 nm LUVs containing POPC and increasing concentrations of unlabeled PI(4,5)P_2. The experiments were carried out for the three PI(4,5)P_2 species in study and figures are labeled according to the acyl-chain composition of the unlabeled PI(4,5)P_2 employed. TF-PI(4,5)P_2 fluorescence intensity weighed lifetime ($\bar{\tau}$) was measured in the presence (400 μM Ca^{2+}, red) and absence (5 mM EDTA, blue) of calcium. It should be noted, that since fluorescence anisotropies of fluorophores are inversely dependent on their respective fluorescence lifetimes[17], that the decrease in lifetime observed during calcium dependent clustering of the analogue, leads to a slightly attenuated drop in <*r*> values.



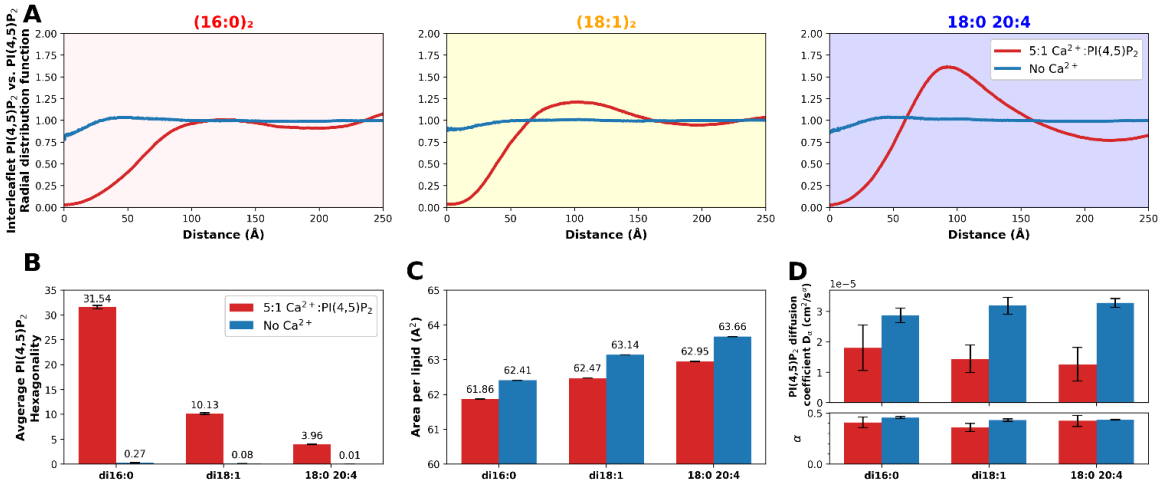
Supplementary Figure 2. TF-PI(4,5)P_2 is incorporated into cation-induced PI(4,5)P_2 clusters regardless of acyl-chain composition. Diffusion coefficients obtained for 0.01% TF- PI(4,5)P_2 in 97.5:2.5 POPC:PI(4,5)P_2 GUVs, in the presence and absence of calcium. The experiments were done for the three PI(4,5)P_2 species in study and figures are labeled according to the acyl-chain composition of the employed unlabeled PI(4,5)P_2. Error bars represent the standard deviation from measurements obtained from at least N=4 independent GUVs.



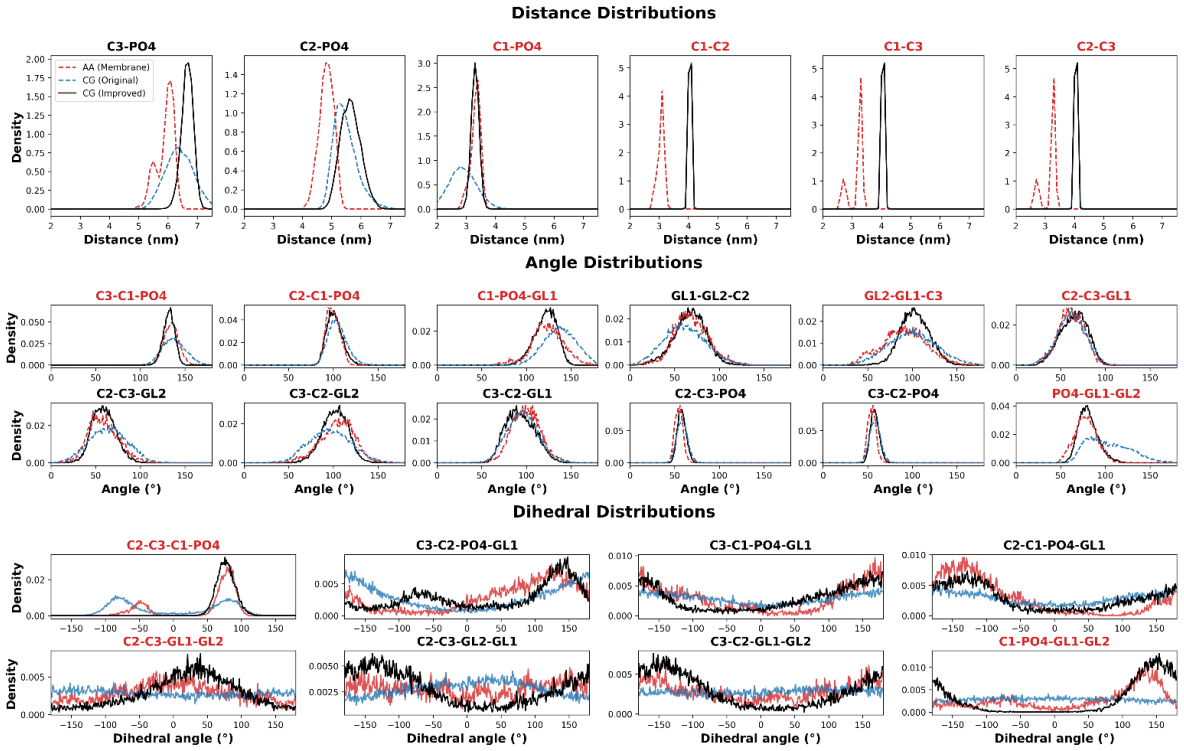
Supplementary Figure 3. PI(4,5)P_2 unsaturated acyl-chain compositions appear to promote the formation of smaller nanodomains as seen by AFM. PI(4,5)P_2 cluster formation was detected through the AFM measurement of SLBs containing DOPC and 5% PI(4,5)P_2 . The experiments were carried out for the three PI(4,5)P_2 species in study and the histograms are labeled according to the acyl-chain composition of the PI(4,5)P_2 species employed. Topographical images were acquired and analyzed with first or second level flattening, using the JPK data processing software, from which the domain width was determined. The sizes of the domains were evaluated by several cross sections on the topographical images. The number of cross sections varied from 70 to approximately 300 to obtain representative data of the domains formed in each SLB. The frequency count was normalized to the highest bin.



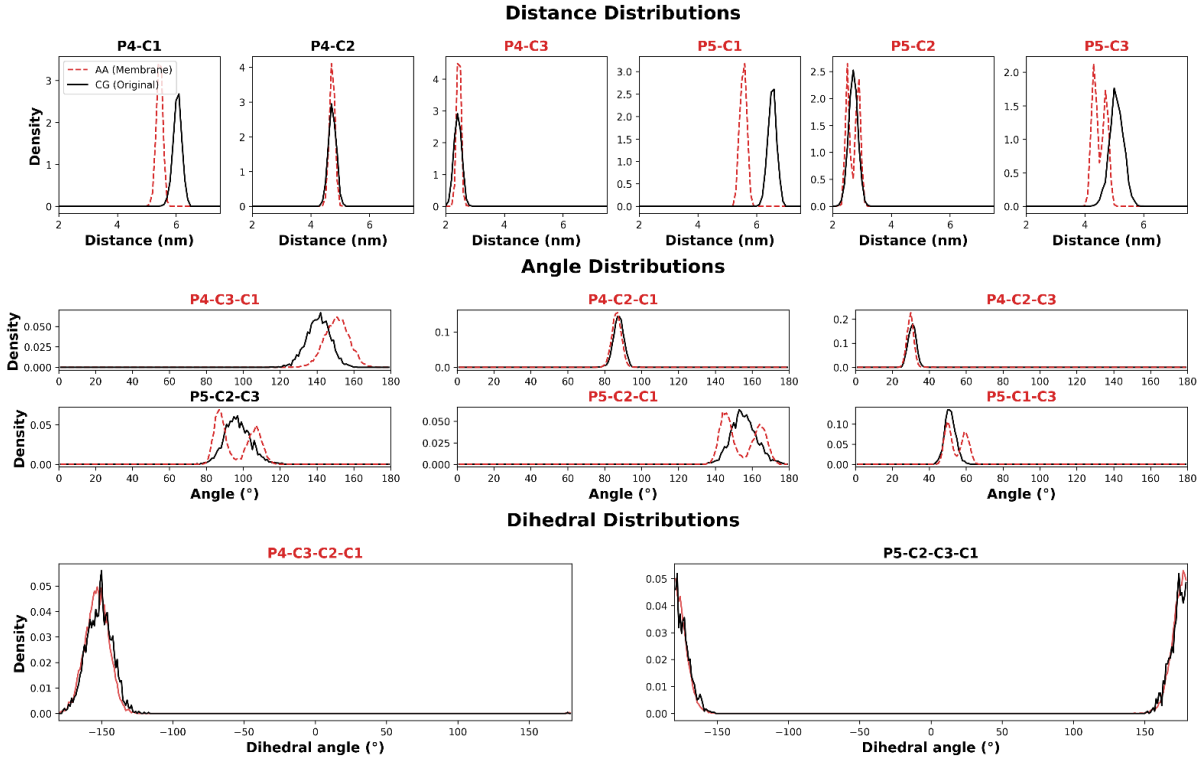
Supplementary Figure 4. TMA-DPH probe incorporation shows that calcium-induced PI(4,5)P_2 nanodomains are significantly more ordered than monodisperse PI(4,5)P_2 even for unsaturated acyl-chain compositions. PI(4,5)P_2 local membrane order was determined through the incorporation of TMA-DPH at a 1:300 lipid ratio in MLVs containing POPC and increasing concentrations of unlabeled PI(4,5)P_2. The experiments were done for the three acyl-chain compositions in study. TMA-DPH fluorescence intensity weighed lifetime was measured in the presence (400 µM Ca2+, red) and absence (5 mM EDTA, blue) of calcium.



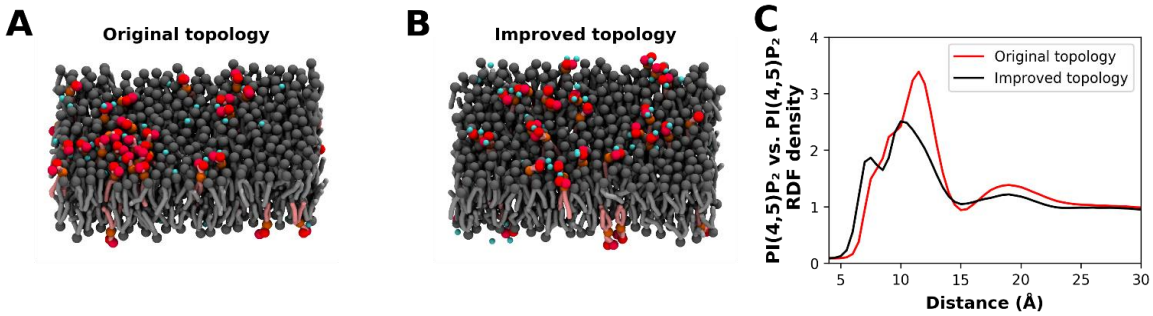
Supplementary Figure 5. CG MD simulations showcase the impact of acyl-chain composition on PI(4,5)P_2 and PI(4,5)P_2 nanodomain biophysical properties. (A) Interleaflet PI(4,5)P_2 vs PI(4,5)P_2 radial distribution function for each composition in the presence and absence of calcium. (B) Average PI(4,5)P_2 acyl-chain hexagonality for each composition in the presence and absence of calcium. (C) Average membrane area per lipid for each system in the presence and absence of calcium. (D) PI(4,5)P_2 diffusion coefficient and scaling exponent recovered from an anomalous diffusion analysis of the simulation, in the presence and absence of calcium.



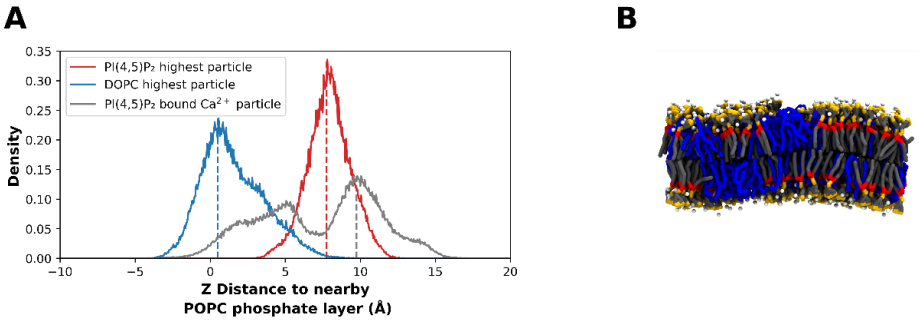
Supplementary Figure 6. Bonded parameters from the original and improved PI topologies. Tracked distance, angle and dihedral distributions from Martini 2.2 PI models as well as the reference atomistic distributions. These were obtained from mapping AA simulations (red line) and compared to the original (blue line) and the improved (black line) CG topologies. The distributions that have an applied potential in the improved topology are marked in red.



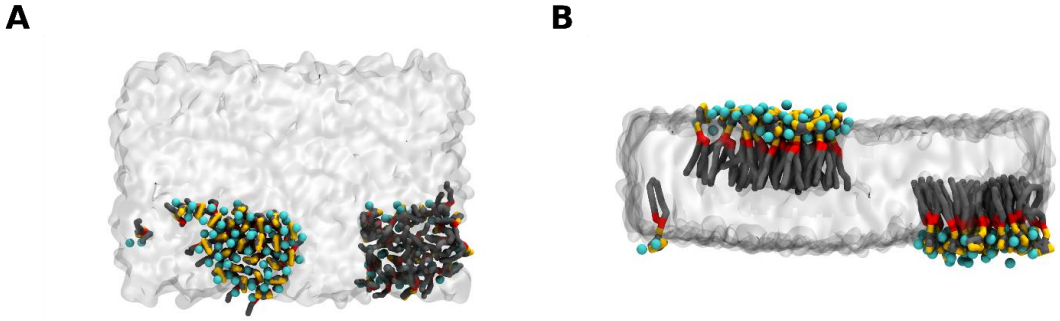
Supplementary Figure 7. Phosphate bonded parameters from the original and improved PI(4,5)P_2 topologies. Tracked distance, angle and dihedral distributions of the phosphate particles from Martini 2.2 PI(4,5)P_2 models as well as the reference atomistic distributions. These were obtained from mapping AA simulations (red line) and compared to the original (blue line) and the improved (black line) CG topologies. The distributions that have an applied potential in the improved topology are marked in red.



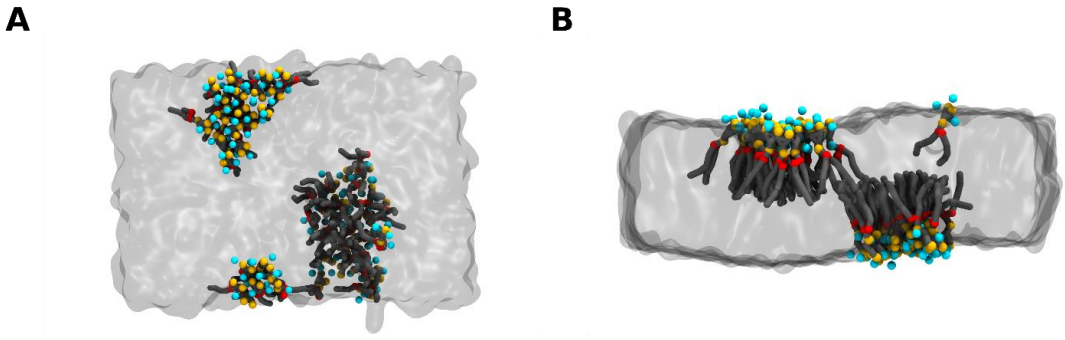
Supplementary Figure 8. Influence of the topology on PI(4,5)P_2 non-specific aggregation. Final simulation snapshots of a membrane simulation containing a 90:10 POPC:PI(4,5)P_2 lipid mixture in the presence of 140 mM NaCl. This simulation was performed with the original (A) and the updated (B) headgroup parameters and ran for at least 40 µs. POPC is depicted in grey, PI(4,5)P_2 in red and sodium atoms within 7 angstroms of PI(4,5)P_2 in blue. Radial distribution functions (RDFs) of the PI(4,5)P_2 headgroup center of mass against itself (C), calculated from the same membrane simulations. The result obtained by the original and improved parameters are depicted in red and black, respectively.



Supplementary Figure 9. Impact of $(16:0)_2$ PI(4,5)P_2 on bilayer thickness in fluid CG MD simulations. Density profile (A) and final simulation snapshot (B), of a membrane simulation containing a 50:50 DOPC: $(16:0)_2$ PI(4,5)P_2 lipid mixture in the presence of calcium, showcasing the difference in height between $(16:0)_2$ PI(4,5)P_2 nanodomains and the bulk membrane.



Supplementary Figure 10. Calcium-induced aggregation was already observed for the original PI(4,5)P_2 Martini 2 CG MD topology. Final simulation snapshots of a membrane simulation containing a 90:10 POPC: $(16:0)_2$ PI(4,5)P_2 lipid mixture in the presence of calcium. This simulation was performed with the original headgroup parameters and ran for 4 µs. PI(4,5)P_2 lipid headgroups and acyl-chains are depicted in grey, with the phosphates discriminated in orange and the glycerol backbone in red. Ca^{2+} ions are represented in blue. The bulk POPC lipids are represented by the translucent grey surface.



Supplementary Figure 11. Calcium-induced PI(4,5)P_2 aggregation is observed with particle mesh ewald (PME) electrostatic treatment. Final simulation snapshots of a membrane simulation containing a 90:10 POPC: $(16:0)_2$ PI(4,5)P_2 lipid mixture in the presence of calcium. This simulation was performed with particle mesh ewald electrostatics, the updated PI(4,5)P_2 headgroup parameters and ran for 4 µs. PI(4,5)P_2 lipid headgroups and acyl-chains are depicted in grey, with the phosphates discriminated in orange and the glycerol backbone in red. Ca^{2+} ions are represented in blue. The bulk POPC lipids are represented by the translucent grey surface.

SUPPLEMENTARY REFERENCES

(1) López, C. A.; Sovova, Z.; van Eerden, F. J.; de Vries, A. H.; Marrink, S. J. Martini Force Field Parameters for Glycolipids. *J. Chem. Theory Comput.* **2013**, *9* (3), 1694–1708.

(2) Bulacu, M.; Goga, N.; Zhao, W.; Rossi, G.; Monticelli, L.; Periole, X.; Tieleman, D. P.; Marrink, S. J. Improved Angle Potentials for Coarse-Grained Molecular Dynamics Simulations. *J. Chem. Theory Comput.* **2013**, *9* (8), 3282–3292.

(3) Marrink, S. J.; Risselada, H. J.; Yefimov, S.; Tieleman, D. P.; De Vries, A. H. The MARTINI Force Field: Coarse Grained Model for Biomolecular Simulations. *J. Phys. Chem. B* **2007**, *111* (27), 7812–7824.

(4) 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.

(5) De Jong, D. H.; Baoukina, S.; Ingólfsson, H. I.; Marrink, S. J. Martini Straight: Boosting Performance Using a Shorter Cutoff and GPUs. *Comput. Phys. Commun.* **2016**, *199*, 1–7.

(6) Martini Coarse Grain Forcefield for Biomolecules http://cgmartini.nl/.

(7) Pérez, F.; Granger, B. E. *IPython: A System for Interactive Scientific Computing*.

(8) Harris, C. R.; Millman, K. J.; van der Walt, S. J.; Gommers, R.; Virtanen, P.; Cournapeau, D.; Wieser, E.; Taylor, J.; Berg, S.; Smith, N. J.; Kern, R.; Picus, M.; Hoyer, S.; van Kerkwijk, M. H.; Brett, M.; Haldane, A.; del Río, J. F.; Wiebe, M.; Peterson, P.; Gérard-Marchant, P.; Sheppard, K.; Reddy, T.; Weckesser, W.; Abbasi, H.; Gohlke, C.; Oliphant, T. E. Array Programming with NumPy. *Nature* **2020**, *585* (7825), 357–362.

(9) Virtanen, P.; Gommers, R.; Oliphant, T. E.; Haberland, M.; Reddy, T.; Cournapeau, D.; Burovski, E.; Peterson, P.; Weckesser, W.; Bright, J.; van der Walt, S. J.; Brett, M.; Wilson, J.; Millman, K. J.; Mayorov, N.; Nelson, A. R. J.; Jones, E.; Kern, R.; Larson, E.; Carey, C. J.; Polat, İ.; Feng, Y.; Moore, E. W.; VanderPlas, J.; Laxalde, D.; Perktold, J.; Cimrman, R.; Henriksen, I.; Quintero, E. A.; Harris, C. R.; Archibald, A. M.; Ribeiro, A. H.; Pedregosa, F.; van Mulbregt, P. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nat. Methods* **2020**, *17* (3), 261–272.

(10) Fabian, P.; Michel, V.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Vanderplas, J.; Cournapeau, D.; Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Thirion, B.; Grisel, O.; Dubourg, V.; Passos, A.; Brucher, M.; Perrot, M.; Duchesnay, É. *Scikit-Learn: Machine Learning in Python*; 2011; Vol. 12.

(11) Hunter, J. D. Matplotlib: A 2D Graphics Environment. *Comput. Sci. Eng.* **2007**, *9* (3), 90–95.

(12) Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual Molecular Dynamics. *J. Mol. Graph.* **1996**, *14* (1), 33–38.

(13) Ingólfsson, H. I.; Melo, M. N.; Van Eerden, F. J.; Arnarez, C.; Lopez, C. A.; Wassenaar, T. A.; Periole, X.; De Vries, A. H.; Tieleman, D. P.; Marrink, S. J. Lipid Organization of the Plasma Membrane. *J. Am. Chem. Soc.* **2014**, *136* (41), 14554–14559.

(14) Goose, J. E.; Sansom, M. S. P. Reduced Lateral Mobility of Lipids and Proteins in Crowded Membranes. *PLoS Comput. Biol.* **2013**, *9* (4), e1003033.

(15) Reddy, T.; Shorthouse, D.; Parton, D. L.; Jefferys, E.; Fowler, P. W.; Chavent, M.; Baaden, M.; Sansom, M. S. P. Nothing to Sneeze At: A Dynamic and Integrative Computational Model of an Influenza A Virion. *Structure* **2015**, *23* (3), 584–597.

(16) Salfer, M.; Collado, J. F.; Baumeister, W.; FernándezBusnadiego, R.; Martínez-Sánchez, A. Reliable Estimation of Membrane Curvature for Cryo-Electron Tomography. *PLoS Comput. Biol.* **2020**, *16* (8), e1007962.

(17) Lakowicz, J. R. *Principles of Fluorescence Spectroscopy*; Springer, 2006.

(18) Gendron, P.-O.; Avaltroni, F.; Wilkinson, K. J. Diffusion Coefficients of Several Rhodamine Derivatives as Determined by Pulsed Field Gradient–Nuclear Magnetic Resonance and Fluorescence Correlation Spectroscopy. *J. Fluoresc.* **2008**, *18* (6), 1093–1101.

C | Supplementary Material for Chapter 4

Supporting Information

Improved parameterization of phosphatidylinositide lipid headgroups for the Martini 3 coarse grain force field

L. Borges-Araújo [†,‖], P. C. T. Souza [§], F. Fernandes [†,‡], M. N. Melo[‖,*]

† iBB-Institute for Bioengineering and Biosciences, Instituto Superior Técnico, Universidade de Lisboa, Lisbon, Portugal

‡ Department of Bioengineering; Instituto Superior Técnico, Universidade de Lisboa, 1049-001 Lisbon, Portugal;

§ Molecular Microbiology and Structural Biochemistry, UMR 5086 CNRS & University of Lyon, 7 Passage du Vercors, F-69367, Lyon, France;

‖ Instituto de Tecnologia Química e Biológica António Xavier, Universidade Nova de Lisboa, Av. da República, 2780-157 Oeiras, Portugal.

* Author to whom correspondence should be addressed; email: m.n.melo@itqb.unl.pt

Supplementary Table of Contents

SUPPLEMENTARY METHODS

Molecular dynamics simulation analysis. All simulations were analyzed making use of in-house developed Python3 programs using the MDAnalysis package. We also used the IPython[1], numpy[2], SciPy[3], scikit-learn[4] and matplotlib[5] packages for scientific computing in Python. Visualization and rendering of the simulations was performed with the molecular graphics viewer VMD[6]. The last 2 μs of each CG simulation and 1 μs of each AA simulation were used for analysis.

