Role of noncovalent interactions in protein peripheral membrane binding

Computational perspectives

Hanif Muhammad Khan

Dissertation for the degree of philosophiae doctor (PhD)
at the University of Bergen

2016

Dissertation date: 19.02.2016

Year: 2016

Title: Role of noncovalent interactions in protein peripheral membrane binding

Computational perspectives

Author: Hanif Muhammad Khan

Print: AIT OSLO AS / University of Bergen

Scientific environment

The research presented in this thesis was carried out during my employment as a research fellow (1st February 2013- 31st January 2016) at the Department of Molecular Biology (MBI), University of Bergen (UiB). Both MBI and the Computational Biology Unit (CBU), at the Department of Informatics, UiB hosted me during this time period. I was supervised by Prof. Nathalie Reuter, and co-supervised by Prof. Inge Jonassen from CBU and Prof. Mathias Ziegler from MBI. These works were part of a research project funded by the Norwegian Research Council (FRIMEDBIO #214167). NOTUR (Norwegian metacenter for computational science) provided computational resources.

I have had the opportunity to closely collaborate with Dr. Cédric Grauffel, Academia Sinica, Taiwan; Prof. Anne Gershenson, University of Massachusetts-Amherst, U.S.A.; Prof. Mary F. Roberts, Boston College, U.S.A.; Prof. Ria Broer and Dr. Remco W.A. Havenith, University of Groningen, Netherlands; Prof. Alexander D. MacKerell Jr., University of Maryland, U.S.A. I have also had the opportunity to collaborate within the group with Dr. Edvin Fuglebakk and Dr. Anne-Sophie Schillinger.

I was also part of the Molecular and Computational Biology Research School (MCB) and of the Norwegian Graduate School in Structural Biology (BioStruct).

Acknowledgements

I would like to thank Nathalie for supervising me and providing critical feedback particularly on my writing. Thanks goes to Reuter group members for interesting discussions and their relentless help. I would like to express my gratitude to my office mates for their patience with me. I am also grateful to all those who read the thesis drafts and provided very useful corrections and suggestions.

Thanks to Cédric for his continuous help. I would also like to thank Anne and Mary for very interesting discussions, Ria and Remco for hosting me once again in Groningen.

Last but not the least, I would like to thank my family for their encouragement and support.

Abstract

Noncovalent forces are important driving forces in nature particularly in biology, and they dictate many biological processes including the binding of peripheral protein to the cell membrane. The widely acknowledged models describe this process as electrostatics driven membrane adsorption followed by short-range protein-lipid interactions i.e. hydrogen bonds, hydrophobic interactions. Some of the key elements in such models are: clusters of basic residues are essential for electrostatic adsorption, and basic residues contribute equally to the membrane binding. Nevertheless, none of these models account for the role of cation-π interactions in membrane binding. With selected protein candidates, we further explore these models and work towards a generalized description of protein peripheral binding to membranes in terms of noncovalent forces.

Our investigation highlights the limitations of these existing descriptions. We demonstrate that the requirement of having a cluster of basic residues is not essential. Further, we show that the contributions of basic residues are distance dependent. In other words, their localization in the membrane-water interface determines their strength and hence is not equal. We also establish the role of tyrosine-choline cation-π interactions in membrane binding of peripheral proteins.

We explore in detail the nature of tyrosine-choline mediated cation-π interactions using high-level quantum mechanical calculations. Later, this information is used to improve the description of cation-π interactions in molecular simulation models. These improvements of force field parameters are further tested using molecular dynamics simulations.

Finally, we used this information to build an interaction diagram that can be used to better describe the binding of peripheral proteins to the cell membrane. Future testing and the generalization of this diagram will further establish this as a common model.

List of publications

List of publications and manuscripts included in thesis

1. A.-S. Schillinger, C. Grauffel, H.M. Khan, Ø. Halskau, N. Reuter, "Two homologous neutrophil serine proteases bind to POPC vesicles with different affinities: When aromatic amino acids matter", *Biochimica et Biophysica Acta (BBA) – Biomembranes,* vol. 1838, pp. 3191-3202, 2014.

2. B. Yang, M. Pu, H.M. Khan, L. Friedman, N. Reuter, M.F. Roberts, A. Gershenson, "Quantifying Transient Interactions between Bacillus Phosphatidylinositol-Specific Phospholipase-C and Phosphatidylcholine-Rich Vesicles", *Journal of the American Chemical Society,* vol. 137, pp. 14-17, 2015.

3. H.M. Khan, T. He, E. Fuglebakk, C. Grauffel, B. Yang, M.F. Roberts, A. Gershenson, N. Reuter, "A role for weak electrostatic interactions in peripheral membrane protein binding", to be submitted.

4. H.M. Khan, C. Grauffel, R. Broer, A.D. MacKerell Jr., R.W.A. Havenith, N. Reuter, "Improving the force field description of tyrosine-choline cation-π interactions: QM investigation of phenol-$N(Me)_4^+$ interactions", to be submitted.

Related manuscript not included in thesis

5. N.L. Chon, J.R. Osterberg, J. Henderson, H.M. Khan, N. Reuter, J.D. Knight, H. Lin, "Membrane Docking of the Synaptotagmin 7 C2A Domain: Computation Reveals Interplay between Electrostatic and Hydrophobic Contributions", *Biochemistry*, vol. 54, pp. 5696-5711, 2015.

Contents

1. Noncovalent interactions

Noncovalent interactions are ubiquitous in nature. Although they are not as strong as covalent forces, which originate from electron sharing and pairing; they play a very important role in physical processes such as molecular recognition[1]. A textbook definition and classification of noncovalent interactions can be found in the Physical Chemistry book by Atkins and Paula[1]. Lodish et al. categorized "noncovalent bonds" as four distinct types: hydrogen bonds, ionic interactions (attractions between oppositely charged ions), van der Waals interactions (caused by transient dipoles), and hydrophobic bonds (between nonpolar molecules)[2]. At a quick glance, this listing looks complete. However, electrostatic repulsion and exchange-repulsion are overlooked, while hydrophobic bonds are a special case of van der Waals interactions. Leckband and Israelachvili tried to probe many possible interactions arising *in vivo* and in laboratory experiments between molecules, surfaces and biomaterials[3]. The authors attempted to provide a detailed generalization of noncovalent forces in biology[3]. Nevertheless, the cation-π interaction is overlooked in all of them and emerged in 1990's as an important noncovalent force in biology[4] and particularly relevant for protein-lipid interactions[5-7].

Interestingly, these interactions can all be decomposed into some fundamental components. It is relatively easy to dissect them in physically meaningful components when we talk about a complex made of dimers in gas phase. When a complex is made of many-bodies (multimers), it is safe to say that it is almost impossible to dissect these components and evaluate these contributions. In solvent or in condensed phase, these interactions are not essentially the same as they are in the gas phase. The medium, be it a solvent or a solid phase, has some unique features (e.g. relative permittivity), which will in turn affect these components differently.

Many biological processes take place in a cellular environment. The cellular environment is highly anisotropic, dissolved in water (if not all, most of it) and contains significant amounts of biomolecules (proteins, carbohydrates, lipids, and nucleic acids) along with ions. Hence, the properties in different cell compartments

are different. These factors make the generalization/grouping of these noncovalent interactions for different biological processes rather difficult. In the following sections, we will first present the fundamental components of noncovalent interactions followed by how they are affected by the environments.

1.1 Fundamental components of noncovalent interactions

Noncovalent interactions are composed of four fundamental components. They are namely electrostatics, exchange, induction, and dispersion. All the noncovalent interactions can be dissected in these four physically meaningful components. For a detailed overview on this topic, the pioneering works by others are highly recommended[8-17]. Let us start with the simple case of two monomers in the gas phase.

1.1.1 Electrostatics

Electrostatics is the interaction between the charge distributions of the molecules. In other words, it is the interaction between the permanent multipole moments of the molecules[18]. The permanent multipole moments are monopole, dipole, quadrupole, octupole, and so on (Figure 1.1). The energy of these interactions between charge distributions is distance-dependent ($\propto 1/r$). The electrostatic interaction between a set of charges (+e,-e or +e,+e/-e,-e) is long-range in gas phase, and electrostatics can be both attractive of repulsive (Figure 1.2). This can range up to 1000 Å or more[19].

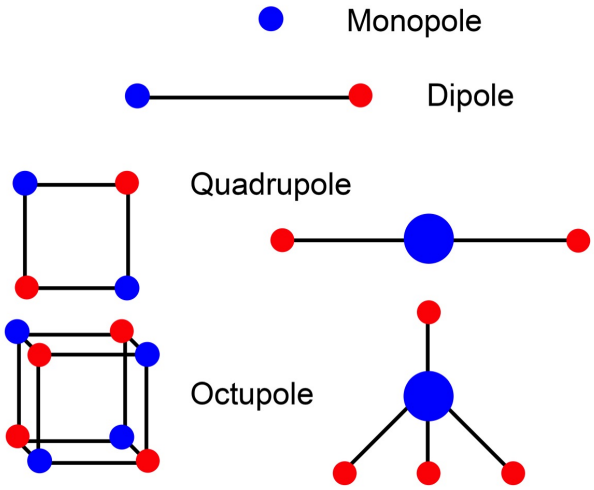


Figure 1.1: Electronic multipole moments in molecules (adapted from [1]).

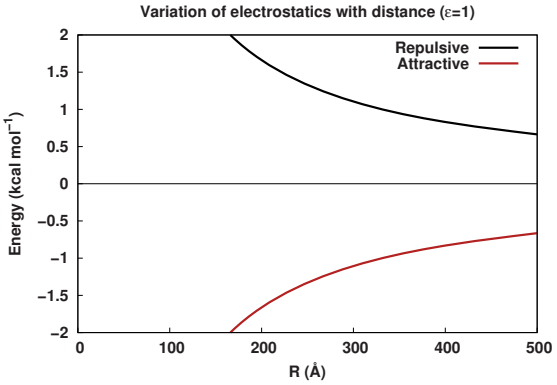


Figure 1.2: Variation of electrostatic interactions between a set of charges. Electrostatic interactions decay very slowly as a function of the distance between particles in gas phase and can be active at very long distances with a contribution of around 1 kcal/mol at 500 Å in the case depicted here for charges of similar signs (+e,+e/-e,-e) and opposite charges (+e,-e).

Here, the term electrostatics is used also to refer to charge-charge, charge-dipole, dipole-dipole interactions [18] (Figure 1.3). Hydrogen bonds and salt bridges/ionic bonds also fall in this category as they are dominated by electrostatics. One can turn the argument around and say, they are special cases of electrostatics.

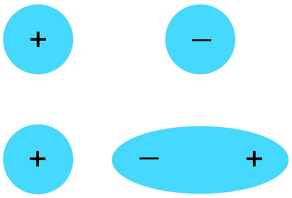


Figure 1.3: Electrostatic interactions between two charges (top), or a charge and a dipole (bottom).

1.1.2 Exchange

Exchange is also known as exchange-repulsion and is not to be confused with electrostatic repulsion. This force is always present between the two interacting monomers when they are close to each other. There is no classical analogue to describe such interaction. It originates mainly from the overlap of molecular orbitals when two monomers are very close, the electrons in the molecular orbitals rearrange themselves following Pauli's exclusion principle. Exchange has an exponential distance-dependence ($\propto 1/e^{r}$).

1.1.3 Induction

Induction is also known as polarization. Induction is the attractive force between the permanent multipole moments of one monomer and the induced multipole moments of other monomer (Figure 1.4). Induction is related to the polarizability of the molecule and the induced moment M can be expressed as[20]:

$$M = \alpha F \qquad (1.1)$$

where α is the polarizability tensor, and F is the external electric field. In case of a dimer, F is the permanent multipole moments of one monomer that generates the external electric field. Often, charge transfer is also attributed to polarization[8].

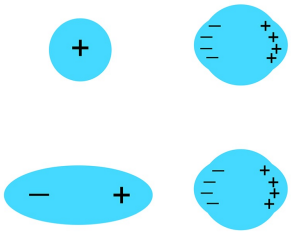


Figure 1.4: Example of induction. Generation of induced dipole moments in a molecule due to the permanent multipole moments of the approaching molecule.

1.1.4 Dispersion

Dispersion is the attractive force between two molecules due to the instantaneous charge fluctuations. In other words, it is the attractive force between the instantaneous multipole moments generated between the monomers. Dispersion is a short-range phenomenon (compared to electrostatics), and is always present although the energetic contribution is weak compared to the other forces. Dispersion is important for particularly the nonpolar molecules (i.e. noble gases, hydrocarbons). The leading contribution to dispersion comes from the generated instantaneous dipole moment and has a distance dependence $\propto 1/r^6$ (Figure 1.5). The higher order poles can contribute and have dependence as $\propto 1/r^8$, $1/r^{10}$, and so on[21].

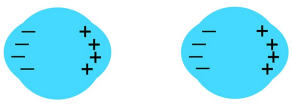


Figure 1.5: Example of dispersion: generation of induced dipole moments in two nonpolar molecules.

1.2 Modeling noncovalent interactions in molecular simulations

Electrostatics is modeled using Coulomb's Law in Class I additive force fields (e.g. additive CHARMM-ff [22]). In the Class I additive force field, point charge (also known as partial atomic charge) description is used to describe electrostatics[22]. The point charge description of electrostatics has a number of limitations [8], but is computationally inexpensive and works remarkably well for many cases. Although exchange has an exponential distance dependence, due to the ease of computation it is often modeled with a ($\propto 1/r^{12}$) term (e.g. CHARMM-ff [23]). This is actually the repulsive part of the Lennard-Jones (LJ) potential. Also, it is not uncommon to use the Buckingham potential that has an exponential term. Dispersion is modeled with the attractive part of the LJ potential[22]. A detailed list of functional forms used to model noncovalent interactions in different additive force fields can be found in the book by Jensen[19]. Induction is often not treated in classical molecular simulations. However, recent advances of polarizable force fields make this possible. There are several ways to model induction. Some of those are: fluctuating charge model[24], multipole moment methods (AMOEBA)[25], and Drude polarizable model (using drude particles- as implemented in CHARMM-ff)[26-30]. Recently, Baker published an excellent review article summarizing the current status of polarizable force fields for biomolecules[31].

1.3 Many-body interactions

As the relationships for a dimer have been established, a many-body (more than two molecules) problem can be explored based on these relationships. A and B are charged, C and D are nonpolar molecules (Figure 1.6). One might wonder why there exists electrostatics interaction between the nonpolar molecules in the figure below. This is due to the interaction between electrons and the nuclei of the participating molecules[16]. Even for dispersion dominated complex such as noble gases (Ar dimer), hydrocarbons: electrostatics exist. However, the contribution is substantially lower than that from dispersion[16].

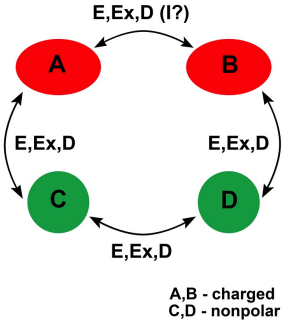


Figure 1.6: Case of a many-body interaction problem and resulting noncovalent components. E, Ex, D, and I stands for electrostatics, exchange, dispersion, and induction, respectively.

The net resulting force will determine the fate of the complex. However, even in gas phase separating these individual contributions due to the presence of another monomer is cumbersome. One knows the contribution of A on B; but due to the presence of C and D, it is difficult to assess in principle how much this contribution will be affected. This is due to the many-body non-additive effects[9]. To accurately treat such problem, one need to add the many-body (three-body, four-body) corrections[9].

1.4 Cooperativity measures

When multiple noncovalent interactions exist simultaneously in a system, if they strengthen each other; they are said to be cooperative. If they weaken each other, they are said to be anti-cooperative. In a biological system, multiple noncovalent interactions act simultaneously.

Cooperative energy for a ternary system ($\Delta E_{coop}(T)$) can be calculated as the difference between the energy of the ternary complex and the energies of the binary complexes[32]. It can be written as follows,

$$\Delta E_{coop}(T) = E_{ABC}(T) - E_{AB}(T) - E_{BC}(T) - E_{CA}(T) + E_A(T) + E_B(T) + E_C(T) \quad (1.2)$$

For a quaternary complex, we can write similarly,

$$\Delta E_{coop}(Q) = E_{ABCD}(Q) - E_{AB}(Q) - E_{BC}(Q) - E_{CD}(Q) - E_{DA}(Q) - E_{CA}(Q) - E_{BD}(Q) + 2E_A(Q) + 2E_B(Q) + 2E_C(Q) + 2E_D(Q) \quad (1.3)$$

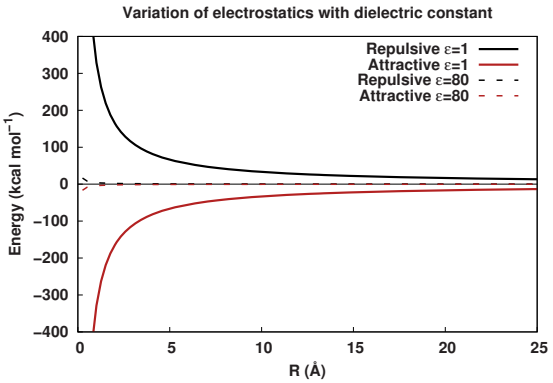
However, in a biological system of hundreds of thousands of molecules, to use cooperativity measure will surely be difficult if not impossible.

1.5 Biologically relevant cases: moving to the condensed phase

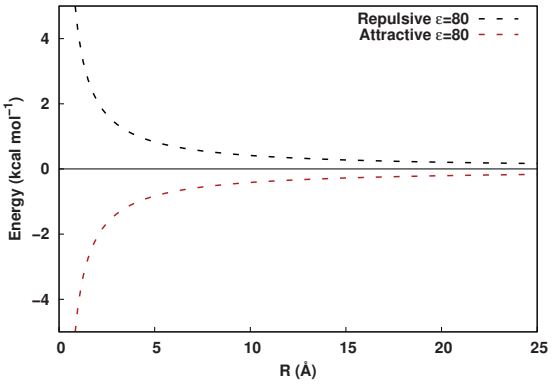
As stated earlier, the cellular environment is often approximated to a liquid phase. The solute-solvent and solvent-solvent interactions play a critical role. When the many monomers are replaced in the many-body problem by water and system is in the liquid phase, the problem becomes more complicated where solvent properties come into play. First, the peculiar property of water will be discussed and how that gives rise to two interesting phenomena: hydrophobic and hydrophilic interactions. Then the special property of membranes will be discussed in the cell membrane section.

Water is a polar solvent with a high dielectric constant (ε=78.5~80) that greatly affects electrostatic interactions by e.g. screening charges. This affects the electrostatic energies greatly (Figure 1.7), and decreases the effective regulatory

distance. In other words, in vacuum where a positive charge (+e) can interact with a negative charge (-e) being 500 Å apart, in water the charge effectively interacts within a shorter distance. Further effect of ionic solvent on charge screening will be discussed in the continuum electrostatics section.



(A)



(B)

Figure 1.7: Variation of electrostatics with the dielectric constant of the medium. (A) Gas phase vs water, (B) zoomed version in water, traces remain around 0.2 kcal/mol even at a distance of 25 Å.

1.6 Hydrophobic and Hydrophilic interactions

Due to the polarity of water, nonpolar molecules will not solubilize well in water; on the other hand, polar molecules easily dissolve in water. Hence, this leads to the origination of hydrophobic (water fearing) and hydrophilic (water loving) interactions. Hydrophobic molecules tend to form aggregates in water. Such as if one puts droplets of oil in water. This phenomenon is often entropy driven. It is difficult to quantify the contributions from hydrophobic and hydrophilic interactions either experimentally or computationally.

Molecules containing a hydrophobic and a hydrophilic part are called amphipathic and phospholipids are examples of such molecules[33].

2. Membranes, proteins, and their interactions

2.1 Cell membrane: its constituents and environment

2.1.1 Phospholipids

Membranes are mainly made of phospholipids[33]. Phospholipids are amphipathic molecules as stated in the previous chapter. They consist of hydrophobic acyl chains and a hydrophilic part made up of glycerol, phosphate, and the headgroup. The acyl chain length can vary from 14 to 24 carbon atoms and can be saturated or unsaturated[33]. The headgroups can be zwitterionic e.g. phosphatidylcholine (PC), phosphatidylethanolamine(PE); or negative e.g. phosphatidylserine (PS), phosphatidylglycerol (PG) (Figure 2.1).

Due to its amphipathic nature, lipids assemble spontaneously in polar solvent like water by exposing the polar portion outwards and the hydrophobic portions inwards. The assembly can take the form of a micelle or a bilayer depending on the shape of the lipids[33].

2.1.2 Bilayers and their properties

The properties along the bilayer dimensions are not uniform. The lateral dimensions are coupled and they act as a two dimensional fluid[33]. The lateral pressure profile is not uniform across the bilayer normal[34]. Also, the properties in the normal direction are anisotropic within that 35~40 Å special arrangement. The dielectric constant of this anisotropic part of the cell is essentially distinct (Figure 2.2). The dielectric constant of the membrane-water interface is lower than that of water and higher than that of the membrane core and has a value between 30 and 40[35, 36].

2.1.3 Membranes and membrane organization models

Membranes are the partitions of the cell and its compartments, typically made up of phospholipid bilayers. They separate the inside of the cell compartments from the outside environments. While the plasma membrane surrounds the whole cell, other

membranes delimit cellular organelles such as the Golgi apparatus, the mitochondria, the endoplasmic reticulum etc.

The phospholipid composition of cell membranes varies from organism to organism [33] and between cellular compartments. A comprehensive overview on such variation for eukaryotic cell membranes[37] and mammalian cell membranes[38] are provided by van Meer and co-workers.

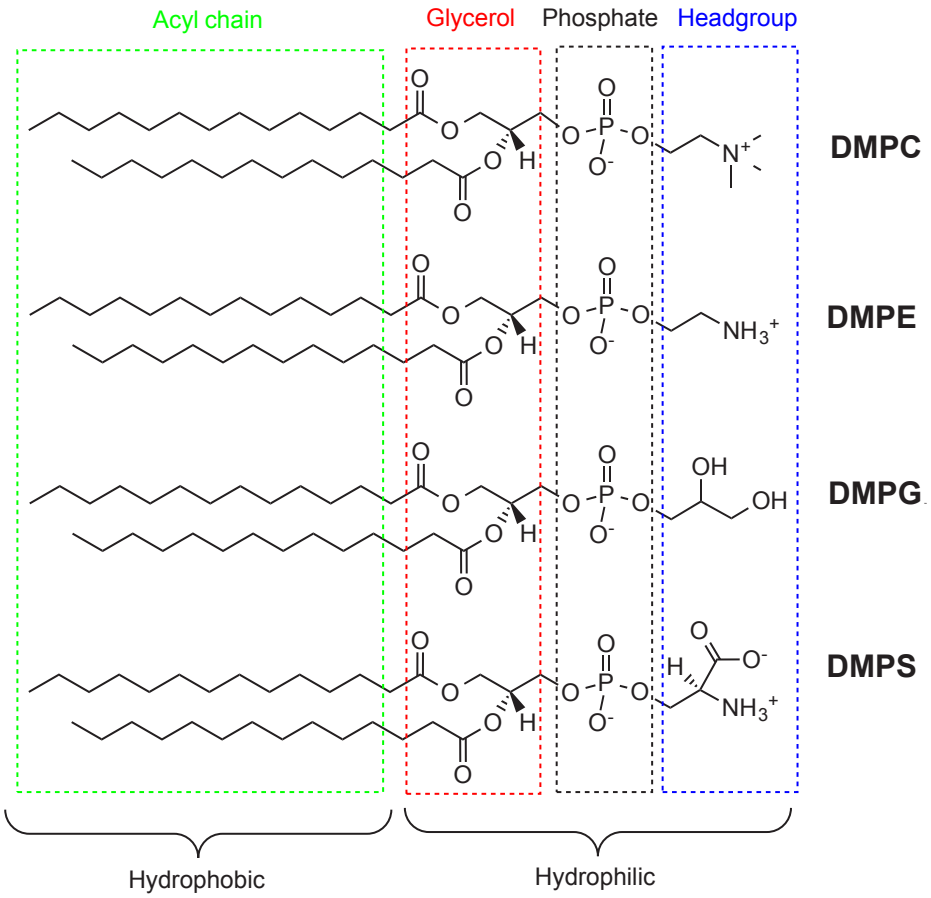


Figure 2.1: Four examples of phospholipids. They are: 1,2-dimyristoyl-sn-glycero-3-phosphocholine (DMPC), 1,2-dimyristoyl-sn-glycero-3-phosphoethanolamine (DMPE), 1,2-dimyristoyl-sn-glycero-3-phosphoglycerol (DMPG), and 1,2-dimyristoyl-sn-glycero-3-phosphoserine (DMPS).

One essential component of eukaryote cell membranes is PC lipids, which can be found abundantly in any cell compartment. The extra-cellular side of the plasma membrane is rich in PC lipid whereas the intra-cellular/cytosolic side is rich in PC along with PE, PS lipids[39]. Note, that the membrane always has an overall negative charge as they are made of negative and zwitterionic phospholipids.

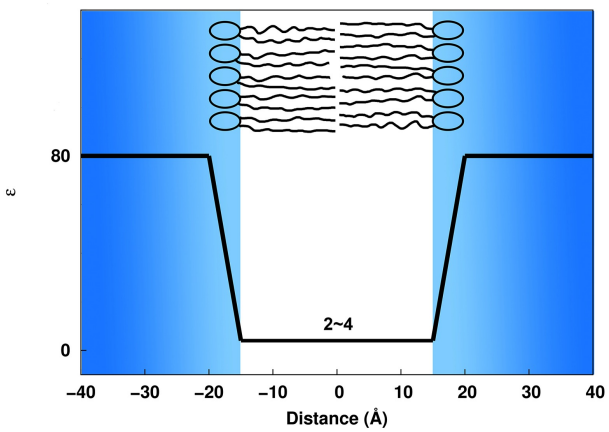


Figure 2.2: Dielectric constant profile across the bilayer normal.

Over the last 100 years, researchers have tried to understand membrane shape and organization. The earlier view dates back to 1925 by Gorter and Grendel[40], where they describe the membrane as composed of two molecules thick layer of fatty substances. The most popular and widely used model of cell membrane is formulated by Singer and Nicolson in 1972[41], also known as “the fluid mosaic model”. In this model, the lipids behave like a fluid at physiological temperature, enabling them to move continuously. The proteins and other biomolecules are embedded in the plasma membrane. This embedding gives the nature of mosaic. The fluid-mosaic mixture can move like a two-dimensional fluid in the lateral dimensions.

Later, in 2005, Engelman provided an updated version where the author describes the membrane as “more mosaic than fluid”[42]. The model tries to emphasize the fact that the membranes are crowded with proteins and that gives rise to the more mosaic nature. In a recent review article[43], Goñi also highlighted the importance of binding

of proteins to the membrane on temporary basis (which is a central topic in this thesis).

Recently, Kusumi et al. proposed a hierarchical “three-tiered” organization of the plasma membranes[44]. The authors highlighted two features to be important: thermal fluctuations/ movements of molecules, and weak cooperativity.

Summarizing all these models, we can essentially focus on some key features in addition to the bilayer form of membranes. These features are:

1. Membranes are crowded[42],

2. Molecules are in motion; let it be lipids, proteins, or other biomolecules. A static representation limits our understanding of membrane as a dynamic entity,

3. There exist proteins that bind to the membrane occasionally[43]. Unlike the soluble or transmembrane proteins, these types of proteins need to bind membrane transiently to perform their function,

4. Membranes are curved, heterogeneous, and often out of equilibrium[43].

2.2 Proteins

2.2.1 Membrane proteins

Membrane proteins are classified in different groups in literature. Interestingly, the categories vary to great extent from one classification to another. Lodish et al. classified membrane proteins in three categories: integral membrane proteins, lipid-anchored membrane proteins, and peripheral membrane proteins[2]. The authors classified peripheral membrane proteins as those that do not interact with the hydrophobic core of the membrane. These proteins bind to membrane either via indirect interaction with the transmembrane proteins or via direct interaction with the lipid headgroups[2]. Alberts et al. described membrane proteins association to membrane in eight different ways[33]. If they can be released from the membrane surface with “gentle extraction procedure” such as the variation of salt concentration

or pH and without disrupting the membrane, they are called as peripheral proteins [33]. Others are categorized as integral membrane proteins. In this description, peripheral proteins interact with other membrane proteins by noncovalent interactions. Petsko and Ringe catagorized membrane proteins in two types: integral membrane proteins and membrane-anchored proteins (peripheral proteins)[45]. Peripheral proteins are defined by the authors as the proteins that are attached to one side of the membrane while the remaining parts of the proteins are soluble in water on the same side of the membrane[45]. Perhaps, this is the most realistic definition of peripheral proteins. Among the membrane proteins, amphitropic proteins are those that bind to the cell membrane reversibly and transiently[46]. In this category, proteins binding to transmembrane proteins (often receptors) embedded in the membrane are not included (Figure 2.3). Such interactions may be categorized as protein-protein interactions.

We will use the term “peripheral protein”/ “amphitropic protein” synonymously. Henceforth the terms will be used only to refer to the proteins that bind to the membrane directly.

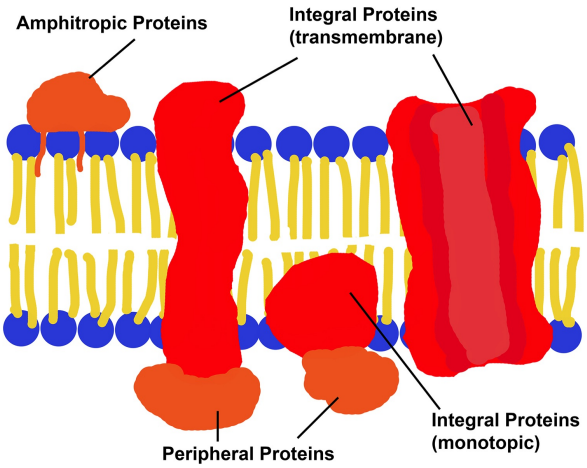


Figure 2.3: Some examples of membrane proteins. (Redrawn, modified and adapted from the source https://en.wikibooks.org/wiki/File:Membrane_protein.png)

2.2.2 Amphitropic proteins and how they bind to the cell membrane

The term amphitropic was first introduced by Burn in 1988[47]. The author pointed out that several proteins which used to be described as soluble cytoplasmic proteins can also localize on the membrane. Later, Johnson and Cornell proposed three distinct strategies for reversible membrane localization of these proteins in 1999[46]. The key features described by Johnson and Cornell are:

1. Equilibrium between soluble (inactive) and the membrane bound (active) forms.

2. Binding is a two-step process where the first step is served by electrostatics for the localization. Second step is the hydrophobic interaction and is "accompanied by intercalation". The binding reaction is fast and there is no delay between the adsorption and intercalation.

3. The off rate or dissociation constant (k_{off}) determines the affinity (K_d). As the on rate is very fast. On rate is determined by association constant (k_{on}).

4. The hydrophobic contributions are dominant.

5. The three strategies for membrane attachment are: a) lipid clamps, b) covalent lipid anchors, and c) amphipathic helices (Figure 2.4).

The model given by Johnson and Cornell is a pioneering model, which opened up a new way of looking at a whole new class of proteins and the importance of their membrane recruitment.

More recently, in 2009, Cho and Stahelin summarized the kinetics and energetics of association of peripheral proteins to the membrane[48]. The authors discussed the cases of membrane targeting domains such as: C1, C2, PH, FYVE, etc. and their host proteins. However, the authors greatly rely on the model proposed by Johnson and Cornell to describe the membrane localization and the bound state. The highlights of the binding mechanism for peripheral proteins are represented below (Figure 2.5):

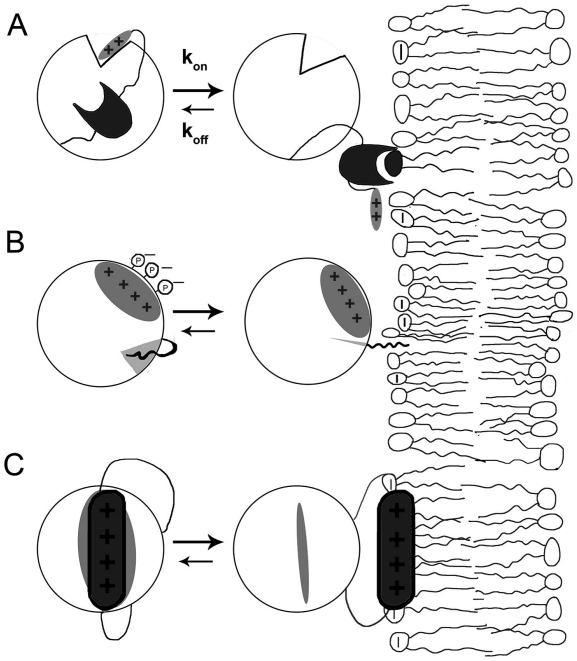


Figure 2.4: Reversible membrane binding strategies for amphitropic proteins described by Johnson and Cornell[46] (redrawn). (A) Lipid clamps, (B) Covalent lipid anchors, and (C) Amphipathic helices.

1. Step A is done "autonomously" or triggered by a "Ca^{2+}-induced electrostatic switch".

2. After localization/adsorption, the protein undergoes any of the steps B, C, or D.

B) Membrane penetration,

C) Binding to a specific lipid, or

D) For many cases, step C is followed by step B.

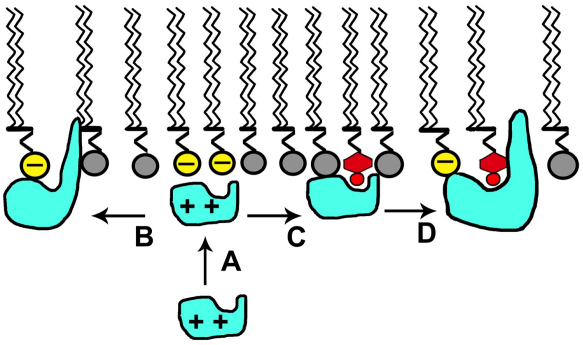


Figure 2.5: Model proposed by Cho and Stahelin[48] (redrawn).

Several attempts have been made to understand these steps, the role of different amino acids and determining the energetic contributions. Inspired by the idea that many cytoplasmic proteins contain clusters of basic amino acids and transiently interact with the membrane, Kim et al. performed experiments including equilibrium binding measurements with Lys_n and Arg_n peptides (n=1-5)[49]. The authors found that the energetic contribution of each lysine residue is ~1.4 kcal/mol irrespective of the negative lipid type and similar contribution is observed for arginine residue[49]. For the hydrophobic interactions with the bilayer, energetic contribution of ~0.8 kcal/mol per acyl chain-protein interaction has been suggested[50].

As stated in earlier sections, plasma membranes are always negatively charged due to the constituent phospholipids. How this negatively charged cell wall attracts proteins towards it is of great interest to the scientific community. Mulgrew-Nesbitt et al. provided a comprehensive review on the role of electrostatics in such attraction[51]. The authors highlighted that the presence of clusters of basic amino acids create a large positive electrostatic potential/patch around the protein interfacial binding site (IBS). The overall charge of the protein is less important as long as it contains a cluster of basic residues. However, when a protein does not contain basic clusters and has a net negative charge, the role of electrostatics is not known.

Murray et al. provided a theoretical insight into the electrostatic partitioning of Src proteins with phospholipid membranes[52]. Based on this work, it has been suggested that if a protein domain is bound to the membrane solely via electrostatic interactions, then it will localize ~3 Å away from the bilayer surface and would not penetrate the bilayer surface[46].

The binding of amphitropic proteins to cell membrane is essential to perform their function. The membrane acts as a meeting point for proteins and lipids[53]. Amphitropic proteins play an important role in cell signaling, metabolism, membrane trafficking[46, 48, 54]. The localization to the membrane serves either of two purposes: 1) to activate the proteins for performing their functions[55], or 2) to bind to the substrates located at the interface[56].

2.2.3 Limitations of existing binding models

Several attempts have been made to understand the binding mechanism of peripheral membrane proteins. Significant effort has been made to understand the role of electrostatics[51, 52, 57-60]. For more details on the overall binding mechanism, I refer the reader to two comprehensive review articles [46, 48]. These reviews do not account for all the peripheral membrane proteins, as stated by the authors[46]. Further enhancements of the models are needed to describe other peripheral membrane binders.

We identify the following limitations in the current peripheral protein-membrane binding models:

- Association of peripheral proteins is mediated by long-range electrostatics, which are interactions of basic clusters with negatively charged lipids. The role of clusters of basic residues is considered as granted[49, 51]. First of all, there might be many proteins that do not contain clusters of basic residues. How they bind to membrane is not known. Secondly, the range for electrostatics to modulate protein binding to membrane periphery is not clearly defined[46]. The estimated contribution of around 1.4 kcal/mol for each basic amino acid[49] seem somewhat arbitrary if one considers

the various types of lipids and the localization of the basic amino acids with respect to the membrane (whether they are localized at the interface or away from the membrane surface).

- There is no distinction between hydrophobic amino acids in these models[46, 48, 51, 59]. In particular aromatic side chains have properties distinct from aliphatic groups, and the three aromatic amino acids themselves have distinct hydrophobicity and polarity. For these reasons, the localization of aromatic amino acids may vary substantially compared to the other hydrophobic amino acids. Hence, these aromatic amino acids may engage in cation-π interaction with some lipids (PC, PE) if they are localized in close proximity of these lipid headgroups.

- Many of these models largely neglect the role of the water-membrane interface, specifically the energetics aspect and also the localization of amino acids at the interface. As it has been discussed earlier, water-membrane interfacial properties are different than bilayer core/bulk water. Hence, interaction of amino acids with the water-membrane interface may play a distinct and important energetic role in the formation of protein-membrane complex.

2.3 Selected protein candidates

We choose several proteins for our investigations. They are serine proteases and bacterial phospholipases. Those that we picked are reported to bind model membranes peripherally. Further details on these candidates are provided in the following sections.

2.3.1 Serine protease homologues: PR3 and HNE

We first choose two homologous serine proteases: Proteinase 3 (PR3) and Human Neutrophil Elastase (HNE) (Figure 2.6). These are enzymes associated with neutrophils (white blood cells). Serine proteases are important as they work as killing agents during infection. However, if not regulated properly they become the cause of

the disease itself by playing a direct role in autoimmune and inflammatory diseases. Such is the case for Rheumatoid arthritis and Wegner's granulomatosis[61, 62].

Figure 2.6: Structures of two homologous neutrophil serine proteases- PR3 (gray) and HNE (cyan). Secondary structure is shown in cartoon representation.

Both *in vitro* and *in silico* studies have shown, that PR3 can bind to the model membranes directly without the aid of any partner protein[63-65]. Human PR3 (221 a.a.) and HNE (218 a.a.) are homologous and have 56% sequence identity and 70% sequence similarity[66]. Most interestingly, PR3 has a net positive charge of +2 and HNE has +11. Also, PR3 has been reported to bind lipid bilayers using cation-π interaction[65].

2.3.2 *Bt*PI-PLC: a bacterial phospholipase

Bacillus thuringiensis phosphatidylinositol specific-phospholipase C (*Bt*PI-PLC) is a 34.8 kDa bacterial virulence factor and contains 296 amino acids (Figure 2.7). The protein is part of the TIM barrel fold super-family[67]. Unlike other proteins that have the TIM barrel fold; this protein is "incomplete" because it contains eight β-strands and only six α-helices[68].

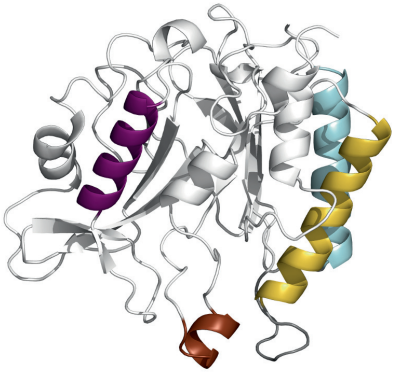


Figure 2.7: Structure of *Bt*PI-PLC. The relevant secondary structure elements are colored as: helix B (αB, brown), helix D (αD, magenta), helix F (αF, cyan), helix G (αG, yellow). A labeled version of *Bt*PI-PLC structure is provided in the Figure 3 of Paper 3.

This protein cleaves GPI-anchored proteins from the outer part of the eukaryotic plasma membranes[69, 70]. *Bt*PI-PLC follows a two step catalytic mechanism[56] and specifically cleaves the sn-3 phosphodiester bond in phosphatidylinositol (PI) [70]. Such activity of *Bt*PI-PLC is observed upon binding to PC or PE interfaces[71]. Wehbi et al. performed binding experiments with non-substrate vesicles[56], and found using Langmuir adsorption isotherm that the *Bt*PI-PLC binds anionic SUV's with higher affinity than the neutral phosphatidylcholine (PC). Such behavior is explained by the significant hydrophobic contribution towards the neutral vesicle and major electrostatics contribution for anionic vesicles[56]. Later, Pu et al. performed Fluorescence Correlation Spectroscopic (FCS) measurements of *Bt*PI-PLC binding and activity with pure and mixed anionic vesicles[70, 72]. Highest activity of *Bt*PI-PLC is observed at X_{PC} 0.5 (PC content in the vesicle) and the tightest binding at around X_{PC} 0.9 where the other partner lipid was phosphatidylglycerol (PG). Vesicles containing phosphatidylmethanol (PMe), phosphatic acid (PA) or phosphatidylserine (PS) did not lead to such tight binding[72].

For the binding of *Bt*PI-PLC, several structural parts and residues were experimentally identified as important: helix B[56, 73], rim loop and tryptophan

residues (W47, W242) [56, 73, 74], a tyrosine residue (Y88)[70] and a lysine residue (K44)[70, 73]. Recently, *Bt*PI-PLC is reported to use cation-π interactions as membrane targeting mechanism to bind lipid bilayers which has been demonstrated by *in silico* and *in vitro* experiments[6]. Three membrane binding regions were identified: helix B, and two other rim loops[6]. However, little is known about the role of electrostatics in *Bt*PI-PLC membrane-binding except a lysine (K44) can play an important role. Mutation of K44 to alanine has been shown to decrease the affinity more than 50 times for slightly anionic vesicles[70]. *Bt*PI-PLC is an overall negatively charged protein (-7). Also, no basic cluster is reported to be associated with *Bt*PI-PLC. Hence, the role of electrostatics in *Bt*PI-PLC membrane binding is still not clear.

