© 2022 Melanie P. Muller

SIMULATING ATOMIC-LEVEL INTERACTIONS IN THE COAGULATION CASCADE

BY

MELANIE P. MULLER

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Biophysics and Computational Biology
in the Graduate College of the
University of Illinois Urbana-Champaign, 2022

Urbana, Illinois

Doctoral Committee:

Professor Emad Tajkhorshid, Chair
Professor James H. Morrissey
Associate Professor Diwakar Shukla
Professor Chad Rienstra, University of Wisconsin-Madison

ABSTRACT

The reactions that take place within the coagulation cascade are essential for formation of clots. They are critical for wound healing when activated properly, and when improperly activated can result in devastating health consequences. The biochemical reactions that make up the cascade have been extensively mapped experimentally, but the detailed mechanisms of coagulation modulation and protease activation have remained elusive. This is partly due to the atomic-level nature of the highly specific interaction between coagulation proteins and membrane phospholipids which allow spontaneous binding from plasma for a number of coagulation proteins at the initiation of clotting. Several important coagulation proteins are also highly flexible, which has made experimental characterization of their complete structures at atomic resolution prohibitively difficult.

Computational methodologies provide a unique means of investigating atomic-level interactions of the blood coagulation cascade that are out of reach with experimental methods. Here, we present our findings probing protein-lipid interactions and complex formation of blood coagulation proteins using computational means. Molecular dynamics simulations have been used to model spontaneous membrane binding and protein-lipid interactions of the membrane binding domains of factor X, factor VII, and factor IX at an atomic level. Building on these simulations, we developed an atomic-level, membrane-bound model of the ternary extrinsic complex of blood coagulation using a novel computational methodology combining nonequilibrium molecular dynamics, specialized membrane bilayer representations, and protein-protein docking. Finally, we will present our atomic-level model of tissue factor:factor VIIa bound to XK1, a tissue factor protease inhibitor and factor X hybrid molecule, and discuss potential for future approaches combining experimental and computational methodologies to probe interactions in the coagulation cascade.

ACKNOWLEDGMENTS

I'd like to first thank my advisor, Professor Emad Tajkhorshid, for his untiring support throughout the process of completing this dissertation. Support of my fellow lab group members past and present has also been vital, and I'd like to particularly thank Dr. Tao Jiang, Dr. Po-chao Wen, Dr. Josh Vermaas, Dr. Christopher Mayne, and Dr. Mark Arcario for their help and encouragement over the years. I'd like to thank Professor James H. Morrissey and his lab members Dr. Yan Wang, Dr. Fabienne Birkle and Dr. Divyani Paul for their many insightful discussions with me, without which this work would not have been possible. I'd like to thank my family members Lisa, Douglas, Andrew, William, and Opal Muller for their encouragement and understanding, particularly during these challenging pandemic years.

TABLE OF CONTENTS

1 INTRODUCTION 1

2 METHODS 7
2.1 Molecular dynamics simulations 7
2.2 Membrane modeling with the HMMM 15
2.3 Nonequilibrium, or biased, simulations 18
2.4 Refinement of experimental structures 20
2.5 Rigid-body protein-protein docking 20

3 LIPID SPECIFICITY OF COAGULATION FACTOR X 22
3.1 Introduction 22
3.2 Methods 25
3.3 Spontaneous membrane insertion of the FX-GLA domain . . . 31
3.4 Comparison of FX-GLA to other GLA domains 40
3.5 Conclusions 43

4 LIPID SPECIFICITY OF COAGULATION FACTOR VII 45
4.1 Introduction 45
4.2 Methods 48
4.3 Binding of FVII-GLA to PS and PA-containing lipid bilayers . 51
4.4 FVII-GLA lipid interactions in PA:PC and PS:PC membranes 54
4.5 Conclusions 59

5 MODELING FORMATION OF THE EXTRINSIC COMPLEX . . 61
5.1 Introduction 61
5.2 Methods 66
5.3 Overview 80
5.4 Individual equilibrium simulations of TF:FVIIa and FX 81
5.5 Protein-protein rigid-body docking and colvar design 82
5.6 Results of biased simulations 84
5.7 Complex characteristics following equilibration 96
5.8 Conclusions 104

6 ADDITIONAL SIMULATIONS OF COAGULATION PROTEIN INTERACTIONS 105
6.1 Introduction 105
6.2 Methods 106

6.3 Spontaneous membrane insertion of FIX-GLA 111
6.4 Modeling the TF:FVIIa:XK1 complex 116

7 REFERENCES 121

1 INTRODUCTION

Lipid-protein interactions are a crucial facet of regulatory methods used to control the biological processes of the cell [1, 2]. Cellular membranes contain a dizzyingly complex array of components with a high variability depending on location.[1, 3]. Charged phospholipids, in particular, are asymmetrically distributed over the bilayer and concentrated in the inner leaflet of the cell. Active regulation by membrane proteins is required to maintain these asymmetric conditions on either side of the barrier, as well as to make the dynamic changes required as the cell and organism face changing environmental situations [4, 5, 6, 7, 8].

Changes in the concentrations of specific lipids allows initiation of biochemical processes which have direct implications for broader human health and disease [9, 10, 11]. One such process is the initiation of thrombosis through the clotting cascade.

The plasma membrane is of central importance to the proper activation of the clotting cascade, as most of the reactions involved cannot occur without coagulation factors first binding to and finally forming macromolecular complexes on the membrane surface [12, 13]. Figure 1.1 illustrates steps in the extrinsic and common pathways of coagulation which are dependent on such complexes. Following vascular damage, coagulation factors are drawn to the membrane by exposure of anionic phospholipids [10, 13, 14]. In vitro experiments have demonstrated that removal from the membrane surface makes coagulation factors thousands of times less active [12, 13]. Further, coagulation factor behavior is highly influenced by the local lipid environment, with many factors showing highly differential membrane affinity and activity in the presence of different anionic phospholipids [15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26]. Major events in coagulation, such as macromolecular complex formation and zymogen activation, are influenced by this lipid dependence [12, 27, 28]. Several reactions in the cascade follow the same model: a) binding of enzyme to membrane, b) complex formation with cofactor, c)

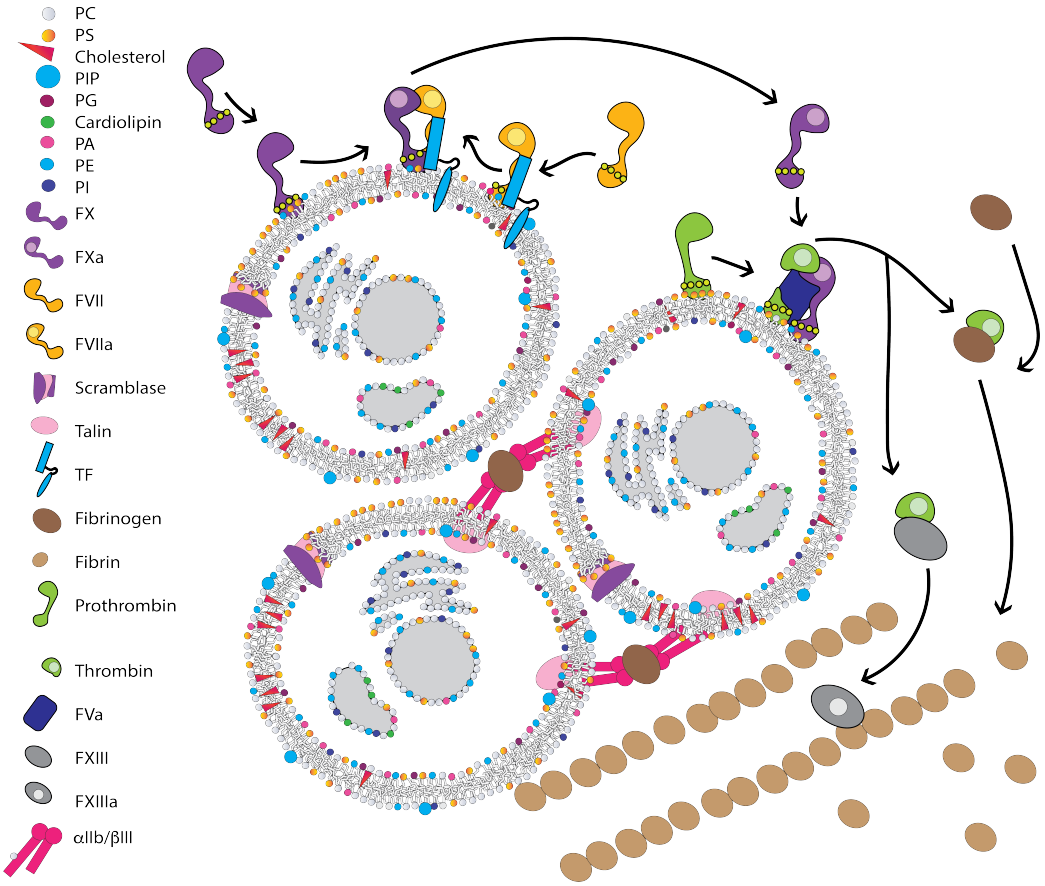


Figure 1.1: Illustration of the extrinsic and common portions of the coagulation cascade. The name of each protein shown is indicated in the key on the left side of the image. Arrows are used to indicate proteins binding the membrane, forming complexes, and completing reactions in the cascade.

ternary complex formation with zymogen substrate, d) catalytic activation of substrate (zymogen) to active enzyme (Figure 1.1) [29, 19, 30, 31]. In many cases, the membrane binding affinities of coagulation enzyme and zymogen are influenced differently by anionic phospholipids, and each protein may be specific to a different anionic lipid [23, 22]. Defects in regulation of the clotting cascade directly impact human health. Thrombotic disorders represent a national and international concern of great magnitude, as thrombosis is a common disease pathway in ischemic heart disease, stroke, and venous thromboembolism [32]. The conditions caused by unwanted clot formation remain among the leading causes of mortality and morbidity, resulting in approximately 10 million cases of venous thromboembolism worldwide and $\sim$2 million venous thromboembolisms per year in the United States alone [32, 33]. While there are effective antithrombotic treatments available currently, such as the standard drugs Warfarin and Heparin, side effects can be severe, the most serious being uncontrolled bleeding [34, 35]. The newer generation of antithrombotic drugs also carry risk of serious complications [34, 35]. Lack of clotting due to hemophilia and other clotting factor deficiencies also present a complex ongoing treatment challenge [36, 37, 38, 39]. More readily controlled treatments with increased speed and specificity in inducing and reducing clotting are needed, requiring significant improvements in understanding of key interactions in the coagulation cascade.

There are two primary areas which must be elucidated if control of the clotting cascade is to be understood: membrane binding and complex formation. Tissue factor (TF) triggers thrombosis in many, if not most, disease settings [40, 41], and understanding the extrinsic complex [TF:factor VIIa (FVIIa) : factor X (FX)] is therefore of particular interest for thrombosis treatment development. Activation of FX through the formation of the extrinsic complex is the initial step in thrombosis [42, 41, 43], and the primary way in which the explosive burst of thrombin needed for clot formation is generated [13]. Both the lipid-protein interactions which modulate binding of FX and FVII and the protein-protein interactions in the extrinsic complex which activate FX remain mysterious, however.

FX and FVII are attracted to the membrane from solution after anionic lipids are exposed to the outer cell membrane following vascular damage, and bind using specialize membrane binding domains known as GLA domains [10]. These coagulation factor GLA domains are the key element of

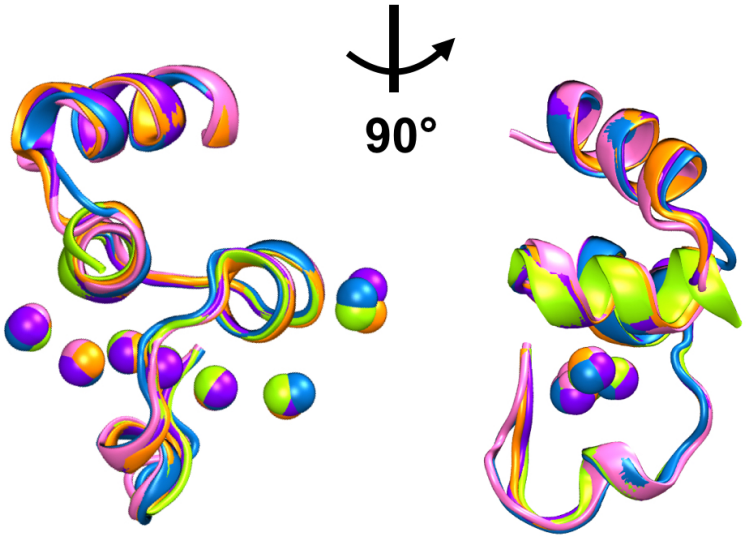


Figure 1.2: GLA domains of coagulation proteins demonstrate a high degree of structrual conservation: structural alignment shown for the GLA domains of five coagulation proteins. The protein backbone is shown for each domain using a cartoon representation, Ca^{2+} ions shown as balls. Bovine factor X (PDB 1IOD [56]) is shown in purple, factor IX (PDB 1J34) [57]) is shown in pink, bovine prothrombin (PDB 1NL2 [52]) is shown in blue, protein C (PDB 1LQV [58]) is shown in light green, and factor VII (1DAN [59]) is shown in orange.

the protein generates lipid specificity. Five structurally aligned coagulation protein GLA domains are shown Figure 1.2.

GLA domains are membrane-binding domains present in coagulation factors known as vitamin K dependent proteins [13, 44, 45]. Vitamin K is required in the reaction that post-translationally generates the specialized γ-carboxyglutamic acid (GLA) residues necessary for proper folding and activity of GLA domains [46]. Each GLA residue carries a -2 charge, and most of them coordinate Ca^{2+} ions as a part of proper folding to expose the hydrophobic ω-loop [10, 47, 48, 49, 50, 51, 52]. Upon GLA domain membrane binding, the set of residues termed the keel penetrate into the hydrophobic core of the membrane [53, 54]. While structure and sequence among GLA domains is highly conserved [45], GLA domains from different coagulation factors show a marked variation in lipid specificity [13, 23, 22, 55]. FX and factor IX (FIX) will only show optimal activation in a bilayer with a critical

concentration of phosphatidylserine (PS) [22, 15]. FVII, by contrast, shows minimal protease activation in the presence of PS and, like the anticoagulant protein C, demonstrates strong activation in the presence of phosphatidic acid (PA) [23]. These highly specific atomic-level lipid interactions are key to control of coagulation factors by the membrane, yet they remain poorly understood.

Similarly, there is very little atomic-level information available regarding macromolecular structures of ternary coagulation complexes. Despite a critical need for understanding biochemistry of coagulation proteins, no crystal structure has been produced for any ternary coagulation complex. Indeed, very few crystal structures of coagulation complexes have ever been developed [59, 27]; efforts have been thwarted by the membrane dependence of complex formation [27]. While models have been developed using protein-protein docking, with the addition of minimal molecular dynamics refinement in one case [60, 61], the role of the membrane has been universally ignored despite the critical involvement of lipid interactions in modulation of complex formation. The contemporary paucity of atomic-level structural information on the extrinsic complex and relevant lipid interactions is due to the inherent difficulty of studying such large complexes with flexible components, particularly in the presence of the membrane [27]. While X-ray crystallography [62], nuclear magnetic resonance (NMR) [63] and electron cryo-microscopy (cryo-EM) [64, 65, 66, 67] can yield atomic details of protein-lipid interactions, and cryo-EM in particular shows potential for yielding high-quality structures of membrane proteins, thus far structures of clotting complexes have proved elusive.

Molecular dynamics (MD) simulation allows for characterization of the structural and dynamic aspects of lipid-protein interactions critical to membrane protein function [68, 69]. Atomistic MD simulations allow for temporal resolutions of as high as 1 fs and spatial resolutions of sub-angstrom level. The method has been successfully employed for the identification of novel drug target sites in other systems [70, 71, 72, 73], and novel techniques offer the ability to overcome traditional difficulties in membrane modeling [74, 75, 76, 77]. While challenges remain, particularly in the ability to obtain sufficient sampling of processes in the cellular environment, simulations have already allowed us to characterize critical lipid-protein interactions in membrane environments [68]. Nonequilibrium and biasing methods such

as steered molecular dynamics (SMD) [78, 79] have been developed to allow for more extensive sampling of slow events. Computational methods such as MDFF can be used in concert with cryo-EM or X-ray crystallography to fit atomic-level structures to lower-resolution experimental data [80, 81, 82].

In the following chapters, results of simulation studies characterizing lipid-protein and protein-protein interactions of coagulation proteins will be described. After an introduction to the background of the techniques used, three main studies will be discussed. First, MD simulations of the GLA domain of coagulation factor X, and their use in describing potential phosphatidylserine specific interactions, will be described. Next, we discuss use of MD simulations to compare interactions of the coagulation factor VII GLA domain with two lipid species with which it has differential lipid specificity. We describe use of novel computational methodologies in modeling the extrinsic coagulation complex, in particular to characterize putative interactions of the potential exosite. Finally, we will discuss characterization of lipid-protein interactions for the coagulation factor IX GLA domain, as well as development of an atomic-level model of TF:FVIIa in complex with a hybrid protein in which a TFPI inhibitor domain is combined with the lightchain of factor X.

2 METHODS*

2.1 Molecular dynamics simulations

In the following chapters, molecular dynamics simulations were employed to study a number of biomolecular phenomena, including membrane binding of peripheral proteins and complex formation. In each case, we employ classical molecular dynamics (MD) simulations, which were originally developed in the late 1950s [83]. As its foundation, classical MD utilizes Newton's equations of motion to describe the dynamics of a set of predefined interaction sites. In biomolecular simulation, these interaction sites may represent either individual atoms or a distinct group of atoms (such as an amino acid residue). The instantaneous forces applied to each interaction site are computed over time, allowing for their trajectory within the simulation system to be determined [85, 86]. A variety of molecular representations, including atomistic (all-atom and united-atom), coarse-grained (CG) and multiscale hybrid models, have been employed to investigate lipid-protein interactions. All of these methods use the basic approximation of representing molecules as a set of interaction sites with the number of atoms in each interaction site

*Parts of this chapter have been adapted with permission from the following articles: (1) M.P. Muller, T. Jiang, C. Sun, M. Lihan, S. Pant, P. Mahinthichaichan, A. Trifan, and E. Tajkhorshid, "Characterization of lipid-protein interactions and lipid-mediated modulation of membrane protein function through molecular simulations," *Chem. Rev.*, vol. 119, pp. 6086–6161, 2019. Copyright 2019 American Chemical Society. (2) J. L. Baylon, J. V. Vermaas, M. P. Muller, M. J. Arcario, T. V. Pogorelov, and E. Tajkhorshid, "Atomic-level description of protein-lipid interactions using an accelerated membrane model," *Biochim. Biophys. Acta, Biomembr.*, vol. 1858, pp. 1573-1583, 2016. (3) J. V. Vermaas, J. L. Baylon, M. J. Arcario, M. P. Muller, Z. Wu, T. V. Pogorelov, and E. Tajkhorshid, "Efficient exploration of membrane associated phenomena at atomic resolution," J. Membr. Biol., vol. 248, no. 3, pp. 563–582, 2015. Copyright 2015 Elsivier Inc. (4) J. V. Vermaas, N. Trebesch, C. Mayne, S. Thangapandian, M. Shekhar, P. Mahinthichaichan, J. L. Baylon, T. Jiang, Y. Wang, M. P. Muller, E. Shinn, Z. Zhao, P. Wen, and E. Tajkhorshid, Microscopic characterization of membrane transporter function by *in Silico* modeling and simulation," in *Computational Approaches for Studying Enzyme Mechanism Part B*, G. A. Voth, Ed. Academic Press, 2016, vol. 578, ch. 16, pp. 373-428.

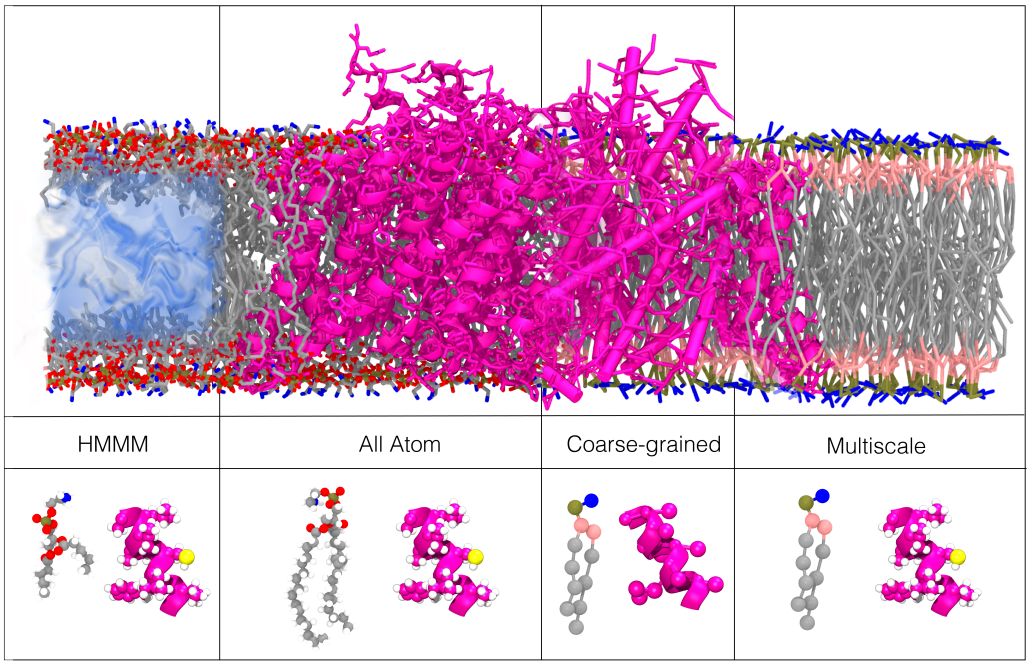


Figure 2.1: Examples of common resolutions/representations used in the investigation of lipid-protein interactions. The upper panels illustrate different representations for a membrane-embedded ClC channel (PDB:1OTS). The lower panels illustrate each of the representations using a single phosphatidylserine (PS) lipid and a short alpha helix (ClC channel, PDB:1OTS). All-atom (AA) simulations use one interaction site per atom. In coarse-grained (CG) simulations, several atoms are grouped into one interaction site. Multiscale simulations use a combination of resolutions either in the same simulation or in sequence. The HMMM (highly mobile membrane mimetic) model uses truncated lipids and a membrane core of inorganic solvent to increase lateral diffusion of membrane lipids. Reprinted with permission from reference [84]. Copyright 2019 American Chemical Society.

defining the resolution of the method. An inherent trade-off exists in choosing simulation method; methods with high resolution provide a higher degree of accuracy in predicting detailed interactions, while lower-resolution methods allow sampling of longer timescale phenomena. All-atom (AA) models, for example, employ one interaction site per atom and thereby provide high resolution information on atomic-level interactions. They are limited, however, to microsecond (μs) timescale and to system sizes of tens of nanometers. As the lipid-protein and protein-protein interactions investigated in this work are atomic-level interactions, simulations were carried out using all-atom models. NAMD, a highly parallelized MD program [86] was used for all simulations.

Membrane simulation poses a set of challenges distinct to those of biomolecules in aqueous solution. The earliest simulations of membranes were perfomed in the 1980s, using highly simplified bilayers and special methods to avoid need for explicit solvent [87, 88, 89] Advances in methodology later allowed for more detailed and realistic bilayers, additions of proteins, and larger system sizes, although simulation time was sill restrained to 100s of picoseconds [90, 91, 92, 93]. Characterization of lipid-protein interactions was still hampered by the short timescales accessible to atomistic simulations. In the intervening period, advances in simulation methodology have allowed for simulations in the 100s of nanoseconds to the microsecond timescale. In addition, advances in force-field parametrization have allowed for more accurate description of bilayer behavior over longer timescales. These simulation times are still relatively short to allow for capturing of lipid-protein interactions, given that protein-lipid interactions generally take place over the millisecond timescale. In this work a specialized method, the highly mobile membrane mimetic (HMMM) was employed to allow for enhance lipid diffusion. The remainder of this section will discuss will discus further details of force-field parametrization, simulation conditions, and the HMMM. An empasis will be placed on simulation aspects which are specifically required for membrane simulation.

2.1.1 Force fields for atomic-level simulation

A perennial area of interest for the development of MD simulations are the force fields used to set characteristics of the molecules being simulated, such

as bonds, angles, dihedrals, and partial charges. Parameters are derived by first using quantum mechanical (QM) calculations, and then fitting parameters with molecular mechanics (MM) [94, 95, 96, 97] to be consistent with the QM and experimental results. Molecular mechanics, fundamentally, describes describes forces between interaction sites as the sum of potential energies. These include many terms which are molecule or atom specific, such as bond-length:

$$
\begin{aligned}
U_{total} &= U_{bonded} + U_{non-bonded} + U_{extra} \\
U_{bonded} &= U_{bonds} + U_{angles} + U_{dihedrals} + U_{impropers} \\
&= \sum_{i \in bonds} k_i \left(b_i - b_0\right)^2 + \sum_{j \in angles} k_j \left(a_j - a_0\right)^2 \\
&+ \sum_{k \in dihedralterms} k_k \left(1 + \cos\left(n_k \chi_k + \delta_k\right)\right) + \sum_{l \in impropers} k_l \left(\chi_l - \chi_0\right)^2 \\
U_{non-bonded} &= U_{VDW} + U_{electrostatic} \\
&= \sum_{i,j \in pairlist} \epsilon_{ij} \left(\left(\frac{R_{ij}^{min}}{r_{ij}} \right)^{12} - 2 \left(\frac{R_{ij}^{min}}{r_{ij}} \right)^{6} \right) + \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}}
\end{aligned}
\tag{2.1}
$$

The quality of the force field determines the quality of the simulation results. Development of AA force fields for proteins can be traced to the 1980s [98, 95, 97]. Force fields developed during this period which reproduced important characteristics of proteins and are still used for modern simulations, include CHARMM [99, 100, 101], AMBER [102, 103], GROMOS [104], and OPLS [105, 106]. Simulations in this work were performed with the most recent CHARMM force field, CHARMM36.

Inclusion of lipids into CHARMM and other AA force fields proved initially challenging, due to both the liquid-crystalline characteristics of lipid bilayers and the paucity of high-resolution structural information for lipids [95]. A number of simulations using CHARMM22 for lipids [107] reproduced experimental results, including a membrane channel within a bilayer [92]. However, the initial MD simulations for lipids in the CHARMM22 force field were 100 ps in length, and a number of limitations became evident upon extending the timescale to 800 ps [108]. The surface tension of the CHARMM22 bilayer was too high, causing the surface area to collapse and the bilayer to become

gel-like unless constant area was imposed during the simulation or a surface tension designed to yield the experimental surface area applied. In addition, order parameters of lipid atoms near the water-bilayer interface were in error for CHARMM22 bilayer simulations [108].

CHARMM27 [109] improved on CHARMM22 lipids through refinement of the Lennard-Jones and torsional parameters of the alkane moiety, as well as the torsional and partial charge parameters of the phosphate moiety [108]. It was used for a range of benchmark studies with simulation times reaching 100 ns [108]. CHARMM27 was also found to describe lipid-protein interactions [108]. At longer timescales, however, a systematic overestimation of the chain order parameters became evident. CHARM27r [110] resolved this problem, but issues with C2 order parameters were still evident. It was still not advisable to carry out simulations using the NPT ensamble. CHARMM36 [111, 112] yielded the correct head-group surface area and correct chain order parameters using NPT conditions. In addition, a 2016 study which compared CHARMM36, Berger, and Slipids in studying a microbial peptide found that membranes with Berger lipids were prone to pore formation, an effect not found in the newer CHARMM and Slipids [113]

AMBER force field lipid parameters, currently a major competitor of CHARMM, have also been improved in recent years. The Lipid11 modular lipid force field for AMBER [114] was developed using the General AMBER Force Field (GAFF [115]). It is designed to function more in the manner of a protein force field; rather than developing parameters for entire lipids, parameters were designed for head groups and tails which could then be combined [114]. Use of an additional surface tension term was required for Lipid11 in order to prevent a phase transition during simulations. The Lipid14 updated modular force field no longer required this additional term [116, 117]. Lipid head-group and tail charges were modified, as well as Lennard-Jones and torsion parameters for alkane chains [116]. The first systematic parameterization of lipid parameters in OPLS-AA was provided at the same time Lipid14 was released [118, 119]. In the initial parameterization effort, lipids showed a transition to crystalline phase at temperatures above the main phase transition temperature, but reparameterization of the hydrocarbon torsional potentials and Lennard-Jones parameters resolved this issue [119].

One significant approximation made by AA force fields is that atoms op-

erate as interaction sites with fixed charges. In reality, these atoms are electronically polarized. Polarizable force fields such as DRUDE, in comparison, include polarization effects. Models used to incorporate these effects include the inducible point dipole model and the fluctuating charge model [120]. Early versions of polarizable force fields were implemented in the 1970s [121, 122] and began to be studied "intensively" in the 1990s [120]. The multipole-based polarizable force field AMOEBA was found to provide improved descriptions of structural and thermodynamic properties of peptides and proteins [123, 124]. The DRUDE force field has been successfully implemented for both proteins [125] and zwitterionic PC and PE lipids [126, 127] for simulations up to μs in length [128]. While DRUDE-polarizable CHARMM has been used to study membrane interactions of ionizable arginine side chains [129], polarizable force fields are yet to be extensively used to study lipid-protein interactions. This is likely due to the increased computational expense in conducting such simulations.

Decades of research have produced large parameter sets are now available for many of the most commonly studied biomolecules, including proteins [101, 130, 131, 103, 132, 133, 134, 106], DNA/RNA [135, 136, 137, 138, 139, 140], carbohydrates [141, 142, 143, 144, 145, 146, 147], and lipids [148, 111, 114, 116, 119]. However, in some cases molecules may not be available in the force field. This is true of two non-standard amino acids required in this work, gamma-carboxy-glutamic acid residues (GLA) and (3S)-3-hydroxy-L-aspartic acid (BHD) residues. Parameters for these residues were developed using parametrzation by analogy, a strategy effective for molecules which contain functional groups similar those in already available in the parameter set. Several tools have been developed to aid in the analogy process, including CGenFF [149, 150] and MATCH [151] for the Charmm Generalized Force Field (CGenFF). For molecules which include functional groups which are not already included in a parameter set, the parameters for the new functional groups using QM and MM simulations. Packages such as The Force Field Toolkit (ffTK) [152] have been designed to streamline and aid in this process.

2.1.2 Numerical integration

In its most basic conception, Newton's equations of motion, $dv/dt = a(t)$ and $dx/dt = v(t)$ are solved during MD simulations for interaction sites (is) at each time step. During AA simulations, these interaction sites are individual atoms. Using a truncated Taylor expansion of $x_{is}(t)$, position at a timestep $t + dt$ can be represented as:

$$x_{is}(t + dt) = x_{is}(t) + v_{is}(t)dt + \frac{a_{is}(t)dt^2}{2} \tag{2.2}$$

Where dt is the magnitude of the time step. Acceleration, velocity, and position can be updated every time step, allowing movement of interaction sites (e.g., atoms) to be tracked over the course of the simulation. A number of numerical integration schemes have been developed for calculating these quantities, including the relatively simple Euler Method. Contemporary MD simulations use more sophisticated numerical integrators which allow for lesser computational expense. In the simulations described here, the standard Verlet integrator[153] was used, which is derived using a central difference approximation of d^2x/dt^2. The Velocity Verlet algorithm, which is typically used in simulation implementations, is shown below.

$$v_{is}\left(t + dt\right) = v_{is}\left(t\right) + a_{is}(t)\frac{dt}{2} + a_{is}(t + dt)\frac{dt}{2} \tag{2.3}$$

Critically, use of the central difference approximation significantly reduces the global error ($O(dt^2)$) incurred calculating each time step as compared to simpler integrators such as Euler's Method. This allows larger time steps to be used in simulation without loss of accuracy, significantly reducing computational costs. As stable integration of position requires time steps between force evaluations to be roughly ten times shorter than the timescale of the fastest degree of freedom [154], a standard approximation to is to fix the length of hydrogen bonds. With hydrogen bond-lengths fixed with SHAKE[155, 156], RATTLE[157], SETTLE[158], or LINCS[159] algorithms, time steps can be increased from 1 fs to approximately 2 fs each.

2.1.3 Temperature and pressure control

One way in which a biomolecular system may be simulated is by placing a set of atoms in a simulation cell of constant size, thus producing a NVE (constant number of particles, constant volume, constant energy) ensameble. However, a simulation constrained to constant energy may convert energy back and forth between kinetic and potential energy, potentially causing large fluctuations in both temperature and pressure. Biomolecular experimental measurements are generally performed at constant temperature and pressure, and it is desirable for MD simulations to be performed under the same conditions. Computational thermostats and barostats have therefore been developed were to allow simulation systems to be treated as NPT (constant particle, temperature, and pressure) ensembles. In NAMD [86] simulations, the standard Langevin thermostat is implemented. This method modifies velocities of interaction sites so as to take dissipative and stochastic forces into account[160], causing the simulation system to trend toward a set reference temperature [161].

The membrane is an anisotropic, experiencing differential stresses along the vectors parallel and normal to the membrane [162]. For membrane simulations, it is therefore also necessary to implement barostats to allows for contraction and expansion of the membrane within a periodic box while maintaining constant system pressure. NAMD[86] implements two barostats to regulate system pressure: the Berendsen barostat[163] and the Nosé-Hoover Langevin piston pressure control[164, 165, 166]. The Beresden barostat allows for contraction and expansion of the membrane by rescaling simulation box dimensions along a constant ratio. If the dimension of the box in the plane of the membrane contracts, the dimension normal to the membrane expands. The piston method prevents rapid oscillations in volume through addition of a dampening term [164, 165, 166].

2.1.4 Approximating non-bonded interactions

When referring to the equation detailing potential energy terms for forcefields (equation Eq. (2.1)), it is clear that the sums for most terms grow linearly with the number of atoms. Non-bonded terms, however, do not. Non-bonded terms grow grow as the square of the number of atoms ($O(N^2)$).

This difference is due to the fact that even distant interaction sites may exert force on each other. In most cases of very distant interaction sites in a system, any force exerted is deminimus. Methods have thus been developed to estimate long-range interactions effects without performing complete non-bonded interaction sums at each step, thus increasing efficiency of the calculation. A cutoff distance is defined, and if a pair of interaction sites is within that distance the summation described in equation Eq. (2.1) is performed. If a pair of interaction sites is further than the cutoff, the calculation is truncated. In performing an MD simulation, it is necessary to implement a system to track atom pairs and determine which fall within the cutoff distance. Since this operation scales linearly with system size, a significant saving in computational cost is achieved.

In the simulations described here, van der Waals interactions are truncated at 12Å allowing minimal error to be incurred. As the potential is dependent on distance r between interaction as r^{-6} and r^{-12}, the term converges rapidly. The electrostatic term, however, is dependent on r as r^{-1} , and does not conversge with distance. More complex methods are thus required to approximate long-distance contributions to the electrostatic potential term. All simulations described here use the particle mesh Ewald (PME) [167, 168] method, which necessitates use of periodic boundary conditions. Interaction sites engaged in long-range interactions are represented as components of a mesh, and regularity and periodicity of the mesh allows Ewald sums to be calcuated using Fast Fourier Transform (FFT) methods. Cost of the method is $O(NlogN)$ where N is to total number of interaction sites/particles.

2.2 Membrane modeling with the HMMM

Lipid diffusion at ambient temperatures is on the order of $10^{-8}\mathrm{cm}^2\mathrm{s}^{-1}$. As AA simulations of membrane proteins are typically run in the 100 ns–1 μs range, orders of magnitude greater sampling would be required to allow diffusive equilibration to occur. This makes many membrane-associated processes, such as spontaneous lipid mixing and membrane insertion of peripheral proteins, difficult to simulate. The highly mobile membrane mimetic (HMMM) model was developed to accelerate lipid diffusion while maintaining atomic-level details for the head-group region [169]. Since the introduction of the

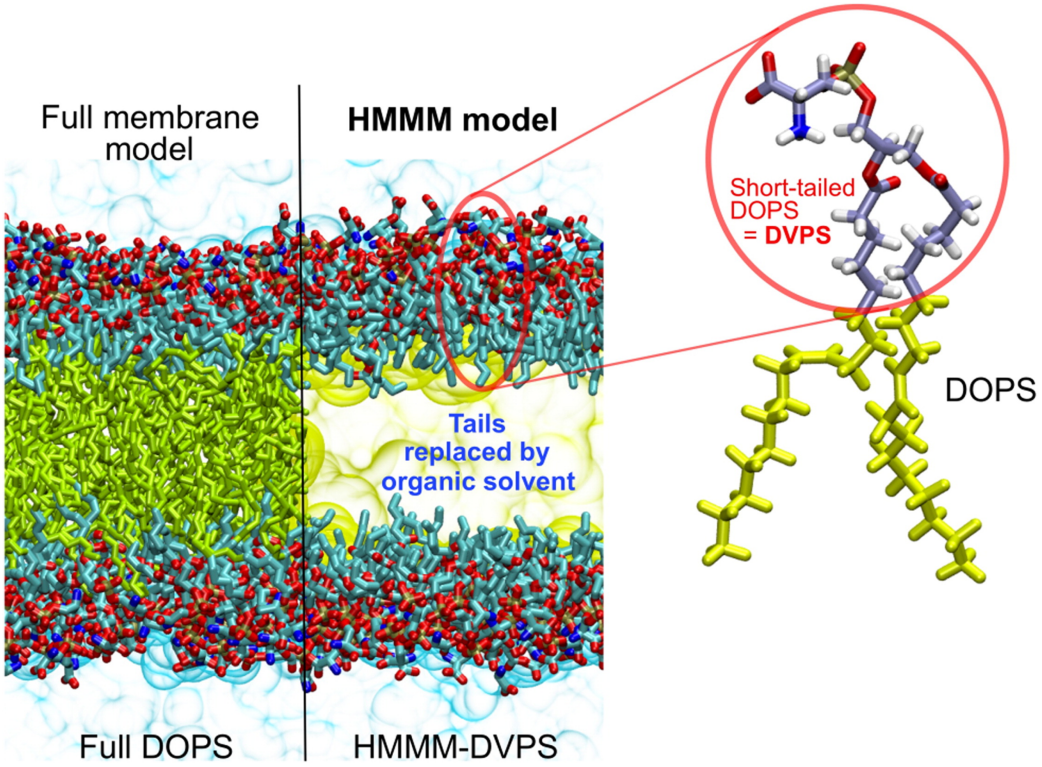


Figure 2.2: Comparison of HMMM and full membrane. Left-hand panel shows a full membrane. Lipid atoms removed in HMMM are shown in yellow, the rest of the lipid is shown in cyan, red, and blue. Hydrogen atoms not shown. Middle panel shows HMMM lipids. DCLE solvent illustrated as yellow fog. The right-hand panel shows and individual DOPS lipid. Atoms absent in HMMM lipids are shown in yellow. Atoms present in HMMM lipid are circled in red. Figure reprinted with permission from reference [169]. Copyright 2012 Biophysical Society.