—*Area per lipid (ApL)* was calculated based on Voronoi tessellation of the membrane surface, as described elsewhere[7]. 2D Voronoi tessellation of the membrane surface was carried out, considering periodic boundary conditions, and using the first bead of each acyl chain as the cell centers. For the AA simulations, the center of geometry of the atoms corresponding to the first CG beads of each acyl chain were used. The Voro++ open source library[8] was used to this end, with modified bindings for python and for generation of 2D tesserae. The area per lipid was obtained by summing the Voronoi cell areas corresponding to the two lipid acyl chains and then averaging over lipid type and simulation time. *System-wide ApL* was calculated from the average box area divided by the number of lipids in each leaflet.

—*Radial distribution functions (RDF)* were obtained in 3 dimensions, from either the centers of geometry (c.o.m.) of the inositol sugar molecules — when calibrating the inositol model — or from the the minimum distance between each lipid headgroup pair — when analyzing lipid–lipid contacts. In the latter case, a 3D RDF was taken even though phosphoinositides are dispersed in the approximately 2D membrane matrix, because the objective was to discriminate short-range interaction preferences that could otherwise be masked by a projection of phosphoinositide ring structures onto the membrane plane. In any case, our parameterization analysis is mostly robust to the exact choice of method since i) RDFs are only employed semi-quantitatively to compare behavior to AA references, and ii) those references' RDFs are obtained in the same conditions.

—*Octanol–water partition free-energies* were calculated from the individual CG solvation free-energies into each solvent, as described elsewhere[9]. Interaction parameters and pressure coupling were kept as for normal CG runs. A stochastic dynamics scheme[10] was employed at a time step of 20 fs and a friction constant of 1 ps^{-1}, coupling the temperature to 298 K. Pressure was isotopically coupled to 1 bar using the Parrinello-Rahman barostat[11]. Single solute molecules were decoupled from solvent boxes of either 898 Martini waters or of 431 octanol molecules hydrated with 38 Martini waters. Decoupling was performed in 11 steps of 25 ns each by scaling down solute–solvent Lennard-Jones interactions (Martini water and octanol are not explicitly charged, and therefore do not establish Coulombic interactions with solutes). Partition free-energies and associated error estimates were obtained by the multistate Bennett acceptance ratio (MBAR) method[12], using the pymbar v3.03 Python package[13].

—*Solvent-accessible surface area (SASA) and Connolly surfaces* were determined using the GROMACS[14] tool *gmx sasa*. For atomistic references the Van der Waals (VdW) radii from the work of Rowland and Taylor[15] were used, whilst for Martini molecules, the VdW radii for R-, S-, and T- beads are 0.264, 0.230, and 0.191 nm, respectively[16].

— *Membrane z density profiles* were calculated by determining the target atom's distance to the nearby POPC phosphate layer and histogramming the *z* distance component. The nearby POPC phosphate layer was defined by any POPC phosphate within 15 Å of a given phosphoinositide,

updated at each timestep (in AA simulations these were taken to be the POPC phosphorus atoms, and in CG simulations the POPC PO4 beads).

— *Membrane thickness* was determined by calculating the distance between the centers of geometry of each leaflet's phosphate layer. Membrane thickness was defined as the z component of this distance.

— The *order parameter, S,* is a geometric parameter that expresses lipid tail order as a measure of average alignment with the membrane normal. The order parameter used here is the extension to coarse-grain of the concept of lipid C-H bond order parameter and is defined as: $S = \frac{1}{2}(3\langle(\cos\theta)^2\rangle - 1)$, where θ is the angle between the vector along a particular lipid tail bead pair and the normal of the bilayer, which we approximate as the system's z-axis. Lipid tail order parameters were calculated for each lipid tail bond, of each independent lipid as described previously[7].

— The *second virial coefficient,* B_{22}, describes the deviation from ideal behavior of a solution with a certain concentration of solute. In our case, it can be used to estimate the self-interaction of solute molecules. B_{22} can be derived from MD simulations, based on the cumulative solute-solute radial distribution function, as described elsewhere[17]. It is defined as: $B_{22}(r') = N_A\,(\frac{2\pi}{3}r'^3 - \frac{N(r')}{2\rho})$, where N_A is Avogadro's constant, r' is the distance and $N(r')$ is the number increment at distance r'. To be of use, this method requires sufficient solute molecules over the separation distance of interest, a prerequisite which is only fulfilled when the free energy landscape does not contain wells so deep that the Boltzmann-distributed solute molecules are effectively depleted from other regions. Others have found this limit to be in the order of -1 RT (-2.5 kJ/mol)[17]. The RDFs were obtained from simulations of 100 nM of myo-inositol in a solvated cubic box of volume ~19^3 nm^3. These trajectories were split in 200 ns segments and for every segment the cumulative inositol-inositol c.o.m. RDF was computed. The resulting curves were averaged to yield the final RDF and the standard error.

—*Lipid lateral diffusion* was calculated from the mean square displacement (MSD) of the molecules in the membrane plane over a range of window sizes: 1, 3, 5, 10, 25, 50, 100, 200, 300, 400, 500 ns. Diffusion constants were then estimated from non-linear least squares fitting to the two parameter equation described previously[18,19]: $MSD = ND_\alpha t^\alpha$, where N represents the degrees of freedom coefficient, which in the case of planar phospholipid bilayers is 2, t represents time and D_α is measured in units of length^2/time$^\alpha$. Standard deviation for both parameters was taken from the square root of the diagonal of the covariance matrix obtained from the non-linear squares fit. Lateral diffusion was calculated for each lipid independently considering the position of its PO4 bead.

Table S2. CHARMM36 Atomistic membrane reference systems.

	Membrane constitution				
Parameterization target	**Lipid**	**CHARMM Topology**	**mol%**	**Simulation time (μs)**	**Protonation**
PI	POPC	POPC	95.00	2.479	
	PO-PI	POPI	5.00		
PI(3)P	POPC	POPC	95.00	2.558	3-Phosphate fully deprotonated
	PO-PI(3)P	POPI13	5.00		
PI(4)P	POPC	POPC	95.00	2.5	4-Phosphate fully deprotonated
	PO-PI(4)P	POPI14	5.00		
PI(5)P	POPC	POPC	95.00	2.538	5-Phosphate fully deprotonated
	PO-PI(5)P	POPI15	5.00		
$PI(3,4)P_2$	POPC	POPC	95.00	2.5245	Half the PI lipids are protonated on the 3-phosphate and the other half on the 4-phosphate.
	$PO\text{-}PI(3,4)P_2$	POPI2C	2.50		
	$PO\text{-}PI(3,4)P_2$	POPI2D	2.50		
$PI(3,5)P_2$	POPC	POPC	95.00	2.197	Half the PI lipids are protonated on the 3-phosphate and the other half on the 5-phosphate.
	$PO\text{-}PI(3,5)P_2$	POPI2A	2.50		
	$PO\text{-}PI(3,5)P_2$	POPI2B	2.50		
$PI(4,5)P_2$	POPC	POPC	95.00	2.252	Half the PI lipids are protonated on the 4-phosphate and the other half on the 5-phosphate.
	$PO\text{-}PI(4,5)P_2$	POPI24	2.50		
	$PO\text{-}PI(4,5)P_2$	POPI25	2.50		
$PI(3,4,5)P_2$	POPC	POPC	94.00	2.524	PI lipids are mono-protonated on one the phosphate groups whilst being fully deprotonated on the two remaining.
	$PO\text{-}PI(3,4,5)P_3$	POPI33	2.00		
	$PO\text{-}PI(3,4,5)P_3$	POPI34	2.00		
	$PO\text{-}PI(3,4,5)P_3$	POPI35	2.00		

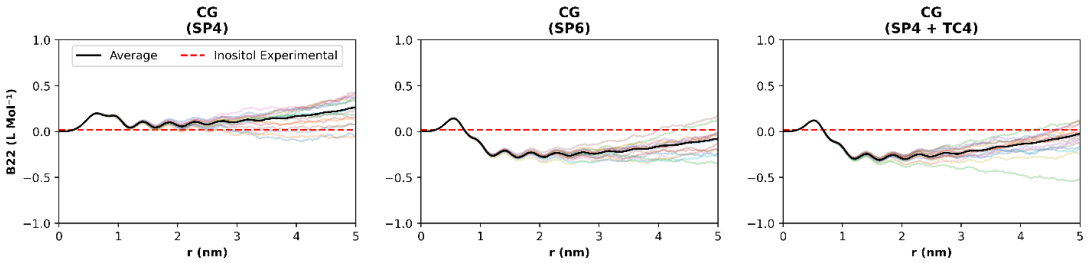


Figure S3. B_{22} vs r profiles of all 200 ns segments for the different CG inositol models. The colored lines represent the several 200 ns segments while the full black line represents the average. The experimental value for myo-inositol is represented by the dashed red line.

Bonded parameter calibration. Figures S1 through S9 below track distance, angle and dihedral distributions of *myo*-inositol and the eight phosphoinositides modeled for Martini 3. They are compared to reference atomistic distributions, both in water and (with the exception of inositol) membrane systems (with composition 90:10 POPC:lipid-of-interest). Reference AA distributions, in dashed lines, were obtained from CG-mapping AA simulation trajectories. CG distributions that have an applied potential in the CG topologies have their title in red. All simulations were run at 310 K.

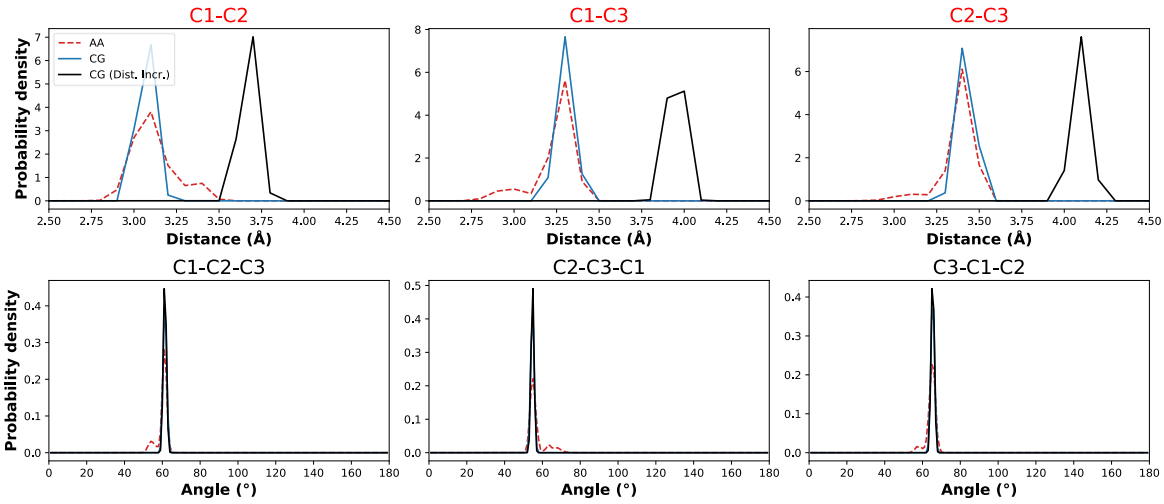


Figure S4. *Myo*-inositol CG bonded parameters. A second CG distribution, in black, represents the 20% increased distances employed in the final inositol model.

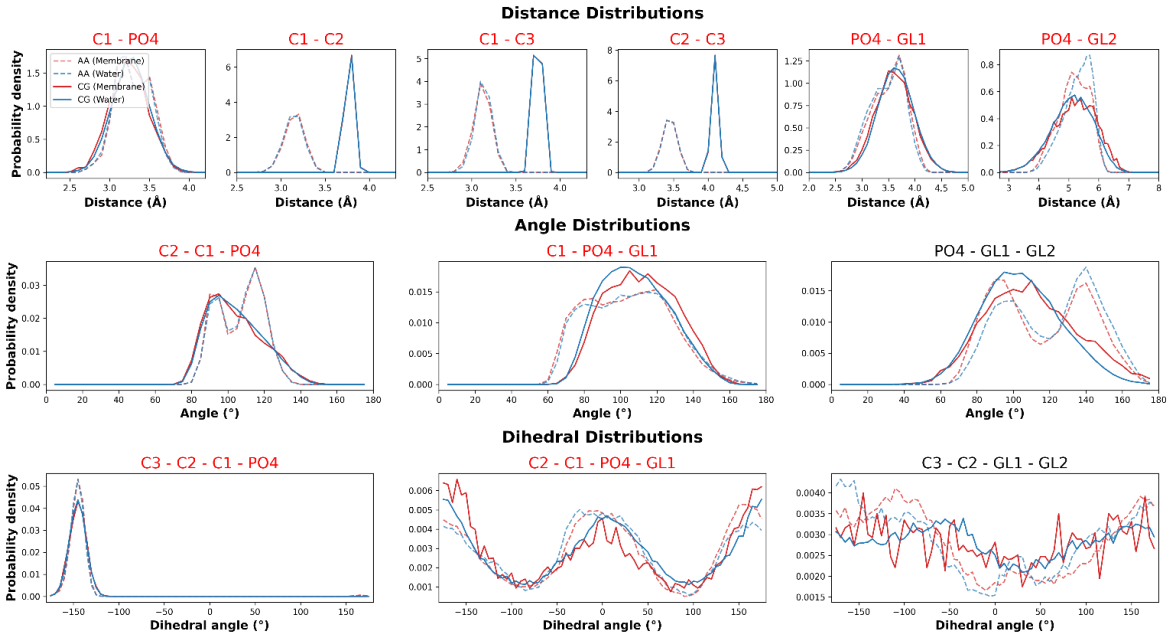


Figure S3. Phosphatidylinositol CG bonded parameters.

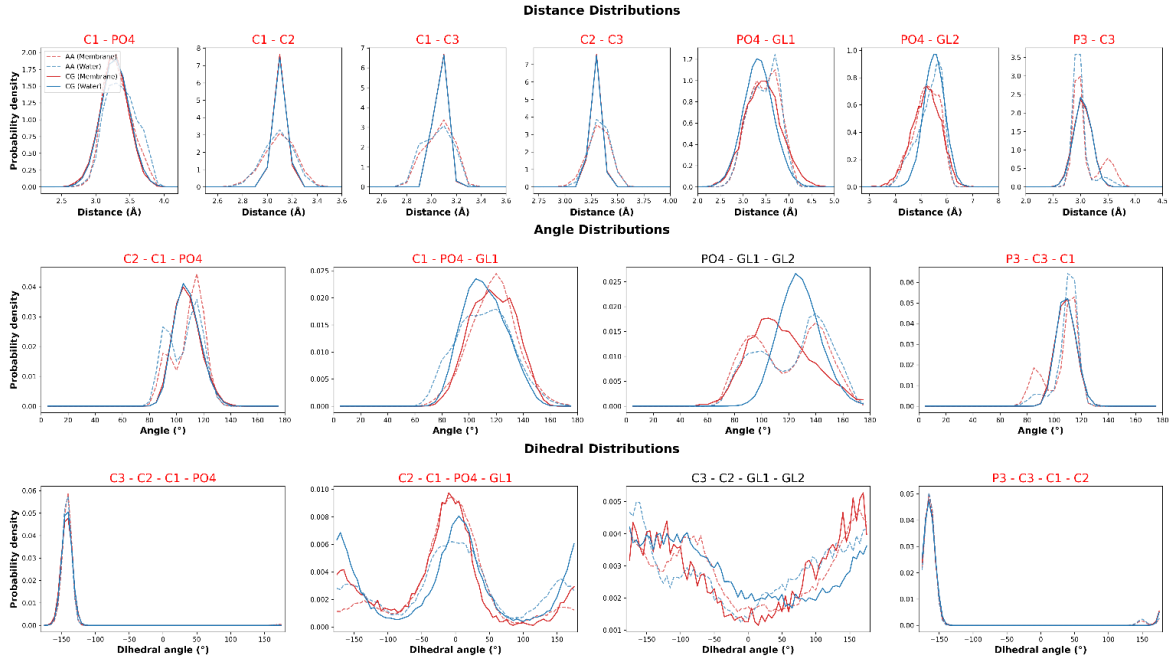


Figure S4. phosphatidylinositol 3-phosphate CG bonded parameters.

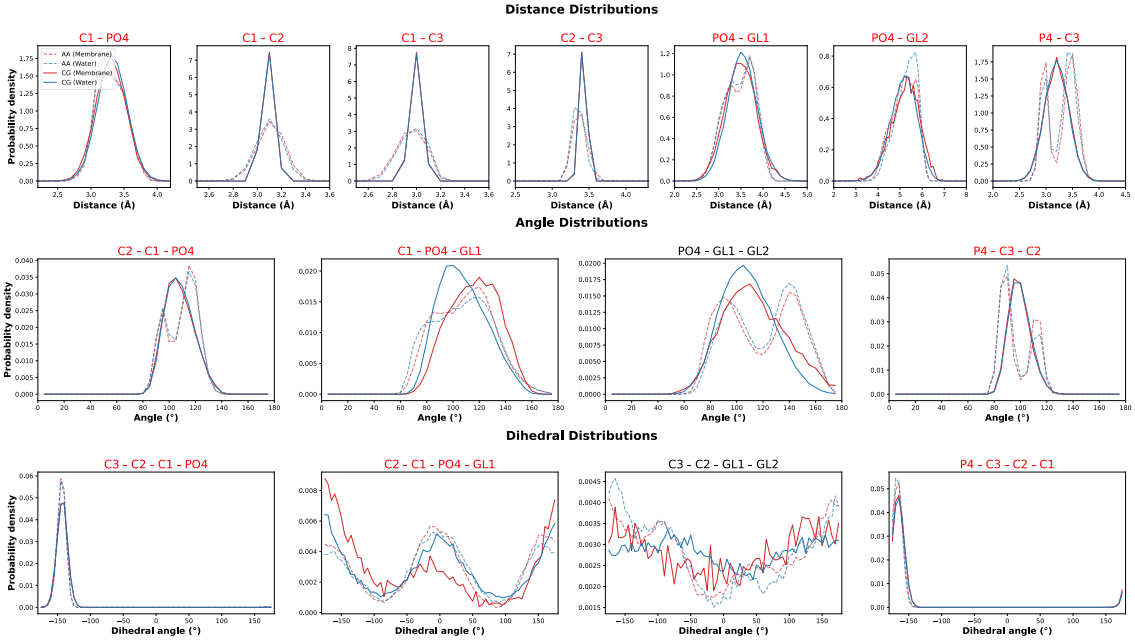


Figure S5. phosphatidylinositol 4-phosphate CG bonded parameters.

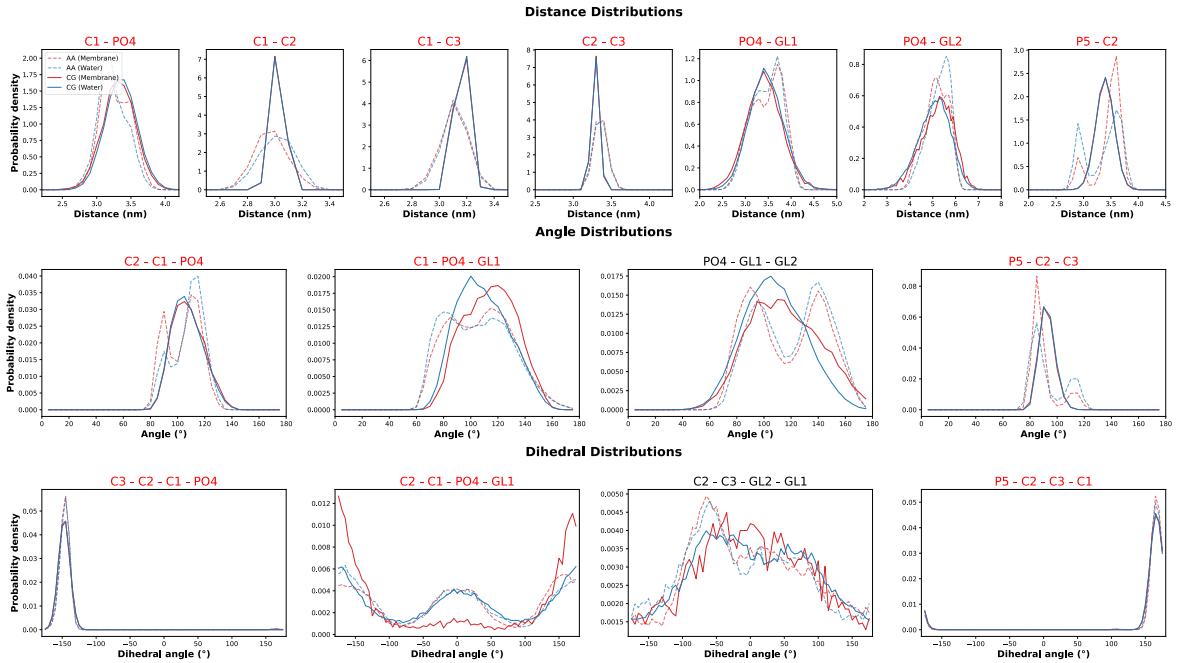


Figure S6. phosphatidylinositol 5-phosphate CG bonded parameters.

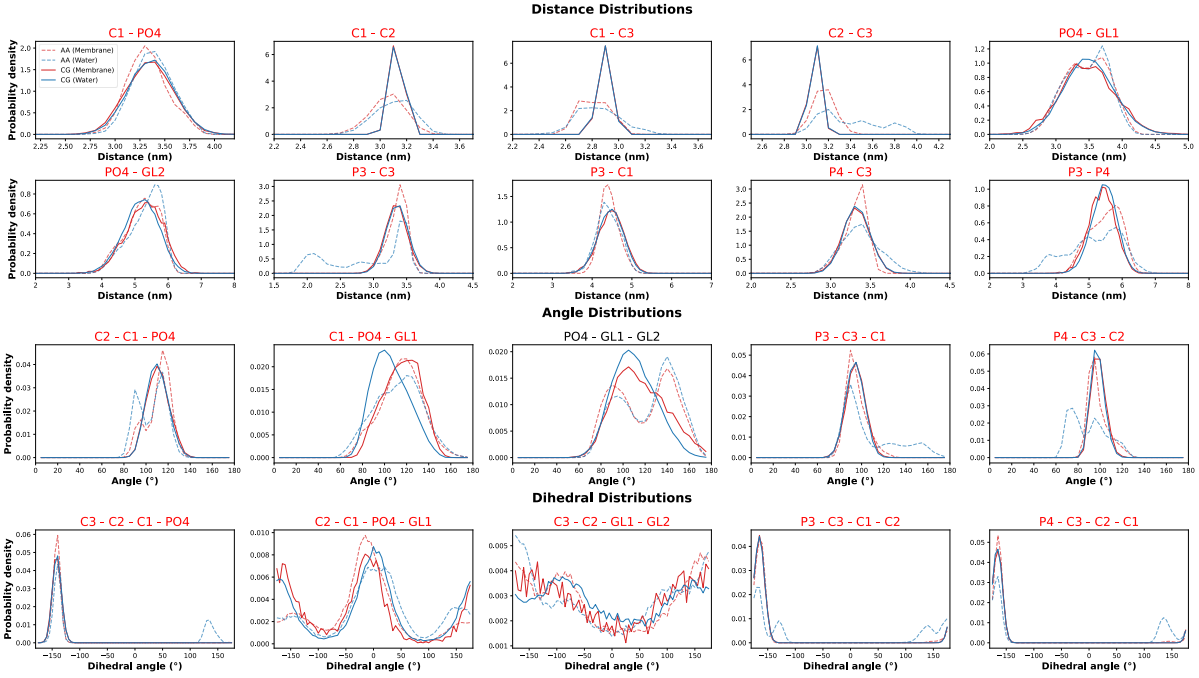


Figure S7. phosphatidylinositol 3,4-bisphosphate CG bonded parameters.

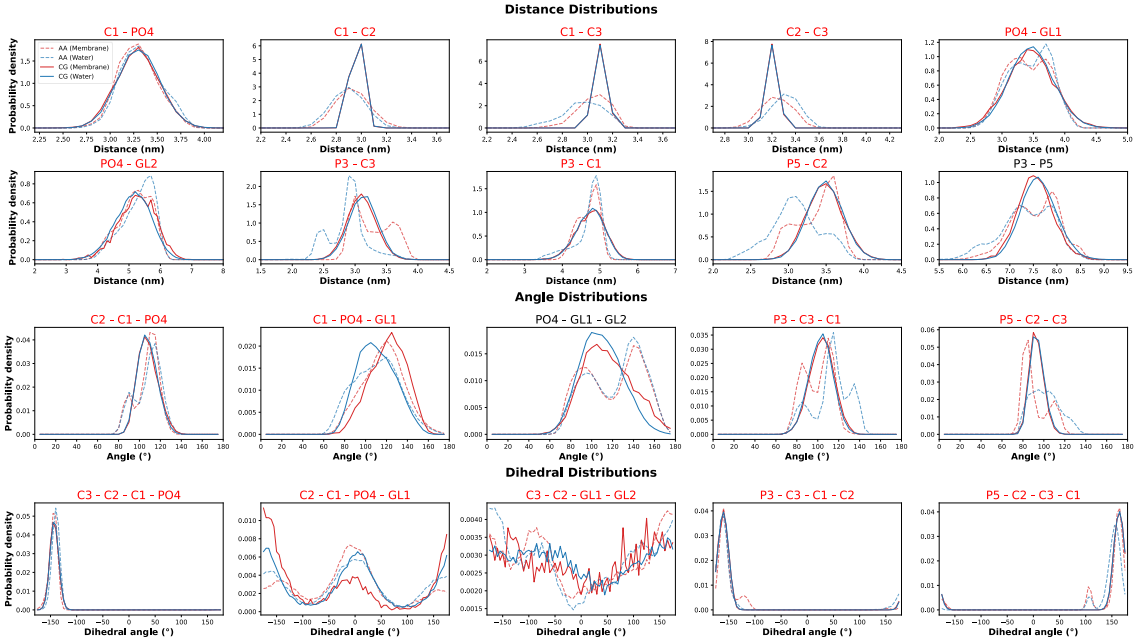


Figure S8. phosphatidylinositol 3,5-bisphosphate CG bonded parameters.