3. Methods

This chapter is divided in four sections: Quantum mechanics, molecular mechanics, continuum electrostatics, system setup and analysis.

3.1 Quantum mechanics

3.1.1 Many electron wavefunction and Schrödinger equation

The non-relativistic time-independent Schrödinger equation has the form:

$$\hat{H}\psi = E\psi \tag{3.1}$$

where ψ is a many electron wavefunction, and $\hat{H}$ is the Hamiltonian operator. In atomic units $\hat{H}$ can be given by[75],

$$\hat{H} = -\frac{1}{2}\sum_i \nabla_i^2 - \frac{1}{2}\sum_A \frac{1}{M_A}\nabla_A^2 - \sum_{A,i}\frac{Z_A}{r_{Ai}} + \sum_{i>j}\frac{1}{r_{ij}} + \sum_{A>B}\frac{Z_A Z_B}{R_{AB}} \tag{3.2}$$

where Z_A is the nuclear charge, M_A is the ratio of mass of nucleus A to the mass of an electron, R_{AB} the distance between nuclei A and B, r_{ij} is the distance between electrons i and j, and r_{Ai} is the distance between nucleus A and electron i.

The many electron Schrödinger equation cannot be solved exactly. Hence, approximations are introduced.

3.1.2 Born-Oppenheimer Approximation

Nuclei are heavier than electrons (almost 2000 times). Hence they move slower than electrons. So, the seperation of nuclear and electronic motion is possible. This is called the “Born-Oppenheimer Approximation”[76]. Due to the slow nuclei movement compared to electrons, we can come to the so-called “electronic” Schrödinger equation[75],

$$\hat{H}^{el}\psi^{el} = E^{el}\psi^{el} \tag{3.3}$$

where,

$$\hat{H}^{el} = -\frac{1}{2}\sum_{i}\nabla_i^2 - \sum_{A,i}\frac{Z_A}{r_{Ai}} + \sum_{i>j}\frac{1}{r_{ij}} \tag{3.4}$$

Nuclear kinetic energy is not included in the electronic Schrödinger equation. The nuclei-nuclei repulsion term should be added to the electronic energy E^{el} to obtain the total energy E for the system. Hence,

$$E = E^{el} + \sum_{A>B}\frac{Z_A Z_B}{R_{AB}} \tag{3.5}$$

This is the energy value one obtains by performing typical electronic structure calculations.

3.1.3 Hartree products and Slater determinants

A wavefunction for N-electrons is a function of 4N coordinates. For each electron, one can describe them by three Cartesian coordinates (x,y,z) and spin function (up or down). Here, the Cartesian coordinate for electron i is collectively referred as r_i and the combination of Cartesian coordinates with the spin coordinates (function) as χ_i.

To describe the N-electron wavefunction, one can write it in terms of the product of one-electron functions, as orbitals are one electron wavefunctions:

$$\psi(\chi_i, \chi_j, \ldots, \chi_N) = \varphi_i(\chi_i)\varphi_j(\chi_j)\ldots\varphi_N(\chi_N) \tag{3.6}$$

This is known as a Hartree product (Equation (3.6)). These notations $\varphi_i(\chi_i)$ are dependent on both spatial and spin functions (as written with χ_i), so they are called the spin orbitals. These spin orbitals can be written as the product of the spatial orbitals and the spin functions (α or β). For example,

$$\varphi_i(\chi_i) = \phi_i(r_i)\left|\alpha\right\rangle \tag{3.7}$$

However, the problem with the Hartree products is that it does not satisfy the anti-symmetry principle.

Considering the case of two electrons, the anti-symmetric wavefunction can be written as,

$$\psi(\chi_1,\chi_2)=\frac{1}{\sqrt{2}}\left[\varphi_1(\chi_1)\varphi_2(\chi_2)-\varphi_1(\chi_2)\varphi_2(\chi_1)\right] \tag{3.8}$$

In a determinant form,

$$\psi(\chi_1,\chi_2)=\frac{1}{\sqrt{2}}\begin{vmatrix}\varphi_1(\chi_1) & \varphi_2(\chi_1)\\ \varphi_1(\chi_2) & \varphi_2(\chi_2)\end{vmatrix} \tag{3.9}$$

With this notation, one satisfies the anti-symmetry principle. The anti-symmetry principle results in the Pauli exclusion principle (the principle quantum numbers cannot be the same for two electrons). If one tries to put two electrons in the same orbital together then $\varphi_1=\varphi_2$, so the resulting wavefunction is zero ($\psi(\chi_1,\chi_2)=0$).

Now, if one writes the wavefunction for N-electrons following the same notation which satisfies the anti-symmetry principle then it leads to

$$\psi=\frac{1}{\sqrt{N!}}\begin{vmatrix}\varphi_i(\chi_i) & \varphi_j(\chi_i) & \cdots & \varphi_N(\chi_i)\\ \varphi_i(\chi_j) & \varphi_j(\chi_j) & \cdots & \varphi_N(\chi_j)\\ \vdots & \vdots & \ddots & \vdots\\ \varphi_i(\chi_N) & \varphi_j(\chi_N) & \cdots & \varphi_N(\chi_N)\end{vmatrix} \tag{3.10}$$

This is known as “slater determinants” that are made up from spin orbitals. A single slater determinant is a single population pattern of molecular orbitals (or a configuration). By expanding this determinant, one obtains N! Hartree products. The N-electrons are arranged in all N! possible ways in the N spin orbitals. This also ensures the indistinguishability of electrons. This part is largely adapted from the lecture notes of David Sherrill[77].

To construct a determinant, one needs the list of occupied orbitals. The occupied orbitals can be obtained using the Hartree-Fock (HF) method, which is also known as Self-consistent field (SCF) method.

3.1.4 HF method

Using the HF procedure, one can obtain the molecular orbitals lowest in energy (i.e. ground state orbitals) and hence the minimized electronic energy. This is done using the variational principle, which states that the expectation value of the Hamiltonian is an upper bound to the exact ground state energy[78].

These molecular orbitals (φ) are written as a linear combination of the atomic orbitals (ϕ) as in Eqn. (3.11). This expansion is known as Linear Combination of Atomic Orbitals (LCAO). The atomic orbitals are also known as basis functions.

$$\varphi = \sum_i c_i \phi_i \tag{3.11}$$

The HF method is an approximation to solve the many-electron Schrödinger equation. In this method, a single Slater determinant is used as the wavefunction. The single determinant is the function of molecular orbitals. Such formulation leads to an integro-differential equation, which can be written as an eigenvalue problem:

$$F\varphi = \varepsilon\varphi \tag{3.12}$$

where F is the Fock operator which is a one-electron operator, φ is the molecular orbital, and ε is the orbital energy. The Fock operator contains terms for the kinetic energy of the electron, potential energy of electron with respect to the nuclei, and the electron-electron repulsion that is treated in an average fashion.

3.1.5 Electron correlation

The Hartree-Fock approximation is very useful as a good starting point for solving the many-electron Schrödinger equation. What lacks in the HF method is electron correlation as they are treated in a mean field approach. The variational principle states that the solution will always be an upper bound of the exact solution. The

difference between the exact solution to the nonrelativistic energy (E) of the system and HF-limit energy (E_0) is known as the correlation energy (E_{corr}):

$$E_{corr} = E - E_0 \tag{3.13}$$

There are several post-HF methods to include the electron correlation effect. They are: Configuration interaction (CI), Many-body perturbation theory (MBPT), and Coupled cluster theory.

3.1.6.1 Møller-Plesset perturbation theory

Many-body perturbation theory (MBPT) can be used to estimate the correlation energy. Møller-Plesset perturbation theory in different orders is a popular electron correlation methods[75, 78, 79]. The most computationally viable one is the MP2 method, which takes into account correlation effects up to the second order.

In the perturbation approach, it is assumed that the Hartree-Fock wavefunction and the energy are very close to the exact wavefunction and the energy[75]. Then one can write the exact Hamiltonian composing of the zeroth-order Hamiltonian $\hat{H}^{(0)}$ and a perturbation $\hat{V}$ that is modulated by a parameter λ. So,

$\hat{H} = \hat{H}^{(0)} + \lambda\hat{V}$ and $\hat{V} << \hat{H}^{(0)}$.

Hence one can write the Schrödinger equation as,

$$\hat{H}\left|\Psi_n\right\rangle = \left(\hat{H}^{(0)} + \lambda\hat{V}\right)\left|\Psi_n\right\rangle = E_n\left|\Psi_n\right\rangle \tag{3.14}$$

where, $\left|\Psi_n\right\rangle = \left|\Psi_n(\lambda)\right\rangle$ and $E_n = E_n(\lambda)$.

Eqn. (3.14) can be expanded in a Taylor series and the Schrödinger eqn. for different orders can be obtained. From there one can calculate the energy corrections for different orders[78]. In this formulation, the sum of the zeroth and the first order gives the HF energy[75]. The correction to the HF energy to recover correlation energy starts at the second order.

3.1.6.2 Coupled cluster method

In the framework of coupled cluster theory, the exact many electron wavefunction can be generated by using an exponential excitation operator to the reference wavefunction/determinant. In single reference CC method, it is usually the HF determinant. The exact many electron wavefunction (Ψ) can be expressed with the exponential excitation operator ($e^{\hat{T}}$), and single determinant (Φ_0),

$$\Psi = e^{\hat{T}}\Phi_0 \tag{3.15}$$

The excitation operator can be written as the combination of single, double, ….. , upto N fold excitations for an N electron system [80],

$$\hat{T} = \hat{T}_1 + \hat{T}_2 + \hat{T}_3 + \ldots\ldots\ldots\ldots + \hat{T}_N \tag{3.16}$$

Unlike the perturbation methods where one adds all types of corrections (S,D,T,Q etc.) to the reference wavefunction up to a given order (2,3,4, etc.); in CC framework one adds all the corrections of given types (S,D,T,Q etc.) up to infinite order[19].

3.1.6 Density functional theory

In the framework of Density functional theory (DFT), one can determine the ground state electronic energy from the electron density (ρ)[81]. Later, Kohn and Sham introduced the Kohn-Sham orbitals in the density functional theory[82]. The authors provided the relation between Kohn-Sham orbitals and the densities. In this formulation, the exchange and the correlation terms are treated with an additional effective potential. This gave birth to the self-consistent Kohn-Sham equation, which can be solved like the HF equation[82]. The limitations of DFT lie in the exact description of this exchange-correlation part, which is not known. Nevertheless, DFT is one of the most widely used and popular computational methods in different branches of science and engineering disciplines.

Three types of exchange-correlation functionals are commonly used: 1) Functionals based on local density approximations (LDA), 2) Functionals based on general

gradient approximations (GGA), and 3) Hybrid functionals that use Hartree-Fock exchange[75].

3.1.7 Interaction energy using SAPT and energy decomposition

In Symmetry-Adapted Perturbation Theory (SAPT), the dimer Hamiltonian has the following form:

$$H = F_A + F_B + V + W_A + W_B \tag{3.17}$$

The dimer Hamiltonian (H) is partitioned into contributions from the monomer Fock operators (F), the interaction between the monomers (V), and the fluctuation potential of each monomer (W). A and B refer to the monomers respectively.

The interaction energy can be written as a perturbation series,

$$E_{\text{int}} = \sum_{n=1}^{\infty}\sum_{k=0}^{\infty}\sum_{l=0}^{\infty}(E_{pol}^{(nkl)} + E_{exch}^{(nkl)}) \tag{3.18}$$

Here, n denotes the order in V, k and l denotes the order in W_A and W_B, respectively. The E_{pol} terms are due to the polarization expansion and E_{exch} are due to the repulsive terms which result from the anti-symmetry of the wavefunction with respect to the exchange of electrons between the monomers. These formulations are adapted from the article by Hohenstein and Sherrill[17], and from the PSI4[83] manual. For a detailed overview of the SAPT methods, refer to the article by Jeziorski et al.[15]; and for the original formulations of the methods refer to the article by Rybak et al.[14]. There are different variants of SAPT method available for both wavefunction based and DFT based calculations. Wavefunction based SAPT variants are discussed in the article by Hohenstein and Sherrill[17]. Using SAPT, one can decompose the interaction energy into physically meaningful components: electrostatics, exchange, induction, and dispersion.

3.2 Molecular mechanics

3.2.1 Molecular Dynamics simulations

As stated earlier, nuclei of atoms experience an average field/force generated by the electronic configuration or arrangement around the nuclei. The average field/force is referred here synonymously. This average field creates so-called "Potential energy surface" over which nuclei of atoms move. In MD simulations, Newton's second law of motion describes the dynamics of a given system. In its simplest form one can write the relation between the potential energy and motion of the system as:

$$-\nabla V(r) = F_i(r) = m_i \frac{d^2 r_i}{dt^2} \tag{3.19}$$

where, the gradient of the energy gives the force acting on a particle (atom). With double integration we can obtain the position of the atoms after certain time. This collection of positions is often called a "trajectory". The motions of atoms are not independent from each other meaning they are coupled. Hence, they need to be solved numerically. Usually, the integration is carried out using finite difference methods.

3.2.2 Monte Carlo Simulations

Simulation technique such as Monte Carlo Simulations also uses the potential energy; however they do not use time as a variable. Hence, often direct time dependent properties cannot be calculated using MC simulations. Instead of time dependent evolvement of the system, in MC it uses random numbers to generate the system configuration and use different criteria to accept or reject the configuration, i.e. based on energy criteria using Metropolis algorithm[84].

3.2.3 Ensembles

Ensemble is a collection of microscopic states[85]. In other words, using ensembles one can calculate macroscopic properties from the microscopic properties using statistical mechanics[86]. Using computer simulations one can generate microscopic

observables such as coordinates, velocities etc. and using these microscopic variables one can measure the macroscopic properties e.g. pressure, temperature etc.

The common statistical ensembles are microcanonical (NVE), canonical (NVT), isothermal-isoberic (NPT), grand canonical (μVT) etc. For each ensemble, some thermodynamic variables are constant and other thermodynamic variables can be measured by ensemble averaging or time averaging. A computer simulation can be characterized by different thermodynamic variables such as V, P, E, T etc. These variables are connected. For example, for a constant number of particles (N), either the volume (V) or pressure (P) can be fixed. Both volume and pressure cannot be fixed together[19].

3.2.4 Ergodic hypothesis

The ergodic hypothesis assumes that the time average of a set of particles is equal to its ensemble average[84]. In other words, independent of the starting configuration of the system we can sample other points of the phase space. This hypothesis enables us to use MD simulations to extract the thermodynamic and structural properties of a system. Monte Carlo simulations calculate ensemble average whereas Molecular Dynamics simulations calculate time averaged properties.

3.2.5 Force field

Irrespective of the simulation technique, one needs to get the potential energy of the system. The potential energy consists of bonded and nonbonded terms (at least this is true for biomolecules). Often this set of energy functions is coined as "force fields". The potential energy form of a typical biomolecular force field[87] (e.g. CHARMM which is class I additive force field[88]) looks like following:

$$V(r) = \sum_{i=1}^{N_{bonds}} k_{bond,i}(d - d_0)^2 + \sum_{i=1}^{N_{angles}} k_{angle,i}(\theta - \theta_0)^2$$

$$+ \sum_{i=1}^{N_{torsions}} k_{torsion,i}[1 + \cos(n\phi - \phi_0)] \qquad (3.20)$$

$$+\sum_{i=1}^{N_{atoms}-1}\sum_{j=i+1}^{N_{atoms}}\left(4\varepsilon_{LJ,ij}\left[\left(\frac{\sigma_{LJ,ij}}{r_{ij}}\right)^{12}-\left(\frac{\sigma_{LJ,ij}}{r_{ij}}\right)^{6}\right]+\frac{1}{4\pi\varepsilon_0}\frac{q_i q_j}{r_{ij}}\right)$$

where, V is the potential energy, d is the distance between two bonded atoms, θ is the angle between three bonded atoms, ϕ is the torsion/dihedral angle between two bonds, ε is the well depth for LJ potential, σ is the equilibrium distance for LJ potentail, q is the charge of an atom, r is the distance between two atoms or between two charges.

The bonded terms are usually obtained from spectroscopic measurements. They can also be obtained from quantum mechanical calculations. The nonbonded terms are usually fitted to quantum mechanical calculations. Sometimes, the torsional parameters are also fitted to QM calculations[22, 88, 89]. Such is the case for CHARMM force field used in this thesis work.

3.2.6 Some essential practical considerations for MD simulations

3.2.6.1 Integrators and timestep

As stated earlier, the integration of equation of motion (EOM) in MD simulations are done numerically. The idea behind is to break down the integration in small parts. These small parts are separated by a fixed time δt[90]. This fixed time is commonly known as the "timestep". The choice of timestep is important for obtaining a realistic dynamics. In terms of numerical methods, a very small timestep is better. However, that will lead to very slow sampling of the phase space. On the other hand, a long timestep will lead to numerical instabilities. Appropriate timestep will lead to a better and smooth sampling of the phase space. For typical all atom MD simulations without any bond constrains, the choice of time step is 1 fs. This is because the timestep should be such so that one can capture the fastest motions of the system which are the bond vibrations. However, this is not true for the "Coarse-grained" MD simulations.

To integrate the EOM, many popular choices are available such as Verlet, velocity verlet, leap-frog, gear's predictor-corrector (with higher order variants) algorithm[84-86]. These equation of motion algorithms all have their advantages and limitations.

3.2.6.2 Boundary conditions

Boundary conditions are needed to remove the surface effects and sometimes to mimic the bulk properties. Both of the effects can be mimicked depending on the boundary conditions. If the surface effects are not removed then the atoms or molecules will evaporate from the simulation system. Boundary condition can be non-periodic, i.e. spherical boundary conditions. This boundary condition is also known as "liquid drop". A harmonic restraint force is applied at the border of the liquid sphere to prevent evaporation of the liquid. However, this boundary condition cannot mimic an infinite system because the sphere cannot be replicated infinitely. To mimic an infinite large system (bulk), periodic boundary conditions are used. Usually, cubic, rectangular, or octahedron prism/box (i.e. the unit cell to replicate) is used depending on the types of molecules one wish to simulate. In periodic boundary conditions, if a particle leaves the unit cell it enters through the opposite face of the cell. Hence, the number of particles in the simulation system is conserved and there are no evaporations of particles.

3.2.6.3 Cut-off

In a MD simulation, to evaluate all the nonbonded forces it would require ½ N (N-1) calculation for a system of N particles at every step. This is time consuming and some nonbonded interactions have strong distance dependence. Hence, one can introduce the scheme where the nonbonded calculations will be done only up to a certain distance from a particle. Typically, such distance is 12 Å. This dramatically reduces the computational cost.

3.2.6.4 Switching and shifting

Sometimes there might be discontinuity in the potential function due to the cut-off. This will lead to inconsistency in the energy calculations. To overcome this problem,

there are several approaches. Two of them are switching and shifting the potential[90].

In the switching approach, the potential is smoothly truncated from the switch distance up to the cut-off distance. Using the switching function, the potential becomes zero at the cut-off and there is no discontinuity.

In the shifting approach, the whole potential is shifted to make the potential zero at the cut-off distance.

3.2.6.5 Long-range electrostatics

The electrostatic interactions decay very slowly with the distance. Hence, use of cut-off does not allow capturing all the electrostatics interactions. To overcome this, one needs to treat the long-range electrostatic interactions or in other words what is left behind beyond the cut-off distance.

To treat the long-range electrostatics effects, some popular approaches exist such as-reaction field electrostatics[86, 90], Ewald summations[84, 90], Fast multipole moment method[84], Multilevel summation method[91]. In the reaction field approach, a uniform dielectric constant is assigned to the simulation zone beyond the cut-off value, which is treated as a dielectric continuum[86]. Within the cut-off zone, all the coulombic interactions are calculated explicitly. In the Ewald summation technique, the calculation is split in two parts. The direct evaluation is carried out within the cut-off distance. The calculation of the long-range part is carried out in the reciprocal space using Fourier transform method. This long range part can also be calculated in real space. However, the calculation becomes computationally costly.

An important feature in the Ewald sum methods is the efficient evaluation of fourier transform. Such efficiency will reduce the computational cost. Different variants of Ewald summation exists such as Particle mesh Ewald (PME), or Particle-particle-mesh Ewald method. One of the most widely used variants is the PME method.

Two drawbacks of PME methods (usually any FFT-based electrostatics method) are:

1. Calculation needs to be performed in 3D-periodic system.

2. FFT calculation is the performance bottleneck for parallel scalability.

To overcome these drawbacks, multilevel summation method (MSM) has been developed recently[91]. In this method, the short-range part is calculated similar to the PME method. The long-range part is interpolated from the grids. The long-range part is done in real space, hence no FFT calculations needed. This method can be used also for semi-periodic and nonperiodic systems.

3.2.6.6 Temperature and pressure coupling

In MD simulations, solving the equation of motion alone leads to the use of NVE ensemble. The system is isolated from the surroundings. However, seldom such conservation of energy mimics a real system. Often, there is exchange of heat with the surroundings in a real system. To mimic such behavior, one needs to use NVT or NPT ensemble during MD simulations. NVT or NPT ensembles can be achieved by coupling the system to the surroundings. For the temperature coupling, the system is connected to an "external heat-bath" with which the system exchanges heat and keeps the temperature constant. There exist several "Thermostatting Algorithms" to achieve such effect e.g. Anderson thermostat [92], Nose-Hoover thermostat [93, 94], Berendsen thermostat[95].

Similarly, pressure coupling with the surrounding can be achieved by using "Barostat Algorithm" such as Berendsen barostat, Parinello-Rahman barostat, Nose-Hoover Langevin piston barostat etc. For example, simulation of lipid bilayers is usually carried out in NPT ensemble (pressure coupling). Such choice of ensemble allows the simulation box to expand (change of volume) and relieve the excess pressure to mimic the real life cell membrane behavior.

3.2.6.7 Multiple timesteps

To capture the realistic dynamics of a system, one needs to choose the timestep in a way so the fastest motion of the system can be captured. However, the time scale for bond vibrations are fast comparing to vdW interactions (slow). Using a small

timestep to calculate all the interactions will limit the phase-space smapling as the calculations are more time consuming. Multiple timesteps algorithm can be used to overcome this problem[85].

In multiple timesteps algorithm, the forces acting on the system are classified into different groups according to their variation with time. Different timesteps are used to update and calculate these forces. The fast varying forces (bonded) are calculated more often than the slow varying forces (electrostatics).

3.3 Continuum electrostatics

Poisson-Boltzmann (PB) method is one of the continuum electrostatics methods. There also exist other approaches such as Generalized Born method (GB). In continuum electrostatics methods, the averaged electrostatics contribution of solvent and ions are treated. In an explicit solvent with ions, one can see the electrostatics contribution of solvents and ions on the biomolecules. However, it is difficult to isolate the contribution, as it is often the problem of many-body interactions. Nowadays, with the advent of free energy techniques, it is possible to estimate the solvent contribution. However, often it is not computationally cheap and straightforward. On that essence, continuum electrostatics is relatively cheap.

3.3.1 Poisson-Boltzmann Theory

PB method will be discussed here as this technique is applied in this thesis to evaluate protein-lipid interactions.

We will start with a molecule/system that is isotropic, meaning the dielectric constant of the system is fixed (dielectric constant, ε) and with no charges. Using the Laplace equation one can obtain the electrostatic potential, ϕ(r):

$$\nabla^2\phi(r) = 0 \tag{3.21}$$

where ∇ is the differential operator and can be expressed by,

$\nabla = \frac{\delta}{\delta x} + \frac{\delta}{\delta y} + \frac{\delta}{\delta z}$ in a three dimensional space.

Now, if the molecule is charged then the Laplacian equation becomes the Poisson equation which is one of the fundamental equation of classical electrostatics and can be expressed by:

$$\nabla^2 \phi(r) = \frac{-4\pi\rho(r)}{\varepsilon} \tag{3.22}$$

where, $\rho(r)$ is the fixed charge density of the molecule/system.

For an anisotropic or inhomogeneous system (i.e. solvent effect), the dielectric polarization at the interface will be different and to take that into account one can rearrange the above equation as follows:

$$\nabla\left[\varepsilon(r)\nabla\phi(r)\right] = -4\pi\rho(r) \tag{3.23}$$

Here, one can achieve that using a distance dependent dielectric constant, ε(r).

In a system where all the charges are present (explicit representation), the Poisson equation reduces to the Coulomb equation,

$$\phi = \frac{1}{\varepsilon}\sum\frac{q_i}{d_i} \tag{3.24}$$

Note, that the dielectric constant ε=1. Such a choice is reasonable as the explicit representation will generate the dielectric screening.

This is the relation between classical Coulomb's equation and the Poisson's equation. They are actually synonymous in explicit representation of all the charges. To account for ionic solvents, the Poisson-Boltzmann equation needs to be solved.

In a real solution, often there are ions present. As stated above, in an explicit model (when one uses Coulomb's equation) it is not needed to treat anything else to take the ion effects into account. To treat the ionic effect into Poisson's equation, the Boltzmann's ionic distribution is added and the equation becomes:

$$\nabla\left[\varepsilon(r)\nabla\phi(r)\right] = -4\pi\rho(r) - \quad ionic \quad contribution \tag{3.25}$$

or,

$$\nabla\left[\varepsilon(r)\nabla\phi(r)\right] = -4\pi\rho^{f}(r) - 4\pi\sum_{i} c_i^{\infty} z_i q e^{\frac{-z_i q\phi(r)}{kT}} \lambda(r) \tag{3.26}$$

where,

$\rho^{f}(r)$ = molecular fixed charges,

c_i^{∞} = the concentration of ion i at an infinite distance from the molecule,

z_i = valency of the ion,

q = proton charge,

k = Boltzmann constant,

T = Temperature,

$\lambda(r)$ = accessibility of ions at point r.

This is the so-called "Non-linear Poisson-Boltzmann equation". This equation can be linearized under the assumption that the electrostatic potential $\phi(r)$ is small compared to kT. And thus one obtain the "linearized Poisson-Boltzmann equation":

$$\nabla\left[\varepsilon(r)\nabla\phi(r)\right] = -4\pi\rho^{f}(r) + 4\pi\frac{\sum_{i} c_i^{\infty} z_i^2 q^2 \phi(r)}{kT}\lambda(r) \tag{3.27}$$

This equation contains an important parameter commonly known as Debye screening parameter (K_D) and is expressed by,

$$K_D^2 = 8\pi \frac{\sum_i c_i^\infty z_i^2 q^2}{2\varepsilon kT} \tag{3.28}$$

which is inversely related to the "Debye length" by:

$$K_D^2 = \frac{1}{l_D^2} \tag{3.29}$$

where l_D is the Debye length.

The Debye screening parameter provides an interesting physical insight. It refers that the electrostatic potential of a molecule in an ionic solvent reduces exponentially. As showed earlier in chapter 1, the dependence of electrostatic energy is 1/r without the presence of any ion using the Coulomb's law. With the use of continuum electrostatics one can estimate the range of electrostatic forces in ionic solution. Using this theory, one can deduce the range of electrostatics. The range reduces to 10-20 Å[96]. This is actually the reality for biomolecules in a cellular environment. This section is largely adapted from a review article by Fogolari et al. [96]. There are some excellent review articles and pioneering work by the group of Barry Honig on the use of Poisson-Boltzmann equation as a continuum electrostatics tool[59, 97, 98]. Readers are referred to those articles for further details of this method.

Once the electrostatic potential $\phi(r)$ is obtained, it is easy to estimate classical electrostatic free energy. The formulation is as follows:

$$\Delta G_{el} = \frac{1}{2}\int_V \rho^f(r)\phi(r)dV \tag{3.30}$$

$$\Rightarrow \Delta G_{el} = \frac{1}{2}\sum \rho^f(r)\phi(r) \tag{3.31}$$

One can simply rewrite this as,

$$\Delta G_{el} = \frac{1}{2}\sum (q_i\phi_i) \tag{3.32}$$

where i represents the atomic position, and q_i, ϕ_i is the fixed charge and the potential at that position respectively.

3.4 System setup and analysis

3.4.1 Continuum electrostatics calculations

3.4.1.1 System setup for continuum electrostatics calculations

We extracted the structures of the equilibrated bilayers after MD simulations to perform the continuum electrostatics calculations by solving the Poisson-Boltzmann equation. The protein-membrane complex is prepared by placing the protein with its membrane binding orientation above the bilayer upper leaflet. Further details are provided in the attached paper (Paper 3).

3.4.1.2 Calculating electrostatic potentials and electrostatic free energy

The parameters for protein and ions are adapted from the CHARMM all atom force field (c22[23] including CMAP correction)[99] and for the lipids from the force field update for lipids (Charmm36)[100]. Poisson-Boltzmann calculations are carried out using APBS (version 1.3) [101]. The non-linear Poisson-Boltzmann equation is solved using the focusing technique[97].

The electrostatic free energy for the protein $G_{el}(P)$, membrane $G_{el}(M)$ and the protein membrane complex $G_{el}(P.M)$ is calculated. The electrostatic free energy contribution to the free energy of binding due to protein membrane interactions can then be evaluated as follows:

$$\Delta G_{el} = G_{el}(P.M) - [G_{el}(P) + G_{el}(M)] \tag{3.33}$$

3.4.2 MD simulations of protein-bilayer complex

3.4.2.1 System setup for protein-bilayer complex

The proteins are manually docked to the pre-equilibrated bilayers for the MD simulations of protein-bilayer complex. The starting orientations are taken from implicit membrane simulations (IMM1-GC). Simulations are performed using

NAMD (v2.9, v2.10)[102]. The CHARMM all-atom force field[23] (c22 including CMAP correction)[99] and the force field update for lipids (CHARMM36)[100] were used for protein-membrane simulations.

3.4.2.2 Trajectory analysis

Analysis were performed on the simulation trajectories using CHARMM (v33b1, v38b2) [103] and VMD (v1.9.1)[104]. Detailed descriptions for analysis of hydrogen bonds, hydrophobic contacts, cation-π interactions, electron denisty profiles (EDP) are provided in the attached papers (Paper 2 and 3).

3.4.3 Quantum mechanical calculations

Quantum mechanical calculations are carried out to improve the force field description of cation-π interactions between tyrosine and choline. Phenol and tetramethylammonium (TMA) are used, as they are analogues of tyrosine and choline, respectively. We built a potential energy surface (PES). The details are provided in the accompanying paper (Paper 4). The calculations are performed using NWCHEM (v6.3) [105] and PSI4 (v 4.0b5)[83].

4. Aims of the thesis

Our main goal is to address the limitations of existing canonical binding models for amphitropic proteins and work towards the construction of a generalized model for peripheral membrane binding in terms of noncovalent interactions.

Using a selection of proteins we intend to bring new insights to the commonly acknowledged view and improve the current models. The selected proteins are:

- two neutrophil serine proteases: Proteinase 3 (PR3) and human neutrophil elastase (HNE);

- a bacterial phospholipase: *Bacillus thuringiensis* phosphatidylinositol-specific phospholipase C (*Bt*PI-PLC).

Each of these proteins has a pathophysiological relevance; PR3 and HNE are drug targets in autoimmune and chronic inflammatory diseases, *Bt*PI-PLC is a virulence factor.

Besides their function, the main differences between these proteins are their interfacial binding site (IBS) and their overall net charge:

- PR3 and HNE are both positively charged. PR3 binds to lipid vesicles much more strongly than HNE does although they are homologous. The IBS of PR3 contains three phenylalanines whereas the IBS of HNE has fewer hydrophobic amino acids and more positively charged residues.

- *Bt*PI-PLC is negatively charged and its IBS is rich in tyrosines. Nevertheless, *Bt*PI-PLC does not display any obvious cluster of basic amino acids at its IBS as other peripheral proteins described in the literature (including PR3).

Our main goal will be achieved in three steps:

1. We will first examine how electrostatics and hydrophobic interactions contribute to the binding affinity of PR3 to lipid vesicles. Comparison between PR3 and its close

homologue HNE allows exploring the effect of substituting hydrophobic amino acids at the IBS and in particular the role of phenylalanine in membrane binding. Furthermore we will investigate the role of electrostatics and the role of the balance between hydrophobicity and electrostatics.

2. With *Bt*PI-PLC, negatively charged and with virtually no basic cluster, we will investigate the role of weak electrostatics. We will also characterize choline-tyrosine cation-π interactions and their role for specific PC binding.

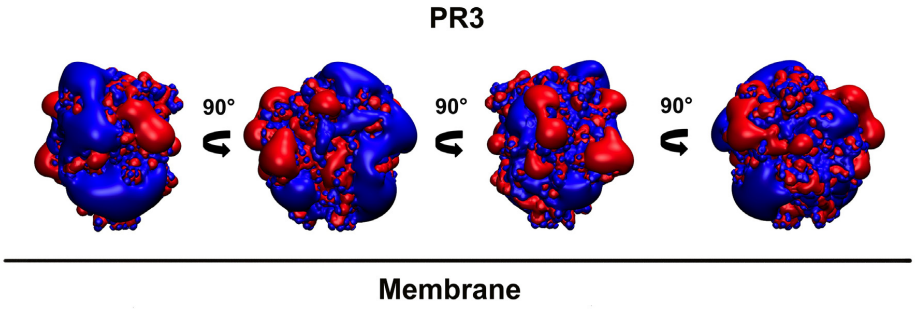
3. We wish to benchmark and if necessary improve the description of cation-π interactions in simulations using molecular mechanics force fields. To achieve this, we investigate the nature of interactions between phenol and tetramethylammonium through the calculation of potential energy surfaces (PES) using high-level electronic structure calculations. Results from these calculations will be confronted to PES calculated with molecular mechanics force fields.

5. Results and discussions

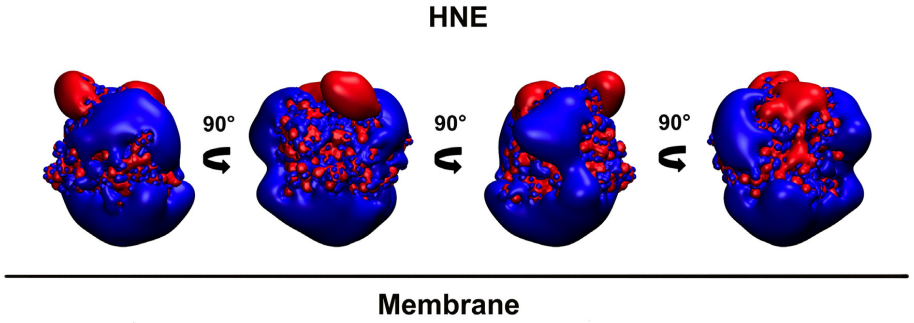
5.1 On the association of peripheral proteins: Case study of PR3, HNE, and *Bt*PIPLC

5.1.1 Electrostatics and hydrophobic interactions contribute differently to the membrane binding of two homologous proteins

Electrostatic potential is a useful property to obtain an idea about electrostatic partitioning of a protein with respect to the membrane [51]. The electrostatic potentials of PR3 and HNE both have positive groove/surface on their IBS (Figure 5.1.1). However, HNE has more positively charged surface around its IBS compared to PR3. This is an indication that HNE will respond more to increasing anionic lipid fraction compared to PR3.



(A)



(B)

Figure 5.1.1: Electrostatic potential of (A) PR3, and (B) HNE. The isocontours at +1 k_BT/e (blue) and -1 k_BT/e (red) are shown. Membrane plane is represented only to highlight the protein IBS and not the binding itself. Hence, a distance is maintained between the proteins and the membrane.

To elucidate both the electrostatic (desolvation penalty of polar residue + protein-membrane coulombic interaction) and the nonpolar contributions (aliphatic + aromatic) to the binding of PR3 and HNE for different membrane composition, we choose a cost-effective implicit membrane model (IMM1-GC)[106]. The negative lipid fraction of the membrane is varied from 0 to 100%.

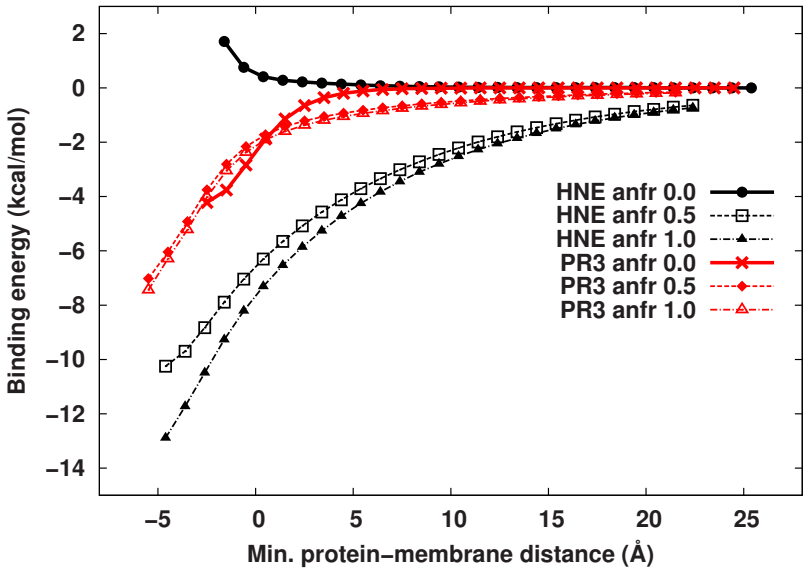


Figure 5.1.2: Variation of binding energy of PR3 and HNE with increasing fraction of anionic lipid (anfr). When anfr= 0.0 the model bilayer is neutral and it is fully anionic when anfr= 1.0. HNE is represented with black lines and PR3 with red lines.

For PR3, the change of anionic lipid fraction has a less profound effect compared to HNE (Figure 5.1.2). For HNE, the binding energy increases dramatically by the

change of membrane from neutral to a 50% negative membrane. However, further increase in the total negative charge density does not increase the binding energy proportionally. The favorable energetic contribution of HNE with an anionic membrane compared to a neutral membrane is interesting. This indicates that HNE binds to a membrane with electrostatics and the nonpolar contribution is not significant. On the other hand, PR3 binds to neutral membrane favorably thus indicating PR3 has nonpolar binding contribution. We have demonstrated this using energy decomposition analysis [107]. For HNE, the desolvation penalty is so high that it has slightly unfavorable total interaction energy. Finally, we would like to highlight that anionic lipid fraction (surface charge density) is sensed by these proteins at almost 20~25 Å. This is in fact the electrostatic contribution. HNE responds more to anionic lipid fraction than PR3. This is due to HNE containing more positively charged amino acids at the IBS. On the other hand, PR3 start sensing a neutral lipid when it is very close to the membrane surface (~5 Å). Similarly, unfavorable contributions of HNE start increasing when it is very close to the surface of a neutral membrane.

The binding energetics of PR3 and HNE analyzed here also correlate well with the experimental binding data obtained using Surface Plasmon Resonance (SPR) measurements[107]. Schillinger et al. showed by varying the salt concentration with a neutral bilayer (POPC liposome) that PR3 binding is less altered by increasing salt concentration[107]. However, HNE binding is significantly affected by the increasing salt concentration[107].

5.1.2 A cluster of basic amino acids is not essential for membrane binding

Not all amphitropic proteins contain well-defined basic clusters. Such an example is *Bt*PI-PLC. The *Bt*PI-PLC has five lysines (K38, K44, K122, K201, K279) and one arginine (R71) at or nearby the IBS. Surprisingly, they do not form a basic cluster as they are not close to each other. The simplest way to confirm such observation is to look at the electrostatic potential (Figure 5.1.3). The electrostatic potential does not show any large blue patch/surface on its IBS. However, *Bt*PI-PLC has been shown to

bind vesicles with an affinity of ca. 3.5 ± 0.7 μM where the vesicles were prepared using different phospholipids (DMPC: DMPG 80: 20 mixture) [6, 7]. This raises the question of how *Bt*PI-PLC reaches the membrane as the requirement of having clusters of basic residues for cytoplasmic proteins are highlighted in a series of previous works[49, 51, 52, 57-59, 108]. Moreover, *Bt*PI-PLC attaches to the extracellular side of the cell, which makes it an even more interesting case as the extracellular side of a cell has less anionic lipid fraction due to the higher PC content compared to the cytoplasmic side[37-39].

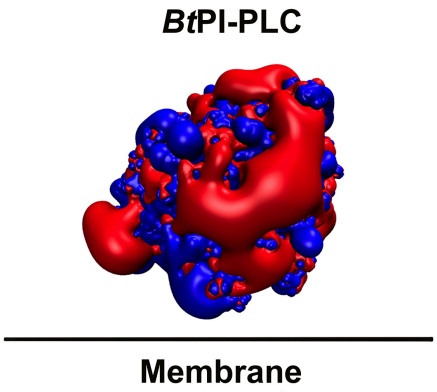


Figure 5.1.3: Electrostatic potential of *Bt*PI-PLC. The isocontours at +1 k_BT/e (blue) and -1 k_BT/e (red) are shown.

To explore the role of electrostatics on a protein like *Bt*PI-PLC, we used a continuum electrostatics approach. The calculated electrostatic free energy (ΔG_{el}) profile shows that *Bt*PI-PLC has a very low electrostatic partitioning towards a slightly negative membrane (negative charge fraction equivalent to 20%). The electrostatic contribution at the minima is -0.25 kcal/mol (Figure 5.1.4) which is almost 12 to 20 times lower than what has been reported for other peripheral proteins (from -3 to -5 kcal/mol) using the same computational approach [51]. Compared to other reported peripheral binders *Bt*PI-PLC stands out. Thus we believe that the role of electrostatics is not only to drag the proteins towards the membrane, but also to maintain the

adequate equilibrium between the soluble and the membrane-bound states. Low electrostatic partitioning means that the protein-membrane coulombic interaction is only slightly higher than the desolvation penalty for the polar and the charged residues[51]. Hence, the protein can easily go back to the soluble state from the membrane-bound state due to environmental changes or external stimuli.