HMMM model, the method has been applied to a wide variety of membrane-associated systems [169, 170, 171, 172, 173, 174, 175, 176, 177, 178, 74, 75, 179, 180, 181, 182, 162, 183, 184, 185, 186]. The HMMM uses lipids that are identical to AA lipids except that the acyl tail is truncated. The space between the truncated lipid tails is then filled with an organic solvent such as DCLE (dichloroethane) to reproduce some of the characteristics of the hydrophobic core of a lipid bilayer [169].

It was demonstrated that a triphasic system containing truncated PS lipids, DCLE solvent, and water would spontaneously form a core layer of DCLE with lipids at the perimeter and water in the external region [169]. The spontaneous bilayer formation occurred within 20 ns, with over 80% of ini-

tially water-submerged lipids reaching the DCLE interface within 10 ns [169]. Five systems were tested with area per lipid ranging from 68 Å^2 to 294 Å^2, yielding lipid diffusion constants at least 10 times that of full-length lipids [169]. Structural analysis showed that the HMMM model reproduced key membrane elements, such as degree of hydration and counter-ion penetration [169]. To test efficacy of the method with a protein system, the coagulation factor VII GLA domain, which had previously been studied using AA simulation with the goal of understanding its membrane binding [53]. (Figure 2.2) was chosen. In AA simulation, however, it was not possible to model spontaneous binding of the protein. Using HMMM, the GLA domain bound spontaneously to PS HMMM membranes in ten independent simulations, resulting in a converged model of the membrane-bound GLA domain [169].

Furthermore, the energetics of amino acid partitioning into the bilayer were assessed [173]. PMFs calculated for side chain insertion into the interfacial region showed that HMMM reproduced accurate results of AA and CG PMFs obtained for insertion into full lipid bilayers [173]. The HMMM with a DCLE core, however, did not accurately describe core energetics for protein partitioning [173]. HMMM has been integrated into the CHARMM-GUI input generator, allowing for convenient generation of HMMM bilayers with a variety of lipid compositions [180]. HMMM has now been used to study a wide variety of peripheral membrane proteins, including coagulation factor GLA domains [179] (discussed in Chapter 3), talin [174], and cytochrome P450 [170]. It has also been employed to study transmembrane domains [175] and the insertion process of lipids into the membrane [172]. Use of HMMM has allowed for extensive sampling of lipid-protein interactions following spontaneous binding of proteins with a high degree of lipid specificity [179], (Chapter 3) something that would be difficult to sample with full lipids in either AA or CG simulations.

While HMMM was highly proficient in modeling of peripheral proteins, energetic differences between DCLE solvent and a natural membrane core region can result in deformations of complex, multi-helix transmembrane proteins [162]. Interactions between protein side chains and the solvent are overly favorable, resulting in intercalation of solvent molecules between protein structural elements [162]. Vermaas et al, attempting to overcome this difficulty, developed *in Silico* solvents for use in HMMM simulations. These molecules, which do not exist in physical reality, are custom-parameterized

to mimic the membrane core while retaining liquid properties [162]. The solvents were demonstrated to allow for improved simulation of transmembrane proteins in HMMM, although some problems still remain [162]. Currently, HMMM still has limited ability to simulate transmembrane proteins with multiple transmembrane helices. Furthermore, HMMM cannot be used to accurately represent energetics and processes at the membrane core, as the properties of core solvents still differ significantly from those of the acyl chains in the lipid bilayer core. In addition, certain types of lipids, such as sphingolipids and sterols, have not been tested in HMMM to ensure it reproduces characteristics of full membranes composed of these lipids.

2.3 Nonequilibrium, or biased, simulations

A variety of methods have been developed to push or bias a simulation toward a desired endpoint. This allows processes to be investigated which would otherwise occur on too long a timescale to be captured at the selected method's resolution level. Biased simulations can be used as a starting point for calculating potential of mean force [187, 188] or free energy [189, 190, 191] associated with a transition process. They have also been used to study ligand binding, transition pathways for transporters, and reproduce AFM or optical tweezer experimental data.[79, 191]. Methods developed include SMD [78, 79], adaptive biasing force [192], and application of a harmonic potential to a collective variable as in targeted MD [193]. These methods all work using the same principle, which is to apply an external, time-dependent force. The force pulls the system along a selected degree of freedom, overcoming local energy barriers. Collective variables (colvars) are used to define the degrees of freedom [194].

In its generalized theoretical form, a collective variable c_v can be defined as a differentiable function of the vector of 3N particle coordinates X.

$$c_v(X) = c_v(x_1, x_2, x_3...x_n) \tag{2.4}$$

Depending on the structure of the system may be dependent on fewer terms

than 3N or expressed as a combination of several such functions.

$$c_v(X) = c_v(z^1(X), z^2(X), z^3(X)...z^\alpha...) \tag{2.5}$$

Number of basis functions α is much smaller than number of particles in the system, and in many cases one basis function suffices [194]. Common basis functions are distances, angles, or coordination numbers. In the work discussed here, coordination number and distances were primarily used. The value of the colvar is determined by summing its components.

$$c_v(X) = \sum_i e_i[z^i(X)]^{p_i} \tag{2.6}$$

Where e_i is the unique coefficient for each component and p_i is a positive integer exponent. Both are one by default. Harmonic potentials can be applied to either maintain the current value of a colvar (a fixed restraint) or move it toward a new value (moving restraint). The form of the potential is shown is the equation below,

$$V(c_v^n) = \frac{1}{2}k\,(\frac{c_v^n - c_v^0}{\Delta c_v^n})^2 \tag{2.7}$$

where k is the force constant and Δc_v^n is the predefined width parameter, which is set to zero. The units of k are are $kcal/mol/\Delta c_v^n$. Chapter 5 describes simulations which apply harmonic potentials to colvars. Colvars used in these simulations are either distance based, and change distance between center of mass of two groups of atoms along a specified vector, or coordnum colvars. Coordnum induces contacts within a specified cutoff to be formed between a two groups of atoms. Coordination number (C) is calculated according to the following formula, where x and y are the two groups of atoms:

$$C(x, y) = \sum_{i \in x} \sum_{j \in y} \frac{1 - (|x_i - x_j|/d_0)^n}{1 - (|x_i - x_j|/d_0)^m} \tag{2.8}$$

Where d_0 is the predefined cutoff distance. The exponents n and m control long-range behavior and by default are set to 1 and 2, respectively.

2.4 Refinement of experimental structures

All of the simulations discussed in the upcoming chapters use x-ray crystallographic structures from the protein data bank[195, 196] as a basis for initial protein structure. X-ray crystallography depends upon developing crystals containing the protein of interest. X-rays are then passed through the crystal, and the structure of the protein within is deduced from its diffraction pattern. Homogeneous structures yield more regular diffraction patterns, and so mobile elements may be poorly resolved or omitted. Here, missing segments of proteins in Chapter 6 were added using MOE loop homology modeling tools [197] In simulating truncated proteins, a neutral cap is added to replicate a charge environment more similar to that of the *in vivo* protein. A variety of such caps are provided in the CHARMM force-field. pK_a of ionizeable side chains of residues such as histidine and aspartate was predicted using PROPKA [198, 199], allowing protonation state to be assigned.

2.5 Rigid-body protein-protein docking

In Chapter 5, rigid-body docking is used to predict contacts between proteins in the extrinsic coagulation complex. The goal protein-protein docking is to predict binding of two molecules starting from their isolated structures [200, 201]. While some methods allow the proteins to change structure once in contact [202], this adds to computational expense, and may not yield superior predictions [201]. Maintaining each structure as a rigid-body decreases the number of conformations to be evaluated and thus reduces the computational cost. Transformations on three rotational and three translational degrees of freedom are evaluated. These searches can be performed very efficiently through Fast Fourier Transform (FFT) [203] methods which map molecular properties onto 3D-grids. These algorithms currently allow a systematic search of 100 billion conformations to be performed in a few hours [201, 204].

The study described in Chapter 5 used the rigid-body program DOT2 [204]. In DOT2, a stationary and moving molecule are defined. Stationary and moving molecules are treated differently in determining their electrostatic potentials. The moving molecule is treated as a set of point charges. The

electrostatic potential of the stationary molecule, on the other hand, is calculated using Poisson-Boltzmann methods using the program APBS [205]. The vander Waals energy term is calculated using the program MSMS [206] from heavy atoms for the stationary molecule. The shape potential is defined as an excluded volume based on a 1.4 Å probe sphere surrounded by a 3 Å favorable layer. The shape potential of the moving molecule is defined by the position of its heavy atoms.

3 LIPID SPECIFICITY OF COAGULATION FACTOR X*

3.1 Introduction

The coagulation cascade is a critical pathway triggered by vascular damage, terminating in the formation of a blood clot [207, 29]. Coagulation is extensively regulated through membrane interactions; the cascade is greatly accelerated upon binding of coagulation factors to negatively charged lipids of the membrane [13, 16]. Once bound, coagulation proteases become thousands of times more active [16, 12] metamorphosing from extremely poor to highly effective enzymes.

Although membrane interactions play an important role in the coagulation cascade, the atomic-level details of how this regulation takes place are not yet fully elucidated. Factor X (FX) is a serine protease zymogen in the final common pathway of blood coagulation, activation of which is required to generate the "explosive" thrombin burst necessary for blood clotting [207, 13, 208, 209]. Formation of complexes which include FX in this pathway, including the extrinsic coagulation complex, are highly membrane dependent [208] FX is one of several coagulation proteins which bind the cellular membrane using a specialized membrane-binding "GLA" domain. These proteins are collectively known as vitamin K dependent proteins [13, 44, 45], as this vitamin is required for post-translational modification of multiple glutamates into γ-carboxyglutamic acid (GLA) residues, necessary for proper folding and activity of GLA domains [46]. Each GLA residue carries a -2 charge, and typically coordinates one or more Ca^{2+} (or, sometimes, Mg^{2+}) ions [210, 211]. This is necessary for proper folding of the GLA domain resulting in a motif where divalent metal ions and

*This chapter is reproduced with permission from the article: (1) M. P. Muller, Y. Wang, J. H. Morrissey, and E. Tajkhorshid, "Lipid specificity of the membrane binding domain of coagulation factor X," J. Thromb. Haem., vol. 15, pp. 2005–2016, 2017. Copyright 2017 International Society on Thrombosis and Haemostasis.

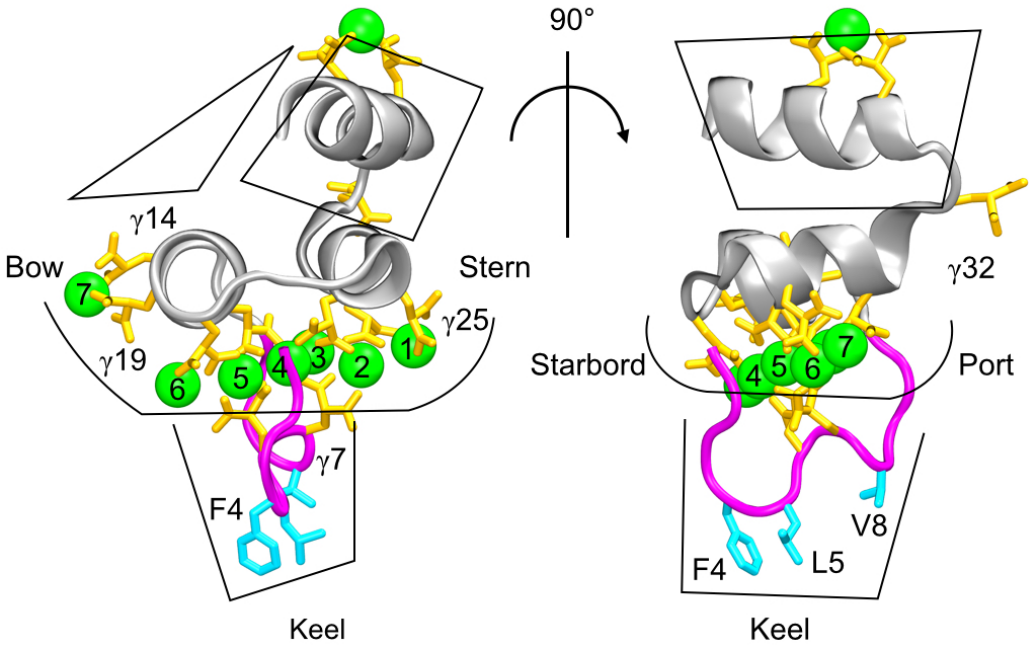


Figure 3.1: The bovine FX GLA domain (FX-GLA) from protein databank structure 1IOD [27]. Significant features of the GLA domain are: 1) Ca^{2+} ions, shown as green balls, 2) specialized γ-carboxyglutamic acid residues, shown in orange, 3) the ω-loop, highlighted in magenta, and 4) the hydrophobic keel residues, high- lighted in blue. In order to refer to orientation of the GLA domain, we have appropriated the nautical terminology of "stern", "bow", "port" and "starboard" to refer to four sides of the protein.

highly charged protein residues are packed into the center of the GLA domain, while three hydrophobic residues in the ω-loop are exposed on the exterior [10, 47, 48, 49, 50, 51, 52]. Upon GLA domain binding to membranes, these three hydrophobic residues on the ω-loop, which we term the "keel," penetrate into the hydrophobic core of the membrane [54, 53]. The bovine FX-GLA domain (PDB entry 1IOD [56]) contains 8 Ca^{2+} ions and 11 GLA residues, mostly located close to the keel. Previous computational work on a similar GLA domain, FVII-GLA, found that this region interacts with the charged head groups of membrane phospholipids [54]. Structure of the FX GLA domain shown in Figure 3.1.

While structure and amino acid sequences among GLA domains are highly conserved [45], GLA domains from different coagulation factors show a marked variance in lipid specificity and membrane binding affinity [13, 23, 22]. FX and most other vitamin K-dependent clotting proteins bind preferentially to membranes containing exposed phosphatidylserine (PS), while the GLA domains of the closely related vitamin K-dependent proteins, FVII and protein C bind preferentially to phosphatidic acid (PA) [22, 15, 55]. The head groups of PS and PA are extremely dissimilar on the atomic level, although both have a net negative charge. While PA consists of a phosphate group alone, the PS head-group includes three charged groups: a phosphate, a carboxy, and an amino group. The atomic-level interactions responsible for this difference in head-group specificity are unknown. In the following chapter, we will discuss simulations which elucidate lipid specificity of the FX-GLA domain. In Chapter 5, we will discuss simulations which study lipid specificity of FVII-GLA domain.

Only one well-folded structure of the FX-GLA has been experimentally characterized with all Ca^{2+} resolved, a bovine FX-GLA bound to a snake venom protein (PDB 1IOD) [56, 13]. All structures characterized for human FX either lack the GLA domain or, when the GLA domain is resolved, show disordered protein folds and missing or unresolved Ca^{2+} ions [212]. Given the importance of Ca^{2+} ions in mediating membrane binding [53], as well as their importance in maintaining a fold with exposed hydrophobic residues [52], it is unwise to draw any conclusions on membrane binding from such disordered structures. All of these structures were characterized in solution, and thus can give little insight into membrane binding interactions. While the prothrombin GLA domain, which is also PS-specific, was crystallized with

a bound lyso-PS, an interaction outside of the context of a bilayer can give limited information on protein-lipid interactions relevant to the membrane-bound state required for coagulation. All-atom MD simulations to study GLA domains have been impeded in the past by the slow timescales required to sample equilibrium protein-lipid interactions [19]. Simulations described here therefore use the HMMM to speed lipid diffusion. Please refer to the methods section for a detailed discussion of the HMMM.

This chapter describes the first model of membrane-bound FX-GLA using extensive molecular dynamics (MD) simulations, and we use this model to investigate protein-lipid interactions between the FX-GLA and PS lipids. Starting from the solution crystal structure (PDB 1IOD [56]), we obtained spontaneous binding of FX-GLA to a PS HMMM in 14 independent, unbiased simulations. These membrane-bound states were converted to conventional membrane representations upon converged binding. The results of 200 ns for each of the 14 FX-GLA structures discussed here allowed for detailed characterization of membrane penetration depth and membrane orientation of the FX-GLA domain. Further, the extensive sampling of lipid-protein interactions has allowed for detailed analysis of GLA-lipid interaction patterns and identification of potential PS-specific binding sites. All three charged functional groups of PS lipids were found to interact extensively with charged moieties of the protein, including GLA side chains, bound Ca^{2+} ions, and basic protein side chains.

3.2 Methods

3.2.1 Membrane binding simulations

A crystal structure of bovine FX-GLA (which we will refer to hereafter as FX-GLA) from Protein Data Bank entry 1IOD [56] was used as the starting structure for our simulations. Using the PSFGEN plugin of VMD [213], an N-terminal ammonium group and N-methyl amide C-terminal group were added, along with hydrogen atoms. The N-methyl amide group was chosen rather than a standard C-terminus because the C-terminus in the full-length protein would be connected to the first FX epidermal growth factor (EGF1) domain, and a C-terminus with a default charge would introduce an artificial

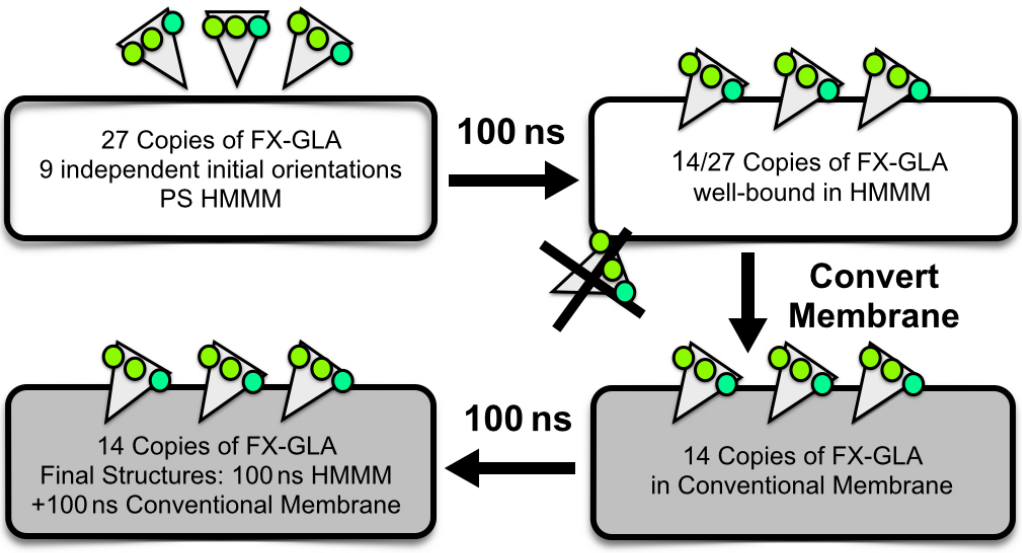


Figure 3.2: Illustration of the simulation protocol for developing the FX-GLA membrane-bound model. After the initial 27 membrane-binding simulations were run for 100 ns, the models with keel residues inserted below the level of the phosphate plane in the membrane (14 systems) were converted to conventional full-tailed membrane models and simulated for an additional 100 ns.

charged group to this region. FX-GLA was next placed in a 60×60×60 Å water box using the SOLVATE plugin of VMD. This box was neutralized with K^+ ions added using the AUTOIONIZE plugin of VMD and equilibrated for 1 ns after 2000 steps of energy minimization. This solvated model was used as the starting point for all the subsequent membrane-binding simulations. Our HMMM membrane construct was used in the binding simulations to increase lipid dynamics and allow for both increased lipid interaction sampling and ease of GLA domain insertion. A 50×50×114 Å solvated HMMM-PS lipid patch was constructed using the HMMM Builder in CHARMM-GUI [180]. Nine independent orientations were generated by placing FX-GLA 3 Å above the membrane. Each FX-GLA membrane orientation was at least a rotation of 15° different in the x-z or y-z plane from every other orientation. Finally, each of the nine independent systems was neutralized with K^+ ions using the AUTOIONIZE plugin of VMD. The final systems were each approximately 24,000 atoms in size. The z-position of HMMM short-tailed lipids was restrained harmonically (k=0.1 kcal/mol/Å^2 applied to lipid tail atoms with CHARMM force-field numbers C31 and C21) throughout the simulations in order to more closely reproduce the natural distribution of head-group positions. Each of the 9 independent FX-GLA systems was simulated 3 times with different initial velocities for a total of 27 independent simulations of 100 ns each (see Figure 3.2).

3.2.2 Conversion of HMMM model to conventional full membrane

In order to ensure that the final FX-GLA membrane-bound model was not unduly affected by the use of the HMMM as opposed to a full-tailed lipid construct, all bound systems were converted to conventional (full-tailed) membrane, using a PDB generated from the final frame of each 100 ns HMMM simulation. Conversion to a full-tailed membrane was only made for FX-GLA structures in which the keel domain penetrated below the level of phosphates in the membrane, which we will hereafter refer to as "properly bound." Examples of FX-GLA which was not properly bound can be found in Figure 3.5 full-tailed lipid was overlaid onto each short-tailed lipid, and coordinates of the missing atoms were then copied from the full-tailed lipid, thus al-

lowing for head-group interactions obtained from HMMM simulations to be preserved. The layer of 1,1-dicholoroethane (DCLE) solvent molecules were then removed from the center of the HMMM bilayer. After a 2,500-step energy minimization, each FX-GLA membrane-bound system was simulated for another 100 ns.

3.2.3 Simulation protocols

All simulations were performed using NAMD2 [86] and the CHARMM36 [111] set of force-field parameters. Parameters of GLA residues were those previously developed by our lab [53] with atom types renamed as appropriate for the conversion from CHARMM27. The TIP3P water model was used for all simulations. Constant pressure was maintained at a target of 1 atm using the Nosé-Hoover-Langevin piston method [164]. Constant temperature was maintained by Langevin dynamics with a damping coefficient, γ, of $1\,\mathrm{ps}^{-1}$ applied to all atoms [165]. All simulations used a time step of 2 fs and a temperature of 310 K. The particle mesh Ewald (PME) method [167] was used for long-range electrostatic calculations with a grid density greater than $1/\mathrm{Å}^3$. Non-bonded interactions were cut after 12 Å with a smoothing function after 10 Å. System setup, visualization, and analysis were carried out using VMD [213].

3.2.4 Analysis

Depth of GLA domain insertion into the membrane was assessed by calculating the z-component of the center of mass (COM) of backbone atoms in residues PHE4, LEU5, and VAL8 in the keel of FX-GLA relative to the COM average plane of phosphate groups (phosphorous atoms and four bonded oxygen atoms) of the PS lipids. The z-position of the COM of Ca^{2+} ions 1-7 (Ca-1 to Ca-7), the z-position of the COM of the lipid carboxy groups (carbon and two bonded oxygen atoms), and the z-position of the COM of the amino groups (nitrogen and three bonded hydrogen atoms) were also calculated relative to the phosphate plane; from this point forward the terms Ca^{2+} ions, phosphate, carboxy and amino groups will refer to the atoms specified above.

Orientation of the GLA domain relative to the membrane plane as assessed using two angles defined in two perpendicular planes, termed "port-to-starboard" and "bow-to-stern." The bow-to-stern angle is the angle between the membrane normal (z-axis) and the vector connecting the COM of Ca-1,2,3 and Ca-4,5,6 (see Figure 3.1). The port-to-starboard angle is the angle between the membrane normal and the vector connecting the COM of backbone atoms in LEU13, GLA14, GLA29, and VAL30 and the COM of backbone atoms in GLA19, CYS22, SER23 and LEU24. These two sets of residues represent opposite ends of the two helices directly above the level of the coordinated Ca^{2+} ions. Thus, the bow-to-stern angle is a measure of degree of the stern and bow ends of FX-GLA and the membrane normal, while the port-to-starboard angle represents relative tilt of the port and starboard ends relative to the membrane normal. Refer to Figure 3.1 for the location of the port, starboard, stern and bow ends of FX-GLA.

Contacts between lipid head-group charged moieties and FX-GLA (bound Ca^{2+} or protein residues) were calculated based on specific distance cutoffs which were determined using radial pair distribution function calculations. Number of contacts here is the number of atoms on one group (e.g., a carboxy group) within a specified cutoff distance of the second group (e.g., a specific protein residue). Distance between all protein and Ca^{2+} atoms and phosphate, amino, and carboxy atoms were calculated, and distances equal to or less than the cutoff were determined to be a contact. The radial pair distribution functional analysis was performed separately for distances between the residue type of interest (protein or Ca^{2+} ions) and the lipid head-group moiety of interest (carboxy, amino, or phosphate). The calculation sampled every 10th frame of the 100 ns full-membrane simulations. Graph of the radial distribution shown in Figure 3.3.

For protein residues and Ca^{2+} ions, all atoms were included in the contact number calculation. Cutoffs determined were 2.50 Å for phosphate-Ca^{2+}, 3.00 Å for carboxy-Ca^{2+}, 2.25 Å for phosphate-protein, 2.50 Å for carboxy-protein, and 2.25 Å for amino-protein contacts. To ensure that all atoms interacting with the protein were included in the calculation, for the contact analysis the periodic box was centered around the protein and all atoms wrapped into the same periodic box.

To screen for potential putative PS-specific binding interactions, average number of protein and Ca^{2+} contacts per lipid were calculated for the full-

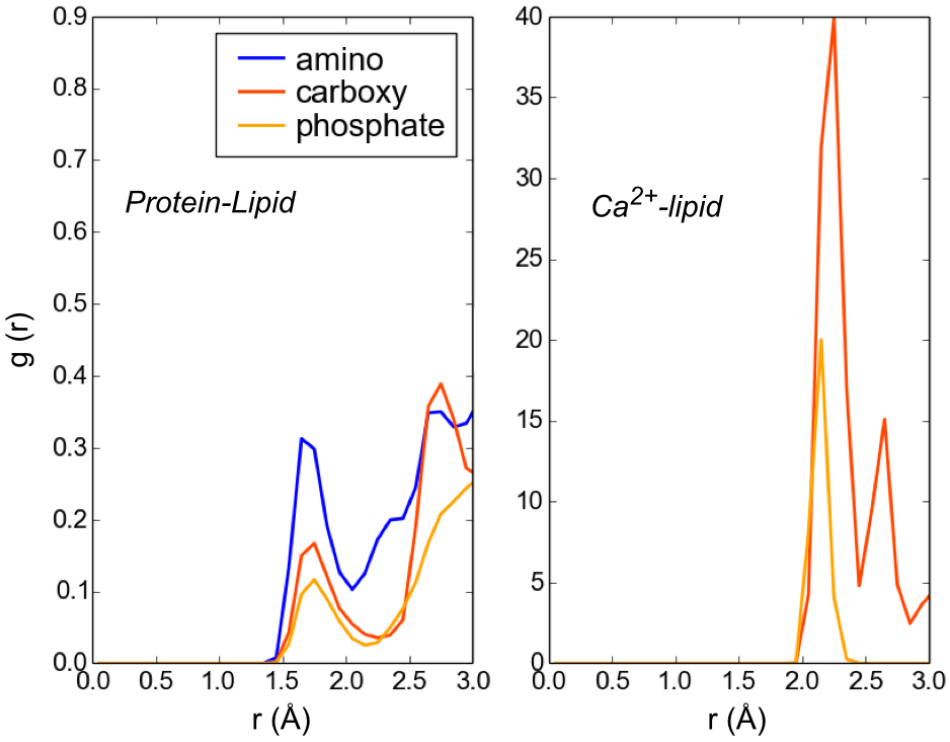


Figure 3.3: The radial distribution function for (A) distances between protein residues and lipid charged functional groups and (B) distances between Ca^{2+} ions (bound to the GLA domain) and lipid charged groups. All contact calculations used the cutoff values derived from this analysis, namely 2.25 Å for protein–phosphate contacts, 2.5 Å for protein–carboxy contacts, 2.25 Å for protein–amino contacts, 2.5 Å for Ca^{2+}–phosphate contacts and 3.0 Å for Ca^{2+}-carboxy contacts. Every 10th frame of a representative 100 ns conventional simulation was used to calculate the radial distribution function. Atoms used for each moiety were: all Ca^{2+} ions for Ca^{2+}, all protein atoms for protein residues, the carbon and two bound oxygen atoms for carboxy groups, the phosphorus atom and four bound oxygen atoms for phosphate groups, and the nitrogen atom and three bound hydrogen atoms for amino groups.

membrane simulations. Lipids with high average protein contacts were visually inspected to determine whether more than one charged group on the lipid was interacting simultaneously with FX-GLA. Distance between the lipid charged groups and protein residues or Ca^{2+} ions were then calculated to determine the lifetime of each interaction during the course of the full-membrane simulation.

3.3 Spontaneous membrane insertion of the FX-GLA domain

In order to characterize the membrane-bound configuration of the FX-GLA domain, we performed membrane-binding simulations starting with 27 copies of the FX-GLA in 9 independent orientations and placed above a 100% PS HMMM membrane. After 100 ns, 14 of 27 FX-GLAs were observed to insert keel residues spontaneously into the membrane (Figure 3.4). Following keel insertion, each GLA domain quickly established a stable binding pose which is conserved across all 14 FX-GLA domains. As can be seen in Figure 3.4, after the keel loop crosses the charged layer of lipid head groups, insertion depth of both the keel residues and Ca^{2+} ions converges within 10 ns. For equilibrated trajectories, the COM of the keel residues reaches a depth of 5 Å below the PS phosphate plane, penetrating well into the membrane's hydrophobic core. The COM of the Ca^{2+} ions, which interact with the membrane during binding, converges to a z-position 2 Å or less above of the carboxy plane, and approximately 5 Å above the phosphate plane. Similarly, the 14 FX-GLAs, in spite of highly varied initial orientations, all converged to similar port-to-starboard and bow-to-stern orientations relative to the membrane subsequent to keel insertion (Figure 3.6). A well-defined binding pose can thus be characterized using these 14 trajectories in terms of both depth of insertion and orientation. The remaining FX-GLA simulations failed to insert their keel domains into the hydrophobic core within 100 ns of HMMM simulation (11 simulations), or showed instability in the GLA domain fold (2 simulations), and thus are not included in further simulations and analysis. Examples of improperly bound trajectories can found in Figure 3.5.

The 14 properly bound FX-GLA models were converted from HMMM to a conventional membrane representation. As can be seen in Figures 3.4 and

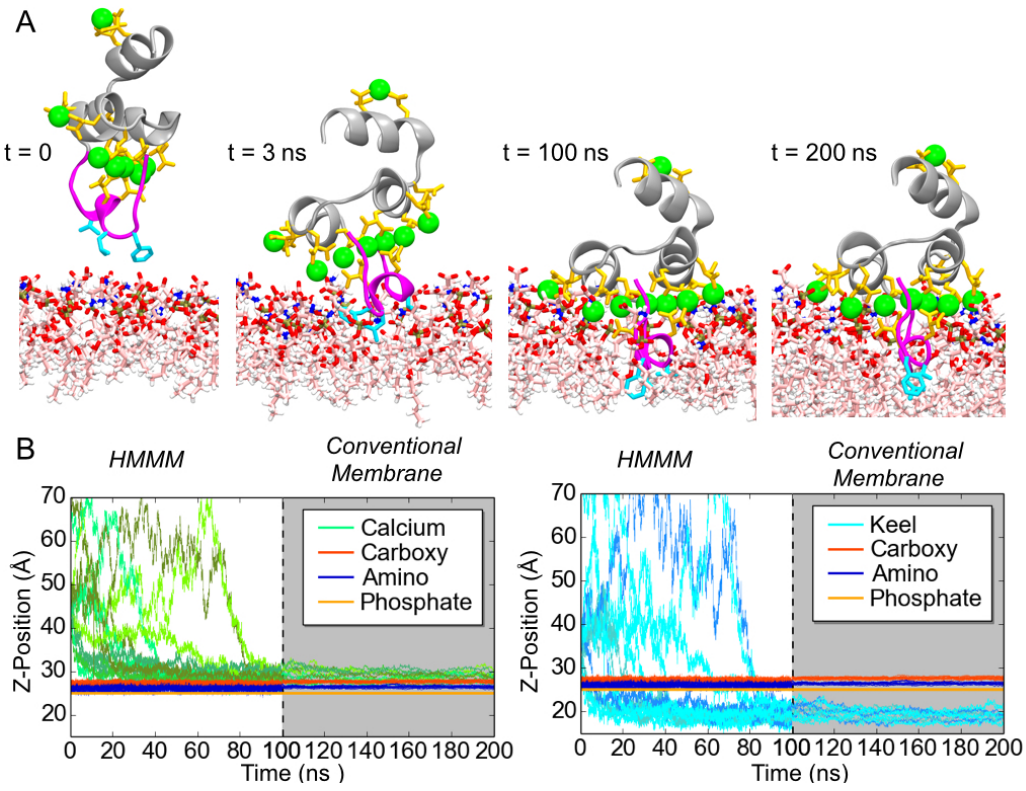


Figure 3.4: Membrane insertion and binding of the FX-GLA domain. A) Binding of the FX-GLA domain at different time points along a representative trajectory. Until 100 ns, the domain is in an HMMM membrane representation. After 100 ns, indicated by the dashed line, the membrane is converted to a conventional membrane. B) Relative z-positions of the COM of relevant moieties, including: lipid phosphate, carboxy and amino groups (both panels); Ca^{2+} ions (right panel); and keel protein (PHE4, LEU5, and VAL8) backbone atoms (left panel). See Methods for atoms included in phosphate, carboxy, and amino groups. Grey background indicates portions of the trajectories with a conventional membrane. Data in this figure refer only to the 14 properly bound FX-GLA domains. Refer to Figure 3.5 for information on the 13 trajectories with improperly bound trajectories. See Methods for definition of "properly" and "improperly" bound.

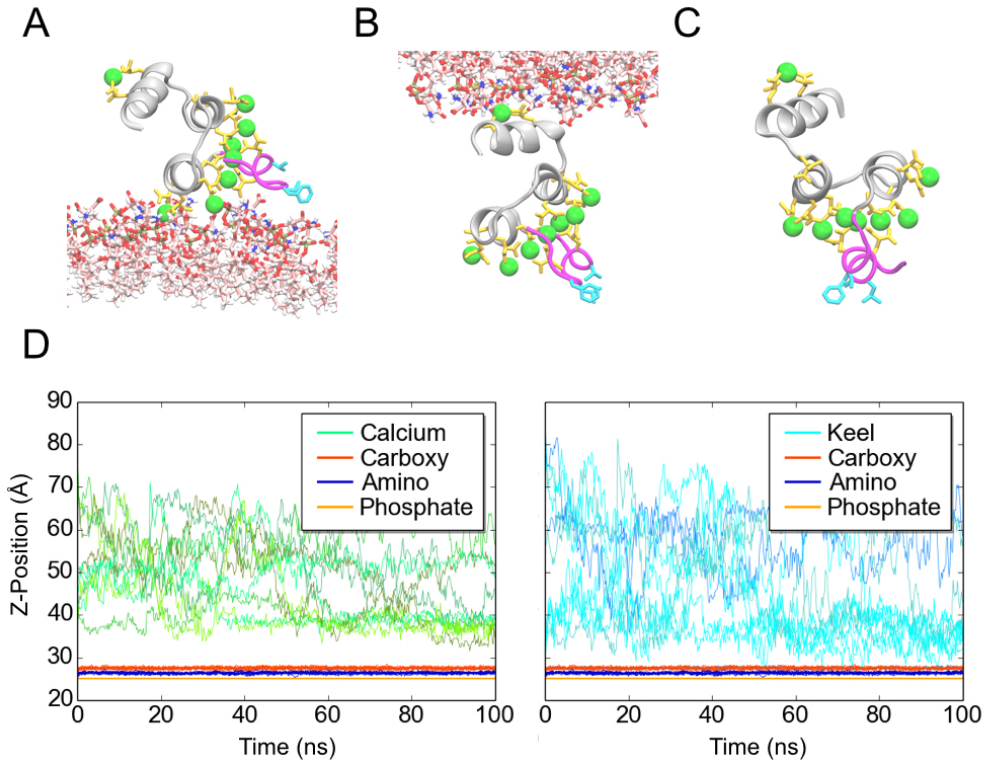


Figure 3.5: Membrane insertion and binding for improperly bound and discarded trajectories. Binding and penetration depth relative to the membrane for these trajectories are shown (A). (B and C) Examples of binding from these trajectories, which are representative of binding found in nine trajectories discarded for failing the criteria of membrane penetration of the keel residues below the phosphate head-group region. Of the simulations with keels above the hydrophobic membrane core, nine FX-GLAs interacted with the membrane but represented non-specific binding poses, which would be incompatible within the context of the full-length coagulation FX. Poses of these nine simulations were either bound sideways to the membrane, as in (B), or upside down, as in (C). Two simulations showed no membrane interaction. An additional two trajectories were rejected as they showed partial unfolding of the omega loop (D).