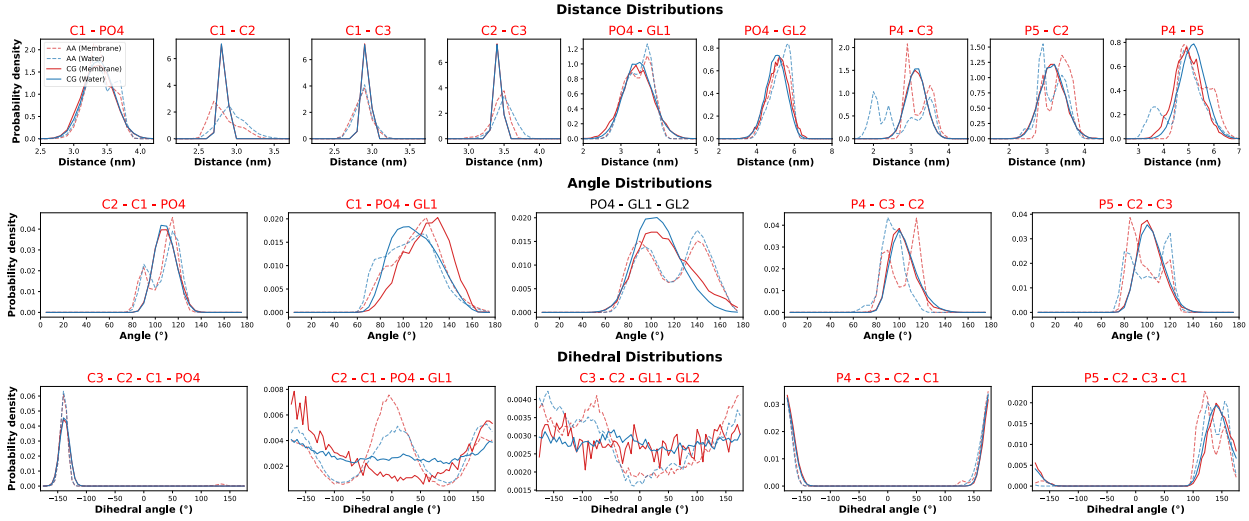


Figure S9. phosphatidylinositol 4,5-bisphosphate CG bonded parameters.

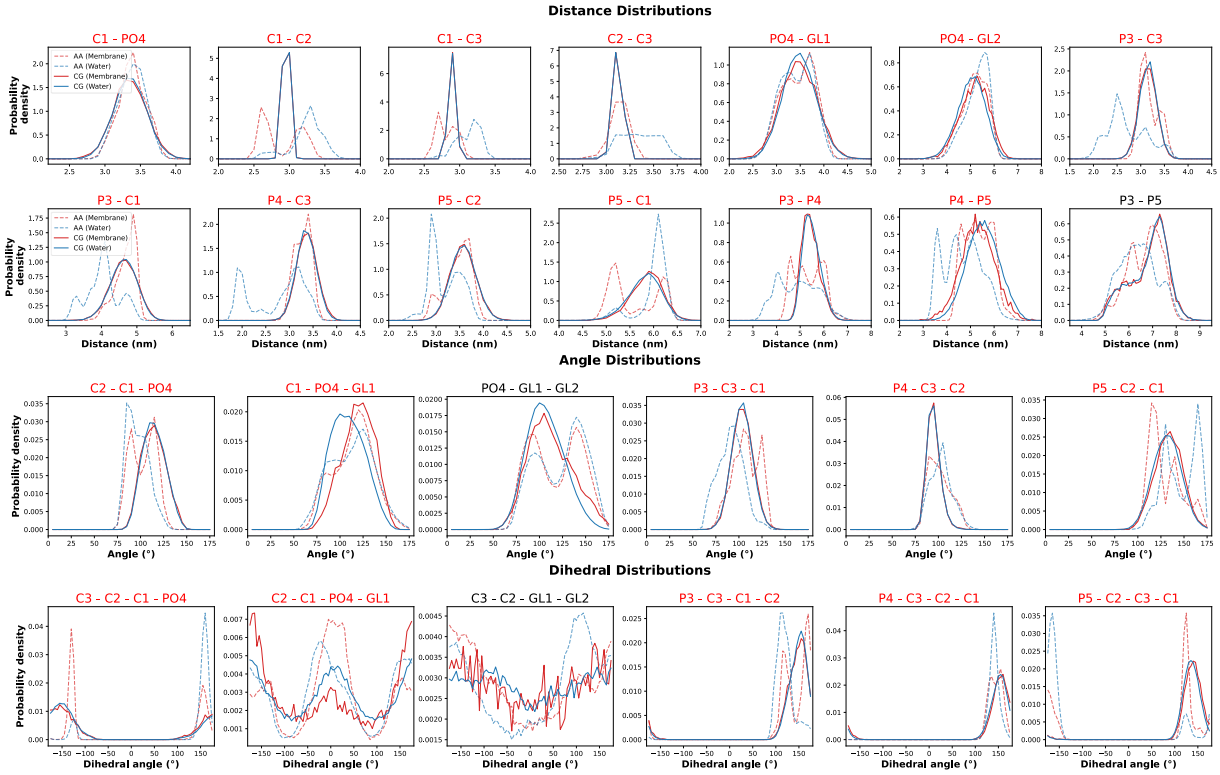


Figure S10. phosphatidylinositol 3,4,5-trisphosphate CG bonded parameters.

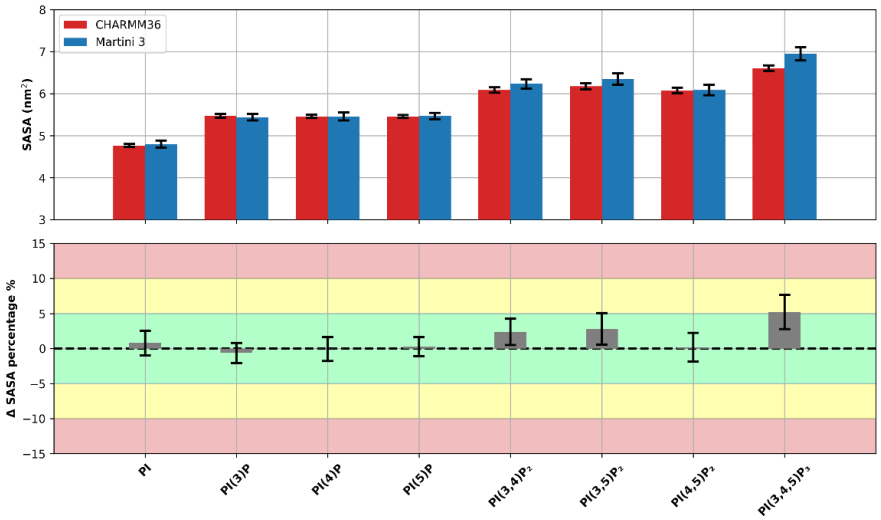


Figure S11. Phosphoinositide lipid headgroup SASA. Phosphoinositide headgroup SASA values measured for the Martini 3 and AA models, as well as the percent difference (ΔSASA) between the two.

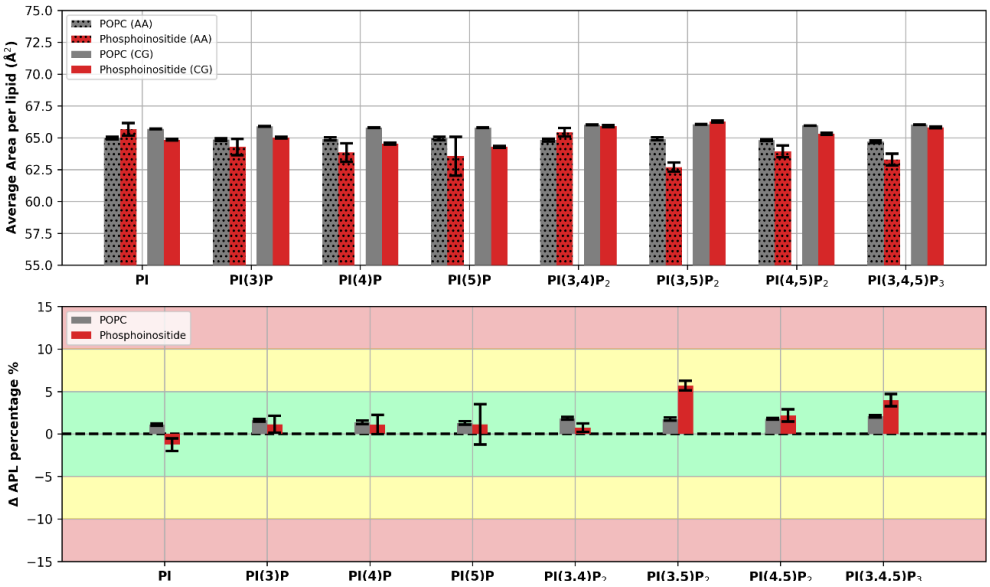


Figure S12. Phosphoinositide area per lipid. Phosphoinositide area-per-lipid values measured for the Martini 3 and AA models, as well as the percent difference (ΔApL) between the two.

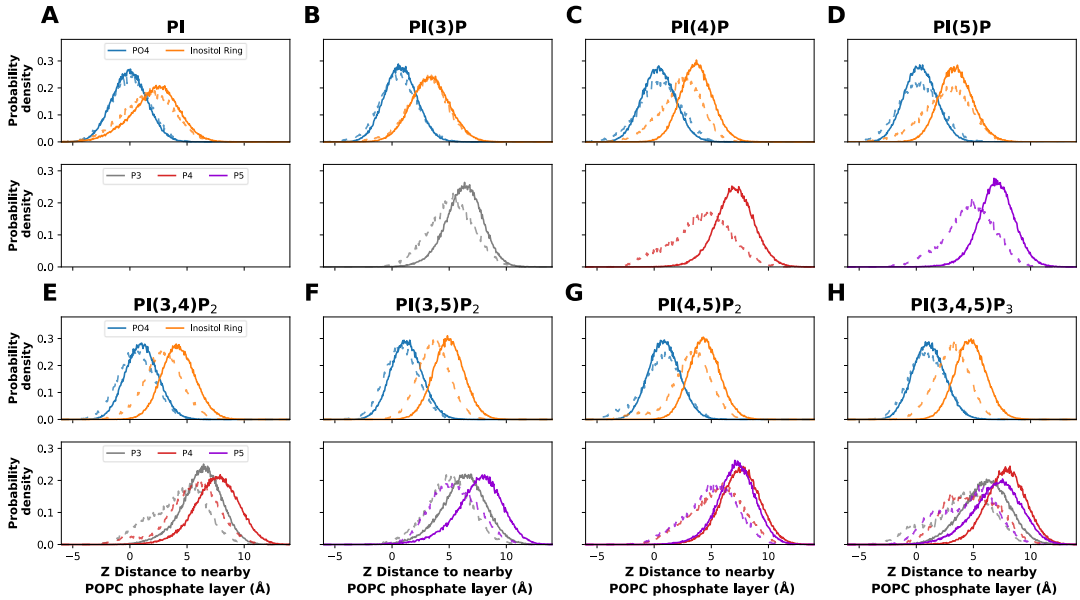


Figure S13. Phosphoinositide membrane *z* density plots. Phosphoinositide membrane density profiles for the phosphodiester group (blue), inositol ring (orange), P3, P4 and P5 (gray, red, and purple, respectively) phosphates. Profiles were determined for the Martini 3 CG models (solid line) as well as the reference AA models (dashed line), and more positive values indicate a positioning farther from the membrane center.

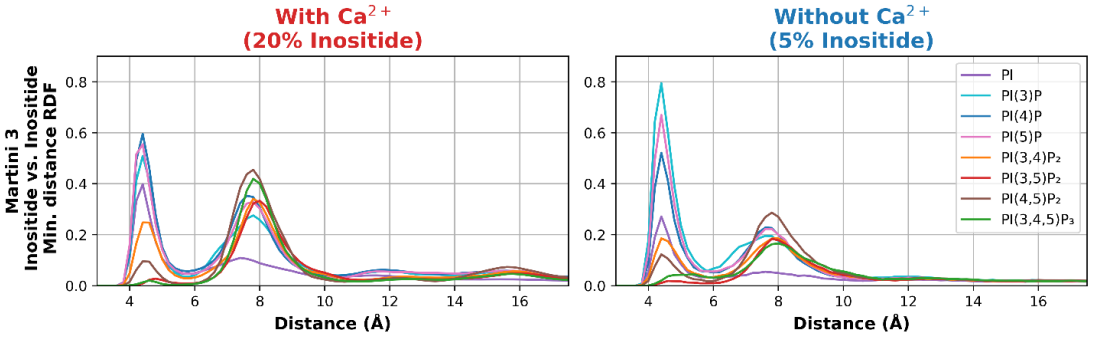


Figure S14. Martini 3 phosphoinositide models' aggregation in response to calcium. Inositide vs inositide minimum distance RDFs for the 8 inositide lipids. These were obtained from systems in the presence (left) and absence (right) of calcium. Systems in the presence of calcium were built at a 20:80 inositide:POPC ratio, mimicking high inositide concentrations in the local vicinity of phosphoinositide domains, and Ca^{2+} was added at a 2.5:1 Ca^{2+}: inositide ratio.. Systems in the absence of calcium were built at a 5:95 inositide ratio, mimicking monodisperse inositide conditions.

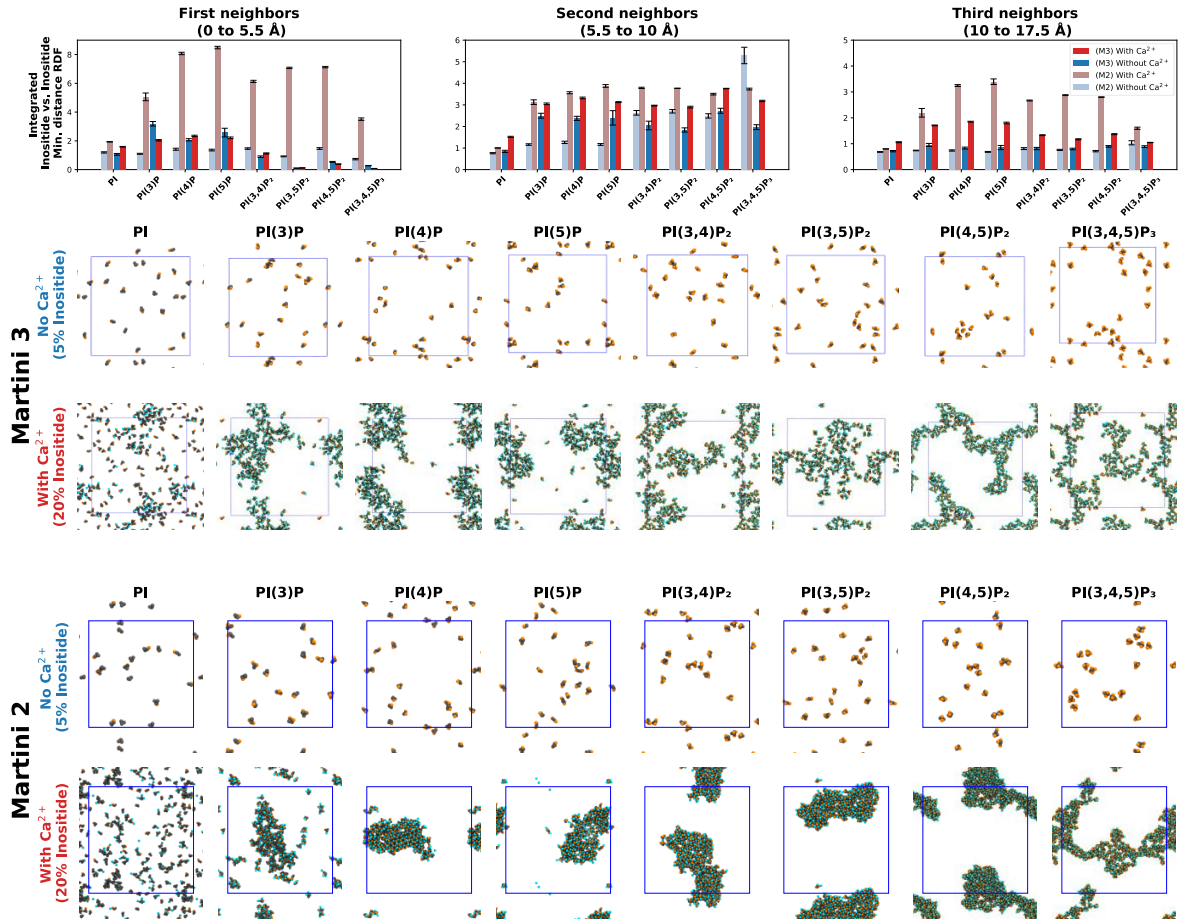


Figure S15. Comparison between Martini 2 and Martini 3 phosphoinositide model aggregation in response to calcium. Integrated inositide vs inositide minimum distance RDFs corresponding to the first, second and third local neighbors, for the 8 inositide lipids with both Martini 2 and 3 models. These were obtained from systems in the presence (red) and absence (blue) of calcium. The full RDFs can be seen in Figures S13 and S15. Final simulation snapshots of the systems are also shown, viewing the top leaflet along the membrane normal, with phosphoinositide lipid headgroups depicted in gray, phosphates in orange, and Ca^{2+} ions in blue.

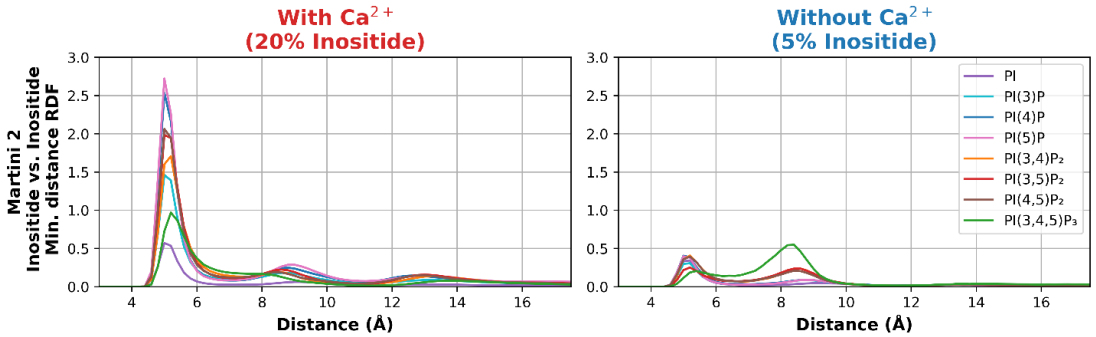


Figure S16. Martini 2 phosphoinositide models' aggregation in response to calcium. Inositide vs inositide minimum distance RDFs for the 8 inositide lipids. These were obtained from systems in the presence (left) and absence (right) of calcium.

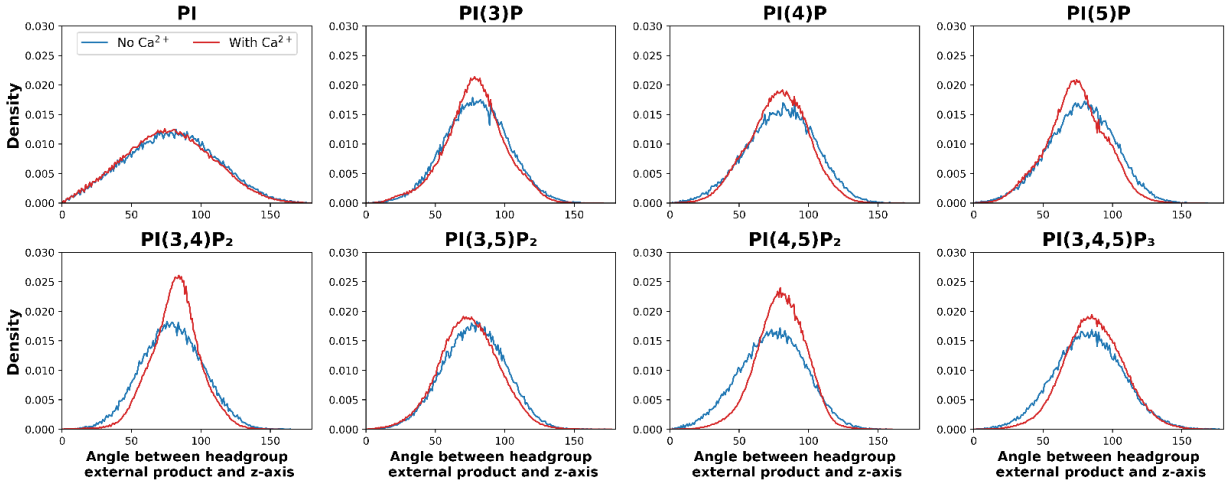


Figure S17. Impact of calcium on phosphoinositide headgroup organization. Distribution of angles between the inositol headgroup external product and the simulation box z-axis. These were obtained from systems in the presence (red) and absence (blue) of calcium.

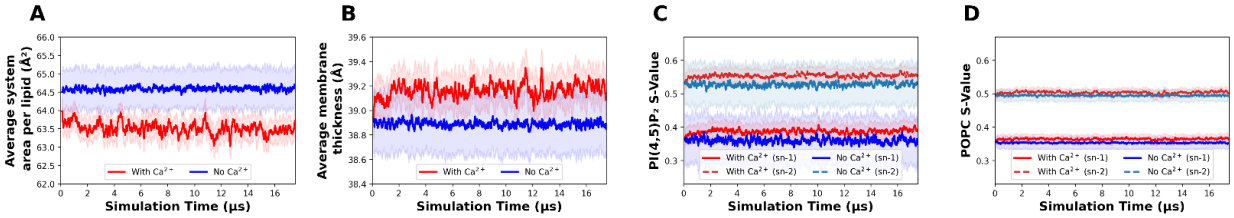


Figure S18. Impact of PI(4,5)P$_2$ calcium-induced clustering on membrane biophysical properties. System ApL (A), membrane thickness (B), and order parameters for PI(4,5)P$_2$ (C) and POPC (D) over the course of a simulation. Values were measured on 80:20 POPC:PI(4,5)P$_2$ systems containing Ca^{2+} at a 2.5:1 Ca^{2+}:PI(4,5)P$_2$ ratio, at 300 K. The values shown are the running average over the course of 200 frames, with the shaded areas representing the running standard deviation.

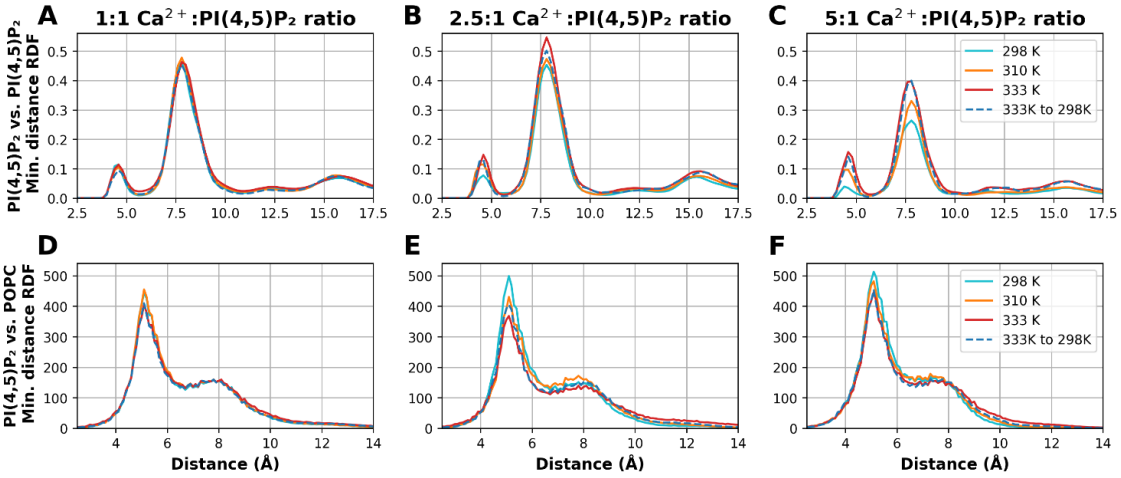


Figure S19. Effects of temperature and Ca^{2+} concentration on calcium-induced PI(4,5)P$_2$ clusters. Inositide vs inositide (A to C) and inositide vs POPC (D to F) minimum distance RDFs in the presence of varying Ca^{2+}:PI(4,5)P$_2$ ratios as well as at increasing temperatures. Results from control runs, where the final conformations from the simulations running at 333 K were used to restart the simulation at 298 K (333 K to 298 K), are also shown.