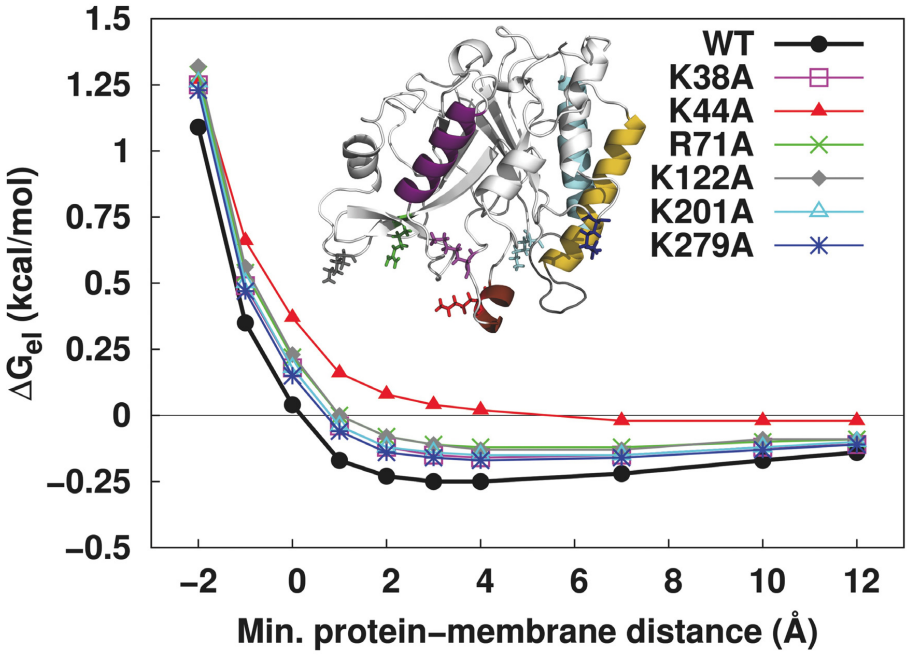


Figure 5.1.4: Electrostatic free energy (ΔG_{el}) profile of wild-type and mutant *Bt*PI-PLC.

5.1.3 Contributions of positively charged residues are not equal and in fact are distance dependent

To understand the positively charged residues contribution to membrane binding, we mutated all the charged residues at the IBS to alanine (Figure 5.1.4). The calculated electrostatic free energy shows that *Bt*PI-PLC membrane electrostatic partitioning is largely governed by K44. Mutating it to an alanine completely abolishes the favorable electrostatic free energy. Mutation of other charged residues has a less

dominant effect. The effect seems to be correlated to the distance between the amino acids and the phosphate plane of the membrane (Table 5.1). Hence, if a basic residue is closer to membrane its contribution is substantially higher than other basic residues localized further away from the membrane.

5.1.4 Non-monotonic effect of membrane anionic charge density on the electrostatic free energy

To examine the effect of surface charge density on membrane binding of *Bt*PI-PLC, we used continuum electrostatics. We prepared four different bilayer compositions where we vary the DMPC lipid compositions (X_{PC} 1.0, 0.8, 0.5, and 0). The other lipid partner is DMPG. The minimum protein-membrane distance (d) is 3 Å, which was found to be the minima in the energy profile. The electrostatic free energy is slightly unfavorable for a neutral membrane with a value of 0.1 kcal/mol (Figure 8 (A) of Paper 3). The profile decreases monotonically up to a charge density of 50% (ΔG_{el} = -0.68 kcal/mol). Increasing negative charge does not make the ΔG_{el} more favorable; in fact it makes ΔG_{el} slightly less favorable (ΔG_{el} = -0.42 kcal/mol).

To explore the effect of salt concentration on ΔG_{el}, we perform calculations with a slightly anionic membrane (X_{PC} 0.8), placing the protein at d= 3 Å and varying the salt concentration from 0.025 to 0.7 M. We found a quasi-parabolic ΔG_{el} profile for *Bt*PI-PLC (Figure 8 (B) of Paper 3). The minimum is located at salt concentration of 0.1M (ΔG_{el} = -0.25 kcal/mol), which is also approximately the physiological ionic strength on the extracellular side of eukaryotic cell.

5.1.5 Final note on the role of electrostatics in membrane association

The protein candidates described in the literature and explored here in this thesis give a basis to further extend the role of electrostatics. We see two types of protein candidates. One type contains large positive patch/groove at their IBS due to clusters of basic amino acids. Some of these proteins are negatively charged[51, 57], and the charge distribution determines the ΔG_{el}. This category responds monotonically to the increasing surface charge density[59]. Also, increasing salt concentrations monotonically increase ΔG_{el} and make it less favorable [59].

On the other hand, proteins that do not contain basic clusters at the IBS and have a net negative charge have large negative surface. Hence, it is very subtle how they interact with an anionic membrane. One of them (*Bt*PI-PLC) is explored here. We found that the electrostatic free energy is 12~20 times lower than the first category. However, this small contribution is essential for membrane binding. Such proteins neither respond monotonically to the increasing surface anionic charge density, nor to the salt concentration. This phenomenon is due to the unfavorable interactions between the large negative surface electrostatic potentials and the negatively charged membrane. It is indeed fascinating to observe how electrostatics partitioning towards membrane is tuned for such type of proteins (considering other proteins with similar characteristics exist). This is an interesting example of nature's negative design principle[109].

We hypothesize the consequences of having such low electrostatic partitioning for a protein by trying to relate it with the function of the protein. *Bt*PI-PLC cleaves GPI-anchored proteins off from the extracellular side of the eukaryotic plasma membrane. To find the substrate on the cell surface, *Bt*PI-PLC uses 3-D (hopping) and 2-D (scooting) search strategies[7]. During this substrate searching process, *Bt*PI-PLC needs to frequently dissociate to overcome the obstacles on the cell surface. For this reason, we believe low electrostatic partitioning helps in the hopping process. It would indeed be interesting to explore if other amphitropic proteins/enzymes use similar substrate searching mechanism. Last but not the least, it would also be interesting to check whether all or some of them have low electrostatic partitioning like *Bt*PI-PLC.

5.2 Membrane bound state of *Bt*PI-PLC and involved interactions

To investigate the membrane bound state of *Bt*PI-PLC, the protein is manually docked to pre-equilibrated bilayers in its membrane binding orientation. The membrane binding orientation was obtained from the implicit membrane simulations (IMM1-GC).

5.2.1 Cation-π interactions play an important role

Grauffel et al. reported the existence of tyrosine-choline cation-π interactions in the membrane bound state of *Bt*PI-PLC[6]. The authors also suggested that the cation-π interactions aid in recognition of PC head groups of DMPC lipids. Here, by varying the PC content of the bilayer we further investigate the findings and show that cation-π interactions between tyrosine and PC headgroups indeed play a role in lipid specificity. The tyrosine residues of *Bt*PI-PLC that engage in cation-π interactions with high occupancies with a pure DMPC lipids also engage in cation-π interactions in mixed bilayers. The cation-π interactions between these tyrosine residues and the bilayer vary as a function of X_{PC} (Figure 5 (B) of Paper 3). Interestingly, none of the tyrosine residues that mediate cation-π interaction, partition below the phosphate plane of the bilayer (Table 5.1 and Figure S10 (B-D) of Paper 2). Also, the occupancies obtained from the MD simulations varying the PC content of the bilayer correlate quite well with the experimentally determined effect of mutating these tyrosine residues (mutated to alanine) on the *Bt*PI-PLC affinities towards SUVs and hence on the evaluated ΔΔG (Figure S2 of Paper 3). The K_d values are taken from Grauffel et al. [6].

There is evidence that multiple aromatic amino acids can form a cation-π cage to bind PC lipid mimics [110]. However, it has not been described whether multiple PC lipids can engage in cation-π interaction with the same aromatic residue at the same time. Our analysis reveals that the cation-π interaction can occur simultaneously between two PC lipids with one tyrosine residue. For example, one of the tyrosine residues mediates cation-π interactions with more than one PC lipid simultaneously

(Table S6 of Paper 3). This identification of cation-π complexes is based on molecular mechanics force fields. Further studies are required to investigate the stability of such adducts using more accurate methods such as QM. However, recent work on model compounds using QM calculations also suggest that the cation-π interactions do not need to be 1:1, and they can indeed form sandwich complexes of cation-π-cation [111].

Table 5.1: Anchoring depth of *Bt*PI-PLC residues in the bilayer core averaged over simulation time. Positive values indicate that the centers of mass of the amino acids are on average buried below the phosphate group of the lipids during the simulation. The average phosphate plane is used as the reference plane. Replicates are denoted as r1 and r2.

SSE	aa	X_{PC} = 1.0		X_{PC} = 0.8		X_{PC} = 0.5		X_{PC} = 0
		r1[a]	r2[b]	r1	r2	r1	r2	r1
β1-αB	K38	-6.8±2.2	-7.6±2.3	-8.1±2.0	-6.7±2.5	-6.8±3.1	-9.3±3.1	-8.4±3.0
αB	Q40	0.6±2.3	-0.2±2.4	-1.1±2.1	-0.4±2.5	0.5±3.3	-2.8±3.1	-2.0±3.1
	N41	3.3±1.8	2.8±2.0	2.3±1.7	3.4±2.3	3.1±2.4	0.7±2.4	1.3±2.6
	P42	4.2±1.9	3.7±1.9	3.0±1.8	4.0±2.2	3.8±2.5	1.5±2.3	1.7±2.7
	I43	4.3±1.7	3.9±1.8	3.7±1.6	4.3±2.1	4.2±1.9	2.5±1.8	2.3±2.4
	K44	0.4±1.7	-0.2±1.8	-0.3±1.6	0.5±2.2	0.4±1.9	-1.1±1.8	-1.3±2.3
	Q45	-0.6±1.9	-1.0±1.9	-1.8±1.8	-0.8±2.1	-1.0±2.4	-3.3±2.0	-3.1±2.7
	V46	0.4±1.9	0.3±1.8	-0.3±1.7	0.3±2.0	0.3±1.9	-2.1±1.9	-1.8±2.7
	W47	-0.8±1.9	-0.7±2.2	-0.7±1.9	-0.8±2.1	0.1±1.7	-2.3±1.8	-2.6±2.6
β2	R71	-10.1±2.0	-10.9±2.4	-10.0±2.3	-9.1±3.1	-9.6±1.8	-10.2±2.3	-10.3±2.5
β2-αD	P84	-3.2±1.9	-4.5±2.3	-2.5±2.0	-2.2±2.7	-1.3±1.6	-3.8±1.7	-4.0±2.4
	L85	-3.6±1.6	-4.2±2.0	-3.7±1.7	-2.9±2.4	-2.9±1.7	-4.8±1.7	-4.7±2.3
	Y86	-5.5±1.9	-7.0±2.2	-6.1±1.9	-5.0±2.8	-2.0±2.1	-6.8±2.5	-6.7±2.4
	Y88	-4.9±2.0	-5.8±2.3	-5.9±2.0	-4.4±2.6	-6.3±2.4	-6.4±3.0	-6.2±2.7
β3-αE	Y118	-8.7±2.1	-8.1±2.6	-8.2±2.3	-7.4±3.1	-8.3±1.8	-7.0±2.2	-7.9±2.7

	K122	-6.1±2.9	-5.3±3.3	-5.2±3.3	-5.1±4.2	-5.0±2.5	-5.3±3.0	-5.8±3.2
β6-αF	Y200	-10.9±2.2	-10.8±2.5	-10.4±1.8	-10.7±2.5	-10.5±2.2	-11.8±3.3	-12.2±4.2
	K201	-7.6±2.6	-7.5±2.8	-7.2±2.1	-7.4±2.8	-7.0±2.6	-8.7±4.0	-9.7±4.4
αF	Y204	-7.3±2.6	-7.5±2.5	-8.3±2.7	-8.6±2.5	-8.7±2.8	-11.6±4.1	-12.6±4.0
β7-αG	S236	-6.6±2.0	-6.5±2.2	-6.1±1.8	-6.5±2.2	-6.4±2.0	-7.8±3.1	-8.9±3.6
	S237	-3.0±1.9	-3.0±2.1	-2.8±1.8	-2.9±2.1	-2.7±1.9	-5.5±3.9	-5.6±3.8
	G238	0.1±2.0	0.2±2.2	0.4±1.8	0.0±2.3	0.4±2.1	-2.5±3.6	-3.8±4.2
	G239	2.2±1.9	2.1±2.0	2.2±1.7	2.2±2.1	2.4±2.0	0.3±3.1	-1.4±4.2
	T240	3.3±2.0	3.3±2.0	2.9±1.9	3.0±2.0	3.1±2.2	1.2±3.1	-0.8±3.5
	A241	2.6±2.2	2.5±1.9	1.7±2.1	2.0±2.1	2.0±2.5	0.4±2.6	0.1±3.3
	W242	3.9±2.4	4.1±2.1	3.3±2.2	3.3±2.2	3.3±2.4	1.3±3.3	3.1±3.2
	N243	-0.3±2.3	0.0±2.1	-0.4±2.0	-0.6±2.2	-0.5±2.3	-2.5±3.4	-2.0±3.4
	S244	-2.3±2.1	-2.3±1.9	-2.9±1.9	-2.7±2.0	-2.7±2.2	-4.4±2.6	-4.4±3.2
αG	Y246	-4.5±2.4	-4.3±2.1	-6.0±2.2	-5.2±2.2	-5.4±2.5	-8.2±2.4	-8.8±3.1
	Y247	-2.3±2.5	-2.0±2.2	-3.4±2.4	-3.2±2.2	-3.2±2.5	-5.5±2.9	-6.1±3.6
	S250	-7.2±2.7	-6.9±2.4	-8.7±2.5	-8.1±2.4	-8.3±2.7	-11.3±2.9	-11.7±3.4
	Y251	-6.7±3.2	-5.8±2.5	-7.6±2.7	-7.4±2.6	-7.2±2.9	-10.3±3.7	-10.9±4.0
β8-αH	K279	-8.1±3.1	-8.1±3.0	-10.1±2.9	-9.3±3.1	-9.5±3.6	-12.6±2.9	-12.7±3.5

[a] values taken from Grauffel et al. [6], [b] kindly provided by Cédric Grauffel for comparison

5.2.2 Hydrophobic contacts and hydrogen bonds do not explain membrane-binding affinity

Hydrophobic contacts mediated by the aliphatic groups of the amino acids and the lipid tails strengthen the binding. We also see the formation of hydrogen bonds between the side chain and the backbone of amino acids with the phosphates, head groups, and the glycerols of the lipids. However, for *Bt*PI-PLC most of the hydrogen bonds form between the amino acid side chains and the phosphate groups of the lipids. The analysis of hydrogen bonds does not show any clear trend in explaining binding affinities for *Bt*PI-PLC towards SUVs (Table S3 of Paper 3). Likewise,

hydrophobic contacts mediated by the enzyme are similar and do not vary significantly with the change of bilayer compositions (Figure 5 (A) of Paper 3). Three structural elements of *Bt*PI-PLC penetrate the phosphate plane and mediate hydrophobic contacts with the acyl chains of the bilayers (Figure 3 of Paper 3). Also the position of *Bt*PI-PLC does not vary significantly with the change of bilayer PC content (Figure 4 of Paper 3). This explains the almost equal hydrophobic contacts irrespective of membrane composition.

5.2.3 Nature of the interactions and role of lipid dynamics

Interactions between amino acids and the lipids are transient in nature (Figure 4 (A) and S8 of Paper 2). These interactions (hydrogen bonds, hydrophobic contacts, and cation-π) form and break within 500 ns time scale. The membrane residence time of *Bt*PI-PLC is 303±30 ms[7] and 500 ns simulation time is a fraction of that residence time. Within this simulation time of 500 ns, many of the interactions have occupancies between 30-60%. In addition, there is significant amount of lipid exchange as well (Figure S9 of Paper 2). Lipid exchange takes place at a time scale of 100-200 ns. The analysis also suggests that these 500 ns simulations manage to reach equilibrium, as there is lipid exchange and the interactions patterns are conserved between simulation replicas (Figure 4 (A) and S8 of Paper 2). Constant lipid exchange within such a time scale for cation-π interactions highlights the dynamic nature of these interactions in protein-membrane binding.

5.2.4 Effect of K44A mutation results in more than electrostatics effect

To understand the effect of mutation of charged residues on the membrane bound *Bt*PI-PLC, we choose the K44A mutant of the same protein. As demonstrated in the previous section, K44A mutation has the most dramatic effect on the membrane association among all the charged residue mutation close to the IBS. Mutated *Bt*PI-PLC (K44A) was manually docked to a mixed bilayer (X_{PC}=0.8). The mutant *Bt*PI-PLC remains membrane bound without any significant altering of secondary structure with an average RMSD of the backbone equal to 1.5±0.2 Å. The RMSD is calculated along the simulation time (500 ns) with the minimized *Bt*PI-PLC mutant (K44A)

structure. The average depth of anchoring is similar to the wild type simulations (Table S7 of Paper 3). The cation-π interactions are also comparable. The difference arises from the loss of hydrogen bond network between the helix B and the lipid phosphate groups of the bilayer. Particularly, the hydrogen bond between backbones of two polar residues (Q40 and N41) with the bilayer is lost. And the long-lasting hydrogen bond of K44 side chain with the lipid phosphate groups is not present in the mutant simulation. This is an interesting example of a cooperative effect; how mutation can affect the cooperativity and as a consequence perturb the membrane binding affinity severely.

5.2.5 Lessons learned

Some of the interesting observations from *Bt*PI-PLC membrane bound simulations are listed below.

Aromatic residues mediate cation-π interactions when they are localized above or close to the phosphate plane (Table 5.1), only localization just at the phosphate plane (W47) or above the phosphate plane (Y246, Y88, Y251, and Y204) ensures cation-π interactions. However, localization at the interface is not the only determining factor for cation-π interactions. Localization of tryptophan at the interface for mediating cation-π interaction is consistent with previous observation for interfacial tryptophan mediated cation-π interaction for transmembrane protein[5]. The localization of the residues below the phosphate plane indicates hydrophobic contact, which is observed for one of the tryptophan residues (W242).

We do not observe any recruitment (tight binding) of PC lipids with the aromatic residues. If that were the case, the occupancy would most probably not be linear with X_{PC}. This implies that the cation-π interaction is opportunistic and occurs stochastically in the presence of PC lipids.

Phenylalanine can in principle engage in cation-π interaction with PC lipids. The IBS of *Bt*PI-PLC contains only one phenylalanine (F37) at close sequential proximity. However, this residue is localized above the phosphate plane (more than 10 Å above)

in all the simulations we performed irrespective of the bilayer composition, and interacts with the solvent. Protein IBS containing phenylalanine have been reported to insert below the phosphate plane and mediate hydrophobic contact[107, 112]. For PR3, it has been observed by our group that phenylalanine engaged in cation-π interaction with pure DMPC or DMPC:DMPG 50:50 mixture[65]. The occupancies for phenylalanine-choline cation-π interactions were at most 20% with pure DMPC bilayer (simulation time 200 ns, analysis on last 100 ns) and at most 10% with DMPC:DMPG 50:50 mixture (simulation time 300 ns, analysis on last 100 ns) [65]. Later work on PR3 binding with POPC lipids (again by our group) has shown that the occupancies (ca. 5%, simulation time 500 ns and analysis on last 300 ns) are very low and hence considered not significant[107]. Further work is needed to investigate the role of phenylalanine residue on cation-π interactions. However, based on our work and literature, we do not find the phenylalanines to engage in long lasting cation-π interactions that would indeed strengthen the binding. Rather they partition in the hydrophobic core and strengthen the binding by means of hydrophobic contact/interaction[107, 112].

Hydrophobic contacts are mediated not only by hydrophobic (and aromatic) residues, but also by other residues i.e. polar and charged residues. We find that the charged residues can mediate hydrophobic contacts with the bilayer core if they are localized at the interface (K44). This observation is also consistent with the previous work on PR3 [65, 107]. Such behavior of charged residues can be attributed to "snorkeling" [113-115] where the charged part of the residues localizes at the interfacial region of the membrane and the aliphatic part interacts with the hydrocarbon core of the membrane.

One of the most interesting observations is the large effect of K44A mutation on *Bt*PI-PLC binding. First, this indicates that the contributions from basic amino acids are not equal. Other charged residue mutations (those localized far from the membrane interface- K38, R71, K279) to alanine have almost the same effect among themselves on the measured binding affinity (K_d) using FCS experiments (by our collaborators- Table S8 of Paper 3) and hence on the calculated ΔΔG. What makes

K44 different is the localization with respect to the membrane, in other words it localizes at the interface. Because of that, it also participates in hydrophobic contact with the lipid bilayer core. Also, the dielectric of the interfacial region should play a role. So, the effect of mutating this K44 has a striking effect on the binding affinity. One can also try to correlate such behavior with the transfer free energy of lysine residues to bilayer interface[116, 117]. The transfer free energy of lysine side chain from water to the interface of DOPC bilayer is -4.4 kcal/mol, and for alanine side chain is -1.6 kcal/mol [117]. Thus, the energetic cost associated with the mutation of a lysine to an alanine at the interface is approximately 2.8 kcal/mol, which is the difference between the two above-mentioned values. The calculated $\Delta\Delta G$ value is 2.4 kcal/mol for mutating lysine to alanine (K44A mutant) compared to wild type *Bt*PI-PLC binding with mixed anionic vesicle (X_{PC}=0.8). With these comparisons, we confirm that the basic residues contribution are position-dependent and are not equal.

5.3 Towards an accurate modeling of cation-π interactions between tyrosine and choline

To understand the nature of cation-π interaction between tyrosine and choline, we use a model system consisting of phenol and tetramethylammonium (TMA). Phenol and TMA are analogues of tyrosine and choline, respectively. We build a potential energy surface (PES) (Figure 1 of Paper 4). First, we benchmark different wavefunction and density-function based methods varying the basis sets against large basis set Coupled Cluster (CCSD(T)) calculations. Further, we perform energy decomposition analysis using Symmetry Adapted Perturbation Theory (SAPT) on the whole PES. We also examine the solvent effect using highly polar solvent (i.e. water). To check the reliability of force fields, we compare the gas phase QM PES with the PES generated using the CHARMM additive force field and the Drude polarizable force field. We propose a modified set of parameters for CHARMM additive force field. A choline bound *S. aureus* PI-PLC structure[110], and a membrane bound phospholipase[6] are used to test these different force fields using molecular dynamics (MD) simulations.

5.3.1 Benchmarking QM level of theory

5.3.1.1 BSSE corrections are important

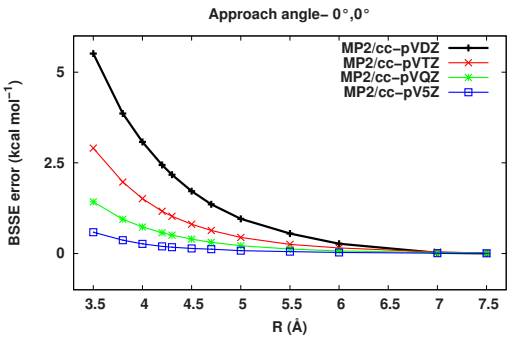
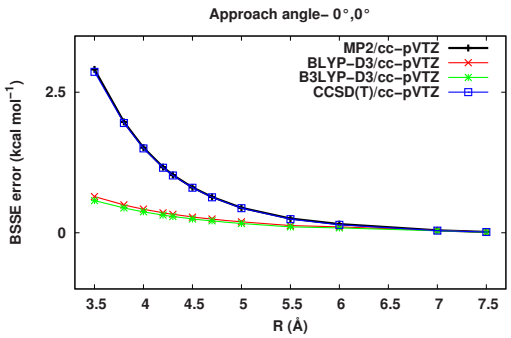
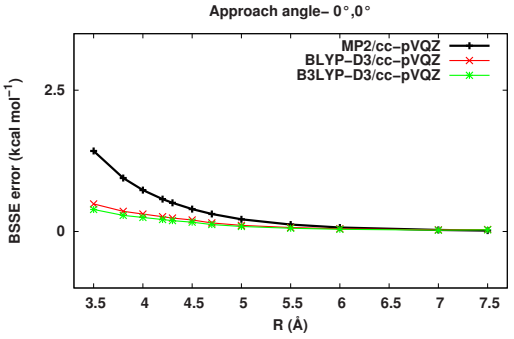


Figure 5.3.1: BSSE error for MP2 methods using different basis sets.



(A)



(B)

Figure 5.3.2: BSSE error for different wavefunction and density function based methods, (A) for cc-pVTZ basis sets, and (B) for cc-pVQZ basis sets.

We benchmark different QM methods before generating the PES. We show the importance of BSSE error for wavefunction based methods such as MP2 (Figure 5.3.1). Interaction energy evaluated without BSSE correction and a small basis set (cc-pVDZ) will lead to an overestimation of interaction energy of almost 2.5 kcal mol^{-1} near the minima (Figure 5.3.1). The smaller the basis set is, the larger the BSSE error, as expected. Caldwell and Kollman calculated the interaction energy of benzene-TMA complex using MP2/6-31G* level of theory, compared to experimental enthalpy of complex formation and found that the interaction energy is overestimated 6 kcal mol^{-1} [118]. The authors did not use BSSE corrections.

Comparison with literature reveals that the interaction energy calculated by Caldwell and Kollman was overestimated by around 6 kcal mol^{-1}, i.e. see the works by Rapp et al.[119], Kim et al.[120], or Pullman et al.[121]. By systematically varying the basis sets, we demonstrate the importance of BSSE correction for wavefunction based methods. Our comparison of different density function based methods using different basis sets shows that DFT methods are less affected by BSSE error when used with dispersion corrections (Figure 5.3.2). For cc-pVTZ basis set, BSSE error at the energy minima for wavefunction based method is ~1.02 kcal mol^{-1}, and for DFT based method is ~0.32 kcal mol^{-1} (Figure 5.3.2. (A)). For cc-pVQZ basis set, BSSE errors are 0.51 kcal mol^{-1} and 0.23 kcal mol^{-1} respectively for wavefunction and density function based methods (Figure 5.3.2 (B)). Based on these comparisons, we recommend to use a large basis set for both wavefunction or DFT based methods- at least cc-pVQZ basis set or equivalent for wavefunction based methods and at least cc-pVTZ basis set or equivalent for DFT based methods including dispersion correction.

5.3.1.2 SAPT2+/aug-cc-pVDZ level of theory performs well

We systematically compare interaction energy converged to complete basis limit obtained from different methods and benchmark against large basis set CCSD(T) results. Here we also compare variants of wavefunction based symmetry adapted perturbation theory (SAPT). We show that MP2 and BLYP-D3 (B3LYP-D3 results are equivalent to BLYP-D3) methods overestimate the interaction energy at the minima by around 1-1.5 kcal mol^{-1}. SAPT2+/aug-cc-pVDZ performs best compared to CCSD(T) (Figure 5.3.3). The interaction energy at the minima obtained by CCSD(T)/cc-pV5Z level (extrapolated) is -9.39 kcal mol^{-1} and SAPT2+/aug-cc-pVDZ level is -9.23 kcal mol^{-1}.

Calculations using SCF methods (HF) largely underestimate the interaction energy (> 6 kcal mol^{-1}, data not shown). Due to lack of electron correlation and hence dispersion, HF calculations are not reliable on such system. MP2 or dispersion corrected DFT is recommended for a qualitative overview on such systems if large basis CCSD(T) calculations are not possible to perform.

5.3.2 Dispersion contributes significantly

Using SAPT2+/aug-cc-pVDZ, we generate the potential energy curves for different approach angles. We perform energy decompositions and also identify the location of the minima. Dispersion contributes more than 35% for any approach angle (Table 1 of Paper 4). The ranking of energetic components at the minima for different approach angles are provided in Figure 5.3.4. By performing this energy decomposition analysis, we demonstrate that the dispersion can indeed contribute most predominantly in cation-π complex of phenol-TMA. We also quantify the dispersion contributions for different approach angles.

Previously, role of induction has been highlighted and quantified by Soteras et al.[122] and Marshall et al.[123] for other cation-π complexes. Role of dispersion in benzene-TMA [120, 121] and phenol-TMA [121] complex has been reported by taking the energy difference between SCF and MP2 energy values. However, no systematic energy decompositions are available for phenol-TMA complex. We have performed a systematic decomposition for phenol-TMA complex, which can be used to improve the force field parameters.

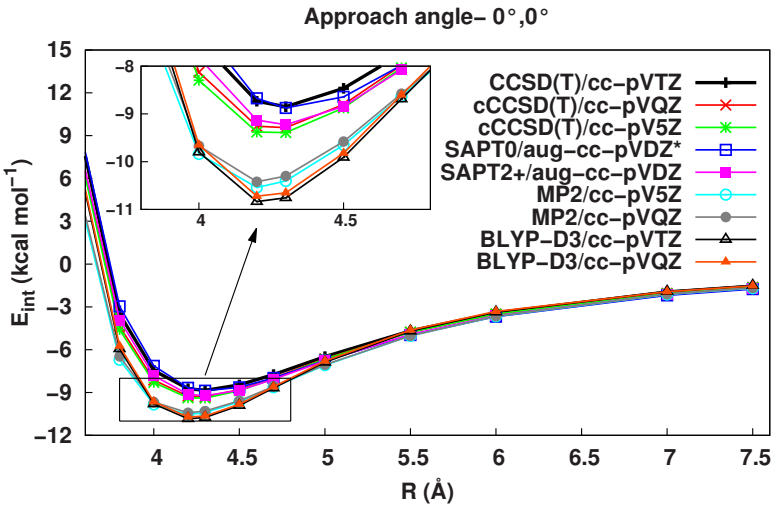


Figure 5.3.3: Interaction energies obtained using wavefunction based SAPT methods are compared to different levels of theory, both DFT and wavefunction based methods. SAPT2+/aug-cc-pVDZ nicely reproduce the large basis set CCSD(T) interaction energies.

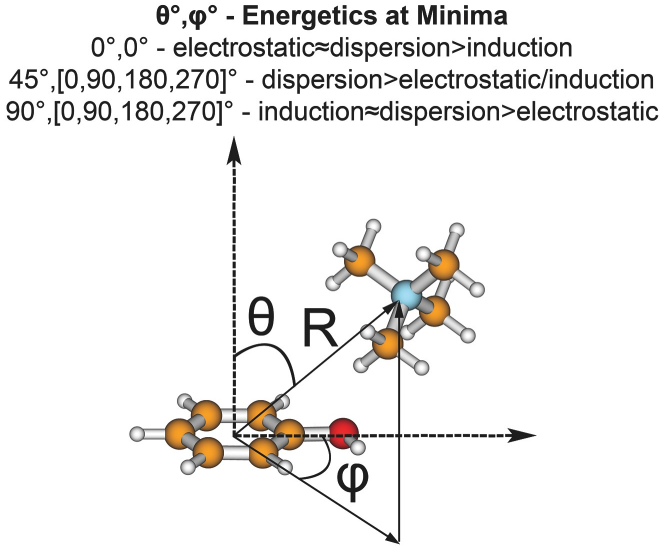


Figure 5.3.4: Contribution of different energetic components at the minima for different approach angles.

5.3.3 Phenol-TMA cation-π complex is stronger than Phenol-ammonium in a polar solvent

To understand the effect of solvation on cation-π complex of phenol-TMA, we used the COSMO solvation model and water as the solvent. We check the solvation effect on the minima for different approach angles and find that the diagonal approach (θ=45°) is favorable in solution phase. For the perpendicular (also the most favorable approach angle) approaching, the strength of interaction in water is -3.33 kcal mol^{-1} (Table 2 of Paper 4). The interaction energy profiles for both gas phase and in water indicate that the interaction energy will attenuate close to 10 Å (Figure 5.3.5).

Rapp et al. showed that the gradual methylation of ammonium in ammonium-benzene complex reduces the interaction strength by more than half from -19 kcal mol^{-1} to -8.8 kcal mol^{-1} in gas phase [119]. So, the benzne-TMA (-8.8 kcal mol^{-1}) and phenol-TMA (-9.23 kcal mol^{-1}) interaction energies at gas phase are comparable. Based on these results and comparing them to other literature[123], we show that the solution phase cation-π interaction of TMA is stronger than ammonium as interaction of ammonium-benzene complex in solution phase (water) is reported to be -0.4 kcal mol^{-1}. So, we expect that cation-π interaction between aromatic amino acids with PC lipids will be stronger than PE lipids in water. Cheng et al. also reported such behavior performing binding experiments of protein to SUVs [110].

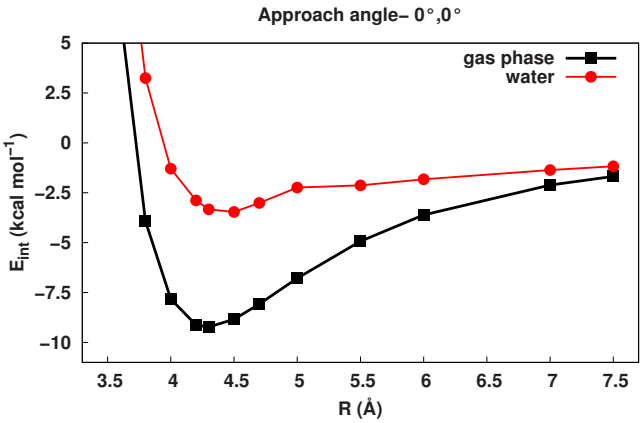


Figure 5.3.5: Interaction energy profile both in gas phase and in water. Energy profile in gas phase is calculated at SAPT2+/aug-cc-pVDZ level, and in water is calculated at BLYP/cc-pVTZ level using COSMO solvation model.

5.3.4 Additive CHARMM force field underestimates the interaction energy and Drude polarizable force field mostly overestimates the geometry

Based on the comparison of the PES generated using QM calculations and with MM force fields, we show that additive CHARMM force field underestimates the interaction energy between phenol and TMA (Figure 4-6 & Table 3 of Paper 4). However, the location of the energy minima calculated is close to the energy minima

identified by QM calculations. Drude polarizable force field performs better than the CHARMM force field in terms of interaction energy. However, the locations of the minima are not preserved and are shifted towards the phenol in most cases (Figure 4-6 & Table 3 of Paper 4).

5.3.5 Improvement of force field parameters for cation-π interactions in CHARMM additive force field and testing

To improve the force field parameters for CHARMM additive force field, we use altered Lennard-Jones potential and specifically deepen the potential well depth for specific atoms (Table 4 & Figure S8 of Paper 4). With this approach we make sure that the minima location is preserved and other interactions are not affected. The resulting force field is termed as "CHARMM-ff-mod". Then we generate the PES with CHARMM-ff-mod force field and find that the interaction energy profile is improved for most of the approach angles (Figure 7-9 and Table 5 of Paper 4). Previous works to improve the force field parameters for cation-π interactions always attempted to reproduce only the perpendicular approach of cation on aromatic rings [118, 124-126]. With the combination of high-level QM calculations and solvation effects on energy minima we show that the other approach angles are still favorable and worth consideration during parameterizations or improvements of force fields. Later, we perform MD simulations with these new parameters and compare them with additive CHARMM force field and Drude polarizable force field.

We choose choline bound structure of engineered *S. aureus* PI-PLC, where there are two cholines bound to the engineered tyrosine cage (Figure 5.3.6). The simulation results show that the CHARMM-ff-mod outperforms the additive CHARMM and Drude polarizable force field (Figure 11 of Paper 4). The observed ligand residence time from the MD simulations suggests that the binding site 2 has higher affinity for choline than the binding site 1. This observation is also in agreement with the experimental observations that the site 2 has higher affinity than the site 1[110].

Also, we perform test on a membrane bound phospholipase using CHARMM-ff and CHARMM-ff-mod parameters. We compared the cation-π occupancies from these

simulations with the experimental observable. The experimental observable was evaluated ΔΔG calculated from apparent K_d values measured by FCS for *Bt*PI-PLC WT and single tyrosine mutants to alanines[6]. The comparison between these data sets is qualitative with the CHARMM-ff whereas it is near quantitative with the CHARMM-ff-mod parameter sets (Figure 12 of Paper 4).

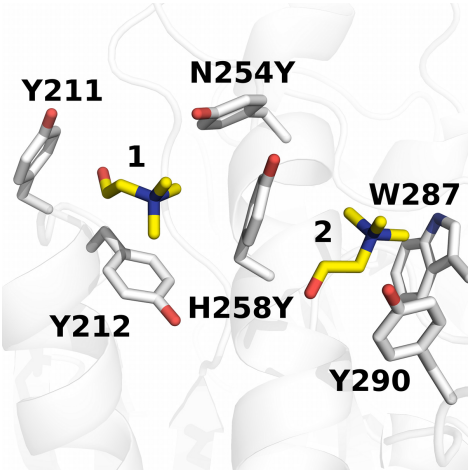


Figure 5.3.6: Choline bound structure of S. aureus PI-PLC (engineered) used for testing force fields. For clarity, waters and ions are not shown. Both the binding sites and the ligands are represented with the numbers 1 and 2.

5.3.6 Summary of the findings

From the generated PES, we find that dispersion contributes significantly for any approach angle of TMA with respect to the phenol ring. Due to this dispersion contribution, cation-π interaction between tyrosine-PC lipids will be stronger than tyrosine-PE lipids in solution phase. The interaction between tyrosine-choline attenuates to zero at around a distance of 10 Å.

Our comparison of variants of CHARMM force fields (additive and polarizable) with QM results reveal that additive CHARMM force field underestimates the interaction energies and Drude polarizable force field overestimates the complex geometries in

gas phase. Based on our analysis, we propose a modified set of parameters for the additive CHARMM force field to model cation-π interactions. These parameters perform well in the MD simulations of test cases. It would be indeed interesting to further test these parameters on a wide range of cation-π complex.

6. Conclusions and future perspectives

6.1 Conclusions

In this thesis, we have examined different proteins, which brings new insights in protein peripheral membrane binding. The investigations revealed limitations of the existing peripheral protein-membrane binding models to describe these peripheral binders. For this reason, we incorporate additional information from our work to the existing description of protein peripheral membrane binding. The outcome is an energy/interaction diagram based on noncovalent interactions. The works described in the previous chapters aid in constructing this diagram (Figure 6.1).

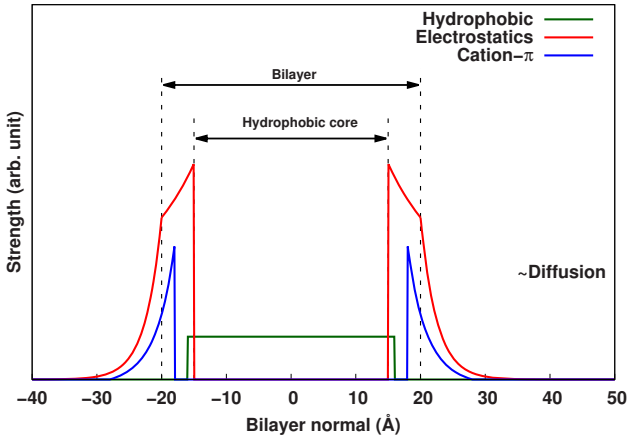


Figure 6.1: Energy diagram in terms of protein-lipid interactions.

The differences in the previous models and the improvements introduced by us are:

1. Cation-π interactions are established as a membrane-binding mechanism. The distances up to which cation-π interaction can regulate are also demonstrated using QM calculations both in the gas phase and the solution phase. During the course of

these works, a modified set of parameters is proposed to accuratly model tyrosine-choline cation-π interactions using CHARMM force field.

2. Hydrophobic interactions are not limited to hydrophobic/aromatic amino acids. In the previous works[46, 48, 51, 59], hydrophobic interactions between the protein and the membrane are described to take place by only hydrophobic/aromatic amino acids. Here, we demonstrate that other type of amino acids can mediate hydrophobic contacts using aliphatic groups of the side chains with the bilayer core if they are localized at the interface. Such observation is consistent with the "snorkeling" of charged residues[113-115]. Snorkeling behavior is mostly demonstrated for transmembrane helices or proteins. Our investigations highlight that it is not unlikely for peripheral proteins.

3. Membrane-water interface is an important region of interest in terms of peripheral proteins membrane binding. In previous works, the role of the membrane-water interface was largely neglected or under represented. However, we find the importance of the interface to be quite substantial in terms of residue localizations. In fact, this is the place where lipid targeting and head groups recognition takes place which has been highlighted by both Cho and Stahelin[48], and Johnson and Cornell[46] for different membrane targeting domains. However, localization of different types of residues and their energetic contributions are not highlighted. We find that the charged residues localize at the interface and contribute significantly in binding energetics, which we describe as more than electrostatic effects in previous chapter. Also, the aromatic residues (Trp, Tyr) localize at the interface or slightly above the interface and can mediate cation-π interactions with particular lipid headgroups.

4. Electrostatics play a key role in peripheral protein membrane binding, which has been shown by several pioneering works by others[49, 51, 52, 57-59, 127]. Earlier models of peripheral protein membrane binding describe electrostatics as a long-range interaction[46, 48]. In principle, this is true in gas phase. Here, we show that in an ionic solution similar to cellular environment, electrostatic interactions attenuate

faster. In earlier works, as described, proteins reach the membrane surface using diffusion and long-range electrostatic forces[46, 48]. The range is not clearly defined. The electrostatic forces were described between negatively charged phospholipids and the positively charged amino acids (Arg, Lys)[51, 59] and their contributions are suggested to be equal (~1.4 kcal/mol)[49]. Again, in these models, electrostatics regulates close to the interface (separation of ~3 Å)[46, 52] or at best up to the interface[48].

We show that electrostatics attenuate within ~ 20 Å of the membrane surface, and our observation is consistent with the range of electrostatic forces for biomolecules by Fogolari et al.[96] using continuum electrostatics (10~20 Å). Moreover, the electrostatics contributions are not equal. It greatly depends on the position of the charged amino acids with respect to the membrane. If a charged residue is localized at the interface, it will have a higher interaction strength than the one localized in the water. Previous works were mostly focused on proteins containing basic amino acids cluster. As a consequence, it was unlikely to observe the position dependent behavior of basic residues in those works.

Table 6.1: Role of amino acid residues in protein peripheral membrane binding.

Interactions	Previous works	Our investigations
Electrostatics	Arg, Lys	Arg, Lys
Hydrophobic	Hydrophobic/Aromatic residues	Hydrophobic, Aromatic (Phe, Trp), Polar (Gln, Asn, Thr), Charged (Lys), Other (Pro) residues
Cation-π	Trp*	Tyr, Trp

* work by others, for transmembrane protein[5]

5. Finally, we summarize the role of amino acid residues and compare them to the previous works in Table 6.1. We incorporate the information from our investigations. This need to be further characterized and extended.

6.2 Future perspectives

Here we outline possible future works:

1. Generalizations: The proposed interaction diagram needs to be generalized. The generalization will establish this diagram as a model. In that respect, it would be indeed interesting to test other amphitropic proteins (those with a net negative charge and low electrostatic partitioning) whether they follow the proposed description. Also it would be interesting to check if their function correlates well with the low electrostatic partitioning. In other words, if these proteins need to have low electrostatic partitioning to frequently dissociate from the membrane to perform their functions.