3.6, both the orientation relative to the membrane plane and the insertion depth within the membrane plane were stable in the 100 ns of simulation

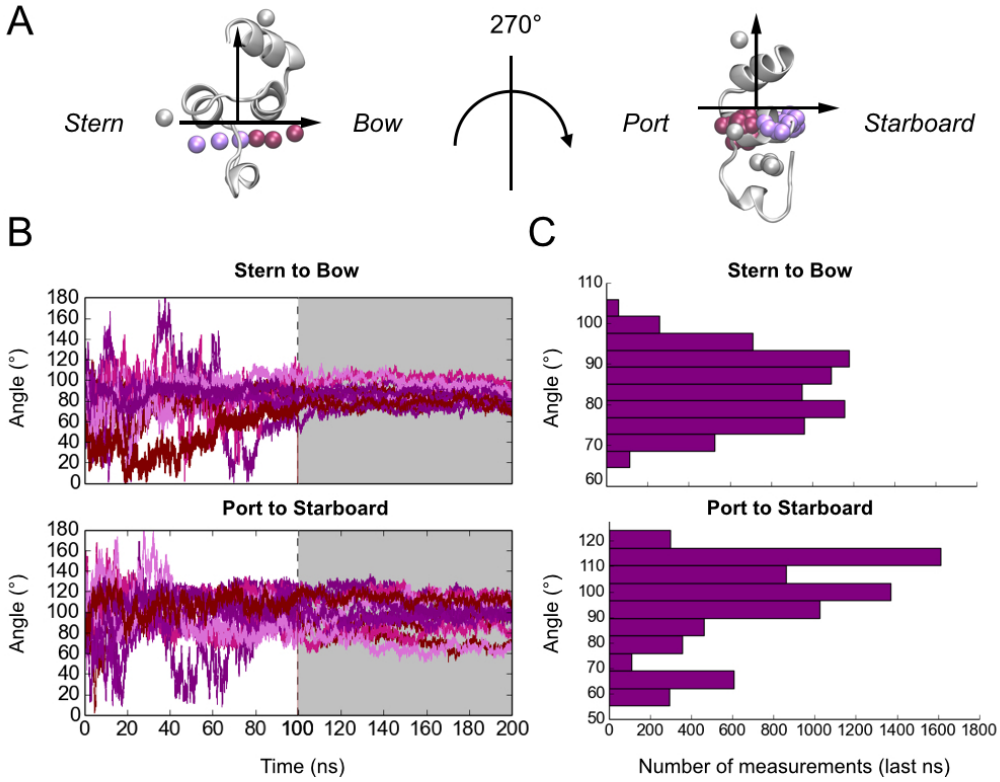


Figure 3.6: FX-GLA orientation relative to the membrane normal. A) Illustration of the "bow-to-stern" and "port-to-starboard" angles to the membrane normal. Atoms used to define opposite ends the vectors are shown in violet and maroon balls. B) Data tracking the bow-to-stern (upper panel) and port-to-starboard (lower panel) angles for all 14 properly bound trajectories. The portion of each panel with a grey background is the 100 ns of simulations in conventional membrane. C) Histogram of bow-to-stern (upper panel) and port-to-starboard (lower panel) angles during the last 10 nanoseconds of simulation in the conventional membrane representation. All 14 properly bound trajectories are included.

following the membrane conversion. In fact, the membrane binding depth and orientation among the 14 trajectories continues to converge further during the 100 ns conventional membrane simulations. The distribution of bow-to-stern angles (Figure 3.6 C, upper panel) follows a classic normal distribution centered at 85°. This suggests the FX-GLA is either parallel to the membrane or tilted slightly toward the bow end. The distribution of port-to-starboard angles is less regular, with the majority of angles lying between 90-120°, and

a small sub-population at about 65°. This indicates that in most bound FX-GLAs the starboard end of the protein is tilted toward the membrane at about 10°, which would potentially enable the GLA32 residue to interact with the membrane (see Figure 3.1). In a small minority of trajectories with a more acute angle, the port-to-starboard angle trends toward an increased value at the end of the trajectory.

The fluctuations in binding height and angle were found, as expected, to decrease upon conversion of the membrane to a conventional representation. This is consistent with perturbation of the binding pose being lessened

Table 3.1: Protein-lipid contacts. Average contact values ≥ 0.05 reported, rounded to two decimal places. The standard deviation is reported in parentheses

Residue	Carboxy	Phosphate	Amino
ASN 2	0.12(0.08)	0.16(0.08)	0.14(0.11)
SER 3	0.12(0.07)	0.27(0.11)	0.08(0.08)
PHE 4		0.09(0.07)	
LEU 5		0.08(0.06)	
GLA 6			0.06(0.04)
GLA 7			0.5(0.09)
LYS 9		0.50(0.12)	
GLN 10	0.09(0.09)	0.43(0.15)	0.12(0.12)
GLY 11		0.06(0.05)	
ASN 12	0.07(0.07)	0.05(0.06)	
GLA 14			0.26(0.07)
ARG 15		0.37(0.10)	
GLA 19			0.21(0.06)
GLA 20			0.29(0.08)
SER 23	0.08(0.05)		
GLA 25			0.29(0.07)
ARG 28	0.14(0.06)		
GLA 29			0.32(0.08)
GLA 32			0.23(0.08)

through slower lateral diffusion of surrounding lipids. Our membrane-bound model of the FX-GLA domain is thus stable under normal membrane conditions, and shows minimal perturbation upon the shift from HMMM to conventional membrane representation. This strongly suggests that our HMMM

model captures essential characteristics of the membrane-bound FX-GLA. PS-Specific Binding of the FX-GLA Domain Protein-lipid interactions were analyzed using all 100 ns trajectories obtained from conventional simulations to determine the nature of PS interactions with the protein, as well as the location and characteristics of potential PS-specific binding sites on the FX-GLA. FX is highly PS specific, even stereo-chemically specific [26], while other structurally and sequentially similar GLA domains (FVII, protein C) are specific to other negatively charged lipids [22, 15, 26] It is unknown how atomic-level interactions give rise to this specificity. Given the importance of PS interactions in the binding of FX, we determined contacts between FX-GLA residues and each of the PS charged groups: phosphate (-1), carboxy (-1), and amino ($+1$).

Examination of these protein-lipid contacts reveals an interesting pattern (Figure 3.7). The significant amino group contacts (average contact number greater than 0.2) occur with the specialized GLA residues. All but one of the GLA residues show significant interactions with amino groups, meeting the cutoff distance criteria in 25-40% of all trajectory frames with low standard deviation values. This indicates that at any given time, approximately two GLA residues on the FX-GLA interact with amino groups. GLA7, which has backbone atoms in the ω-loop, is most likely to be engaged in this interaction (0.45 average contact number), while GLA residues in the two helices located above the level of Ca^{2+} ions 1-7 (residues 13 to 32) have roughly equal likelihood of interacting with an amino group ($\sim$0.3 average contact number) (Figure 3.7, Table 3.1). GLA32 shows significant amino interactions, consistent with the observation that the starboard end of the FX-GLA generally tilts toward the membrane. It is also the last residue in the 44-residue sequence to show any membrane contacts.

The majority of contacts between Ca^{2+} ions and lipids are with carboxy groups; at any given time there is at least one carboxy group bound to each of the bow Ca^{2+} ions (Ca-1 and Ca-2) and stern Ca^{2+} ions (Ca-6 and Ca-7). Ca-6 also frequently interacts with phosphate (0.52 average contact number) (Figure 3.7, Table 3.2). The three center Ca^{2+} ions (Ca-3, Ca-4, and Ca-5) show no charge-charge interactions, indicating that their role is purely structural and that they do not participate in mediating membrane binding. Ca^{2+} ions which interact with PS carboxy groups are all coordinated by GLA residues with significant amino group interactions. This suggests that

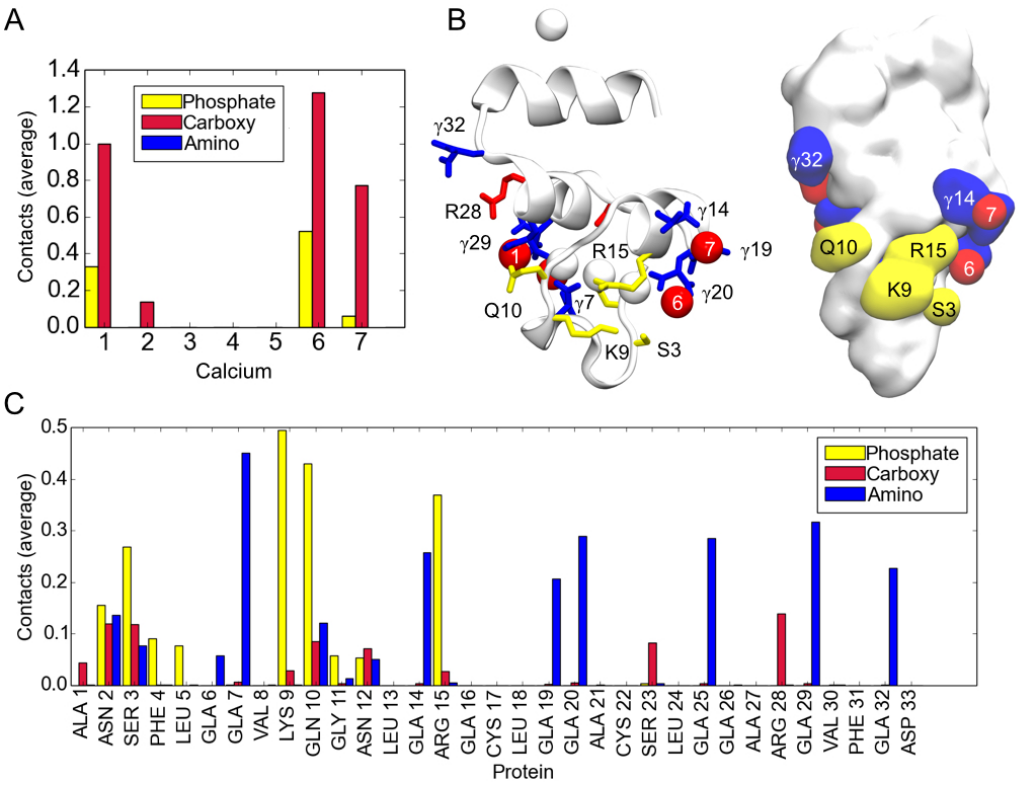


Figure 3.7: Lipid contacts with the GLA domain. A) Histogram of average contacts between Ca^{2+} ions on the GLA domain and carboxy and phosphate groups in PS lipids. B) Illustration of the FX-GLA domain indicating protein residue and Ca^{2+} ion numbering, along with the charged group with majority of contacts to the residue. Blue indicates an amino interaction, red indicates a carboxy interaction, and yellow indicates a phosphate interaction. Only average contact numbers 0.1 or greater are illustrated with a color code. C) Histogram of protein residue contacts to lipid charged groups. Average contact values are calculated across all 14 trajectories for the entire 100 ns of the conventional membrane simulations. Refer to Methods for details on contact calculation.

Table 3.2: Ca^{2+}-Lipid Contacts. Average contact values rounded to two decimal places. The standard deviation is reported in parentheses.

Ca^{2+}	Carboxy	Phosphate
1	1.00(0.14)	0.33(0.08)
2	0.14(0.04)	0.00
3	0.00	0.00
4	0.00	0.00
5	0.00	0.00
6	1.23(0.19)	0.52(0.10)
7	0.77(0.10)	0.06(0.03)

the same PS lipid may contribute the carboxy and amino groups for both interactions. It would be difficult sterically for separate PS lipids to satisfy both interactions simultaneously, as the amount of space between a Ca^{2+} ion and its coordinating GLA residue is small.

The primary protein-phosphate binding sites on FX-GLA consists of two clusters of residues on the port and starboard faces of the protein, with the cluster consisting of LYS9, GLN10, and ARG15 on the starboard side being particularly consistent and prominent throughout the trajectories (0.50, 0.43, and 0.37 average contact number, respectively). As can be seen in Figure 3.7 B, the side chains of these three residues have a very similar z-position relative to the keel and are all located on the starboard side of the protein, forming a phosphate binding patch. SER3 and ASN2 on the port side can also anchor the GLA through phosphate binding, but this is observed less frequently than binding to the port cluster (0.16 and 0.27 average contact numbers, for SER3 and ASN2, respectively). The overall binding pattern for the FX-GLA is tiered; charged residues close to the keel primarily form lipid phosphate contacts, whereas those surrounding the Ca^{2+} ions participate in interactions with lipid amino and, more rarely, lipid carboxy charged groups (SER23 and ARG28).

We identified a number of putative PS-specific binding sites, or sites where PS lipids were bound using multiple charged groups. In some instances, all three functional groups on a PS lipid simultaneously engaged with protein residues and Ca^{2+} ions. Interactions with multiple charged groups on the same lipid are more likely to be indicative of PS-specific sites. The two puta-

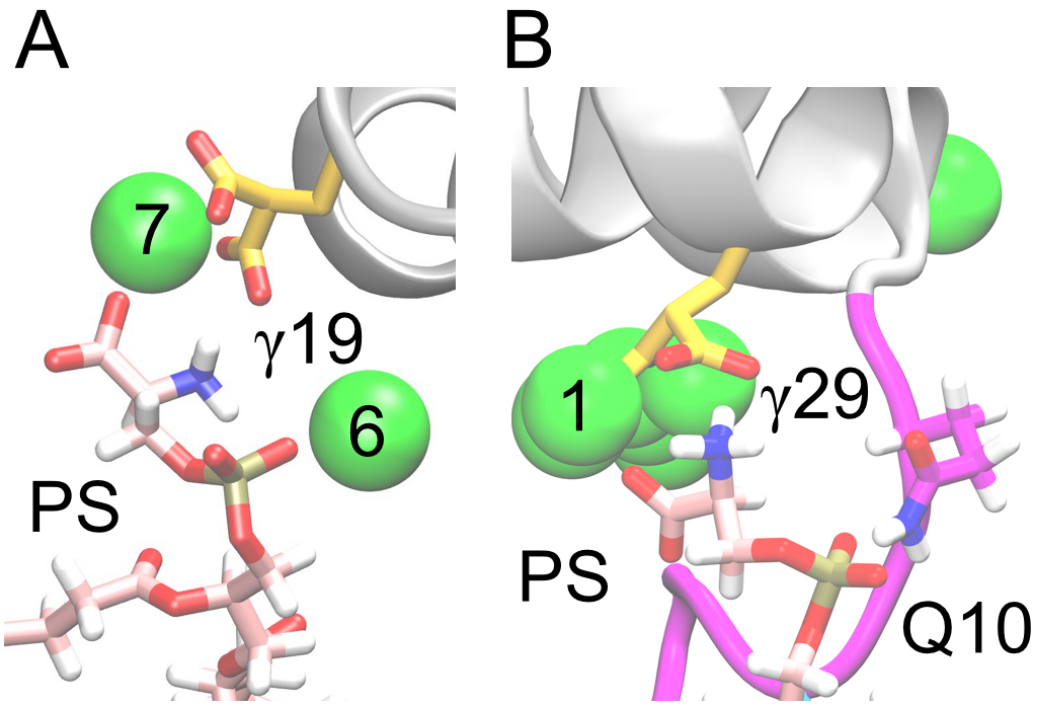


Figure 3.8: Putative PS binding sites. We present here the two putative PS binding sites which correlate most strongly with our contacts frequency data (Figure 3.7). A) Putative site 1: interaction between PS carboxy and Ca-7, amino and GLA19, and phosphate and Ca-6. B) Putative site 2: interaction between carboxy and Ca-1, amino and GLA29, and phosphate and GLN10. the charged residues close to the keel primarily form phosphate contacts, while those surrounding the Ca^{2+} ions participate in interactions with amino and, more rarely, caboxy charged groups (SER23, ARG28).

tive PS-specific binding sites we consider most likely based on the performed contact analysis are shown in Figure 3.8 These are interactions between: (i) PS, GLA19, Ca-6 and Ca-7; and (ii) PS, GLA29, GLN10 and Ca-1 (Figure 3.8). Additional PS binding modes observed in our simulations are shown Figure 3.9. Each mode of interaction reported in Figure 3.8 and Figure 3.9 was stable for at least 50 ns in the conventional membrane simulations. In the majority of cases, a PS carboxy group interacts with a Ca^{2+} ion and the PS amino group interacts simultaneously with one of the GLA residues that coordinates that same Ca^{2+} ion.

3.4 Comparison of FX-GLA to other GLA domains

The source of the marked differences in binding affinity for PS exhibited by the GLA domains of blood clotting proteins remains to be investigated. Clearly, specific residues in FX-GLA actively engage all three important charged groups on PS lipids. However, many of these residues are identical to residues in other GLA domains which bind with low or lesser affinity to PS. For example, the FVII-GLA, which shows low lipid affinity for PS [23], has a nearly identical placement of GLA residues to FX in the portion of the domain which actively interacts with the membrane. The sole exception is GLA32, the last of the 44 residues in FX-GLA to show interaction with the membrane according to our analysis, and which in FVII is LYS32. The sequence identity between the two domains is 55%, with half of the dissimilar residues located in 33-44, a helix that does not interact with the membrane. According to our analysis, residues which are not identical between the two domains and which are involved in PS charged group interactions for FX-GLA residues are SER3, LYS9, GLN10, and GLA32. In FVII-GLA, these residues are ALA3, ARG9, PRO10, and LYS32. (Figure 3.10). The change from GLA to LYS at position 32 reverses the charge at that position, changing it from a -2 to a +1. Interestingly, FIX, known like FX to have a high PS specificity [10], also has a GLA residue in this position, while protein C, which like FVII binds with higher affinity to phosphatidic acid but with low affinity to PS [17] has a GLN residue in the same position. This residue would be a good target for mutational studies, as based on our analysis as it may be involved in PS-specific interactions but is not involved in the structural

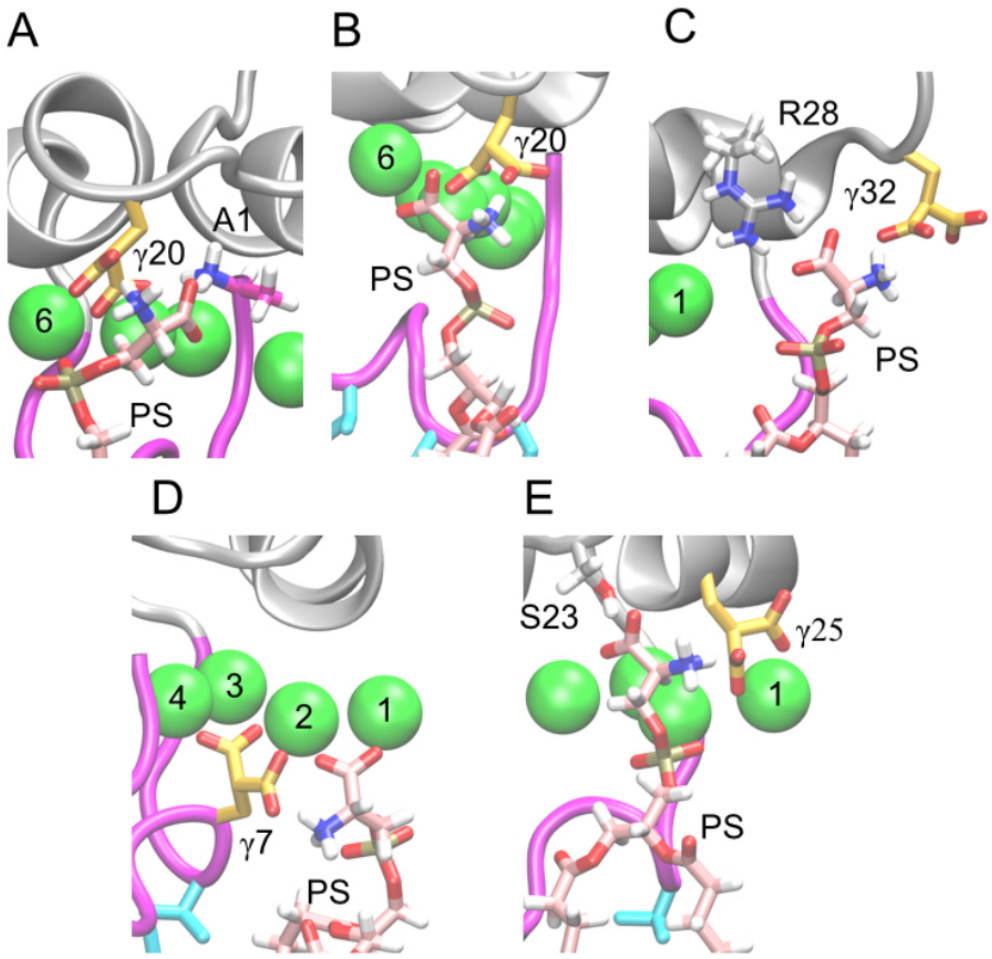


Figure 3.9: Phosphatidylserine (PS) binding modes. A–E show various PS binding modes we identified in which two or more charged groups on the same lipid interact with the protein/Ca^{2+}. Specific residues interacting are (A) N-terminal (ALA1) with carboxy, GLA420 with amino, and Ca-6 with phosphate, (B) GLA20 with amino, Ca-6 with carboxy, (C) GLA32 with amino, ARG28 with carboxy, (D) GLA7 with amino, Ca-1 and 2 with carboxy, (E) GLA25 with amino and SER23 with carboxy.

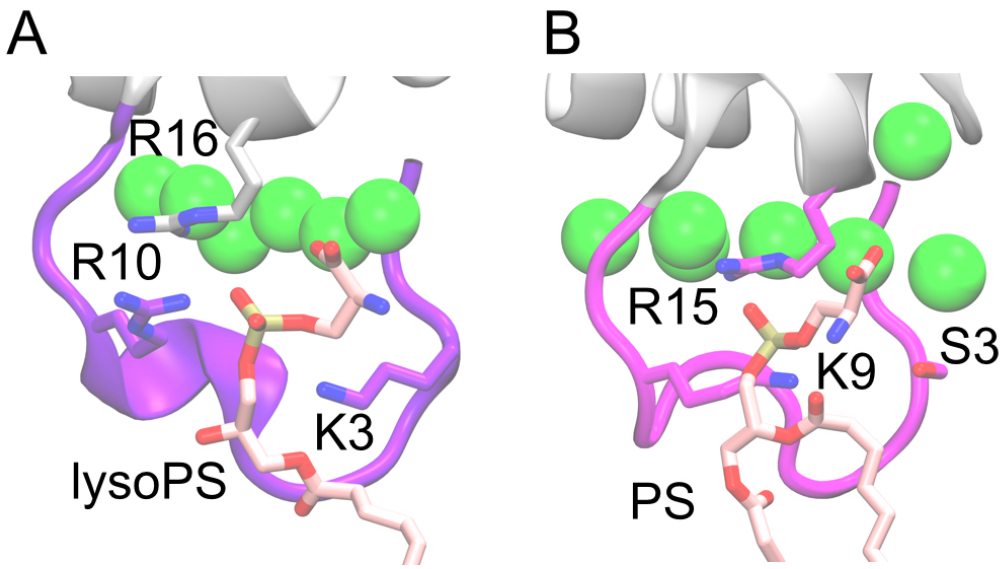


Figure 3.10: Comparison of phosphate binding site for A) prothrombin-GLA co-crystallized with a single lysoPS (PDB 1NL2) [52], and B) FX-GLA bound to PS from one of our conventional membrane simulations. Binding of PS is similar in both cases, with the equivalent starboard face residues (ARG10 and ARG16 on prothrombin-GLA, LYS9 and ARG15 on FX-GLA) interacting with the phosphate group. SER3 in FX-GLA does not interact with the phosphate, however, unlike LYS3 in the prothrombin structure.

coordination of Ca^{2+} ions.

This change of SER3 and GLN10 in FX to hydrophobic residues in FVII may change the pattern of phosphate-protein interactions, which could potentially also influence favorability of PS binding to the protein. Interestingly, one of the equivalent residues on the prothrombin GLA domain was found in a combined NMR and X-ray crystallography study [52] to bind PS head groups. The lipid carboxy group interacts with Ca^{2+} ions while the phosphate group binds to the charged patch formed by LYS3, ARG16, and ARG10 [52] (Figure 3.10) The PS pocket on the prothrombin-GLA is located in nearly the same position as the SER3, LYS9, GLN10, ARG15 phosphate binding patch found on our FX-GLA domain, although the SER3 side chain is likely too short to participate with LYS9 and ARG16 in simultaneously binding a PS as in the prothrombin-GLA domain. It is possible that these sites on the various GLA domains form a non-specific phosphate binding patch which is involved in modulating the orientation of the protein. Mutations at this site could thus indirectly affect lipid affinity, and even activation rates of FX when bound to PS.

3.5 Conclusions

In this chapter, we present the first structural model for the membrane-bound state of FX-GLA, the membrane-binding domain of a critical coagulation protein, and propose two putative PS-specific binding sites based on the simulation results. Experimental data showing binding of lysoPS to the prothrombin GLA domain are consistent with our results [52], as the single lipid in this experiment bound with its phosphate group to a site identified on FX-GLA as a phosphate binding site. Further experimental studies of the effects of mutating GLA32 and residues of the phosphate binding pocket will be informative as they can validate the proposed model of binding in this study. Mutation of additional GLA residues might also be useful if effected without modifying the FX-GLA fold and Ca2+ coordination. It is possible that GLA7 and GLA25 provide sufficient structural support to coordinate Ca^{2+} ions 1 and 2 without GLA29.

In order to enhance sampling of PS–GLA interactions, an artificially high PS concentration (100% PS) was used in combination with a membrane

representation with enhanced lipid dynamics (HMMM). Provided there is sufficient computational power, future studies will examine the protein in a more native membrane environment, composed of a mixed lipid bilayer and taking into account the impact of bulk solution divalent ions such as Ca^2+ and Mg^{2+}, thereby improving the realism of the environment and its relevance to physiological conditions. Another natural extension of the present study is to employ models of human coagulation factors.

Despite these technical limitations, the structural model developed in the present study will serve as the first structural framework, allowing for more rationally designed experiments investigating, for example, the impact of specific mutations on the activity of FX, or the development of small molecules for potential drug targets that would block lipid-specific binding of FX-GLA domains with the goal of diminishing thrombosis.

4 LIPID SPECIFICITY OF COAGULATION FACTOR VII*

4.1 Introduction

It is well established that lipid composition of the platelet surface has a profound effect on blood clotting [13, 16, 22, 214]. In order to initiate clotting, negatively charged lipids are exposed to the outer surface of the plasma membrane, allowing for recruitment of coagulation factors. Factor VII (FVII) is the catalytic subunit of the enzyme that triggers the extrinsic cascade of blood coagulation [13, 207]. The formation of the extrinsic complex, consisting of tissue factor, FVIIa, and FX, is integral to generation of the explosive burst of thrombin that allows a clot to form [207, 215].

FVII membrane binding depends on a specialized, post-translationally modified domain known as the GLA domain. Coagulation factors like FVII which contain GLA domains are referred to as vitamin K dependent, as this vitamin is necessary to synthesizing γ-carboxyglutamic acid (GLA) residues in these domains [216]. GLA domain containing coagulation factors are members of the extrinsic, intrinsic, and anticoagulant pathways [207, 217]. The negatively charged GLA residues allow for coordination of multiple Ca^{2+} and Mg^{2+} ions, which are critical for the correct folding of the structure and its interaction with negatively charged surfaces [53, 218]. The GLA domains of vitamin K-dependent coagulation factors are structurally and sequentially similar to each other (Figure 4.1), but they exhibit widely varying affinities to anionic lipids [17, 219, 10, 220, 221]. The FVII GLA domain (FVII-GLA), for example, is structurally similar to the GLA domains of factor X (FX-GLA) and prothrombin, but FVII has only weak affinity to phosphatidylserine (PS) while FX and prothrombin show strong affinities to this lipid. FVII and protein C, on the other hand, has been found to have con-

*This chapter is reproduced from the article:
(1) M. P. Muller, J. H. Morrissey, and E. Tajkhorshid, "Molecular View into Preferential Binding of the Factor VII GLA Domain to Phosphatidic Acid," (*Submitted*).

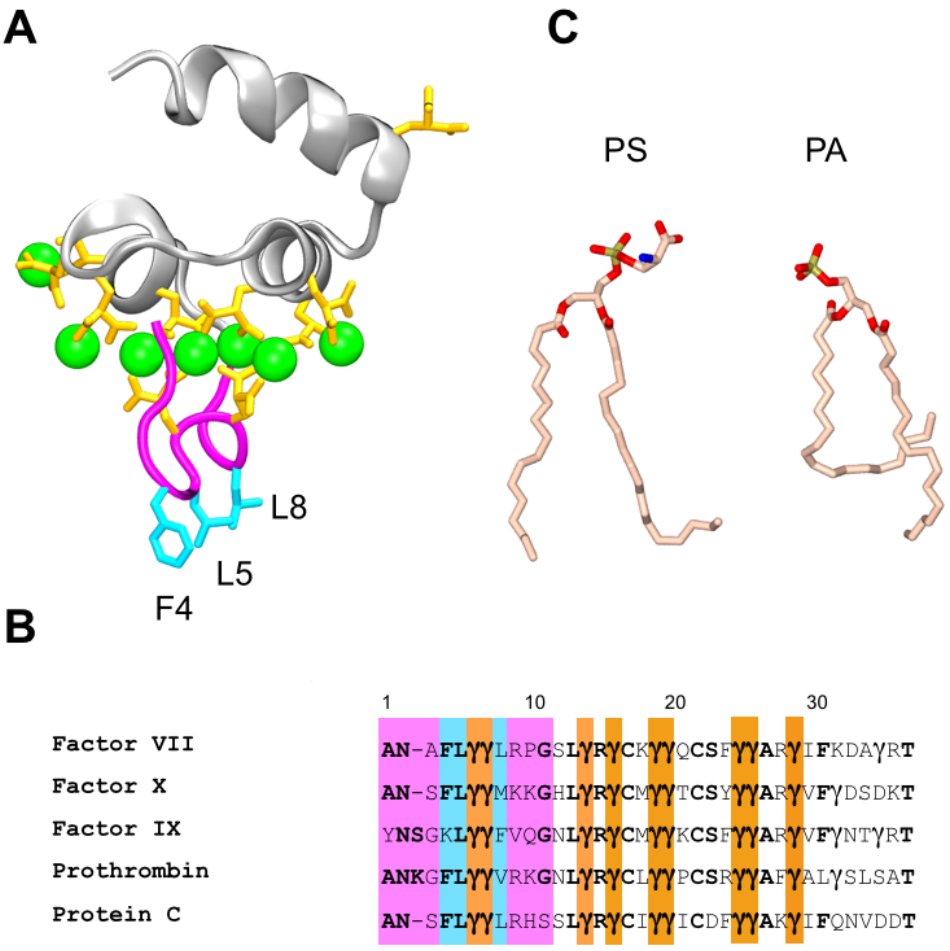


Figure 4.1: FVII-GLA, which is structurally and sequentially similar to other coagulation factor GLA domains, binds preferentially to PA-containing lipid bilayers. A) Structure of the FVII-GLA domain [59]. Key elements shared by all GLA domains are highlighted. Specialized GLA residues each with a -2 charge are shown in gold. The ω loop is shown in magenta. Hydrophobic "Keel" residues, which penetrate into the membrane core, are shown in light blue. Ca^{2+} ions are shown in green. B) PS and PA lipid structures. FVII shows preferential binding to PA-containing liposomes as compared to those with PS. Phosphorus atoms are shown in gold, oxygen in red, carbon in beige, and nitrogen atoms in blue. Hydrogen atoms are not drawn for simplicity. C) Comparison of sequences of GLA domains from various coagulation factors. GLA residues (γ) in the sequences are highlighted in gold, keel residues in blue, and the ω-loop residues in magenta.

siderably higher affinities to membranes containing phosphatidic acid (PA) compared to PS. For example, FVIIa has been shown to exhibit increased enzymatic activity when bound to liposomes containing PA [23].

PA is a rare anionic lipid, making up approximately 1-2 % of the cell membrane, compared to 2-10 % for PS [222]. PA is the simplest anionic phospholipid and contains a single negatively charged phosphate group in its structure, although the overall charge can vary from -1 to -2 at physiologic pH (Figure 4.1). PS carries a net charge of -1, but it contains multiple charged groups in its structure, a positively charged amino group, a negatively charged phosphate group, and a negatively charged carboxy group.

Here, we use molecular dynamics simulations to study the membrane binding of FVII-GLA to PA and PS-containing lipid bilayers (with the balance being phosphatidylcholine, PC). A number of MD studies have previously probed GLA domain lipid interactions using full lipid models [223, 220, 53], but did not sufficiently reproduce spontaneous membrane binding of the studied domains. We use a specialized membrane model, termed the highly mobile membrane mimetic (HMMM), in which lipid tails are truncated and replaced with a solvent, thus allowing for substantially increasing lipid diffusion, enhancing the sampling of lipid-protein interactions, and capturing their spontaneous binding [169]. This method has been used previously in a wide variety of membrane systems [169, 170, 171, 172, 173, 174, 175, 176, 177, 178, 74, 75, 179, 180, 181, 182, 162, 183, 184, 185, 186] including FX-GLA and FVII-GLA membrane binding [179] to 100% PS bilayers [169]. To our knowledge this is the first simulation study to examine binding of FVII-GLA to PA-lipids. As the atomic-level MD model requires a fixed charge to be assigned all atoms, PA lipids were assigned a -1 charge as the most likely at physiological pH citeKOOI2005. We generate 48 independent systems of FVII-GLA over PA:PC and PS:PC membranes, and allow FVII-GLA to spontaneously bind from solution. Each membrane patch is either 1:1 PA:PC or 1:1 PS:PC, consistent with concentrations used in experimental studies [23]. This allows us to characterize the membrane binding mode of FVII-GLA domain and assess possible molecular level sources of its differential affinity to PA and PS-containing membranes.

4.2 Methods

4.2.1 Spontaneous membrane-binding simulations

A crystal structure of FVII-GLA from Protein Data Bank entry 1DAN [59] was used as the starting structure for all GLA domain simulations. Using the PSFGEN plugin of VMD [213], an N-terminal ammonium group and a C-terminal N-methylamide C-terminal group were added, along with hydrogen atoms. The neutral N-methylamide group was chosen because the C-terminus of the FVII-GLA in the full-length FVII protein would be connected to the first epidermal growth factor (EGF) domain, and a C-terminal patch with a default charge would introduce an artificial charge to this region. FVII-GLA was next placed in a $60\times60\times60\,\text{Å}^3$ water box using the SOLVATE plugin of VMD and neutralized with Na^+ ions added using the AUTOIONIZE plugin of VMD. This solution system was then equilibrated for 1 ns after 2,000 steps of energy minimization.

Our HMMM membrane construct was used in the binding simulations to increase lipid dynamics and allow for increased sampling and GLA domain insertion into the membrane. Three independent 1:1 PA:PC HMMM lipid patches as well as three independent 1:1 PS:PC HMMM patches were generated. PA lipid headgroups were assigned a charge of -1. Each patch was $50\times50\times114\,\text{Å}^3$ in size and was constructed and solvated using the HMMM Builder [180] in CHARMM-GUI [224, 225].

Nine independent orientations of FVII-GLA (a rotation of 15° in the xz or yz plane) were generated for each of the three PA:PC and three PS:PC membrane patches. In the beginning of each simulation FVII-GLA was placed at least 3 Å above the membrane patch. This resulted in generation of 27 unique FVII-GLA systems for each lipid composition. Each system was neutralized with Na^+ ions using the AUTOIONIZE plugin of VMD. The final systems were each approximately 24,000 atoms. The z-position of HMMM short-tailed lipids was restrained harmonically ($k = 0.1\,\text{kcal/mol/Å}^2$ applied to C21 and C31 lipid tail atoms) throughout the simulations in order to more closely reproduce the natural distribution of the head-group atoms. Each system was then simulated for 100 ns collecting 2.7 μs of sampling for both FVII-GLA PS:PC and FVII-GLA PA:PC systems.

4.2.2 Conversion of HMMM membranes to conventional full membranes

In order to ensure that the final FVII-GLA membrane-bound model was not unduly affected by the use of the HMMM as opposed to a full-tailed lipid construct, all bound systems were converted to conventional (full-tailed) membrane, using a PDB generated from the final frame of each 100 ns membrane-binding HMMM simulation. We used 1-palmitoyl-2-oleoyl (PO) tails to produce complete POPC, POPA, and POPS lipids. Conversion to a full-tailed membrane was only made for the systems in which the keel penetrated below the level of phosphorus atoms in the membrane. A full-tailed lipid was overlaid onto each short-tailed lipid, and coordinates of the missing atoms were then copied from the full-tailed lipid, thus allowing for head-group interactions obtained from the HMMM simulations to be preserved. The layer of 1,1-dicholoroethane (DCLE) solvent molecules were then removed from the center of the HMMM bilayer. After a 2,500-step energy minimization, each membrane-bound FVII-GLA system was simulated in a full membrane for additional 100 ns.