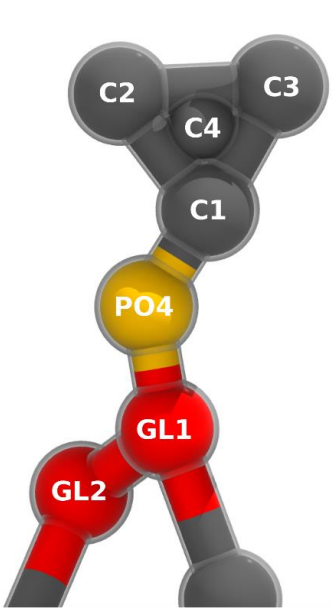


Figure S20. Structure and bead names of the phosphatidylinositol CG model parameterized in this work. Acyl-chain beads connected to the glycerol backbone (GL1 and GL2) are not shown and their names follow the standard martini 3 topology for *1*-palmitoyl-*2*-oleoyl phosphoacylglycerol.

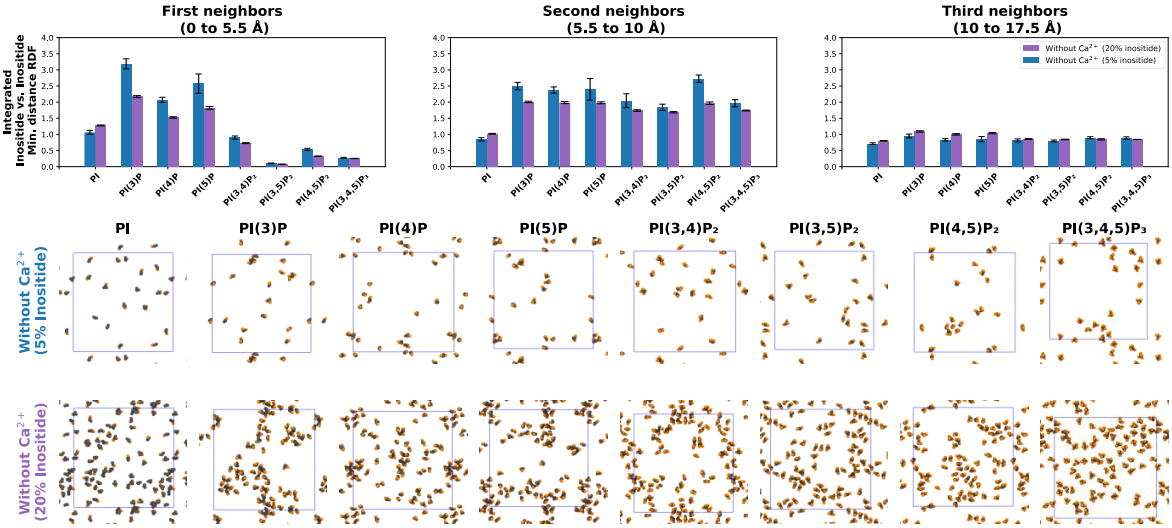


Figure S21. Effect of concentration on phosphoinositide aggregation in the absence of divalent cations. Integrated inositide vs inositide minimum distance RDFs corresponding to the first, second and third local neighbors, for the 8 inositide lipids in the absence of calcium at both 5 and 20% lipid concentration (blue and purple, respectively). Final simulation snapshots of the systems are also shown, viewing the top leaflet along the membrane normal, with phosphoinositide lipid headgroups depicted in gray, phosphates in orange, and Ca^{2+} ions in blue.

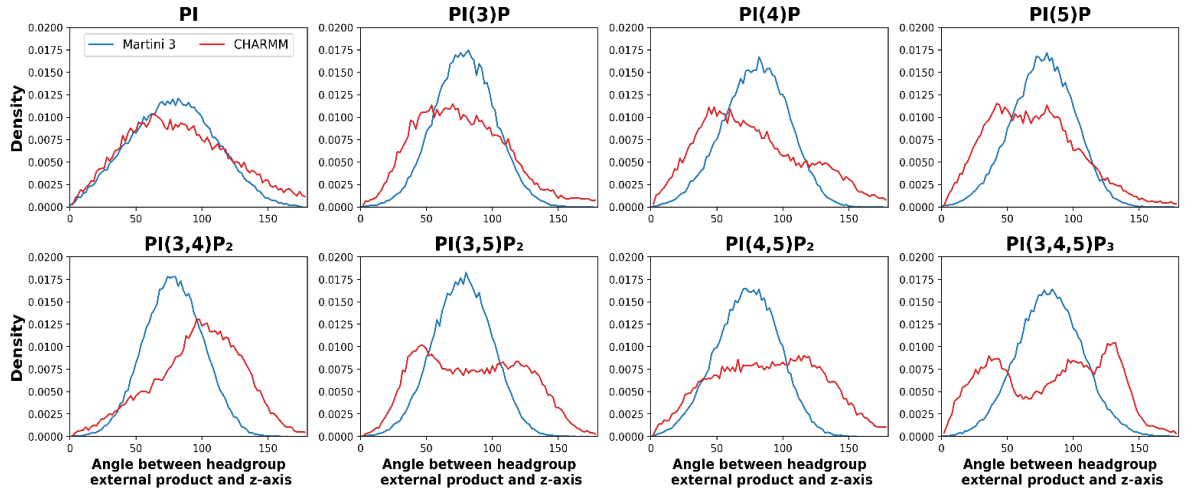


Figure S22. Comparison between CG and AA inositide headgroup tilt when inserted in a POPC bilayer. Angle between the external product defined by the headgroup beads and the simulation box z axis for the 8 CG inositide lipids and the atomistic references (blue and red, respectively).

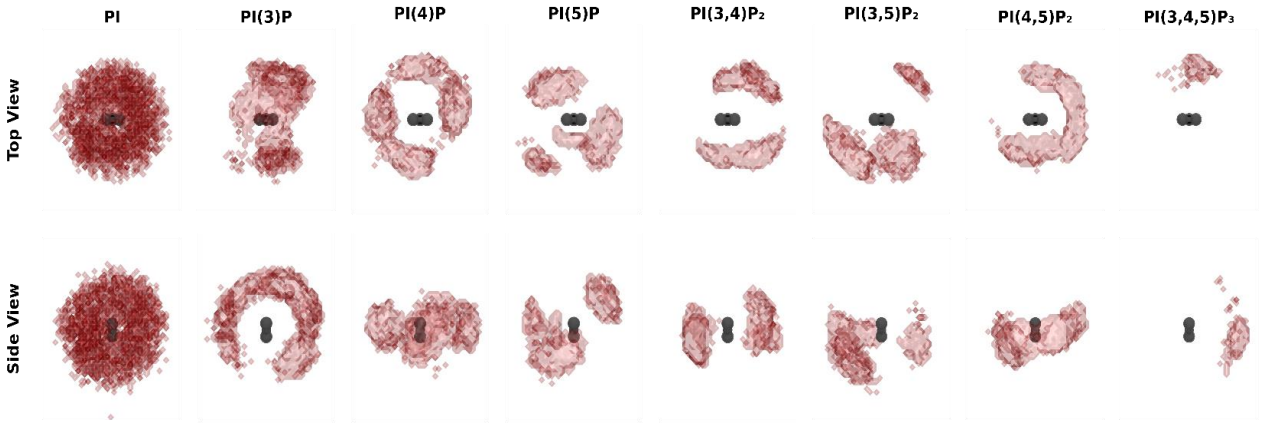


Figure S23. Inositide headgroup occupancy maps in Ca^{2+}-induced clusters. Occupancy map of inositide headgroup particles (*C1 C2 C3* and *C4* beads) within 12 Å of a reference headgroup. These were obtained from 20:80 inositide:POPC systems containing Ca^{2+} at a 2.5:1 Ca^{2+}:inositide ratio at 300 K. Views are shown with the reference aligned so that the inositol is coplanar with the *xz* plane and beads *C2* and *C3* lie parallel to the *x* axis.

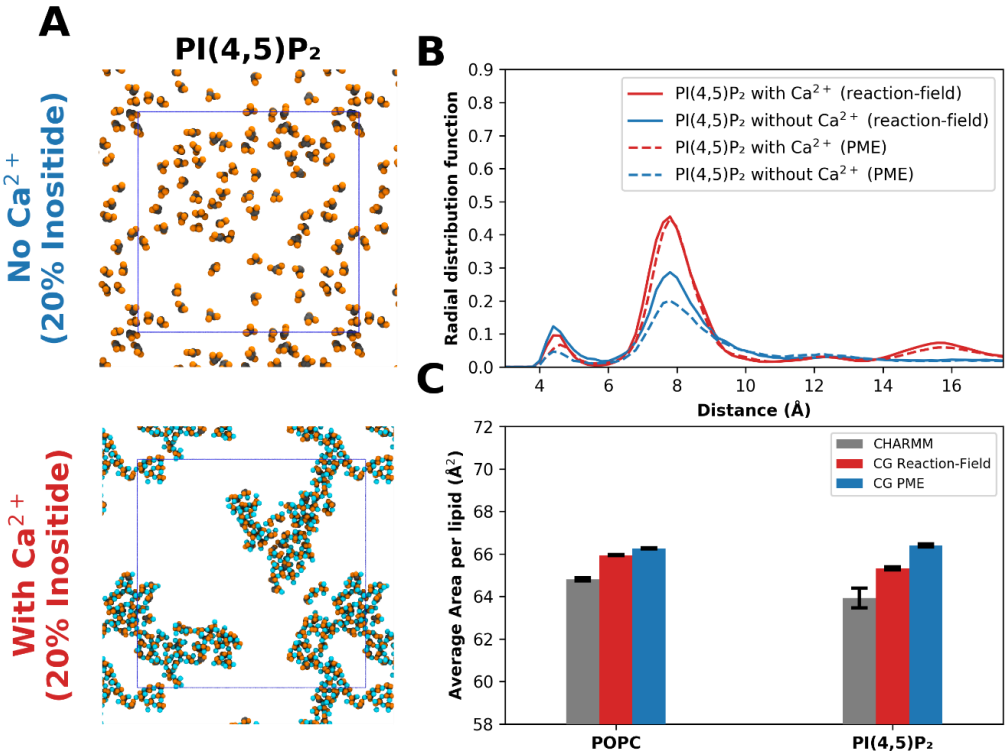


Figure S24. Effect of PME on PI(4,5)P₂ Ca^{2+}-induced aggregation and area per lipid. Final simulation snapshots (A) and minimum distance PI(4,5)P₂ vs PI(4,5)P₂ RDFs (B) obtained from systems in the presence (red) and absence (blue) of calcium at 300 K and ran with PME electrostatics treatment. For comparison, RDFs from systems run with reaction-field electrostatics are also shown (dashed lines). Simulation snapshots show the top leaflet along the membrane normal, with phosphoinositide lipid headgroups depicted in gray, phosphates in orange, and Ca^{2+} ions in blue. Average area per lipid (C) recovered for POPC and PI(4,5)P₂ in the absence of calcium, for AA (grey), CG M3 ran with reaction field (red) and CG M3 ran with PME (blue).

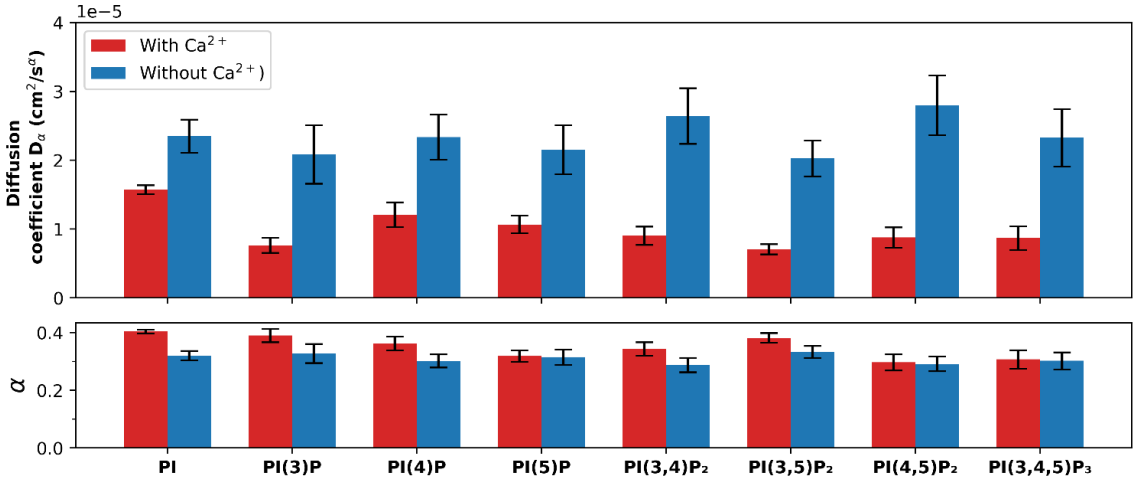


Figure S25. Inositide diffusion coefficients. Inositide diffusion coefficients and scaling exponents recovered from anomalous diffusion analysis of simulation systems, in the presence and absence of calcium (calculation details in the Supplementary Methods, above). These were obtained from systems in the presence (20:80 inositide:POPC, red) and absence (5:95 inositide:POPC, blue) of calcium at 300K.

SUPPLEMENTARY REFERENCES

(1) Pérez, F.; Granger, B. E. *IPython: A System for Interactive Scientific Computing*.

(2) Harris, C. R.; Millman, K. J.; van der Walt, S. J.; Gommers, R.; Virtanen, P.; Cournapeau, D.; Wieser, E.; Taylor, J.; Berg, S.; Smith, N. J.; Kern, R.; Picus, M.; Hoyer, S.; van Kerkwijk, M. H.; Brett, M.; Haldane, A.; del Río, J. F.; Wiebe, M.; Peterson, P.; Gérard-Marchant, P.; Sheppard, K.; Reddy, T.; Weckesser, W.; Abbasi, H.; Gohlke, C.; Oliphant, T. E. Array Programming with NumPy. *Nature* **2020**, *585* (7825), 357–362.

(3) Virtanen, P.; Gommers, R.; Oliphant, T. E.; Haberland, M.; Reddy, T.; Cournapeau, D.; Burovski, E.; Peterson, P.; Weckesser, W.; Bright, J.; van der Walt, S. J.; Brett, M.; Wilson, J.; Millman, K. J.; Mayorov, N.; Nelson, A. R. J.; Jones, E.; Kern, R.; Larson, E.; Carey, C. J.; Polat, İ.; Feng, Y.; Moore, E. W.; VanderPlas, J.; Laxalde, D.; Perktold, J.; Cimrman, R.; Henriksen, I.; Quintero, E. A.; Harris, C. R.; Archibald, A. M.; Ribeiro, A. H.; Pedregosa, F.; van Mulbregt, P. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nat. Methods* **2020**, *17* (3), 261–272.

(4) Fabian, P.; Michel, V.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Vanderplas, J.; Cournapeau, D.; Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Thirion, B.; Grisel, O.; Dubourg, V.; Passos, A.; Brucher, M.; Perrot, M.; Duchesnay, É. *Scikit-Learn: Machine Learning in Python*; 2011; Vol. 12.

(5) Hunter, J. D. Matplotlib: A 2D Graphics Environment. *Comput. Sci. Eng.* **2007**, *9* (3), 90–95.

(6) Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual Molecular Dynamics. *J. Mol. Graph.* **1996**, *14* (1), 33–38.

(7) Ingólfsson, H. I.; Melo, M. N.; Van Eerden, F. J.; Arnarez, C.; Lopez, C. A.; Wassenaar, T. A.; Periole, X.; De Vries, A. H.; Tieleman, D. P.; Marrink, S. J. Lipid Organization of the Plasma Membrane. *J. Am. Chem. Soc.* **2014**, *136* (41), 14554–14559.

(8) Rycroft, C. H. VORO++: A Three-Dimensional Voronoi Cell Library in C++. *Chaos*. American Institute of Physics Inc. December 27, 2009, p 041111.

(9) Sousa, F. M.; Lima, L. M. P.; Arnarez, C.; Pereira, M. M.; Melo, M. N. Coarse-Grained Parameterization of Nucleotide Cofactors and Metabolites: Protonation Constants, Partition Coefficients, and Model Topologies. *J. Chem. Inf. Model.* **2021**.

(10) Goga, N.; Rzepiela, A. J.; De Vries, A. H.; Marrink, S. J.; Berendsen, H. J. C. Efficient Algorithms for Langevin and DPD Dynamics. *J. Chem. Theory Comput.* **2012**, *8* (10), 3637–3649.

(11) Parrinello, M.; Rahman, A. Polymorphic Transitions in Single Crystals: A New Molecular Dynamics Method. *J. Appl. Phys.* **1981**, *52* (12), 7182–7190.

(12) Klimovich, P. V.; Shirts, M. R.; Mobley, D. L. Guidelines for the Analysis of Free Energy Calculations. *J. Comput. Aided. Mol. Des.* **2015**, *29* (5), 397–411.

(13) Shirts, M. R.; Chodera, J. D. Statistically Optimal Analysis of Samples from Multiple Equilibrium States. *J. Chem. Phys.* **2008**, *129* (12), 124105.

(14) 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.

(15) and, R. S. R.; Taylor*, R. Intermolecular Nonbonded Contact Distances in Organic Crystal Structures: Comparison with Distances Expected from van Der Waals Radii. **1996**.

(16) Souza, P. C. T. T.; Alessandri, R.; Barnoud, J.; Thallmair, S.; Faustino, I.; Grünewald, F.; Patmanidis, I.; Abdizadeh, H.; Bruininks, B. M. H. H.; Wassenaar, T. A.; Kroon, P. C.; Melcr, J.; Nieto, V.; Corradi, V.; Khan, H. M.; Domański, J.; Javanainen, M.; Martinez-Seara, H.; Reuter, N.; Best, R. B.; Vattulainen, I.; Monticelli, L.; Periole, X.; Tieleman, D. P.; de Vries, A. H.; Marrink, S. J. Martini 3: A General Purpose Force Field for Coarse-Grained Molecular Dynamics. *Nat. Methods* **2021**, *18* (4), 382–388.

(17) Schmalhorst, P. S.; Deluweit, F.; Scherrers, R.; Heisenberg, C.-P.; Sikora, M. Overcoming the Limitations of the MARTINI Force Field in Simulations of Polysaccharides. *J. Chem. Theory Comput.* **2017**, *13* (10), 5039–5053.

(18) Goose, J. E.; Sansom, M. S. P. Reduced Lateral Mobility of Lipids and Proteins in Crowded Membranes. *PLoS Comput. Biol.* **2013**, *9* (4), e1003033.

(19) Reddy, T.; Shorthouse, D.; Parton, D. L.; Jefferys, E.; Fowler, P. W.; Chavent, M.; Baaden, M.; Sansom, M. S. P. Nothing to Sneeze At: A Dynamic and Integrative Computational Model of an Influenza A Virion. *Structure* **2015**, *23* (3), 584–597.

Supporting Information

Revisiting the impact of acyl-chain composition on phosphatidylinositol 4,5-bisphosphate nanodomains using the Martini 3 force field

Luís Borges-Araújo [†,§,⊥], Fábio Fernandes [†,‡,⊥,*], Manuel N. Melo[§,*]

† iBB-Institute for Bioengineering and Biosciences, Instituto Superior Técnico, Universidade de Lisboa, Lisbon, Portugal
⊥ Associate Laboratory i4HB—Institute for Health and Bioeconomy at Instituto Superior Técnico, Universidade de Lisboa, Lisbon, Portugal
‡ Department of Bioengineering; Instituto Superior Técnico, Universidade de Lisboa, 1049-001 Lisbon, Portugal;
§ Instituto de Tecnologia Química e Biológica António Xavier, Universidade Nova de Lisboa, Av. da República, 2780-157 Oeiras, Portugal.

SUPPLEMENTARY METHODS

Molecular dynamics simulation analysis. All simulations were analyzed making use of in-house developed Python3 programs using the MDAnalysis package. We also used the IPython[1], numpy[2], SciPy[3], scikit-learn[4] and matplotlib[5] packages for scientific computing in Python. Visualization and rendering of the simulations was performed with the molecular graphics viewer VMD[6]. Unless otherwise stated, the last 2 μs of each CG simulation were used for analysis.

—*System-wide Area per lipid (ApL)* was calculated from the average box area divided by the number of lipids in each leaflet.

— *Membrane thickness* was determined by calculating the distance between the centers of geometry of each leaflet's phosphate layer. Membrane thickness was defined as the *z* component of this distance.

— The *S-value order parameter* is a geometric parameter which can be used to compare the different lipid tails' overall order/flexibility. Lipid tail S-value order parameters were calculated for each lipid tail bond, of each independent lipid as described previously[7]. The S-value used here is the extension to coarse-grain of the concept of lipid C-H bond order parameter and is defined

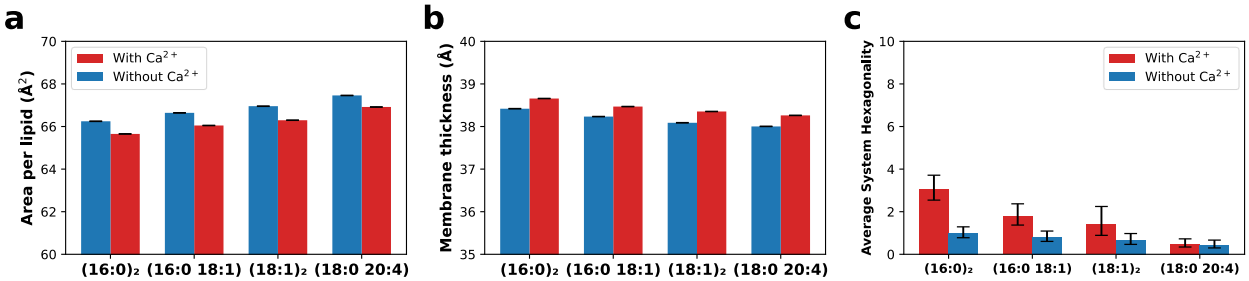
as: $S = \frac{1}{2}(3\langle(\cos\theta)^2\rangle - 1)$, where θ is the angle between the vector along a particular lipid tail bead pair and the normal of the bilayer, which we approximate as the system's z-axis. When calculating the S-Value vs. the distance to a reference target, each S-Value obtained was then binned according to their distance to the reference in the xy plane. Bins with less than analyzed 200000 elements over the analysed simulation time (2 μs) were removed from analysis.

—The *hexagonality index* was used in comparing the ability of the different lipid tails to organize into a hexagonal matrix. Hexagonality was calculated for the first lipid tail bead of each individual lipid. A lipid tail bead was considered to be within a hexagonal matrix if the following criteria were met: i) the tail bead is within 6 Å of the 6 closest neighbors; and ii) each of the 6 closest neighbors to the target bead also has at least 2 neighbors within that cutoff. To eliminate some thermal fluctuation noise the hexagonality index was calculated from a trajectory with coordinates averaged over a 5-frame running window.

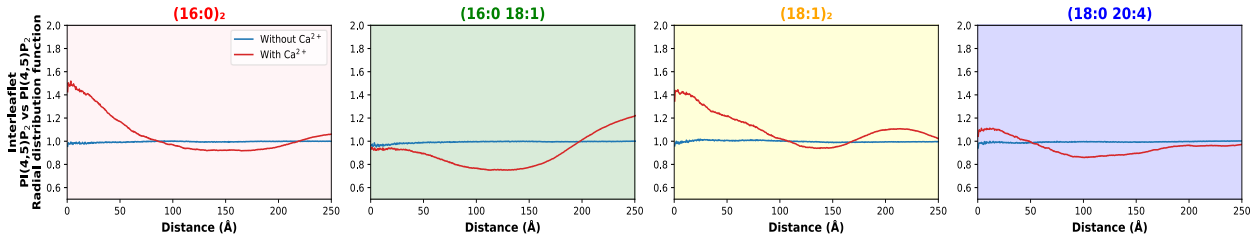
—*Radial distribution functions (RDF)* are obtained by processing the distance matrix from one reference set of positions to one or more sets of positions. The distance arrays are histogrammed and averaged over the simulation time. The RDF is then obtained by normalizing the histogram by the number of reference positions, volume of the bin and average particle density of the positions. In our cases the RDFs are calculated over the xy membrane plane and thus the z component of the positions is not taken into consideration. Additionally, normalization employs the area of the bin instead of its volume.