2. Nature of interaction between cation-π complexes greatly depends on the cation and the aromatic ring involved. For example, cation-π interactions between benzene and Na^+ or benzene and TMA complexes are not the same. The first one is dominated by electrostatics, and the second one is a mixed influence complex. Because of this reason, systematic checking of other cation-π interactions e.g. tryptophan-choline, PS mediated cation-π with other aromatic amino acids is necessary. Again, ability of coarse-grained (CG) models to capture cation-π interactions should be tested. Coarse-grained (CG) MD simulations of *Bt*PI-PLC with the Martini CG model and a modified Martini version do not lead to membrane binding with a PC:PG mixture of 80:20 (results not shown). Preliminary attempt using the modified Martini version by changing the nonbonded interaction (LJ potential) between choline headgroup and the aromatic ring did not lead to complex formation. It would be interesting to systematically investigate further the Martini and other CG models.

References

[1] P. Atkins and J. d. Paula, *Physical Chemistry*, Eighth ed. New York: W. H. Freeman and Company, 2006.

[2] H. Lodish*, et al.*, *Molceular Cell Biology*, 4th ed. New York: W.H. Freeman, 2000.

[3] D. Leckband and J. Israelachvili, "Intermolecular forces in biology," *Quarterly Reviews of Biophysics,* vol. 34, pp. 105-267, 2001.

[4] J. C. Ma and D. A. Dougherty, "The Cation-π Interaction," *Chemical Reviews,* vol. 97, pp. 1303-1324, 1997.

[5] F. N. R. Petersen*, et al.*, "Interfacial Tryptophan Residues: A Role for the Cation-π Effect?," *Biophysical Journal,* vol. 89, pp. 3985-3996, 2005.

[6] C. Grauffel*, et al.*, "Cation−π Interactions As Lipid-Specific Anchors for Phosphatidylinositol-Specific Phospholipase C," *Journal of the American Chemical Society,* vol. 135, pp. 5740-5750, 2013.

[7] B. Yang*, et al.*, "Quantifying Transient Interactions between Bacillus Phosphatidylinositol-Specific Phospholipase-C and Phosphatidylcholine-Rich Vesicles," *Journal of the American Chemical Society,* vol. 137, pp. 14-17, 2015.

[8] A. J. Stone, "Intermolecular Potentials," *Science,* vol. 321, pp. 787-789, 2008.

[9] A. Stone, *The Theory of Intermolecular Forces*: OUP Oxford, 2013.

[10] A. van der Avoird*, et al.*, "Ab initio studies of the interactions in Van der Waals molecules," in *Van der Waals Systems*. vol. 93: Springer Berlin Heidelberg, 1980, pp. 1-51.

[11] P. E. S. Wormer and A. van der Avoird, "Intermolecular Potentials, Internal Motions, and Spectra of van der Waals and Hydrogen-Bonded Complexes," *Chemical Reviews,* vol. 100, pp. 4109-4144, 2000.

[12] P. Hobza and R. Zahradník, "Van der Waals systems: Molecular orbitals, physical properties, thermodynamics of formation and reactivity," in *Van der Waals Systems*. vol. 93: Springer Berlin Heidelberg, 1980, pp. 53-90.

[13] W. Klopper*, et al.*, "Ab initio computations close to the one-particle basis set limit on the weakly bound van der Waals complexes benzene-neon and benzene-argon," *The Journal of Chemical Physics,* vol. 101, p. 9747, 1994.

[14] S. Rybak*, et al.*, "Many-body symmetry-adapted perturbation theory of intermolecular interactions. H2O and HF dimers," *The Journal of Chemical Physics,* vol. 95, p. 6576, 1991.

[15] B. Jeziorski*, et al.*, "Perturbation Theory Approach to Intermolecular Potential Energy Surfaces of van der Waals Complexes," *Chemical Reviews,* vol. 94, pp. 1887-1930, 1994.

[16] E. G. Hohenstein, "Implementation and applications of density-fitted symmetry-adapted perturbation theory," PhD Thesis, School of Chemistry and Biochemistry, Georgia Institute of Technology, 2011.

[17] E. G. Hohenstein and C. D. Sherrill, "Wavefunction methods for noncovalent interactions," *Wiley Interdisciplinary Reviews: Computational Molecular Science,* vol. 2, pp. 304-326, 2012.

[18] A. D. Buckingham and B. D. Utting, "Intermolecular Forces," *Annual Review of Physical Chemistry,* vol. 21, pp. 287-316, 1970.

[19] F. Jensen, *Introduction to computational chemistry*, Second ed.: John Wiley & Sons, Ltd, 2007.

[20] F. London, "The general theory of molecular forces," *Transactions of the Faraday Society,* vol. 33, pp. 8-26, 1937.

[21] F. London, "100 Years of Physical Chemistry," in *100 Years of Physical Chemistry: A Collection of Landmark Papers*: The Royal Society of Chemistry, 2003, pp. 1-22.
[22] A. D. Mackerell, "Empirical force fields for biological macromolecules: Overview and issues," *Journal of Computational Chemistry,* vol. 25, pp. 1584-1604, 2004.
[23] A. D. MacKerell*, et al.*, "All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins," *J. Phys. Chem. B,* vol. 102, pp. 3586-3616, 1998.
[24] S. Patel and C. L. Brooks, "CHARMM fluctuating charge force field for proteins: I parameterization and application to bulk organic liquid simulations," *Journal of Computational Chemistry,* vol. 25, pp. 1-16, 2004.
[25] J. W. Ponder*, et al.*, "Current Status of the AMOEBA Polarizable Force Field," *Journal of Physical Chemistry B,* vol. 114, pp. 2549-2564, 2010.
[26] G. Lamoureux*, et al.*, "A polarizable model of water for molecular dynamics simulations of biomolecules," *Chemical Physics Letters,* vol. 418, pp. 245-249, 2006.
[27] G. Lamoureux and B. Roux, "Modeling induced polarization with classical Drude oscillators: Theory and molecular dynamics simulation algorithm," *Journal of Chemical Physics,* vol. 119, p. 3025, 2003.
[28] A. Savelyev and A. D. MacKerell, "Balancing the Interactions of Ions, Water, and DNA in the Drude Polarizable Force Field," *The Journal of Physical Chemistry B,* vol. 118, pp. 6742-6757, 2014.
[29] J. Chowdhary*, et al.*, "A Polarizable Force Field of Dipalmitoylphosphatidylcholine Based on the Classical Drude Model for Molecular Dynamics Simulations of Lipids," *The Journal of Physical Chemistry B,* vol. 117, pp. 9142-9160, 2013.
[30] P. E. M. Lopes*, et al.*, "Polarizable Force Field for Peptides and Proteins Based on the Classical Drude Oscillator," *Journal of Chemical Theory and Computation,* vol. 9, pp. 5430-5449, 2013.
[31] C. M. Baker, "Polarizable force fields for molecular dynamics simulations of biomolecules," *Wiley Interdisciplinary Reviews: Computational Molecular Science,* vol. 5, pp. 241-254, 2015.
[32] S. Saha and G. N. Sastry, "Cooperative or Anticooperative: How Noncovalent Interactions Influence Each Other," *Journal of Physical Chemistry B,* 2015.
[33] B. Alberts*, et al.*, *Molecular Biology of the Cell*, Fifth ed. New York: Garland Science, 2008.
[34] P. A. Janmey and P. K. J. Kinnunen, "Biophysical properties of lipids and dynamic membranes," *Trends in Cell Biology,* vol. 16, pp. 538-546, 2006.
[35] Y. Kimura and A. Ikegami, "Local dielectric properties around polar region of lipid bilayer membranes," *J. Membr. Biol.,* vol. 85, pp. 225-231, 1985.
[36] G. Gramse*, et al.*, "Nanoscale Measurement of the Dielectric Constant of Supported Lipid Bilayers in Aqueous Solutions with Electrostatic Force Microscopy," *Biophysical Journal,* vol. 104, pp. 1257-1262, 2013.
[37] G. van Meer*, et al.*, "Membrane lipids: where they are and how they behave," *Nat Rev Mol Cell Biol,* vol. 9, pp. 112-124, 2008.
[38] G. van Meer and A. I. P. M. de Kroon, "Lipid map of the mammalian cell," *Journal of Cell Science,* vol. 124, pp. 5-8, 2011.
[39] H. I. Ingólfsson*, et al.*, "Lipid Organization of the Plasma Membrane," *Journal of the American Chemical Society,* vol. 136, pp. 14554-14559, 2014.
[40] E. Gorter and F. Grendel, "ON BIMOLECULAR LAYERS OF LIPOIDS ON THE CHROMOCYTES OF THE BLOOD," *The Journal of Experimental Medicine,* vol. 41, pp. 439-443, 1925.
[41] S. J. Singer and G. L. Nicolson, "The Fluid Mosaic Model of the Structure of Cell Membranes," *Science,* vol. 175, pp. 720-731, 1972.

[42] D. M. Engelman, "Membranes are more mosaic than fluid," *Nature,* vol. 438, pp. 578-580, 2005.
[43] F. M. Goñi, "The basic structure and dynamics of cell membranes: An update of the Singer–Nicolson model," *Biochimica et Biophysica Acta (BBA) - Biomembranes,* vol. 1838, pp. 1467-1476, 2014.
[44] A. Kusumi*, et al.*, "Hierarchical mesoscale domain organization of the plasma membrane," *Trends in Biochemical Sciences,* vol. 36, pp. 604-615, 2011.
[45] G. A. Petsko and D. Ringe, *Protein Structure and Function*: New Science Press, 2004.
[46] J. E. Johnson and R. B. Cornell, "Amphitropic proteins: regulation by reversible membrane interactions (Review)," *Molecular Membrane Biology,* vol. 16, pp. 217-235, 1999.
[47] P. Burn, "Talking point Amphitropic proteins: a new class of membrane proteins," *Trends in Biochemical Sciences,* vol. 13, pp. 79-83, 1988.
[48] W. Cho and R. V. Stahelin, "Membrane-Protein Interactions in Cell Signaling and Membrane Trafficking," *Annual Review of Biophysics and Biomolecular Structure,* vol. 34, pp. 119-151, 2005.
[49] J. Kim*, et al.*, "Binding of peptides with basic residues to membranes containing acidic phospholipids," *Biophysical Journal,* vol. 60, pp. 135-148, 1991.
[50] M. Luckey, *Membrane Structural Biology: With Biochemical and Biophysical Foundations*. Cambridge: Cambridge University Press, 2008.
[51] A. Mulgrew-Nesbitt*, et al.*, "The role of electrostatics in protein–membrane interactions," *Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids,* vol. 1761, pp. 812-826, 2006.
[52] D. Murray*, et al.*, "Electrostatics and the Membrane Association of Src: Theory and Experiment," *Biochemistry,* vol. 37, pp. 2145-2159, 1998.
[53] P. V. Escribá*, et al.*, "Membranes: a meeting point for lipids, proteins and therapies," *Journal of Cellular and Molecular Medicine,* vol. 12, pp. 829-875, 2008.
[54] R. V. Stahelin, "Lipid binding domains: more than simple lipid effectors," *Journal of Lipid Research,* vol. 50, pp. S299-S304, 2009.
[55] M. N. Teruel and T. Meyer, "Translocation and Reversible Localization of Signaling Proteins: A Dynamic Future for Signal Transduction," *Cell,* vol. 103, pp. 181-184, 2000.
[56] H. Wehbi*, et al.*, "Investigating the Interfacial Binding of Bacterial Phosphatidylinositol-Specific Phospholipase C," *Biochemistry,* vol. 42, pp. 9374-9382, 2003.
[57] D. Murray, "The Role of Electrostatic Interactions in the Regulation of the Membrane Association of G Protein beta gamma Heterodimers," *Journal of Biological Chemistry,* vol. 276, pp. 45153-45159, 2001.
[58] D. Murray*, et al.*, "Electrostatic Properties of Membranes Containing Acidic Lipids and Adsorbed Basic Peptides: Theory and Experiment," *Biophysical Journal,* vol. 77, pp. 3176-3188, 1999.
[59] D. Murray*, et al.*, "The role of electrostatic and nonpolar interactions in the association of peripheral proteins with membranes," in *Current Topics in Membranes*. vol. Volume 52, T. J. M. Sidney A. Simon, Ed.: Academic Press, 2002, pp. 277-307.
[60] S. M. Singh and D. Murray, "Molecular modeling of the membrane targeting of phospholipase C pleckstrin homology domains," *Protein Science,* vol. 12, pp. 1934-1953, 2003.

[61] H. L. Wright, *et al.*, "Neutrophil function in inflammation and inflammatory diseases," *Rheumatology,* vol. 49, pp. 1618-1631, 2010.
[62] V. Witko-Sarsat, *et al.*, "A Large Subset of Neutrophils Expressing Membrane Proteinase 3 Is a Risk Factor for Vasculitis and Rheumatoid Arthritis," *Journal of the American Society of Nephrology,* vol. 10, pp. 1224-1233, 1999.
[63] W. H. Goldmann, *et al.*, "Interaction of purified human proteinase 3 (PR3) with reconstituted lipid bilayers," *European Journal of Biochemistry,* vol. 261, pp. 155-162, 1999.
[64] E. Hajjar, *et al.*, "Computational prediction of the binding site of proteinase 3 to the plasma membrane," *Proteins: Structure, Function, and Bioinformatics,* vol. 71, pp. 1655-1669, 2008.
[65] T. Broemstrup and N. Reuter, "How does Proteinase 3 interact with lipid bilayers?," *Physical Chemistry Chemical Physics,* vol. 12, p. 7487, 2010.
[66] T. Broemstrup, "Peripheral membrane binding of Proteinase 3: In silico description of amino acid specific binding interactions and their lipid type dependency," PhD Thesis, Department of Informatics, University of Bergen, 2010.
[67] J. Cheng, *et al.*, "Does Changing the Predicted Dynamics of a Phospholipase C Alter Activity and Membrane Binding?," *Biophysical Journal,* vol. 104, pp. 185-195, 2013.
[68] X. Shi, *et al.*, "Modulation of Bacillus thuringiensis Phosphatidylinositol-specific Phospholipase C Activity by Mutations in the Putative Dimerization Interface," *Journal of Biological Chemistry,* vol. 284, pp. 15607-15618, 2009.
[69] S. Hirose, *et al.*, "Mammalian glycosylphosphatidylinositol-anchored proteins and intracellular precursors," in *Methods in Enzymology*. vol. 250: Academic Press, 1995, pp. 582-614.
[70] M. Pu, *et al.*, "Fluorescence Correlation Spectroscopy of Phosphatidylinositol-Specific Phospholipase C Monitors the Interplay of Substrate and Activator Lipid Binding," *Biochemistry,* vol. 48, pp. 6835-6845, 2009.
[71] C. Zhou, *et al.*, "Allosteric Activation of Phosphatidylinositol-Specific Phospholipase C: Specific Phospholipid Binding Anchors the Enzyme to the Interface," *Biochemistry,* vol. 36, pp. 10089-10097, 1997.
[72] M. Pu, *et al.*, "Correlation of Vesicle Binding and Phospholipid Dynamics with Phospholipase C Activity: insights into phosphatidylcholine activation and surface dilution inhibition," *Journal of Biological Chemistry,* vol. 284, pp. 16099-16107, 2009.
[73] S. Guo, *et al.*, "Role of Helix B Residues in Interfacial Activation of a Bacterial Phosphatidylinositol-Specific Phospholipase C," *Biochemistry,* vol. 47, pp. 4201-4210, 2008.
[74] J. Feng, *et al.*, "Role of Tryptophan Residues in Interfacial Binding of Phosphatidylinositol-specific Phospholipase C," *Journal of Biological Chemistry,* vol. 277, pp. 19867-19875, 2002.
[75] W. J. Hehre, *A guide to molecular mechanics and quantum chemical calculations*: Wavefunction Irvine, CA, 2003.
[76] M. Born and R. Oppenheimer, "Zur Quantentheorie der Molekeln," *Annalen der Physik,* vol. 389, pp. 457-484, 1927.
[77] D. Sherrill, "Lecture notes in Computational Chemistry," *Georgia Institute of Technology*.
[78] A. Szabo and N. S. Ostlund, *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*: Dover Publications, 1989.
[79] C. Møller and M. S. Plesset, "Note on an Approximation Treatment for Many-Electron Systems," *Physical Review,* vol. 46, pp. 618-622, 1934.

[80] J. F. Harrison, "Lecture notes in Computational Chemistry," *Michigan State University,* 2012.
[81] P. Hohenberg and W. Kohn, "Inhomogeneous Electron Gas," *Physical Review,* vol. 136, pp. B864-B871, 1964.
[82] W. Kohn and L. J. Sham, "Self-Consistent Equations Including Exchange and Correlation Effects," *Physical Review,* vol. 140, pp. A1133-A1138, 1965.
[83] J. M. Turney*, et al.*, "Psi4: an open-source *ab initio* electronic structure program," *Wiley Interdisciplinary Reviews: Computational Molecular Science,* vol. 2, pp. 556-565, 2012.
[84] D. Frenkel and B. Smit, *Understanding Molecular Simulation, Second Edition: From Algorithms to Applications (Computational Science)*: Academic Press, 2001.
[85] M. Tuckerman, *Statistical Mechanics: Theory and Molecular Simulation*: Oxford University Press, USA, 2010.
[86] M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids (Oxford Science Publications)*: Oxford University Press, 1989.
[87] M. Feig, "Biomolecular Solvation in Theory and Experiment," in *Modeling Solvent Environments*: Wiley-VCH Verlag GmbH & Co. KGaA, 2010, pp. 1-29.
[88] K. Vanommeslaeghe and A. D. MacKerell, "CHARMM additive and polarizable force fields for biophysics and computer-aided drug design," *Biochimica et Biophysica Acta, General Subjects,* vol. 1850, pp. 861-871, 2015.
[89] K. Vanommeslaeghe*, et al.*, "CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields," *Journal of Computational Chemistry,* vol. 31, pp. 671-690, 2010.
[90] A. Leach, *Molecular Modelling: Principles and Applications (2nd Edition)*: Prentice Hall, 2001.
[91] D. J. Hardy*, et al.*, "Multilevel Summation Method for Electrostatic Force Evaluation," *Journal of Chemical Theory and Computation,* vol. 11, pp. 766-779, 2015.
[92] H. C. Andersen, "Molecular dynamics simulations at constant pressure and/or temperature," *The Journal of Chemical Physics,* vol. 72, p. 2384, 1980.
[93] S. Nosé, "A unified formulation of the constant temperature molecular dynamics methods," *The Journal of Chemical Physics,* vol. 81, pp. 511-519, 1984.
[94] W. G. Hoover, "Canonical dynamics: Equilibrium phase-space distributions," *Physical Review A,* vol. 31, pp. 1695-1697, 1985.
[95] H. J. C. Berendsen*, et al.*, "Molecular dynamics with coupling to an external bath," *The Journal of Chemical Physics,* vol. 81, pp. 3684-3690, 1984.
[96] F. Fogolari*, et al.*, "The Poisson–Boltzmann equation for biomolecular electrostatics: a tool for structural biology," *Journal of Molecular Recognition,* vol. 15, pp. 377-392, 2002.
[97] M. K. Gilson and B. H. Honig, "Calculation of electrostatic potentials in an enzyme active site," *Nature,* vol. 330, pp. 84-86, 1987.
[98] K. A. Sharp and B. Honig, "Calculating total electrostatic energies with the nonlinear Poisson-Boltzmann equation," *J. Phys. Chem.,* vol. 94, pp. 7684-7692, 1990.
[99] A. D. Mackerell*, et al.*, "Extending the treatment of backbone energetics in protein force fields: Limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations," *Journal of Computational Chemistry,* vol. 25, pp. 1400-1415, 2004.
[100] J. B. Klauda*, et al.*, "Update of the CHARMM All-Atom Additive Force Field for Lipids: Validation on Six Lipid Types," *J. Phys. Chem. B,* vol. 114, pp. 7830-7843, 2010.

[101] N. A. Baker, *et al.*, "Electrostatics of nanosystems: Application to microtubules and the ribosome," *Proceedings of the National Academy of Sciences,* vol. 98, pp. 10037-10041, 2001.
[102] J. C. Phillips, *et al.*, "Scalable molecular dynamics with NAMD," *Journal of Computational Chemistry,* vol. 26, pp. 1781-1802, 2005.
[103] B. R. Brooks, *et al.*, "CHARMM: The biomolecular simulation program," *Journal of Computational Chemistry,* vol. 30, pp. 1545-1614, 2009.
[104] W. Humphrey, *et al.*, "VMD - Visual Molecular Dynamics," *J Molec Graphics,* vol. 14, pp. 33-38, 1996.
[105] M. Valiev, *et al.*, "NWChem: A comprehensive and scalable open-source solution for large scale molecular simulations," *Computer Physics Communications,* vol. 181, pp. 1477-1489, 2010.
[106] T. Lazaridis, "Implicit solvent simulations of peptide interactions with anionic lipid membranes," *Proteins: Structure, Function, and Bioinformatics,* vol. 58, pp. 518-527, 2005.
[107] A.-S. Schillinger, *et al.*, "Two homologous neutrophil serine proteases bind to POPC vesicles with different affinities: When aromatic amino acids matter," *Biochimica et Biophysica Acta (BBA) - Biomembranes,* vol. 1838, pp. 3191-3202, 2014.
[108] N. Ben-Tal, *et al.*, "Binding of small basic peptides to membranes containing acidic lipids: theoretical models and experimental results," *Biophysical Journal,* vol. 71, pp. 561-575, 1996.
[109] I. Gitlin, *et al.*, "Why Are Proteins Charged? Networks of Charge–Charge Interactions in Proteins Measured by Charge Ladders and Capillary Electrophoresis," *Angewandte Chemie International Edition,* vol. 45, pp. 3022-3060, 2006.
[110] J. Cheng, *et al.*, "The Cation-π Box Is a Specific Phosphatidylcholine Membrane Targeting Motif," *Journal of Biological Chemistry,* vol. 288, pp. 14863-14873, 2013.
[111] S. A. Abraham, *et al.*, "Do Cation···π Interactions Always Need to be 1:1?," *ChemPhysChem,* vol. 13, pp. 695-698, 2012.
[112] B. Rogaski and J. B. Klauda, "Membrane-Binding Mechanism of a Peripheral Membrane Protein through Microsecond Molecular Dynamics Simulations," *Journal of Molecular Biology,* vol. 423, pp. 847-861, 2012.
[113] E. Lindahl and M. S. P. Sansom, "Membrane proteins: molecular dynamics simulations," *Current Opinion in Structural Biology,* vol. 18, pp. 425-431, 2008.
[114] J. A. Killian and G. von Heijne, "How proteins adapt to a membrane-water interface," *Trends in Biochemical Sciences,* vol. 25, pp. 429-434, 2000.
[115] E. Strandberg and J. A. Killian, "Snorkeling of lysine side chains in transmembrane helices: how easy can it get?," *FEBS Letters,* vol. 544, pp. 69-73, 2003.
[116] D. Bonhenry, *et al.*, "Effects of Phospholipid Composition on the Transfer of a Small Cationic Peptide Across a Model Biological Membrane," *Journal of Chemical Theory and Computation,* vol. 9, pp. 5675-5684, 2013.
[117] J. L. MacCallum, *et al.*, "Partitioning of Amino Acid Side Chains into Lipid Bilayers: Results from Computer Simulations and Comparison to Experiment," *The Journal of General Physiology,* vol. 129, pp. 371-377, 2007.
[118] J. W. Caldwell and P. A. Kollman, "Cation-π Interactions: Nonadditive Effects Are Critical in Their Accurate Representation," *Journal of the American Chemical Society,* vol. 117, pp. 4177-4178, 1995.
[119] C. Rapp, *et al.*, "Cation-π interactions of methylated ammonium ions: A quantum mechanical study," *Proteins: Structure, Function, and Bioinformatics,* vol. 82, pp. 1494-1502, 2014.

[120] K. S. Kim*, et al.*, "On Binding Forces between Aromatic Ring and Quaternary Ammonium Compound," *Journal of the American Chemical Society,* vol. 116, pp. 7399-7400, 1994.
[121] A. Pullman*, et al.*, "Interaction of the Tetramethylammonium Ion with the Cycles of Aromatic Amino Acids beyond the SCF Ab Initio Level," *Journal of the American Chemical Society,* vol. 120, pp. 8553-8554, 1998.
[122] I. Soteras*, et al.*, "Induction effects in metal cation-benzene complexes," *Physical Chemistry Chemical Physics,* vol. 10, pp. 2616-2624, 2008.
[123] M. S. Marshall*, et al.*, "Potential Energy Curves for Cation-π Interactions: Off-Axis Configurations Are Also Attractive," *Journal of Physical Chemistry A,* vol. 113, pp. 13628-13632, 2009.
[124] K. Ansorg*, et al.*, "Cation−π Interactions: Accurate Intermolecular Potential from Symmetry-Adapted Perturbation Theory," *Journal of Physical Chemistry B,* vol. 117, pp. 10093-10102, 2013.
[125] H. Minoux and C. Chipot, "Cation-π Interactions in Proteins: Can Simple Models Provide an Accurate Description?," *Journal of the American Chemical Society,* vol. 121, pp. 10366-10372, 1999.
[126] E. A. Orabi and G. Lamoureux, "Cation−π and π–π Interactions in Aqueous Solution Studied Using Polarizable Potential Models," *Journal of Chemical Theory and Computation,* vol. 8, pp. 182-193, 2012.
[127] S. McLaughlin and A. Aderem, "The myristoyl-electrostatic switch: a modulator of reversible protein-membrane interactions," *Trends in Biochemical Sciences,* vol. 20, pp. 272-276, 1995.

I

Biochimica et Biophysica Acta 1838 (2014) 3191–3202

Contents lists available at ScienceDirect

Biochimica et Biophysica Acta

journal homepage: www.elsevier.com/locate/bbamem

Two homologous neutrophil serine proteases bind to POPC vesicles with different affinities: When aromatic amino acids matter

Anne-Sophie Schillinger [a,b], Cédric Grauffel [a,b,1], Hanif Muhammad Khan [a,b], Øyvind Halskau [a], Nathalie Reuter [a,b,*]

[a] Department of Molecular Biology, University of Bergen, Pb. 7803, N-5020 Bergen, Norway
[b] Computational Biology Unit, Department of Informatics, University of Bergen, Pb. 7803, N-5020 Bergen, Norway

ARTICLE INFO

Article history:
Received 20 February 2014
Received in revised form 29 August 2014
Accepted 3 September 2014
Available online 16 September 2014

Keywords:
Amphitropic protein
Large unilamellar vesicles
Molecular dynamics simulations
SPR: surface plasmon resonance
Proteinase 3
Neutrophil elastase

ABSTRACT

Neutrophil serine proteases Proteinase 3 (PR3) and human neutrophil elastase (HNE) are homologous antibiotic serine proteases of the polymorphonuclear neutrophils. Despite sharing a 56% sequence identity they have been shown to have different functions and localizations in the neutrophils. In particular, and in contrast to HNE, PR3 has been detected at the outer leaflet of the plasma membrane and its membrane expression is a risk factor in a number of chronic inflammatory diseases. Although a plethora of studies performed in various cell-based assays have been reported, the mechanism by which PR3, and possibly HNE bind to simple membrane models remains unclear. We used surface plasmon resonance (SPR) experiments to measure and compare the affinity of PR3 and HNE for large unilamellar vesicles composed of 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC). We also conducted 500-nanosecond long molecular dynamics simulations of each enzyme at the surface of a POPC bilayer to map the interactions between proteins and lipids and rationalize the difference in affinity observed in the SPR experiment. We find that PR3 binds strongly to POPC large unilamellar vesicles ($K_d = 9.2 \times 10^{-7}$ M) thanks to the insertion of three phenylalanines, one tryptophan and one leucine beyond the phosphate groups of the POPC lipids. HNE binds in a significantly weaker manner ($K_d > 10^{-5}$ M) making mostly electrostatic interactions via lysines and arginines and inserting only one leucine between the hydrophobic lipid tails. Our results support the early reports that PR3, unlike HNE, is able to directly and strongly anchor directly to the neutrophil membrane.

1. Introduction

Neutrophils are the most abundant type of leukocytes and are key components of the innate immune system, able to mediate both anti-infectious and pro-inflammatory effects [1,2]. Neutrophil serine proteases (NSP) Proteinase 3 (PR3, EC 3.4.21.76) and human neutrophil elastase (HNE) are homologous antibiotic serine proteases of the polymorphonuclear neutrophils (PMNs) which can be considered as important cellular targets in a number of chronic inflammatory diseases [3]. PR3 and HNE are mainly localized within the azurophilic granules of resting neutrophils and can be exposed at the cell surface when the neutrophils are activated. Despite sharing a 56% sequence identity (Fig. 1) and a high structural similarity (Cf. Fig. 2A) [4] PR3 and HNE have been shown to have different functions and localizations in the neutrophils. In particular, and in contrast to HNE, PR3 has been detected in secretory vesicles and on the outer leaflet of the plasma membrane [5,6]. Because of its peculiar localization PR3 has been suggested to play a role in the pathophysiology of various chronic inflammatory diseases involving neutrophils and especially in granulomatosis with polyangeitis, a systemic vasculitis associated with autoantibodies against PR3 [6–10].

Witko-Sarsat et al. first reported a specific association of PR3 to the plasma membrane, which they described as stronger "than only an ionic interaction" [11]. On the other hand Campbell et al. argued in favor of a weak charge-dependent mechanism similar for both

Abbreviations: PR3, proteinase 3; HNE, human neutrophil elastase; LUV, large unilamellar vesicle; POPC, 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine; SPR, surface plasmon resonance; MD, molecular dynamics

* Corresponding author at: University of Bergen, Department of Molecular Biology, Pb. 7803, N-5020 Bergen, Norway. Tel.: +47 555 84040, fax: +47 555 89683.
E-mail addresses: Anne-Sophie.Schillinger@mbi.uib.no (A.-S. Schillinger), cedric@ibms.sinica.edu.tw (C. Grauffel), Hanif.Khan@mbi.uib.no (H.M. Khan), Oyvind.Halskau@mbi.uib.no (Ø. Halskau), nathalie.reuter@mbi.uib.no (N. Reuter).
[1] Present address: For Cédric Grauffel: Institute of Biomedical Sciences, Academia Sinica, Taipei 115, Taiwan.

http://dx.doi.org/10.1016/j.bbamem.2014.09.003
0005-2736/

```
PR3      IVGGHEAQPHSRPYMASLQMRGNPGSHFCGGTLIHPSFVLTAAHCLRDIP  63
HNE      IVGGRRARPHAWPFMVSLQLRG---GHFCGATLIAPNFVMSAAHCVANVN  63
         ****:.*:**: *:*.***:**    .****.*** *.**::****: ::

PR3      QRLVNVVLGAHNVRTQEPTQQHFSVAQVFLNNYDAENKLNDILLIQLSSP 111
HNE      VRAVRVVLGAHNLSRREPTRQVFAVQRIFENGYDPVNLLNDIVILQLNGS 111
         * *.*******:  :***:* *:* ::* *.**. * ****:::**...

PR3      ANLSASVATVQLPQQDQPVPHGTQCLAMGWGRVGAHDPPAQVLQELNVTV 162
HNE      ATINANVQVAQLPAQGRRLGNGVQCLAMGWGLLGRNRGIASVLQELNVTV 162
         *.:.*.* ..*** *.: : :*.******** :* :   *.*********

PR3      VTFFCRPHNICTFVPRRKAGICFGDSGGPLICDGIIQGIDSFVIWGCATR 222
HNE      VTSLCRRSNVCTLVRGRQAGVCFGDSGSPLVCNGLIHGIASFVRGGCASG 222
         ** :**  *:**:*  *:**:******.**:*:*:*:** ***  ***:

PR3      LFPDFFTRVALYVDWIRSTLR 243
HNE      LYPDAFAPVAQFVNWIDSIIQ 243
         *:** *: ** :*:** * ::
```

Fig. 1. Sequence alignment of PR3 and HNE. The two sequences share 56% sequence identity. We use boxes to highlight amino acids of HNE aligning with the predicted PR3 IBS (blue fonts for basic residues and orange for hydrophobic amino acids). Amino acids forming the catalytic triad are labeled with green fonts.

proteases [12]. In agreement with the work of Witko-Sarsat et al., Goldman et al. showed using spectrophotometry techniques that PR3 and HNE bind with different affinities to reconstituted lipid bilayers [13]. Using bilayers with different ratios of zwitterionic (DMPC[2]) and anionic (DMPG[3]) phospholipids, they showed that while PR3 binds to DMPC vesicles with an estimated K_d of 85 μM. HNE wasn't observed to bind to pure DMPC vesicles but binds to mixed DMPC:DMPG 1:1 liposomes with a K_d of 14.5 μM. The authors also show that PR3 binds best to DMPC:DMPG 1:1 vesicles with a K_d of 4.5 μM. Moreover results from differential scanning calorimetry and hydrophobic photolabelling indicate that PR3 inserts amino acids into the hydrophobic region of the lipid vesicles, while HNE does it to a lesser extent. In 2004, Durant et al. used PR3 and HNE cDNA transfected mast cell lines and showed that PR3 was expressed at the cell surface after induced degranulation while HNE was released into the extra-cellular medium. Using molecular dynamics simulations with at first a simple membrane model [14], we have reported that PR3 inserts aromatic and aliphatic amino acids into the hydrophobic core of the bilayer models, while HNE interacts mostly via electrostatic interactions to the bilayer interface. We further used relatively short MD simulations (50 ns) of PR3 with explicit DMPC bilayers to describe the protein–lipid interactions at the atomic level of detail [15]. We reported an interface-binding site (IBS) composed of a few basic amino acids (R177, R186A,[4] R186B, K187, R222) that ensure proper orientation of PR3 towards the membrane to allow for the insertion of a hydrophobic patch (V163, F165, F166, I217, W218, L223, F224) (Fig. 2B). Mutations of four hydrophobic (F165, F166, L223, F224) or four basic amino acids (R186A, R186B, K187, R222) significantly affects the membrane expression of PR3 in a cell-based assay, thus validating the role of the predicted IBS for PR3 membrane expression [16]. Taken altogether these studies indicate that despite their high sequence similarity, PR3 and HNE interact with lipid membranes using different types of interactions. This is further supported by the amino acid substitutions in the region of the PR3 IBS (Cf Fig. 2) where in particular two of the three phenylalanines, as well as tryptophan W218, are substituted by non-aromatic residues in HNE.

The available data to date has been obtained by different teams using a wide range of approaches and models, from cell-based assays to molecular modeling, but only one study using standard biophysics methods in 1999 [13].

It is generally acknowledged that the association of peripheral proteins with lipid bilayers is fast while the dissociation is slow meaning that the dissociation rate constant is the main determinant of the binding strength. As a consequence, in simple systems, the affinity for the membrane is mostly accounted for by interactions between the protein interfacial binding site and lipids. Such interactions can be investigated by molecular dynamics (MD) simulations using the protein of interest docked on an all-atom lipid bilayer; this will provide a detailed map of the lipid–protein interactions, and for example the respective contributions of hydrophobic and polar amino acids (see Refs. [17–19]). Surface plasmon resonance has been successfully used to study interactions between lipid vesicles and proteins and can be used to obtain accurate values of equilibrium dissociation constants (K_d) [20,21]. Both SPR and MD methods are thus complementary as they provide information at different levels of resolution.

We here report a combined in vitro and in silico study investigating the affinity of PR3 and HNE for POPC bilayers. We chose POPC vesicles as the thickness of POPC bilayers is more relevant to model the plasma membrane than those of DMPC. The choice of zwitterionic PC headgroups is relevant to the main lipid constituents of the plasma membrane and allows us to focus on short-range interactions with the bilayers, and in particular on the energetic contribution of the insertion of hydrophobic amino acids. We use SPR to characterize the affinity of PR3 and HNE for large unilamellar vesicles (LUVs) and 500-nanosecond long molecular dynamics simulations of the two enzymes at the surface of a pre-equilibrated POPC bilayer to map the interactions between proteins and lipids at the atomic level of detail. The use of SPR allows us to obtain insights, albeit limited, into the kinetics of PR3 and HNE binding to POPC vesicles.

2. Material and methods

2.1. Molecular modeling

2.1.1. Molecular dynamics simulations

We performed MD simulations of both PR3 and HNE inserted in POPC lipid bilayers using the following procedure: (1) equilibration of the lipid bilayer, (2) insertion of the protein in the lipid bilayer and

[2] DMPC:dimyristoylphosphatidylcholine.

[3] DMPG:dimirystoylphosphatdylglycerol.

[4] We use the chymotrypsin numbering for both PR3 and HNE. It presents the advantage of providing a consistent numbering for all enzymes of the family but introduces letters in addition to the numbering (e.g. the two consecutive arginines labeled as 186A and 186B).

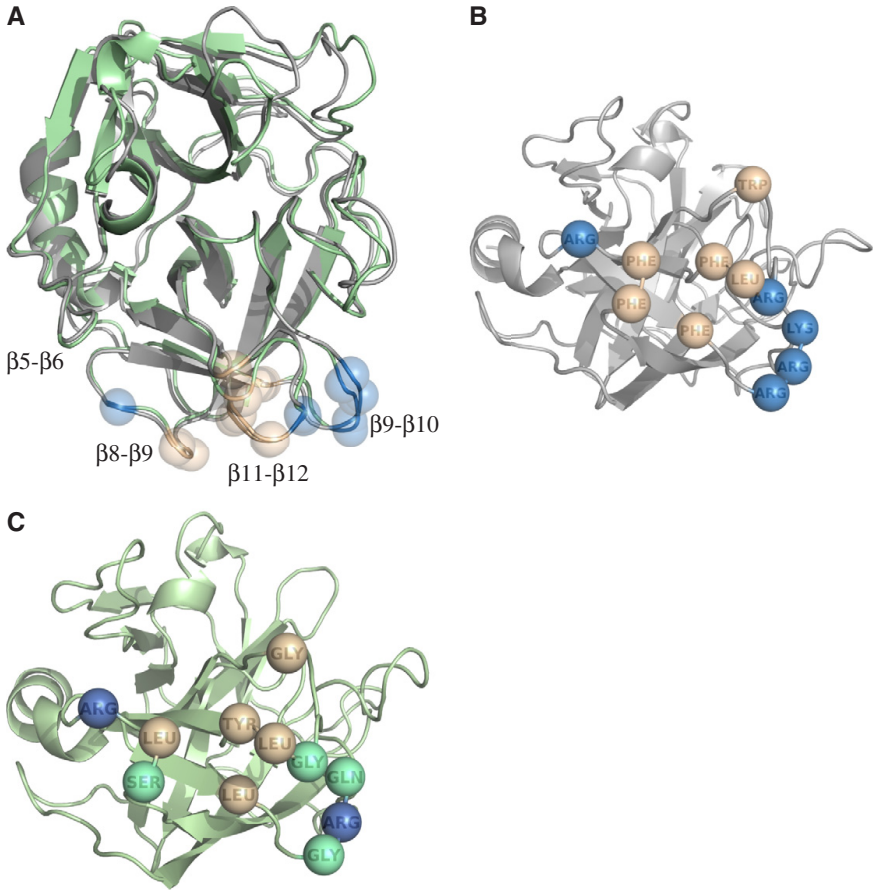


Fig. 2. (A) Structural alignment of PR3 (gray) and HNE (green). The protein secondary structure elements are represented using cartoons while each amino acid forming the PR3 IBS, as well as the amino acids of the HNE sequence aligning with the PR3 IBS (Cf. Fig. 1), are represented using balls (blue and orange for basic and hydrophobic amino acids, respectively, green for others). The nature of these amino acids in PR3 and HNE is shown on panels (B) and (C), respectively.

(3) simulation of the protein-bilayer complex and subsequent analysis of the resulting trajectories (Table 4).

In this manuscript we consequently use the chymotrypsin numbering for both PR3 and HNE. It presents the advantage of providing a consistent numbering for all enzymes of the family but introduces letters in addition to the numbering (e.g. the two consecutive arginines labeled as 186A and 186B).

2.1.1.1. POPC bilayer. A lipid bilayer made of 256 POPC was built using the CHARMM-GUI [22–24]. The lipid bilayer was subjected to energy minimization using NAMD [25] and the CHARMM36 force field update for lipids [26]. The system was then equilibrated without surface tension for 300 ps at 310 K using a time step of 2 fs and velocities reassignment every 500 fs, and subsequently run into production for 80 ns. The SHAKE algorithm was applied to constrain bonds between hydrogen and heavy atoms [27]. Non-bonded interactions were truncated using a cutoff of 12 Å, using a force-based switch function for van der Waals and a shift function for electrostatics. For estimating long-range electrostatic forces, the particle-mesh-Ewald (PME) algorithm was used [28,29]. The Langevin algorithm was used to control temperature (310 K, damping coefficient: 1.0) and pressure (target pressure: 1 atm, oscillation period: 75 fs, oscillation decay time: 25 fs) [30]. The area per lipid and the order parameters were monitored along the simulation to assess the properties of the bilayer. The order parameters S_{CD} were calculated with VMD [31] from the mean value of the angle between each C – H bond of the lipid tails and the normal to the membrane. The profiles are consistent with those in Ref. [26]. The surface area was calculated to be 65.5 ± 0.8 Å^2 on average during the simulation, close to that reported by Klauda et al. [26] (64.7 ± 0.2 Å^2) for a POPC bilayer simulated using the same CHARMM 36 force field. Kučerka et al. report an estimate of 68.3 ± 1.5 Å^2 using hybrid electron density models [32].

2.1.1.2. Insertion of Proteinase 3 and HNE at the interface of the lipid bilayer. The cartesian coordinates of PR3 were taken from chain A of the X-ray structure referenced 1FUJ [33] in the RSCB Protein Data Bank [34] and those of HNE from the 1PPF structure [35]. PR3 and HNE were then oriented with respect to, and inserted at, the interface of the equilibrated POPC lipid bilayer as described previously for PR3 [15]. Briefly, each of the enzymes was positioned at the surface of a POPC lipid bilayer in the orientation predicted by implicit bilayer simulations for HNE [14] and using earlier all-atom simulations with a DMPC bilayer for PR3

Table 1
Anchorage of PR3 in a POPC lipid bilayer: inventory of interactions and depth of anchorage.