4.2.3 Simulation protocols

All simulations were performed using NAMD [86, 226] and the CHARMM36 [111] force-field parameters. Parameters of GLA residues were those previously developed by our lab [53] with atom types renamed as appropriate for the conversion from CHARMM27. The NPT conditions and TIP3P water model was used for all simulations. Constant pressure was maintained at a target of 1 atm using the Nosé-Hoover-Langevin piston method [164]. Constant temperature was maintained by Langevin dynamics with a damping coefficient, γ, of $1\,\mathrm{ps}^{-1}$ applied to all atoms [165]. All simulations used a timestep of 2 fs and a temperature of 310 K. The particle mesh Ewald (PME) method [167] was used for long-range electrostatic calculations with a grid density greater than $1/\,\mathrm{Å}^3$. Non-bonded interactions were cut after 12 Å with a smoothing function after 10 Å. System setup, visualization, and analysis were carried out using VMD [213].

4.2.4 Analysis

Depth of the GLA domain insertion into the membrane was assessed by calculating the z components of center of mass (COM) of these groups, individually, relative to the midplane of phosphate groups (phosphorus atoms and four bonded oxygen atoms) of lipids: (1) backbone atoms in residues 4, 5, and 8 in the keel of FVII-GLA and (2) Ca^{2+} ions 1-7 (Ca-1 to Ca-7). Similarly, the z components of lipid carboxy groups (carbon and two bonded oxygen atoms), lipid amino groups (nitrogen and three bonded hydrogen atoms) were also calculated relative to the phosphate plane.

Orientation of the GLA domain relative to the membrane plane was assessed using two angles defined in two perpendicular planes, termed "port-to-starboard" and "bow-to-stern." The bow-to-stern angle is the angle between the membrane normal (z-axis) and the vector connecting the COM of Ca-1 to 3 and Ca-4 to 6. These are the same measures as previously used to analyze orientation of the domain on the membrane in our previous study of the FX-GLA domain membrane binding [179]. The port-to-starboard angle is the angle between the membrane normal and the vector connecting the COM of backbone atoms in residue 13, 14, 29, and 30 and the COM of backbone atoms in 19, 22, 23 and 24. These two sets of residues represent opposite ends of the two helices directly above the level of the coordinated Ca^{2+} ions. Thus, the bow-to-stern angle measures the angle between the stern and bow ends of FVII-GLA and the membrane normal, while the port-to-starboard angle represents the tilt of the port and starboard ends relative to the membrane normal.

Contacts between lipid head-group charged moieties and FVII-GLA (bound Ca^{2+} or protein residues) were calculated based on distance cutoffs which were determined using radial pair distribution function calculations, as in our previous study of the FX-GLA domain [179]. Number of contacts here is the number of atoms on one group (e.g., a carboxy group) within a specified cutoff of the second group (e.g., a specific protein residue). Distance between all protein and Ca^{2+} atoms and phosphate, amino, and carboxy atoms were calculated, and distances equal to or less than the cutoff were determined to be a contact. The radial pair distribution function was calculated separately for distances between the residue type of interest (protein or Ca^{2+} ions) and the lipid head-group moiety of interest (carboxy, amino, or phosphate). The

calculation sampled every 10th frame of the 100 ns full-membrane simulations. For protein residues and Ca^{2+} ions, all atoms were included in the contact calculation. Cutoffs determined were 2.50 Å for phosphate-Ca^{2+}, 3.00 Å for carboxy-Ca^{2+}, 2.25 Å for phosphate-protein, 2.50 Å for carboxy-protein, and 2.25 Å for amino-protein contacts.

4.3 Binding of FVII-GLA to PS and PA-containing lipid bilayers

We performed 54 independent HMMM simulations to capture spontaneous binding of FVII-GLA to anionic membranes, 27 copies of PA:PC (1:1) membranes, and 27 PS:PC (1:1) membranes. The FVII-GLA domains bound successfully to 24 of 27 PA:PC bilayers, while in only 17 out of 27 PS:PC systems did FVII-GLA bind successfully to the membrane. As described in Methods, binding was deemed successful if the keel residues of FVII-GLA penetrated beneath the phosphate level in the lipid head groups. Figure 4.2 shows examples of binding poses of FVII-GLA to PA:PC and PS:PC bilayers, as well as a graphical representation of the membrane depth of each FVII-GLA over the course of both the HMMM and full-membrane simulations. As shown in Figure 4.2, once the keel of FVII-GLA penetrated below the phosphate level, it remained stably bound for the remainder of the simulation.

Interestingly, we found a deeper penetration of FVII-GLA into PA:PC membranes, compared to the PS:PC ones. When bound to the PA:PC membranes, the keel height (averaged over the last 10 ns of the full-membrane simulations) was 17.7 ± 1.0 Å and the GLA Ca^{2+} ions were at 27.2 ± 0.8 Å, where the average height of phosphate atoms in the membrane was 25 Å. When bound to the PS:PC lipids, the average keel height was 19.7 ± 1.7 Å, and the Ca^{2+} ions were at 29.3 ± 1.7 Å. FVII-GLA is thus binding with approximately 2.0 Å deeper penetration into PA:PC bilayers compared to PS:PC bilayers, as well as to a more converged depth. This is particularly notable as there was more sampling for FVII-GLA when bound to PA:PC.

The orientation of the FVII-GLA domain relative to the membrane was assessed using the "bow-to-stern" and "port to starboard" tilt angles previously developed in our study of the FX-GLA membrane binding [179]. As described in Methods, the bow-to-stern tilt assesses the angle between the

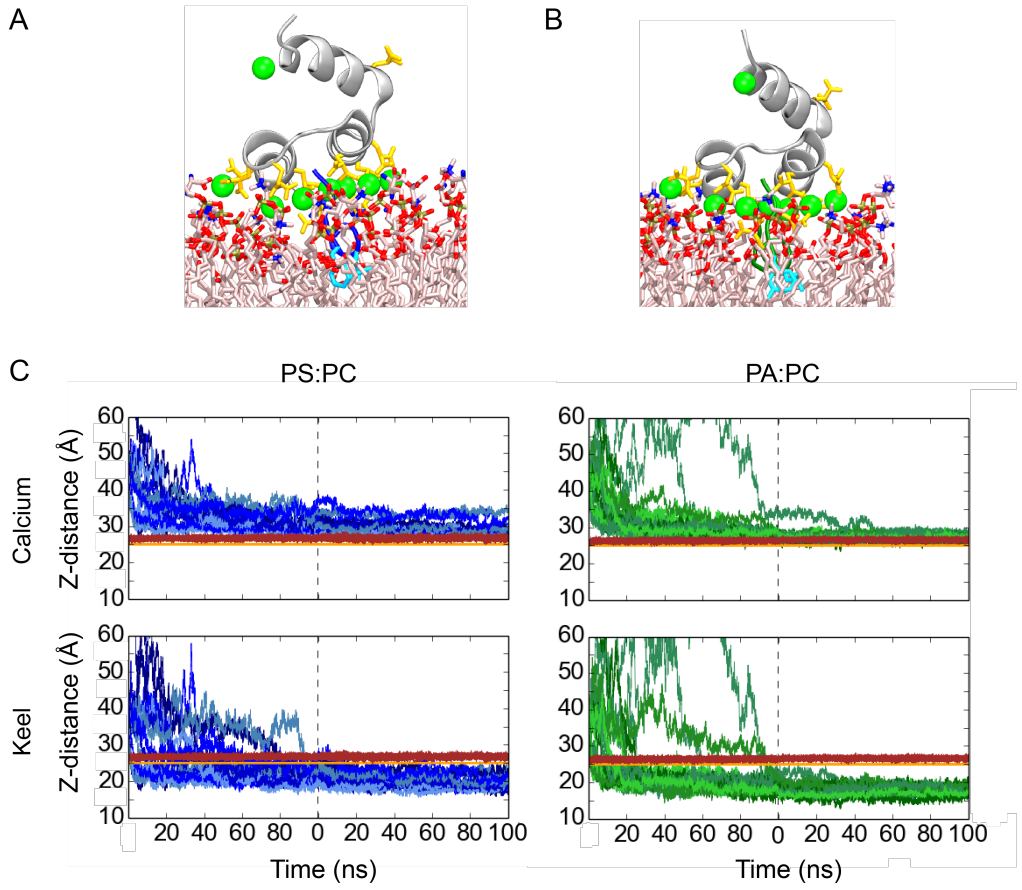


Figure 4.2: FVII-GLA binds spontaneously to PS:PC and PA:PC membranes. A) Representative FVII-GLA bound to a PS:PC full membrane. Ca^{2+} ions are shown in green, GLA residues in orange, keel backbone in dark blue, and keel residues in light blue. For lipids, carbon is shown in beige, oxygen in red, nitrogen in blue, and phosphorus atoms in gold. B) Representative FVII-GLA bound to a PA:PC full membrane. Same color scheme as in A is used. C) Membrane depth of insertion of the FVII-GLA measured by the z-position of the COM of GLA Ca^{2+} ions (upper panel) and the COM of keel CA atoms (lower panel) relative to the membrane. Different shades of blue (for PS:PC simulations) and green (for PA:PC simulations) are used to show different traced COMs of the FVI-GLA components, orange shows the position of the phosphate layer used as the reference ($z = 25$ Å), and maroon shows the z-component of the positions of the choline nitrogens in lipids. The dashed vertical line indicates the transition from HMMM simulations to full-membrane simulations.

membrane normal vector and a vector through the FVII-GLA Ca^{2+} ions 1-6.

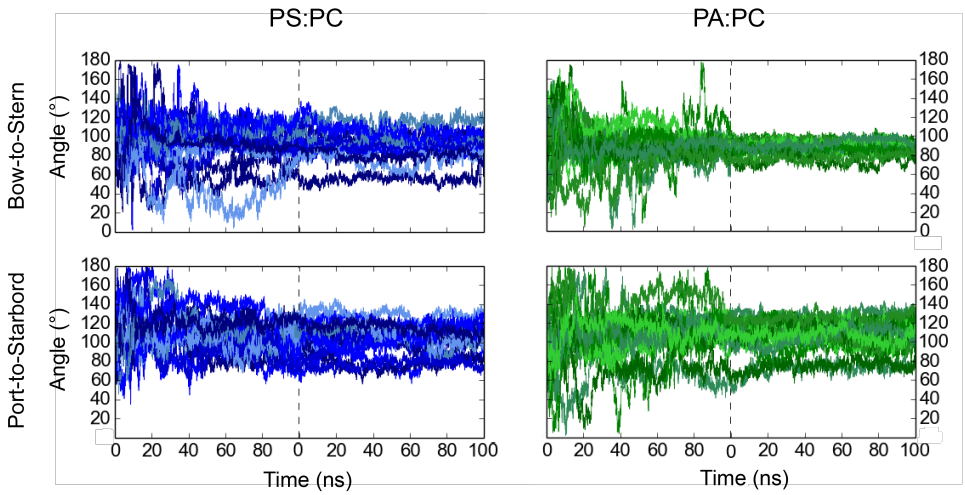


Figure 4.3: The orientation of the membrane-bound FVII-GLA relative to the lipid bilayer. The "Bow-to-stern" orientation is shown in the upper panels and the "port-to-starboard" orientation in the lower panels. Traces for FVII-GLA binding to PS:PC bilayers are shown in different shades of blue, and those for binding to PA:PC bilayers are shown in green shades. The vertical dashed line shows the transition from the HMMM simulations to full-membrane simulations.

The port-to-starboard tilt, in the other hand is the angle between the membrane normal and a vector that runs parallel to the α-helices in the GLA domain closest to the keel. Figure 4.3 shows the time series of these angles throughout the HMMM and full-membrane simulations.

In addition to the difference in depth of membrane insertion, the orientation of the FVII-GLA upon binding to the membrane is also different between the PA:PC and PS:PC membranes. We found an average "port-to-starboard" tilt angle (averaged over the last 10 ns of the full-membrane simulations) of $109 \pm 15°$ for the FVII-GLA domain when bound to PA:PC membranes, and $105 \pm 14°$ when bound to PS:PC bilayers. The bow-to stern tilt angle for the PA:PC membranes was $88 \pm 5°$, while for the FVII-GLA domain bound to the PS:PC membranes it was $91 \pm 13°$. While the average orientation between the PA-bound and PS-bound GLA domains is similar, PS-bound GLA domains took on a wider range of "bow-to-stern"' tilts in the same 10 ns of simulation. When bound to PA-containing bilayers, the standard deviations of both "port-to-starboard" and "bow-to-stern" tilts are similar to those found for the FX-GLA domain previously [179]. Although the units

may not be directly compared, the orientation of the protein on the surface of the membrane represents a softer degree of freedom is expected to fluctuate more widely compared to the depth of penetration into the membrane.

4.4 FVII-GLA lipid interactions in PA:PC and PS:PC membranes

In order to further compare the interactions of the FVII-GLA domain with the two studied anionic membranes, we quantified individual interactions between the residues in the FVII-GLA domain and any charged phospholipid moieties during the full-membrane simulations. We assessed contacts between either protein residues or bound Ca^{2+} ions and the phosphate, carboxy, or amino moieties of lipids. The average contact numbers between these lipid moieties and individual FVII-GLA residues are shown in Figure 4.4.

As the only charged group able to form strong electrostatic interactions with the FVII-GLA domain in the PA:PC bilayers is the lipid phosphate, significant interactions are found between both these groups and the positive protein residues or the bound Ca^{2+} ions of the FVII-GLA domain in the PA:PC simulations. Particularly, Ca-1 and Ca-6 (individual Ca^{2+} ions that are tightly bound to the FVII GLA domain), interact frequently with the phosphate groups, with an average contact number of 1.4 and 1.2, respectively (Figure 4.4), indicating that they interact with one or more phosphate groups for the majority of the simulation. In the simulations with PS:PC membranes, the positive sites of the protein instead interact largely with the carboxy groups of lipids, which can be attributed to the larger exposure of the carboxy groups in PS lipids, compared to the buried position of their phosphate groups. The overall number of contacts for Ca-1 and Ca-6 with PS and PC phosphate are also generally decreased to 1.0 and 0.8 (combined for both phosphate and carboxy contacts).

A possible explanation for the lesser number of interactions in PS-containing bilayers may lie in an examination of representative binding modes. Figure 4.5 illustrates representative lipid-Ca^{2+} interactions for PA and PS lipids at the end of 100 ns full-membrane simulations. Representative structures used had average binding depth, as well as average "port-to-starboard" and

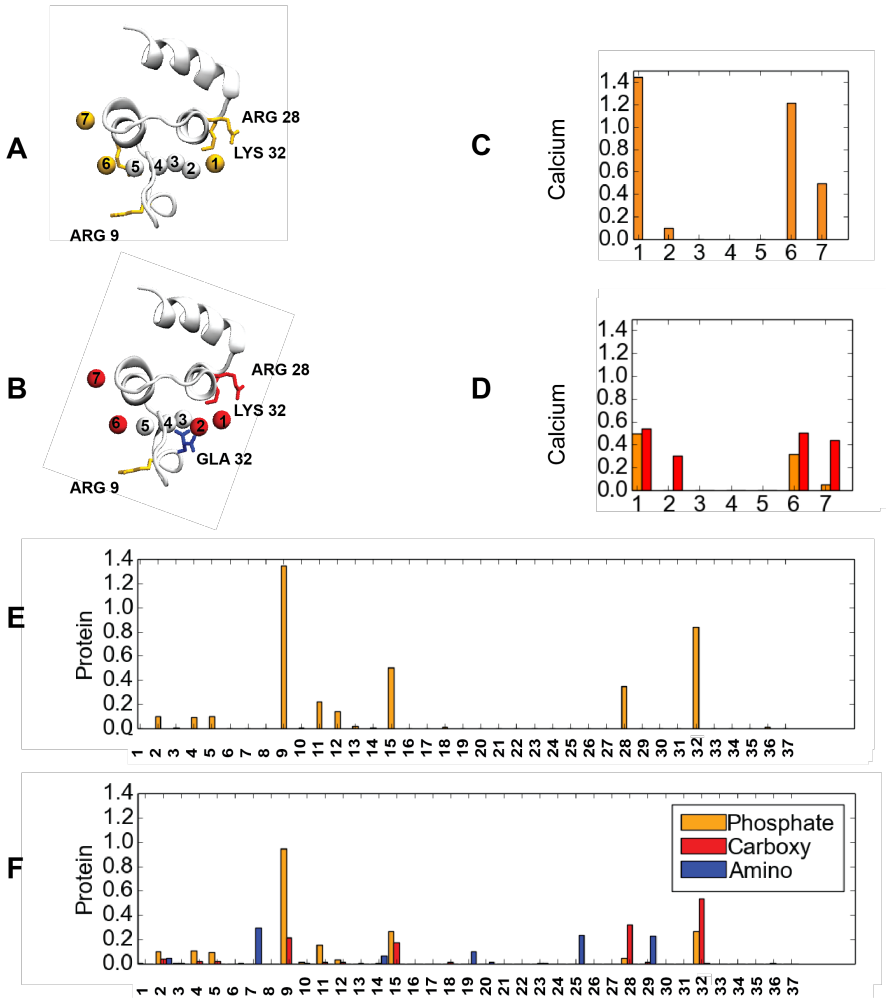


Figure 4.4: Specific contacts between lipids and FVII-GLA. A) FVII-GLA residues most frequently contacting PA phosphates are highlighted in orange. B) FVII-GLA domain with most frequent contacts to PS lipid groups are highlighted, with red indicating most frequent contacts to PS carboxy groups, yellow for most frequent contacts to the phosphate groups, and blue most frequent contacts to PS amino groups. C-F) Histograms of average contacts to lipid charged groups. Average contact values are calculated for the entire 100 ns of the full-membrane simulations. C) FVII-GLA/Ca^{2+} contacts to lipid charged groups in 1:1 PA:PC membranes. Orange indicates contact to the phosphate group of either PA or PC lipid. D) Histogram of Ca^{2+} residue contacts to lipid charged groups in 1:1 PS:PC membranes. Orange indicates contact to PS or PC phosphate groups, red indicates contact with PS carboxy groups. E) Histogram of protein residue contacts to lipid charged groups in 1:1 PA:PC membranes. Orange indicates contacts to phosphate group of either PA or PC lipid. F) Histogram of protein residue contacts to lipid charged groups in 1:1 PS:PC membranes. Orange indicates contact to PS or PC phosphate group, red indicates contact with PS carboxy group, and blue indicates contact with PS amino group.

"bow-to-stern" orientations. The interactions shown were stable for the entirety of the 100 ns full-membrane simulation. For the PA:PC membrane, three PA lipids and PC lipid contribute to Ca^{2+} binding, with two lipids interacting with each of Ca^{2+} residue 1 and 6. In contrast, three PS lipids interact with Ca^{2+} ions, one each with Ca^{2+} 6 and 7, and the third coordination both Ca^{2+} residue 1 and 2. While multiple phosphate groups coordinate with Ca^{2+} ions in the PA:PC binding simulation, no more than on PS lipid carboxy group interacts simultaneously with each Ca^{2+} ion. It is possible that the truncated PA structure allows both PA and PC phosphate to pack more closely around the Ca^{2+} ions, thus increasing the average number of total lipid contacts for Ca^{2+} ions in PA:PC membrane simulations.

For protein-lipid interactions, residues with placement at a greater height from the membrane are found to have a higher number of carboxy interactions when in PS:PC membranes than do residues those on the ω-loop. In the PA:PC membranes, the protein residues that have significant interactions with the lipid phosphate groups include ARG9, ARG15, ARG28, and LYS32. In the PS:PC membranes, only ARG9 and ARG 15 have a majority of lipid interactions charged-group interactions with phosphate, while ARG28 and LYS32 have a majority of interactions with carboxy groups. This may be due to the shallower binding depth of FVII-GLA in PS:PC compared to PA:PC, which would make these two relatively membrane-distant residues less accessible to phosphate groups.

Representative membrane-bound configurations of FVII-GLA are shown in Figure 4.6 where important lipid-protein interactions are highlighted. In the snapshot shown in Figure 4.6 for FVII-GLA bound to PA:PC membranes, the basic side chain of ARG28 directly interacts with the phosphate group of a PC, while the LYS32 residue is in contact with a PA phosphate group. In the representative snapshot shown in Figure 4.6A, PS carboxy groups interact with both ARG28 and LYS32. In Figure 4.6C-D, ARG9 is shown to bind to PA phosphate and PS phosphate, respectively.

In examining sequence conservation between GLA-domains among these residues, we find a high degree of conservation for residue 28, which is ARG for FVII, FX, FIX, LYS for protein C. Thus, the residue is a basic amino acid for both PS-specific FX and FIX and PA-specific FVII and protein C. This appears to be a fairly non-specific interaction which does not seem to be an element of the lipid-specific affinity to PA for FVII-GLA. Similarly, residue

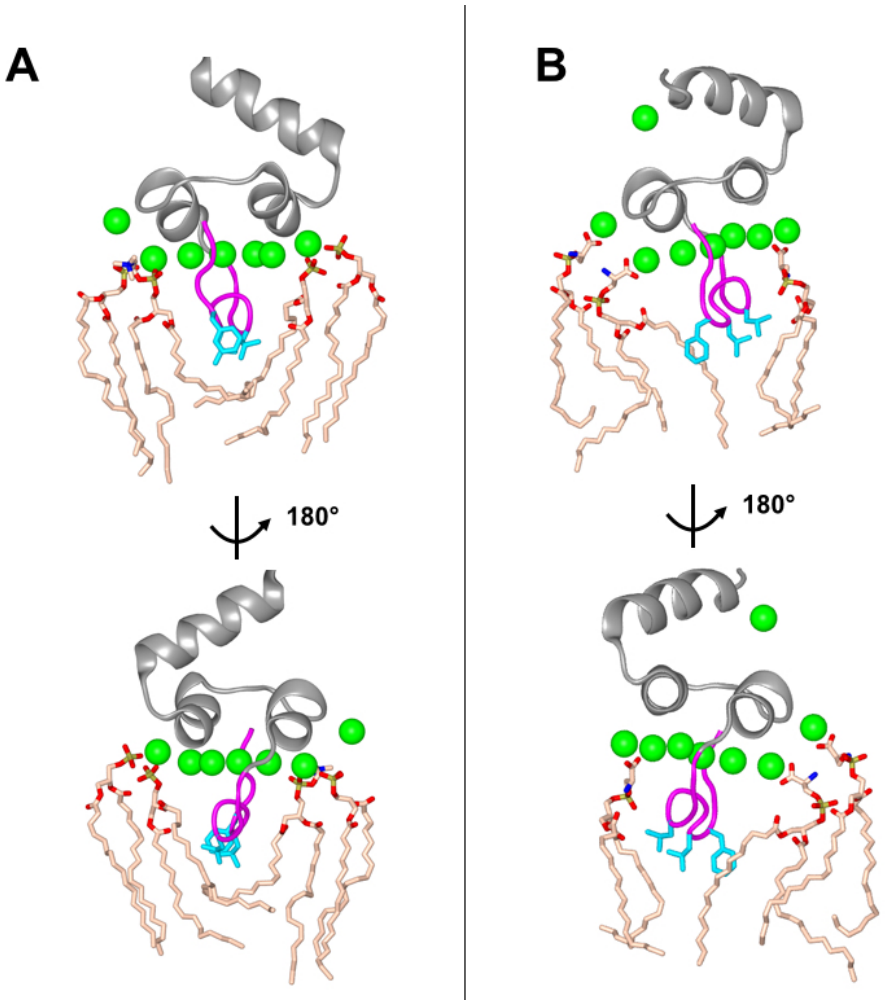


Figure 4.5: Characteristic lipid-Ca^{2+} interactions from full-membrane simulations for FVII-GLA. See methods for distance cutoffs. Ca^{2+} ions are shown in green, oxygen atoms in red, nitrogen in blue, phosphorus in gold. Lipid carbon atoms are shown in beige. ω-loop of the GLA domain is drawn in magenta, and keel residues in light blue. Remainder of FVII-GLA shown in silver. A) Lipids interacting with FVII-GLA Ca^{2+} in PA:PC membrane-binding simulation. One PA and one PC interact with Ca6 through their phosphate groups, while two PA lipids interact with Ca-1. B) Lipids interacting with FVII-GLA Ca^{2+} in PS:PC membrane-binding simulation. One PS lipid interacts with each of Ca-6 and Ca-7, while one PS lipid interacts with Ca-1 and Ca-2 simultaneously. All PS lipids interact interact through their phosphate groups.

9 is positively charged LYS for FX, and ARG for prothrombin, FVII and protein C. In our previous study of FX, LYS9 interacted similarly to ARG9 for FVII here, providing a phosphate anchor. Residue 15 is a conserved ARG for all five GLA domains we assessed (Figure 4.1). For residue 32,

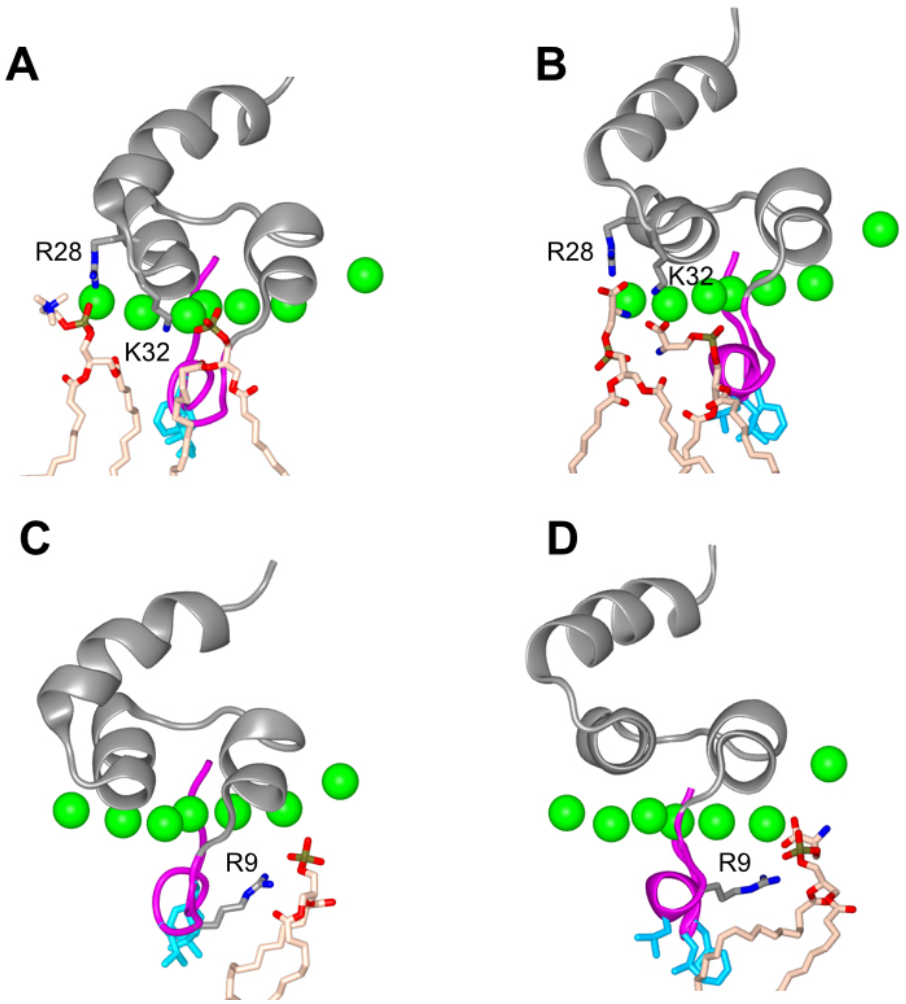


Figure 4.6: Characteristic lipid-protein interactions from simulations with full-length fatty acyl chains. Ca^{2+} ions are shown in green, oxygen in red, nitrogen in blue, and phosphorus in gold. Lipid carbon atoms shown in beige. ω-loop is drawn in magenta, and keel residues in light blue. Remainder of FVII-GLA is shown in silver. Four simulation snapshots are shown in A-D. A)PA:PC membranes: R28 and K32 residues interact with PC and PA lipid phosphate groups. B) PS:PC membranes: R28 and K32 residues interact with PS lipids through their carboxy groups. C) PA:PC membranes: K9 interacting with PA lipid through the phosphate group. D) PS:PC membranes: R9 interacts with phosphate of PS, while the carboxy group of same lipid interacts with Ca6.

however, sequences are clearly different between PA-specific and PS-specific GLA domains. In FX, FIX, and prothrombin, residue 32 is a negatively charged GLA residue. In PA-specific FVII and protein C, however, it is either a positively charged LYS or a polar GLN residue. While electrostatically, the

positive charge of a lysine can interact favorably with both lipid phosphate (present in both PA and PS) and lipid carboxy groups (only present in PS), judged by the sequence, the LYS in this location in the FVII GLA domain is a strong candidate for mediating the preferential binding to PA.

4.5 Conclusions

Experimental measurements have established improved binding of FVII to PA-containing membranes relative to PS-containing membranes [23]. In order to investigate the molecular basis of this lipid effect, we have used molecular dynamics simulations to capture spontaneous binding of the FVII-GLA domain to PA:PC (1:1) membranes and compared the process to that in PS:PC (1:1) membranes. Our analysis suggests that FVII-GLA binds the membrane with deeper penetration to PA:PC lipid bilayers as compared to PS:PC ones. FVII-GLA keel residues penetrated $\sim$2.0 Å deeper below the level of the lipid phosphates when bound to PA-containing bilayers as compared to PS-containing bilayers. The binding of the FVII-GLA domain to the PA-containing membranes was also more homogeneous than in PS-containing membranes. The standard deviation in the z-position (membrane normal) was narrower for the PA-associated FVII-GLA domains than the PS-associated ones. In addition, the standard deviations of the tilt angles (orientation) exhibited by the FVII-GLA domain were also narrower for the PA:PC simulations for the "bow-to-stern" angle. FVII-GLA also interacts with different phospholipid moieties in the PA:PC bilayers as compared to PS:PC bilayers. In the PA-containing bilayers, FVII-GLA residues interacted primarily with the phosphate groups and with a higher average number of lipid groups. This was particularly notable when examining contacts between the lipid charged groups and FVII-GLA Ca^{2+} ions.

A possible explanation for the obtained findings may be differential steric hindrance of the two bilayers. PA, with its truncated phosphate head-group, is quite short compared to PS, which is more comparable in head group size to PC. The bulky choline head-group is likely to sterically hinder the access of charged FVII-GLA residues to available phosphate moieties. In the PA:PC membranes, the steric hindrance incurred by the choline groups may be reduced, as the truncated PA head-group allows it more space to

move within the head group layer. FVII-GLA Ca^{2+} ions are thus able to penetrate more deeply into the membrane and coordinate with neighboring PC phosphate groups . In the PS:PC bilayers, the carboxy groups of the PS lipids are more available to coordinate the Ca^{2+} ions but may be unable to coordinate charged FVII-GLA residues as effectively. This would explain the narrower range of "bow-to-stern" tilts visited in the PA-containing bilayers, and why the tilt is dependent on the relative position of the Ca^{2+} ions.

The position of the tightly bound Ca^{2+} ions in GLA domains are structurally similar, however. What differences could account for the differential affinity between FVII and FX? One notable sequence difference between the FVII-GLA domain and other GLA domains which are generally more PS specific, lies in the keel residues which penetrate into the core of the membrane (see Figure 4.1). At residue 8, both FVII and protein C, which have high PA affinities, have a LEU. In FX and prothrombin the equivalent residues are MET and VAL, respectively, which are bulkier hydrophobic residues. The other two side chains that make up the keel are otherwise identical (PHE and LEU) in these three proteins. The PA-preferential GLA domains thus have a somewhat less bulky keel residue than PS-preferential ones, which may necessitate a deeper binding depth for equivalently strong hydrophobic interactions.

A limitation in our study is that PA lipids are modeled with a fixed single negative charge. As one of the deprotonation pKa values for PA is close to physiological pH [227, 228], it is possible that *in vivo* PA exhibits a mix of -1 and -2 charged states, especially upon binding to the FVII-GLA domain (or other largely charged species). This aspect has not been taken into account in our fixed-charge simulation system. Our conclusions are also limited by use of a simplified two-lipid bilayer system. Despite the limitations imposed by the model, we believe the molecular insights into lipid-protein interactions provided by the study can be used as a basis for further experimental exploration of protein-lipid modulation of FVII ad other clotting proteins.

5 MODELING FORMATION OF THE EXTRINSIC COMPLEX

5.1 Introduction

Formation of the extrinsic coagulation complex is crucial to initiating the series of reactions which result in the burst of thrombin necessary to form a blood clot [229, 230, 41, 43, 207, 231]. Coagulation proteases factor VII (FVII) and factor X (FX) are recruited to the membrane following exposure of anionic lipids to the surface of the cell [41, 13, 22, 210]. Activated FVII (FVIIa) then binds its cofactor, tissue factor (TF), on the membrane surface, forming a binary complex that allows for binding and activation of zymogen FX [41, 231, 207] (Figure 5.1 illustrates TF:FVIIa and FX on the membrane surface). Once activated, FX participates in reactions that activate other proteases in the coagulation cascade, ultimately resulting in the formation of a clot [231, 230]. As formation of the TF:FVIIa:FX complex represents the first step of the extrinsic coagulation pathway, this ternary complex (termed the extrinsic complex) is of great significance to development of potential treatments for blood disorders such as thrombosis and hemophilia.

Despite its great relevance to coagulation, there is little atomic-level information available regarding the macromolecular structure of the extrinsic complex. Indeed, very few atomic-level structures of any coagulation complexes have ever been developed [59, 27]. Efforts to do so have been thwarted by the strong membrane dependence of complex formation between the coagulation proteins and their cofactors [27]. An additional factor which further complicates characterization of individual coagulation proteases is their substantial flexibility [169]. No complete structure of the FX protease has been resolved, for example, although each domain of the protease has been crystalized separately [232, 56]. In order to understand the interactions that regulate the formation of the extrinsic complex, and thus the entire coagulation cascade, a greater understanding of the atomic-level structural details of the complex is required.

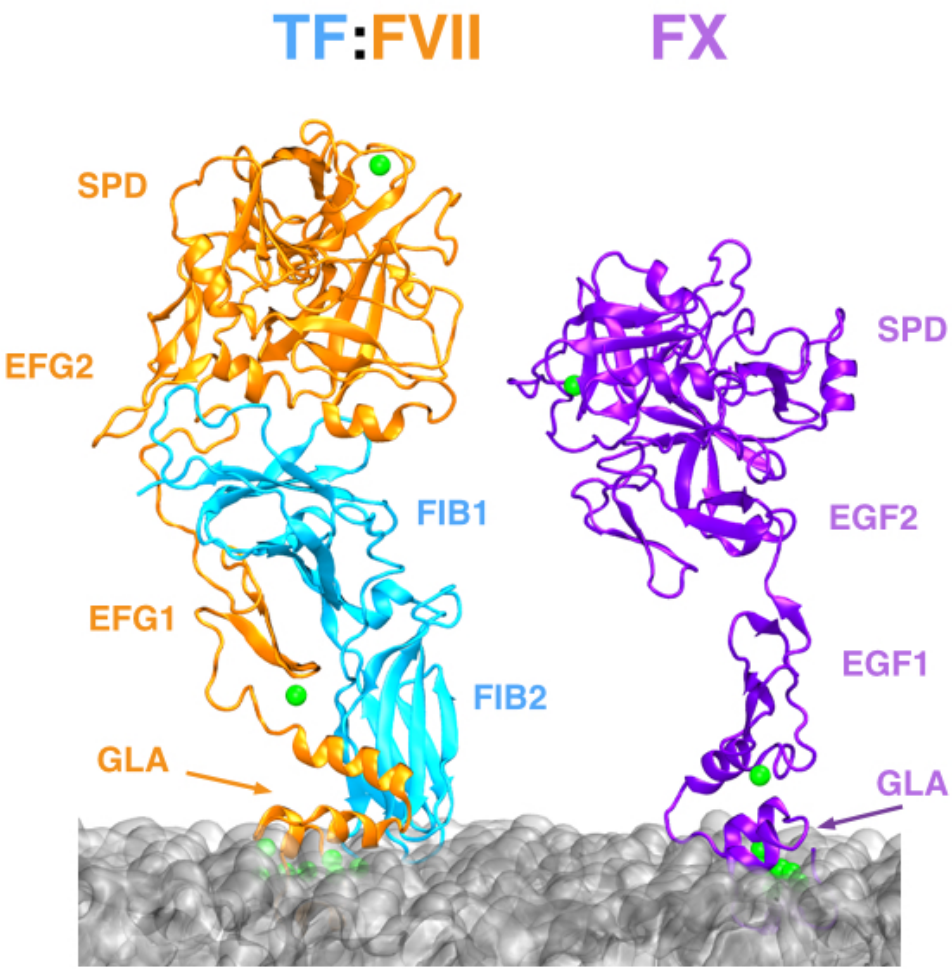


Figure 5.1: Illustration of the components of the extrinsic complex on the surface of a membrane prior to its ternary complex formation. FVIIa is shown in orange, TF in blue, and FX in purple. Bound Ca^{2+} ions are shown in green. FVII and TF first form a binary complex, which increases the enzymatic activity of FVIIa, and then they bind to FX. The process is catalyzed significantly by their association with active areas of the membrane.