—*Membrane curvature analysis* was performed adapting the MDAnalysis tool to calculate membrane curvature (https://github.com/MDAnalysis/membrane-curvature). In summary, the GL1 and GL2 beads from one of the bilayer leaflets were selected and used to derive a surface from which the curvature can be calculated. This surface is divided in a grid and the mean curvature (H nm^{-1}) is calculated for each individual bin. Square and rectangular membranes were divided in 35 × 35, and 35 × 11 grids, respectively. Curvature values are averaged over every 5 consecutive frames. For visualization, the bins are interpolated using Gaussian interpolation.

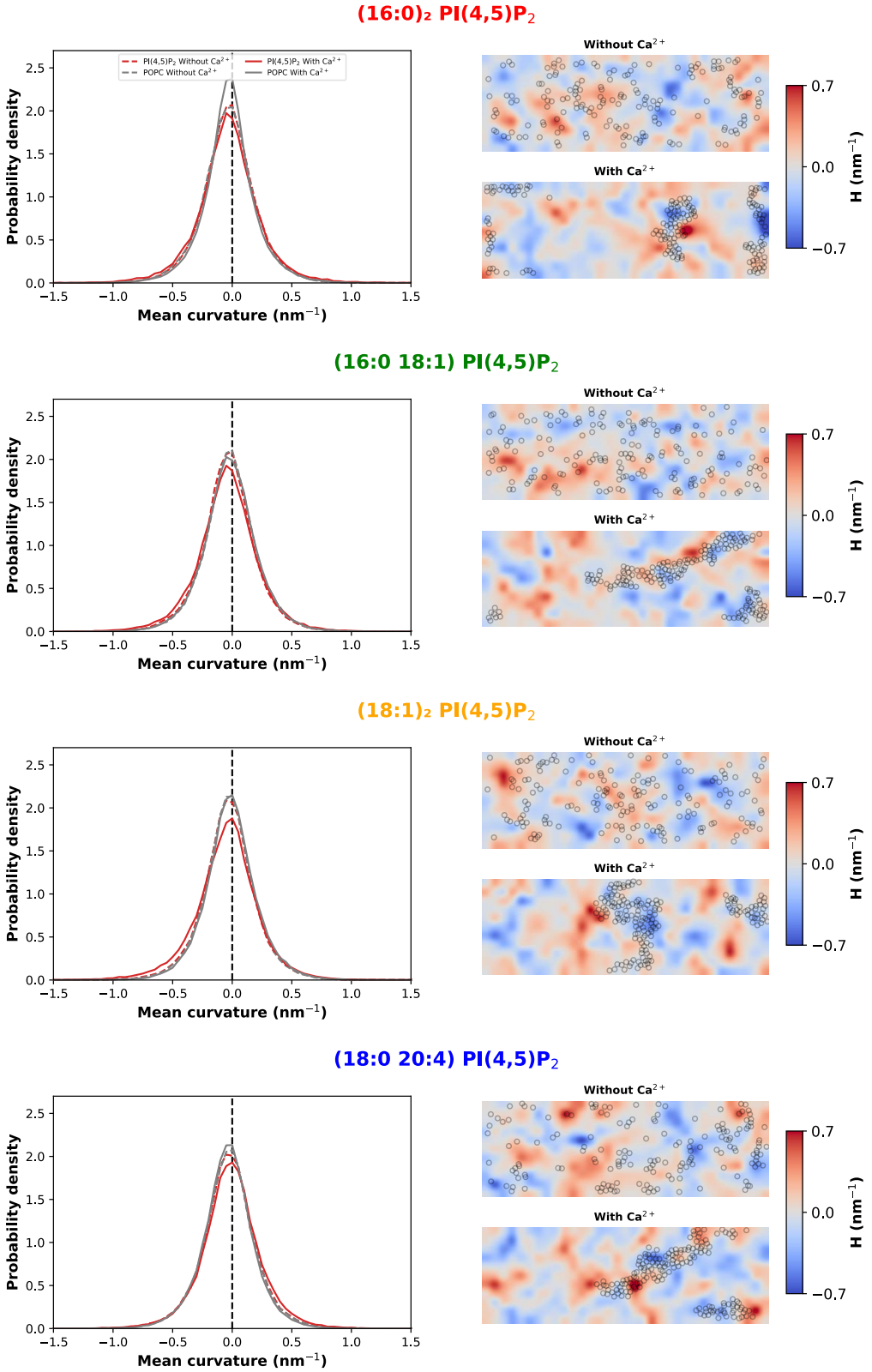
—*Membrane bending rigidity (K_c)* was calculated using simple Helfrich-Canham elastic theory as described by Fowler *et al*[8], by adapting the code they have freely shared, which can be found at https://github.com/philipwfowler/calculate-bilayer-power-spectrum . In summary, in the absence of surface tension, the power spectrum of the height fluctuations, $\langle|h(\mathbf{q})^2|\rangle$, can be given by: $q^4\langle|h(\mathbf{q})^2|\rangle = \frac{k_B T}{K_c}$, where q is the magnitude of the wavevector $\mathbf{q}$ (nm^{-1}), K_c is the bending rigidity, T is the temperature and k_B is the Boltzmann's constant. In practice, the surface of the bilayer is defined by the position of the PO4 beads, and the height of the bilayer interpolated onto a grid of size 0.5 nm using cubic splines. The resulting arrays are transformed into Fourier space using the FFTW routines present in the SciPy python module. K_c is then recovered by plotting the power spectrum equation and fitting the initial constant portion of the function at low q values. For simplicity, we set this to be $q < 0.5$ nm^{-1}, following the recommendations of Fowler *et al*[8]. The last 4 μs of simulations were divided into 1 μs bins. Power spectrum was calculated for every bin and fitted as previously described. Average K_c and error values were determined from this data pool by bootstrapping.



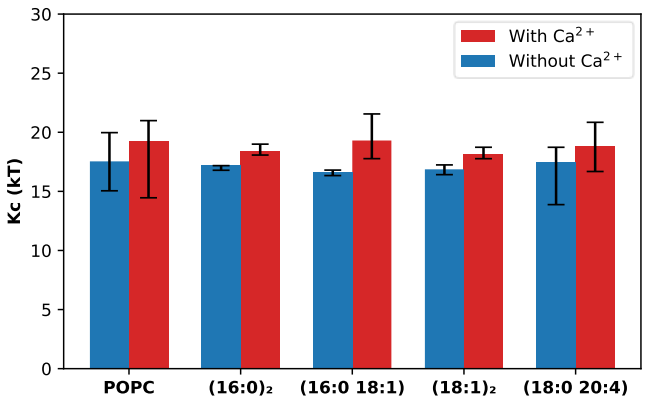
Supplementary figure S1. Martini 3 CG MD simulations showcase the impact of acyl-chain composition on PI(4,5)P_2 and PI(4,5)P_2 nanodomain biophysical properties. (a) Average membrane area per lipid, (b) membrane thickness, and (c) system hexagonality for each system in the presence and absence of Ca^{2+}.



Supplementary figure S2. Impact of Ca^{2+} and acyl-chain composition on PI(4,5)P2 interleaflet registration. Interleaflet PI(4,5)P2 vs PI(4,5)P2 radial distribution function (RDF) for each composition in the presence and absence of Ca^{2+}.



Supplementary figure S3. Association of Ca^{2+}-induced PI(4,5)P_2 nanodomains with membrane curvature fluctuations using rectangular membrane patches. Histogram of mean curvature associated with either PI(4,5)P_2 (red) or POPC (grey), from asymmetric membrane simulations containing 10 mol% PI(4,5)P_2 in the inner leaflet, for each of the four studied acyl chain compositions in the presence (solid line) and absence (dashed line) of Ca^{2+}. For each system, top view snapshots are shown, colored by the local curvature value. PI(4,5)P_2 phosphodiester beads are represented by the black circles.



Supplementary figure S4. Impact of PI(4,5)P_2 composition and Ca^{2+}-induced nanodomain formation on bilayer bending rigidity. Bending rigidity, K_c, for a pure POPC bilayer system and asymmetric POPC bilayers containing 10% PI(4,5)P2 in the lower leaflet, for each acyl-chain composition in the presence and absence of Ca^{2+}. The power spectra from which the K_c values were recovered are shown in Supplementary figure S5.

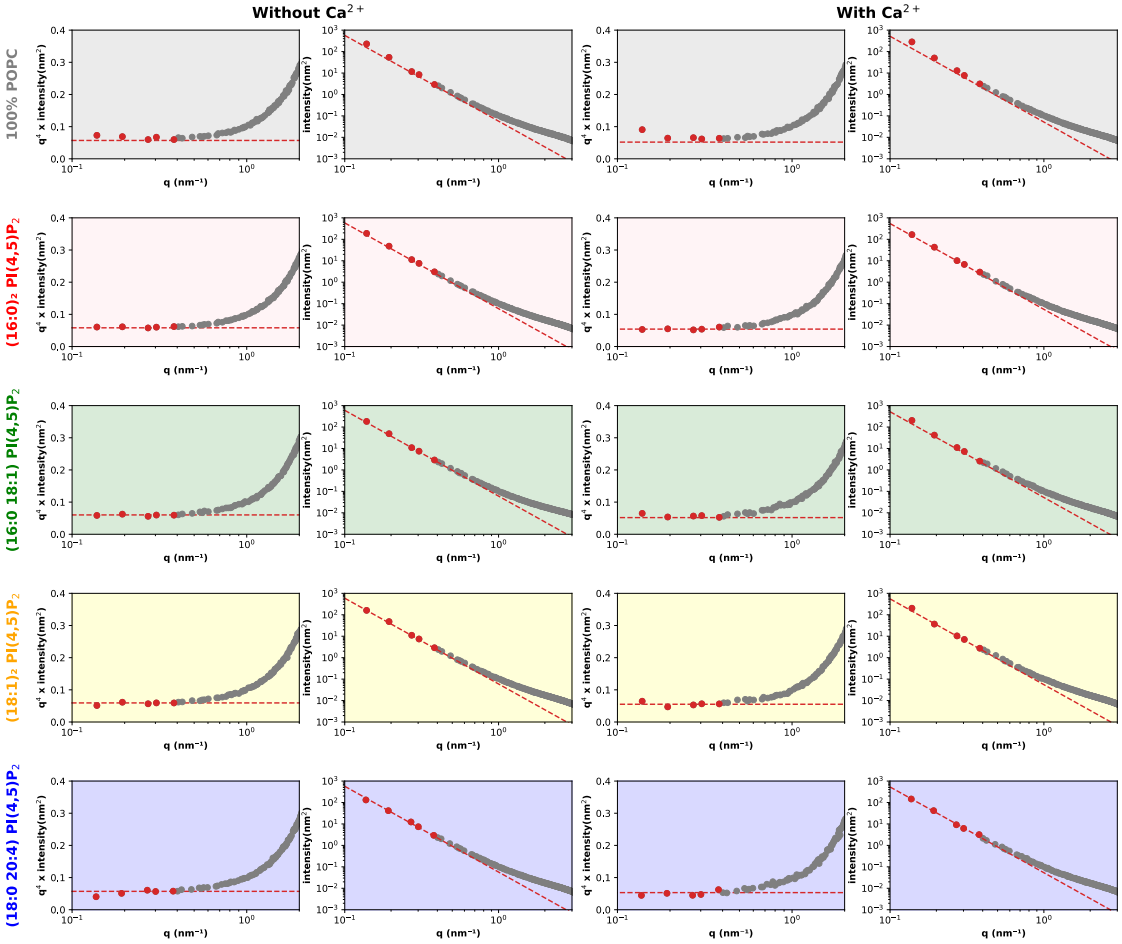


Figure S5. Impact of PI(4,5)P_2 and Ca^{2+}-induced PI(4,5)P_2 clustering on the power spectra of height simulations. Power spectra of the height simulations for our asymmetric membrane simulations containing 10 mol% PI(4,5)P_2 in the inner leaflet, for each of the four studied acyl chain compositions in the presence and absence of Ca^{2+}. Spectra shown (grey dots) were obtained from the first of the 4 bins analyzed. The fit shown (red) is the average for all bins.

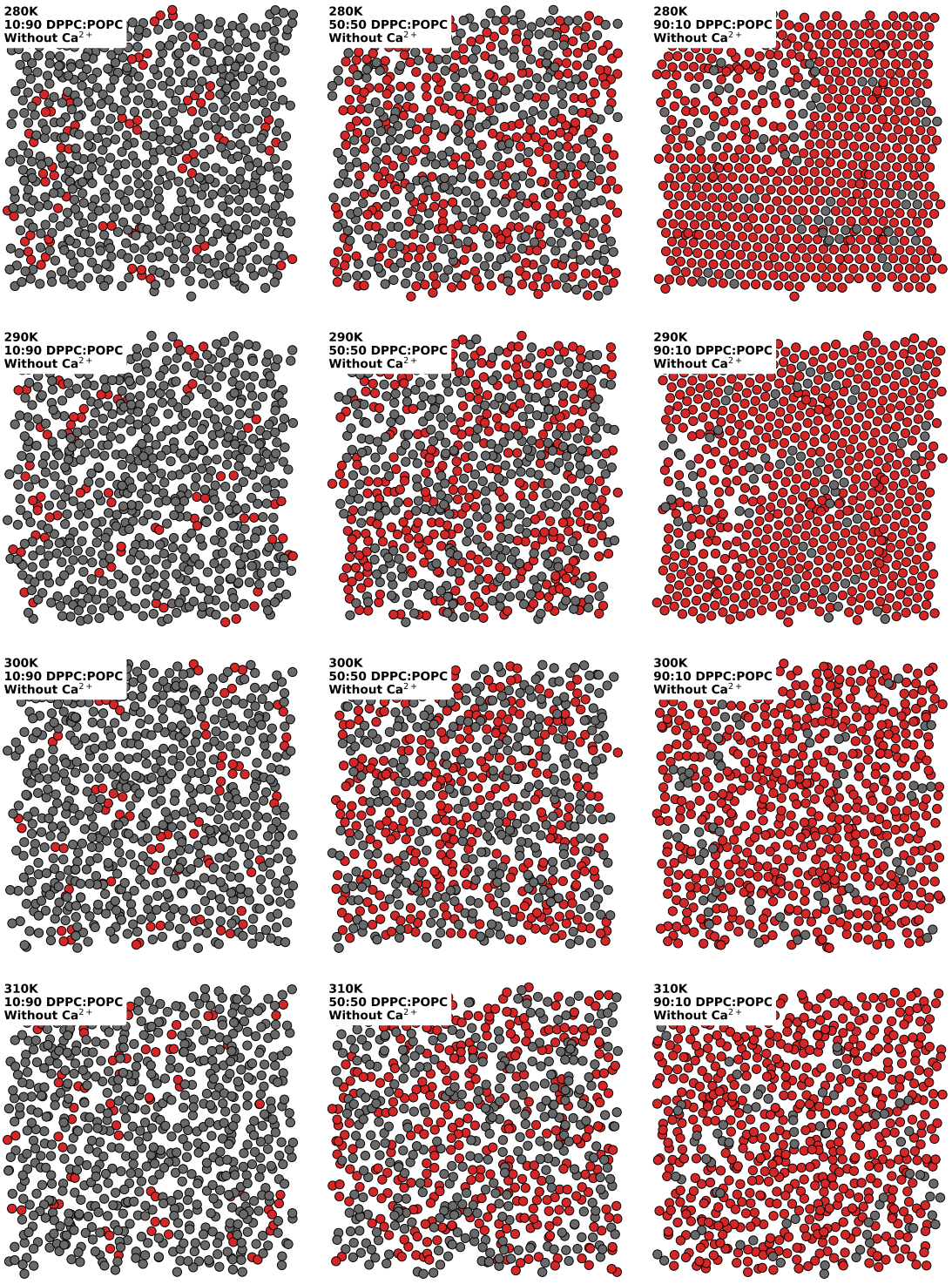


Figure S6. Final snapshots from DPPC:POPC gel-phase formation simulations. Top view snapshots from DPPC:POPC gel-phase formation simulations ran at multiple temperatures and molar ratios. The first acyl-chain beads for both DPPC and POPC are shown in red and grey, respectively.

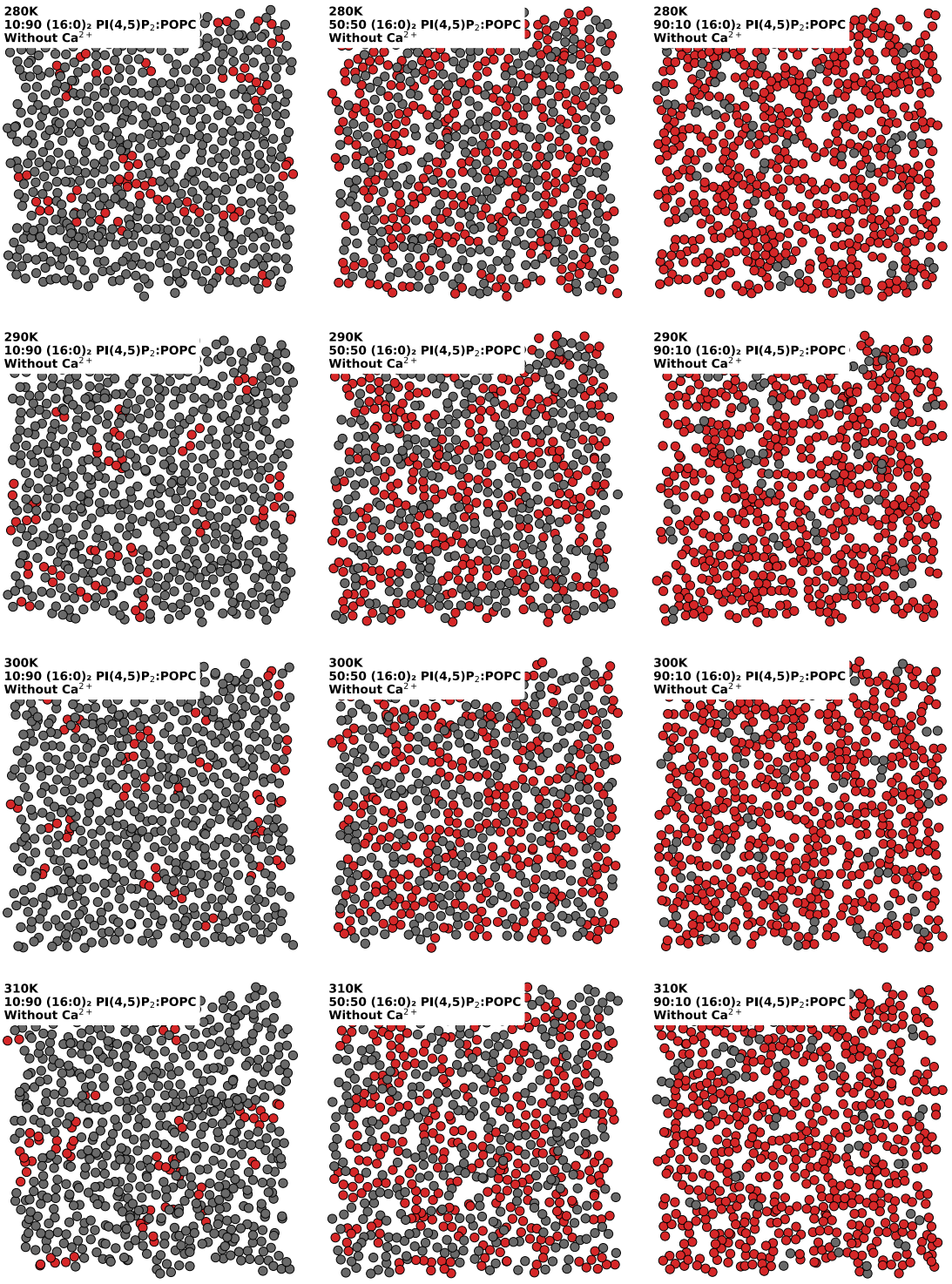


Figure S7. Final snapshots from $(16:0)_2$ PI(4,5)P_2 : POPC gel-phase formation simulations, in the absence of Ca^{2+}. Top view snapshots from $(16:0)_2$ PI(4,5)P_2 : POPC gel-phase formation simulations ran at multiple temperatures and molar ratios, in the absence of Ca^{2+}. The first acyl-chain beads for both PI(4,5)P_2 and POPC are shown in red and grey, respectively.

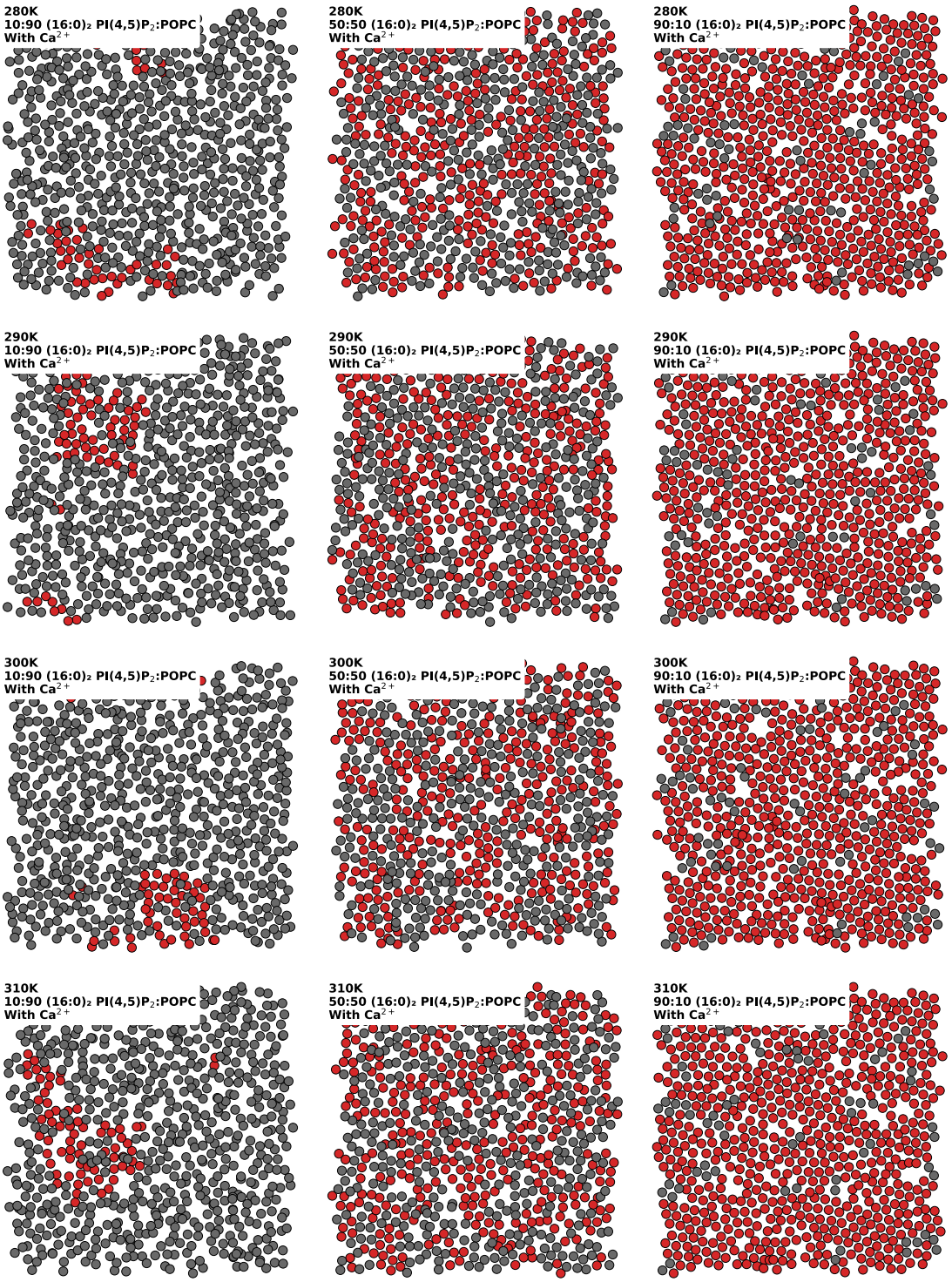


Figure S8. Final snapshots from $(16:0)_2$ PI(4,5)P_2 : POPC gel-phase formation simulations, in the presence of Ca^{2+}. Top view snapshots from $(16:0)_2$ PI(4,5)P_2 : POPC gel-phase formation simulations ran at multiple temperatures and molar ratios, in the presence of Ca^{2+}. The first acyl-chain beads for both PI(4,5)P_2 and POPC are shown in red and grey, respectively.