Loop	Amino acid	Depth[a] (Å)	Hydrophobic contacts[b]	Hydrogen bonds[c] (%)	Cation–π[e] (%)
β5–β6	K99	−11.2 2.8		20.1	
β8–β9	V163	−3.0 ± 1.9	2.3		
	T164	−2.8 ± 2.0		45.6	
	F165	+ 1.3 ± 1.8	1.5		
	F166	+ 1.7 ± 1.7	2.5	**28.1**	
	R177	−4.8 ± 2.5	1.1	87.9	
β9–β10	R186A	+ 0.0 ± 2.3	4.2	85.9	
	R186B	−2.2 ± 2.7	1.5	58.7[d]/82.9	
	K187	−1.2 ± 2.4	1.6	**90.4**/74.5	
β11–β12	F215	−10.0 ± 2.2			5.9
	W218	+ 0.8 ± 2.8	1.9	**26.6**	5.7
	T221	−0.9 ± 2.4	1.6		
	R222	−0.2 ± 2.0		49.0[d]	
	L223	+ 2.5 ± 2.2	5.8		
	F224	+ 0.3 ± 1.8	1.2		
	P225	−2.0 ± 1.6	1.9		

[a] Positive values indicate that the center of mass of the amino acid is buried in the bilayer beyond the plane defined by the phosphate groups.
[b] Average number of hydrophobic contacts per frame (listed if above 1).
[c] Occupancies of hydrogen bonds with POPC phosphate groups in % (if >20; bold numbers are for hydrogen bonds involving the protein backbone).
[d] Hydrogen bond between Arg186B or Arg222 and POPC glycerols.
[e] Occupancy of cation–π adducts (if >5%).

[15]. PR3 was then translated 2 Å above its initial position to account for the difference in width between POPC and DMPC bilayers. Six lipids overlapping with the proteins were removed, in both the cases of PR3 and HNE and as reported earlier for phosphatidylinositol specific phospholipase C [36]. The starting conformation for HNE is shown on Fig. 3A, the starting conformation of PR3 is not represented but displays a depth of anchorage similar to that of HNE.

2.1.1.3. Simulations PR3-POPC and HNE-POPC. The systems were then minimized with CHARMM (v33b1) [37] using the following harmonic restraints: 150 kcal/mol/Å^2 on the protein backbone, water and ion molecules, 100 kcal/mol/Å^2 for membrane located further than 5 Å and 75 kcal/mol/Å^2 less than 5 Å from the protein and 10 kcal/mol/Å^2 for protein side chains located at a distance of 5 Å or less from the membrane. The minimization consisted of 20 cycles of 500 steps of steepest descent and 100 steps of conjugate gradients algorithms with restraints being scaled by 0.65 after each cycle. The systems were then solvated in a cubic box of TIP3 water molecules [38] using VMD (version 1.8.7) [31]. Two and eleven chloride ions were added by replacing random water molecules to neutralize the system for PR3 and HNE, respectively. The system was subsequently equilibrated using NAMD [25] and the CHARMM force field (c22 with CMAP corrections [39]) with two short runs of 400 ps in NVT ensemble, with velocities reassigned every 50 fs and 500 fs, and then further equilibrated for 2 ns. The integration of the equations of motion was done using a Multiple Time Step algorithm [40]; bonded interactions and short-range nonbonded forces were evaluated in every step and long range electrostatics every second step. The system was then run into production for 500 ns in the NPT ensemble. Pressure and temperature control, as well as the cutoff scheme and treatment of long-range electrostatic interactions are the same as for the equilibration of the bilayer.

Table 2
Anchorage of HNE in a POPC lipid bilayer: inventory of interactions and depth of anchorage.

Loop	Amino acid	Depth[a] (Å)	Hydrophobic contacts[b]	Hydrogen bonds[c] (%)	Cation-π[e] (%)
β5–β6	P96	−7.2 ± 5.8	1.0		
	V97	−3.8 ± 5.3	2.6		
β7–β8	R146	−9.4 ± 2.5		69.0	
β8–β9	T164	−5.7 ± 2.7		25.8	
	S165	−2.9 ± 2.5		36.0	
	L166	−0.6 ± 2.1	5.9	25.4	
	R177	−4.3 ± 3.1	1.1	64.8	
	R178	−4.3 ± 3.2		**23.2**/84.9	
β9–β10	R186	−6.7 ± 3.8		62.2/20.2[d]	
	G186A	−6.9 ± 3.5		**41.7**	
	F192	−11.6 ± 3.8			13.2
β11–β12	R217	−7.2 ± 4.4		49.0	
	S221	−3.2 ± 2.4		28.0	
	G222	−2.6 ± 2.1		**28.4**	
	L223	+0.4 ± 2.0	6.3		
	Y224	−1.3 ± 1.9		31.1[d]	

[a] Mean values and standard deviations. Positive values indicate that the center of mass of the amino acid is buried in the bilayer beyond the plane defined by the phosphate groups.
[b] Average number of hydrophobic contacts per frame (listed if above 1).
[c] Occupancies of hydrogen bonds in % (occupancies less than 20% are omitted; bold numbers for backbone hydrogen bonds).
[d] Hydrogen bond between R186 or Y224 and POPC glycerols.
[e] Occupancy of cation–π adducts (occupancies less than 5% are omitted).

Table 3
LUV immobilization levels and chip coverage accession by BSA binding (BSA is used at 0.1 mg/ml and is injected 60 s at 10 μl · min^{-1}). Values reported are the means and standard deviations of four experiments.

Immobilization level (RU)	BSA binding level (RU)
8669 ± 95	43 ± 2.6

2.1.1.4. Analysis. Based on the evolution of the root mean square deviation (RMSD) between the trajectory conformations and the conformation of the enzymes before the MD simulations (Cf. supplementary data, Fig. S1), we decided to use the trajectories between 200 and 500 ns as sampling windows. All analyses are thus performed on the last 300 ns of the simulations.

The occupancies of hydrogen bonds were calculated with the CHARMM program [37] using a 2.4 Å cutoff distance between hydrogen and acceptor and a 130° donor–hydrogen–acceptor angle criterion. The donor and acceptor definition are taken from the CHARMM force field [38]. Hydrophobic contacts were defined using a 3 Å cutoff distance between aliphatic groups of the lipids and of the enzymes (Charmm atom types ca; cb; cg1; cg2; cg2; ha*; hb*; hg; hg2*; type cg except for hsd, hse, asn, asp; type hg1 except for cys, thr, ser; type cd except for arg, gln, glu; type cd1; type cd2 except for hsd, hse; type ce1, ce2, cz and associated hydrogens of phe, tyr, type cd1, cd2, ce2, ce3, cz2, cz3 and associated hydrogen of trp, type cay and type hy*). Cation–π interactions between aromatic rings (phenylalanine, tyrosine and tryptophan) are considered to exist when all distances between the heavy atoms of the aromatic ring and choline nitrogen are below 7 Å and when these distances do not differ by more than 1.5 Å [41,42].

We evaluated the depth of anchorage of the proteins as described in Grauffel et al. [36]. Briefly we used the mean z coordinate of the phosphorus atoms as a reference plane. The center of mass of each residue was calculated and its difference to the reference plane was calculated. The corman module of the Charmm program was used for coordinate statistics. Values reported are means of the distances of the last 300 ns of simulations. Averaged electron density profiles (EDP) were calculated using the VMD Density Profile plugin [43], at 0.5 Å resolution using a structure per nanosecond (on sampling window). Residue profiles were normalized with respect to the peak of the phosphate profile.

Table 4
K_d values for PR3 and HNE with POPC LUVs. Values reported are the means and standard deviations of six experiments for PR3 and four for HNE.

K_d ($\times 10^{-7}$ M)	
PR3	HNE
9.22 ± 0.4	>100

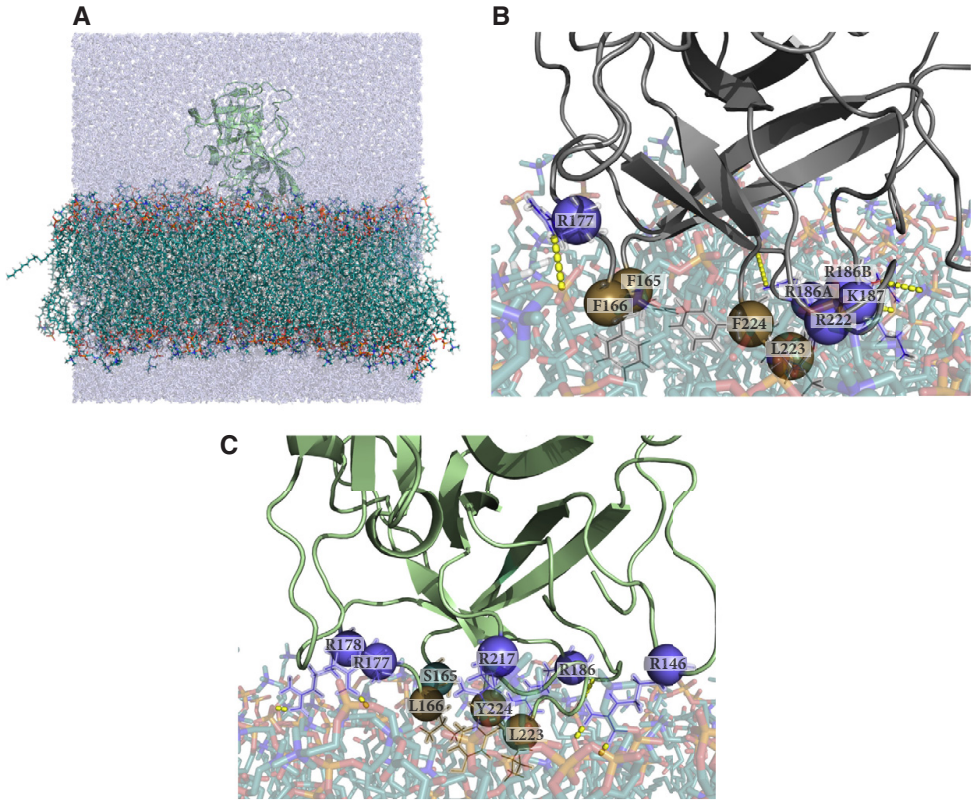


Fig. 3. Molecular dynamics simulations of PR3 and HNE at the surface of a POPC bilayer. (A) Simulated system with HNE represented using green cartoons, the POPC bilayer using sticks colored by atom types and the water molecules in light blue. Randomly picked snapshots of (B) PR3 and (C) HNE interface binding sites (at 286 and 343 ns, respectively). Amino acids mediating either hydrogen bonds of at least 45% occupancy for hydrogen bonds or at least one hydrophobic contact on average, are highlighted with balls. Yellow dashed lines represent hydrogen bonds.

2.1.2. Energetic contributions to membrane binding using IMM1

We used the IMM1 membrane model [44], as implemented in Charmm [37] (v38) to evaluate the binding energy of PR3 and HNE to a zwitterionic membrane as a function of their distance to the membrane. IMM1 is based on the EEF1 model for water-soluble proteins, which uses a linear distance-dependent dielectric constant, neutralizes the ionic side chains, and adds a Gaussian solvent exclusion term to the CHARMM 19 energy function [45]. In IMM1 the membrane is represented as a hydrophobic slab. The solvation parameters change smoothly from aqueous values outside the membrane to values corresponding to a nonpolar solvent inside the membrane. Briefly, the effective energy of the protein is given by:

$$W_{IMM1} = E_{intra} + \Delta G_{solv} \quad (1)$$

where E_{intra} is the intramolecular energy of the protein given by the CHARMM 19 and ΔG_{solv} its solvation free energy calculated as described in Ref. [44].

We calculated W_{IMM1} of the protein anchored at the membrane [14] and then every angstrom along an axis parallel to the membrane normal, up to a distance of 15 Å above the membrane plane; the distance is defined by the shortest distance between protein and membrane surface. The starting geometries are chosen as the bound structures from earlier IMM1 simulations [14] and are kept rigid during the procedure. The binding energy of PR3 (and HNE) to membranes could thus be calculated as the difference between the effective energy with the membrane model IMM1 and the effective energy in water (EEF1.1), which was also calculated:

$$\Delta W_{IMM1} = \Delta E_{intra} + \Delta\Delta G_{solv} \quad (2)$$

The total binding energy ΔW_{IMM1} of each of the proteins can be decomposed into terms that represent the contribution of each type of atom to the change in solvation energy plus the electrostatic term:

$$\Delta W_{IMM1} = \Delta E_{elec} + \Delta W_{hydrophobic} + \Delta W_{polar} \quad (3)$$

ΔE_{elec} is the change in the intra-molecular coulombic term when transferring the protein from water to the membrane. Since the same structure is used to calculate W_{IMM1} (water) and W_{IMM1} (membrane) the bonded terms and the van der Waals contribution cancel out. The next two terms represent the contribution to $\Delta\Delta G_{solv}$ (Eq. (3)); $\Delta W_{hydrophobic}$ is the contribution from the aromatic groups (atoms of type CR* in Charmm19) and the aliphatic groups (CH* atom types in Charmm 19) and ΔW_{polar} is the contribution from the polar groups.

2.2. Sample preparation

2.2.1. Proteins

PR3 and HNE were purchased from Athens Research & Technology and fatty acid free bovine serum albumine (BSA) from Sigma. According to the manufacturer, the purity of PR3 and HNE is higher than 95%. We assessed it using SDS-page and also have assessed the enzyme activity of the enzymes in kinetic assays.

2.2.2. Liposomes

The lipids (POPC) were purchased from Avanti Polar Lipids. Liposomes were prepared as reported in [46]. Lipids solvated in chloroform were added in glass tubes in the prerequisite amount. Lipids were handled and kept out of light and reactive atmosphere as much as possible by operation in hoods, flushing reagent bottles with dry N_2, and using glass containers wrapped in aluminum foil. The chloroform solutions were dried under dry N_2 pressure. Traces of chloroform were removed by subjecting the samples to vacuum for at least 2 h. Lipid cakes were rehydrated with HBS-N buffer (HBS-N: 0.100 mM HEPES, 150 mM NaCl, pH 7.4) and vortexed vigorously until all films were suspended as slurry. For liposome-preparation, solutions were subjected to seven freeze–thaw cycles using liquid N_2 and a water bath. The hydrated multilamellar structures were then extruded at room temperature and well above the lipid T_m using a Mini-Extruder (Avanti Polar Lipids) assembled using two Millipore filters of 100 nm pore size. Samples were forced through the filters 10 times using Hamilton syringes and the resulting solutions were transferred to clean, foil wrapped glass tubes and stored at 4 °C. Final liposome composition was 100% POPC and the total lipid concentration was 2.5 mM.

2.3. Surface plasmon resonance

The SPR analyses were carried out on a BIAcore T200 (BIAcore, GE Healthcare) and Biacore T200 Control Software. All experiments were carried at 25 °C. Protein and lipid interactions were monitored using a L1 sensor chip. A preparation procedure was performed before each experiment. The surface of the L1 sensor chip was first cleaned with a 1 min injection of 40 mM octylglucoside at a flow rate of 10 μL/min. Liposome solutions were diluted to 1 mM concentration with running buffer and injected at a flow rate of 1 μL/min for 10 min until maximum binding was reached. Liposome maximum deposition was about 8500 response units (RU) for POPC. The surface of the L1 chip was then washed with a solution of 10 mM NaOH for 1 min at a flow rate of 10 μL/min. The completeness of the chip coverage was assessed by injection of bovine serum albumin (BSA) at 0.1 mg/mL and at a rate of 10 μL/min for 60 s. Generally this injection did not perturb the lipid-covered chip by more than 43 RU, and it rapidly fell back to its original value when injection of BSA stopped. Binding assays were then performed on the validated chips. The two proteins (PR3 and HNE) were diluted to sets of at least 5 different concentrations ranging from 0.125 μM to 3 μM (0.125, 0.5, 1, 2, 3) for PR3 (two additional concentrations for HNE, 6 and 9 μM), and were injected over the immobilized liposomes at a flow rate of 5 μL/min for 120 s and 180 s (for HNE and PR3 respectively) until equilibrium was reached. The dissociation phase was measured for at least 420 s after the addition of the sample. At the end of the binding assay, the surface of the sensor chip was regenerated with a solution of octylglucoside 40 mM for 30 s at a flow rate of 30 μL/min. No reference channel was used due to non-specific binding of PR3 on the chip [47]. Instead we focused on achieving maximal coverage of the chip with liposomes and in this way ensure that the resulting SPR signal was completely dominated by the protein interacting with the lipid membrane [48]. The SPR data were analyzed with the Biacore T200 Evaluation Software. Binding affinities were calculated using the steady state affinity model (Langmuir model) and maximal resonance unit (RU) was plotted against concentration. Additional experiments were performed at increasing salt concentrations to evaluate its effect on the binding of PR3 and HNE. We measured the binding responses of PR3 and HNE (0.5 μM) with [NaCl] = 150 mM, [NaCl] = 300 mM and [NaCl] = 700 mM.

3. Results

3.1. Molecular modeling

3.1.1. Molecular dynamics simulation with an explicit bilayer model

PR3 and HNE were positioned at the interfacial region of POPC lipid bilayers as described in the Material and methods section and illustrated on Fig. 3A. Each system was simulated for 500 ns and analyzed in order to characterize the interactions between the enzyme interfacial binding sites and the lipids. We report in Tables 1 and 2 the occupancy of significant hydrogen bonds along the sampling window (occupancy above 20%), as well as the average number of hydrophobic contacts for the amino acids that achieve on average more than one contact per frame of the trajectory. On Fig. 3 we show a snapshot of the simulations of PR3 (Fig. 3B) and HNE (Fig. 3C).

3.1.1.1. PR3. The simulation indicates that the structure of PR3 is not affected by the presence of the membrane; the average RMSD between the conformations in the trajectory is 1.54 ± 0.19 Å on the sampling window (Cf Fig. S1, Supplementary Data). PR3 remains at a stable depth of anchorage at the bilayer interface; we calculate a distance of 19.7 ± 1.6 Å between the center of mass of PR3 and the average plane of the phosphorus atoms (Cf. Supplementary Data, Fig. S2). Interactions between PR3 and the POPC bilayer are mediated almost exclusively by amino acids located on three different loops: β8–β9 (amino acids 163 to 180), β9–β10 (184–197), β11–β12 (215–225). The positions of PR3 amino acids with respect to the average plane of the phosphorus atoms gives an indication of their depth of anchorage in the lipid bilayer. Two loops are anchored significantly beyond the phosphorus atoms; loop β11–β12 appears to be the one that has the deepest anchorage. The β8–β9 loop with F165 and F166 is also anchored beyond the plane of the phosphorus atoms. Yet the average depths of anchorage measured are not characteristic of deep anchoring, and suggest a positioning of the loops at the interface. The β9–β10 loop carries most of the basic cluster identified in our early implicit membrane simulations [14] and is positioned slightly above the two other ones.

Most of the hydrogen bonds we observe involve the phosphate groups of POPC lipids. A low number of hydrogen bonds with occupancies below 20% involve glycerol groups, only Arg186B[5] (58.7%) and Arg222 (49.0%) have occupancies of hydrogen bonds with glycerol above 20% indicating that they are buried somewhat deeper in the interface than the other basic amino acids. The strongest hydrogen bonds involve basic amino acids (R177, R186A, R186B and K187) and have occupancies above 80%. Remarkably Lys187 is involved in hydrogen bonds through its side chain (74.5%) and backbone (90.4%). We have earlier predicted, using simulations with an implicit membrane model and mutagenesis experiments [16], that R186A, R186B, K187 and R222 play a major role in PR3 interaction with cell membranes. In particular mutating these four amino acids into four alanines would abrogate PR3 membrane expression in Rat Basophil Leukemia (RBL) cells. Our results confirm the importance of this cluster of basic amino acids constituted of four arginines (R177, R186A, R186B, R222) and one lysine (K187). Besides these, lysine 99 (K99) reported to be important for ligand binding [49,50] mediates hydrogen bonds with the lipids. It is also the case of F166 and W218 although they mediate interactions via their backbone atoms while their side chains are heavily involved in hydrophobic contacts with the lipid tails.

[5] PR3 and HNE amino acids are numbered according to the chymotrypsin convention, which is common for all serine proteases of the family. To account for insertions this convention includes letters and numbers (Cf Materials and methods section).

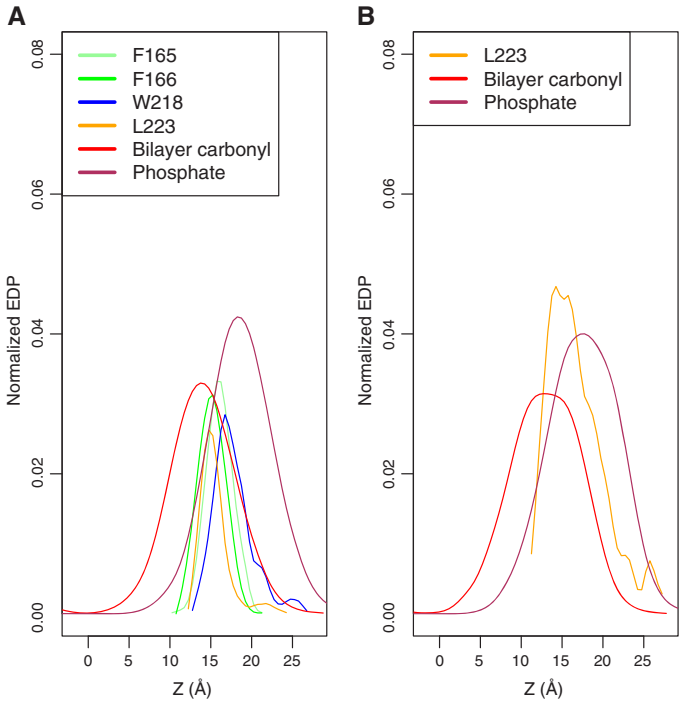


Fig. 4. PR3 and HNE hydrophobic amino acid anchoring. Normalized electron density profiles (EDP) for selected amino acids (F165, F166, W218, and L223 for PR3, L223 for HNE) calculated from the MD simulations of (A) PR3 and (B) HNE. EDPs of the carbonyl (C atom) and phosphate groups (phosphorus atoms) are also plotted. EDPs are averaged over the last 300 ns of the simulations during which the penetration of the above-mentioned amino acids remains unchanged.

We calculated the average number of hydrophobic contacts per frame along the sampling window (Cf. Table 1). Several amino acids of the predicted interfacial binding site display hydrophobic contacts with the POPC lipid bilayer. As expected these are aromatic (F165, F166, W218, F224) and hydrophobic amino acids (V163, T221, L223 and P225). Among these, V163, F166, L223 have particularly high average number of contacts (2.3, 2.5 and 5.8, respectively). Yet their anchorage is not very deep within the hydrophobic tails but rather at the lowest level of the interfacial region. Indeed the density profiles (Fig. 4A) show expected values for F165, F166 and L223 at 15.7, 15.2 and 15.2 Å, respectively. These are between the expected values for the phosphates ($z = 18.2$ Å) and the carbonyl groups ($z = 13.7$ Å). The positioning of the phenylalanine is comparable to previously reported anchorage of phenylalanines of Osh4 [18]. Simultaneous mutations of the four amino acids F165, F166, L223 and F224 did impair membrane expression of PR3 on RBL cells [16]. Interestingly the basic cluster involved in strong hydrogen bonding (R177, R186A, R186B, K187) is also involved in hydrophobic contacts with the lipid tails. In agreement with its involvement in hydrogen bonds with POPC glycerol groups, R186A is the basic amino acid with the higher number of hydrophobic contacts. The aromatic residues F165, F166, W218 and F224 are actually embedded in the bilayer.

While we observed strong cation–pi interactions between W218 and DMPC lipids in our previous work, the occupancy of this interaction is of only 5.7% in the present simulation, which we do not consider as being significant.

3.1.1.2. HNE. Using an implicit membrane model, we previously predicted that HNE would bind to cell membranes using the same interfacial binding site as PR3 [14]. We therefore inserted HNE in the POPC bilayer similarly to Proteinase 3; using the same orientation and the same depth of anchoring. With the implicit membrane model, we also observed a higher electrostatic contribution than in the case of PR3 and fewer contributions from hydrophobic amino acids.

The structure of HNE is unaffected by the POPC bilayer (RMSD 1.26 ± 0.17 Å) and as PR3, it remains stably anchored at the bilayer interface (Cf. Supp Mat, Fig. S2) although a visual inspection of the trajectories indicate that the orientation of HNE with respect to the membrane plane varies more than that of PR3. Most of the interactions with POPC lipids are achieved by amino acids carried by the same three loops as in PR3 (β8–β9, β9–β10, β11–β12) (Cf Table 2) plus an additional interaction through R146 (loop β7–β8). In fact basic residues located on the loops β7–β8, β8–β9 and β9–β10 seem to alternate as anchors with the protein tilting around an axis perpendicular to the bilayer along the simulation. This is well illustrated by the variation along time of the depth of anchoring of amino acids R146 (β7–β8), R177, R178 (β8–β9) and R186 of loop β9–β10 (Cf Fig. S3 in Supplementary data). The proline and valine numbered 96 and 97, respectively (P96, V97, on loop β5–β6), are involved in interactions with the lipid tails as illustrated by their number of hydrophobic contacts (1.0 and 2.6, respectively. Cf Table 2). V97 is only two amino acids away from L99, which interestingly is not observed to interact with the lipids. Its equivalent in PR3 is a lysine (K99) and is observed to interact with the lipid heads via hydrogen bonds.

Hydrogen bonds between HNE and the lipids are mediated by more amino acids than in the case of PR3 (12 against 8) but only one out of the twelve, R178, has an occupancy above 80%. R177 and R178 are the arginines that on average are the most deeply inserted into the interface during the simulation (Cf Table 2 and Fig. S3 in supplementary data). Of the basic amino acids involved in hydrogen bonds with lipids, only

R177 is also involved in the PR3 IBS (Cf. Fig. 1 and Table 1). Yet HNE achieves a high number of hydrogen bonds with lipids when positioned on the bilayer surface, using other arginines not conserved in PR3 (R146, R178, R186, R217).

Only five amino acids mediate an average number of hydrophobic contacts above or equal to 1.0 (V97, P96, L166, R177 and L223). This is strikingly less than in PR3 for which 12 amino acids had a higher number of hydrophobic contacts than this threshold. The two leucines mediate the highest number of contacts and L223 is the only residue anchored beyond the phosphate plane (Cf Table 2). The electron density profile of L223 shows an anchorage comparable to that observed for PR3 (CF. Fig. 4B), between the carbonyl and phosphate groups, with an expected anchorage value of $z = 14.2$ Å. Carbonyl and phosphate groups have expected values of 13.2 Å and 17.7 Å, respectively. Amino acid K99 of PR3 forms hydrogen bonds with POPC lipids (Cf previous paragraph), while it is not the case of its equivalent in HNE (L99) which cannot form hydrogen bonds; instead two other amino acids of the same loop (β5–β6, P96 and V97) mediate hydrophobic contacts with the lipids.

All together the simulation results indicate that HNE interacts with the bilayer using mostly hydrogen bonds and very few hydrophobic anchors, suggesting a looser binding to lipid membranes than PR3.

3.1.2. Energetic contribution of hydrophobic amino acids to membrane binding

Using an implicit membrane model, we could estimate the contribution of hydrophobic amino acids to the binding energy of PR3 and HNE. These contributions, as well as the total binding energy are plotted on Fig. 5 for distances between the membrane and the proteins ranging from the protein being anchored to a distance of 15 Å between the protein and the membrane surface. Both PR3 (Fig. 5A) and HNE (Fig. 5B) show comparable unfavorable contributions (ca. 5 kcal/mol) from the sum of their polar amino acids and the intra-molecular electrostatics term (Cf Material and methods section). The contribution from aliphatic and aromatic residues, on the other hand, is significantly larger for PR3 (−9 kcal/mol) than for HNE (at most −4 kcal/mol) when the proteins are anchored. This results in an overall favorable binding of PR3 with an optimum energy of ca. −4.5 kcal/mol and a slightly unfavorable binding energy for HNE. The resolution of the model we used and the fact that we used a rigid protein do not allow us to obtain accurate evaluation of the total binding energy and in particular of the short-range interactions that are dependent on atomic level of description. The HNE binding energy could still be slightly favorable albeit far from being as favorable as the one from PR3. Overall these values indicate that the affinity of PR3 for zwitterionic membranes is strongly dominated by its hydrophobic amino acids.

3.2. Surface plasmon resonance

To experimentally verify the hypotheses resulting from the MD simulations, we conducted SPR assays to compare the affinity of PR3 and HNE for large unilamellar vesicles (LUVs) constituted of POPC lipids (Table 3).

3.2.1. Liposome immobilization

Liposomes were immobilized on the surface of the L1 sensor chip at a low flow rate ($1\ \mu L \cdot min^{-1}$) until the maximal amount of deposition was reached. Liposome immobilization levels were monitored over time and the mean immobilization level for POPC LUVs was 8669 ± 95 RU calculated on four different experiments (Cf. Table 1). To avoid non-specific binding of proteins to the surface of the L1 chips, special care was taken to cover the chip surface at the highest possible levels of liposomes. The level of the coverage sensor chip was assessed with BSA injections ($0.1\ mg \cdot mL^{-1}$). Resulting signals from BSA of around

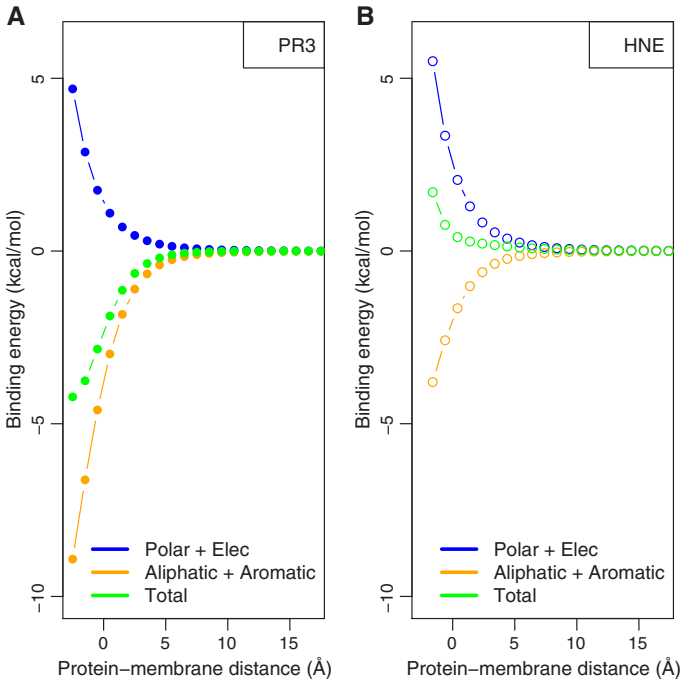


Fig. 5. Hydrophobic and electrostatic contributions to the binding energy of PR3 (A) and HNE (B). The contributions from polar amino acids and intra-molecular electrostatics are plotted in blue, the contribution from aromatic and aliphatic amino acids in orange and the total energy in green.

100 RU or less indicate a sufficient coverage [51]. In our case, BSA binding amounts to 43 ± 2 RUs and allowed us to pursue experiments further with POPC.

3.2.2. Binding of PR3 to POPC LUVs

We investigated the interaction of PR3 with neutral liposomes made of POPC using SPR. Liposomes were immobilized on the surface of the L1 sensor chip as described above. Binding assays were performed by injecting protein samples at increasing concentrations and affinity calculations were carried out by steady state analysis. We monitored the association phase for 180 s and the dissociation phase for 420 s. The sensorgrams (Fig. 6) show that the protein response is concentration dependent and is reaching equilibrium towards the end of each injection. The calculated K_d between PR3 and POPC is 9.22×10^{-7} M. During the dissociation phase, we also observed that the response signal of PR3 does not return to zero and thus demonstrates a persistent binding of PR3 to the liposomes.

3.2.3. Binding of HNE to POPC LUVs

The binding of HNE towards POPC was monitored using the same procedure as for PR3, but using a higher maximum concentration (9 μM). The association of the protein to the LUVs was monitored for 120 s (shorter than for PR3) and the dissociation for 420 s. The sensorgrams are presented on Fig. 6B and show that HNE can bind to liposomes made of POPC in a concentration-dependent manner which indicates a direct binding of the protein to the liposomes. During the dissociation phase, the signal drops immediately and returns to the baseline value. This is in contrast to the behavior of PR3. The kinetics of the protein–membrane interaction seems to be different for the two proteins. For the K_d calculation, the data collected for HNE clearly show that equilibrium was not reached even at an enzyme concentration of 9 μM. It was therefore not possible to calculate the affinity accurately but we can evaluate a lower limit for the K_d value of 1×10^{-5} M.

3.2.4. Effect of salt on HNE and PR3 binding to POPC LUVs

Comparison of binding responses of PR3 and HNE at different salt concentrations (Fig. 7) shows that, at 300 mM NaCl concentration in the running buffer, the PR3 response is less altered than that of HNE as it maintains a rather high binding response. HNE response, however, is considerably lowered compared to the one at [NaCl] = 150 mM. In addition, HNE seems to partially dissociate before the end of the 120 s of injection, indicating a rather unstable binding and difficulties maintain the protein on the lipid vesicles. At 700 mM of NaCl, about a third of the PR3 response is maintained, whereas the binding of HNE is practically abolished.

4. Discussion and conclusion

The SPR experiments yield a K_d in the low micromolar range ($9.22 \pm 0.4 \times 10^{-7}$ M) for PR3, while the binding of HNE is weaker and its K_d is not within the range of concentrations tested; we estimate it to be above 10^{-5} M. For comparison, K_d measured for hIIa-PLA_2 with the same method was 6.8 10^{-8} M [20] and for the lactadehrin C2 domain 3.2 10^{-7} M [21]. The difference we observe between PR3 and HNE is consistent with the results of Goldman et al. although we measure a K_d for PR3 with POPC that is lower by two orders of magnitude than the value they determined in 1999 (85×10^{-6} M) using spectrophotometric measurements and DMPC vesicles. This indicates that PR3 binds best to bilayers of unsaturated lipids suggesting that it is sensitive to lipid packing. The study from Goldman et al. appears to have been conducted using multilamellar vesicles, whereas ours were extruded to produce monodisperse unilamellar vesicles.

Although the use of a low flow rate for the injection of protein on the SPR chip prevents us from determining accurate rate constants, the SPR sensorgrams show significantly different dissociation rates for PR3 and HNE, indicating that both enzymes are bound differently to the lipid bilayer with PR3 binding being perpetuated after the flow is interrupted, while HNE transfers back to the bulk almost immediately. Generally, long-range non-specific interactions (typically electrostatics) accelerate the association of peripheral membrane proteins and short range interactions (typically van der Waals) slow the dissociation [52]. While our SPR sensorgrams show no visible differences in the association phase, they show a slower dissociation for PR3, which would therefore imply that short-range interactions are fewer or weaker in HNE. The

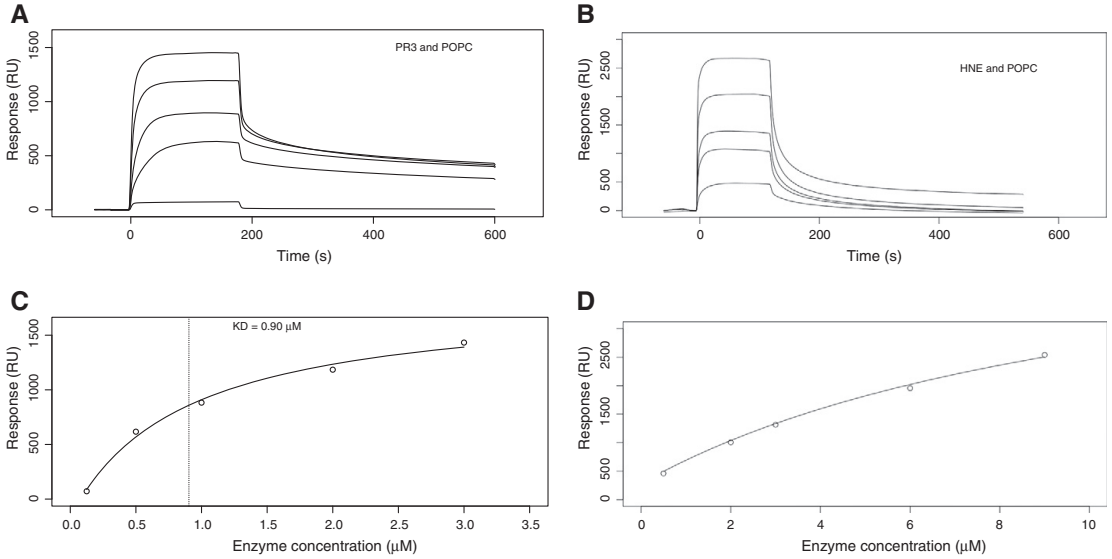


Fig. 6. Binding assay of PR3 and HNE to LUVs from surface plasmon resonance. PR3 (A) and HNE (B) binding responses, and respective affinity data below (C. PR3 and D. HNE) over immobilized POPC. All data are blank subtracted. No double referencing has been done due to high non-specific binding to the reference channel (L1 chip with no liposomes—data not shown).

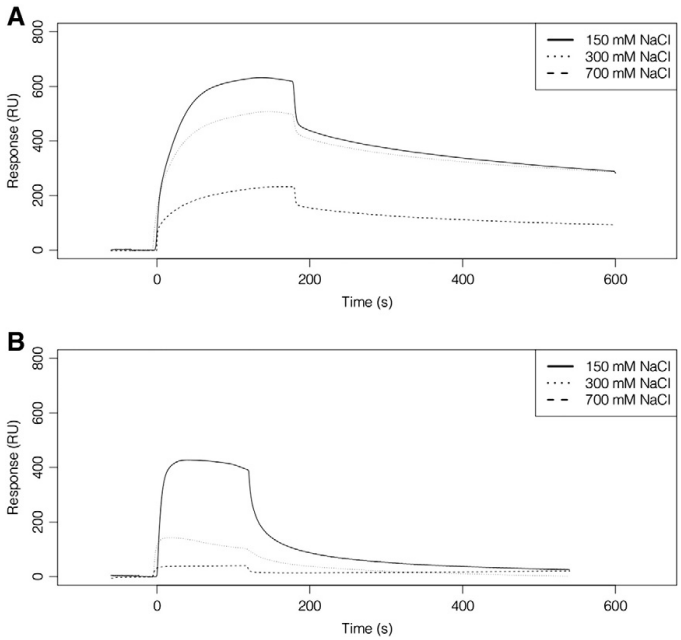


Fig. 7. Binding assay of PR3 (A) and HNE (B) to LUVs from surface plasmon resonance at salt concentrations ranging from 150 mM to 700 mM.

difference between PR3 and HNE in terms of short-range van der Waals interactions is clearly characterized by our MD simulations. The IBS for HNE and PR3 are different (Tables 1 and 2, Fig. 3) leading to a higher average number of hydrophobic contacts per simulation frame between the lipids and PR3 (27.1) than between the lipids and HNE (16.9). On the other hand, and although both HNE and PR3 have about the same number of basic amino acids at their IBS, HNE achieves a higher number of hydrogen bonds with the lipid phosphates. Particularly relevant to the difference in dissociation rates is the fact that PR3 inserts a higher number of aromatic amino acids (F165, F166, W218, F224 in PR3 vs. F192 in HNE) below the plane of the phosphorus atoms. This may also explain the higher shift in the membrane transition temperature observed for PR3 vs HNE by differential scanning calorimetry in the study from Goldman et al. According to the Wimley–White interfacial hydrophobicity scale, based on the transfer free energy of pentapeptides (AcWL–X–LL) from water to a POPC bilayer [53,54], aromatic residues have the most favorable partitioning energies, while charged amino acids have large unfavourable energies. Other amino acids make relatively small contributions. The difference in the number of aromatic residues inserted by both enzymes in the bilayer thus explains the difference in membrane affinity for PR3 and HNE. Anchorage of hydrophobic aromatic residues for anchoring into lipid bilayers has been reported for other amphitropic proteins or peptides [18,55].

The SPR sensorgrams and the difference in the number of hydrophobic and aromatic amino acids anchored in the hydrophobic region of the bilayer between both enzymes is thus consistent with a specific association of PR3 to the plasma membrane, "stronger than only an ionic interaction" [11] with insertion of hydrophobic amino acids [14] while HNE has a more shallow interaction based on hydrogen bonds and short-range electrostatic interactions with LUVs. A qualitative interpretation of SPR measurements using different concentrations of salt (NaCl) shows key differences in the way the binding responses of PR3 and HNE are affected. Whereas PR3 displays a slowly decreasing binding at increasing concentrations of salt, HNE is rapidly and dramatically affected to become nearly abrogated at high salt concentration.

As PR3 is known to bind to the exoplasmic side of the neutrophils plasma membrane, the choice of POPC for the lipid vesicles was motivated by their thickness, closer to plasma membranes than DMPC lipids. Moreover the PC headgroups are relevant with lipid components of the plasma membrane [56–58]. Conveniently the use of vesicles constituted of zwitterionic lipids only allows us to distinguish between the effect of long-range electrostatics and short-range amino acid–lipid interactions. Our results show that PR3 is able to bind to pure PC vesicles, indicating that it could bind to PC domains at the plasma membrane. We expect that adding a fraction of anionic lipids in the bilayer would improve the electrostatic interaction between the lipids and the enzyme in a roughly comparable manner, with a slight advantage for HNE. Yet, and because of its ability to insert 5 hydrophobic amino acids into the lipid bilayer against one for HNE, we would expect PR3 to still be a stronger membrane binder than HNE on membranes containing a fraction of anionic lipids. Studies using vesicles with some anionic lipid content will need to be used to investigate this aspect.

A number of proteins have been shown to be co-localized or co-immunoprecipitate with membrane-expressed PR3 and have been proposed as partners of PR3 at the neutrophil membrane (reviewed in refs [9,59]): CD177 (NB1) [59,60], Fcgamma receptor FcγRIIIb and $p22^{phox}$ subunit of cytochrome b558 [61], β2 integrin adhesion molecule CD11b/CD18 [62], Protease Activated Receptor 2 (PAR2) [63,64], CD177 [65], Phospholipid Scramblase 1 [66] and calreticulin [67]. Yet, to our knowledge, there exist little evidence of a physical interaction between membrane-bound PR3 and the identified partners, except for calreticulin for which a nanomolar affinity has has been reported using SPR experiments. Direct interaction of PR3 with the membrane phospholipids and interaction with protein partners are not mutually exclusive if we consider the formation of a protein complex. Interactions of membrane-bound PR3 with proteins, possibly transmembrane receptors, are indeed necessary for the function of PR3 and would stabilize its interactions with the membrane.