Biochemical mutagenesis experiments and molecular dynamics (MD) simulations have both greatly contributed to our knowledge of the complex and its component proteins. A number of residues in the catalytic, serine-protease domain (SPD) of FVIIa (FVII-SPD) have been shown to affect FX activation without an effect on the TF:FVIIa interaction [233, 234]. Strikingly, mutagenesis experiments have also demonstrated that residues as much as $\sim$60 Å away from the catalytic site of FVIIa, and near the membrane interface, have a significant effect on activation of FX [235, 236, 237, 238]. These residues, termed the potential exosite, are contained in TF and located close to the membrane surface. An antibody which inhibits FX activation is shown to bind to this potential exosite of TF without affecting FVIIa binding [239]. Mutagenesis studies suggest that the potential exosite may be a binding site for the specialized FX membrane-binding GLA domain (FX-GLA) [236, 240, 241, 242]. Both complete removal of the FX-GLA domain [241] and its substitution with a different GLA domain [242] have been shown to compromise FX activation. FVII GLA domain (FVII-GLA) residues may also contribute to the formation of the extrinsic complex independent of interactions with TF [240, 12]. Figure 5.2 illustrates the residues that contribute to the potential exosite.

The role of the first epidermal growth factor (EGF)-like domain of FX (FX-EGF1) in its binding to TF:FVIIa is somewhat less clear. While several mutations on factor IX EGF1 are found in hemophilia B [244], mutagenesis studies have formed different conclusions regarding importance of FX-EGF1 for extrinsic complex formation [241, 245]. Zhong et al. found that replacing FX-EGF1 with the EGF1 domain of protein C resulted in a 6-10 fold reduction in affinity for soluble TF (sTF) [245]. A study that compared full FX, FX without GLA domain, and FX without GLA or EGF1, found that activation rates were not affected by removal of EGF1 [241].

Although these results offer valuable information on the extrinsic complex, they alone do not provide sufficient details required to construct an atomic-level structure of the complex. In some cases, the effects of the studied mutations may even be indirect. Computational methods offer a complementary approach that allows for integrating experimental data into structural models for the complex. Two previous models of the extrinsic complex have been developed based on protein-protein docking. [60, 61] However, neither has taken into account the crucial role of the membrane and its lipids. Explicit

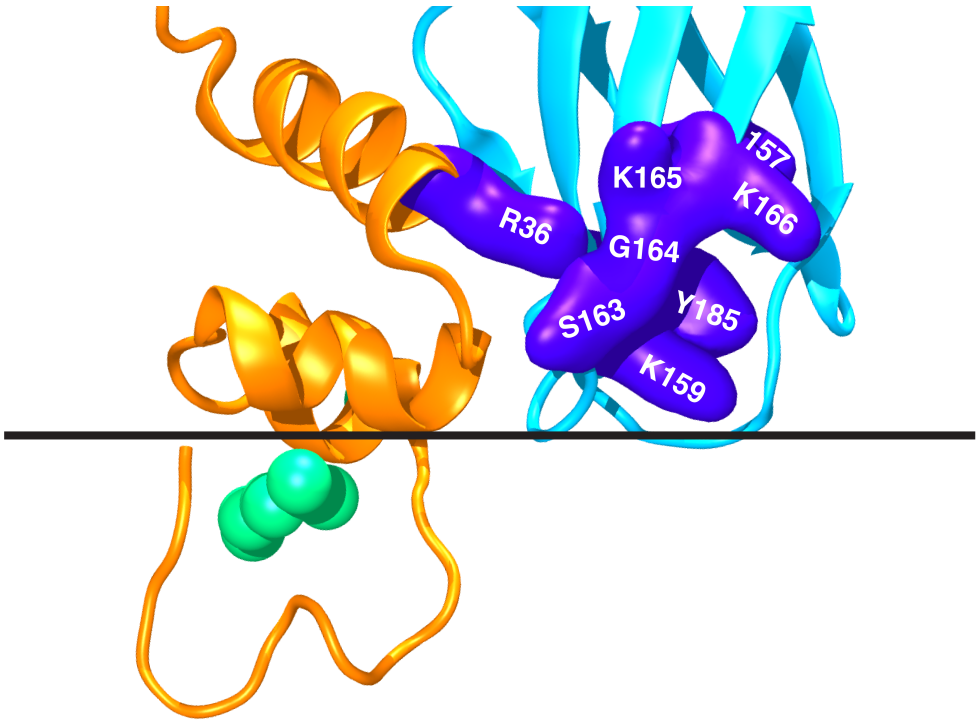


Figure 5.2: Potential TF Exosite. TF:FVIIa residues identified through mutagenesis experiments as putative residues interacting with the FX-GLA domain. Potential exosite residues of TF [235] are shown in purple, as is a nearby residue of FVIIa (R36) which affects FX activation [240].
FVII-GLA is shown in orange, Ca^{2+} ions in green, and TF in light blue. A black line depicts the approximate location of the membrane surface, as seen in MD simulations of the TF:FVIIa complex [243]
.

consideration of lipid interactions is particularly important in predicting interactions of the FVII-GLA and FX-GLA domains within the ternary complex, as lipid-protein interactions determine their orientation and positioning on the surface of the membrane [235, 53, 179]. In addition, effectiveness of modeling through protein-protein docking alone is largely hampered by the inherent flexibility of FX, as rigid-body docking can only utilize snapshots of a protein's position as input.

MD simulations can provide concurrent spatial and temporal resolutions that cannot be matched experimentally or through rigid-body docking. MD simulations have previously been used to study atomic-level protein-lipid interactions for the membrane binding domains of several coagulation proteins, [178, 246, 53, 179] including FVII-GLA and FX-GLA [53, 179]. MD simulations have also been used to characterize membrane interactions and dynamic behavior of the TF:FVIIa complex [243]. Combined with nonequilibrium techniques, the method has also been successfully employed to capture conformational changes in complex membrane proteins such as receptors and transporters, [247, 84, 248, 249, 250] and custom membrane models have been developed to overcome traditional difficulties in membrane simulations [169, 170, 171, 172, 173, 174, 175, 176, 177, 178, 74, 75, 179, 180, 181, 182, 162, 183, 184, 185, 186, 251, 252].

In this study, we develop the first atomic structure of the extrinsic coagulation complex on the surface of a lipid bilayer. The model was developed using specialized MD simulation techniques to model the formation of the extrinsic complex on the surface of a lipid bilayer. We employed a novel protocol which includes protein-protein docking results as an approximate target for nonequilibrium MD, performed on the surface of a specialized membrane model known as the HMMM, with enhanced lipid dynamics [169, 74]. In our simulations, formation of the extrinsic complex is slowly induced on the surface of a lipid bilayer using target docked poses filtered according to the best experimental information available. We provide the most detailed and complete model of the extrinsic complex to date and describe details of the putative binding interface between the experimentally identified exosite and the FX-GLA domain.

5.2 Methods

The general strategy for constructing the atomic model for the extrinsic complex is to first collect information on the interactions between individual domains of the protein components, either from docking of individual domains, from available experimental data, or from mechanistic implications, and then use these inter-domain interactions to guide a simulation in which membrane-bound models of FX and TF/FVII are brought together on the surface of a membrane and while they remain bound to the membrane. These steps are described in more detail in the following sections.

5.2.1 Modeling individual membrane-bound models of FX and TF/FVIIa

Full models of FX and TF:FVIIa were constructed in their bound form to a 50:50 PC:PS (phosphatidylcholine:phosphatidylserine) HMMM. Starting with membrane-bound FVII-GLA and FX-GLA domains [53, 179], we aligned the complete TF:FVII and FX structures in order to construct the complete membrane-bound models. Each membrane-bound, full model was then simulated for 100 ns under equilibrium conditions to obtain a relaxed structure to be used for the complex formation simulation. Details of construction and equilibration of both models follow below.

In order to construct the complete structure of FX, bovine FX-GLA found in protein database structure 1IOD [56] was mutated to the human FX-GLA using the PSFGEN plugin in VMD. The human FX-GLA structure was then placed with all atoms at least 3 Å above a 1:1 PS:PC $50x50$ HMMM membrane. Using the autoionize plugin of VMD, the simulation system was neutralized with Na^+ ions. The system was next energy minimized for 2500 steps and simulated for 100 ns. During this simulation the FX-GLA domain spontaneously inserted into the membrane bilayer with keel residues below the level of lipid phosphate atoms.

The structures of FX-SPD and FX-EGF1 and FX-EGF2 were obtained from the protein database structure 1XKA [232]. The FX-GLA domain from the last frame of the 100 ns simulation described above and the FX-EGF1 were aligned to the FVII structure in protein database structure 1DAN, and the entire protein connected through the addition of four missing residues

in the sequence (residues 45-48) using PSFGEN. PSFGEN was also used to add residue R194, creating a truncated activation peptide, which was capped with a methylated C-terminal. Finally, a Ca^{2+} ion bound to EGF1 was added based on alignment to 1DAN.

The neutrally-charged methylated C-terminal was used because in the FX zymogen additional residues of the activation peptide would be covalently bound, starting with the neutral T193. The entire structure was solvated in water and a Na^+/Cl^- concentration of 0.250 mM. A 1 ns simulation was then performed with all Ca^{2+} ions and backbone atom of FX restrained except residues 194 to 198 with a force constant of 2.5. (2500 minimization steps).

With the entire protein structure complete, final steps were taken to complete the simulation box. The entire structure was then solvated with the SOLVATE plugin of VMD and Na^+, and Cl^- ions sufficient for a concentration of 250 mM were added with the AUTOIONIZE pluggin of VMD. The system was energy minimized for 2500 steps and run for 100 ps with FX protein and Ca^{2+} ion position harmonically restrained with a force constant of 2.5 kcal/mol.

The Ca^{2+} ion was placed at the FX-EGF1 domain based on its location in the FVII-EGF1 1DAN structure, where it is coordinated by residues 46, 47, 49, 63, and 64. Following the 1 ns simulation, simulations were run to induce proper coordination of the ion by protein sidechains. Extrabonds in NAMD was used to induce coordination of the specific residues known to bind to the Ca^{2+} ion. Atoms induced to bond with the Ca^{2+} ions were chosen based on those binding to the Ca^{2+} ion in FVIIa structure in 1DAN. NMR data suggests the FX-EGF1 equivalent to the FVII-EGF1 resiudes in 1DAN interact with Ca^{2+} in FX [253]. Using CHARMM atom names, the atoms with extrabonds specified were: ASP 46 residue atom OD2, GLY 47 residue atom O, GLN residue 49 atom OE1, BHD 63 residue atoms OD1 and OD2, and GLY 64 residue atom O. Eleven short simulations were run in which the strength of the extrabond was gradually decreased. Each simulation was 100 ps, and the force constant was decreased from 10 kcal/mol to 0 kcal/mol. During these simulations, Ca^{2+} atoms as well as backbone atoms of all protein residues except 45 to 65 were restrained with a force constant k=2.5 kcal/mol.

The full FX protease was placed on the membrane by aligning it with the FX-GLA domain in the last frame of the 100 ns membrane-binding simulation. Sufficient HMMM PS:PC (1:1 ratio) lipids were added to bring mem-

brane area to 100×100 Å^2. Sufficient water was added to bring the system size to 100×100×215 Å^3. The AUTOIONIZE plugin of VMD was used to add Na^+ and Cl^- ions sufficient for a concentration of 250 mM. First, a 100 ps simulation was run with all protein and Ca^{2+} ions restrained with force constant k=2.5. The solvated FX structure was then simulated for 100 ns with no restraints after 2500 steps of energy minimization. Extrabonds was used to maintain crystallographic interactions between Ca-6 and atoms within 2.5 anstroms in the PDB 1IOD. A force constant of k=2 kcal/mol was used.

In building the TF:FVIIa structure, a structure of the isolated FVII-GLA domain was first derived from protein databank structure 1DAN. The FVII-GLA domain was placed 3 Å above a 1:1 PS:PC 50x50 HMMM membrane, and the system was neutralized with Na^+ ions using the autoionize pluggin of VMD. The system was next energy minimized for 2500 steps and simulated for 100 ns. During the simulation, the FVII-GLA domain spontaneously bound to the membrane with keel residues inserting below the level of the lipid phosphate groups.

The TF:FVIIa complex structure used was the same as that modeled in our group's previous study of membrane-bound TF:FVIIa [243]. The structure was aligned with the FVII-GLA bound to 1:1 PS:PC HMMM from the last frame of the simulation described above and placed on the membrane. Sufficient HMMM PS:PC (1:1 ratio) lipids were added to bring membrane area to 100×100 Å^2. Sufficient water was added to bring the system size to 100×100×160 Å^3. The AUTOIONIZE plugin of VMD was used to add Na^+ and Cl^- ions sufficient for a concentration of 250 mM. During a 100 ps simulation, protein and Ca^{2+} atoms were restrained at k=2.5 force constant. The entire system was then simulated for 100 ns after 2500 steps of energy minimization.

5.2.2 Protein-protein docking

To gain insight into potential binding modes of the TF:FVIIa:FX complex, rigid-body protein-protein docking was performed using DOT2 [204]. Three protocols were performed to predict binding of: 1) the FX-SPD and FVII-SPD domains, 2) the FX-EGF1 domain and TF:FVIIa, and 3) the FX-GLA domain and TF:FVIIa. (See Figure 5.3 for illustration) The GLA and BHD

residues were added to the molecular library used for preparing the structures by analogy to standard amino acids already available in the library. Ca^{2+} ions were added to the library with a radius of 1.326 [254].

In protocol 1, the moving molecule was the FX-SPD (FX protein residues 194 to 429) domain with FX-EGF2 (residues 91 to 152) and the stationary molecule was FVII-SPD (residues 153 to 406) with FVII-EGF2 (residues 91 to 152). In protocol 2, the moving molecule was FX-GLA (protein residues 1-44 and all associate calcium residues), and the stationary molecule was TF-FIB2 (residues 106-213). FX and TF:FVIIa structures used in docking were derived from 100 ns equilibrium simulations of membrane-bound FX and TF:FVIIa. Snapshots for docking were taken from the simulations at 60 ns, 70 ns, 80 ns, 90 ns and 100 ns. Each protocol was repeated with each of these structures, producing 5 docking runs per protocol.

For all protocols, grid spacing was set to 1 Å, and number of degrees per rotation was set to 6°. Ten clashes between the two molecules, or "bumps," were allowed for output structures. DOT2 was set to output the maximum number of rotations, which is 200,000. Since each protocol was repeated 5 times, a total of 1,000,000 structures were output per protocol. Docked structures were filtered to retain only those structures satisfying experimental constraints on complex binding. Detailed protocols for filtering of docked structures follow below.

5.2.3 Filtering docked structures from protocol 1

The first filter for protocol 1 is based on the fact that the residues whose peptide bond is to be cleaved by FVII at the P1 pocket are FX residues I195 and R194 [255]. R194 was added later to the FX structure used in simulation, but in docking we used the FX-SPD structure in 1XKA, which is active and lacking R194. To satisfy the first filter, C_α atom of FX residue I195 must be within 15 Å of the P1 pocket residue S203 C_α atom. This cuttoff represents the distance between the R194 to and FVII S203 C_α atoms based on inhibitor binding in PDB 1DAN plus our estimate of possible displacement of I195 from its position in active FX-SPD (which was used in docking) to its position in an zymogen FX-SPD. Detailed information on calculation of this cuttoff follows.

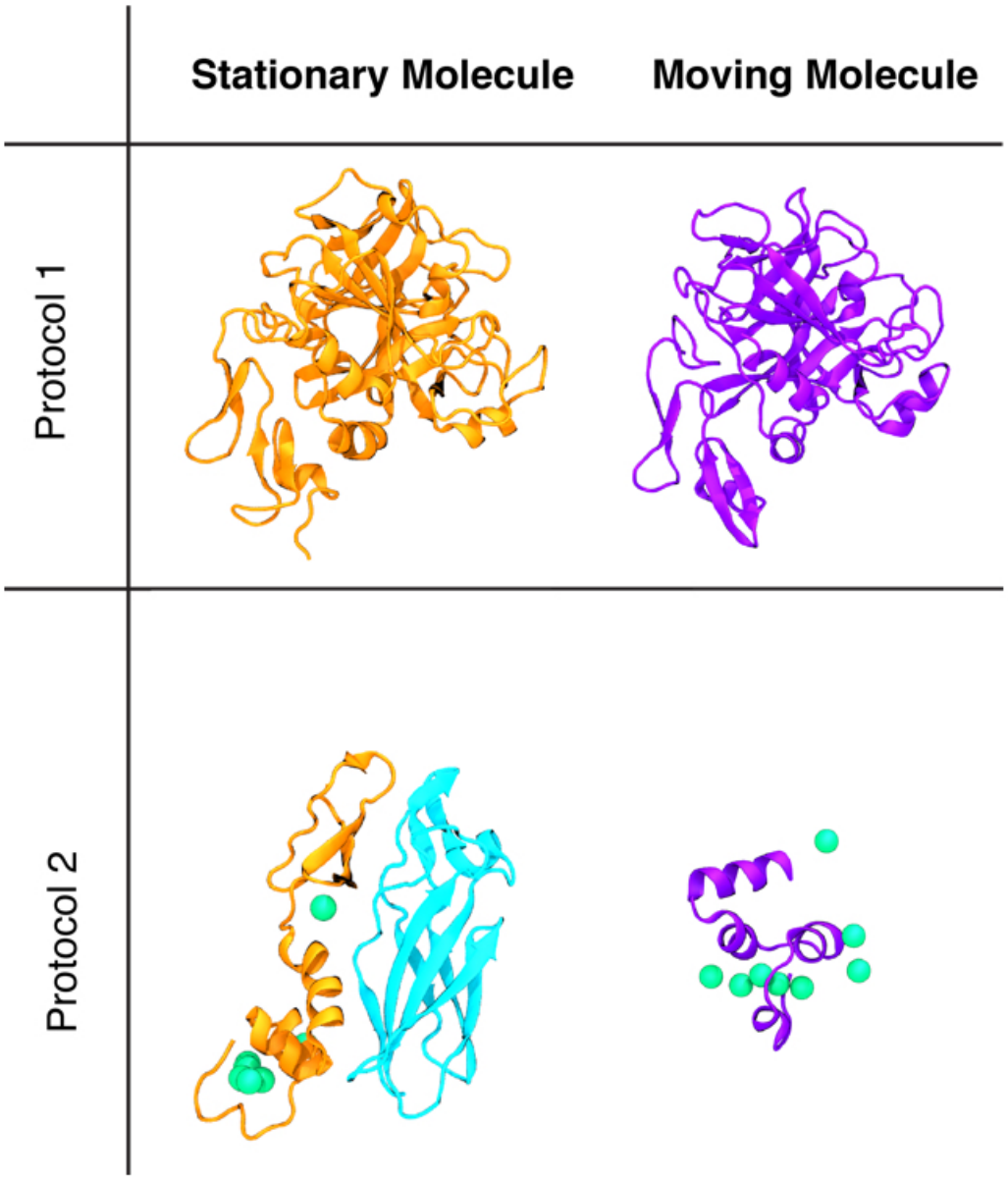


Figure 5.3: Structures used in the three protein-protein docking protocols. In protocol 1, the stationary part is FVII-SPD and FVII-EGF2, and the moving part is FX-SPD and FX-EGF2. In protocol 2, the stationary part is the FVII-GLA, FVII-EGF1, and TF-FIB2, while the moving part is FX-GLA. Domains of FX are shown in purple, domains of FVII in orange, TF domains in blue, and Ca^{2+} ions in green.

We estimated active-zymogen displacement of R194 from inactive to active state by assessing structures of zymogen FVII-SPD (PDB 1JBU [256]), FXI-SPD (PDB 2F83 [257]), CR1 complement (PDB [258]), and MASP-1 (4IGD [259]). For each structure, we aligned the domain to FX-SPD and calculated displacement of equivalent cleaved residues to the position of R194. For use in calculating the cutoff distance, we selected the shortest displacement seen among available structures to ensure the most conservative estimate of possible orientations of FX-SPD, and minimize the possibility of including unfavorable binding poses.

In the PDB structure 1DAN, the C_α atom of the arginine-based inhibitor side chain was within 6Å of the S203 C_α atom. Thus, the rough estimate of the distance between I195 C_α and and S203 C_α would be 10Å+6Å−1.5 (1 amino residue backbone length)=14.5Å, which was rounded up to the nearest full integer, 15Å.

The second filter ensured that the C_α atom of FX-SPD residue C124 was at a higher z-position than the C_α atom of residue C89 (preventing FX orientations incompatible with simultaneous membrane and FVII-SPD binding). The third and final filter ensured the C_α atom of FVII-SPD residue C127 had a higher z-position than the C_α atom of FX-SPD residue C89 (preventing FX orientations incompatible with simultaneous membrane and FVII-SPD binding). The "clustering" command in VMD was used to cluster the filtered set based on RMSD. Five final clusters were produced with an RMSD cutoff of 2 Å. Of the top three clusters, the one with the most contacts within 8 Å between any FX residue C_α atom and FVII residues 21, 24, 39, 60, 75, 76, 77, 78, 82, 103, 153, 156, 158, 159, 161, 195, 196, 197, 199, 204, 229, and 304 C_α atoms was selected for further analysis. These FVII residues were found experimentally to affect FX activation while not affecting TF:FVIIa association [233]. The "measure contacts" command in VMD was used for this analysis.

5.2.4 Filtering docked structures from protocol 2

For the first filter, the C_α atoms of seven residues on tissue factor, (157, 159, 163, 164, 165, 166, 185), previously found upon mutation to affect activity [235], as well as of residue 36 of FVII which has been identified as

interacting with FX-GLA [240], were required to be within 15 Å of at least one C_α atom of an FX-GLA residue (residues 1 to 44). To satisfy the second filter, the z-component of the center of mass (COM) was required to be within 5 Å for the C_α atoms of FX and FVII keel residues (residues 4, 5, 8). The z-component of the center of mass (COM) was also required to be within 5 Å for the C_α atoms of FX and FVII aromatic stack residues (32-44). Clustering in VMD was used to calculate the five top RMSD clusters within a 2 Å RMSD. The cluster with the greatest number of frames was selected for use in later analysis.

5.2.5 Analysis of docked structures for colvar design

Docked structures which satisfied the filtering protocols were clustered using the Clustering command in VMD with an RMSD cutoff of 2 Å, and ranked based on number of poses grouped. For protocol 1, five total clusters were produced with an RMSD cutoff of 2 Å. Of the top three most populated clusters, the one with the the most contacts within 8 Å between any FX residue C_α atom and FVIIa residues 21, 24, 39, 60, 75, 76, 77, 78, 82, 103, 153, 156, 158, 159, 161, 195, 196, 197, 199, 204, 229, and 304 C_α atoms (FVIIa residues known experimentally to affect activity [233]) was selected for use in collective variable development. These were residues of FVII-SPD determined through mutagenesis to affect FX activation without impairing the ability of FVII-SPD to bind TF, while also being solvent accessible [233]. For protocol 2, the most populous cluster was selected.

Docked structures in the top selected clusters were analyzed to determine which sets of C_α atoms between the stationary and moving molecules were most frequently in contact (at least 0.05 normalized probability). The measure contacts method in VMD was used in this analysis. Two C_α atoms were considered to be in contact if they were within 8 Å. The coordination for C_α number within 8 Å was then calculated for the sets of residues most frequently in contact for each pose in the top cluster. This analysis was used to determine target values for coordnum collective variables SPDCON and GLACON. Once atoms were selected for use in collective variables, distance and coordnum were calculated for the final cluster of docked poses, and cutoff for the collective variable determined by selecting the longest distance

and smallest coordination number for the cluster, respectively.

5.2.6 Extrinsic complex formation: nonequilibrium simulations

The input structures for the TF:FVIIa:FX complex components were derived from five snapshots from the FX and TF:FVIIa membrane-bound simulations. Snapshots were taken from 60 ns, 70 ns, 80 ns, 90 ns, and 100 ns of the equilibrium simulations. The TF:FVIIa and FX systems were combined and placed so as to ensure that the residues of FX-GLA identified (through analysis of rigid-body docking output) as most likely to be in contact with the TF potential exosite were facing TF:FVIIa.

Simulations to induce binding of the extrinsic complex used a total of 11 collective variables: 6 coordnum collective variables (used to induce or discourage contacts between specific sets of atoms or residues), and 5 distance collective variables. Simulations were induced in a multi-phase process, the primary steps of which are illustrated in Figure 5.4. Two collective variables were used to induce binding between the FX-GLA domain and TF:FVIIa, 3 for binding of FVII-SPD and FX-SPD. Four additional collective variables were used to induce docking of FX residue R194 to the FVII activation site. The system was moved along the path of the collective variables using moving harmonics.

In brief, GLA domains are first induced to bind, followed by the SPD domains. Finally, a truncated FX activation peptide is docked in to the FVIIa P1 pocket. A detailed description of this protocol and the collective variables used follows below.

GLADIS: distanceXY collective variable used to reduce distance between the FX-GLA domain and TF:FVIIa in the XY plane. It reduces the distance between the COM of TF-FIB2 C_α, FVII-GLA C_α, and FVII-GLA Ca^{2+} atoms and the COM of FX-GLA Ca^{2+} and C_α atoms of residues 405 to 414, 430, 431, and 432. These residues are either embedded within the bilayer, or part of the side of FX-GLA determined most likely to bind through docking. Including membrane-embedded atoms prevents potential distortion of the membrane binding pose as the GLA domain is pushed through the bilayer toward TF:FVIIa. Center of mass distance was reduced to 22 Å. This cutoff

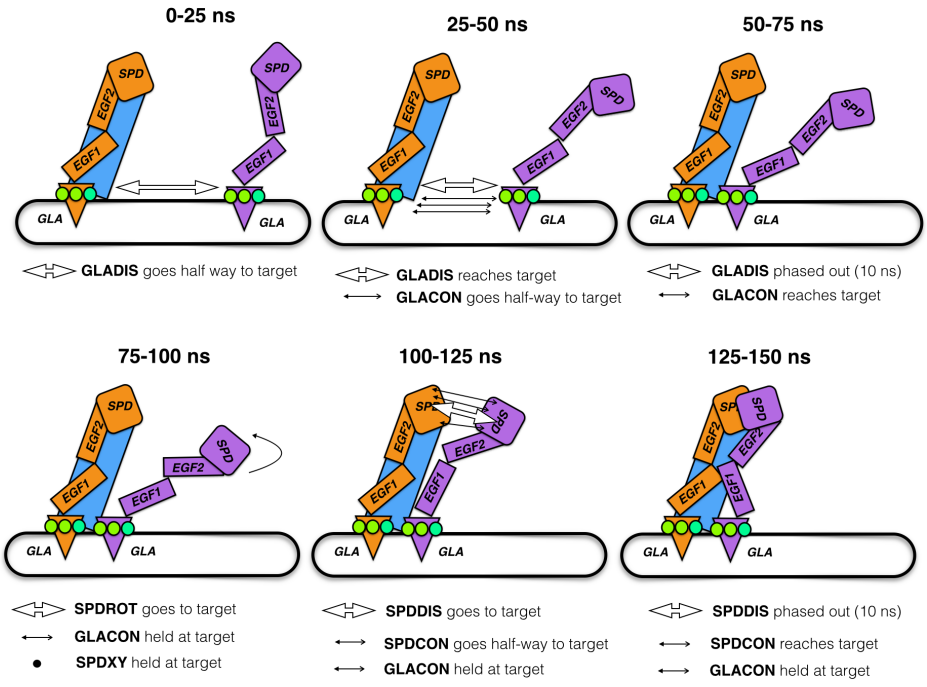


Figure 5.4: Illustration of the effects and timing of collective variables which induce binding between major domains during the first 150 ns of the simulation. TF shown in blue, FVIIa in orange, and FX in purple with major domains labeled. Arrows are used to show when domains are moving under influence of collective variables. Large white arrows with a black outline indicate a distance collective variable, while small black arrows indicate a coordum collective variable.

was determined by calculating distance between the two groups of atoms described above in the final cluster of docked poses and selecting the longest distance.

GLACON: coordnum collective variable used to induce contacts between the FX-GLA and potential exosite. The group of atoms on FX used were C_α atoms of residues 12, 13, 14, 27, 30, 31, 32, 33, and 36. TF:FVIIa residues in the second group for the collective variable were C_α atoms of FVII-GLA residue 36 and TF residues 157, 159, 163, 164, 165, 166, and 185. These residues were found to have the highest contacts numbers of contacts in analysis of the final cluster of docked poses as described in the methods and results sections. Coordination number was increased to 11.5, with a cutoff distance of 8 Å. The 11.5 cutoff was determined by calculating the coordnum value between the two groups in the final cluster of docked poses as described above, and selecting the smallest coordnum.

SPDROT: coordum collective variable used to rotate the FX-SPD domain and provide a clear line of sight between the groups of atoms used in SPDDIS. Coordination number is increased between C_α atoms of FX-SPD residues 194-197, 324-327, 334, 372-375, 378 and C_α atoms of FVII-SPD residues 16, 17, 148, 149, 152, 201-206, and 231. These groups were selected because they are the same as will be used in SPDDIS. Coordination number was increased to 168 with a cutoff distance equal to the center of mass distance between FVII-SPD and FX-SPD when the collective variable is first initiated.

SPDXY: distanceXY collective variable used to keep distance between FX-SPD and FVII-SPD domains constant in the XY plane while SPDROT is active. Distance between COM of FX-SPD C_α atoms and COM FVII-SPD C_α in the XY plane is centered at distance in XY at when collective variable initiated.

SPDNON: coordnum collective variable eliminate contacts between FX-SPD and TF:FVIIa. Coordination number between FX-SPD C_α atoms and FVII-SPD, TF C_α atoms is reduced to zero. Cutoff used was 8 Å .

SPDDIS: distance collective variable used to pull the two SPD domains together. The COM distance between FX-SPD C_α atoms 194-197, 324-327, 334, 372-375, 378 and C_α atoms of FVII-SPD residues 16, 17, 148, 149, 152, 201-206, and 231 of was reduced to 17 Å. The atom groups were chosen by selecting residues used for groups in SPDCON and all other C_α atoms within 5 Å.

SPDDIS2: same as SPDDIS except FX-SPD residues 194-196 are not included in calculating the COM distance. This allows activation peptide docking when ACTCON, ACTDIS, and ACTNON are active.

SPDCON: coordnum collective variable used to induce contacts between the two SPD domains. Coordination increased between C_α atoms of FVII-SPD residues 39, 60, 153, 156, 204, 229 and C_α atoms of FX-SPD residues 195, 196, 250, 257, and 332. These are the residues found in analysis of the final cluster of docked poses to be in most contact as described above. Contacts were increased to 5, with a cutoff of 8 Å. The 5 cutoff was determined by calculating the coordnum value between the two groups in the final cluster of docked poses as described above, and selecting the smallest coordnum.

ACTCON: collective variable used increase coordination number of the NE, NH1, NH2, and CZ atoms of FX-SPD residue R194 and the OG, O, OD1, OD2 atoms of FVII-SPD residues 201 and 202. and the COM of all FX-SPD C_α atoms. Coordination number is increased to 9 with a cutoff of 4 Å . Groups and cutoffs are determined through analysis of inhibitor binding in PDB 1DAN. In PDB 1DAN, the arginine side chain interacts with the serine side chain and backbone oxygen of residue 202 (name O and OG in the CHARMM naming system) and the oxygen of atoms of the ASP 201 side chain atoms OD1 OD2. The coordination number between the arginine side chain atoms and FVII-SPD residues 201 202 were calculated in the 1DAN structure, and determined to be 9 with a 4 Å cutoff.

ACTDIS: distance is decreased between the FX-SPD residue R194 side chain CD atom and FVII-SPD residue 203 CYS backbone atom C_α. Distance between the two atoms is decreased to 4 Å. Cutoffs and groups determined through analysis of inhibitor binding in PDB 1DAN. Inhibitor arginine side chain atom CD (CHARMM naming system) is within 4 Å of protein FVII-SPD residue 203 and name C_α.

ACTNON: coordnum collective variable which reduces contacts between side chain atoms of FX-SPD residues R194 and C_α atoms of all FX-SPD residues except 194 and 195. Coordination number is reduced to zero with a cutoff of 8 Å .

Collective variables described above are used in each phase of the complex formation simulations as follows:

1) From 0-25 ns, GLADIS was initiated and completed half of the distance to its target. The force constant for GLADIS was 10 kcal/mol.

2) From 25-50 ns, GLADIS was continued and GLACON initiated. GLADIS reached its target, while GLACON reached half its target. The force constant for GLACON was 10 kcal/mol.

3) From 50-75 ns, GLADIS is held at its target value. GLACON reached its target value. The force constant of GLACON was held at 10 kcal/mol, while the force constant of GLADIS was gradually decreased to zero over 10 ns. SPDNON was initiated and reached its target value, and the force constant was 10 kcal/mol.

4) From 75-100 ns, GLACON was held harmonically at its target value. SPDROT was initiated and reached its target coordination number. Force constant for all collective variables was 10 kcal/mol. SPDXY was held at its center. The force constant for all collective variable was 10 kcal/mol.

5) From 100-125 ns, GLACON was held harmonically at its target value. SPDDIS was initiated and set to reach its target distance. SPDCON was initiated and set to move to half its target coordination number. The force constant for all collective variable was 10 kcal/mol.

6) From 125-150 ns, GLACON and was held harmonically at its target values. SPDDIS2 was initiated and held at its target value while the force constant was gradually reduced to 0 over 10 ns. SPDCON reached its target value. ACTNON and ACTDIS were initiated and reached their targets in 10 ns. ACTCON was initiated and reached its target value in 25 ns. The force constant was 10 kcal/mol for each collective variable.

7) From 150-200 ns, GLACON, SPDCON, ACTCON, and ACTDIS held at their targets and their force constants gradually decreased to zero.

8) From 200-400 ns, the complex is simulated without collective variables on an HMMM. From 400-550, the complex is simulated on a full membrane model.

5.2.7 Equilibration of complexes

After biased simulations are completed, all harmonic potentials are phased out over 50 ns. Force constant k is gradually reduced to zero for each. The complexes are then simulated for an addition 200 ns without bias. In order to ensure that the final TF:FVIIa:FX complex was not unduly affected by the use of the HMMM as opposed to a full-tailed lipid construct, all bound

systems were converted to conventional (full-tailed) membrane, using a PDB generated from the final frame of each 200 ns equilibrium HMMM simulation. A full-tailed lipid was overlaid onto each short-tailed lipid, and coordinates of the missing atoms were then copied from the full-tailed lipid, allowing head-group interactions obtained from HMMM simulations to be preserved. The layer of 1,1-dicholoroethane (DCLE) solvent molecules were then removed from the center of the HMMM bilayer. After a 10,000-step energy minimization, each TF:FVIIa:FX complex was simulated for another 150 ns in the full membrane.

5.2.8 Simulation protocol

All MD simulations were performed with NAMD2 [260, 86, 226] and CHARMM36 force-field parameters for lipids [111] and proteins [131]. Parameters for GLA (γ-carboxy-glutamic acid) residues were those previously developed by our laboratory with atom types renamed from CHARMM27 to CHARMM36 [53]. The special amino acid residue BHD ((3S)-3-hydroxy-L-aspartic acid) was developed by analogy to ASP and SER parameters in CHARMM36. The TIP3P water model was used for all simulations [261]. Constant pressure was maintained at a target of 1 atm using the Nosé-Hoover-Langevin piston method [164]. Constant temperature was maintained at 310 K by Langevin dynamics with a damping coefficient, γ, of $1\,\mathrm{ps}^{-1}$ applied to all atoms [165]. All simulations used a 2 fs timestep, as well as the vdWForceSwitching scheme. Covalent bonds involving hydrogens and the H-O-H angle were fixed using the SHAKE and RATTLE algorithms, respectively [156, 157]. The particle mesh Ewald (PME) method [167] was used for long-range electrostatic calculations with a grid density greater than $1/\,\text{Å}^3$. Non-bonded interactions were truncated after 12 Å with a smoothing function starting at 10 Å. System setup, visualization and analysis were carried out using VMD [213].

In HMMM simulations, the z-positions of the short-tailed lipids were restrained harmonically with a force constant $k = 0.05$ kcal/mol/Å^2) applied to lipid tail atoms C21 and C31 (CHARMM topology naming convention). Lipid harmonic restraints were maintained throughout the simulations in order to reproduce the natural distribution of head-group posi-

tions. The z-positions of the 1,1-dichloroethane (DCLE) solvent atom C3 (CHARMM topology naming convention) was restrained with a force constant $k = 0.1\,\text{kcal/mol/Å}^2$. Area per lipid for all HMMM lipids was $73\,\text{Å}^2$. The flexibleCell option in NAMD was used for HMMM simulation with constantArea applied. For full membrane simulations, flexibleCell with constantRatio was used.

5.2.9 Simulation analysis

Angles between FX domains were calculated during the FX and complex formation simulations. Two angles were defined using the centers of masses (COMs) of C_α atoms of selected residues. The first angle, α, is the angle between the vector connecting the COM of FX-SPD to the COM of FX-EGF2, and the vector defined by the COMs of FX-EGF2 and FX-EGF1. Angle β is defined as the angle between the vector defined by COM of FX-SPD and FX-EGF1 and the vector defined by the COM of FX-EGF1 and FX-GLA domains. CA atoms were used to calculate each center of mass using the measure center command in VMD. Tilt of FX-GLA and FVII-GLA domains relative to the membrane were analyzed using the same methodology described in our previous study of FX-GLA domains [179].