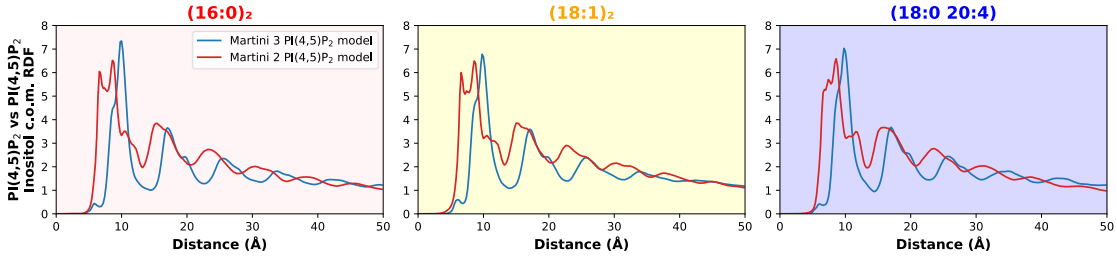


Figure S9. Radial distribution functions (RDFs) of the PI(4,5)P_2 headgroup center of mass (c.o.m.) against itself, for both Martini 2 and Martini 3 PI(4,5)P_2 models. PI(4,5)P_2 vs PI(4,5)P_2 c.o.m. RDF, from 50 × 50 nm^2 POPC membrane patches containing 10 mol% PI(4,5)P_2 for three acyl-chain compositions in the presence of Ca^{2+}. RDFs were obtained for both the Martini 2 (red line) and Martini 3 (blue line) models.

SUPPLEMENTARY REFERENCES

(1) Pérez, F.; Granger, B. E. *IPython: A System for Interactive Scientific Computing*.

(2) Harris, C. R.; Millman, K. J.; van der Walt, S. J.; Gommers, R.; Virtanen, P.; Cournapeau, D.; Wieser, E.; Taylor, J.; Berg, S.; Smith, N. J.; Kern, R.; Picus, M.; Hoyer, S.; van Kerkwijk, M. H.; Brett, M.; Haldane, A.; del Río, J. F.; Wiebe, M.; Peterson, P.; Gérard-Marchant, P.; Sheppard, K.; Reddy, T.; Weckesser, W.; Abbasi, H.; Gohlke, C.; Oliphant, T. E. Array Programming with NumPy. *Nature* **2020**, *585* (7825), 357–362.

(3) Virtanen, P.; Gommers, R.; Oliphant, T. E.; Haberland, M.; Reddy, T.; Cournapeau, D.; Burovski, E.; Peterson, P.; Weckesser, W.; Bright, J.; van der Walt, S. J.; Brett, M.; Wilson, J.; Millman, K. J.; Mayorov, N.; Nelson, A. R. J.; Jones, E.; Kern, R.; Larson, E.; Carey, C. J.; Polat, İ.; Feng, Y.; Moore, E. W.; VanderPlas, J.; Laxalde, D.; Perktold, J.; Cimrman, R.; Henriksen, I.; Quintero, E. A.; Harris, C. R.; Archibald, A. M.; Ribeiro, A. H.; Pedregosa, F.; van Mulbregt, P. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nat. Methods* **2020**, *17* (3), 261–272.

(4) Fabian, P.; Michel, V.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Vanderplas, J.; Cournapeau, D.; Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Thirion, B.; Grisel, O.; Dubourg, V.; Passos, A.; Brucher, M.; Perrot, M.; Duchesnay, É. *Scikit-Learn: Machine Learning in Python*; 2011; Vol. 12.

(5) Hunter, J. D. Matplotlib: A 2D Graphics Environment. *Comput. Sci. Eng.* **2007**, *9* (3), 90–95.

(6) Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual Molecular Dynamics. *J. Mol. Graph.* **1996**, *14* (1), 33–38.

(7) Ingólfsson, H. I.; Melo, M. N.; Van Eerden, F. J.; Arnarez, C.; Lopez, C. A.; Wassenaar, T. A.; Periole, X.; De Vries, A. H.; Tieleman, D. P.; Marrink, S. J. Lipid Organization of the Plasma Membrane. *J. Am. Chem. Soc.* **2014**, *136* (41), 14554–14559.

(8) Fowler, P. W.; Hélie, J.; Duncan, A.; Chavent, M.; Koldsø, H.; Sansom, M. S. P. Membrane Stiffness Is Modified by Integral Membrane Proteins. *Soft Matter* **2016**, *12* (37), 7792–7803.

D | Supplementary Material for Chapter 5

Supplementary Material

Genetic algorithm-guided optimization of anti-SARS-CoV-2 biomolecules.

Luís Borges-Araújo [†,⊥,‡], Diana Lousa[‡], Cláudio M. Soares[‡], Manuel N. Melo[‡,*]

† iBB-Institute for Bioengineering and Biosciences, Instituto Superior Técnico, Universidade de Lisboa, Lisbon, Portugal
⊥ Associate Laboratory i4HB—Institute for Health and Bioeconomy at Instituto Superior Técnico, Universidade de Lisboa, Lisbon, Portugal

‡ Instituto de Tecnologia Química e Biológica António Xavier, Universidade Nova de Lisboa, Av. da República, 2780-157 Oeiras, Portugal.

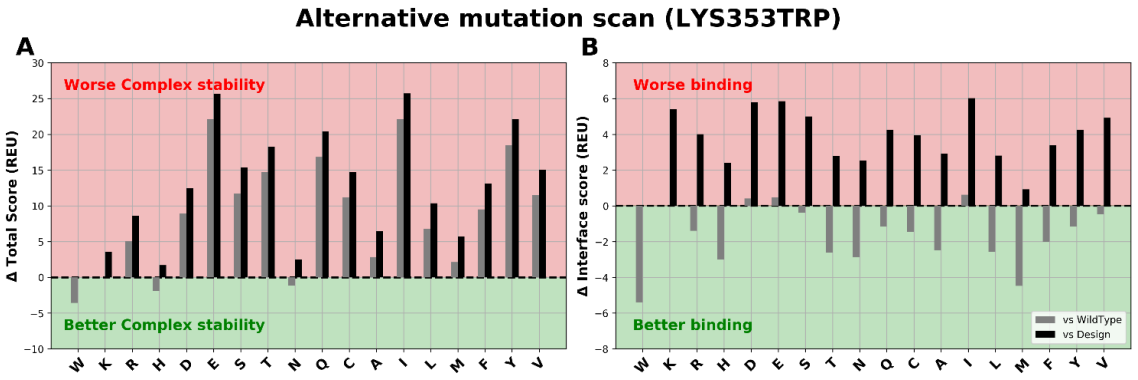


Figure S1. Alternative mutation scan for LYS353TRP. Difference in total score (A) and interface score (B) relative to the wt and designed sequences are shown. At least 200 docking runs were performed for each sequence.

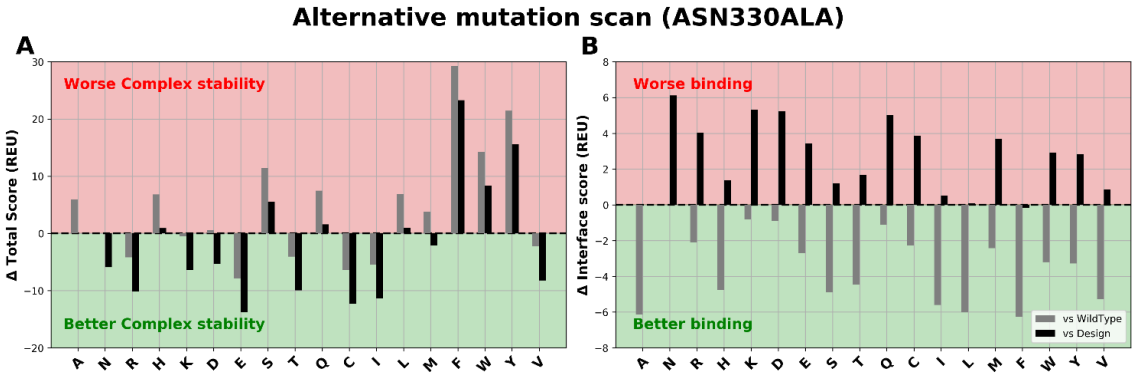


Figure S2. Alternative mutation scan for ASN330ALA. Difference in total score (A) and interface score (B) relative to the wt and designed sequences are shown. At least 200 docking runs were performed for each sequence.

Table 1. Several statistics concerning the *wt*, HH and Designed ACE2 variants. The values correspond to the average of the averages obtained in the five replicas for each condition. Errors shown are standard error of the mean, obtained by bootstrapping, expressed as 2σ. All statistics are calculated from the 2nd μs of simulation.

	RMSD from X-ray structure (nm)	Contact surface between ACE2 and RBD (nm^2)	Average number of residue contact pairs	Number of hydrogen bonds
wt	0.250 +- 0.016	10.156 +- 0.059	145	9.52 +- 1.13
H374N H378N	0.227 +- 0.020	10.141 +- 0.065	145	9.87 +- 0.20
Designed	0.273 +- 0.045	9.929 +- 0.638	143	5.40 +- 1.26

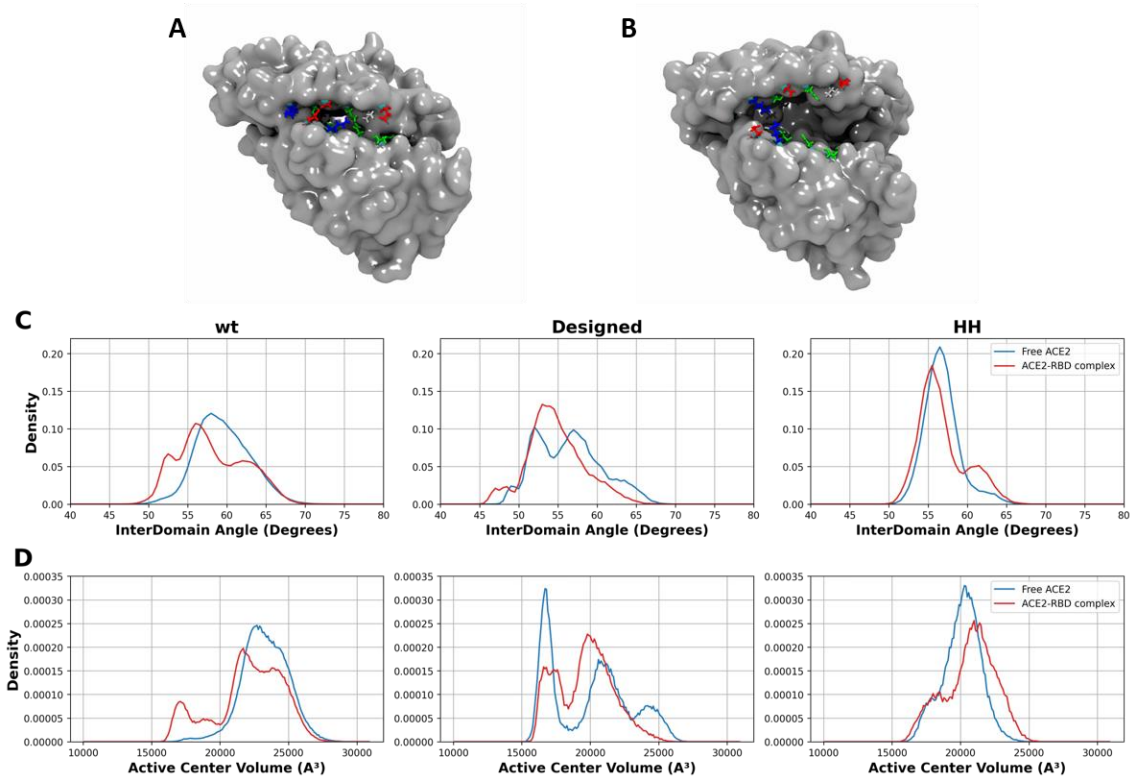


Figure S3. ACE2 active center hinge-bending dynamics. ACE2 structure with the active center hinge open (A) and closed (B). Interdomain angle populations (C) and active center volume (D) for the *wt*, designed and HH mutant sequences. ACE2 parameters were determined for both free ACE2 in solution (blue line) and the ACE2-RBD complexes (red line).

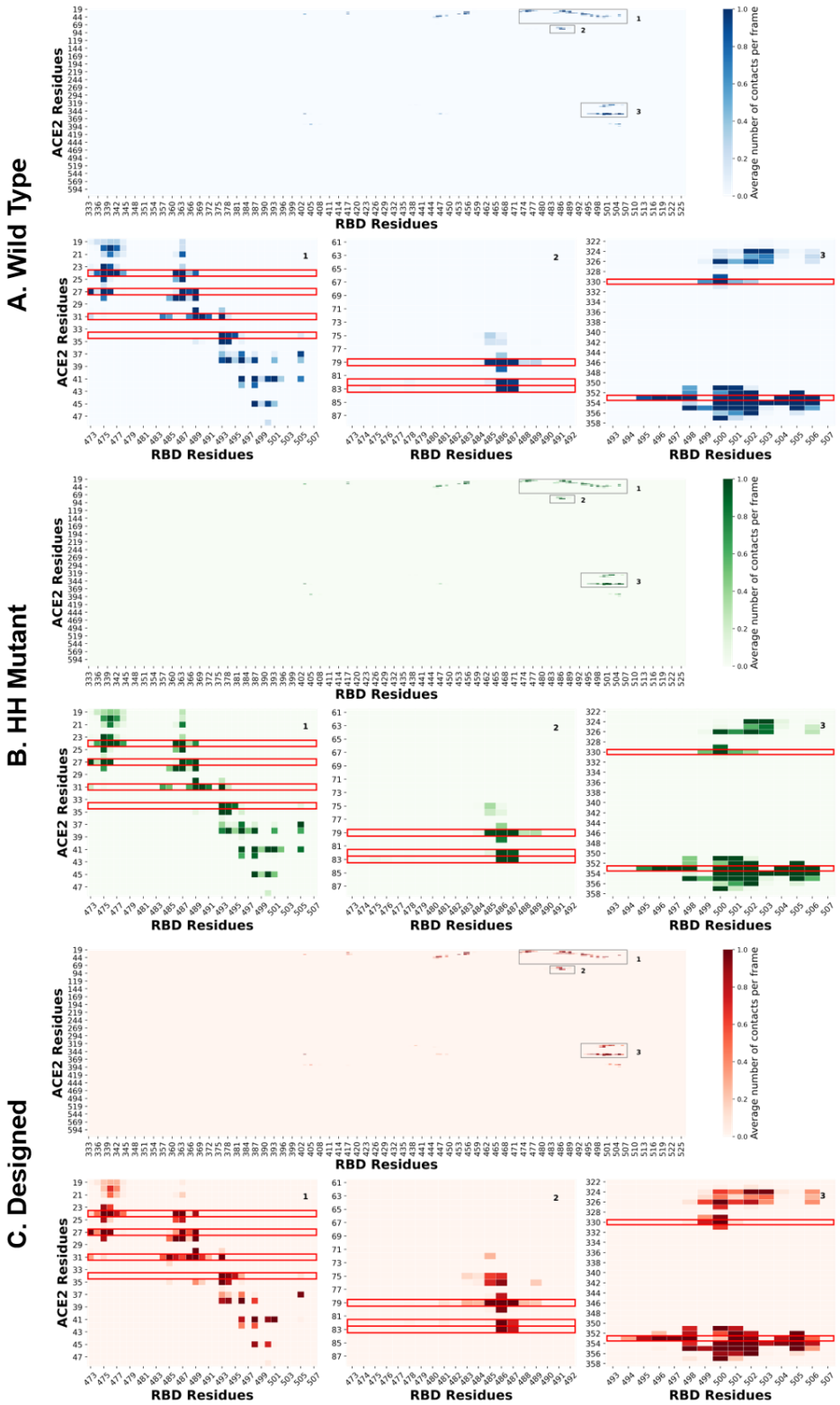


Figure S4. ACE2-RBD contact matrixes. ACE2-RBD contact matrixes for the *wt* (A), HH mutant (B) and designed (C) sequences. Contacts are defined as having two residues' center of geometry within 10 Å, and are averaged over the course of the simulations. The mutated residues are marked by the red boxes.

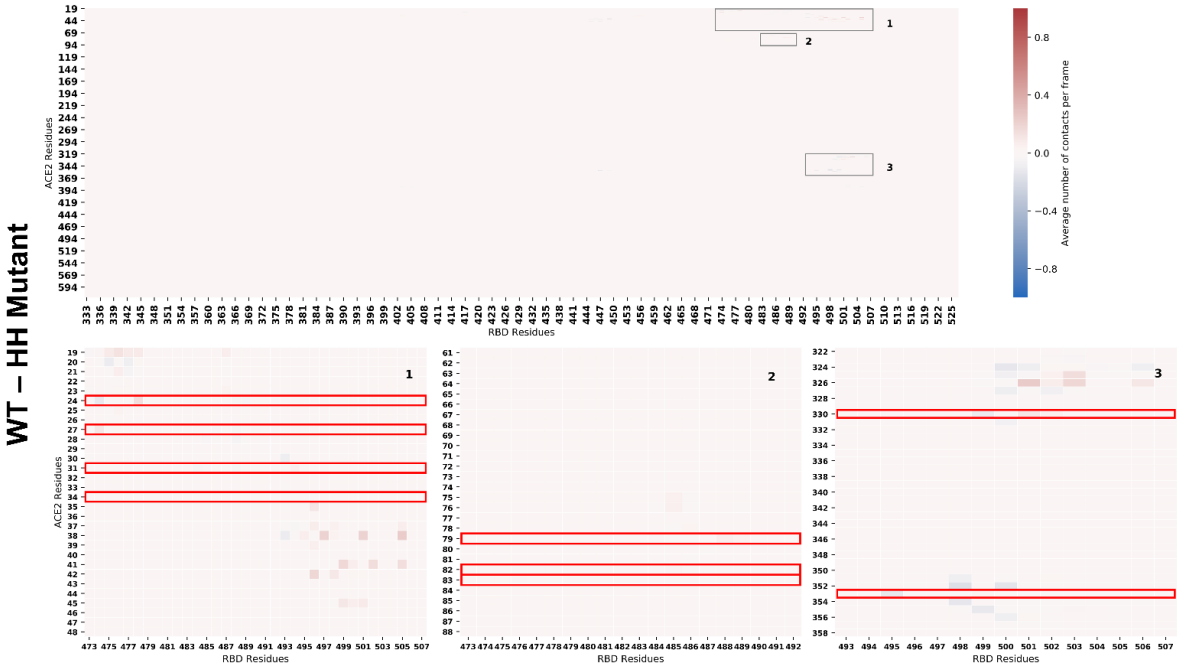


Figure S5. Difference in ACE2-RBD contacts between the wt and HH mutant sequences. HH mutant contact matrix relative to the wt sequence. Contacts are defined as having two residues' center of geometry within 10 Å, and are averaged over the course of the simulations. The mutated residues are marked by the red boxes.

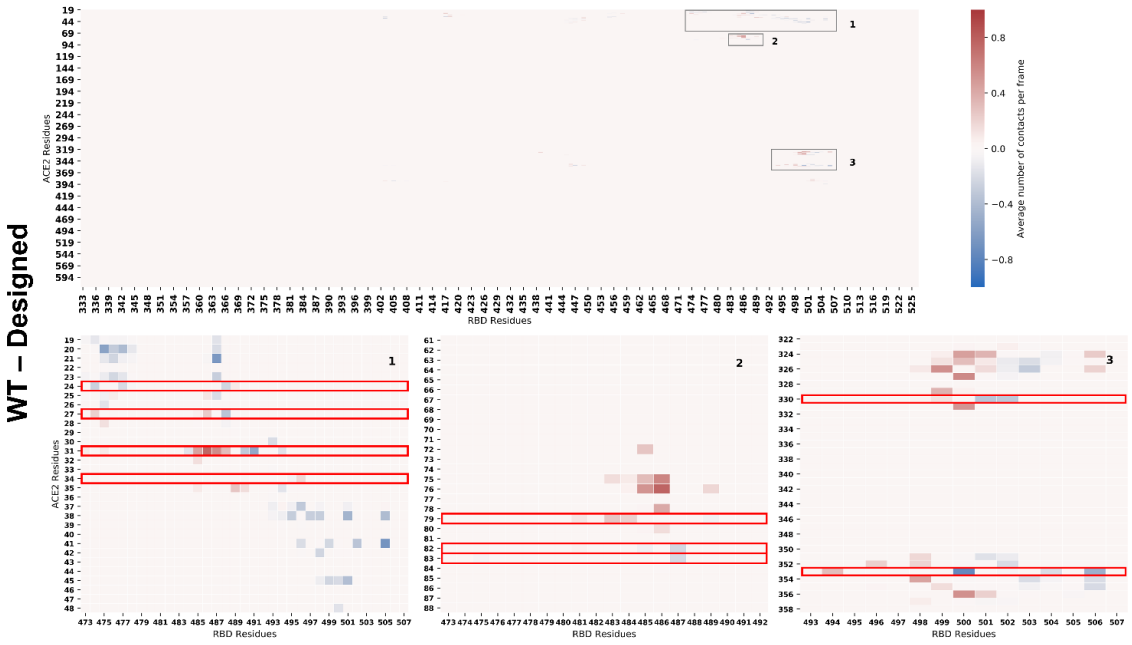


Figure S6. Difference in ACE2-RBD contacts between the wt and designed sequences. Designed sequence contact matrix relative to the wt sequence. Contacts are defined as having two residues' center of geometry within 10 Å, and are averaged over the course of the simulations. The mutated residues are marked by the red boxes.

Supporting Information

SARS-CoV-2 variants impact RBD conformational dynamics and ACE2 accessibility

Mariana Valério[†,§], Luís Borges-Araújo[†,⊥,‡,§], Manuel N. Melo[†], Diana Lousa[†], Cláudio M. Soares[†,*]

† Instituto de Tecnologia Química e Biológica António Xavier, Universidade Nova de Lisboa, Av. da República, 2780-157 Oeiras, Portugal.

⊥ iBB-Institute for Bioengineering and Biosciences, Instituto Superior Técnico, Universidade de Lisboa, Lisbon, Portugal

‡ Associate Laboratory i4HB—Institute for Health and Bioeconomy at Instituto Superior Técnico, Universidade de Lisboa, Lisbon, Portugal

§ M.V. and L.B.A. contributed equally to this work.

Videos S1 to S4. RBD and RBM conformation dynamics for the wt, alpha, beta, and delta variants. Trajectory samples recovered from the AA MD simulations of the *wt*, alpha, beta and delta RBD in water. The ridge region of the RBD is colored in red. Residues of interest are labelled at the start of the video. Renderization done with VMD, with the positions averaged over 5 consecutive frames.

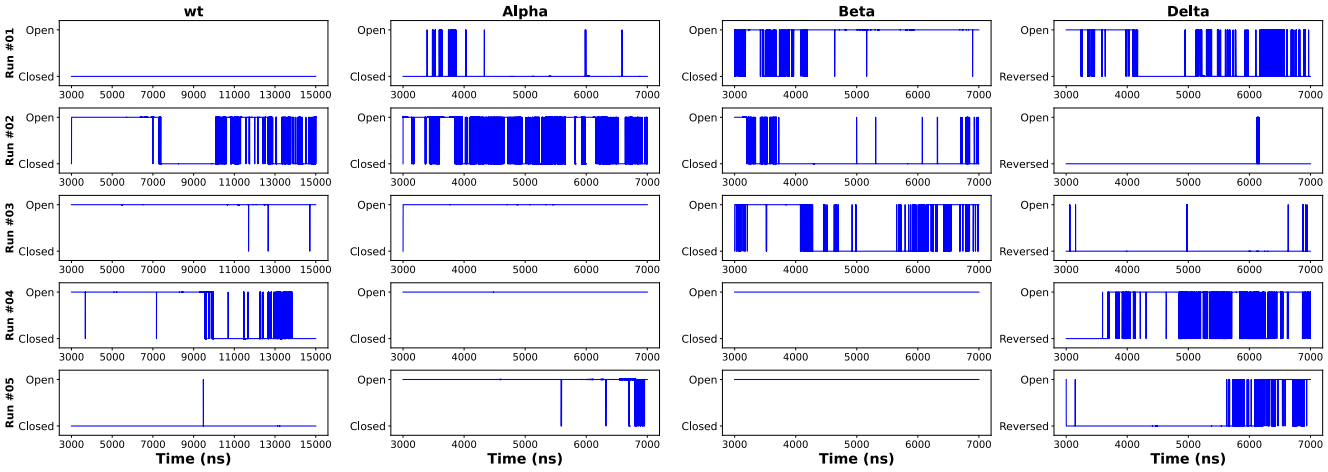


Figure S1. RBD open – closed dynamics over simulation time. RBD open – closed dynamics as determined by analysis of the conformational basins (Supplementary Table S1) recovered by PCA. Data shown for the five replicas for each variant tested. The first 3 μs of simulation were used for equilibration.

Table S1. Energy surface landscape analysis from 2D PCA of SARS-CoV-2 RBD conformational dynamics in water. Energy surface landscape analysis and defined basins for each of the tested RBD variants. Energy minima, frame percentage and loop conformation for each of the basins is also given. Overall analysis of "open" vs. "closed" conformation is also shown. 95 % Confidence intervals (CI) were calculated with bootstrap resampling from the frame percentages recovered from the individual simulation replicas. Representative structures can be seen in figures S8, S9, S10 and S11.