To summarize, using long molecular dynamics simulations of PR3 at the surface of POPC bilayers and SPR experiments following the binding of PR3 to POPC LUVS, we have demonstrated that PR3 can bind directly to POPC lipid bilayers by inserting one aliphatic and four aromatic amino acids into the hydrophobic core of the bilayer. Our results thus indicate that PR3 is able to mediate direct interactions with the exoplasmic leaflet of the neutrophil membrane, which is mostly constituted of PC lipids. HNE interacts with the same LUVs in a shallower manner dominated by short-range electrostatic interactions. The difference in affinity between the two proteins can be explained by the difference in the nature of their IBS, namely the number of hydrophobic aromatic amino acids present.

Acknowledgements

This work was supported by grants from the Bergen Research Foundation and the Norwegian Research Council (FRIMEDBIO #214167). Parallab (High Performance Computing Laboratory at the University of Bergen) and NOTUR (Norwegian metacenter for computational science) are thankfully acknowledged for the provision of CPU time.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.bbamem.2014.09.003.

References

[1] C. Nathan, Neutrophils and immunity: challenges and opportunities, Nat. Rev. Immunol. 6 (2006) 173–182.
[2] B. Amulic, C. Cazalet, G.L. Hayes, K.D. Metzler, A. Zychlinsky, Neutrophil function: from mechanisms to disease, Annu. Rev. Immunol. 30 (2012) 459–489.
[3] B. Korkmaz, M.S. Horwitz, D.E. Jenne, F. Gauthier, Neutrophil elastase, proteinase 3, and cathepsin G as therapeutic targets in human diseases, Pharmacol. Rev. 62 (2010) 726–759.
[4] E. Hajjar, T. Broemstrup, C. Kantari, V. Witko-Sarsat, N. Reuter, Structures of human proteinase 3 and neutrophil elastase—so similar yet so different, FEBS J. 277 (2010) 2238–2254.
[5] E. Csernok, M. Ernst, W. Schmitt, D.F. Bainton, W.L. Gross, Activated neutrophils express proteinase 3 on their plasma membrane in vitro and in vivo, Clin. Exp. Immunol. 95 (1994) 244–250.
[6] V. Witko-Sarsat, P. Lesavre, S. Lopez, G. Bessou, C. Hieblot, B. Prum, L.H. Noel, L. Guillevin, P. Ravaud, I. Sermet-Gaudelus, J. Timsit, J.P. Grunfeld, L. Halbwachs-Mecarelli, A large subset of neutrophils expressing membrane proteinase 3 is a risk factor for vasculitis and rheumatoid arthritis, J. Am. Soc. Nephrol. 10 (1999) 1224–1233.
[7] A.A. Rarok, C.A. Stegeman, P.C. Limburg, C.G. Kallenberg, Neutrophil membrane expression of proteinase 3 (PR3) is related to relapse in PR3-ANCA-associated vasculitis, J. Am. Soc. Nephrol. 13 (2002) 2232–2238.
[8] A. Schreiber, A. Busjahn, F.C. Luft, R. Kettritz, Membrane expression of proteinase 3 is genetically determined, J. Am. Soc. Nephrol. 14 (2003) 68–75.
[9] V. Witko-Sarsat, N. Reuter, L. Mouthon, Interaction of proteinase 3 with its associated partners: implications in the pathogenesis of Wegener's granulomatosis, Curr. Opin. Rheumatol. 22 (2010) 1–7.
[10] A. Schreiber, F.C. Luft, R. Kettritz, Membrane proteinase 3 expression and ANCA-induced neutrophil activation, Kidney Int. 65 (2004) 2172–2183.
[11] V. Witko-Sarsat, E.M. Cramer, C. Hieblot, J. Guichard, P. Nusbaum, S. Lopez, P. Lesavre, L. Halbwachs-Mecarelli, Presence of proteinase 3 in secretory vesicles: evidence of a novel, highly mobilizable intracellular pool distinct from azurophil granules, Blood 94 (1999) 2487–2496.
[12] E.J. Campbell, M.A. Campbell, C.A. Owen, Bioactive proteinase 3 on the cell surface of human neutrophils: quantification, catalytic activity, and susceptibility to inhibition, J. Immunol. 165 (2000) 3366–3374.
[13] W.H. Goldmann, J.L. Niles, M.A. Arnaout, Interaction of purified human proteinase 3 (PR3) with reconstituted lipid bilayers, Eur. J. Biochem. 261 (1999) 155–162.
[14] E. Hajjar, M. Mihajlovic, V. Witko-Sarsat, T. Lazaridis, N. Reuter, Computational prediction of the binding site of proteinase 3 to the plasma membrane, Proteins 71 (2008) 1655–1669.
[15] T. Broemstrup, N. Reuter, How does proteinase 3 interact with lipid bilayers? Phys. Chem. Chem. Phys. 12 (2010) 7487–7496.
[16] C. Kantari, A. Millet, J. Gabillet, E. Hajjar, T. Broemstrup, P. Pluta, N. Reuter, V. Witko-Sarsat, Molecular analysis of the membrane insertion domain of proteinase 3, the Wegener's autoantigen, in RBL cells: implication for its pathogenic activity, J. Leukoc. Biol. 90 (2011) 941–950.
[17] S. Jaud, D.J. Tobias, J.J. Falke, S.H. White, Self-induced docking site of a deeply embedded peripheral membrane protein, Biophys. J. 92 (2007) 517–524.
[18] B. Rogaski, J.B. Klauda, Membrane-binding mechanism of a peripheral membrane protein through microsecond molecular dynamics simulations, J. Mol. Biol. 423 (2012) 847–861.
[19] C.N. Lumb, J. He, Y. Xue, P.J. Stansfeld, R.V. Stahelin, T.G. Kutateladze, M.S. Sansom, Biophysical and computational studies of membrane penetration by the GRP1 pleckstrin homology domain, Structure 19 (2011) 1338–1346.
[20] R. Stahelin, W. Cho, Differential roles of ionic, aliphatic, and aromatic residues in membrane–protein interactions: a surface plasmon resonance study on phospholipases A2, Biochemistry 40 (2001) 4672–4678.
[21] R.V. Stahelin, Surface plasmon resonance: a useful technique for cell biologists to characterize biomolecular interactions, Mol. Biol. Cell 24 (2013) 883–886.
[22] S. Jo, T. Kim, W. Im, Automated builder and database of protein/membrane complexes for molecular dynamics simulations, PLoS ONE 2 (2007) e880.
[23] S. Jo, T. Kim, V.G. Iyer, W. Im, CHARMM-GUI: a web-based graphical user interface for CHARMM, J. Comput. Chem. 29 (2008) 1859–1865.
[24] S. Jo, J.B. Lim, J.B. Klauda, W. Im, CHARMM-GUI membrane builder for mixed bilayers and its application to yeast membranes, Biophys. J. 97 (2009) 50–58.
[25] J.C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R.D. Skeel, L. Kale, K. Schulten, Scalable molecular dynamics with NAMD, J. Comput. Chem. 26 (2005) 1781–1802.
[26] J.B. Klauda, R.M. Venable, J.A. Freites, J.W. O'Connor, D.J. Tobias, C. Mondragon-Ramirez, I. Vorobyov, A.D. MacKerell Jr., R.W. Pastor, Update of the CHARMM all-atom additive force field for lipids: validation on six lipid types, J. Phys. Chem. B 114 (2010) 7830–7843.
[27] H.C. Andersen, Rattle—a velocity version of the shake algorithm for molecular-dynamics calculations, J. Comp. Physiol. 52 (1983) 24–34.
[28] T. Darden, D. York, L. Pedersen, Particle mesh Ewald—an N. Log(N) method for Ewald sums in large systems, J. Chem. Phys. 98 (1993) 10089–10092.
[29] U. Essmann, L. Perera, M.L. Berkowitz, T. Darden, H. Lee, L.G. Pedersen, A smooth particle mesh Ewald method, J. Chem. Phys. 103 (1995) 8577–8593.
[30] S.E. Feller, Y.H. Zhang, R.W. Pastor, B.R. Brooks, Constant-pressure molecular-dynamics simulation—the Langevin piston method, J. Chem. Phys. 103 (1995) 4613–4621.
[31] W. Humphrey, A. Dalke, K. Schulten, VMD—visual molecular dynamics, J. Mol. Graph. 14 (1996) 33–38.
[32] N. Kucerka, S. Tristram-Nagle, J.F. Nagle, Structure of fully hydrated fluid phase lipid bilayers with monounsaturated chains, J. Membr. Biol. 208 (2005) 193–202.
[33] M. Fujinaga, M.M. Chernaia, R. Halenbeck, K. Koths, M.N. James, The crystal structure of PR3, a neutrophil serine proteinase antigen of Wegener's granulomatosis antibodies, J. Mol. Biol. 261 (1996) 267–278.
[34] H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, P.E. Bourne, The protein data bank, Nucleic Acids Res. 28 (2000) 235–242.
[35] W. Bode, A. Wei, R. Huber, E. Meyer, J. Travis, S. Neumann, X-ray crystal structure of the complex of human leukocyte elastase (PMN elastase) and the third domain of the turkey ovomucoid inhibitor, EMBO J. 5 (1986) 2453–2458.
[36] C. Grauffel, B. Yang, T. He, M.F. Roberts, A. Gershenson, N. Reuter, Cation-pi interactions as lipid-specific anchors for phosphatidylinositol-specific phospholipase C, J. Am. Chem. Soc. 135 (2013) 5740–5750.
[37] B.R. Brooks, C.L. Brooks III, A.D. Mackerell Jr., L. Nilsson, R.J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch, A. Caflisch, L. Caves, Q. Cui, A.R. Dinner, M. Feig, S. Fischer, J. Gao, M. Hodoscek, W. Im, K. Kuczera, T. Lazaridis, J. Ma, V. Ovchinnikov, E. Paci, R.W. Pastor, C.B. Post, J.Z. Pu, M. Schaefer, B. Tidor, R.M. Venable, H.L. Woodcock, X. Wu, W. Yang, D.M. York, M. Karplus, CHARMM: the biomolecular simulation program, J. Comput. Chem. 30 (2009) 1545–1614.
[38] A.D. MacKerell, D. Bashford, M. Bellott, R.L. Dunbrack, J.D. Evanseck, M.J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F.T.K. Lau, C. Mattos, S. Michnick, T. Ngo, D.T. Nguyen, B. Prodhom, W.E. Reiher, B. Roux, M. Schlenkrich, J.C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiorkiewicz-Kuczera, D. Yin, M. Karplus, All-atom empirical potential for molecular modeling and dynamics studies of proteins, J. Phys. Chem. B 102 (1998) 3586–3616.
[39] A.D. Mackerell, M. Feig, C.L. Brooks, Extending the treatment of backbone energetics in protein force fields: limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations, J. Comput. Chem. 25 (2004) 1400–1415.
[40] J.A. Izaguirre, S. Reich, R.D. Skeel, Longer time steps for molecular dynamics, J. Chem. Phys. 110 (1999) 9853–9864.
[41] H. Minoux, C. Chipot, Cation–pi interactions in proteins: can simple models provide an accurate description? J. Am. Chem. Soc. 121 (1999) 10366–10372.
[42] F.N.R. Petersen, M.O. Jensen, C.H. Nielsen, Interfacial tryptophan residues: a role for the cation–pi effect? Biophys. J. 89 (2005) 3985–3996.
[43] T. Giorgino, Computing 1-D atomic densities in macromolecular simulations: the density profile tool for VMD, Comput. Phys. Commun. 185 (2014) 317–322.
[44] T. Lazaridis, Effective energy function for proteins in lipid membranes, Proteins 52 (2003) 176–192.
[45] E. Neria, S. Fischer, M. Karplus, Simulation of activation free energies in molecular systems, J. Chem. Phys. 105 (1996) 1902–1921.
[46] O. Halskau, M. Ying, A. Baumann, R. Kleppe, D. Rodriguez-Larrea, B. Almas, J. Haavik, A. Martinez, Three-way interaction between 14-3-3 proteins, the N-terminal region of tyrosine hydroxylase, and negatively charged membranes, J. Biol. Chem. 284 (2009) 32758–32769.
[47] R. Mouri, K. Konoki, N. Matsumori, T. Oishi, M. Murata, Complex formation of amphotericin B in sterol-containing membranes as evidenced by surface plasmon resonance, Biochemistry 47 (2008) 7807–7815.
[48] A. Olaru, M. Gheorghiu, S. David, C. Polonschii, E. Gheorghiu, Quality assessment of SPR sensor chips; case study on L1 chips, Biosens. Bioelectron. 45 (2013) 77–81.

[49] E. Hajjar, B. Korkmaz, F. Gauthier, B.O. Brandsdal, V. Witko-Sarsat, N. Reuter, Inspection of the binding sites of proteinase 3 for the design of a highly specific substrate, J. Med. Chem. 49 (2006) 1248–1260.
[50] S. Narawane, A. Bundjo, C. Grauffel, B.E. Haug, N. Reuter, In silico design, synthesis and assays of specific substrates for proteinase 3: influence of fluorogenic and charged groups, J. Med. Chem. 57 (2014) 1111–1115.
[51] E.M. Erb, X. Chen, S. Allen, C.J. Roberts, S.J. Tendler, M.C. Davies, S. Forsen, Characterization of the surfaces generated by liposome binding to the modified dextran matrix of a surface plasmon resonance sensor chip, Anal. Biochem. 280 (2000) 29–35.
[52] W. Cho, R.V. Stahelin, Membrane–protein interactions in cell signaling and membrane trafficking, Annu. Rev. Biophys. Biomol. Struct. 34 (2005) 119–151.
[53] S.H. White, W.C. Wimley, Hydrophobic interactions of peptides with membrane interfaces, Biochim. Biophys. Acta 1376 (1998) 339–352.
[54] W.C. Wimley, S.H. White, Experimentally determined hydrophobicity scale for proteins at membrane interfaces, Nat. Struct. Biol. 3 (1996) 842–848.
[55] S. Vanni, L. Vamparys, R. Gautier, G. Drin, C. Etchebest, P.F. Fuchs, B. Antonny, Amphipathic lipid packing sensor motifs: probing bilayer defects with hydrophobic residues, Biophys. J. 104 (2013) 575–584.
[56] G. van Meer, D.R. Voelker, G.W. Feigenson, Membrane lipids: where they are and how they behave, Nat. Rev. Mol. Cell Biol. 9 (2008) 112–124.
[57] D.L. Daleke, Regulation of transbilayer plasma membrane phospholipid asymmetry, J. Lipid Res. 44 (2003) 233–242.
[58] R.F.A. Zwaal, A.J. Schroit, Pathophysiologic implications of membrane phospholipid asymmetry in blood cells, Blood 89 (1997) 1121–1132.
[59] N. Hu, J. Westra, C.G. Kallenberg, Membrane-bound proteinase 3 and its receptors: relevance for the pathogenesis of Wegener's granulomatosis, Autoimmun. Rev. 8 (2009) 510–514.
[60] S. von Vietinghoff, G. Tunnemann, C. Eulenberg, M. Wellner, M.C. Cardoso, F.C. Luft, R. Kettritz, NB1 mediates surface expression of the ANCA antigen proteinase 3 on human neutrophils, Blood 109 (10) (2007) 4487–4493.
[61] A. David, R. Fridlich, I. Aviram, The presence of membrane proteinase 3 in neutrophil lipid rafts and its colocalization with FcgammaRIIIb and cytochrome b558, Exp. Cell Res. 308 (2005) 156–165.
[62] A. David, Y. Kacher, U. Specks, I. Aviram, Interaction of proteinase 3 with CD11b/CD18 (beta2 integrin) on the cell membrane of human neutrophils, J. Leukoc. Biol. 74 (2003) 551–557.
[63] B. Jiang, E. Grage-Griebenow, E. Csernok, K. Butherus, S. Ehlers, W.L. Gross, J.U. Holle, The role of proteinase 3 (PR3) and the protease-activated receptor-2 (PAR-2) pathway in dendritic cell (DC) maturation of human-DC-like monocytes and murine DC, Clin. Exp. Rheumatol. 28 (2010) 56–61.
[64] C.J. Kuckleburg, P.J. Newman, Neutrophil proteinase 3 acts on protease-activated receptor-2 to enhance vascular endothelial cell barrier function, Arterioscler. Thromb. Vasc. Biol. 33 (2013) 275–284.
[65] B. Korkmaz, A. Kuhl, B. Bayat, S. Santoso, D.E. Jenne, A hydrophobic patch on proteinase 3, the target of autoantibodies in Wegener granulomatosis, mediates membrane binding via NB1 receptors, J. Biol. Chem. 283 (2008) 35976–35982.
[66] C. Kantari, M. Pederzoli-Ribeil, O. Amir-Moazami, V. Gausson-Dorey, I.C. Moura, M.C. Lecomte, M. Benhamou, V. Witko-Sarsat, Proteinase 3, the Wegener autoantigen, is externalized during neutrophil apoptosis: evidence for a functional association with phospholipid scramblase 1 and interference with macrophage phagocytosis, Blood 110 (2007) 4086–4095.
[67] J. Gabillet, A. Millet, M. Pederzoli-Ribeil, P. Tacnet-Delorme, L. Guillevin, L. Mouthon, P. Frachet, V. Witko-Sarsat, Proteinase 3, the autoantigen in granulomatosis with polyangiitis, associates with calreticulin on apoptotic neutrophils, impairs macrophage phagocytosis, and promotes inflammation, J. Immunol. 189 (2012) 2574–2583.

Two homologous neutrophil serine proteases bind to POPC vesicles with different affinities: when aromatic amino acids matter

Anne-Sophie SCHILLINGER[1,2], Cédric GRAUFFEL[1,2], Hanif Muhammad KHAN[1,2],Øyvind HALSKAU[1], Nathalie REUTER[1,2*]

SUPPORTING INFORMATION

Figure S1
RMSD along the MD trajectory calculated for PR3 (grey) and HNE (green), with respect to the enzyme structure used as starting point for the simulation (enzyme structure solvated and minimized).

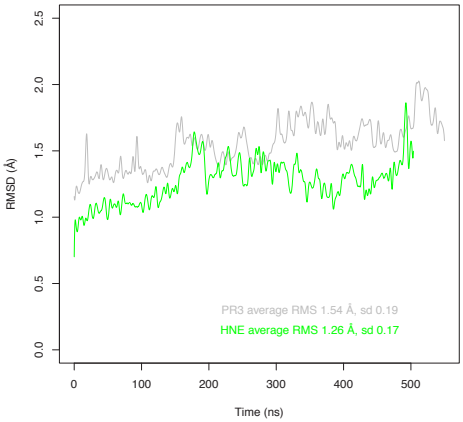


Figure S2
Distance of the center of mass of PR3 (grey) and HNE (green) to the average plane of the phosphorus atoms. Negative values indicate that the center of mass lies in the water slab above the phosphorus atoms.

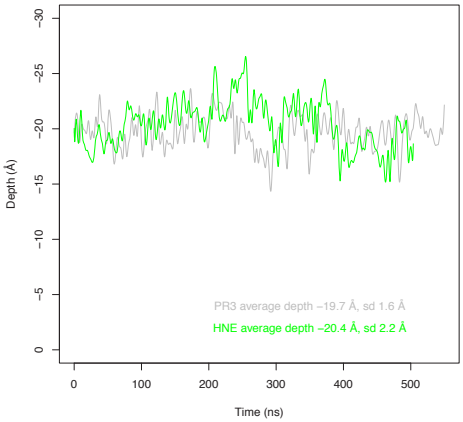
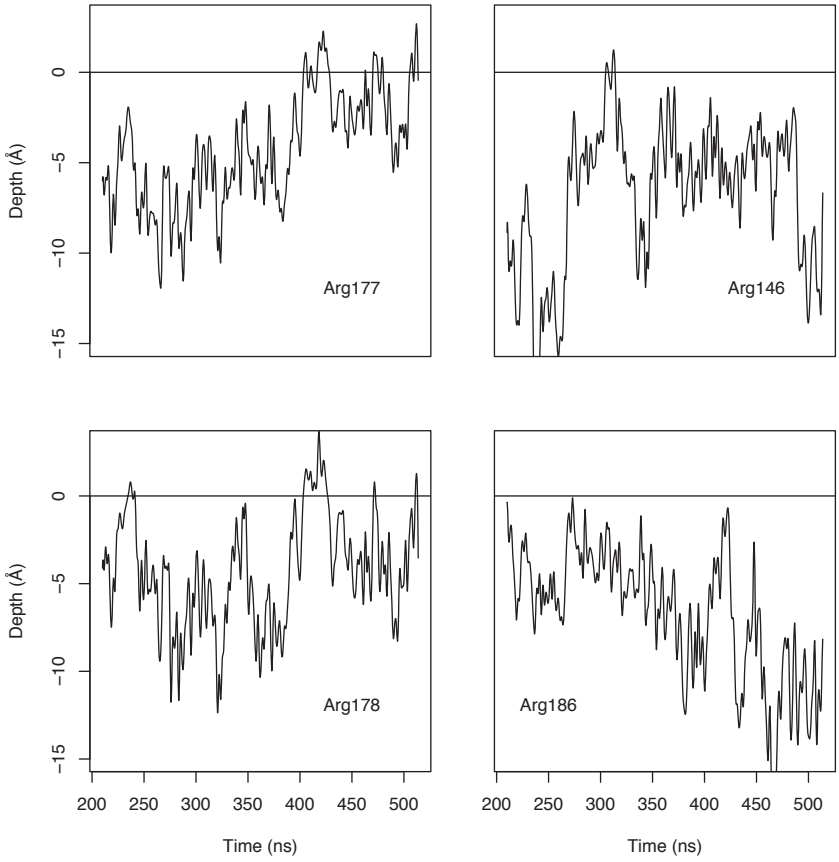


Figure S3
Depth of anchorage along simulation time calculated for amino acids Arg 146 (loop β7-β8), Arg177 (loop β8-β9), Arg178 (loop β8-β9) and Arg186 (loop β9-β10). The depth is calculated as described in the Methods section, positive values indicate that the center of mass of the amino acid is buried in the bilayer beyond the plane defined by the phosphate groups.



III

A role for weak electrostatic interactions in peripheral membrane protein binding

H. M. Khan[1,2], T. He[3], E. Fuglebakk[1,2], C. Grauffel[1,2], B. Yang[3,4], M. F. Roberts[3], A. Gershenson[4], N. Reuter[1,2*]

[1] Department of Molecular Biology, University of Bergen, Norway;
[2] Computational Biology Unit, Department of Informatics, Bergen, Norway;
[3] Department of Chemistry, Boston College, Chestnut Hill, U.S.A.;
[4] Department of Biochemistry and Molecular Biology, University of Massachusetts, Amherst, U.S.A.
[*] Corresponding Author- NR: Universitetet i Bergen, Computational Biology Unit, Institutt for informatikk, Pb. 7803 N-5020 Bergen, nathalie.reuter@uib.no.

ABSTRACT

Bacillus thuringiensis phosphatidylinositol specific phospholipase C (*Bt*PI-PLC) is a secreted virulence factor that binds specifically to negatively charged lipid bilayers containing phosphatidylcholine (PC). *Bt*PI-PLC carries a negative net charge and its interfacial binding site (IBS) has no obvious cluster of basic amino acids. Continuum electrostatics calculations show that, as expected, nonspecific electrostatic interactions between the protein and membranes vary as a function of the fraction of anionic lipids present in the bilayers. Yet they are strikingly weak with a calculated ΔG_{el} below 1 kcal/mol, largely due to a single lysine residue (K44). When K44 is mutated to alanine the *Bt*PI-PLC equilibrium dissociation constant for small unilamellar vesicles increases more than fifty times (ca. 2.4 kcal/mol) suggesting that interactions between Lys44 and lipids are not merely electrostatic. Comparison between trajectories obtained with different lipid compositions reveal that neither hydrogen bonds nor hydrophobic contacts between the *Bt*PI-PLC IBS and bilayers are affected by the bilayer composition. The occupancies of cation-π interactions between choline groups of PC lipids and protein tyrosines, however, vary as a function of the PC content. Interestingly there is no recruitment of choline to PC-high-affinity binding sites. Basic amino acids can contribute to both nonspecific electrostatic and short-range protein-lipid interactions. Their overall contribution to binding affinity is context-dependent and cannot be approximated by a rule-of-thumb value. Further, statistics on the distribution of basic amino acids in a dataset of membrane-binding domains reveal that weak electrostatics might be a more common mechanism for peripheral membrane binding than generally thought.

Keywords

amphitropic membrane proteins, molecular dynamics simulations, electrostatics, fluorescence correlation spectroscopy, phosphatidylinositol-specific phospholipase C (PI-PLC)

INTRODUCTION

The association of peripheral membrane proteins with biological membranes is classically described as an electrostatically-driven approach followed by the intercalation of hydrophobic side groups into the lipid bilayer. Long-range nonspecific electrostatic forces between the negatively charged membrane and clusters of basic amino acids bring the protein into a binding-competent orientation relative to the lipid bilayer and play a major role for numerous prototypical peripheral membrane proteins(1-4). Experimental and computational pioneering studies have evaluated nonspecific electrostatics to contribute a few kilocalories per mole to the overall affinity and estimated that each basic amino acid contributes up to 1 kcal/mol to the binding free energy(2, 5).

Bacillus thuringiensis phosphatidylinositol specific phospholipase C (*Bt*PI-PLC) is a 34.8 kDa secreted virulence factor that carries a negative net charge (-7 e) and binds specifically to phosphatidylcholine-containing negatively charged lipid bilayers. Its interfacial binding site (IBS) consists of a small α-helix, helix B(6), and two neighboring loops rich in tyrosines that we have shown engage in cation-π interactions with the choline groups of DMPC lipids in neutral bilayers(7). These cation-π interactions provide a likely molecular mechanism for *Bt*PI-PLC phosphatidylcholine (PC) specificity(7, 8) but do not account for its preference for bilayers containing a small fraction of anionic lipids.

The affinity of *Bt*PI-PLC has been measured to be tightest for vesicles containing 20% anionic lipids. In these conditions the affinity is ca. 4 times more favorable than for neutral vesicles(9) while higher anionic lipid content decreases the affinity considerably. Surprisingly, for the tightest binding conditions, mutating a single lysine (K44) decreases the affinity to $1/55^{th}$ of that of the wild type (WT) protein corresponding to ca. 2.4 kcal/mol. This effect is ca. 14 times greater than the fourfold effect obtained when adding 20% anionic lipids to PC vesicles. If mutating K44 affected only the nonspecific electrostatic forces involved in the adsorption of protein onto the phospholipid bilayer the effect of removing the anionic lipids should be comparable to that of mutating K44(6, 10). Furthermore the effect of the K44A mutation on the affinity of *Bt*PI-PLC towards negatively charged vesicles is strikingly high compared to values found in the literature(2, 5).

Here we investigate the forces driving *Bt*PI-PLC specificity for negatively charged PC-containing vesicles by separately addressing the two steps governing membrane affinity: association (k_{on}) and dissociation (k_{off}). While the former is fast, the dissociation that follows is comparatively slow and constitutes the rate-limiting step. As a consequence, in systems where the protein binds to the membrane without undergoing significant conformational changes or interacting with another protein, the affinity for the membrane is mostly accounted for by interactions between the protein interfacial binding site (IBS) and lipids. As these are difficult to assess experimentally molecular dynamics (MD) simulations are widely used for that purpose (11-14).

Using multiple 500 nanosecond-long MD simulations of *Bt*PI-PLC at the surface of pure dimyristoylphosphatidylcholine (DMPC), pure dimyristoylphosphatidylglycerol (DMPG), and mixed DMPC:DMPG bilayers we map the specific protein-lipid

interactions and investigate how these are influenced by the anionic lipid content. We also evaluate the nonspecific electrostatic contributions of key basic amino acids to the association step using continuum electrostatics calculations and experimental determinations of the effect of their mutation to alanine. The latter is achieved by using fluorescent correlation spectroscopy (FCS) to measure the affinity of *Bt*PI-PLC variants for small unilamellar vesicles. Combining the computational and experimental data allows us to formulate a complete model of *Bt*PI-PLC membrane binding. A novel characteristic of this model is the unusually weak nonspecific electrostatics contribution due to the curious number and distribution of lysines and arginines on the *Bt*PI-PLC surface. In an attempt to evaluate how (un)common such a distribution is, we finally analyze the number and distribution of basic amino acids in a database of peripheral membrane binding proteins and find that a significant number of peripheral membrane proteins are highly likely to also display weak nonspecific electrostatics. We propose that such weak interactions could be beneficial for the function of peripheral membrane proteins, particularly those that must exhibit a quick response to environmental changes.

MATERIALS AND METHODS

Molecular dynamics simulations

As there is no X-ray structure available for wild type *Bt*PI-PLC, we built a model as described by Grauffel et al.(7), using the X-ray structure of the *B. thuringiensis* Y247S/Y251S(15) mutant (PDB ID: 3EA1) and W47A/W242A(16) mutant (PDB ID: 2OR2). The starting orientation of bilayer-bound *Bt*PI-PLC was obtained from the same study, and based on implicit membrane simulations(7).

The protein was manually docked on the pre-equilibrated mixed bilayers. The protocol used for preparation of the bilayers is described in Supporting Information. The membrane orientation was obtained from *Bt*PI-PLC implicit simulations(7). Lipids were removed to avoid coordinate overlapping and steric clashes. Then the system was minimized as described by Grauffel et al.(7) and solvated with TIP3P water molecules using VMD(17). After solvation, additional sodium ions were added by randomly replacing water molecules to achieve an overall charge neutral system where necessary.

MD Simulations of bilayer bound BtPI-PLC The combined protein-lipid system was then subjected to two short 400 ps equilibrations in the NVT ensemble with constraints on the protein backbone. Subsequently, the system was equilibrated for 2 ns in the NPT ensemble without any constraint before finally performing the 500 ns NPT simulation. The temperature was set to 310 K during the simulation with a 2 fs integration time step in NAMD (v2.9)(18). The temperature was controlled as described above for the mixed bilayer. The pressure was set to 1 atm with an oscillation period of 200 fs and a damping time scale of 50 fs. The CHARMM all-atom force field(19) (c22 including CMAP correction)(20) and the force field update for lipids (CHARMM36)(21) were used. Trajectory conformations were saved every 10 ps. Two simulations were performed, and in the second replicate, the protein was rotated by 180 degree around the bilayer normal (z-axis) to allow for different initial protein-lipid contacts. This is done to avoid bias in protein-lipid interactions due to the initial distribution of lipids under or around the protein, and also to improve the sampling. Analyses of the trajectories were performed

using CHARMM (v33b1)(22) and VMD (v1.9.1)(17) on the last 450 ns of each simulation in a similar fashion as in Grauffel et al.(7). Hydrophobic contacts, hydrogen bonds and cation-π interactions are averaged over the two replicas. Hydrophobic interactions and hydrogen bonds were assigned as described in Supporting Information. Cation-π interactions between the aromatic amino acids (Tyr, Phe and Trp) and the choline group of the DMPC lipids were assigned as described earlier(7, 23).

Continuum electrostatics calculations

We extracted the structures of the bilayers after simulations to perform the continuum electrostatics calculations by solving the Poisson-Boltzmann equation. These equilibrated bilayers provide better description of the membrane than ideal bilayers. The protein-membrane complex is prepared by placing the protein with its membrane binding orientation above the bilayer upper leaflet. We define the minimum protein-membrane distance (d) as the z-distance between the HD1 atom of Ile43 and the maximum z-coordinate of the bilayer. HD1 is a hydrogen atom on the Cδ atom of the Ile43 chosen because it is the most deeply anchored atom in the membrane binding conformation. The initial position of the protein is set to d=12 Å. The protein is then gradually translated towards the membrane. Both the protein and the membrane are kept rigid. Poisson-Boltzmann calculations are carried out using APBS (version 1.3) (24). Once the electrostatic potential is calculated, the electrostatic free energy can be calculated from the electrostatic potential(25, 26). A similar scheme has been used for other proteins before(27-29). The non-polar contribution to the binding free energy is not evaluated, as that is not the goal here.

Database Statistics

The statistics for net charges and the distributions of basic amino acids in peripheral membrane proteins was obtained from a dataset based on the OPM-database(30). This database contains membrane associated proteins, and predictions of their orientation in a DOPC bilayer model. We selected all protein families of the OPM-classification classified as type *Monotopic/Peripheral*, for a total of 337 families. For the selected families, we obtained average net charges, $\hat{c}_F$, for each family. We defined solvent exposed residues to be residues that have at least 20% of their surface contributing to the surface of the protein its maximal exposure. Using MMTK(31), we calculated surface contributions from each atom using a radius around the atom equal to the sum of its van der Waals radius and a fixed 1.4 Å term to mimic water accessibility. Maximal exposure for each residue type is obtained from a random coil model(32). We defined the insertion coordinate i, of a point to be the distance from that point to the end of the hydrocarbon region of the membrane in the membrane model employed in the OPM database(33). This distance was made negative for points on a hydrated side of the membrane. We use $n_p^{\text{basic}}(z,x)$ to denote the number of solvent exposed basic amino acids in protein p, that have alpha carbons with insertion coordinates $z \leq i$ and $i<x$. The number of basic amino acids at the IBS is then taken to be $n_p^{\text{basic}}(0,-10\text{Å})$, and we report the family average, w_F^{basic}. We also calculated the average surface density of basic amino acids at varying distances from the membrane ($n_p^{\text{basic}}(z,x)/n_p^{\text{any}}(z,x)$), which we report as d_S^{basic} (More details are provided as Supporting information).

Expression and purification of *Bt*PI-PLC and its variants

All *Bt*PI-PLC mutants were constructed from the plasmid containing the mutant N168C *Bt*PI-PLC gene using a QuikChange site-directed mutagenesis kit from Agilent Technologies. Where the Cys at 168 is used to fluorescently label the proteins for FCS experiments. Mutated genes were sequenced to ensure incorporation of the correct mutations and that no other mutation had arisen. All *Bt*PI-PLC variants expressed at high levels in *E. coli* and were purified with two chromatographic steps (elution on a Q-Sepharose fast flow column followed by a phenyl-Sepharose column as described for the recombinant *Bt*PI-PLC(10)). More than 90% purity was achieved as monitored by SDS-PAGE. Protein solutions were concentrated and the protein concentration was calculated from the absorption at 280 nm using extinction coefficients estimated by ProtParam(34). The composition of secondary structure elements, estimated from far UV circular dichroism (CD) data, were essentially the same for all the variants indicating no significant secondary structure changes (see supporting information Table S9). *Bt*PI-PLC thermal stability was assessed using far UV CD and monitoring the ellipticity at 222 nm while increasing the sample temperature 0.5 degree per minute (6, 35). Specific activities for all variants (see supporting information) were determined towards SUVs prepared by sonication and composed of 2 mM phosphatidylinositol (PI) plus various concentrations of 1-palmitoyl-2-oleoyl-phosphatidylcholine (POPC) in 50 mM HEPES (containing 1 mg/ml bovine serum albumin (BSA)), pH 7.5. For pure PI and PI/PC (4:1) SUVs, the total phospholipid concentration is below or comparable to the K_d, hence mutant specific activities are lower than WT enzyme. For PI/PC SUVs with X_{PC} = 0.5 and 0.8, the total phospholipid concentration is above the apparent K_d for all the Lys/Arg to Ala mutants except K44A. Specific activities were measured with ^{31}P NMR spectroscopy as described previously (9, 36). All phospholipids were obtained from Avanti Polar Lipids.

FCS measurements of *Bt*PI-PLC binding to SUVs

The FCS based SUV binding experiments take advantage of the fact that protein binding to vesicles slows translational diffusion. FCS experiments were performed using *Bt*PI-PLC variants labeled at N168C with Alexa Fluor 488 maleimide and a home-built confocal setup based on an IX-70 inverted microscope (Olympus) as previously described (37) (Cf. Supporting Information).

RESULTS

Influence of PC content on short-range specific protein-lipid interactions

In order to determine the roles of short-range protein-lipid interactions in *Bt*PI-PLC membrane affinity and identify those responsible for its lipid specifity, we performed multiple 500 ns-long MD simulations for *Bt*PI-PLC docked to pre-equilibrated DMPC:DMPG bilayers with four different mole fractions of DMPC, X_{PC} = 0, 0.5, 0.8 or 1.0 (See Methods section and Table S1). The simulation results for X_{PC}=1.0 are taken from earlier reported simulations(7). Although the residence time of *Bt*PI-PLC on small unilamellar vesicles (SUVs) has been measured to be a few 100s of milliseconds(8), the

use of shorter MD simulations to map relevant protein-lipid interactions has proven reliable(7, 8, 38). In order to avoid bias due to the initial distribution of lipids around the protein all simulations were repeated using different initial positions of the protein in the membrane plane. We also performed a simulation using the K44A *Bt*PI-PLC mutant bound to an anionic membrane (X_{PC}=0.8).

Interactions with a pure DMPG bilayer (X_{PC}=0) Both simulations of WT *Bt*PI-PLC docked to pure DMPG bilayers indicate loose binding of the protein to pure DMPG bilayers compared to DMPC-containing bilayers. In one of the simulations, the protein completely detaches from the bilayer within 200 ns (Fig. 1). In the other simulation, the protein remains bound but the structure of the β7-αG loop (also called the rim loop) becomes distorted (Fig. 2). The average backbone RMSD of β7-αG along the simulation time is 1.8±0.4 Å compared to 0.4±0.1 Å in simulations with PC-containing bilayers (see Fig. S1 in Supporting Information). Compared to these simulations we also observe a loss of long-lived backbone hydrogen bonds (N243-G238, N243-T240 and Y248-S244) in the loop. A similar distortion of the β7-αG loop was observed in simulations of the free Y248A variant that shows impaired lipid binding with a K_d ca. 150 times higher than WT *Bt*PI-PLC towards vesicles with X_{PC} = 1(7).

A B

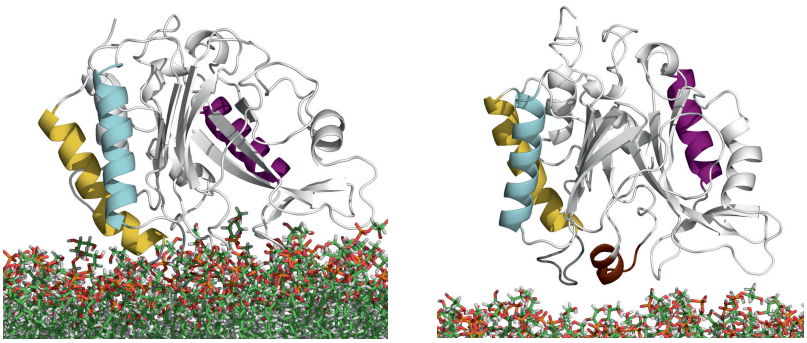


FIGURE 1 Snapshots from one of the MD simulations of WT *Bt*PI-PLC on a DMPG bilayer (X_{PC}=0, replica 2). (A) System at the beginning of the production run, (B) system after 200 ns. *Bt*PI-PLC is represented by cartoons and the lipids with sticks. The secondary structure elements relevant for this study are colored as follows: helix B (αB, brown), helix D (αD, magenta), helix F (αF, cyan), helix G(αG, yellow). For the sake of clarity, water molecules and ions are not shown.

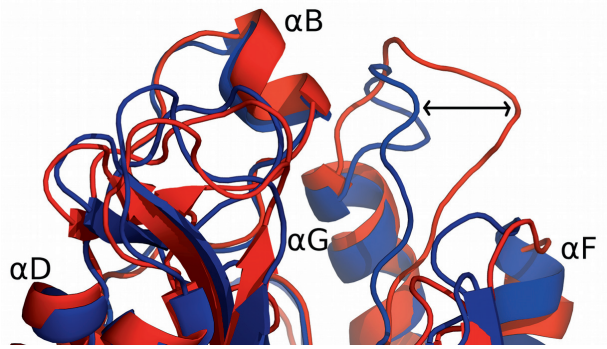


FIGURE 2 Comparison of the *Bt*PI-PLC membrane bound secondary structure after a 500 ns simulation for X_{PC}=0 (red, pure DMPG) and X_{PC}=0.5 (blue). The backbone structures were aligned using Pymol(39). For the sake of clarity, lipids, water molecules and ions are not shown. The essential intramolecular interactions within the rim loop (β7-αG) are lost at X_{PC}=0 causing a distorted loop structure (arrow).

Interactions *with DMPC:DMPG bilayers (X_{PC}=[0.5, 0.8, 1])* All of the simulations with bilayers containing PC lipids yielded stable *Bt*PI-PLC anchorage with no significant structural changes compared to the X-ray crystal structure. The backbone RMSD is at most 1.61 Å (average values for each simulation are provided as Supporting Information in Table S2). The protein anchors helix B and the rim loop in the bilayer interface (Fig. 3) at similar depth (Fig. 4) independent of the PC lipid content.

For each of the three bilayer compositions we inventoried the interactions between bilayer lipids and protein residues located at the interface. We report in Table 1 the numbers of hydrogen bonds, hydrophobic contacts and cation-π interactions per frame, averaged over each of the bilayer compositions and replicas. The corresponding values per amino acid are provided as Supporting Information (Supp. Inf. Table S3). Briefly, three main *Bt*PI-PLC regions mediate most of the interactions with lipids, namely helix B, the β2-αD loop and the β7-αG loop. Helix B anchors deepest with most of its amino acids below the phosphate plane (Fig. 3 and 4). Residues 238-242 of the β7-αG loop are also inserted below the average phosphate plane (see Table S4 in Supporting Information). All three anchored regions mediate hydrophobic interactions with multiple lipid tails; for example, helix B mediates 25 to 26 hydrophobic contacts irrespective of the PC content in the bilayer (Fig. 5A). The hydrophobic contacts mediated by *Bt*PI-PLC with the lipids do not vary significantly with X_{PC}. The same applies to hydrogen bonds. We do observe long-lived hydrogen bonds between lipid phosphate groups and side-chains of charged (K44, R71, K122, K201), polar (S236, S244), and aromatic residues (Y88, Y246, Y247, Y251) as well as backbones of polar residues (Q40, N41). Yet there is no correlation between the number and stability of these hydrogen bonds and the mole ratio of PC lipids in the bilayers.