Tilt of the TF:FVIIa complex was analyzed by measuring the principle axis of rotation through the TF-FIB1 and TF-FIB2 domains. θ_{CM} and θ_{NC} were calculated in the same manner as a previous study of membrane-bound TF:FVIIa [243]. θ_P is the angle between the axis perpendicular to both I_{NCM} in the plane of the membrane and I_C.

Number of lipids in the immediate vicinity of the FX-GLA and FVII-GLA domains were calculated using the "measure contacts" command in VMD. Each GLA C_α or Ca^{2+} atom within $5\,\text{Å}$ of a lipid phosphate atom was counted as one contact.

Number of contacts between protein and Ca^{2+} residues of TF:FVII and FX were assessed using the "measure contacts" command in VMD. Each protein or Ca^{2+} atom of TF:FVIIa within $3\,\text{Å}$ of FX was considered to be in contact with FX.

5.3 Overview

Our goal is to present, on an atomic level, the most complete membrane-bound model of the extrinsic complex possible given currently available experimental data. To that end, we have constructed full-length models of FX and TF:FVIIa on a PS:PC 50:50 HMMM bilayer. Each model was simulated for 100 ns, and snapshots from 60 ns-100 ns were selected to serve as starting conformations for the complex formation simulations and for rigid-body docking calculations.

The construction of the individual FX and TF:FVIIa models is described at length in the methods section. The construction of the FX model required particular care, as no complete FX or FXa structure is available in the protein databank. Our model is membrane bound and near complete, lacking only a full activation peptide segment (residues 183-234). This was necessary as no inactive crystal structures of FX-SPD are available. The 1XKA [232] structure used to construct our FX model is an activated SPD, and activation peptide residues 140 to 194 are therefore absent from the structure.

We determined that the available loop modeling methods would not be sufficient to ensure an accurate representation of the activation peptide's conformation, potentially introducing inaccuracies into our model. We therefore elected to only include a truncated activation peptide, adding only residue R194 so as to include the cleavage site of the FX and ultimately include docking of the cleavage side into the P1 pocket of FVIIa [255].

Our TF:FVIIa complex model is built similarly to our previous membrane-bound TF:FVIIa model [53], with the exception that a PS:PC 50:50 HMMM was used to rather than 100% conventional DOPS. Both our FX and TF:FVIIa models were simulated for 100 ns, and a number of characteristics assessed as described below. We confirmed that the binding poses of both FVII-GLA and FX-GLA were stable, and the orientation of FX-GLA was consistent with our previous study of this domain [179]. We also determined that orientation of TF:FVIIa was consistent with our previous work [243], and that a sufficient variety of confirmations of the extremely flexible FX protease would be represented at the initiation of our complex formation simulation. Extended discussion of this analysis follows.

5.4 Individual equilibrium simulations of TF:FVIIa and FX

HMMM simulations were run for 100 ns for each of the complete structures of TF:FVIIa and FX. These simulations allowed the proteins to equilibrate within the membrane, and to generate a variety of configurations for use as the starting points of the nonequilibrium simulations. Analysis was performed to show degree of variation between starting poses in nonequilibrium simulations as well as establish baseline RMSD and lipid binding information for comparison to nonequilibrium simulations.

We assessed the confirmation of the very flexible FX protease using two metrics, α and β, as described in the methods section. For our 100 ns simulation of membrane-bound FX, α varies 130-180° over the final 40 ns, while β varies 105-160°. By choosing snapshots at 60 ns-100 ns, we select a diverse set of starting poses for FX at the start of each complex formation simulation.

For the FX-GLA domain, orientation to the membrane was assessed using "port-to-starboard" and "bow-to-stern" metrics as described in our previous study of bovine FX-GLA [179]. At and after 60 ns, both angles vary by approximately 20°, with a range of 90-110°. This orientation is also consistent with our previous study of isolated FX-GLA domains [179].

As FVIIa is bound to TF such that its domains do not have freedom to rotate, so an analysis of α and β was not performed for FVII. However, orientation of FVII-GLA to the membrane normal is examined. The "port to starboard" angle varies by by 100-125°, the "bow to stern" angle varies by 40-60° of the 100 ns equilibrium simulation. In the last 60 ns, each angle varies by 25° or less, suggesting a stable bilayer binding pose.

The tilt of TF relative to the membrane was analyzed using two angles, and the angle between the two fibronectin type III domains was also calculated as described in the methods section. Values of θ_{CM} and θ_{NC} are similar to those found in a previous study of TF:FVIIa [243]. In the previous study, θ_{CM} was calculated to be 94.4±4.47° [243] over 100 ns, while here it was calculated to be 95.1±7.0° over 100 ns, nearly identical values. θ_{NC} was calculated at 126.6±5.2° previously, while in our simulation it was calculated at 120.9±4.2°, an ≈5.7 °difference. Value of θ_P was calculated to be 108.6±4.1°, suggesting the surface of the protein containing the potential exosite is angled ≈5-10° away from the membrane plane.

Protein/Domain	RMSD	SD
TF:FVII	2.90	0.47
TF	2.48	0.53
FVII-SPD	1.31	0.17
FVII-EGF1	0.69	0.10
FVII-EGF2	1.69	0.45
FVII-GLA	0.91	0.11
FX	8.20	2.69
FX-SPD	1.63	0.18
FX-EGF1	1.09	0.34
FX-EGF2	0.63	0.12
FX-GLA	0.76	0.10

Table 5.1: Average RMSD values during equilibrium simulations of FX and TF:FVIIa, both for full proteins and key domains.

We also assessed the RMSD of full proteins and key domains of TF, FVIIa, and FX for comparison to nonequilibrium simulations. Results can be found in Table 5.1. RMSD of the the TF:FVIIa complex was consistent with the complex RMSD found in the previous TF:FVIIa study: 2.94±0.41 previously [243] vs. 2.90±0.47 here. This suggests use of the HMMM does not have any significant effect on the TF:FVIIa structure. The RMSD of the full FX structure was 8.20±2.69. This is an indicator of the flexibility of FX on the membrane surface. Number of lipids in the immediate vicinity of the the FX-GLA and FVII-GLA domains during the equilibrium simulations was calculated. We found that the number of phosphate groups within 8 Å of a GLA domain CA atom was 2.92±0.76 for FX-GLA and 1.70±0.77 for FVII-GLA. The lesser number of lipid phosphate in vicinity of the FVII-GLA as compared to FX-GLA is likely due to interaction of FVIIa with TF.

5.5 Protein-protein rigid-body docking and colvar design

Using snapshots from 100 ns simulations of FX and TF:FVIIa as described above, rigid-body docking simulations were performed to provide additional information on potential binding modes for the extrinsic complex. Rather than docking the entire FX molecule to the TF:FVIIa, we elected to dock

individual domains of FX known to bind TF:FVIIa. This reduced the risk of biasing results inherent in using static frames of a very flexible protein as starting points of the docking simulations. A total of 1,000,000 docked poses were produced for each of the protocols as described in the methods section. After filtering according to known orientation parameters for FX and TF on the membrane surface, remaining structures were clustered and the best cluster utilized for development of collective variables.

Two types of collective variables were needed to drive complex formation simulations: distance collective variables, which reduce distance between the center of mass of specific atom groups, and coordnum collective variables, which induce formation of contacts between specific groups of atoms. The first step in development is to determine residues to be included in the coordnum collective variables. The collective variable used to induce contacts between the SPD domain and potential exosite is SPDCON, while the collective variable used to reduce distance between the FX-GLA and TF:FVIIa is SPDDIS. Similarly, the collective variable used to induce contacts between the GLA domain and potential exosite is GLACON, while the collective variable used to reduce distance between the FX-GLA and TF:FVIIa is GLADIS.

Both types of collective variables require "groups," sets of atoms to which force is applied, and a "target," a goal toward which the simulation is biased. In order to determine the correct groups and targets for SPDCON, the top cluster of docked poses from protocol 1 was analyzed to determine the set of residues in most frequent contact as described in the methods. The CA atoms of these residues constitute the groups to which force is applied in the collective variable. The coordination number between the two groups was then calculated for each structure in the cluster, and the smallest coordnum value for the set was used as the target value in the collective variable. As we cannot be sure which docked pose provides the most accurate representation of extrinsic complex binding, using the smallest coordnum target reduces the risk of forcing a unfavorable pose during complex formation. If it is favorable to form additional contacts, they have the potential to form spontaneously during our equilibrium MD simulations. The groups for SPDDIS included the same atoms as used in SPDCON, in addition to all CA atoms within 5 Å of them. To determine the target distance for SPDDIS, we measured the center of mass distance between the groups of atoms, and chose the greatest distance among all the clustered structures as the cutoff.

The same process was used to define the groups and targets for GLACON. Groups for GLACON were determined using the top cluster from protocol 2 with the most frequent contacts as described in the methods. Additional considerations had to be factored into the design of GLADIS so as to reduce potential for membrane binding disruption. TF:FVIIa atom group was composed of CA and Ca^{2+} atoms of TF-FIB2 and FVII-GLA, while the FX residues include any residues on the membrane penetrating loop of FX as well as residues on the same face of FX as residues determined likely to bind TF through docking. By applying force to the likely TF-binding interface of FX-GLA only, we allow GLADDIS to optimally orient the FX-GLA for contact formation by the GLACON colvar.

We created additional collective variables to smooth the process of SPD binding between FVIIa and FX. They are used to rotate the FX-SPD domain so that SPDCON and SPDDIS can function properly, as well as prevent binding of FX-SPD to the TF:FVIIa complex during the first portion of complex formation, when FX-GLA binds TF:FVIIa. Additional details can be found in the methods section.

We designed three collective variables to dock our truncated FX activation peptide into the FVIIa P1 pocket catalytic site during the final segment for the simulation. To define our collective variables, we assessed contacts between between a FVIIa inhibitor and the TF:FVIIa complex in crystal structures. Extended discussion of the details of this design can be found in the methods. Three collective variables were used: one to ensure no contacts between R194 and other FX residues, one to reduce the distance between R194 and the catalytic site, and one to induce specific contacts between R194 and the catalytic site as seen in the inhibitor crystal structure (1DAN). Additional details of these collective variables can be found in the methods section.

5.6 Results of biased simulations

Five independent simulations were performed using as the initial poses snapshots of FX and TF:FVIIa taken every 10 ns from 60-100 ns in the equilibrium simulations. A detailed description of the simulation protocol is described in the methods section. The first step of the simulation was to

bring the the FX-GLA into proximity with the potential exosite using the GLADIS and GLACON collective variables. Nonequilibrium work and collective variable value were assessed for GLACON and GLADIS over steps 1-3 (Figure 5.5, Figure 5.6). GLADDIS shows a linear progression toward its target value over 50 ns. Nonequilibrium work associated with the colvar also increases linearly for the first 25 ns to 20 kcal/mol/Å^2 or less (Figure 5.5). The value of GLADIS continues to linearly decrease during step 2,

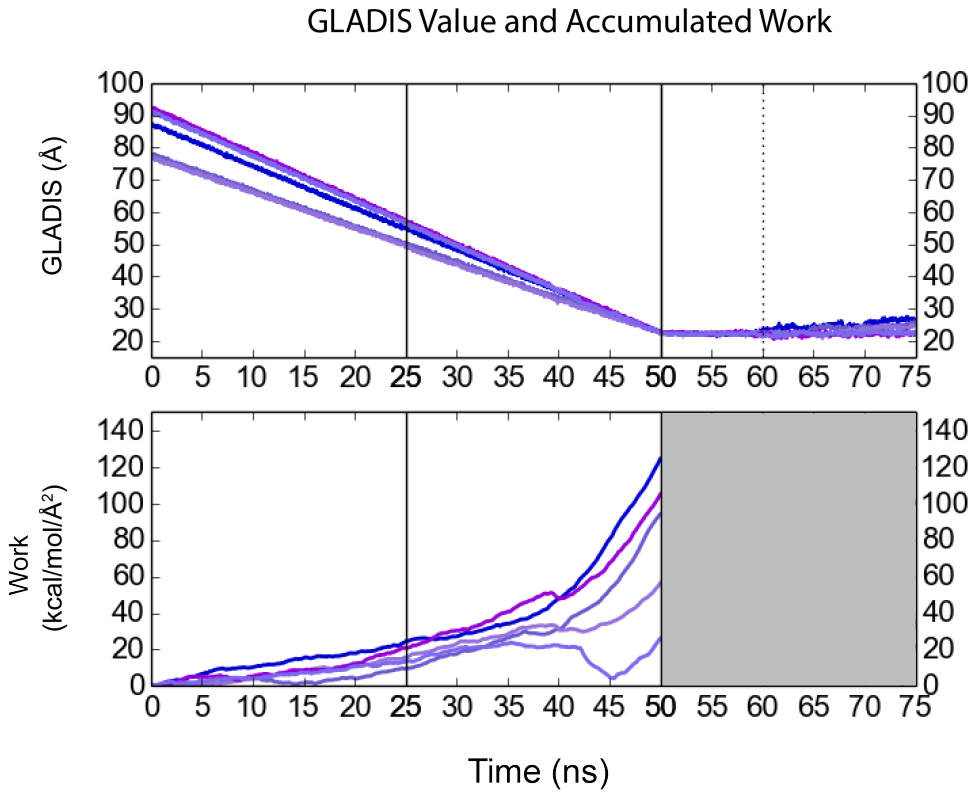


Figure 5.5: GLADIS value and nonequilibrium work in the first 75 ns of simulation (steps 1-3). The top row shows the value of GLADIS, while the bottom row shows the nonequilibrium work. A solid line is used to indicate initiation of a new step in the colvar protocols. For example, a solid line is placed at 25 ns when the GLACON colvar is initiated at step 2 begins. At 50 ns, the GLADIS colvar becomes a stationary rather than moving colvar, and nonequilibrium work is no longer collected. When no nonequilibrium work is collected, this is indicated with a grey box. The dotted line indicates the point at which the force constant k for GLADIS has been reduced to zero.

nonequilibrium work required to reach the target value increases to between

120 kcal/mol/Å^2 and 24 kcal/mol/Å^2.

The GLACON colvar is initiated at 25 ns, (step 2) but does not begin to be effective in exerting force until the value of GLADIS is approximately 40 Å or less. Once this GLADIS value is reached, a marked uptick is observed in the value of GLACON (Figure 5.6). Differences in nonequilibrium work

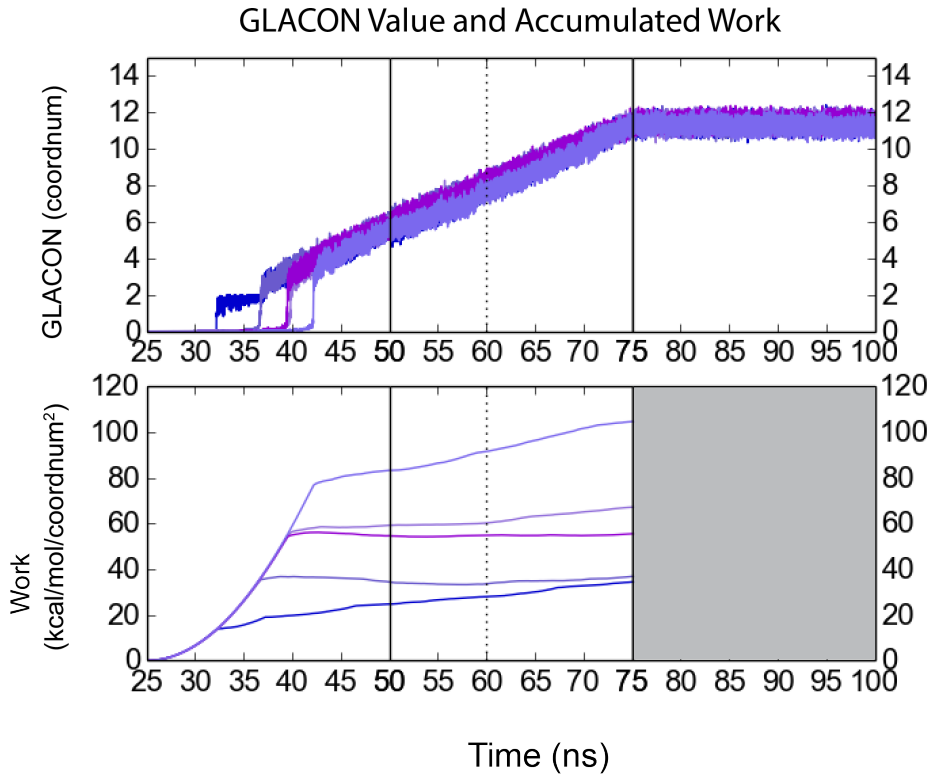


Figure 5.6: GLACON value and nonequilibrium work from 25-100 ns (steps 2-4). The top row shows the value of GLACON, while the bottom row shows the nonequilibrium work. A solid line is used to indicate a change in the colvar protocol, as at 50 ns when the GLADIS colvar becomes a stationary colvar. The dotted line indicates the point at which the force constant k for GLADIS has been reduced to zero. At 75 ns, the GLADIS colvar becomes a stationary rather than moving colvar, and nonequilibrium work is no longer collected. When no nonequilibrium work is collected, this is indicated with a grey box.

between each of the five copies may be caused by differences in how well aligned FX-GLA was with the potential exosite when the GLACON colvar was initiated. At the end of 50 ns, the GLACON collective variable reaches

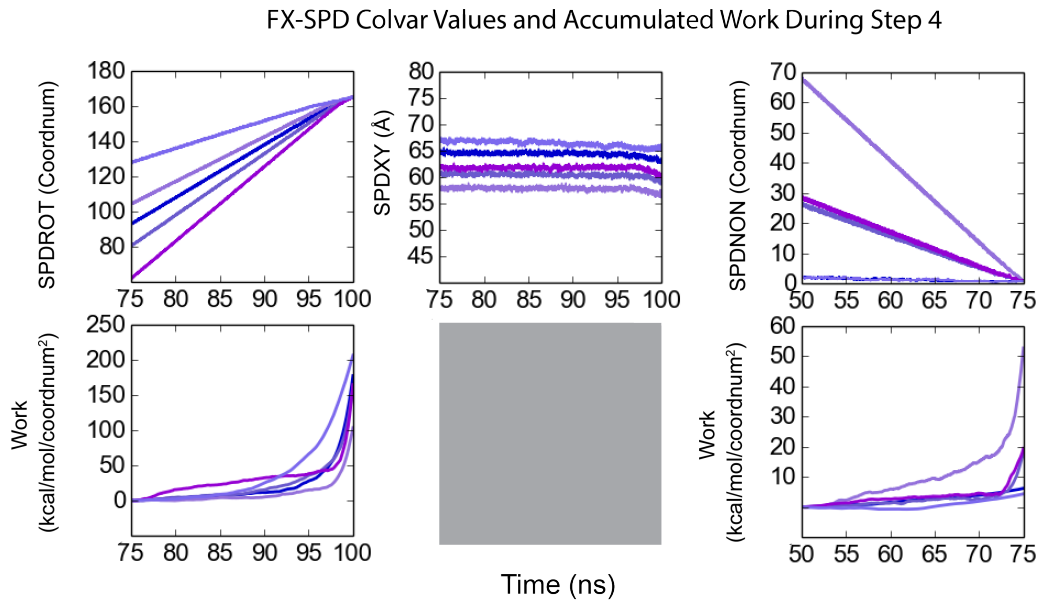


Figure 5.7: Colvar value and accumulated work for SPDROT, SPDXY, and SPDNON. Colvar values are shown in the top row, while the accumulated work is shown in the bottom row. These colvars are used in step 4 to effect rotation of the FX-SPD and align R194 on the FX-SPD domain with the FVII-SPD P1 pocket. SPDXY has no nonequilibrium work reported as it is a stationary colvar, which is indicated using a grey box.

a value halfway to its target, 11.5. GLACON continues to linearly increase until 75 ns, when its target is reached. GLADIS becomes a stationary colvar at 50 ns, and its force constant k is gradually decreased from 10 to 0 over 10 ns.

In summary, both GLADIS and GLACON were progressed linearly toward their targets. Once the force constant of GLADIS is decreased to zero, the value of the distance between GLADIS groups diverges from its target value after 15 ns by no more than 4 Å (Figure 5.5).

To prepare for steps 4-6, in which interactions are induced between FX-SPD and FVII-SPD, the SPDNON colvar is employed to eliminate/prevent contacts forming between FX-SPD and TF:FVIIa (Figure 5.7) In addition, the FX-SPD is rotated so that the R194 residue of FX-SPD faces the FVII-SPD P1 pocket. SPDXY is a stationary molecule which mains the COM distance between the two SPD domains (Figure 5.7).

During step 4, 75-100 ns, SPD is rotated so the groups being biased to form contacts on the SPD domains will have a favorable alignment with unobstructed line of sight between the two groups. During 100-125 ns, (step

5) SPDDIS and SPDCON are activated simultaneously form interactions between the FX-SPD and FVII-SPD. SPDCON is gradually increased over

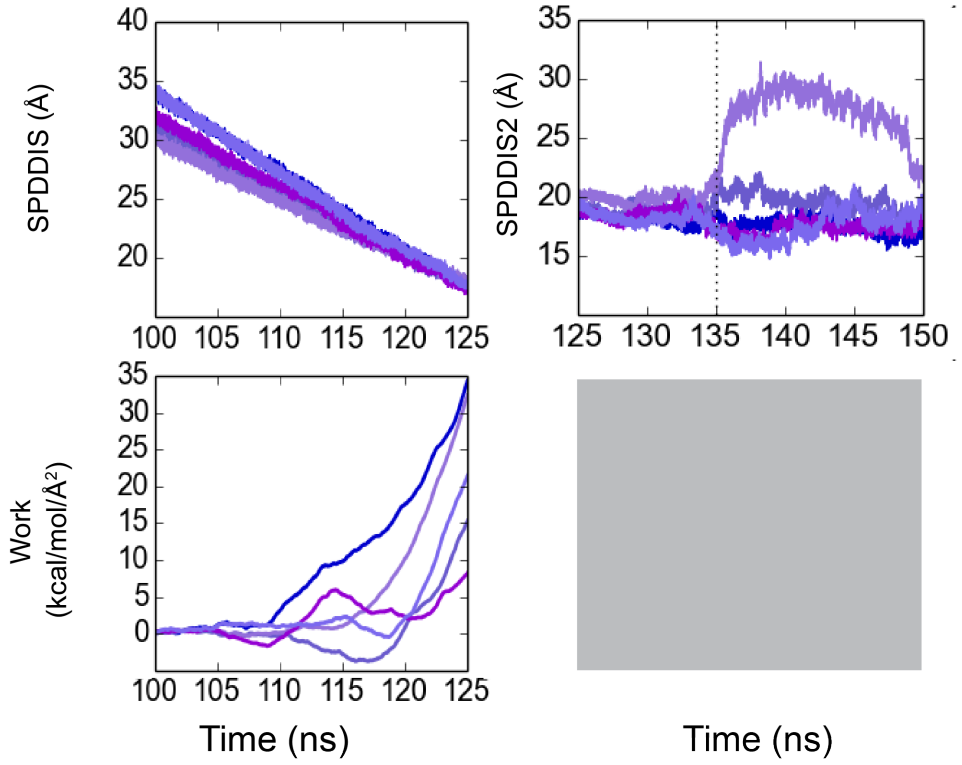


Figure 5.8: SPDDIS value (top left panel) and nonequilibrium work (bottom left panel) from 100-125 ns (step 5). SPDDIS2 value (top right panel) from 125-150 ns. SPDDIS is reduced to 17 Å over 25 ns. SPDDIS2 is held at 19 Å for 10 ns, with force constant gradually reduced to zero. The dotted line indicates the point at which the force constant k for SPDDIS2 has been reduced to zero. SPDDIS2 is a moving colvar, so no nonequilibrium work is collected. A grey box is shown to indicate no nonequilibrium work is collected.

100-150 ns, with a target coordination number of 5. None of the copies reached the target. Three copies reached a value of between 4 and 4.5 Å, while two copies reached coordination values of 3 Å or less. This suggests that the docked confirmation is less favorable for our SPD domains than the FX-GLA and potential exosite confirmation. Nonequilibrium work associated with this colvar was 60 kcal/mol/coordnum^2 or less. SPDDIS is linearly reduced to 17 Å, and accumulated work in each case was 35 kcal/mol/Å^2 or

less. SPDDIS2 is then implemented, which is identical to SPDDIS except for excluding residues used to dock the activation peptide to the P1 pocket of FVIIa (Figure 5.8).

SPDDIS2 is restrained using a stationary harmonic for 10 ns (until 135 ns, while the force constant is gradually reduced from 10 to 0. After the harmonic is phased out, all copies stay within 5 Å of the SPDDIS2 target value except one. The value of SPDDIS2 for this copy rises to 30 Å before lowering to 22 Å by 150 ns. As a stationary harmonic is used, no nonequilibrium work is reported for this colvar.

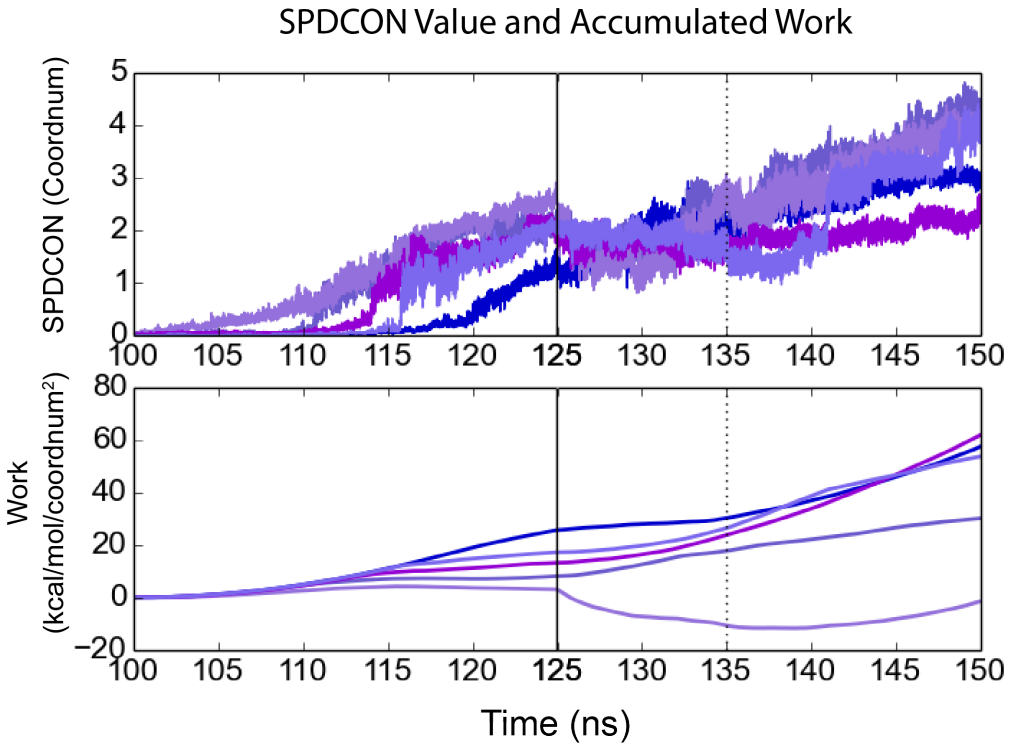


Figure 5.9: SPDCON value and nonequilibrium work from 75 ns-150 ns of simulation (steps 4-6). The top row shows the value of SPDCON, and the bottom row shows the nonequilibrium work. A solid line is used to indicate a change in the colvar protocol, as at 125 ns. The dotted line at 135 ns indicates the point at which the force constant k for SPDDIS2 has been reduced to zero.

While SPDCON and SPDDIS2 are executed (step 6, 125-150 ns), three colvars are used to dock FX-SPD residue R194 to the FVII-SPD P1 pocket. ACTNON is applied in the first 10 ns to eliminate contacts between R194

and the rest of the FX-SPD domain, while ACTDIS simultaneously decreases the distance between R194 and S201 in the P1 pocket. ACTCON is used to induce specific contacts between R194 and specific residues of the P1 pocket. Progress toward centers and nonequilibrium work for these colvars is shown in (Figure 5.10).

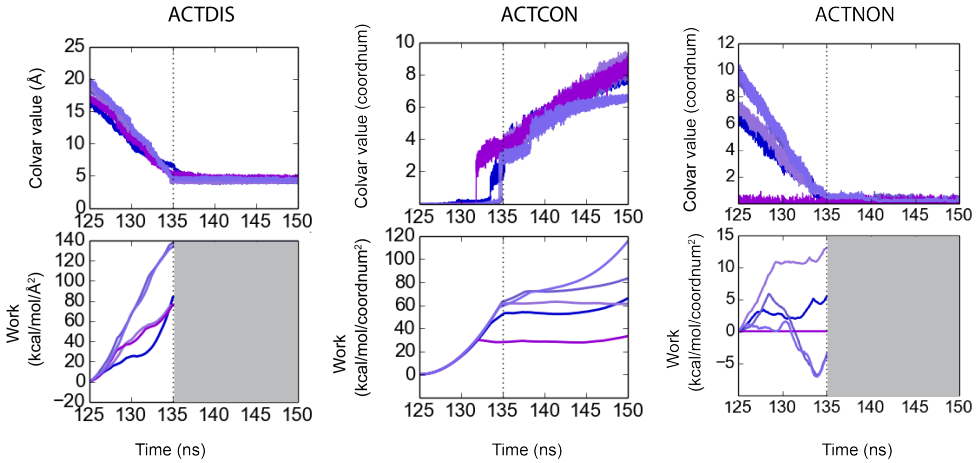


Figure 5.10: Values of ACTDIS, ACTCON, and ACTNON during 125-150 ns. Top row shows the value of the colvar, and the botom row shows the nonequilibrium work associated with that colvar. At 135 ns (indicated with dotted line) ACTDIS and ACTNON become stationary colvars, and so no nonequilibrium work is colected after this point. This is indicated with a grey box.

During this period, GLACON is restrained harmonically to 11.5 and varies between 11 and 12 Å. GLADIS, which has no harmonics applied, gradually converges to approximately 24 Å for all copies except one (initiated from 100 ns pose in the equilibrium simulation) which gradually unbinds (Figure 5.11). No further simulations or were performed with this copy. Analysis of the simulation in which the FX-GLA domain disassociated from TF showed that FX-GLA residues bound in a different location from the stably bound complex copies. While stably bound simulations showed interactions between TF residues 166 and negatively charged residues of the FX-GLA aromatic stack, the unstably associated complex did not. This is suggestive of the importance of interactions between the FX-GLA aromatic stack and residue 166.

The difference in binding mode between stable and unstable complexes may be due to a difference in the initial orientation of TF to the membrane surface.

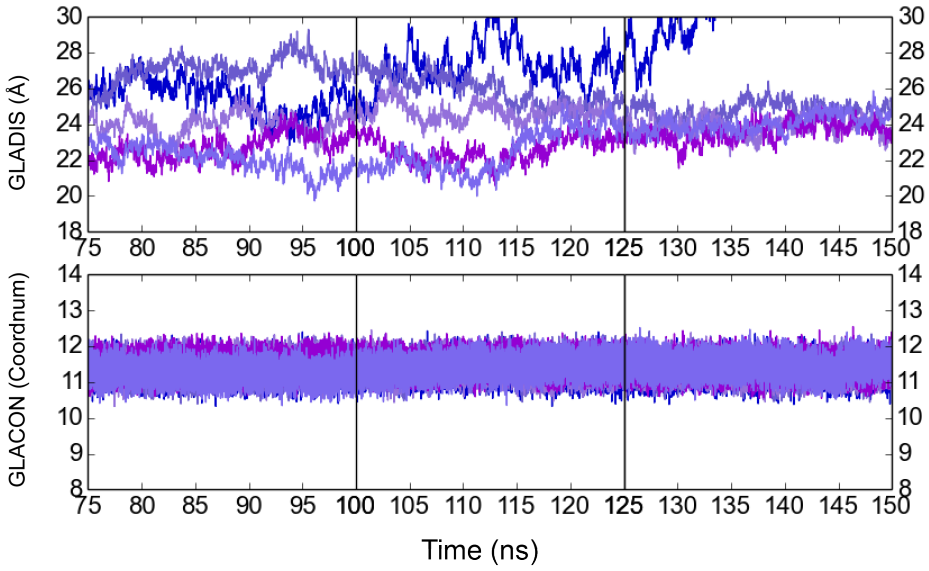


Figure 5.11: GLACON and GLADIS value during 75 ns-150 ns of simulation (steps 4-6). The top row shows the value of GLADIS, while the bottom row shows the value of GLACON. A solid line is used to indicate a change in the colvar protocol. No force is applied to the GLADIS colvar during this part of the simulation.

The value of θ_{CM} for TF in the unstable is copy is higher by approximately 10° than any of the other TF copies during the beginning of step 3, the step in which the colvar protocol for binding FX-GLA to TF reaches its targets for GLACON and GLADDIS (Figure 5.12). During unbinding, θ_{CM} decreases from 110° to 90°. It may be that the increased tilt during step 3 allowed an FX-GLA interaction that was not consistent with stable extrinsic complex binding.

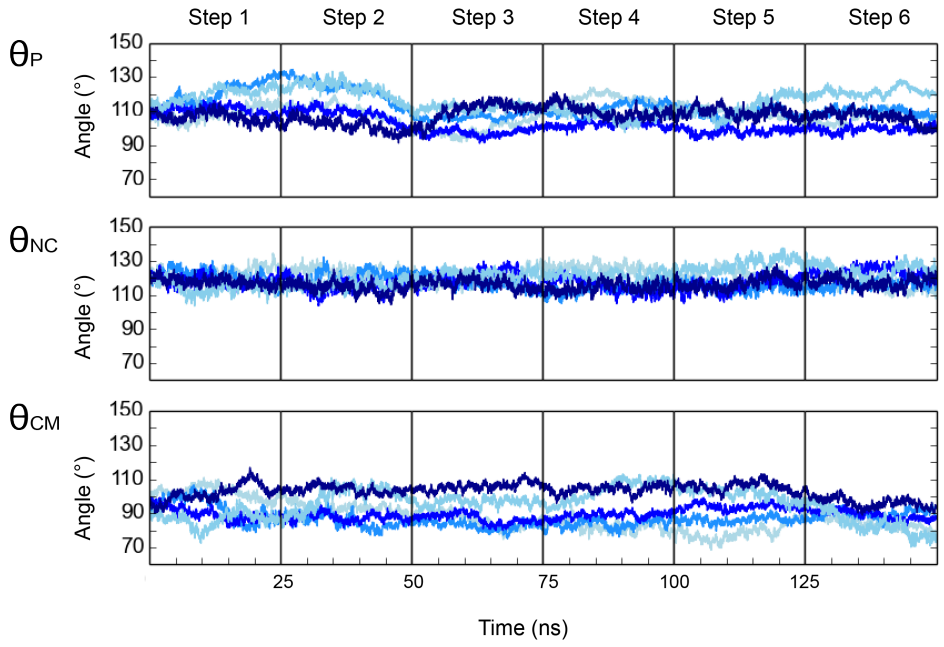


Figure 5.12: Measures of the internal angle between of TF domains and orientation on membrane surface for steps 1-6. Values of θ_{NC}, θ_{CM}, and θ_P over 150 ns of the TF:FVIIa:FX simulations. Data from each simulation is shown using a different shade of blue. Black lines are used to indicate end of each simulation step, when changes are made in the biasing forces and collective variables. Black lines are used to indicate end of each simulation step, when changes are made in the biasing forces and collective variables. All copies stably bound FX-GLA except one. The TF orientation data for the poorly bound complex is shown in dark blue.

Values of the "port-to-starboard" (Figure 5.13) and "bow-to-stern" (Figure 5.14) angles for each GLA domain were tracked during each step of the biased simulation, as were number of lipid phosphate groups within 5 Å FX-GLA residues (Figure 5.15). Bow-to-stern tilt is initially elevated for one copy

to 110°during after step 3. By the end of step 6, all bow-to-stern tilt are within the 90-100° range expected based on our previous simulations [179]. Port-to-starboard tilt remains in the 60-120° expected for that metric in our previous simulations of FX [179] Fluctuations are seen in the orientations of individual domains, but by the end of step 6 however, all FX-GLA have orientation values within 10°, of those observed on initiation at the start of step 1. Number of lipids in vicinity of FX-GLA is relatively constant throughout the simulations, with an average of 6 lipid phosphate within 5 Å of FX-GLA C_α and Ca^{2+} atoms . This is consistent with variation in the last 10 ns of our equilibrium simulation of solitary FX. Orientation and binding mode of FX-GLA following association with TF:FVIIa is thus consistent with previously characterized FX binding orientations.