	Basin	Free Energy	<E>/k_BT	S/R	E_{min}/k_BT	Frame percentage (%)	Loop Conformation	Time in "closed" state (%) ± CI (95%)	Time in "open" state (%) ± CI (95%)	"closing" ΔΔG (kJ/mol)
WT	0	-5.16	1.36	6.52	0.44	37.69	"Closed"	$55.50^{+2.12}_{-2.93}$	$44.49^{+0.93}_{-1.12}$	$0.5^{+0.21}_{-0.13}$
	1	-4.67	0.99	5.67	0.26	23.01	"Open"			
	2	-4.41	0.57	4.98	0.00	17.66	"Closed"			
	3	-4.20	1.04	5.24	0.44	14.42	"Open"			
	4	-3.28	3.62	6.90	2.86	5.77	"Open"			
	5	-1.57	4.07	5.64	3.61	1.04	"Open"			
	6	0.79	5.30	4.51	5.02	0.13	"Open"			
	7	0.83	5.42	4.60	5.21	0.14	"Closed"			
	8	0.91	5.39	4.48	5.01	0.12	"Open"			
	9	2.70	5.06	2.37	5.01	0.01	"Closed"			
					Total	99.99				
Alpha (N501Y)	0	-5.20	1.02	6.22	0.00	39.18	"Open"	$27.31^{+1.79}_{-2.04}$	$72.64^{+1.16}_{-1.43}$	$-2.44^{+0.21}_{-0.23}$
	1	-4.64	1.47	6.11	0.35	22.41	"Open"			
	2	-4.37	1.84	6.21	0.83	17.23	"Closed"			
	3	-3.82	0.96	4.78	0.35	9.85	"Closed"			
	4	-3.10	2.77	5.86	1.83	4.82	"Open"			
	5	-2.71	3.88	6.58	3.01	3.28	"Open"			
	6	-2.34	4.10	6.44	3.11	2.41	"Open"			
	7	-0.28	5.21	5.49	4.93	0.29	"Open"			
	8	-0.11	4.43	4.54	4.05	0.25	"Open"			
	9	0.11	5.46	5.35	5.10	0.23	"Closed"			
					Total	99.95				
Beta (K417N E484K N501Y)	0	-4.95	0.58	5.53	0.00	23.73	"Open"	$30.18^{+1.24}_{-1.74}$	$69.81^{+0.66}_{-0.73}$	$-2.09^{+0.13}_{-0.15}$
	1	-4.70	0.46	5.16	0.09	18.41	"Open"			
	2	-4.38	1.31	5.69	0.90	13.43	"Closed"			
	3	-3.98	1.32	5.30	0.90	9.03	"Closed"			
	4	-3.83	2.70	6.53	1.93	7.72	"Closed"			
	5	-3.78	2.86	6.64	2.02	7.36	"Open"			
	6	-3.36	2.93	6.30	2.30	4.86	"Open"			
	7	-3.20	2.20	5.40	1.68	4.14	"Open"			
	8	-2.93	4.08	7.02	3.32	3.21	"Open"			
	9	-2.71	3.66	6.37	3.11	2.54	"Open"			
	10	-2.58	3.53	6.12	2.98	2.23	"Open"			
	11	-2.14	3.70	5.85	3.28	1.44	"Open"			
	12	-1.74	3.78	5.52	3.48	0.95	"Open"			
	13	-1.72	3.99	5.71	3.70	0.94	"Open"			
					Total	99.99				
Delta (L452R T478K)	0	-5.74	1.03	6.77	0.00	53.51	"Open"	N/A	100.00	N/A
	1	-4.97	2.32	7.29	1.46	24.76	'Reverse Open"			
	2	-4.13	1.70	5.84	1.20	10.75	'Reverse Open"			
	3	-3.46	3.17	6.63	2.62	5.50	"Open"			
	4	-2.93	3.34	6.27	2.63	3.22	"Open"			
	5	-1.88	4.15	6.03	3.77	1.12	"Open"			
	6	-1.85	4.12	5.97	3.60	1.13	"Open"			
					Total	99.99				

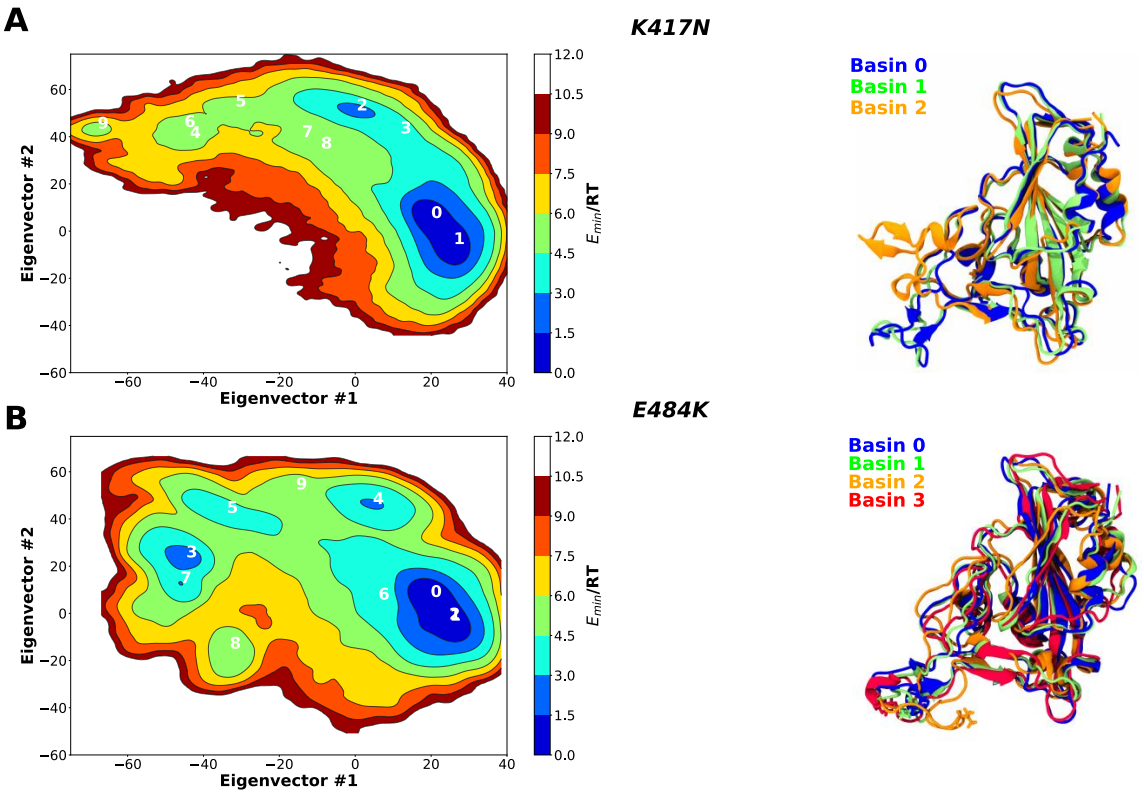


Figure S2. Two-dimension principal component analysis (PCA) of SARS-CoV-2 RBD's mutants conformational dynamics in water. Plots of the first two principal components determined from the Cα backbone of the (A) K417N and (B) E484K RBD mutants. Snapshots of the lowest energy structures for selected basins are also shown.

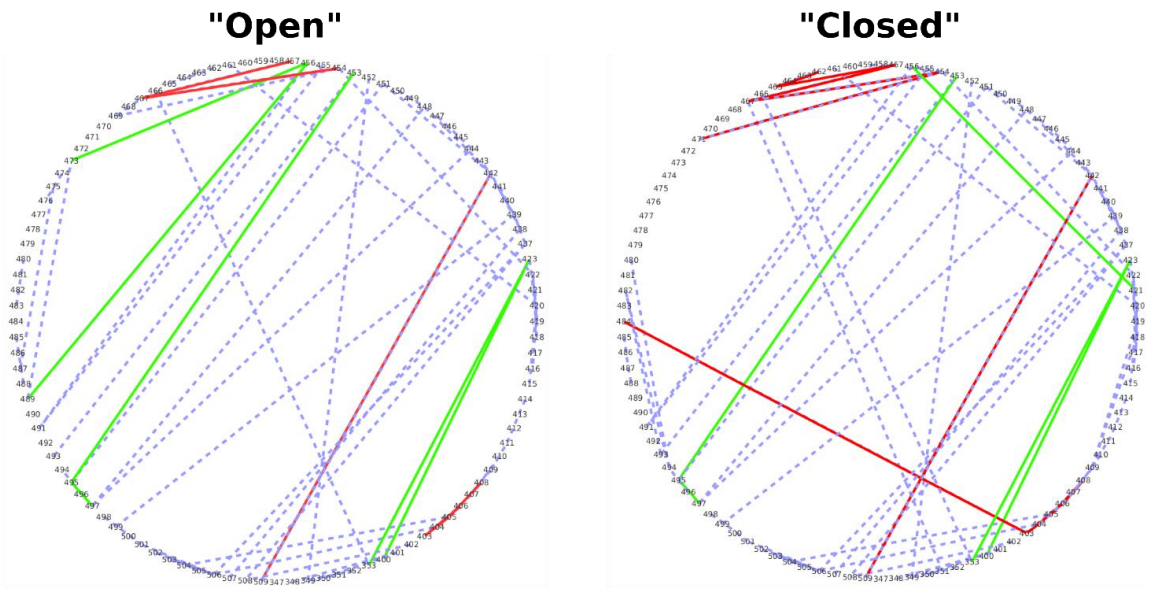


Figure S3. Residue interaction networks (RINs) for the "open" and "closed" conformations of *wt* RBD. RINs determined using RIP-MD for the 5000 lowest energy conformations obtained for the most populated "open" and "closed" basins of *wt* RBD. Hydrogen bonds, salt bridges and pi-pi interactions are shown in blue, red and green, respectively.

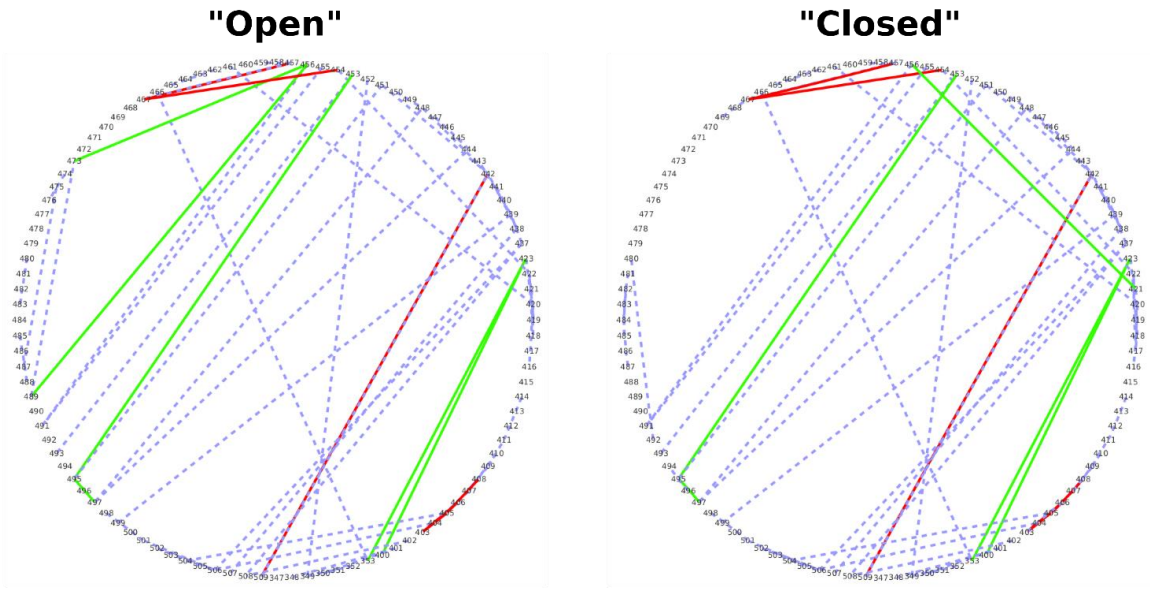


Figure S4. Residue interaction networks (RINs) for the "open" and "closed" conformations of alpha variant. RINs determined using RIP-MD for the 5000 lowest energy conformations obtained for the most populated "open" and "closed" basins of alpha RBD. Hydrogen bonds, salt bridges and pi-pi interactions are shown in blue, red and green, respectively.

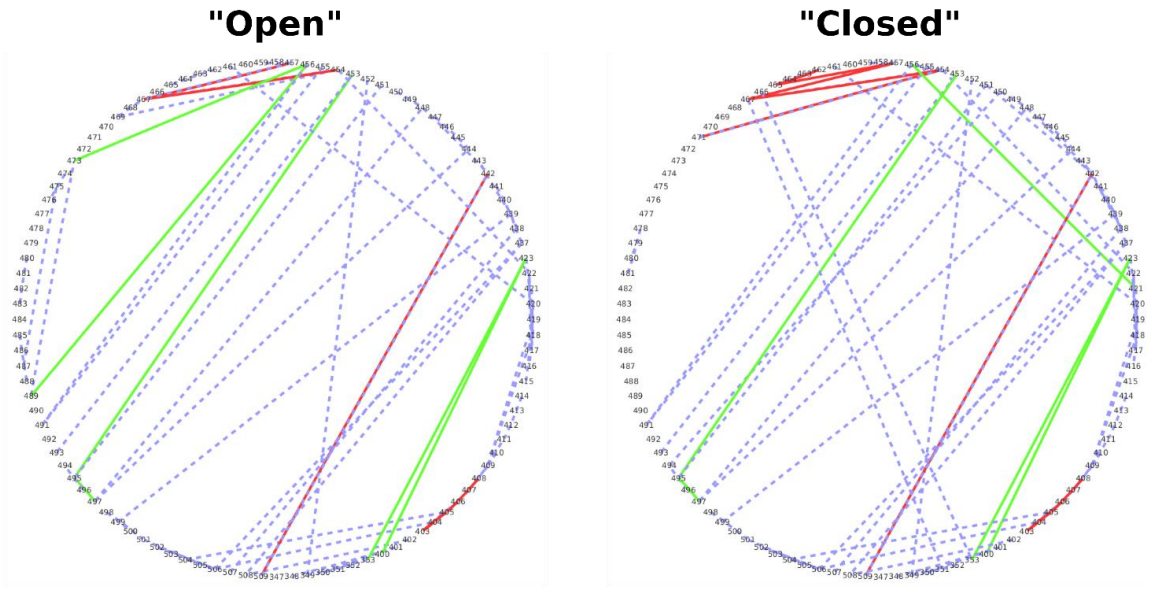


Figure S5. Residue interaction networks (RINs) for the "open" and "closed" conformations of beta variant. RINs determined using RIP-MD for the 5000 lowest energy conformations obtained for the most populated "open" and "closed" basins of beta RBD. Hydrogen bonds, salt bridges and pi-pi interactions are shown in blue, red and green, respectively.

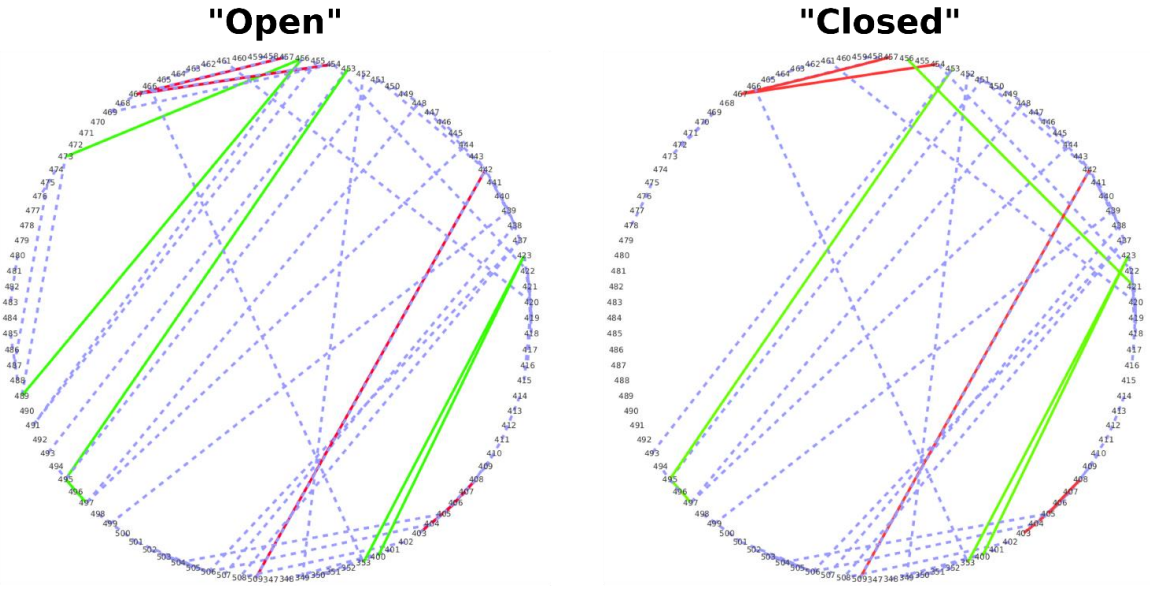


Figure S6. Residue interaction networks (RINs) for the "open" and "reversed" conformations of delta variant. RINs determined using RIP-MD for the 5000 lowest energy conformations obtained for "open" and "reversed" basins of alpha RBD. Hydrogen bonds, salt bridges and pi-pi interactions are shown in blue, red and green, respectively.

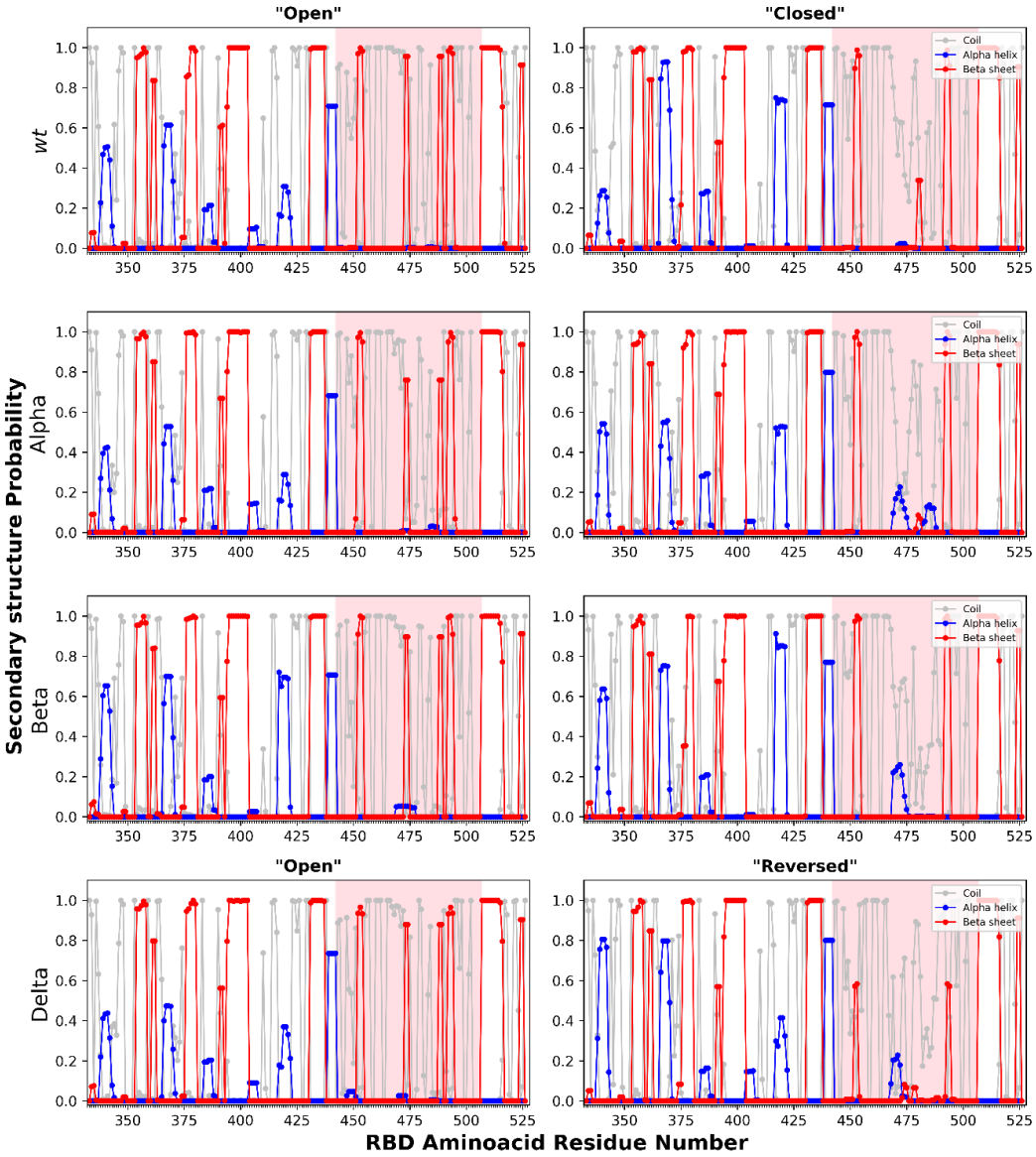


Figure S7. Secondary structure of *wt*, alpha, beta and delta SARS-CoV-2 RBD simulated in water. Probability of coil, α-helix and β-sheet secondary structures was obtained using the GROMACS tool gmx do_dssp[1] for both conformations ("open"/"closed " and "open"/"reversed") of all four variants. RBM region is highlighted in pink.

Table S2. Surface Accessible Surface Area (SASA) analysis of SARS-CoV-2 RBD in water. SASA values were calculated using the GROMACS tool gmx_sasa[1] for the whole trajectory (Entire trajectory) and for the two major conformations ("open"/"closed " and "open"/"reversed"). Results were also divided in the contribution of hydrophobic atoms, which are the ones with charges [-0.2, 0.2], and hydrophilic, those outside of this range. 95 % Confidence intervals (CI) were calculated with bootstrap resampling.

Variant	Protein SASA (nm^2)	Hydrophobic atoms SASA (nm^2)	Hydrophilic atoms SASA (nm^2)
	Average ± CI (95%)	Average ± CI (95%)	Average ± CI (95%)
WT			
Average	$107.72^{+0.07}_{-0.07}$	$55.07^{+0.05}_{-0.05}$	$52.62^{+0.05}_{-0.05}$
"Open"	$109.08^{+0.05}_{-0.05}$	$55.11^{+0.04}_{-0.04}$	$53.97^{+0.04}_{-0.05}$
"Closed"	$105.52^{+0.05}_{-0.05}$	$53.68^{+0.04}_{-0.04}$	$51.57^{+0.04}_{-0.04}$
Alpha			
Average	$108.16^{+0.04}_{-0.04}$	$55.13^{+0.03}_{-0.03}$	$53.04^{+0.03}_{-0.03}$
"Open"	$108.62^{+0.06}_{-0.06}$	$55.26^{+0.04}_{-0.04}$	$53.36^{+0.04}_{-0.04}$
"Closed"	$102.99^{+0.07}_{-0.07}$	$52.69^{+0.04}_{-0.04}$	$50.31^{+0.05}_{-0.05}$
Beta			
Average	$108.10^{+0.01}_{-0.01}$	$54.95^{+0.01}_{-0.01}$	$53.15^{+0.01}_{-0.01}$
"Open"	$108.58^{+0.06}_{-0.06}$	$53.70^{+0.04}_{-0.04}$	$53.88^{+0.04}_{-0.04}$
"Closed"	$107.65^{+0.08}_{-0.08}$	$56.32^{+0.05}_{-0.05}$	$51.33^{+0.05}_{-0.06}$
Delta			
Average	$109.62^{+0.11}_{-0.11}$	$55.46^{+0.08}_{-0.08}$	$54.16^{+0.08}_{-0.07}$
"Open"	$109.58^{+0.05}_{-0.05}$	$54.87^{+0.04}_{-0.04}$	$54.70^{+0.04}_{-0.04}$
"Reversed"	$106.08^{+0.07}_{-0.07}$	$53.65^{+0.04}_{-0.04}$	$52.43^{+0.04}_{-0.04}$

Table S3. Compilation of ACE2-RBD binding kinetics data from recent studies. Kinetic parameters of ACE2 binding to *wt*, alpha, beta and delta RBD/Spike variants data obtained from SPR and BLI[2–8] .