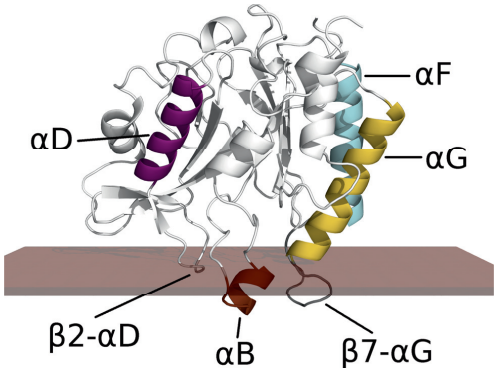


FIGURE 3 *Bt*PI-PLC secondary structure elements and their calculated insertion at the interface of a mixed DMPC:DMPG bilayer (X_{PC}=0.5). *Bt*PI-PLC is represented by cartoons. The relevant secondary structure elements are labeled and colored as in Fig. 1. Helix B (brown) is inserted beyond the phosphate plane (brown slab) of the mixed bilayer, constructed from MD simulation data.

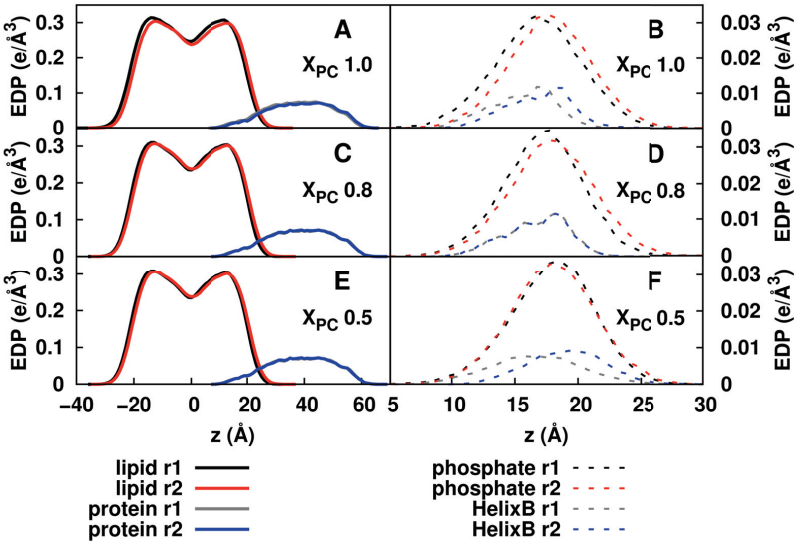


FIGURE 4 Electron Density Profiles (EDP) from MD simulations for three different bilayers. (A), (C), (E) protein, (gray, r1=replica 1 and blue r2=replica 2), and lipids (r1, black; r2, red), for X_{PC}=1, 0.8 and 0.5. (B), (D), (F) deepest helix B (gray and blue) and phosphate groups of the upper lipid leaflet (red and black).

TABLE 1 Inventory of interactions with pure DMPC and mixed bilayers: average number of hydrogen bonds, hydrophobic contacts and cation-π per frame are given for each replica (r1/r2)

X_{PC}	0.5	0.8	1.0
Hydrogen bonds	7.9/4.4	7.0/7.6	7.8/6.9
Hydrophobic contacts	49.2/41.1	41.7/40.2	45.0/42.3
Cation-π	2.5/2.2	2.8/2.9	3.6/3.0

The trend is different for cation-π interactions between choline headgroups of PC lipids and tyrosines. The overall number of cation-π interactions clearly increases with PC enrichment. We reported earlier (7) that Y88, Y204, Y246 or Y251 engage in high-occupancy cation-π interactions in simulations of *Bt*PI-PLC with pure DMPC bilayers. Adding DMPG to the bilayers we observe that the occupancy of cation-π interactions varies with X_{PC} (Fig. 5B). The occupancies of cation-π interactions mediated by Y246, Y251 and Y204 increase from X_{PC}=0.5 to 0.8 but not significantly between X_{PC}=0.8 and 1.0. Although we ensure that the simulations are not biased by the initial lipid distribution, we cannot rule out the possibility that this apparent saturation is an artifact of the relatively short time scale of our simulations. Y88 and Y246 are located close to the IBS and mediate the cation-π interactions with the highest occupancy. They are also the Tyr residues whose mutation to alanine has the largest effect on protein affinity for SUVs(7) (also see Fig. S2 in Supp. Inf.). Overall the occupancy of these interactions for tyrosines 86, 88, 204, 246, 247, 251 correlates qualitatively well with the effects of their mutation to alanine on *Bt*PI-PLC affinity for SUVs (Cf. Ref. (7) and Fig. S2 in Supp. Inf.). Y251 is the only tyrosine that we observe mediating two cation-π interactions simultaneously with an occupancy of up to 11% at X_{PC} = 0.8 (r1), which might be due to the accessibility of its sidechain to PC lipids (see Tables S5 and S6 in Supp. Inf.).

A B

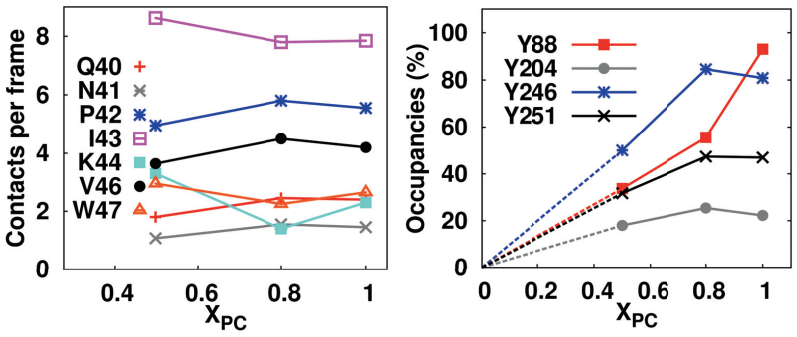


FIGURE 5 Influence of bilayer composition on short-range protein-lipid interactions. (A) Average number of hydrophobic contacts per frame between helix B amino acids and the bilayer lipids as a function of X_{PC}. The membrane composition has no significant effect

on hydrophobic insertion. (B) Cation-π interactions between Tyr amino acids and choline headgroups of DMPC lipids. The occupancies along the simulation are plotted with respect to membrane PC content.

Interactions between BtPI-PLC K44A with a mixed bilayer (X_{PC}=0.8) Experimental data indicate that mutating K44 to alanine drastically affects the affinity of *Bt*PI-PLC for SUVs. We here investigate the effect of this mutation on specific protein-lipid interactions by docking K44A *Bt*PI-PLC on a mixed bilayer (X_{PC}=0.8). Unlike simulations with pure DMPG bilayer, or the Y248A mutant in water(7), K44A remains bound to the membrane surface with an unaltered structure. The average RMSD of the protein backbone along simulation time relative to the energy minimized K44A structure is 1.5±0.2 Å. Analysis of the depth of anchorage of different amino acids shows negligible variation along the simulation but major interactions are lost (Cf. Supporting Information Table S7). In particular, an important protein-lipid hydrogen bond network, involving Q40 and N41 backbone atoms, the K44 side chain and DMPC phosphates, is lost around helix B (7). As a consequence the number of hydrogen bonds mediated by helix B is low compared to WT. Moreover helix B mediates slightly fewer hydrophobic contacts per frame (24.7 instead of 25.7). As might be expected, the occupancies of cation-π interactions are generally comparable to those observed for WT *Bt*PI-PLC.

Non-specific electrostatics interactions upon membrane association

The *Bt*PI-PLC IBS includes five lysines (K38, K44, K122, K201, K279) and one arginine. While K44 is located on Helix B, the four other lysines and one arginine (R71) are carried by other structural elements and do not form the obvious cluster(s) often described for membrane binding domains. This raises the question of the magnitude of the nonspecific electrostatic interactions between proteins and negatively charged membranes.

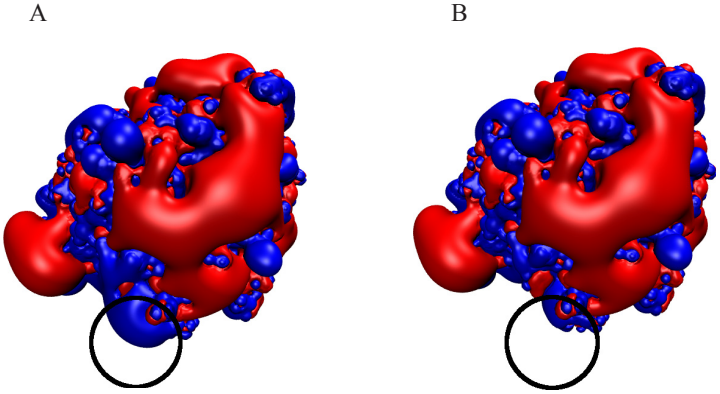


FIGURE 6 Calculated electrostatic potentials of (A) WT *Bt*PI-PLC and (B) K44A. The isocontours at +1 k_BT/e (blue) and -1 k_BT/e (red) are shown. Note the change of the isosurface around the K44A mutation site in helix B (black circle). The images were prepared using VMD (17).

Electrostatic free energy profile We calculate the electrostatic component of the binding free energy using continuum electrostatics and solving the Poisson-Boltzmann (PB) equation (Cf. Methods section and Supp. Inf. Fig. S3). This approach has been successfully used to describe the electrostatic properties of proteins, nucleic acids and membranes (24, 40, 41). Note that the chosen approach only allows us to reasonably estimate the electrostatic contribution when the protein is above the modeled membrane and not in its most favorable depth of anchorage. The latter is measured by the MD simulations presented above to be about 8 to 10 Å below the membrane surface, as measured by the position of the center of mass of isoleucine 43, the deepest anchored residue.

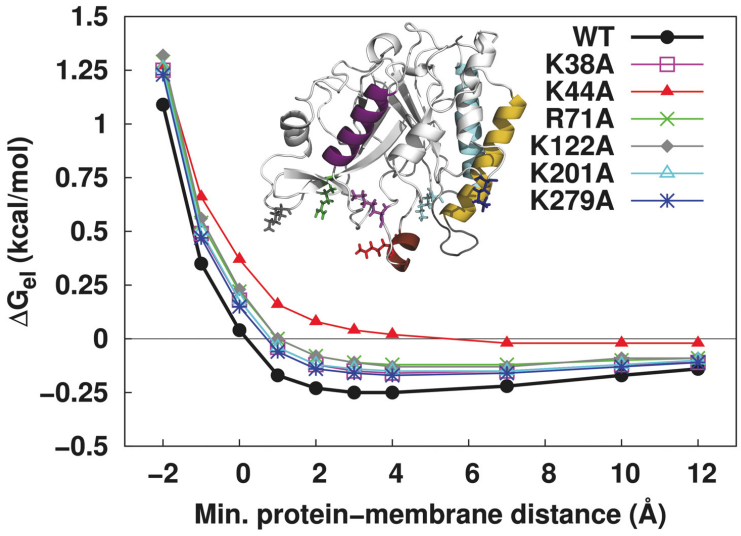


FIGURE 7 Electrostatic free energy profile of WT and mutants with an anionic membrane (X_{PC}=0.8) using a 0.1M salt concentration. A hydrogen atom in Ile43 (Charmm nomenclature HD1) is the closest *Bt*PI-PLC atom to the membrane in the membrane binding orientation where the approach is along the z-axis perpendicular to the membrane. The position of this atom relative to the membrane is decreased in order to drag the protein towards the membrane while keeping both the protein and the membrane rigid.

Isocontours at $\pm 1 k_B T/e$ of the electrostatic potential of *Bt*PI-PLC (Fig. 6A) show large negative regions while the positive regions are more restricted in size with one around helix B. The profile of the electrostatic free energy for WT approaching a negatively charged bilayer (DMPC:DMPG, X_{PC}=0.8) shows a minimum at a minimum protein-membrane distance d of 3 to 4 Å (Fig. 7). The electrostatic free energy is at most -0.25 kcal/mol, much lower than that reported for other membrane binding domains using the same computational approach (from -3 to -5 kcal/mol)(27). The electrostatic free energy

becomes less favorable as the protein moves towards the membrane from its minimum eventually becoming unfavorable when the protein is in contact with the membrane surface (+0.4 kcal/mol) and crosses the upper limit of the membrane (+1.1 kcal/mol) (Fig. 7). The favorable contribution is largely due to K44 and the K44A mutation almost completely abolishes the favorable ΔG_{el} (Fig. 6B and 7). Mutations at other positively charged residues (K38, R71, K122, K201 and K279) affect ΔG_{el} to a lesser extent with a value of ca. -0.1 kcal/mol at the most favorable distance.

Effect of membrane composition and salt concentration We evaluate the dependence of the protein electrostatic free energy on lipid composition by calculating the electrostatic free energy at d=3 Å for decreasing ratios of DMPC lipids (X_{PC}=1.0, 0.8, 0.5 and 0) (Fig. 8A). The electrostatic free energy is slightly unfavorable for a neutral membrane (0.1 kcal/mol) and decreases monotonically with the PC content until X_{PC}=0.5 (ΔG_{el} = -0.68 kcal/mol). It then increases to -0.42 kcal/mol with X_{PC} decreasing to 0, in agreement with experimental data (42).

We further calculated ΔG_{el} for d=3 Å and X_{PC}= 0.8 at salt concentrations ranging from 0.025 to 0.7 M (Fig. 8B). *Bt*PI-PLC shows a quasi-parabolic dependence of the electrostatic free energy on salt concentration. At the lowest ionic strength tested (0.025M), the electrostatic free energy of interaction between the protein and the membrane is slightly unfavorable (0.08 kcal/mol). It then quickly becomes more favorable until a salt concentration of 0.1 M (-0.25 kcal/mol) which is also approximately the physiological ionic strength surrounding the target eukaryotic cell. It then gradually becomes less favorable with increasing salt concentrations. This behavior can be explained by the fact that the unfavorable interactions between the protein and the bilayer are not significantly screened at low salt concentrations (see large negative isocontours of the electrostatic potential on Fig. 6A). However it is important to note that the electrostatic partitioning of *Bt*PI-PLC is low.

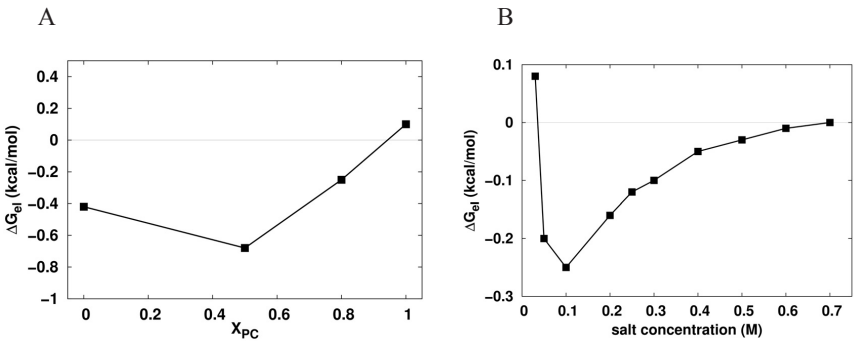


FIGURE 8 Influence of bilayer composition and salt concentration on the calculated electrostatic free energy. (A) The electrostatic free energy becomes more favorable as X_{PC} decreases from 1 to 0.5, and increases from X_{PC}=0.5 to X_{PC}=0. Calculations are done at 0.1 M KCl and with a protein-membrane distance of 3 Å (minimum ΔG_{el} in Fig. 7). (B) Electrostatic free energy as a function of salt concentration where the salt

concentration is varied from 0.025 to 0.7 M KCl. The protein is in the same position with respect to the bilayer as in (A).

The affinity of *Bt*PI-PLC variants for SUVs

We determined, using FCS, SUV binding affinities for a selection of *Bt*PI-PLC mutants (Fig. 9 and Supp. Inf. Table S8). K38, R71 and K279 are basic amino acids located in or close to the IBS of *Bt*PI-PLC. Their mutation to alanine does not significantly affect the protein structure (Cf. Table S9 in Supp. Inf.) and they all show comparable enzymatic activity relative to WT at X_{PC} = 0.5 (see Supporting Information Fig. S4). Lower specific activity towards X_{PC} = 0 and 0.2 SUVs correlates with the loss of binding affinity provided by removal of a given Lys or Arg (Cf. Supp Info Fig. S4B).

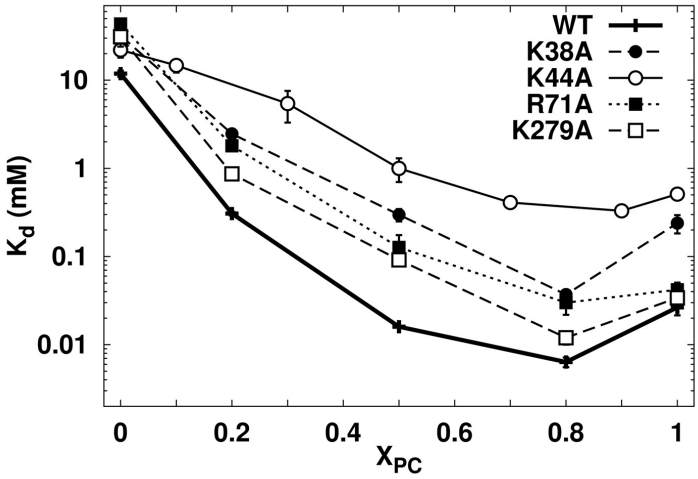


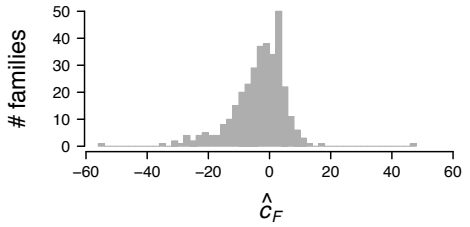
FIGURE 9 Binding of *Bt*PI-PLC WT and cationic amino acids mutants to SUVs. The apparent K_d is plotted as a function of X_{PC}. K_d values are provided as Supporting Information (Table S8).

Experiments measuring affinity of the mutants for PC/PG SUVs show that R71A and K279A have apparent dissociation constants, K_ds, that are 2 to 4 times higher than that for WT towards SUVs containing X_{PC} = 0 to 0.5, and these mutants recover binding affinity at higher X_{PC} presumably due to the increased importance of specific tyrosine mediated cation-π interactions with PC relative to electrostatics. The K38A mutation is more perturbing with an approximately 10 times lower K_d relative to WT for 0.3 to 0.5 X_{PC} SUVs and little to no recovery of affinity at higher X_{PC} suggesting that this mutation may perturb both electrostatic and PC-specific interactions. However, none of these variants show the two orders of magnitude or greater decreases in K_d observed for K44A interactions with SUVs containing 0.1 to 0.9 X_{PC}. Surprisingly, the K44A K_d stayed relatively stable up to X_{PC}=0.8, a region in which the affinity of WT for SUVs increases

substantially. Then at X_{PC}=0.5 the K_d was approximately 50 times higher than that of WT. This is similar to the results for K38A where perturbations near helix B affect both anionic binding and specific PC recognition.

Distribution of basic amino acids in a database of peripheral membrane proteins

A



B C

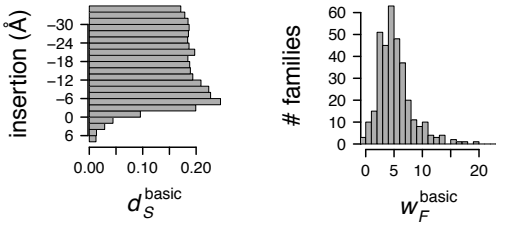


FIGURE 10 Peripheral membrane proteins often lack large basic patches. (A) Distribution of mean net charges ($\hat{c}_F$) for protein families in OPM. (B) The mean surface density of basic amino acids (d_S^{basic}), as a function of position along the membrane normal in predicted membrane-binding orientations from OPM. Positive values along the insertion axis correspond to the hydrocarbon region of the membrane. (C) Distribution of the mean number of basic amino acids close to the membrane (w_F^{basic}).

The low electrostatic partitioning combined with the major role of one particular basic residue is unexpected for a peripheral membrane protein. Our MD simulations and continuum electrostatics calculations indicate that this is due to the small number of basic amino acids at the *Bt*PI-PLC IBS and their spatial distribution. In order to find out whether this feature of *Bt*PI-PLCs is shared by other peripheral membrane proteins we performed a statistical analysis on the predicted membrane orientations of proteins classified as Peripheral/Monotopic in the OPM database(30). Distribution of overall charges (Fig. 10A) reveals that the balance of acidic and basic amino acids in these proteins does not show any over-representation of net positive charge. The mean density of basic amino acids on the protein surface seems to be slightly higher close to membrane head groups (Fig. 10B, where the phosphate density is expected to peak at insertion coordinates around -4 Å(33)). Observing the distribution of the mean number of basic amino acids within 10 Å of the end of hydrocarbon region (Fig. 10C), reveals that it is not uncommon for protein families in the data set to have 3 or fewer basic residues in this

region. These results suggest that *Bt*PI-PLC is not unique and many peripheral membrane proteins may lack large surface clusters of basic amino acids.

DISCUSSION

Using continuum electrostatics calculations, molecular dynamics simulations and FCS affinity measurements we can formulate a complete model of *Bt*PI-PLC specific binding to anionic PC-containing vesicles. Weak nonspecific electrostatics are dominated by lysine 44 which, upon mutation yields a loss of affinity resulting not only from the loss of nonspecific electrostatic interactions but also of short-range interactions with bilayer lipids. Neither the count of hydrogen bonds nor that of hydrophobic contacts between the protein and lipids are affected by lipid composition. Instead the balance between weak nonspecific electrostatics and opportunistic cation-π interactions ensures that the protein interacts preferentially with anionic vesicles containing large amounts of PC lipids.

Opportunistic choline-tyrosine cation-π interactions

Simulations of the enzyme at the interface of pure DMPG bilayers show a loose complex with a clear loss of short-range protein-lipid interactions, and in one of the MD replicas the protein dissociates from the bilayer after 200 ns. This is particularly meaningful for a simulation started with the protein anchored at the bilayer interface. On the other hand the 500 nanosecond-long MD simulations of *Bt*PI-PLC at the surface of mixed PC/PG bilayers result in a stably anchored protein. In agreement with previous experimental and computational studies, both helix B and the rim loop intercalate hydrophobic amino acids in the bilayers(6, 7, 10). However, neither hydrophobic contacts nor hydrogen bonds between the protein and lipids are dependent on the bilayer composition (Cf. Table 1 and Fig. 5A). The simulations show that the occupancy of the cation-π interactions between choline headgroups of DMPC and the tyrosines contributing the most to the affinity (Y88, Y204, Y246, Y251) vary as a function of DMPC content. Furthermore the occupancies of these interactions during the simulations correlate qualitatively well with the effect their mutation has on the binding affinities of WT (Cf. Supporting Information Fig. S2); the higher the measured effect on K_d the larger the occupancy of these interactions during the WT simulations. It is important to note that, unlike cation-π interactions observed in ligand-receptor binding (e.g. acetylcholine esterase (43) or histone tails (44)), *Bt*PI-PLC does not recruit choline groups to one or several PC-high-affinity binding sites. Rather these interactions are opportunistic and stochastically occur in the presence of PC lipids. This might also explain the apparent saturation of cation-π interaction between Y246, Y251 and Y204 at high X_{PC}. The opportunistic character of the choline-protein interactions serves the function of *Bt*PI-PLC well, as it is a virulence factor that recognizes the extracellular leaflets of eukaryotic membranes which have a high PC content.

While the ability of Charmm and other force fields to reproduce cation-π interactions between tryptophans and methylated lysines has been investigated (45), we are not aware of similar studies for interactions involving tyrosines. Even though our results cannot

replace a systematic benchmark they indicate that the force field provides a qualitatively satisfactory description of the tyrosine-choline interactions.

Weak non-specific electrostatic interactions

Seeing that the short-range interactions determining k_{off} do not explain *Bt*PI-PLC specificity for negatively charged vesicles, we turned to evaluating the interactions contributing to k_{on} using continuum electrostatic methods, a qualitative approach that has proven reliable for that purpose. As expected the calculated electrostatic contributions to *Bt*PI-PLC association with bilayers depends on the anionic lipid content and is most favorable at equal amounts of zwitterionic and anionic lipids. It is less favorable for increasing amounts of PG anionic lipids. This is in agreement with experimental data (42) and is due to the unfavorable interactions between the large negative regions of its surface potential and the anionic membrane. Despite an overall negative charge, *Bt*PI-PLC shows favorable electrostatic partitioning towards anionic membranes. In this respect, *Bt*PI-PLC is not unique as other negatively charged amphitropic proteins are known to partition to anionic membranes (29). By contrast, a striking characteristic of *Bt*PI-PLC is the small magnitude of the electrostatic free energy ΔG_{el} (~-0.25 kcal/mol, Fig. 7) compared to what has been calculated for other peripheral membrane proteins using the same method (27, 29). While the calculation of ΔG_{el} using continuum electrostatics calculations has obvious limitations due to the simplicity of the model, comparison of the results obtained for *Bt*PI-PLC to those obtained with the same approach for other proteins remains valid and informative. Furthermore the affinity of *Bt*PI-PLC for neutral pure PC vesicles (K_d=0.026 mM) is only approximately 4.1 times less favorable than that for vesicles with 20% anionic lipids (K_d=0.0064 mM) which is the PC:PG ratio yielding the tightest binding (9). This corresponds to an electrostatic contribution of less than one 1 kcal/mol. This, like the computational evaluation, is significantly lower than reported experimental values for amphitropic proteins (ca. 2 kcal/mol) (2). A marginally favorable ΔG_{el} means that the magnitude of the Coulombic contribution is only slightly higher than the magnitude of the desolvation penalty for the polar and charged residues (27). The low electrostatic interaction energy is explained by the absence of a well-defined cluster of basic amino acids at the *Bt*PI-PLC IBS (46, 47).

Unlike other basic amino acids K44 is an interfacial residue

Another striking feature of *Bt*PI-PLC is the large effect the mutation of one single lysine into an alanine has on K_d. For vesicles with X_{PC}=0.8, the mutation of K44 to alanine decreases the affinity about 55 times compared to WT or a contribution of 2.4 kcal/mol, which is clearly higher than the calculated electrostatics contribution (<1 kcal/mol). It is also higher than early reports of the contribution of basic amino acids to membrane binding. For example Kim and co-workers reported that the binding affinity of short basic peptides for vesicles containing acidic lipids was increased tenfold by the introduction of each additional lysine, independently of the identity of the anionic lipid (5). This represents a contribution to the binding free energy of about 1.4 kcal/mol. Similarly, membrane targeting of the protein kinase pp60src (Src) and mutation of two of the six basic amino acids to neutral asparagines increased the K_d by 100 while mutation of five of the six increased K_d another order of magnitude. The latter is similar to the binding affinity of wild-type Src for neutral PC vesicles showing that in this case mutation of the

basic cluster mainly affects the electrostatic interaction between Src and negatively charged vesicles (47). In addition, White and Wimley reported a free energy of ca. 1 kcal/mol for the transfer of lysines to the interfacial region of POPC large unilamellar vesicles (48).

Yet the large effect of the K44A mutation is in agreement with what is expected from computed transfer free energies of lysine and alanine side chains to the interface of a DOPC bilayer (49); -4.4 and -1.6 kcal/mol, respectively. The difference represents the cost of replacing the lysine by an alanine and amounts to 2.8 kcal/mol. The high contribution of Lys44 to membrane binding is thus due to its position in the headgroup region where it participates in hydrophobic interactions with the lipid tails and long-lasting hydrogen bonds with lipid phosphate groups. K44 is located close to the average plane of the phosphate groups in the different bilayers while other positively charged residues (i.e. K38, R71, K279) are further away from the IBS (Supp. Inf. Table S4). This also correlates with the moderate effects of the K38A, R71A and K279A mutations on *Bt*PI-PLC affinity (Fig. 9).

Looking at the computed free energy of transfer of a lysine side chain through POPC bilayers as a function of the distance from the center of the membrane is very informative. The slope is rather steep along the 5 to 10 Å that separate the bulk from the preferred position of the charged side chain with a difference between the two environments of about 2.5 kcal/mol (49-51). The contribution of basic amino acids at the IBS of peripheral membrane proteins will thus vary significantly depending on their position. The environment of these positive side chains and intramolecular interactions they might engage in are also expected to modulate their contributions. Furthermore our survey of the OPM database (30) shows that the distribution of surface basic amino acids in peripheral membrane proteins with respect to the membrane normal is rather broad (Fig. 10B) indicating a potentially broad range of contributions of those amino acids to the affinity of peripheral membrane proteins for biological membranes. Thus it is difficult to formulate a rule of thumb that would provide an average contribution per basic amino acid to peripheral membrane binding since the magnitude of such contributions is context-dependent.

Other peripheral membrane proteins display comparable distributions of basic amino acids

A survey of the peripheral proteins present in the OPM database allowed us to calculate and plot the net charges of membrane-binding proteins (Fig. 10A). The resulting distribution of charges is comparable to what has been observed for the entire proteome of *Saccharomyces cerevisiae* (52) and indicates that peripheral membrane binding proteins do not display a distribution of net charges skewed towards positive values. In fact, around one-third of the protein families classified as peripheral/monotopic in OPM have on average three or fewer basic amino acids within 10 Å of the membrane surface in their predicted membrane-bound state. It should be kept in mind that these charge estimates are based on simply counting basic and acidic residues, treating histidines as neutrally charged and considering neutral pH. Nonetheless, the results from analyzing the OPM database suggest that weak non-specific electrostatics is common among membrane-binding proteins.

CONCLUSION

Tyrosine-choline cation-π interactions are opportunistic and occur as the result of the presence of PC lipids in the membrane. We propose that they constitute a mechanism for the recognition of PC-rich eukaryotic cell membranes. The highest affinity of *Bt*PI-PLC towards slightly anionic SUV's is the result of weak electrostatic contributions from basic residues, particularly from a key basic residue. Our investigation shows that the energetic contributions of basic amino acids to peripheral membrane binding are dependent on their position with respect to the interfacial binding site and how deep they might anchor in the lipid bilayer. As a result, and given the distribution of lysines and arginines in known amphitropic proteins we suggest that weak nonspecific electrostatics might be more common than generally thought and should be considered as a means for proteins to respond quickly to environmental changes.

SUPPORTING MATERIAL

Detailed computational protocols, detailed material and methods for FCS measurements, area per lipids (Tables S1), enzyme structural variation along MD simulations (Table S2), protein-lipid interactions and depths of anchorage for each relevant amino acid of the WT (Tables S3 and S4) and K44A mutant (Table S7), characteristics of Tyr-Cho interactions (Tables S5 and S6), binding affinities and enzymatic activity (Table S8 and Figure S4), secondary structure content and T_m (Table S9) of WT and mutants *Bt*PI-PLC, flexibility of the β7-αG loop (Figure S1), comparison between Tyr-Cho occupancy data from MD and FCS affinity data on tyrosine mutants (Figure S2), PB electrostatic surface potential of bilayer-bound *Bt*PI-PLC (Figure S3).

ACKNOWLEDGMENTS

NOTUR (Norwegian metacenter for computational science) is thankfully acknowledged for provision of CPU time.

Funding Sources

N.R. and H.M.K.: Norwegian Research Council (FRIMEDBIO #214167). M.F.R. and A.G.: National Institutes of Health R01 GM60418.

ABBREVIATIONS

*Bt*PI-PLC, *Bacillus thurigiensis* phosphatidylinositol phospholipase C; DMPC, dimyristoylphosphatidylcholine; DMPG, dimyristoylphosphatidylglycerol; FCS, fluorescence correlation spectroscopy; MD, molecular dynamics; PC, phosphatidylcholine.

SUPPORTING CITATIONS

References (53-71) appear in the Supporting Material.

REFERENCES

1. Heimburg, T., and D. Marsh. 1996. Thermodynamics of the Interaction of Proteins with Lipid Membranes. In Biological membranes: a molecular perspective from computation and experiments. K. M. Merz, and B. Roux, editors. Birkhäuser Boston, Boston. 405-462.
2. Johnson, J. E., and R. B. Cornell. 1999. Amphitropic proteins: regulation by reversible membrane interactions. Molecular Membrane Biology 16:217-235.
3. Murray, D., A. Arbuzova, G. Hangyas-Mihalyne, A. Gambhir, N. Ben-Tal, B. Honig, and S. McLaughlin. 1999. Electrostatic properties of membranes containing acidic lipids and adsorbed basic peptides: theory and experiment. Biophys J 77:3176-3188.
4. Luckey, M. 2008. Membrane Structural Biology: With Biochemical and Biophysical Foundations. Cambridge University Press.
5. Kim, J. Y., M. Mosior, L. A. Chung, H. Wu, and S. Mclaughlin. 1991. Binding of Peptides with Basic Residues to Membranes Containing Acidic Phospholipids. Biophys J 60:135-148.
6. Guo, S., X. Zhang, B. A. Seaton, and M. F. Roberts. 2008. Role of helix B residues in interfacial activation of a bacterial phosphatidylinositol-specific phospholipase C. Biochemistry 47:4201-4210.
7. Grauffel, C., B. Yang, T. He, M. F. Roberts, A. Gershenson, and N. Reuter. 2013. Cation-pi interactions as lipid-specific anchors for phosphatidylinositol-specific phospholipase C. J Am Chem Soc 135:5740-5750.
8. Yang, B., M. Pu, H. M. Khan, L. Friedman, N. Reuter, M. F. Roberts, and A. Gershenson. 2015. Quantifying transient interactions between Bacillus phosphatidylinositol-specific phospholipase-C and phosphatidylcholine-rich vesicles. J Am Chem Soc 137:14-17.
9. Pu, M., M. F. Roberts, and A. Gershenson. 2009. Fluorescence correlation spectroscopy of phosphatidylinositol-specific phospholipase C monitors the interplay of substrate and activator lipid binding. Biochemistry 48:6835-6845.
10. Feng, J., H. Wehbi, and M. F. Roberts. 2002. Role of tryptophan residues in interfacial binding of phosphatidylinositol-specific phospholipase C. J Biol Chem 277:19867-19875.
11. Rheault, J. F., E. Gagne, M. Guertin, G. Lamoureux, M. Auger, and P. Lague. 2015. Molecular Model of Hemoglobin N from Mycobacterium tuberculosis Bound to Lipid Bilayers: A Combined Spectroscopic and Computational Study. Biochemistry 54:2073-2084.
12. Rogaski, B., and J. B. Klauda. 2012. Membrane-binding mechanism of a peripheral membrane protein through microsecond molecular dynamics simulations. J Mol Biol 423:847-861.
13. Weber, D. K., S. Yao, N. Rojko, G. Anderluh, T. P. Lybrand, M. T. Downton, J. Wagner, and F. Separovic. 2015. Characterization of the Lipid-Binding Site of Equinatoxin II by NMR and Molecular Dynamics Simulation. Biophys J 108:1987-1996.
14. Busse, R. A., A. Scacioc, R. Krick, A. Pérez-Lara, M. Thumm, and K. Kühnel. 2015. Characterization of PROPPIN-Phosphoinositide Binding and Role of Loop 6CD in PROPPIN-Membrane Binding. Biophys J 108:2223-2234.

15. Shi, X., C. Shao, X. Zhang, C. Zambonelli, A. G. Redfield, J. F. Head, B. A. Seaton, and M. F. Roberts. 2009. Modulation of Bacillus thuringiensis phosphatidylinositol-specific phospholipase C activity by mutations in the putative dimerization interface. J Biol Chem 284:15607-15618.
16. Shao, C., X. Shi, H. Wehbi, C. Zambonelli, J. F. Head, B. A. Seaton, and M. F. Roberts. 2007. Dimer structure of an interfacially impaired phosphatidylinositol-specific phospholipase C. J Biol Chem 282:9228-9235.
17. Humphrey, W., A. Dalke, and K. Schulten. 1996. VMD - Visual Molecular Dynamics. J Molec Graphics 14:33-38.
18. Kale, L., R. Skeel, M. Bhandarkar, R. Brunner, A. Gursoy, N. Krawetz, J. Phillips, A. Shinozaki, K. Varadarajan, and K. Schulten. 1999. NAMD2: Greater scalability for parallel molecular dynamics. J Comp Phys 151:283-312.
19. MacKerell, A. D., D. Bashford, M. Bellott, R. L. Dunbrack, J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiorkiewicz-Kuczera, D. Yin, and M. Karplus. 1998. All-atom empirical potential for molecular modeling and dynamics studies of proteins. J Phys Chem B 102:3586-3616.
20. Mackerell, A. D., M. Feig, and C. L. Brooks. 2004. Extending the treatment of backbone energetics in protein force fields: Limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. J Comput Chem 25:1400-1415.
21. Klauda, J. B., R. M. Venable, J. A. Freites, J. W. O'Connor, D. J. Tobias, C. Mondragon-Ramirez, I. Vorobyov, A. D. MacKerell, Jr., and R. W. Pastor. 2010. Update of the CHARMM all-atom additive force field for lipids: validation on six lipid types. J Phys Chem B 114:7830-7843.
22. Brooks, B. R., C. L. Brooks, 3rd, A. D. Mackerell, Jr., L. Nilsson, R. J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch, A. Caflisch, L. Caves, Q. Cui, A. R. Dinner, M. Feig, S. Fischer, J. Gao, M. Hodoscek, W. Im, K. Kuczera, T. Lazaridis, J. Ma, V. Ovchinnikov, E. Paci, R. W. Pastor, C. B. Post, J. Z. Pu, M. Schaefer, B. Tidor, R. M. Venable, H. L. Woodcock, X. Wu, W. Yang, D. M. York, and M. Karplus. 2009. CHARMM: the biomolecular simulation program. J Comput Chem 30:1545-1614.
23. Minoux, H., and C. Chipot. 1999. Cation-pi interactions in proteins: Can simple models provide an accurate description? J Am Chem Soc 121:10366-10372.
24. Baker, N. A., D. Sept, S. Joseph, M. J. Holst, and J. A. McCammon. 2001. Electrostatics of nanosystems: Application to microtubules and the ribosome. Proc Natl Acad Sci USA 98:10037-10041.
25. Sharp, K. A., and B. Honig. 1990. Calculating total electrostatic energies with the nonlinear Poisson-Boltzmann equation. J Phys Chem 94:7684-7692.
26. Gallagher, K., and K. A. Sharp. 1998. Electrostatic Contributions to Heat Capacity Changes of DNA-Ligand Binding. Biophys J 75:769-776.
27. Mulgrew-Nesbitt, A., K. Diraviyam, J. Wang, S. Singh, P. Murray, Z. Li, L. Rogers, N. Mirkovic, and D. Murray. 2006. The role of electrostatics in protein-membrane interactions. Biochim Biophys Acta 1761:812-826.

28. Singh, S. M., and D. Murray. 2003. Molecular modeling of the membrane targeting of phospholipase C pleckstrin homology domains. Protein Sci 12:1934-1953.
29. Murray, D., S. McLaughlin, and B. Honig. 2001. The role of electrostatic interactions in the regulation of the membrane association of G protein beta gamma heterodimers. J Biol Chem 276:45153-45159.
30. Lomize, M. A., A. L. Lomize, I. D. Pogozheva, and H. I. Mosberg. 2006. OPM: orientations of proteins in membranes database. Bioinformatics 22:623-625.
31. Hinsen, K. 2000. The Molecular Modeling Toolkit : a new approach to molecular simulations. J. Comput. Chem. 21:79-85.
32. Miller, S., J. Janin, A. M. Lesk, and C. Chothia. 1987. Interior and surface of monomeric proteins. J Mol Biol 196:641-656.
33. Lomize, A. L., I. D. Pogozheva, and H. I. Mosberg. 2011. Anisotropic solvent model of the lipid bilayer. 2. Energetics of insertion of small molecules, peptides, and proteins in membranes. J Chem Inf Model 51:930-946.
34. Gasteiger, E., C. Hoogland, A. Gattiker, S. Duvaud, M. R. Wilkins, R. D. Appel, and A. Bairoch. 2005. Protein identification and analysis tools on the ExPASY server. In The Proteomics Protocols Handbook. J. M. Walker, editor. Humana Press, New York. 571-607.
35. Cheng, J., S. Karri, C. Grauffel, F. Wang, N. Reuter, M. F. Roberts, P. L. Wintrode, and A. Gershenson. 2013. Does changing the predicted dynamics of a phospholipase C alter activity and membrane binding? Biophys J 104:185-195.
36. Pu, M., A. Orr, A. G. Redfield, and M. F. Roberts. 2010. Defining specific lipid binding sites for a peripheral membrane protein in situ using subtesla field-cycling NMR. J Biol Chem 285:26916-26922.
37. Pu, M., X. Fang, A. G. Redfield, A. Gershenson, and M. F. Roberts. 2009. Correlation of vesicle binding and phospholipid dynamics with phospholipase C activity: insights into phosphatidylcholine activation and surface dilution inhibition. J Biol Chem 284:16099-16107.
38. Schillinger, A. S., C. Grauffel, H. M. Khan, O. Halskau, and N. Reuter. 2014. Two homologous neutrophil serine proteases bind to POPC vesicles with different affinities: When aromatic amino acids matter. Biochim Biophys Acta 1838:3191-3202.
39. Schrodinger, LLC. 2010. The PyMOL Molecular Graphics System, Version 1.3r1.
40. Ben-Tal, N., B. Honig, R. M. Peitzsch, G. Denisov, and S. McLaughlin. 1996. Binding of small basic peptides to membranes containing acidic lipids: theoretical models and experimental results. Biophys J 71:561-575.
41. Murray, D., and B. Honig. 2002. Electrostatic control of the membrane targeting of C2 domains. Mol Cell 9:145-154.
42. Wehbi, H., J. Feng, J. Kolbeck, B. Ananthanarayanan, W. Cho, and M. F. Roberts. 2003. Investigating the interfacial binding of bacterial phosphatidylinositol-specific phospholipase C. Biochemistry 42:9374-9382.
43. Branduardi, D., F. L. Gervasio, A. Cavalli, M. Recanatini, and M. Parrinello. 2005. The role of the peripheral anionic site and cation-pi interactions in the ligand penetration of the human AChE gorge. J Am Chem Soc 127:9147-9155.