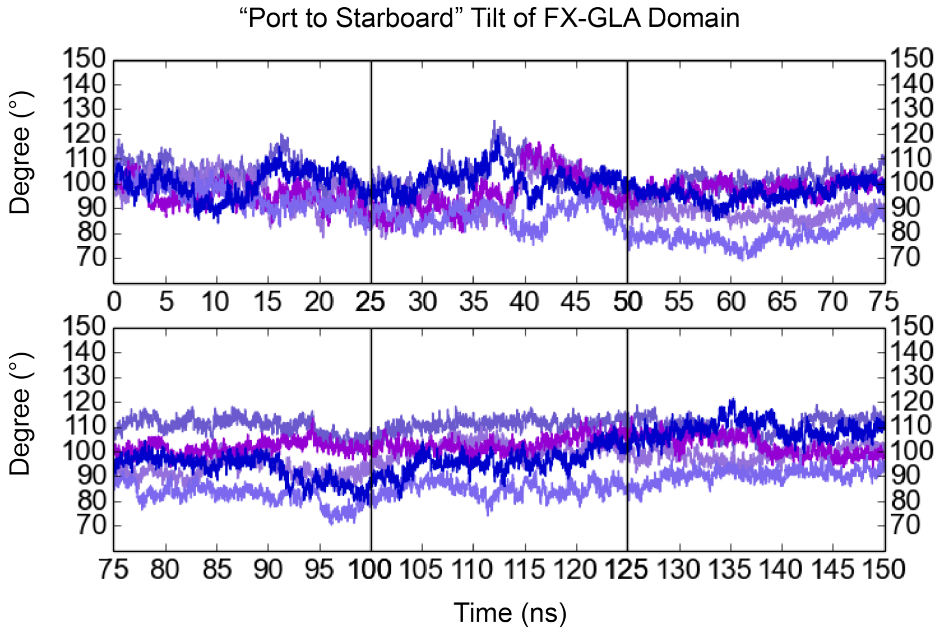


Figure 5.13: "Port-to-starboard" angle of FX-GLA to the membrane normal while in the TF:FVII:FX complex for steps 1-6. Data from each simulation is shown in a different shade of purple. Black lines indicate the end of each step, when collective variables are modified.

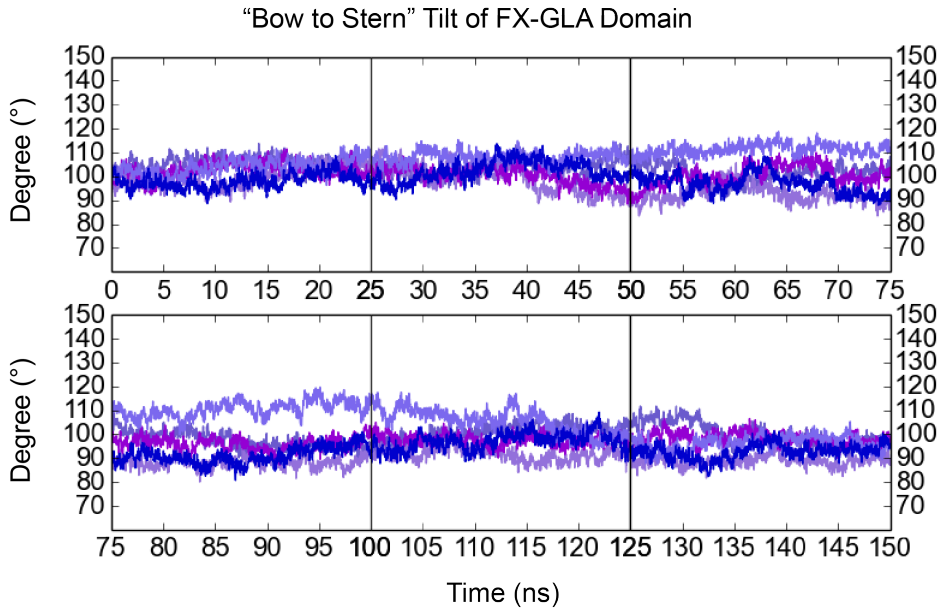


Figure 5.14: "Bow-to-stern" angle of FX-GLA to the membrane normal while in the TF:FVII:FX complex for steps 1-6. Data from each simulation is shown in a different shade of purple. Black lines indicate the end of each step, when collective variables are modified.

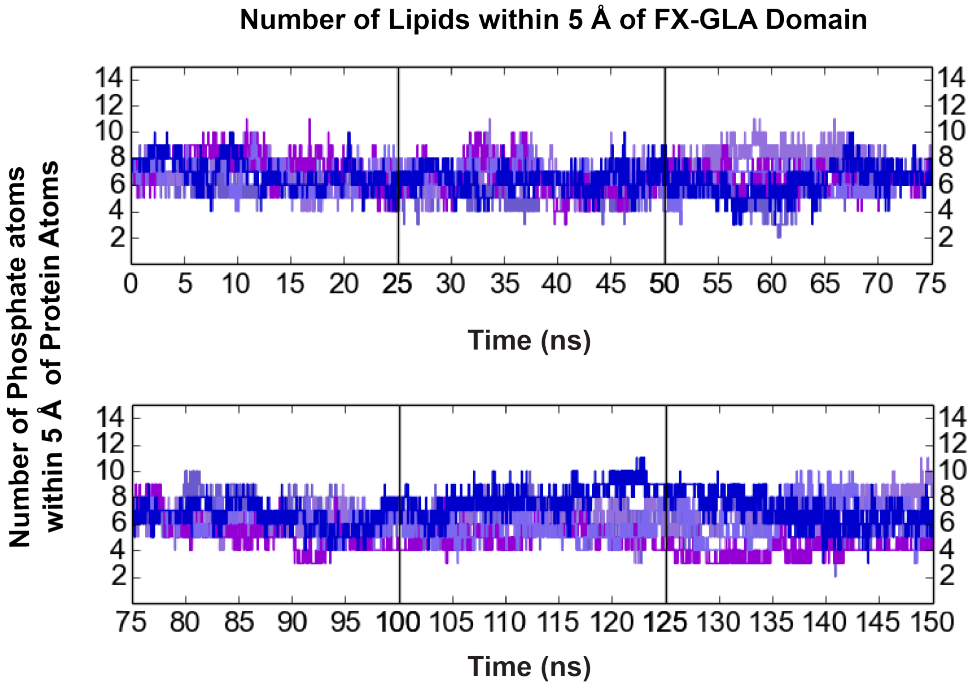


Figure 5.15: Number of phosphate atoms within 5 Å of FX-GLA domain atoms for nonequilibrium simulation steps 1-6. Data from each simulation is shown in a different shade of purple. Black lines are used to indicate end of each simulation step, when changes are made in the biasing forces and collective variables.

5.7 Complex characteristics following equilibration

Over the course of 50 ns, all biases were gradually removed from the simulations (step 7). Each complex was equilibrated for 200 ns in HMMM (step 8). The membrane was then converted to standard lipids and the system simulated an additional 150 ns (step 9). Orientation of FX-GLA, FVII-GLA and TF were assessed during steps 7-9 and was consistent with orientation measures from the end of equilibium simulations for these proteins (within 20°). The values of GLACON, GLADIS, SPDCON, and SPDDIS target distances/contacts were assessed during each of these simulations. Values of GLACON and and GLADIS remained within 4 coordination numbers and 5 Å of their nonequilibrium simulation target values, respectively (Figure 5.16). SPDDIS also remains within 5 Å its target value, while SPDCON in each case decreases to between 2-3 coordinations numbers (Figure 5.16).

The reduction of coordination number in SPDCON is not correlated to reduced total contacts between FX-SPD and FVII-SPD, however. Average contact number (number of residues within 3 Å) was calculated for each residue of FX-SPD and FVII-SPD. A number of residues showed increased contact number from simulation step 7 to 9. Residues with contact number greater that 0.5 are shown in Figure 5.18. FVII-SPD residues in contact with FX-SPD include 17, 153-154, 183, 201-204, and 227-234. Of these, 153, 204, and 229 were included in groups defining the SPDCON collective variable. FX-SPD residues in contact with FVII-SPD included 194-197, 199, 202, 327-328, 332, 336, 341, 367, 370-372, and 386. Of these, 195 and 332 were part of the collective variable. Binding between FX-GLA and TF showed stable interactions between residues which had been used to define GLACON. Stably binding residues are shown in Figure 5.19 C. Residues with significant numbers of contacts included tissue factor residues 162-164 and 166. Residues 163, 164, and 166 were all identified experimentally as causing impaired FX activation when mutated [235]. All of these residues showed greater numbers of contacts in step 9 as compared to step 4. This suggests a favorable initial positioning at the end of the biased simulations which allowed FX-GLA to increase protein-protein contacts with TF. The residues of FX most frequently interacting with FX include GLA residues 13, 30-33, and 36. Contacts with TF were found to increase with TF following removal of biasing forces. All of these residues were identified through docking as part

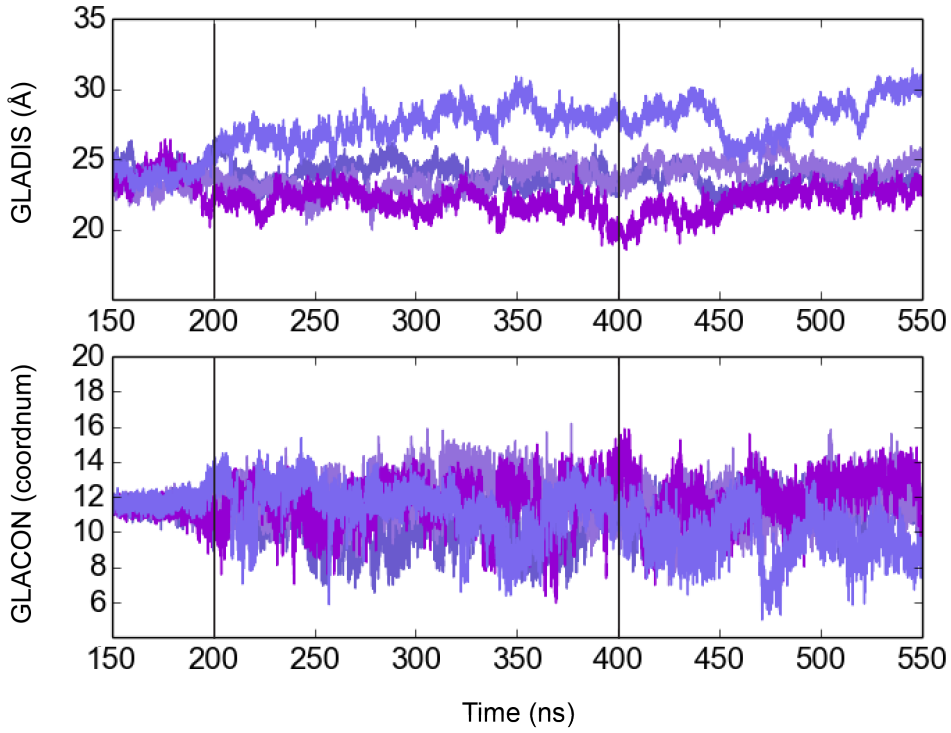


Figure 5.16: GLACON and GLADIS value during simulation steps 7-9. The top row shows the value of GLADIS, while the bottom row shows the value of GLACON. A solid line is used to indicate the end of each step, when the biasing forces are removed (end of step 7) and when the membrane is converted to standard lipid (end of step 8). No force is applied to the colvars during this part of the simulation.

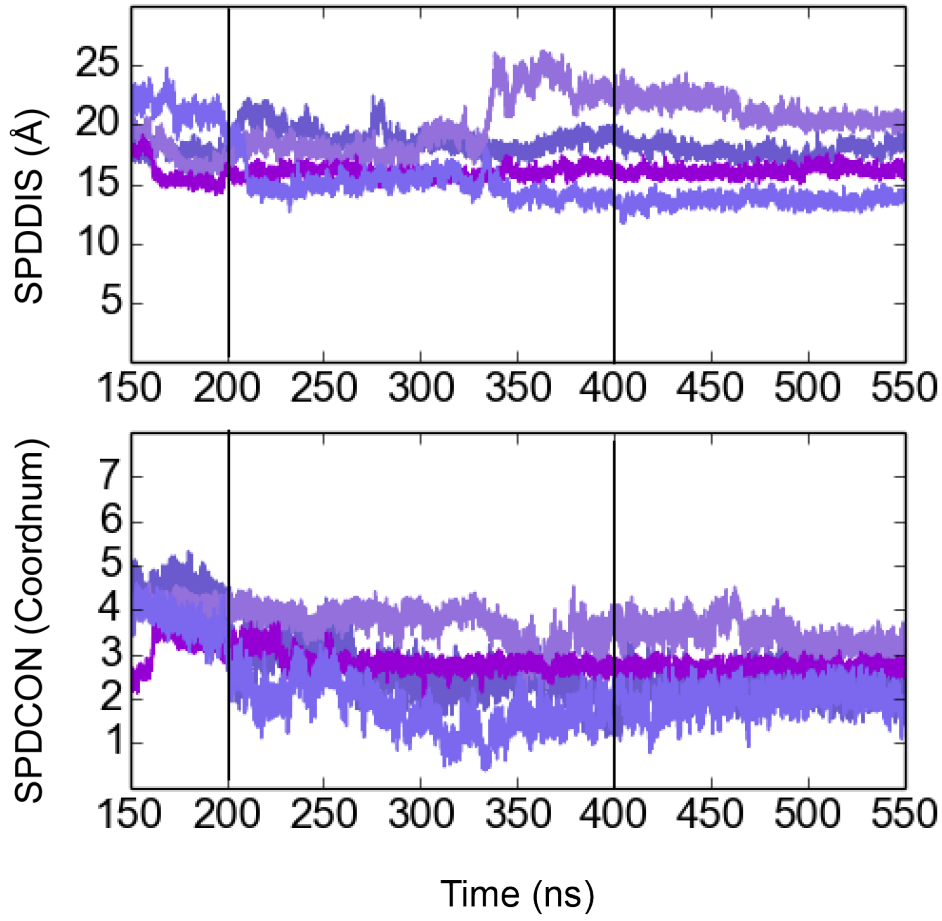


Figure 5.17: SPDCON and SPDDIS value during simulation steps 7-9. The top row shows the value of SPDDIS, while the bottom row shows the value of SPDCON. A solid line is used to indicate the end of each step, when the biasing forces are removed (end of step 7) and when the membrane is converted to standard lipid (end of step 8). No force is applied to the colvars during this part of the simulation.

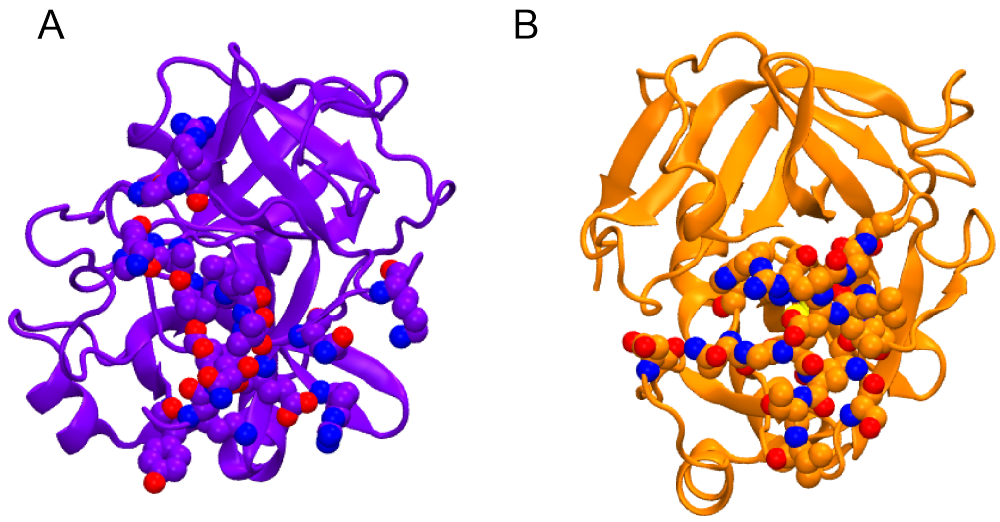


Figure 5.18: Residues of FX-SPD and FVII-SPD which interact. A) FX-SPD, shown in purple. Residues are illustrated in balls having an average contact number (defined as number of FVIIa residues within 3 Å) of 0.5 or greater. Blue balls represent nitrogen, purple balls carbon, and red balls oxygen. B) FVII-SPD, shown in orange. Residues illustrated with balls having an average contact number (defined as number of FX-SPD residues within 3 Å) of 0.5 or greater. Blue balls represent nitrogen, orange balls carbon, and red balls oxygen atoms

of the likely contact site between FX-GLA and the TF potential exosite, and included in the GLACON colvar.

We selected one of the stably bound TF:FVIIa:FX as the best of the four model based the model having the largest number of FX-GLA to potential exosite contacts. Residues of the TF exosite bound primarily to residues of the FX-GLA aromatic stack in what could be described as a "double zipper"; on one side of the 162-166 loop the two domains interact using electrostatic interactions, while on the other they interact using hydrophobic packing (Figure 5.20). Residues GLA32 and ASP33 form the electrostatic component of this interaction with TF residue LYS166. Mutations of GLA32 have been implicated in factor X deficiency syndromes (Tokyo Mutation) [262]. In addition, mutations to the TF residues which pack with FX-GLA residues in the hydrophobic part of the interaction, SER162 and SER163, have recently been shown to decrease FX and FIX activation rates [263]. It appears that TF residue K166 may bind more frequently to FX-GLA than K165. Experimental studies had demonstrated a double mutant of both of these lysine residues significantly affected the extrinsic complex's FX activation rate [264, 236].

Interestingly, experimental mutagenesis studies have shown similar decreases in activation rate for FIX when residues of the TF exosite are mutated [235]. FIX is a worse substrate for FVIIa than FX however, activated at a lower rate by wild-type TF:FVIIa than FX [264, 235]. Additionally, certain residues whose mutation effects FX activation seem to have a lesser effect on FIX activation. This included TF residues 159 and 164, which in our simulations interacted with residues of the FX aromatic stack [235]. It is possible these observed differences are due atomic-level differences between FX and FIX in the aromatic stack region. Figure 5.21 compares the FX, FIX and FVII GLA sequences in the aromatic stack region, residues 31-36. These GLA domains are both structurally and sequentially very similar, but show the highest number of sequence differences in the 32-36 (Figure 5.21B), highlighted in green). A major difference is substitution of FX residue D33, a negative residue, to an amphiphilic neutral residue in FIX (Figure 5.21A).

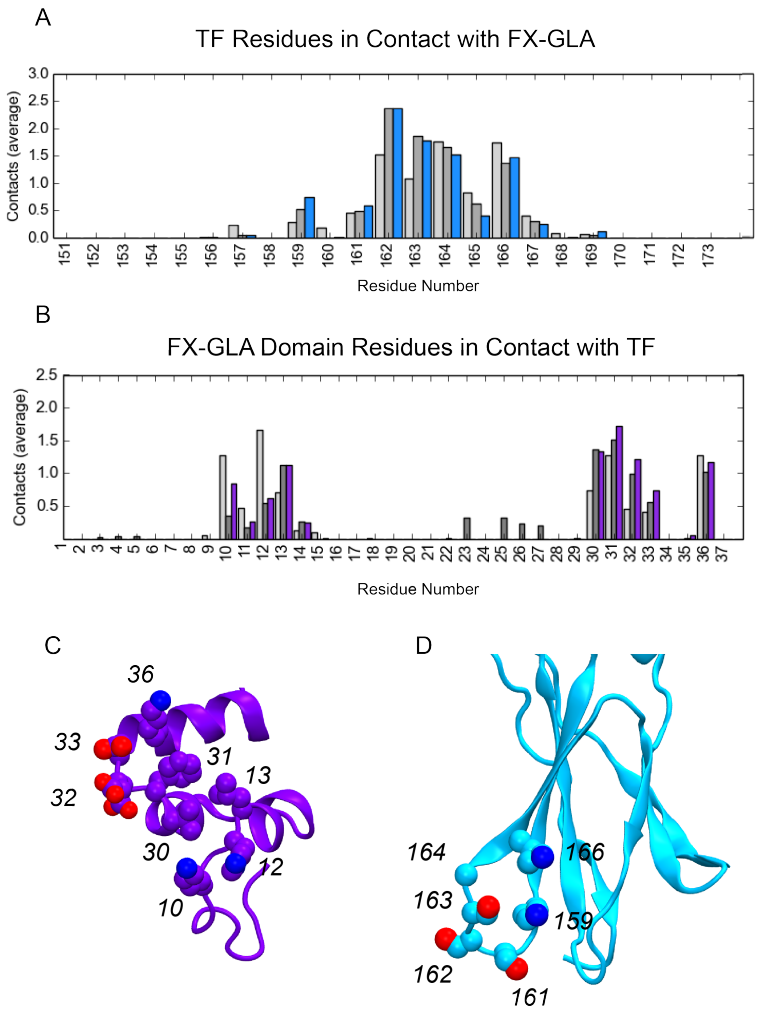


Figure 5.19: Contacts between TF and FX-GLA. A) TF residues interacting with FX-GLA. A contact is defined as number of protein FX-GLA residues within 3 AA of each TF residue. Contacts shown with average value of 0.5 or greater. Blue shows the average contact value over 150 ns is step 9, dark grey shows the average contact value over 200 ns is step 8, and light grey shows the average contact value over 25 ns in step 4. B) FX-GLA residues interacting with TF. A contact is defined as number of protein TF residues within 3 Å of each FX-GLA residue. Contacts shown with average value of 0.5 or greater. Purple shows the average contact value over 150 ns in step 9, dark grey shows the average contact value over 200 ns in step 8, and light grey shows the average contact value over 25 ns in step 4. C)FX-GLA residues with average contact number of 0.5 or greater. Blue balls represent nitrogen, purple balls carbon, and red balls oxygen atoms. D)TF residues with average contact number of 0.5 or greater. Blue balls represent nitrogen, orange balls carbon, and red balls oxygen atoms.

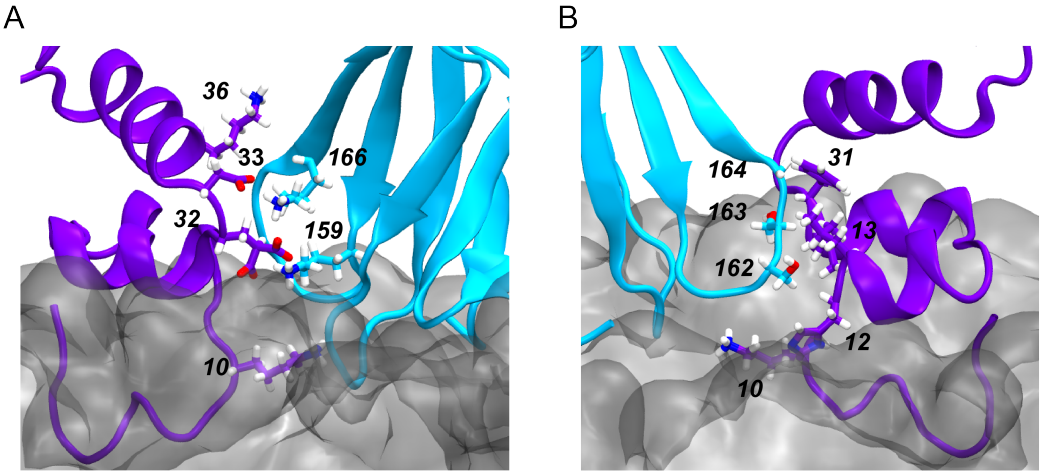
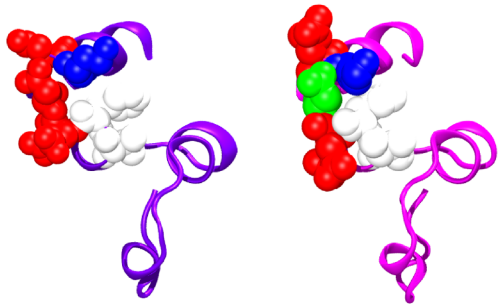


Figure 5.20: Double-zipper interaction of FX-GLA and TF in the TF:FVIIa:FX complex at the end of equilibrium simulations in standard membrane (550 ns total simulation). FX shown in purple, TF in light blue, and membrane lipids in grey space-filling illustration. A) Primarily electrostatic interactions illustrated between residues of FX and TF, with important residues shown with sticks. Carbon atoms shown in purple, oxygen shown in red, nitrogen atoms in blue, and hydrogen atoms in white. Residue numbers listed next to each residue. B) Complex rotated 180°, illustrating primarily hydrophobic interactions between residues of FX and TF, with important residues shown with sticks. Carbon atoms shown in light blue, oxygen atoms in red, nitrogen atoms in blue, and hydrogen atoms in white. Residue numbers listed next to each residue.

A



B

	10	20	30
Factor X	AN-SFLγγMKKGHLγRγCMγγTCSYγγARγVFγDSDKT		
Factor IX	YNSGKLγγFVQGNLγRγCMγγKCSFγγARγVFγNTγRT		
Factor VII	AN-AFLγγLRPGSLγRγCKγγQCSFγγARγIFKDAγRT		

Figure 5.21: Comparison of the TF binding region of FX-GLA with the equivalent region of FIX-GLA. A) Illustration of FX-GLA in purple, FIX-GLA in pink. Residues 30, 31, 32, 33, and 36 are illustrated with balls. Hydrophobic residues shown in white, positively charged residues in blue, negatively charged residues in red, and amphiphilic residues in green. B) Sequence alignment of the GLA domains of FVII, FX, and FIX. Identical residues are bold. Residues 32-36 are highlighted in green.

5.8 Conclusions

We have induced the formation of the extrinsic complex using a novel simulation protocol on the membrane surface. We identified a putative binding surface for the FX-GLA, consisting of residues on the surface of the aromatic stack. We propose these residues may bind to the TF exosite in a "double-zipper" formation of alternating electrostatic and hydrophobic interactions close to the surface of the lipid bilayer. This interaction mode is consistent with previous experimental studies [264, 236, 263].

We have also induced binding of the FX-SPD and FVII-SPD domains, with residues of FVII-SPD interacting with FX-SPD which have been shown to reduce FX activation rates when mutated. Conclusions which can be drawn as to SPD domain binding are limited as an activated FX-SPD was used in construction of our FX model, and residues in activation peptide that might otherwise interact with FVII-SPD are missing from FX-SPD. Binding of the domains was stable for 350 ns of simulation following the elimination of biasing forces, with protein-protein interactions consistent with observations of previous experimental data [233].

This work has allowed for development of the first atomic-level model of the extrinsic complex on a bilayer, with interactions predicted from rigid-body docking calculations both stable within the simulation and consistent with previous experiment. While our simulations are limited by the short timescales provided by simulation methodologies, we believe this study can provide a basis for systematically designed studies probing FX binding in the extrinsic complex, particularly in the region of the potential exosite.

6 ADDITIONAL SIMULATIONS OF COAGULATION PROTEIN INTERACTIONS

6.1 Introduction

Here we present simulations of FIX-GLA spontaneous binding to a 1:1 PS:PC HMMM bilayer, as well as development of an atomic-level model of TF:FVIIa:XK1. While these studies are preliminary, we believe they are valuable in having the potential to form the basis of further investigation in future experimental and computational work.

FIX is an important coagulation factor which is a component of the intrinsic coagulation cascade. Deficiencies of appropriately active FIX are responsible for a common form of hemophilia, hemophila B [265, 266]. Like other coagulation factors studied in the preceding chapters,FIX binds to phospholipid membranes using a GLA domain. The GLA domain of FIX has a high degree of both sequence and structural similarity to FX-GLA and other GLA domains of proteins in the coagulation cascade. Like FX, FIX has been found to be most active in the presence of sufficient concentrations of PS [267]. While both FIX and FX can serve as substrates to TF:FVIIa, FX has been found to be a superior substrate [235]. Our previous work modeling the extrinsic complex suggests binding FX-GLA interacts with tissue factor in the extrinsic complex through its Given the high degree of sequence and structural similarity between the FIX-GLA and FX-GLA.

Here, six copies of FIX-GLA are simulated binding spontaneously to a 1:1 PS:PC HMMM over 100 ns of equilibrium molecular dynamics simulations. We find that all the FIX-GLA converge in terms of binding depth and orientation to the membrane plane. We assess lipid interactions of FIX-GLA residues, and assess for potential source for the differential affinity of FIX and FX as substrates in the extrinsic complex.

TFPI is a potent inhibitor of FXa, and forms a quaternary complex of FXa-TFPI-FVIIa:TF, blocking both FXa and FVIIa activity [268, 269, 270]. The TFPI kunitz domains bind SPD active sites, the first kunitz domain (KD1)

to FVII-SPD, while the second kunitz domain binds to FX-SPD [271]. XK1 is a hybrid construct which has been studied experimentally in which the FX-SPD domain is replaced by the first kunitz domain (KD1) of TFPI [272]. Through the inhibitory binding of the KD1 domain, may stabilize binding of the FX light chain to TF:FVII. This would allow the TF:FVIIa:XK1 complex to provide a superior subject compared to the extrinsic complex for experimental characterization through cryo-EM.

Here we report development of a model of TF:FVIIa:XK1 at an atomic level using molecular dynamics simulations. The TF:FVIIa:XK1 complex was simulated for 500 ns on an HMMM 1:1 PS:PC bilayer. The membrane was then converted to a conventional representation and simulated an additional 195 ns. At the initiation of the simulation, FVII-SPD was bound to KD1. We initially modeled the FX lightchain and TF:FVIIa complex separately on the surface of the bilayer. During the HMMM simulation, the FX lightchain and TF:FVIIa spontaneously bound, an association that was stable for the remainder of the simulation.

6.2 Methods

6.2.1 Simulation protocols

All simulations were performed using NAMD [86, 226] and the CHARMM36 [111] force-field parameters. Parameters of GLA residues were those previously developed by our lab [53] with atom types renamed as appropriate for the conversion from CHARMM27. The NPT conditions and TIP3P water model was used for all simulations. Constant pressure was maintained at a target of 1 atm using the Nosé-Hoover-Langevin piston method [164]. Constant temperature was maintained by Langevin dynamics with a damping coefficient, γ, of $1\,\mathrm{ps}^{-1}$ applied to all atoms [165]. All simulations used a timestep of 2 fs and a temperature of 310 K. The particle mesh Ewald (PME) method [167] was used for long-range electrostatic calculations with a grid density greater than $1/\,\text{Å}^3$. Non-bonded interactions were cut after 12 Å with a smoothing function after 10 Å. System setup, visualization, and analysis were carried out using VMD [213].

6.2.2 Spontaneous membrane-binding simulations of FIX-GLA

A crystal structure of FIX-GLA from Protein Data Bank entry 1J35 [57] was used as the starting structure for all GLA domain simulations. Using the PSFGEN plugin of VMD [213], an N-terminal ammonium group and a C-terminal N-methylamide C-terminal group were added, along with hydrogen atoms. The neutral N-methylamide group was chosen because the C-terminus of the FIX-GLA in the full-length FIX protein would be connected to the first epidermal growth factor (EGF) domain, and a C-terminal patch with a default charge would introduce an artificial charge to this region. FIX-GLA was next placed in a $60\times60\times60\,\text{Å}^3$ water box using the SOLVATE plugin of VMD and neutralized with Na^+ ions added using the AUTOIONIZE plugin of VMD. This solution system was then equilibrated for 1 ns after 2,000 steps of energy minimization.

Our HMMM membrane construct was used in the binding simulations to increase lipid dynamics and allow for increased sampling and GLA domain insertion into the membrane. Three independent 1:1 PS:PC HMMM lipid patches were generated. Each patch was $50\times50\times114\,\text{Å}^3$ in size and was constructed and solvated using the HMMM Builder [180] in CHARMM-GUI [224, 225].

Six independent orientations of FIX-GLA (a rotation of 15° in the xz or yz plane) were generated above two sets of the three PS:PC HMMM membrane patches. In the beginning of each simulation FIX-GLA was placed at least 3 Å above the membrane patch. Each system was neutralized with Na^+ ions using the AUTOIONIZE plugin of VMD. The final systems were each approximately 26,000 atoms. The z-position of HMMM short-tailed lipids was restrained harmonically ($k = 0.05\,\text{kcal/mol/Å}^2$ applied to C21 and C31 lipid tail atoms) throughout the simulations in order to more closely reproduce the natural distribution of the head-group atoms. Each system was then simulated for 100 ns.

6.2.3 FIX-GLA simulation analysis

Depth of the GLA domain insertion into the membrane was assessed by calculating the z components of center of mass (COM) of these groups, in-

dividually, relative to the midplane of phosphate groups (phosphorus atoms and four bonded oxygen atoms) of lipids: (1) backbone atoms in residues 5, 6, and 9 in the keel of FIX-GLA and (2) Ca^{2+} ions 1-7 (Ca-1 to Ca-7). Similarly, the z components of lipid carboxy groups (carbon and two bonded oxygen atoms), lipid amino groups (nitrogen and three bonded hydrogen atoms) were also calculated relative to the phosphate plane.

Orientation of the GLA domain relative to the membrane plane was assessed using two angles defined in two perpendicular planes, termed "port-to-starboard" and "bow-to-stern." The bow-to-stern angle is the angle between the membrane normal (z-axis) and the vector connecting the COM of Ca-1 to Ca-3 and Ca-4 to Ca-6. These are the same measures as previously used to analyze orientation of the domain on the membrane in our previous study of the FX-GLA domain membrane binding [179]. The port-to-starboard angle is the angle between the membrane normal and the vector connecting the COM of backbone atoms in residue 14, 15, 30, and 31 and the COM of backbone atoms in 20, 23, 24 and 25. These two sets of residues represent opposite ends of the two helices directly above the level of the coordinated Ca^{2+} ions. Thus, the bow-to-stern angle measures the angle between the stern and bow ends of FIX-GLA and the membrane normal, while the port-to-starboard angle represents the tilt of the port and starboard ends relative to the membrane normal.

Contacts between lipid head-group charged moieties and FIX-GLA (bound Ca^{2+} or protein residues) were calculated based on distance cutoffs which were determined using radial pair distribution function calculations, as in our previous study of the FX-GLA domain [179]. Number of contacts here is the number of atoms on one group (e.g., a carboxy group) within a specified cutoff of the second group (e.g., a specific protein residue). Distance between all protein and Ca^{2+} atoms and phosphate, amino, and carboxy atoms were calculated, and distances equal to or less than the cutoff were determined to be a contact. The radial pair distribution function was calculated separately for distances between the residue type of interest (protein or Ca^{2+} ions) and the lipid head-group moiety of interest (carboxy, amino, or phosphate). The calculation sampled every 10th frame of the final 10 ns of the 100 ns HMMM simulations. For protein residues and Ca^{2+} ions, all atoms were included in the contact calculation. Cutoffs determined were 2.50 Å for phosphate-Ca^{2+}, 3.00 Å for carboxy-Ca^{2+}, 2.25 Å for phosphate-protein, 2.50 Å for carboxy-

protein, and 2.25 Å for amino-protein contacts.

6.2.4 Constructing the TF:FVII:XK1 complex

To construct our model of the XK1, we derived KD1 from PDB 4BQD [268]. This was placed in a water box of size 57x41x51 $Å^3$, Na^+ and Cl^- ions added to bring concentration to 250 mM, and simulated for 5 ns. The last frame was used as an initial starting structure for the KD1:FVII:SPD complex. The last frame of the TF:FVII 100 ns simulation previously performed in our study of the TF:FVII:FX complex (see Chapter 5) was used to as the starting point for the TF:FVIIa:KD1 complex.

In order to model the binding of FVII-SPD to KD1, FVII-SPD and KD1 were aligned using the 1TFX [273] PDB structure as a template. In 1TFX, porcine trypsin SPD is bound bound the TFPI KD2 domain. Trypsin SPD as a high degree of structural similarity to FVII-SPD, while KD2 has a high degree of structural similarity to KD1. The TF:FVIIa:KD1 complex was then simulated on the surface of the HMMM membrane. Five simulations were performed for 5 ns with force constant k on the backbone atoms and C21 C31 atoms gradually decreased from 2.5 to zero with steps of 0.5.

Model the complete TF:FVIIa:XK1 structure, the FX light chain was constructed using our previously developed FX model. The 63 ns frame of the 100 ns simulation of membrane-bound FX was used as the starting structure for the lightchain and bound lipids. This snapshot was chosen as the light chain is most extended in the z-direction at this point in the simulation, bringing it closest in z-position to KD1 in our TF:FVII:KD1 structure. FX sequence used is 1 to 127, with additional FX residues removed.

In order to connect KD1 to the FX lightchain, the XK1 sequence from residues 128-138 was modeled using the 2016 MOE [197] homology-based loop fitting method.The loop chosen had the highest sequence identity of template loop to XK1 (1st criteria) and rough loop score (second criteria). Two structures were found with 27.3% sequence identity to the 128-138 XK1 loop, and the loop with the lowest S score selected.

The second 10 residue loop for XK1 residues 200-210 was added using MOE denovo loop modeling. The loop was added in PSFGEN and then refined in MOE. The loop structure with the lowest energy was selected. As

no homology loop in the library was had greater than 25% sequence identity, the MOE denovo loop modeler method was used. The top ranked (s rank) one-hundred loops out of 1,000 were built, and of those the loop with the lowest energy (GB/VI value) used in our model.

To complete the TF:FVIIa:XK1 structure, the complete XK1 structure structure was aligned to the KD1 domain in the structure of TF:FVIIa:KD1 described above. The XK1 GLA domain was aligned to the FX-GLA domain from the 63 ns frame of the 100 ns FX equilibration simulation. All lipids within 5 Å of the FX-GLA domain were selected, and combined with the protein into the full TF:FVII:XK1 structure. The same number of lipids were removed from the lipid patch used in the TF:FVII:KD1 structure, and the complete XK1 structure with associated lipids combined with the TF:FVIIa:KD1 lipid box. A 5 ns equilibration simulation was performed with lipids restrained with a 0.05 kcal/mol/Å^2 force constant at the C31 C21 atoms and protein backbone restrained at 3 force constant. The phosphate atoms of lipids within 5 Å of the GLA domain was also restrained at force constant 3 kcal/mol/Å^2.