			WT			Alpha			Beta			Delta		
Reference	RBD/Spike	Technique	K_{on} (M^{-1} s^{-1})	K_{off} (s^{-1})	K_d (nM)	K_{on} (M^{-1} s^{-1})	K_{off} (s^{-1})	K_d (nM)	K_{on} (M^{-1} s^{-1})	K_{off} (s^{-1})	K_d (nM)	K_{on} (M^{-1} s^{-1})	K_{off} (s^{-1})	K_d (nM)
McCallum et al. 2021 [2]	RBD	SPR	7.70×10^4	6.70×10^{-3}	78	7.50×10^4	1.20×10^{-3}	15	-	-	-	5.90×10^4	4.30×10^{-3}	63
Tian et al. 2021 [3]	RBD	SPR	2.50×10^4	2.10×10^{-4}	8.3	7.50×10^4	0.37×10^{-4}	0.5	6.40×10^4	0.30×10^{-4}	0.5	-	-	-
Laffeber et al. 2021 [4]	RBD	SPR	4.50×10^5	7.80×10^{-3}	17	5.70×10^5	1.30×10^{-3}	2.4	7.60×10^5	4.30×10^{-3}	5.8	-	-	-
Supasa et al. 2021 [5]	RBD	SPR	3.88×10^4	2.92×10^{-3}	75.1	6.38×10^4	6.85×10^{-4}	10.7	-	-	-	-	-	-
Wirnsberger et al. 2021 [6]	RBD	SPR	6.86×10^5	11.0×10^{-3}	16.2	4.36×10^5	1.59×10^{-3}	3.72	6.86×10^5	4.42×10^{-3}	6.5	10.1×10^5	8.23×10^{-3}	8.61
de Souza et al. 2021 [7]	RBD	SPR	0.90×10^6	91.6×10^{-4}	10.3	1.30×10^6	15.5×10^{-4}	1.2	1.20×10^6	39.4×10^{-4}	3.3			
de Souza et al. 2021 [7]	Spike	SPR	0.09×10^6	5.80×10^{-4}	6.4	0.10×10^6	1.70×10^{-4}	0.1	0.30×10^6	3.00×10^{-4}	0.3			
Saville et al. 2021 [8]	Spike	BLI	1.40×10^5	7.09×10^{-4}	5.06	-	-	-	-	-	-	1.51×10^5	4.01×10^{-4}	2.65
Yang et al. 2021 [9]	Spike	BLI	3.43×10^4	10.6×10^{-5}	3.1	3.86×10^4	5.25×10^{-5}	1.36	7.32×10^4	1.79×10^{-5}	0.25	4.39×10^4	1.71×10^{-5}	0.39

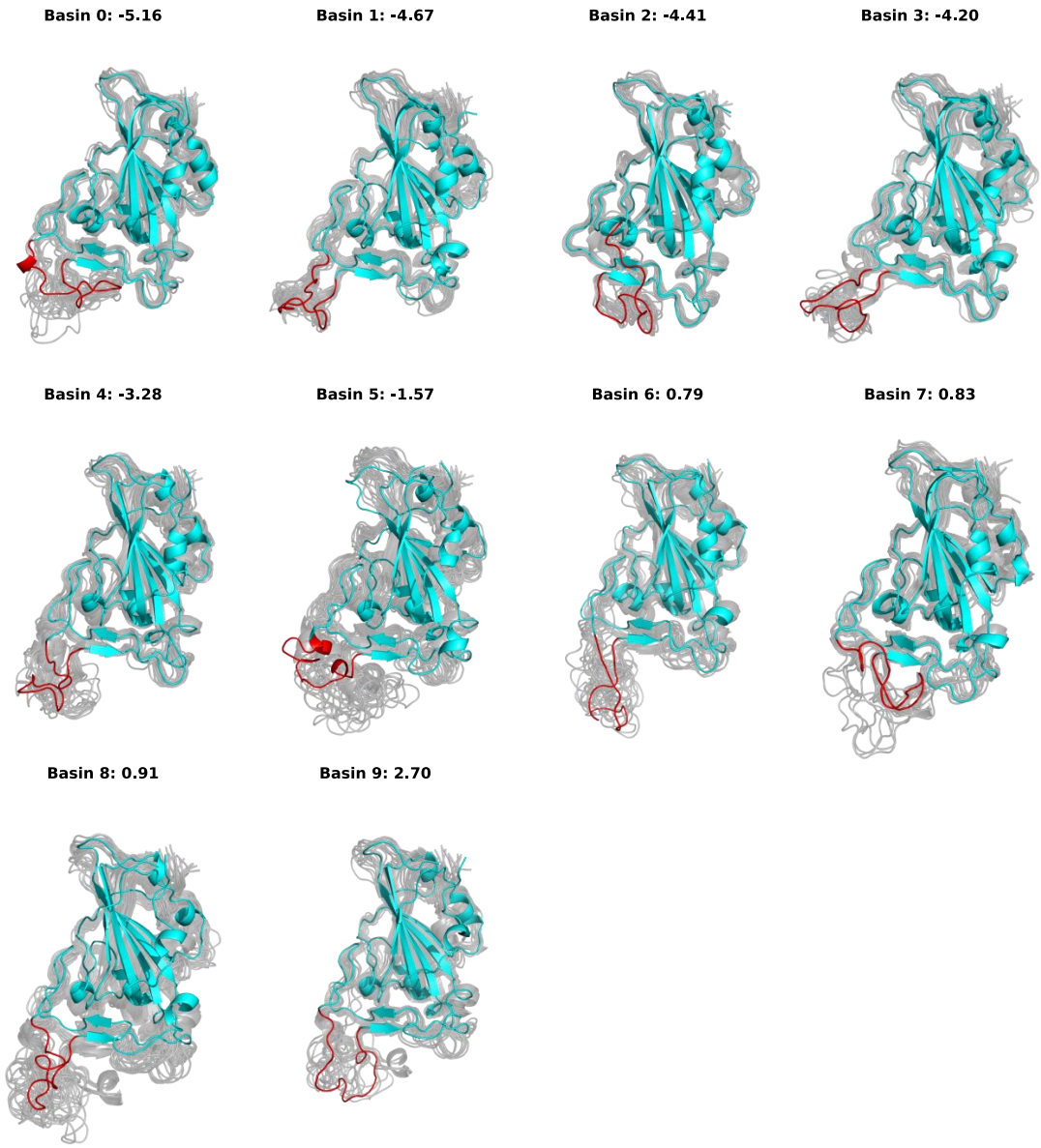


Figure S8. Snapshots representative of all *wt* RBD PCA basins. The structures corresponding to the free energy minima of all conformational basins are represented in blue, with the ridge region highlighted in red, together with structures sampled from the same basin (background, gray colored).

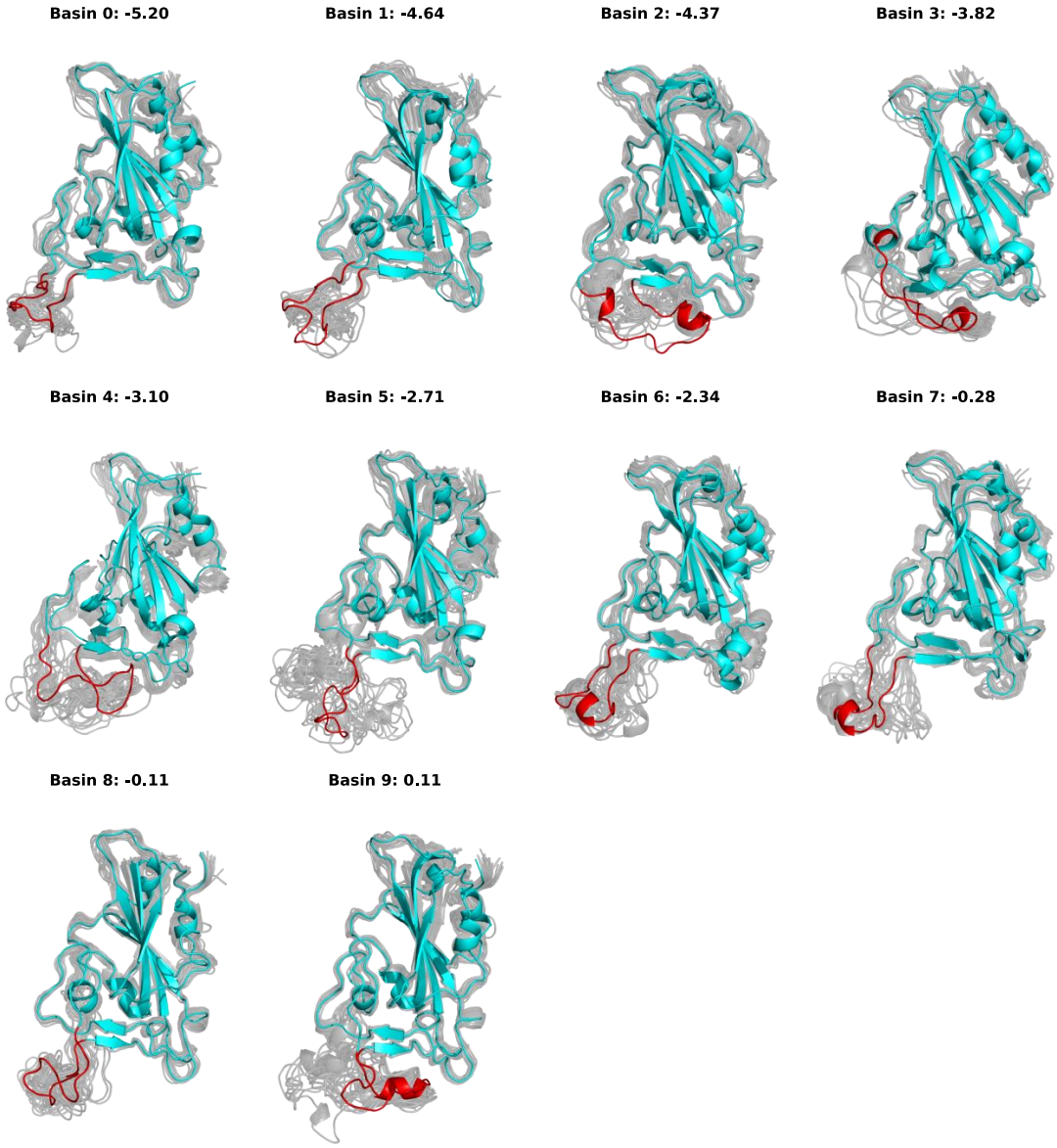


Figure S9. Structures representative of all alpha RBD PCA basins. The structures corresponding to the free energy minima of all conformational basins are represented in blue, with the ridge region highlighted in red, together with structures sampled from the same basin (background, gray colored).

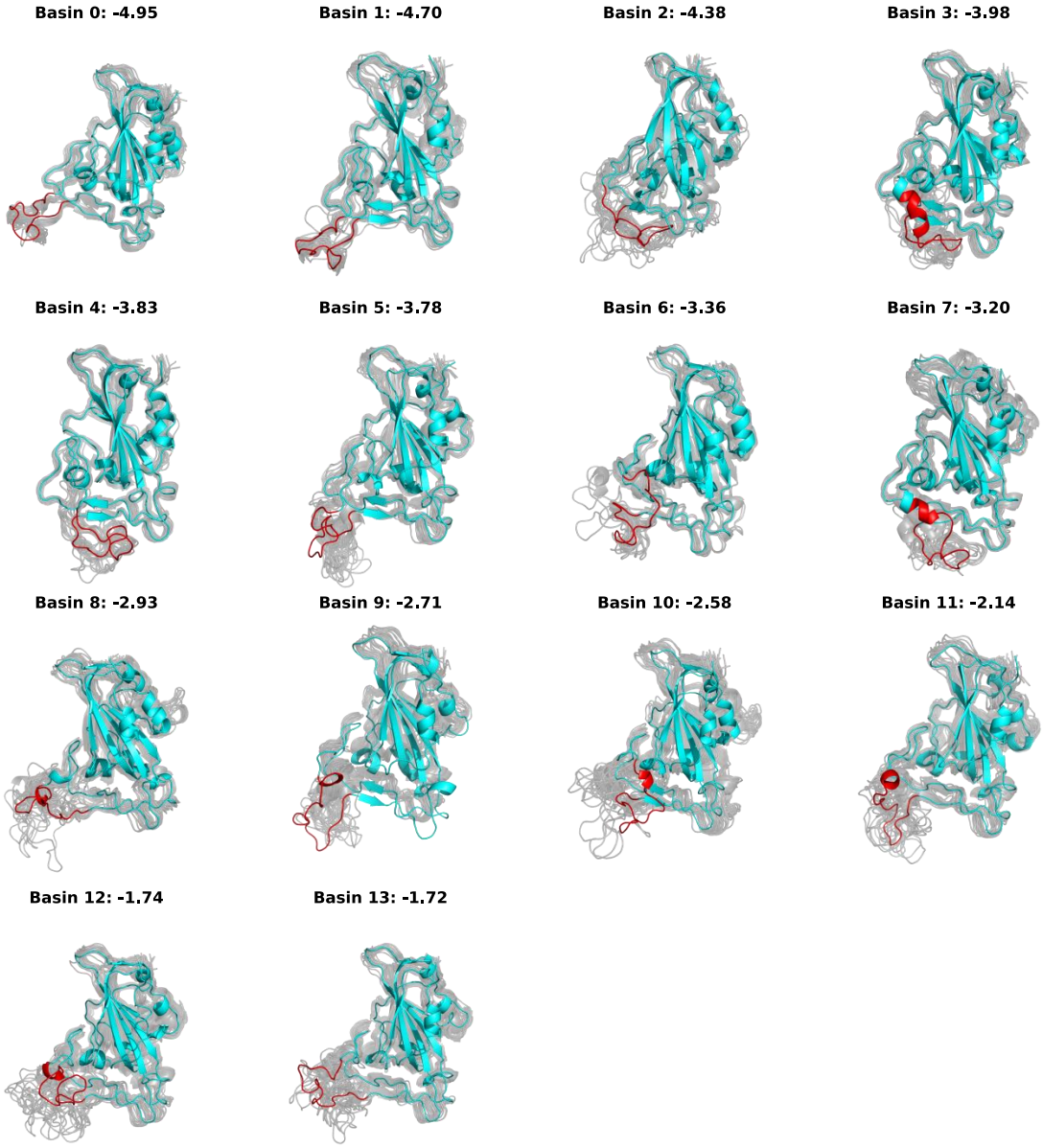


Figure S10. Structures representative of all beta RBD PCA basins. The structures corresponding to the free energy minima of all conformational basins are represented in blue, with the ridge region highlighted in red, together with structures sampled from the same basin (background, gray colored).

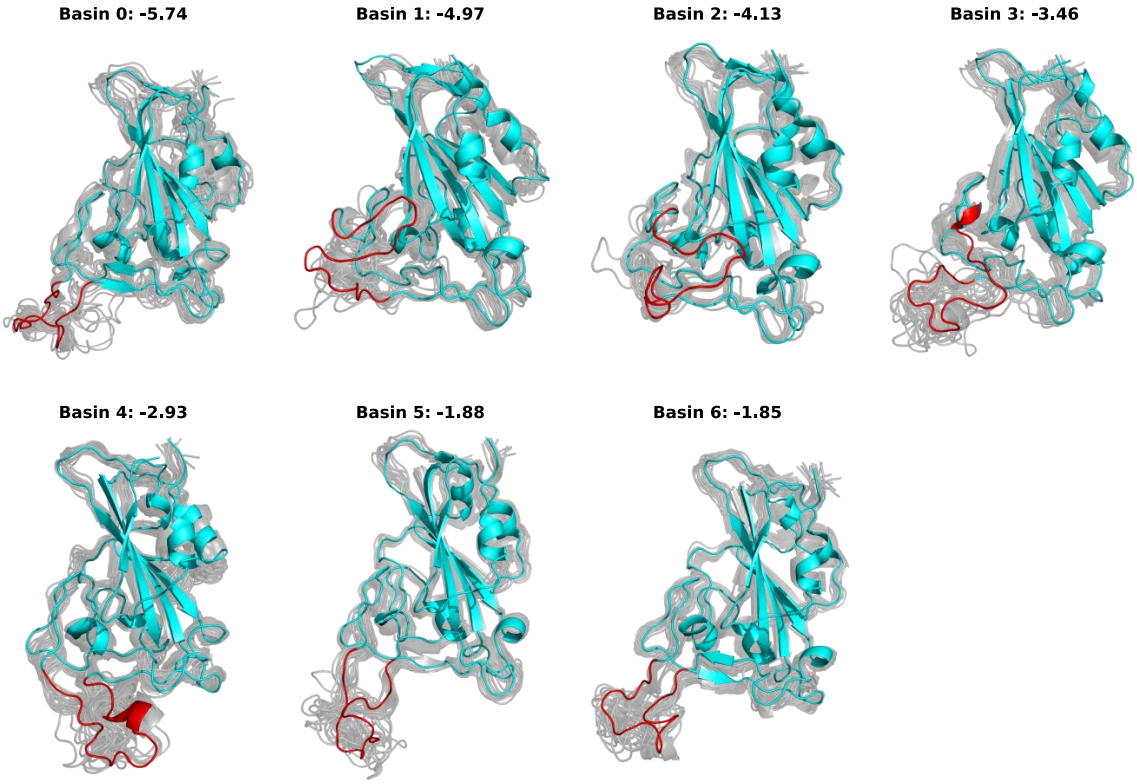


Figure S11. Structures representative of all delta RBD PCA basins. The structures corresponding to the free energy minima of all conformational basins are represented in blue, with the ridge region highlighted in red, together with structures sampled from the same basin (background, gray colored).

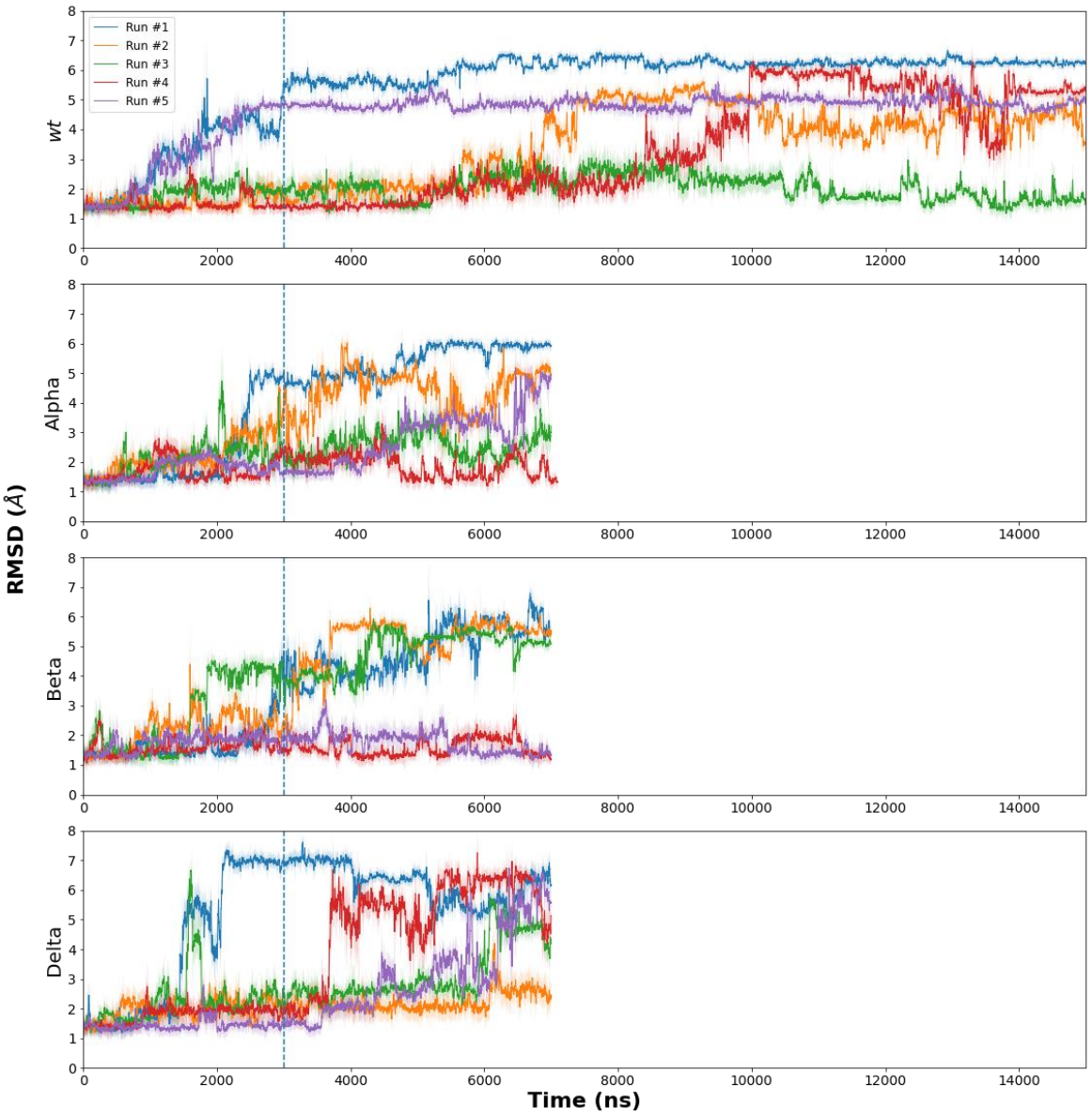


Figure S12. RBD Cα root-mean-square deviation (RMSD) moving average in solution. Data shown for the five replicas for each variant tested. Cα were fitted against the RBD X-ray structure from PDB ID: 6M0J. The moving average was calculated using the neighboring 50 frames. The first 3 μs of simulation were used for equilibration (blue dashed line) and the remaining frames were used for further PCA and RIN analysis.

REFERENCES

(1) 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.

(2) McCallum, M.; Walls, A. C.; Sprouse, K. R.; Bowen, J. E.; Rosen, L.; Dang, H. V.; deMarco, A.; Franko, N.; Tilles, S. W.; Logue, J.; Miranda, M. C.; Ahlrichs, M.; Carter, L.; Snell, G.; Pizzuto, M. S.; Chu, H. Y.; Voorhis, W. C. Van; Corti, D.; Veesler, D. Molecular Basis of Immune Evasion by the Delta and Kappa SARS-CoV-2 Variants. *bioRxiv* **2021**, 2021.08.11.455956.

(3) Tian, F.; Tong, B.; Sun, L.; Shi, S.; Zheng, B.; Wang, Z.; Dong, X.; Zheng, P. N501Y Mutation of Spike Protein in SARS-CoV-2 Strengthens Its Binding to Receptor ACE2. *Elife* **2021**, *10*.

(4) Laffeber, C.; de Koning, K.; Kanaar, R.; Lebbink, J. H. G. Experimental Evidence for Enhanced Receptor Binding by Rapidly Spreading SARS-CoV-2 Variants. *J. Mol. Biol.* **2021**, *433* (15), 167058.

(5) Supasa, P.; Zhou, D.; Dejnirattisai, W.; Liu, C.; Mentzer, A. J.; Ginn, H. M.; Zhao, Y.; Duyvesteyn, H. M. E.; Nutalai, R.; Tuekprakhon, A.; Wang, B.; Paesen, G. C.; Slon-Campos, J.; López-Camacho, C.; Hallis, B.; Coombes, N.; Bewley, K. R.; Charlton, S.; Walter, T. S.; Barnes, E.; Dunachie, S. J.; Skelly, D.; Lumley, S. F.; Baker, N.; Shaik, I.; Humphries, H. E.; Godwin, K.; Gent, N.; Sienkiewicz, A.; Dold, C.; Levin, R.; Dong, T.; Pollard, A. J.; Knight, J. C.; Klenerman, P.; Crook, D.; Lambe, T.; Clutterbuck, E.; Bibi, S.; Flaxman, A.; Bittaye, M.; Belij-Rammerstorfer, S.; Gilbert, S.; Hall, D. R.; Williams, M. A.; Paterson, N. G.; James, W.; Carroll, M. W.; Fry, E. E.; Mongkolsapaya, J.; Ren, J.; Stuart, D. I.; Screaton, G. R. Reduced Neutralization of SARS-CoV-2 B.1.1.7 Variant by Convalescent and Vaccine Sera. *Cell* **2021**, *184* (8), 2201-2211.e7.

(6) Wirnsberger, G.; Monteil, V.; Eaton, B.; Postnikova, E.; Murphy, M.; Braunsfeld, B.; Crozier, I.; Kricek, F.; Niederhöfer, J.; Schwarzböck, A.; Breid, H.; Jimenez, A. S.; Bugajska-Schretter, A.; Dohnal, A.; Ruf, C.; Gugenberger, R.; Hagelkruys, A.; Montserrat, N.; Holbrook, M. R.; Oostenbrink, C.; Shoemaker, R. H.; Mirazimi, A.; Penninger, J. M. Clinical Grade ACE2 as a Universal Agent to Block SARS-CoV-2 Variants. *bioRxiv* **2021**, 2021.09.10.459744.

(7) Souza, A. S. de; Amorim, V. M. de F.; Guardia, G. D. A.; Santos, F. R. C. dos; Santos, F. F. dos; Souza, R. F. de; Juvenal, G. de A.; Huang, Y.; Ge, P.; Jiang, Y.; Paudel, P.; Ulrich, H.; Galante, P. A. F.; Guzzo, C. R. Molecular Dynamics Analysis of Fast-Spreading Severe Acute Respiratory Syndrome Coronavirus 2 Variants and Their Effects in the Interaction with Human Angiotensin-Converting Enzyme 2. *bioRxiv* **2021**, 2021.06.14.448436.

(8) Saville, J. W.; Mannar, D.; Zhu, X.; Srivastava, S. S.; Berezuk, A. M.; Demers, J.-P.; Zhou, S.; Tuttle, K. S.; Sekirov, I.; Kim, A.; Li, W.; Dimitrov, D. S.; Subramaniam, S. Structural and Biochemical Rationale for Enhanced Spike Protein Fitness in Delta and Kappa SARS-CoV-2 Variants. *bioRxiv* **2021**, 2021.09.02.458774.