44. Lu, Z., J. Lai, and Y. Zhang. 2009. Importance of charge independent effects in readout of the trimethyllysine mark by HP1 chromodomain. J Am Chem Soc 131:14928-14931.
45. Zheng, X., C. Wu, J. W. Ponder, and G. R. Marshall. 2012. Molecular dynamics of beta-hairpin models of epigenetic recognition motifs. J Am Chem Soc 134:15970-15978.
46. McLaughlin, S., and A. Aderem. 1995. The myristoyl-electrostatic switch: a modulator of reversible protein-membrane interactions. Trends Biochem Sci 20:272-276.
47. Sigal, C. T., W. Zhou, C. A. Buser, S. McLaughlin, and M. D. Resh. 1994. Amino-terminal basic residues of Src mediate membrane binding through electrostatic interaction with acidic phospholipids. Proc Natl Acad Sci USA 91:12253-12257.
48. Wimley, W. C., and S. H. White. 1996. Experimentally determined hydrophobicity scale for proteins at membrane interfaces. Nat Struct Biol 3:842-848.
49. MacCallum, J. L., W. F. Bennett, and D. P. Tieleman. 2007. Partitioning of amino acid side chains into lipid bilayers: results from computer simulations and comparison to experiment. J Gen Physiol 129:371-377.
50. Bonhenry, D., M. Tarek, and F. Dehez. 2013. Effects of Phospholipid Composition on the Transfer of a Small Cationic Peptide Across a Model Biological Membrane. J Chem Theory Comput 9:5675-5684.
51. MacCallum, J. L., W. F. Bennett, and D. P. Tieleman. 2008. Distribution of amino acids in a lipid bilayer from computer simulations. Biophys J 94:3393-3404.
52. Ke, R., and S. Mitaku. 2005. Local repulsion in protein structures as revealed by a charge distribution analysis of all amino acid sequences from the Saccharomyces cerevisiae genome. J Phys Condens Matter 17:S2825.
53. Jo, S., T. Kim, V. G. Iyer, and W. Im. 2008. CHARMM-GUI: A web-based graphical user interface for CHARMM. J. Comput. Chem. 29:1859-1865.
54. Jorgensen, W. L., J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein. 1983. J. Chem. Phys. 79:926-935.
55. Feller, S. E., Y. H. Zhang, R. W. Pastor, and B. R. Brooks. 1995. Constant-Pressure Molecular-Dynamics Simulation - the Langevin Piston Method. J Chem Phys 103:4613-4621.
56. Essmann, U., L. Perera, M. L. Berkowitz, T. Darden, H. Lee, and L. G. Pedersen. 1995. A smooth particle mesh Ewald method. J Chem Phys 103:8577-8593.
57. Izaguirre, J. A., S. Reich, and R. D. Skeel. 1999. Longer time steps for molecular dynamics. J Chem Phys 110:9853-9864.
58. Andersen, H. C. 1983. Rattle - a Velocity Version of the Shake Algorithm for Molecular-Dynamics Calculations. J Comput Phys 52:24-34.
59. Venable, R. M., Y. Luo, K. Gawrisch, B. Roux, and R. W. Pastor. 2013. Simulations of Anionic Lipid Membranes: Development of Interaction-Specific Ion Parameters and Validation Using NMR Data. Journal of Physical Chemistry B 117:10183-10192.

60. Giorgino, T. 2014. Computing 1-D atomic densities in macromolecular simulations: The density profile tool for VMD. Comput Phys Commun 185:317-322.
61. Gilson, M. K., and B. H. Honig. 1987. Calculation of electrostatic potentials in an enzyme active site. Nature 330:84-86.
62. Middleton, E. R., and E. Rhoades. 2010. Effects of Curvature and Composition on alpha-Synuclein Binding to Lipid Vesicles. Biophys J 99:2279-2288.
63. Rusu, L., A. Gambhir, S. McLaughlin, and J. Radler. 2004. Fluorescence correlation spectroscopy studies of peptide and protein binding to phospholipid vesicles. Biophys J 87:1044-1053.
64. Thompson, N. L. 1991. Fluorescence correlation spectroscopy. In Topics in Fluorescence Spectroscopy. J. R. Lakowicz, editor. Plenum Press, New York. 337-378.
65. Magde, D., E. L. Elson, and W. W. Webb. 1974. Fluorescence Correlation Spectroscopy .2. Experimental Realization. Biopolymers 13:29-61.
66. Elson, E. L., and D. Magde. 1974. Fluorescence Correlation Spectroscopy .1. Conceptual Basis and Theory. Biopolymers 13:1-27.
67. Grauffel, C., B. Yang, T. He, M. F. Roberts, A. Gershenson, and N. Reuter. 2013. Cation−π Interactions As Lipid-Specific Anchors for Phosphatidylinositol-Specific Phospholipase C. J Am Chem Soc 135:5740-5750.
68. Murray, D., A. Arbuzova, G. Hangyás-Mihályné, A. Gambhir, N. Ben-Tal, B. Honig, and S. McLaughlin. 1999. Electrostatic Properties of Membranes Containing Acidic Lipids and Adsorbed Basic Peptides: Theory and Experiment. Biophysical Journal 77:3176-3188.
69. Mulgrew-Nesbitt, A., K. Diraviyam, J. Wang, S. Singh, P. Murray, Z. Li, L. Rogers, N. Mirkovic, and D. Murray. 2006. The role of electrostatics in protein–membrane interactions. Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids 1761:812-826.
70. Murray, D., A. Arbuzova, B. Honig, and S. McLaughlin. 2002. The role of electrostatic and nonpolar interactions in the association of peripheral proteins with membranes. In Current Topics in Membranes. T. J. M. Sidney A. Simon, editor. Academic Press. 277-307.
71. Lumb, C. N., and M. S. P. Sansom. 2012. Finding a Needle in a Haystack: The Role of Electrostatics in Target Lipid Recognition by PH Domains. PLoS Computational Biology 8:e1002617.

Supporting Material for: A role for weak electrostatic interactions in peripheral membrane protein binding

H. M. Khan[1,2], T. He[3], E. Fuglebakk[1,2], C. Grauffel[1,2], B. Yang[3,4], M. F. Roberts[3], A. Gershenson[4], N. Reuter[1,2*]

[1] Department of Molecular Biology, University of Bergen, Norway;
[2] Computational Biology Unit, Department of Informatics, Bergen, Norway;
[3] Department of Chemistry, Boston College, Chestnut Hill, U.S.A.;
[4] Department of Biochemistry and Molecular Biology, University of Massachusetts, Amherst, U.S.A.
[*] Corresponding Author- NR: Universitetet i Bergen, Computational Biology Unit, Institutt for informatikk, Pb. 7803 N-5020 Bergen, nathalie.reuter@uib.no.

Methods.

Bilayer preparation for MD simulations

Bilayers of DMPC, DMPG and DMPC:DMPG lipid mixtures were built using the CHARMM-GUI(1). We considered four bilayers: pure DMPC (X_{PC} = 1.0), an 80:20 (X_{PC} = 0.8) mixed bilayer, a 50:50 (X_{PC} = 0.5) mixed bilayer, and pure DMPG (X_{PC} = 0). Here X_{PC} represents the mole fraction of PC lipids in the mixture. All bilayers contain 256 lipids (128 in each leaflet). TIP3P(2, 3) water molecules were added to hydrate the lipids and sodium was used as a counter-ion to achieve overall charge neutrality of the system when necessary. The CHARMM36 force field, with updates for lipids(4), and NAMD (v2.9)(5) were used for the simulations. The bilayers were first minimized using 4000 steps of a conjugate gradient algorithm. The system was then equilibrated at 310 K, and 1 atm in the NPT ensemble for 400 ps with constant area in the x-y dimension and 2 fs time steps. The temperature was controlled with Langevin dynamics with the temperature damping coefficient set to 1.0 and the velocities periodically reassigned every 1 ps. The pressure was controlled using the Langevin piston method(6) with an oscillation period of 75 fs and a damping time scale of 25 fs. We used anisotropic pressure coupling. The non-bonded interactions cut-off was set to 12 Å and long-range electrostatics corrections beyond the cut-off were modeled using the Particle Mesh Ewald(7) method. We used switching functions for both the electrostatics and van der Waals interactions with the switch distance set to 11 Å. We used the r-RESPA(8) multiple time step algorithm and short-range non-bonded forces were evaluated every 2 fs while long-range electrostatics were evaluated every 4 fs. SHAKE(9) was used to constraint all bonds between hydrogen atoms and heavy atoms. The systems were then further equilibrated for 200 ps after removing the constant area constraint. The bilayers were further simulated for 100 ns in the NPT ensemble without velocity reassignment. The bilayer with a 50:50 PC:PG ratio (X_{PC} = 0.5) was simulated for 200 ns to achieve proper lipid mixing. The pressure control parameters for the Langevin piston method were altered to 200 fs for the oscillation period and 50 fs for the damping time scale. All other simulation parameters remained unchanged. The values of the area per lipid for the bilayers are given in Table S1.

Note that we did not use a correction for the Lennard Jones potential between Na^+ and lipid oxygens(10)for several reasons; our simulations start with the protein inserted in the bilayer, we do not expect the ions to play a significant role in the protein-bilayer interaction as indicated by the low electrostatics partitioning (Cf. Results section, continuum electrostatics calculations), and to our knowledge the correction has not been benchmarked on systems containing surface-bound proteins or peptides. We did prepare a DMPG bilayer using the Lennard-Jones correction though, and then ran test simulations with *Bt*PI-PLC using the correction. In these conditions the binding of *Bt*PI-PLC to the bilayer was unstable as we also observe in the simulations without the NBFIX correction.

MD trajectory analysis

All analyses were performed using CHARMM (v33b1)(11) and VMD (v1.9.1)(12) on the last 450 ns of each simulation. Hydrophobic contacts, hydrogen bonds and cation-π interactions are averaged over the two replicas. Hydrophobic interactions between the

protein and lipid bilayer were assigned if the protein and membrane candidate atoms were within 3 Å of each other for at least 10 ps. For the protein, candidate atoms for hydrophobic contacts are side chain atoms from aliphatic groups, while for the lipid bilayer; the candidate atoms are from the hydrophobic lipid tails. Hydrogen bonds were assigned using the donor and acceptor definitions from the CHARMM(3) force field with a distance cut-off of 2.4 Å and an angle cut-off of 130 degrees. The two cut-off criteria had to be met at least for 10 ps. Cation-π interactions between the aromatic amino acids (Tyr, Phe and Trp) and the choline group of the DMPC lipids were assigned if the nitrogen atom of the choline group was within 7 Å of the aromatic rings. Additionally these distances should not vary more than 1.5 Å (13, 14). Combining these two criteria, we can screen for the possible formation of cation-π adducts during the course of simulation.

Occupancy of a particular interaction is reported as the number of conformations with the interaction present divided by the total number of conformations in the trajectory sampling window.

Electron density profiles (EDP) were generated using the Density Profile Tool(15) VMD plugin. EDPs were calculated from the trajectory at 1 ns intervals. The projection axis is the z-axis, which is also the bilayer normal. The profile resolution is $\Delta z = 1$ Å for the protein, lipid bilayer and water. The profile resolution is $\Delta z = 0.5$ Å for amino acids, membrane binding components of protein, choline and phosphate planes of the membrane.

Continuum electrostatics calculations

Poisson-Boltzmann calculations are carried out using APBS (version 1.3) (16). The parameters for protein and ions are adapted from the CHARMM all atom force field (c22 including CMAP correction)(17) and for the lipids from the force field update for lipids (Charmm36) (4). The non-linear Poisson-Boltzmann equation is solved using the focusing technique (18). Coarse grid dimensions are expanded by a factor of 2.5 as both *Bt*PI-PLC and anionic membranes are highly charged. Fine grid dimensions are obtained by adding 50 Å to the molecular dimensions. We used fine grid spacing of 0.5 and 0.35 Å and the difference in the calculated electrostatic free energy was less than 0.003 kcal/mol; a fine mesh spacing of 0.35 Å was thus chosen for all the calculations. The solute dielectric is set to 2 and the solvent (water) dielectric is set to 80. KCl is used to investigate the effect salt concentration on the electrostatics.

The electrostatic free energy for the protein $G_{el}(P)$, membrane $G_{el}(M)$ and the protein membrane complex $G_{el}(P.M)$ is calculated. The electrostatic free energy contribution to the free energy of binding due to protein membrane interactions can then be evaluated as follows:

$$\Delta G_{el} = G_{el}(P.M) - [G_{el}(P) + G_{el}(M)] \qquad \text{(S1)}$$

Database statistics

Analysis of net charges For the selected families, we obtained average net charges, $\hat{c}_F$, for each family, *F*, as:

$$\hat{c}_F = \frac{1}{|F|}\sum_{p\in F} c_p \qquad \text{(S2)}$$

where $|F|$ is the number of proteins in the family *F*, and c_p is the net charge of a protein obtained as:

$$c_p = K_p + R_p - E_p - D_p \qquad \text{(S3)}$$

where K_p, R_p, E_p, and D_p respectively denote the number of lysines, arginines, glutamates and aspartates in the protein.

Analysis of basic residues For a family F, we calculated the mean number of basic amino acids at the IBS as:

$$w_F^{\text{basic}} = \frac{1}{|F|}\sum_{p\in F} n_p^{\text{basic}}(0, -10\text{Å}) \qquad \text{(S4)}$$

where 'basic' denote amino acids of type Lys or Arg. Please note that the plot on Figure 10 has been truncated to exclude one family with $w_F^{\text{basic}} > 40$.

We also calculated the mean surface density of basic amino acids at varying distances from the membrane. We define this statistic for windows of size *l* around different insertion coordinates. For a family *F*, the mean surface density of basic amino acids is:

$$d_F^{\text{basic}}(i, l) = \frac{1}{|F|}\sum_{p\in F} \frac{n_p^{\text{basic}}\left(i-\frac{l}{2}, i+\frac{l}{2}\right)}{n_p^{\text{any}}\left(i-\frac{l}{2}, i+\frac{l}{2}\right)} \qquad \text{(S5)}$$

where 'any' denote amino acids of any type. We report averages over the entire dataset:

$$d_S^{\text{basic}} = \frac{1}{|S|}\sum_{F\in S} d_F^{\text{basic}}(\text{i}, 2\text{Å}) \qquad \text{(S6)}$$

where $|S|$ denotes the number of families in the set.

FCS measurements of BtPI-PLC binding to SUVs

FCS experiments were carried out at 22 ^{o}C on 300 µL samples in PBS, pH 7.4, plus 1 mg/ml BSA to stabilize PI-PLC, in chambered coverglass wells (LabTek), coated with 10 mg/ml BSA and rinsed with PBS prior to use to prevent protein adhesion to the sides of the wells. 10 nM labeled PI-PLC was titrated with SUVs, and the fraction of protein bound to vesicles was determined from two species fits to the autocorrelations (obtained

in crosscorrelation mode), $G(\tau) = A_p g_p(\tau) + A_v g_v(\tau)$, where p and v denote free protein and SUVs that are fluorescent due to PI-PLC binding, respectively, and A_j is the amplitude of species j (19-22). The correlation function for species j, $g_j(\tau)$, accounts for diffusion of the molecules through the observation volume which depends on the radius and extent of the observation volume, determined from fits to rhodamine 110 calibration data using $D = 280\ \mu m^2 s^{-1}$ at 22 ^oC(23), and D_j, the diffusion coefficient for each species(22, 24). D_p for the free protein was experimentally determined in the absence of vesicles while D_v for the SUVs was determined from global fits to all of the titration experiments for a particular X_{PC} using Origin (OriginLab). The apparent fraction of protein bound to the SUVs, *f*, can be determined from A_p and the time-averaged number of proteins in the observation volume in the absence of vesicles <No>, by $f = 1-A_p\langle N_o\rangle = 1-A_p/A_{p,o}$, where $A_{p,o}=1/\langle N_o\rangle$ is the autocorrelation amplitude for free PI-PLC prior to titration and corrected for volume changes. The apparent dissociation constant, K_d, representing PI-PLC partitioning onto the vesicle, and a cooperativity coefficient, *n*, were determined from fits to the equation $f = f_{max}[PL]^n/(K_d^n+[PL]^n)$, where *f*, is determined for different total lipid concentrations, [PL], at fixed X_{PC}, and f_{max} is the apparent maximum fraction bound. FCS experiments were repeated twice using different vesicle and protein preparations. Assuming that uncertainties are normally distributed and considering the range of standard deviation values from all of the FCS measurements, a 2-3 times increase in apparent K_d relative to N168C PI-PLC (WT*) is considered significant.

Table S1. Area per lipid for different bilayers obtained from MD simulations.

X_{PC}	Area per lipid (Å^2)
1.0	60.9±0.9
0.8	59.6±0.9
0.5	58.5±1.0
0.0	59.1±1.0

Table S2. Average RMSD of the *Bt*PI-PLC backbone along the MD simulations relative to the starting protein structure (minimized X-ray structure, in the presence of the bilayer).

X_{PC}	replicate	RMSD (Å)
1.0	r1	1.44±0.11
	r2	1.34±0.12
0.8	r1	1.43±0.10
	r2	1.17±0.11
0.5	r1	1.43±0.13
	r2	1.61±0.09

Table S3. Inventory of interactions between *Bt*PI-PLC and the bilayers[a].

SSE[b]	aa	Hydrophobic [c]			Hbonds (%) [d]			Cation-π (%) [e]		
		X_{PC}			X_{PC}			X_{PC}		
		0.5	0.8	1	0.5	0.8	1	0.5	0.8	1
αB	Q40	1.8	2.5	2.4	64.4	99.2	92.0			
	N41	1.1	1.5	1.4	62.5	96.8	90.5			
	P42	4.9	5.8	5.6						
	I43	8.6	7.8	7.8						
	K44	3.3	1.4	2.3	85.9	98.3	98.5			
	Q45				19.7	27.2	33.3			
	V46	3.6	4.5	4.2						
	W47	2.9	2.2	2.6				23.4	12.5	3.3
	Y53							1.6	8.1	1.9
β2	R71				49.4	50.5	38.7			
β2-αD	P84	4.9	2.6	2.7						
	L85	3.4	1.2	2.3						
	Y86							9.9	15.6	11.9
	Y88				42.2	58.0	61.6	33.8	55.6	93.1
β3-αE	Y118							1.6	7.8	12.9
	K122				67.8	56.2	56.9			
β6-αF	Y200							49.8	12.1	43.0
	K201				53.7	46.5	45.0			
αF	Y204							18.0	25.5	22.3
β7-αG	S236				38.5	33.4	32.8			
	G238	0.8	0.9	1.1						
	G239	0.9	1.9	2.3						
	T240	2.7	2.5	2.6						
	A241	3.2	3.1	3.2						
	W242	2.9	2.8	2.9						
	S244				43.0	82.0	68.4			
αG	Y246				20.1	12.9	34.6	50.1	84.5	80.9
	Y247				29.2	36.0	43.7	11.1	11.9	12.7
	Y251				35.6	34.0	38.8	31.7	47.4	47.0

[a] DMPG simulations (X_{PC}=0) are not included as (1) they do not show tight binding of the protein, and (2) we want to follow the dependence of the interactions on the PC:PG ratio (some interactions such as cation-π obviously disappear in the absence of PC). The results presented in the table above are averages over two replicas on the last 450ns.. [b] Secondary structure elements; α: helix, β: strand, βi-αX: loop between strand i and helix X. [c] Average number of hydrophobic contacts per frame. [d] Occupancies of hydrogen bonds in %. Only the hydrogen bonds observed in the two replicas at the same X_{PC} are shown. Hydrogen bonds reported are mainly between the lipid phosphate groups and amino acid side chains but a few form with the protein backbone (**bold numbers**). Hydrogen bonds with headgroups or glycerol groups are not stable throughout the simulations. [e] Occupancies of cation-π adducts.

Table S4. Simulations averages of the anchoring depth of *Bt*PI-PLC residues in the bilayer. Positive values indicate that the centers of mass of the amino acids are on average buried below the phosphate group of the lipids during the simulation. The average phosphate plane is used as the reference plane.

SSE	aa	X_{PC} = 1.0		X_{PC} = 0.8		X_{PC} = 0.5		X_{PC} = 0
		r1	r2	r1	r2	r1	r2	r1
β1-αB	K38	-6.8±2.2	-7.6±2.3	-8.1±2.0	-6.7±2.5	-6.8±3.1	-9.3±3.1	-8.4±3.0
αB	Q40	0.6±2.3	-0.2±2.4	-1.1±2.1	-0.4±2.5	0.5±3.3	-2.8±3.1	-2.0±3.1
	N41	3.3±1.8	2.8±2.0	2.3±1.7	3.4±2.3	3.1±2.4	0.7±2.4	1.3±2.6
	P42	4.2±1.9	3.7±1.9	3.0±1.8	4.0±2.2	3.8±2.5	1.5±2.3	1.7±2.7
	I43	4.3±1.7	3.9±1.8	3.7±1.6	4.3±2.1	4.2±1.9	2.5±1.8	2.3±2.4
	K44	0.4±1.7	-0.2±1.8	-0.3±1.6	0.5±2.2	0.4±1.9	-1.1±1.8	-1.3±2.3
	Q45	-0.6±1.9	-1.0±1.9	-1.8±1.8	-0.8±2.1	-1.0±2.4	-3.3±2.0	-3.1±2.7
	V46	0.4±1.9	0.3±1.8	-0.3±1.7	0.3±2.0	0.3±1.9	-2.1±1.9	-1.8±2.7
	W47	-0.8±1.9	-0.7±2.2	-0.7±1.9	-0.8±2.1	0.1±1.7	-2.3±1.8	-2.6±2.6
β2	R71	-10.1±2.0	-10.9±2.4	-10.0±2.3	-9.1±3.1	-9.6±1.8	-10.2±2.3	-10.3±2.5
β2-αD	P84	-3.2±1.9	-4.5±2.3	-2.5±2.0	-2.2±2.7	-1.3±1.6	-3.8±1.7	-4.0±2.4
	L85	-3.6±1.6	-4.2±2.0	-3.7±1.7	-2.9±2.4	-2.9±1.7	-4.8±1.7	-4.7±2.3
	Y86	-5.5±1.9	-7.0±2.2	-6.1±1.9	-5.0±2.8	-2.0±2.1	-6.8±2.5	-6.7±2.4
	Y88	-4.9±2.0	-5.8±2.3	-5.9±2.0	-4.4±2.6	-6.3±2.4	-6.4±3.0	-6.2±2.7
β3-αE	Y118	-8.7±2.1	-8.1±2.6	-8.2±2.3	-7.4±3.1	-8.3±1.8	-7.0±2.2	-7.9±2.7
	K122	-6.1±2.9	-5.3±3.3	-5.2±3.3	-5.1±4.2	-5.0±2.5	-5.3±3.0	-5.8±3.2
β6-αF	Y200	-10.9±2.2	-10.8±2.5	-10.4±1.8	-10.7±2.5	-10.5±2.2	-11.8±3.3	-12.2±4.2
	K201	-7.6±2.6	-7.5±2.8	-7.2±2.1	-7.4±2.8	-7.0±2.6	-8.7±4.0	-9.7±4.4
αF	Y204	-7.3±2.6	-7.5±2.5	-8.3±2.7	-8.6±2.5	-8.7±2.8	-11.6±4.1	-12.6±4.0
β7-αG	S236	-6.6±2.0	-6.5±2.2	-6.1±1.8	-6.5±2.2	-6.4±2.0	-7.8±3.1	-8.9±3.6
	S237	-3.0±1.9	-3.0±2.1	-2.8±1.8	-2.9±2.1	-2.7±1.9	-5.5±3.9	-5.6±3.8
	G238	0.1±2.0	0.2±2.2	0.4±1.8	0.0±2.3	0.4±2.1	-2.5±3.6	-3.8±4.2
	G239	2.2±1.9	2.1±2.0	2.2±1.7	2.2±2.1	2.4±2.0	0.3±3.1	-1.4±4.2
	T240	3.3±2.0	3.3±2.0	2.9±1.9	3.0±2.0	3.1±2.2	1.2±3.1	-0.8±3.5
	A241	2.6±2.2	2.5±1.9	1.7±2.1	2.0±2.1	2.0±2.5	0.4±2.6	0.1±3.3
	W242	3.9±2.4	4.1±2.1	3.3±2.2	3.3±2.2	3.3±2.4	1.3±3.3	3.1±3.2
	N243	-0.3±2.3	0.0±2.1	-0.4±2.0	-0.6±2.2	-0.5±2.3	-2.5±3.4	-2.0±3.4
	S244	-2.3±2.1	-2.3±1.9	-2.9±1.9	-2.7±2.0	-2.7±2.2	-4.4±2.6	-4.4±3.2
αG	Y246	-4.5±2.4	-4.3±2.1	-6.0±2.2	-5.2±2.2	-5.4±2.5	-8.2±2.4	-8.8±3.1
	Y247	-2.3±2.5	-2.0±2.2	-3.4±2.4	-3.2±2.2	-3.2±2.5	-5.5±2.9	-6.1±3.6
	S250	-7.2±2.7	-6.9±2.4	-8.7±2.5	-8.1±2.4	-8.3±2.7	-11.3±2.9	-11.7±3.4
	Y251	-6.7±3.2	-5.8±2.5	-7.6±2.7	-7.4±2.6	-7.2±2.9	-10.3±3.7	-10.9±4.0
β8-αH	K279	-8.1±3.1	-8.1±3.0	-10.1±2.9	-9.3±3.1	-9.5±3.6	-12.6±2.9	-12.7±3.5

Table S5. Number of DMPC lipids mediating one or more cation-π interactions with *Bt*PI-PLC tyrosines (mixed bilayer, X_{PC} = 0.5).

Residue	Replica 1	Replica 2
Y251	4	5
Y246	2	2
Y88	1	2

Table S6. Occupancies of Y251 cation-π interactions involving more than one PC lipid simultaneously. The occupancies (%) are reported for the three different PC bilayers simulated. Reported replicates are chosen based on the presence of cation-π interactions between Y251 and more than one lipid.

X_{PC} & replica	Occupancies (%)
X_{PC} = 1 r2	11.5
X_{PC} = 0.8 r1	11.1
X_{PC} = 0.5 r1	4.8

Table S7. Anchorage depth and inventory of interactions for *Bt*PI-PLC K44A and a mixed bilayer (X_{PC}-0.8).

SSE	aa	depth (Å)	Hydrophobic	Hbonds (%)[a]	Cation-π (%)
αB	Q40	-1.4±1.9	0.9	**7.2**[b]/34.6	
	N41	1.8±1.7	0.5	**7.2**/17.0/9.8[c]	
	P42	2.9±1.6	5.2		
	I43	3.4±1.6	8.2		
	K44A	-0.9±1.6	1.6		
	Q45	-1.9±1.6	0.3	18.5	
	V46	-0.3±1.6	3.9		
	W47	-0.2±1.6	4.1		
β2	R69	-12.3±1.7		80.6	
	R71	-9.4±2.2		73.4	
β2-αD	H82	-7.4±1.7	1.1	6.6	
	P84	-2.2±2.0	6.3		
	L85	-3.3±1.7	5.0		
	Y86	-6.4±2.5		**22.9**/9.3	14.7
	Y88	-6.6±2.2		63.7	38.2
β3-αE	Y118	-7.4±2.3		35.1	36.0
	K122	-8.0±4.0		**7.9**/26.3	
β6-αF	Y200	-9.8±2.1			93.1
	K201	-6.3±2.4		63.8	
αF	N203	-6.5±3.0		12.4	
	Y204	-8.4±2.9		6.9	32.4
β7-αG	S236	-6.7±1.8	0.3	22.9	
	S237	-3.1±1.8			
	G238	0.1±1.9	1.4		
	G239	1.9±1.8	0.5		
	T240	2.6±2.0	2.5		
	A241	1.2±2.1	3.2		
	W242	2.8±2.3	2.7	12.6/11.2[d]	
	N243	-0.8±2.2		16.3	
	S244	-3.4±1.9		79.5	
αG	Y246	-6.3±2.1		8.1	85.7
	Y247	-3.8±2.4		43.5	13.3
	Y248	-6.5±2.2			
	S250	-9.3±2.4			
	Y251	-8.1±2.8		34.2	50.3

[a] Hbonds occur mainly between sidechains and phosphate groups; [b] Hbond between backbone and phosphate (bold numbers) [c] Hbond between sidechain and headgroup; [d] Hbond between sidechain and glycerol

Table S8. Binding affinities of *Bt*PI-PLC variants for PC/PG SUVs.

			K_d (mM)		
*Bt*PI-PLC [a]	X_{PC} = 0	0.2	0.5	0.8	1.0
WT	12±1	0.31±0.01	0.016±0.001	0.0064±0.0008	0.026±0.005
K38A	32±1	2.2±0.2	0.30±0.05	0.037±0.002	0.024±0.06
K44A	22±4	8.8 [b]	1.0±0.3	0.35 [b]	0.51±0.07
R71A	43±5	1.8±0.2	0.13±0.05	0.030±0.008	0.042±0.009
K279A	31±7	0.87±0.04	0.092±0.011	0.012±0.002	0.034±0.008

[a] All of the variants of *Bt*PI-PLC contain the N168C mutation for fluorescent labeling
[b] Data for K44A binding at X_{PC}=0.2 and 0.8 was interpolated from data between X_{PC} = 0.1 and 0.3 and X_{PC} = 0.7 and 0.9, respectively.

The actual K44A K_d values obtained were:

X_{PC}	Kd (mM)
0	22.2±3.9
0.1	14.7±2.5
0.3	5.43±2.13
0.5	1.03±0.28
0.7	0.41±0.06
0.9	0.33±0.03
1.0	0.51±0.07

Table S9. Comparison of secondary structure content and T_m values for WT and mutant *Bt*PI-PLCs from far UV circular dichroism data.

Protein	% Secondary Structure				T_m (^oC)[a]
	α-Helix	*β*-Sheet	*β*-Turn	Random Coil	
WT	20.6	32.2	17.2	30.0	57.1
K38A	20.2	32.5	17.1	30.1	57.3
K44A	21.6	30.9	17.4	30.1	60.3
R71A	21.4	30.9	17.2	30.3	56.8
K279A	20.8	31.8	17.2	30.1	56.0

[a] Loss of secondary structure was assessed by monitoring the ellipticity at 222 nm as a function of temperature

Figure S1. The β7-αG loop is more flexible in simulations without PC lipids. The RMSD of the loop relative to **its initial conformation** is plotted across the simulations for *Bt*PI-PLC interacting with lipid bilayers with different lipid compositions. On a pure DMPG bilayer (X_{PC} = 0) the loop looses intramolecular interactions resulting in more flexibility.

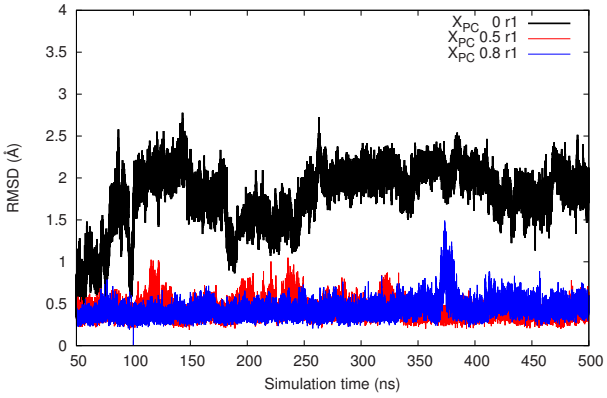
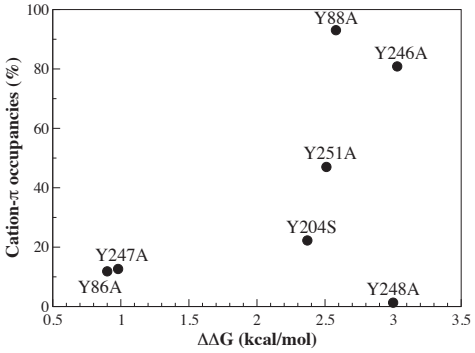
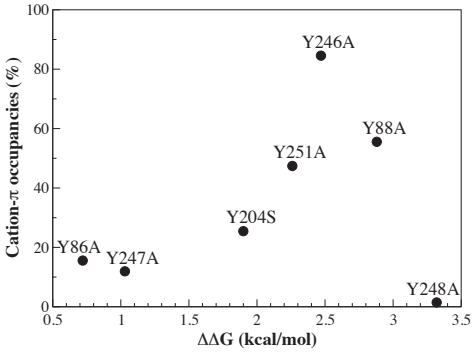


Figure S2. Comparison between occupancies of cation-π interactions observed during the MD simulations and evaluation of ΔΔG calculated from apparent K_d values measured by FCS for *Bt*PI-PLC WT and single tyrosine mutant at X_{PC}=1.0 (A), 0.8 (B) and 0.5 (C). K_d value are from *Grauffel et al.*(25). See same reference for an explanation of the role of Y248 and the effect of its mutation to an alanine.

A



B



C

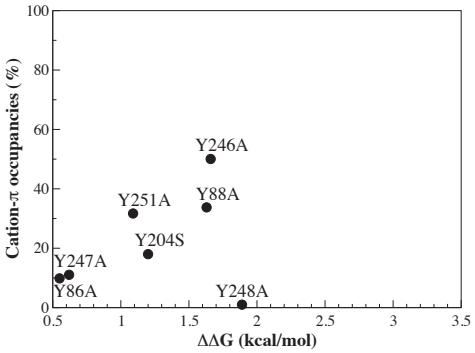

Cation-π occupancies (%)
ΔΔG (kcal/mol)
Y246A
Y88A
Y251A
Y204S
Y247A
Y86A
Y248A

Figure S3. Calculated electrostatic potential the of *Bt*PI-PLC and membrane (X_{PC}=0.8) complex when *Bt*PI-PLC is 4 Å away from the membrane. Isocontours of -/+ 1 (unit KT/e) are mapped on the molecular surface (negative: red; positive: blue). (A) Side view, in an equilibrated bilayer isosurfaces are not essentially flat. (This contrasts with other published views where isosurfaces for model membranes are represented as flat(26-28)) (B) Top view, local zones of blue and red alternate in the membrane. This pattern is also observed for a zwitterionic membrane (i.e. POPC/DMPC) (29).

A **B**

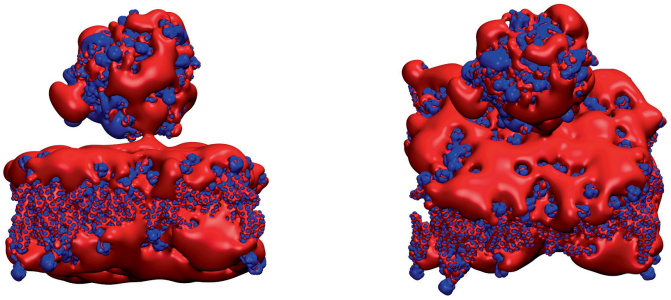
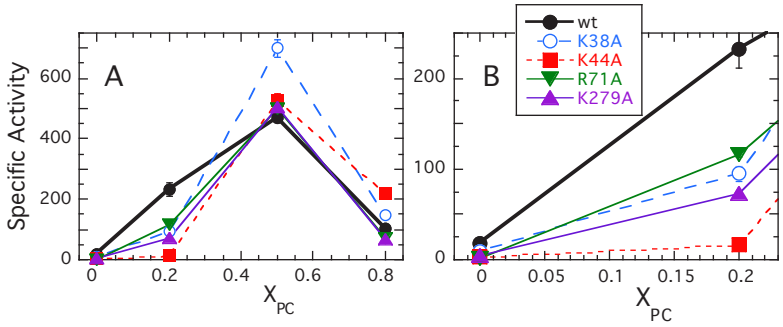


Figure S4. Specific activity of *Bt*PI-PLC variants towards PI/POPC SUVs with 2 mM PI and varying X_{PC}: WT (solid circle), K38A (open circle), K44A (solid square), R71A (inverted triangle), and K279A (triangle). Protein concentrations were adjusted to between 0.15-2 g/ml to ensure less than 20% cleavage of the PI. (A) All cationic residue mutants exhibit enzymatic activity similar to wildtype (WT) for $X_{PC} \geq 0.5$ implying that the active site is not significantly perturbed in these alanine mutants. (B) A zoomed in view of the activity for low X_{PC}. i.e., more anionic SUVs showing the extremely low activity of K44A.



SUPPORTING REFERENCES

1. Jo, S., T. Kim, V. G. Iyer, and W. Im. 2008. CHARMM-GUI: A web-based graphical user interface for CHARMM. J Comput Chem 29:1859-1865.
2. Jorgensen, W. L., J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein. 1983. J. Chem. Phys. 79:926-935.
3. MacKerell, A. D., D. Bashford, M. Bellott, R. L. Dunbrack, J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiorkiewicz-Kuczera, D. Yin, and M. Karplus. 1998. All-atom empirical potential for molecular modeling and dynamics studies of proteins. J Phys Chem B 102:3586-3616.
4. Klauda, J. B., R. M. Venable, J. A. Freites, J. W. O'Connor, D. J. Tobias, C. Mondragon-Ramirez, I. Vorobyov, A. D. MacKerell, Jr., and R. W. Pastor. 2010. Update of the CHARMM all-atom additive force field for lipids: validation on six lipid types. J Phys Chem B 114:7830-7843.
5. Kale, L., R. Skeel, M. Bhandarkar, R. Brunner, A. Gursoy, N. Krawetz, J. Phillips, A. Shinozaki, K. Varadarajan, and K. Schulten. 1999. NAMD2: Greater scalability for parallel molecular dynamics. J Comp Phys 151:283-312.

6. Feller, S. E., Y. H. Zhang, R. W. Pastor, and B. R. Brooks. 1995. Constant-Pressure Molecular-Dynamics Simulation - the Langevin Piston Method. J Chem Phys 103:4613-4621.
7. Essmann, U., L. Perera, M. L. Berkowitz, T. Darden, H. Lee, and L. G. Pedersen. 1995. A smooth particle mesh Ewald method. J Chem Phys 103:8577-8593.
8. Izaguirre, J. A., S. Reich, and R. D. Skeel. 1999. Longer time steps for molecular dynamics. J Chem Phys 110:9853-9864.
9. Andersen, H. C. 1983. Rattle - a Velocity Version of the Shake Algorithm for Molecular-Dynamics Calculations. J Comput Phys 52:24-34.
10. Venable, R. M., Y. Luo, K. Gawrisch, B. Roux, and R. W. Pastor. 2013. Simulations of Anionic Lipid Membranes: Development of Interaction-Specific Ion Parameters and Validation Using NMR Data. Journal of Physical Chemistry B 117:10183-10192.
11. Brooks, B. R., C. L. Brooks, 3rd, A. D. Mackerell, Jr., L. Nilsson, R. J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch, A. Caflisch, L. Caves, Q. Cui, A. R. Dinner, M. Feig, S. Fischer, J. Gao, M. Hodoscek, W. Im, K. Kuczera, T. Lazaridis, J. Ma, V. Ovchinnikov, E. Paci, R. W. Pastor, C. B. Post, J. Z. Pu, M. Schaefer, B. Tidor, R. M. Venable, H. L. Woodcock, X. Wu, W. Yang, D. M. York, and M. Karplus. 2009. CHARMM: the biomolecular simulation program. J Comput Chem 30:1545-1614.
12. Humphrey, W., A. Dalke, and K. Schulten. 1996. VMD - Visual Molecular Dynamics. J Molec Graphics 14:33-38.
13. Grauffel, C., B. Yang, T. He, M. F. Roberts, A. Gershenson, and N. Reuter. 2013. Cation-pi interactions as lipid-specific anchors for phosphatidylinositol-specific phospholipase C. J Am Chem Soc 135:5740-5750.
14. Minoux, H., and C. Chipot. 1999. Cation-pi interactions in proteins: Can simple models provide an accurate description? J Am Chem Soc 121:10366-10372.
15. Giorgino, T. 2014. Computing 1-D atomic densities in macromolecular simulations: The density profile tool for VMD. Comput Phys Commun 185:317-322.
16. Baker, N. A., D. Sept, S. Joseph, M. J. Holst, and J. A. McCammon. 2001. Electrostatics of nanosystems: Application to microtubules and the ribosome. Proc Natl Acad Sci USA 98:10037-10041.
17. Mackerell, A. D., M. Feig, and C. L. Brooks. 2004. Extending the treatment of backbone energetics in protein force fields: Limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. J Comput Chem 25:1400-1415.
18. Gilson, M. K., and B. H. Honig. 1987. Calculation of electrostatic potentials in an enzyme active site. Nature 330:84-86.
19. Cheng, J., S. Karri, C. Grauffel, F. Wang, N. Reuter, M. F. Roberts, P. L. Wintrode, and A. Gershenson. 2013. Does changing the predicted dynamics of a phospholipase C alter activity and membrane binding? Biophys J 104:185-195.
20. Middleton, E. R., and E. Rhoades. 2010. Effects of Curvature and Composition on alpha-Synuclein Binding to Lipid Vesicles. Biophys J 99:2279-2288.

21. Rusu, L., A. Gambhir, S. McLaughlin, and J. Radler. 2004. Fluorescence correlation spectroscopy studies of peptide and protein binding to phospholipid vesicles. Biophys J 87:1044-1053.
22. Thompson, N. L. 1991. Fluorescence correlation spectroscopy. In Topics in Fluorescence Spectroscopy. J. R. Lakowicz, editor. Plenum Press, New York. 337-378.
23. Magde, D., E. L. Elson, and W. W. Webb. 1974. Fluorescence Correlation Spectroscopy .2. Experimental Realization. Biopolymers 13:29-61.
24. Elson, E. L., and D. Magde. 1974. Fluorescence Correlation Spectroscopy .1. Conceptual Basis and Theory. Biopolymers 13:1-27.
25. Grauffel, C., B. Yang, T. He, M. F. Roberts, A. Gershenson, and N. Reuter. 2013. Cation−π Interactions As Lipid-Specific Anchors for Phosphatidylinositol-Specific Phospholipase C. J Am Chem Soc 135:5740-5750.
26. Murray, D., A. Arbuzova, G. Hangyás-Mihályné, A. Gambhir, N. Ben-Tal, B. Honig, and S. McLaughlin. 1999. Electrostatic Properties of Membranes Containing Acidic Lipids and Adsorbed Basic Peptides: Theory and Experiment. Biophysical Journal 77:3176-3188.
27. Mulgrew-Nesbitt, A., K. Diraviyam, J. Wang, S. Singh, P. Murray, Z. Li, L. Rogers, N. Mirkovic, and D. Murray. 2006. The role of electrostatics in protein–membrane interactions. Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids 1761:812-826.
28. Murray, D., A. Arbuzova, B. Honig, and S. McLaughlin. 2002. The role of electrostatic and nonpolar interactions in the association of peripheral proteins with membranes. In Current Topics in Membranes. T. J. M. Sidney A. Simon, editor. Academic Press. 277-307.
29. Lumb, C. N., and M. S. P. Sansom. 2012. Finding a Needle in a Haystack: The Role of Electrostatics in Target Lipid Recognition by PH Domains. PLoS Computational Biology 8:e1002617.