6.2.5 Simulation of the TF:FVII:XK1 structure

A 500 ns HMMM equilibrium simulation was performed lipid C21 C32 atoms restrained in the Z-axis with a 0.05 kcal/mol/Å^2 force constant. The system was then converted to a conventional (full-tailed) membrane, using a PDB generated from the final frame of the 500 ns HMMM simulation. We used 1-palmitoyl-2-oleoyl (PO) tails to produce complete POPS and POPC lipids. A full-tailed lipid was overlaid onto each short-tailed lipid, and coordinates of the missing atoms were then copied from the full-tailed lipid, thus allowing for head-group interactions obtained from the HMMM simulations to be preserved. The layer of 1,1-dicholoroethane (DCLE) solvent molecules were then removed from the center of the HMMM bilayer. After a 2,500-step energy minimization, each membrane-bound FVII-GLA system was simulated in a full membrane for additional 195 ns.

6.2.6 TF:FVIIa:XK1 complex analysis

Center of mas distance was measured between XK1-GLA COM and TF-FIB2 COM and between FVII-SPD to XK1-KD1. C_α atoms were used to calculate each center of mass of each domain using the "measure center" command in VMD. Number of contacts between protein and Ca^{2+} residues of TF:FVIIa and XK1 were assessed using the "measure contacts" command in VMD. Each protein or Ca^{2+} atom of TF:FVIIa within 3 Å of FX was considered to be in contact with FX. Contacts were averaged over the entirety of the 195 ns conventional membrane simulation.

6.3 Spontaneous membrane insertion of FIX-GLA

In order to characterize the membrane binding of the FIX-GLA domain, six copies of FIX-GLA with distinct initial orientations were placed above 1:1 PS:PC HMMM membrane patches and simulated 100 ns. Two distinct membrane patches were used, each contributing initial lipid configurations for three of the six FIX-GLA domain simulations. Five of the six membrane-bound models bound spontaneously and inserted their keel below the level of the phosphate groups by 100 ns. One FIX-GLA bound improperly as defined by the methods and was eliminated from further analysis.

The five properly bound FIX-GLA engaged the HMMM membrane and converged in binding depth by 80 ns of simulation (Figure 6.1). The COM of the keel residues reaches a depth of up to 5 Å below the PS phosphate plane. The COM of the Ca^{2+} ions, which interact with the membrane during binding, converges to a z-position of 5-7 Å above the phosphate plane. FIX-GLA bound more shallowly to the bilayer than FX-GLA by about 5 Å of depth for both Ca^{2+} COM and keel residue COM.

Of the five FIX-GLA simulations, all five converged in terms of their orientation (Figure 6.2). For the five converged FIX-GLA domains, the "bow-to-stern" orientation of 20-60° in the last 10 ns of the 100 ns simulation. FIX-GLA tilt in the "port-to-starboard" orientation ranged from 90-130°. Interestingly, the orientation which the majority of FIX-GLA take on is significantly different than that found for FX-GLA and FVII-GLA in our previous work. For both FX-GLA in PS, the "bow-to-stern" tilt ranged from approximately 75-100°. The "port-to-starboard" tilt for FIX is relatively

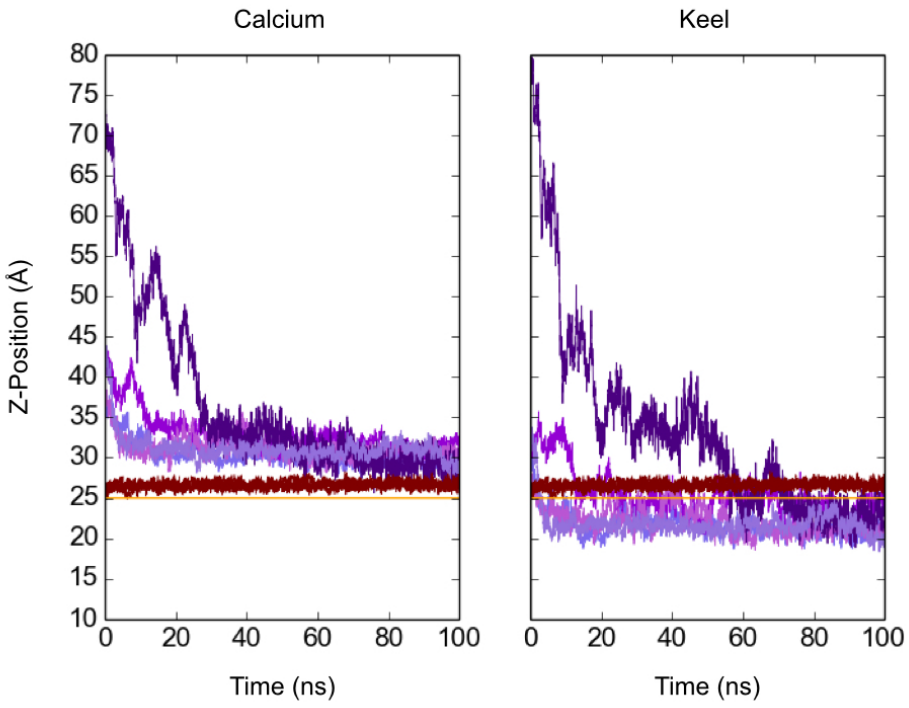


Figure 6.1: Membrane insertion and binding of the FIX-GLA domain. Relative z-positions of the COM of relevant moieties are shown, including: lipid phosphate (orange) and choline groups (maroon); Ca^{2+} ions (left panel); and keel protein (LYS5, LEU6, and PHE9) backbone atoms (right panel). See Methods for atoms included in phosphate and choline groups. Data in this figure refer only to the 5 properly bound FIX-GLA domains. See Methods for definition of "properly" and "improperly" bound.

similar to FX and FVII-GLA by comparison, with approximate values of 90-120° found for FX-GLA. While FX and FIX share similar specificity to PS lipids, we found orientation of FX-GLA in the membrane to be more similar to FVII-GLA in the bilayer.

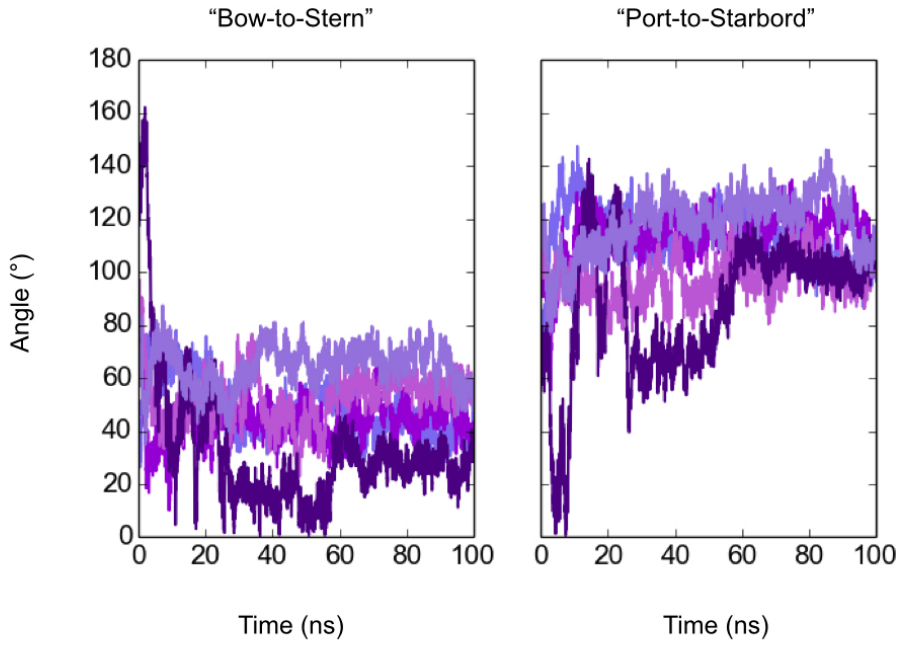


Figure 6.2: FIX-GLA orientation relative to the membrane normal. Left panel shows the "bow-to-stern" angle over 100 ns HMMM simulation, and Right panel shows "port-to-starboard" angles for the 5 properly bound trajectories.

Analysis of lipid-protein contacts for FIX-GLA also shows a differential binding pattern when compared to bFX-GLA. While bFX-GLA demonstrated numerous contacts with residues 20-32, FIX-GLA demonstrates few protein-lipid contacts for residues greater than 20 (Figure 6.3). The primary protein-lipid interactions for FIX-GLA are between LYS5, ARG11, ARG16 and lipid phosphate groups (Figure 6.3). FIX-GLA Ca^{2+} ions primarily interact with lipid carboxy groups.

The differential pattern of lipid head-group interactions between FIX-GLA and FX-GLA is most likely due to the more acute "bow-to-starboard" orientation of FIX-GLA. A possible explanation for the differences in binding

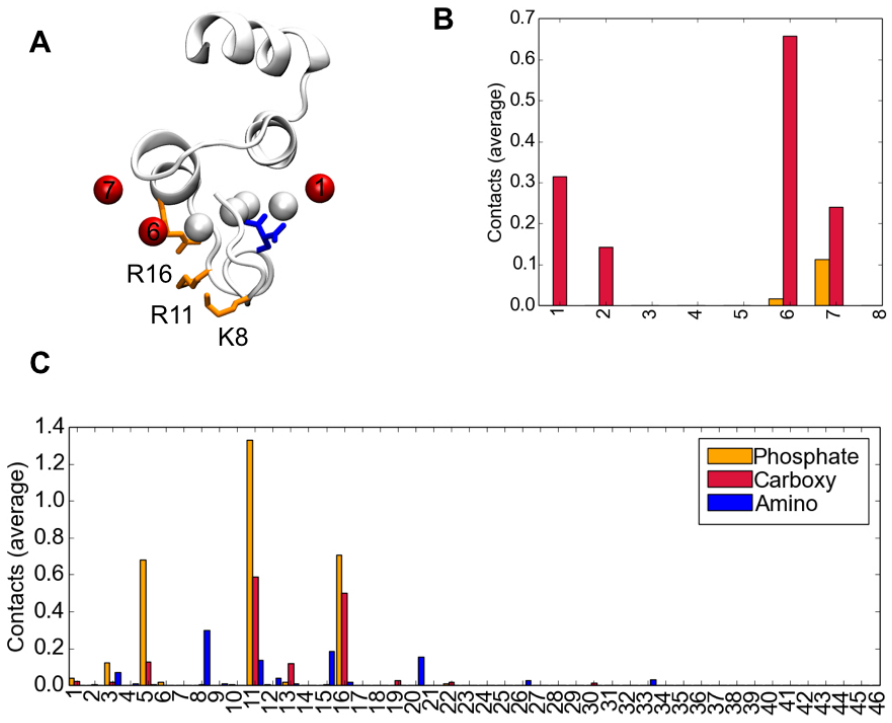


Figure 6.3: Specific contacts between lipids and FIX-GLA. A) FIX-GLA domain with most frequent contacts to anionic lipid groups are highlighted, with red indicating most frequent contacts to PS carboxy groups, orange for most frequent contacts to the phosphate groups, and blue most frequent contacts to PS amino groups. B-C) Histograms of average contacts to lipid charged groups. Average contact values are calculated for the last 10 ns of the HMMM simulations. B) FIX-GLA/Ca^{2+} contacts to lipid charged groups. Orange indicates contact to the phosphate group of either PS or PC lipid. C) Histogram of protein residue contacts to lipid charged groups. Orange indicates contacts to phosphate group of either PS or PC lipid, red indicates contact with PS carboxy group, and blue indicates contact with PS amino group.

orientation may lie in sequence differences between FX and FIX-GLA. While the two GLA domains are largely sequentially conserved, the major residues of FIX-GLA which contribute protein-lipid interactions are LYS5, ARG11, and ARG16. The equivalent residues to ARG11 and ARG16 on bFX-GLA also involved in interactions with phosphate groups (GLN10 and ARG15 respectively). LYS5, on the other hand, is a part of the keel, and in FX-GLA is an entirely hydrophobic residue. Analysis of contacts between FIX-GLA and HMMM lipid headgroups in our simulations shows that the positively charged keel LYS5 residue of FIX frequently interacts with lipid phosphate groups (Figure 6.3). Examples of interactions between LYS5 and lipid moieties in the FIX-GLA simulations can be found in Figure 6.4. Simultaneously,

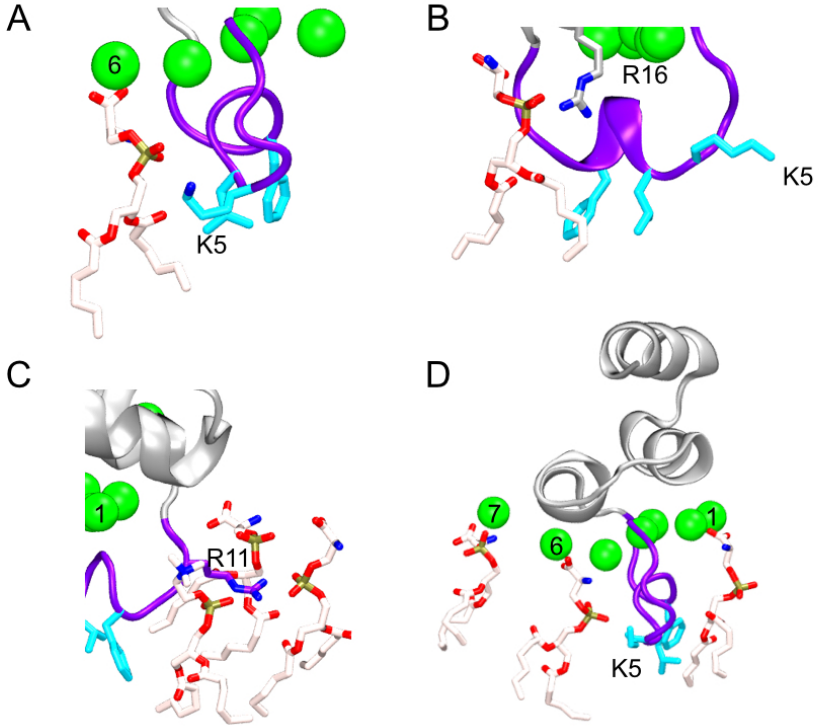


Figure 6.4: Characteristic FIX-GLA interactions with lipids for residues A) LYS5, B) ARG11, C) ARG16, and D) Ca^{2+} ions. Ca^{2+} ions are shown in green, oxygen atoms in red, nitrogen in blue, phosphorus in gold. Lipid carbon atoms are shown in beige. ω-loop of the GLA domain is drawn in dark purple, and keel residues in light blue. Remainder of FIX-GLA shown in silver. Each representative interaction taken from the last 10 ns of the 100 ns HMMM simulation.

while in FVII and FX the residue R28 is positively charged and able to interact with anionic phospholipid head-group moieties, the equivalent residue on FIX-GLA is PHE29, a hydrophobic residue with not change. This residue is located on the GLA domain alphahelix that is most "stern" in position, according to our naming convention. Thus, while FIX-GLA contains an additional positively charged group on the ω-loop compared to FX-GLA, promoting comparatively shallow binding, it also contains one fewer residues on the port end of the protein capable of interacting with lipid headgroups. This could explain the shallower binding observed as well as the more acute "bow-to-stern" orientation.

The differences found between binding orientation of FIX-GLA and FIX-GLA may explain the differential affinity of the FIX and FX proteases to interactions with TF:FVIIa. It has been shown experimentally that both FIX and FX can bind and be activated by TF:FVIIa, but that FX is a much better substrate than FIX [264]. Our previous work as suggested that residues of the FX-GLA aromatic stack (see Chapter 5) are involved in TF:FVIIa:FX binding. FIX-GLA has a very similar sequence to FX-GLA in the aromatic stack, so it is not clear a priori what the source may be of differential binding affinity between FIX and FX for TF:FVIIa.

Our work hear suggest that while the sequence is similar, position of the FIX-GLA aromatic stack will be different than that of FX-GLA following membrane binding. The aromatic stack, located on the GLA domain stern, will be displace upward in z-position relative to the membrane surface. Thus, while containing the same residues, FIX may be less likely to form interactions with the TF potential exosite than FX due to its membrane binding orientation.

6.4 Modeling the TF:FVIIa:XK1 complex

In order to develop atomic-level, membrane-bound model of TF:FVIIa:XK1, we constructed a TF:FVIIa:XK1 model on the surface of an HMMM bilayer and simulated for 500 ns. The membrane was then converted from HMMM to a conventional, full-tail representation, and the complex simulated for an additional 195 ns. At the start of the simulation, the XK1 kunitz domain was bound to FVII-SPD, modeled on binding of porcine trypsin to KD2. The

center of mass between KD1 and FVII-SPD remained constant throughout the HMMM and conventional membrane simulations (Figure 6.5).

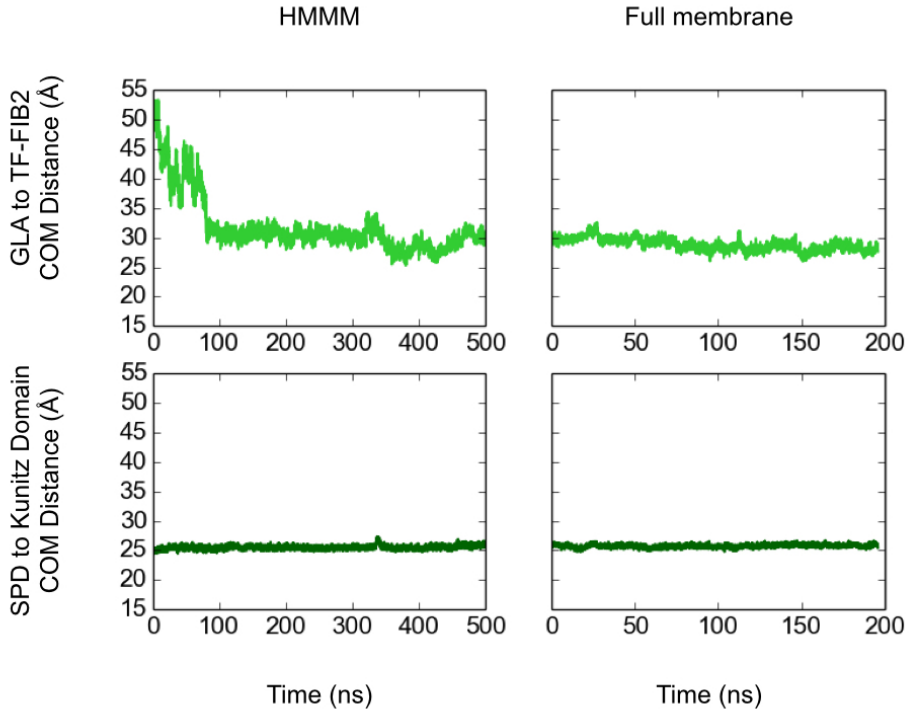


Figure 6.5: Change in center of mass (COM) distance between relevant domains of TF:FVIIa:XK1 complex over the course of HMMM and conventional membrane simulations. COM distance between membrane associating domains XK1-GLA and TF-FIB2 is shown in the upper panels in light green. Distance between the KD1 and FVII-SPD domains are shown in the lower panels in dark green. Data for 500 ns HMMM simulation shown in the left-hand panels, Data for 195 ns conventional membrane simulations on the right.

The XK1 GLA domain was initially placed so that there were no contacts between TF and XK1-GLA. Within 20 ns the XK1-GLA and TF initiated contact, which persisted throughout the rest of the 500 ns HMMM simulation. Center of mass distance between XK1-GLA and TF-FIB2, initially 55Å, decreases to approximately 30 Å within the first 100 ns (Figure 6.5). The GLA-TF-FIB2 center of mass distance remains 35 Å or less for the remainder of the HMMM 500 ns simulation, as well as the 195 ns conventional membrane simulation. Figure 6.6 illustrates the TF:FVIIa:XK1 complex at several points along the HMMM simulation, as well as the final complex at

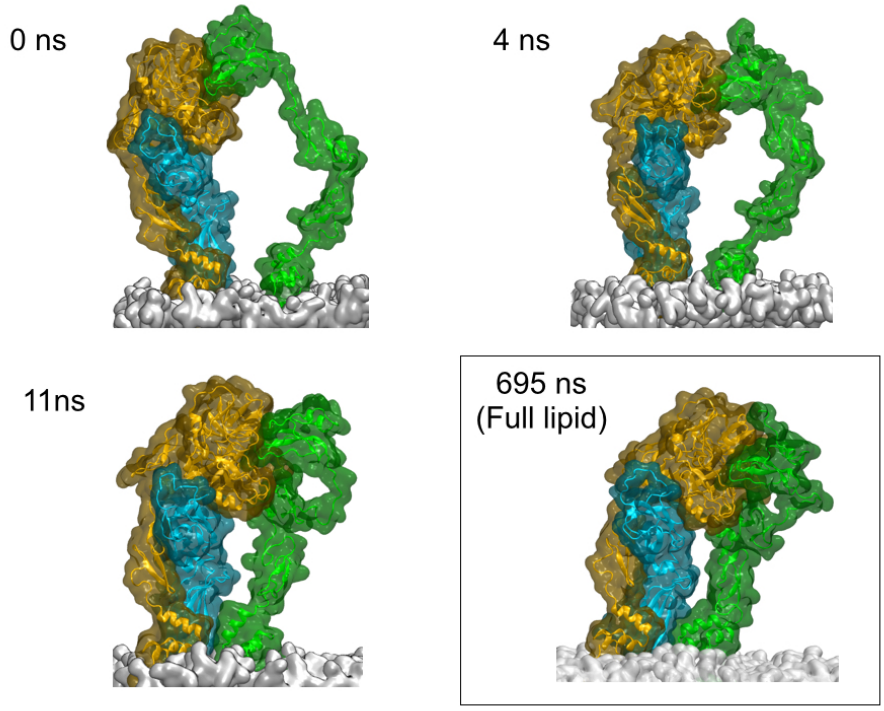


Figure 6.6: Illustration of spontaneous binding of XK1 lightchain to TF:FVIIa. Proteins shown using a cartoon backbone model overlaid with an illustration of the protein surface. XK1 shown in green, TF shown in light blue, FVIIa shown in orange, and membrane lipids shown in silver. XK1 lightchain and TF are in contact within 11 ns or initiation of the equilibrium HMMM simulation. The association persists over 500 ns of HMMM simulation and 195 ns of conventional membrane simulation.

the end of the conventional simulation.

We found that the contact site for XK1-GLA and TF-FIB2 was the previously described potential exosite of TF [239]. Average contacts for XK1-GLA within 3 Å of TF show over one contact on average for each of LYS166 and LYS169 (Figure 6.7B). XK1 residues with frequent contacts to TF are GLA32, ASP33, LYS36, and GLA39. We found that XK1interacts with the positively charged LYS166 of this site through GLA32 and ASP33 (Figure 6.7A). GLA32 simultaneously interacts with a PS lipid and LYS166. This is very similar the interaction between FX-GLA and TF in our model of the extrinsic complex (see Chapter 5). The interaction between XK1-GLA and TF-FIB2 is illustrated in Figure 6.7A. However, the hydrophobic packing interactions found in our TF:FVIIa:FX model are not seen here. The

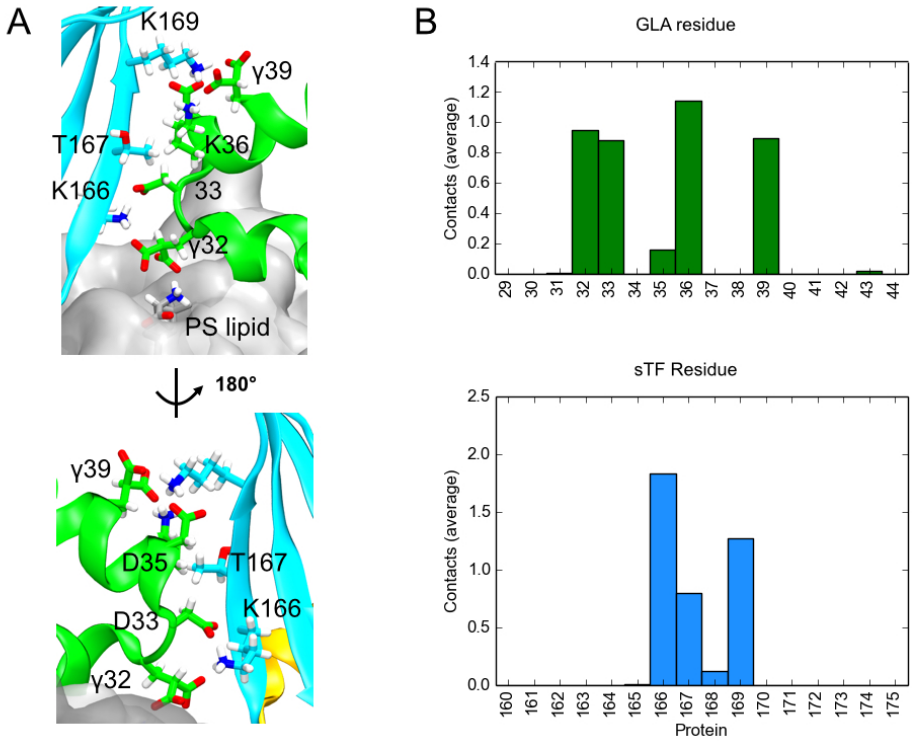


Figure 6.7: Residues involved in FX-GLA and KD1 interaction for TF:FVIIa:XK1. A) Representative interaction between FX-GLA and KD1 from last 20 ns of the conventional membrane simulation. B) Average contact number within 3 Å between XK1-GLA residues and TF-FIB2 residues over entire 195 ns of the conventional membrane simulation. Upper panel (Green) shows average number of TF residues within 3 Å of each GLA domain residue. Lower panel (blue) shows average number of XK1-GLA residues within 3 Å of listed TF residues.

TF:FVIIa:XK1 complex presented here shows stable interactions of both membrane binding segments and enzymatic inhibitory segments over hundreds of nanoseconds. We believe this model could have value as a future starting point for refinement of cryo-EM data using molecular dynamics flexible fitting, which has been used previously to refine lower-resolution experimental data [80].

7 REFERENCES

[1] G. van Meer, D. R. Voelker, and G. W. Feigenson, "Membrane lipids: where they are and how they behave," *Nat. Rev. Mol. Cell Biol.*, vol. 9, pp. 112–124, 2008.

[2] R. B. Gennis, *Biomembranes: Molecular Structure and Function.* New York: Springer-Verlag, 1989.

[3] A. A. Spector and M. A. Yorek, "Membrane lipid composition and cellular function." *J. Lipid Res.*, vol. 26, pp. 1015–1035, 1985.

[4] P. F. Devaux and A. Zachowski, "Maintenance and consequences of membrane phospholipid asymmetry," *Chem. Phys. of Lipids*, vol. 73, pp. 107 – 120, 1994.

[5] W. Dowhan, "Molecular basis for membrane phospholipid diversity: Why are there so many lipids?" *Annu. Rev. Biochem.*, vol. 66, pp. 199–232, 1997.

[6] E. M. Bevers, P. Comfurius, D. W. Dekkers, and R. F. Zwaal, "Lipid translocation across the plasma membrane of mammalian cells," *Biochim. Biophys. Acta*, vol. 1439, pp. 317 – 330, 1999.

[7] M. A. Lemmon, "Membrane recognition by phospholipid-binding domains," *Nat. Rev. Mol. Cell Biol.*, vol. 9, pp. 99–111, 2008.

[8] R. V. Stahelin, "Lipid binding domains: more than simple lipid effectors," *J. Lipid Res.*, vol. 50, pp. S299–S304, 2009.

[9] R. F. A. Zwaal, P. Comfurius, and E. M. Bevers, "Lipid-protein interactions in blood coagulation," *Biochim. Biophys. Acta*, vol. 1376, pp. 433–453, 1998.

[10] C. L. Stace and N. T. Ktistakis, "Phosphatidic acid- and phosphatidylserine-binding proteins," *Biochim. Biophys. Acta*, vol. 1761, pp. 913–926, 2006.

[11] M. P. Wymann and R. Schneiter, "Lipid signalling in disease," *Nat. Rev. Mol. Cell Biol.*, vol. 9, pp. 162–176, 2008.

[12] P. F. Neuenschwander and J. H. Morrissey, "Roles of the membrane-interactive regions of factor VIIa and tissue factor – the factor VIIa Gla domain is dispensable for binding to tissue factor but important for activation of factor X," *J. Biol. Chem.*, vol. 269, pp. 8007–8013, 1994.

[13] R. F. A. Zwaal, P. Comfurius, and E. M. Bevers, "Lipid-protein interactions in blood coagulation," *Biochim. Biophys. Acta*, vol. 1376, pp. 433–453, 1998.

[14] E. M. B. Robert, P. Comfurius, and F. A. Zwaal, "Changes in membrane phospholipid distribution during platelet activation," *Biochim. Biophys. Acta*, vol. 736, pp. 57–66, 1983.

[15] S. Krishnaswamy, K. A. Field, T. S. Edgington, J. H. Morrissey, and K. G. Mann, "Role of the membrane surface in the activation of human coagulation factor X," *J. Biol. Chem.*, vol. 267, pp. 26 110–267 120, 1992.

[16] A. W. Shaw, V. S. Pureza, S. G. Sligar, and J. H. Morrissey, "The local phospholipid environment modulates the activation of blood clotting," *J. Biol. Chem.*, vol. 282, pp. 6556–6563, 2007.

[17] J. F. McDonald, A. M. Shah, R. A. Schwalbe, W. Kisiel, B. Dahlbäck, and G. L. Nelsestuen, "Comparison of naturally occurring vitamin K-dependent proteins: Correlation of amino acid sequences and membrane binding properties suggests a membrane contact site," *Biochemistry*, vol. 36, pp. 5120–5127, 1997.

[18] J. H. Morrissey, E. Tajkhorshid, S. G. Sligar, and C. M. Rienstra, "Tissue factor/factor VIIa complex: Role of the membrane surface," *Thromb. Res.*, vol. 129, pp. S8–S10, 2012.

[19] J. H. Morrissey, E. Tajkhorshid, and C. M. Rienstra, "Nanoscale studies of protein-membrane interactions in blood clotting," *J. Thromb. Haem.*, vol. 9 (Suppl. 1), pp. 167–167, 2011.

[20] J. H. Morrissey, V. Pureza, R. L. Davis-Harrison, S. G. Sligar, Y. Z. Ohkubo, and E. Tajkhorshid, "Blood clotting reactions on nanoscale phospholipid bilayers," *Thrombosis Research*, vol. 122, pp. S23–S26, 2008.

[21] J. H. Morrissey, V. Pureza, R. L. Davis-Harrison, S. G. Sligar, C. M. Rienstra, A. Z. Kijac, Y. Z. Ohkubo, and E. Tajkhorshid, "Protein membrane interactions: blood clotting on nanoscale bilayer," *J. Thromb. Haem.*, vol. 7, pp. 169–172, 2009.

[22] N. Tavoosi, R. L. Davis-Harrison, T. V. Pogorelov, Y. Z. Ohkubo, M. J. Arcario, M. C. Clay, C. M. Rienstra, E. Tajkhorshid, and J. H. Morrissey, "Molecular determinants of phospholipid synergy in blood clotting," *J. Biol. Chem.*, vol. 286, pp. 23 247–23 253, 2011.

[23] N. Tavoosi, S. A. Smith, R. L. Davis-Harrison, and J. H. Morrissey, "Factor VII and protein C are phosphatidic acid-binding proteins," *Biochemistry*, vol. 52, pp. 5545–5552, 2013.

[24] G. L. Nelsestuen and M. Broderius, "Interaction of prothrombin and blood-clotting factor X with membranes of varying composition," *Biochemistry*, vol. 16, pp. 4172–4177, 1977.

[25] P. F. Neuenschwander, E. Bianco-Fischer, A. R. Rezaie, and J. H. Morrissey, "Phosphatidylethanolamine augments factor VIIa–tissue factor activity: Enhancement of sensitivity to phosphatidylserine," *Biochemistry*, vol. 34, pp. 13 988–13 993, 1995.

[26] P. Comfurious, E. F. Smeets, G. M. Willems, E. M. Bevers, and R. F. A. Zwaal, "Assembly of the prothrombinase complex on lipid vesicles depends on the stereochemical configuration of the polar headgroup of phosphatidylserine," *Biochemistry*, vol. 33, pp. 10 319–10 324, 1994.

[27] B. C. Lechtenberg, T. A. Murray-Rust, D. J. Johnson, T. E. Adams, S. Krishnaswamy, R. M. Camire, and J. A. Huntington, "Crytal structure of the prothrombinase complex from the venom of pseudonaja textilis," *Blood*, vol. 122, pp. 2777–2783, 2013.

[28] J. H. Morrissey, P. F. Neuenschwander, Q. Huang, C. D. McCallum, B. Su, and A. E. Johnson, "Factor VII-tissue factor: functional importance of membrane-protein interactions," *Thromb. Haemost.*, vol. 78, pp. 112–116, 1997.

[29] K. G. Mann, M. E. Nesheim, W. R. Church, P. Haley, and S. Krishnaswamy, "Surface-dependent reactions of the vitamin K-dependent enzyme complexes," *Blood*, vol. 76, pp. 1–16, 1990.

[30] K. G. Mann, R. J. Jenny, and S. Krishnaswamy, "Cofactor proteins in the assembly and expression of blood clotting engyme complexes," *Annu. Rev. Biochem.*, vol. 57, pp. 915–956, 1988.

[31] R. J. Baugh, C. D. Dickinson, W. Ruf, and S. Krishnaswamy, "Exosite interactions determine the affinity of factor X for the extrinsic Xase complex," *J. Biol. Chem.*, vol. 275, pp. 28 826–28 833, 2000.

[32] ISTH Steering Commitee for World Thrombosis Day, "Thrombosis: a major contributor to the global disease burden," *J. Thromb. Haem.*, vol. 12, pp. 1580–90, 2014.

[33] C. E. Mahan, E. T. Holdsworth, and S. M. Welsh, "A United States cost model for a preventable and costly adverse event." *Thromb. Haemost.*, vol. 106, pp. 405–415, 2011.

[34] Editorial Board, "New oral anticoagulants for acute venous thromboembolism," *J. Am. Med. Assoc.*, vol. 311, pp. 731–732, 2014.

[35] M. R. Patel, K. W. Mahaffey, J. Garg, G. Pan, D. E. Singer, W. Hacke, G. Breithardt, J. L. Halperin, G. J. Hankey, J. P. Piccini, R. C. Becker, C. C. Nessel, J. F. Paolini, S. D. Berkowitz, K. A. Fox, R. M. Califf, and the ROCKET AF Steering Committee, "Rivaroxaban versus Warfarin in novalvular atrial fibrillation," *N. Engl. J. Med.*, vol. 365, pp. 883–891, 2011.

[36] J. Wight and S. Paisley, "The epidemiology of inhibitors in haemophilia A: a systematic review," *Haemophilia*, vol. 9, pp. 418–435, 2003.

[37] J. M. Soucie, B. Evatt, and J. D., "Occurence of hemophilia in the united states. the hemophilia surveilance system project investigators." *A. J. Hematol.*, vol. 59, pp. 288–940, 1998.

[38] S. Shetty and K. Ghosh, "Novel theraputic approaches for haemophilia," *Haemophilia*, vol. 21, pp. 152–161, 2015.

[39] J. Soucie, C. Miller, A. Payne, P. Bockenstedt, C. Kempton, M. Manco-Johnson, and E. Neff, "A study of prospective surveilance for inhibitors among persons with haemophilia in the united states," *Haemophilia*, vol. 20, pp. 230–7, 2014.

[40] A. D. M. Jr. and L. Nilsson, "Molecular dynamics simulations of nucleic acid-protein complexes," *Curr. Opin. Struct. Biol.*, vol. 18, pp. 194 – 199, 2008.

[41] J. H. Morrissey, "Tissue factor: a key molecule in hemostatic and non-hemostatic systems," *Int. J. Hem.*, vol. 79, pp. 103–108, 2004.

[42] S. I. Rapaport, "Regulation of the tissue factor pathway," *ANYAS*, vol. 614, pp. 51–62, 1991.

[43] Y. Nemerson, "Tissue factor and hemostasis," *Blood*, vol. 71, pp. 1–8, 1988.

[44] B. Furie and B. C. Furie, "The molecular basis of blood coagulation," *Cell*, vol. 53, pp. 505–518, 1988.

[45] G. L. Nelsestuen, A. M. Shah, and S. B. Harvey, "Vitamin K-dependent proteins," *Vitam. Horm.*, vol. 58, pp. 355–389, 2000.

[46] C. Vermeer, "γ-carboxyglutamate-containing proteins and the vitamin K-dependent carboxylase," *Biochem. J.*, vol. 266, pp. 625–636, 1990.

[47] S. J. Freedman, B. C. Furie, B. Furie, and J. D. Baleja, "Structure of the metal-free γ-carboxyglutamic acid-rich membrane binding region of factor IX by two-dimensional NMR spectroscopy," *J. Biol. Chem.*, vol. 270, pp. 7980–7987, 1995.

[48] S. J. Freedman, B. C. Furie, B. Furie, and J. D. Baleja, "Structure of the calcium ion-bound γ-carboxyglutamic acid rich domain of factor IX," *Biochemistry*, vol. 34, pp. 12 126–12 137, 1995.

[49] M. Soriano-García, K. Padmanabhan, A. M. de Vos, and A. Tulinsky, "The Ca^{2+} ion and membrane binding structure of the gla domain of Ca-prothrombin fragment 1," *Biochemistry*, vol. 31, pp. 2554–2566, 1992.

[50] M. Sunnerhagen, S. Forsén, A.-M. Hoffrén, T. Drakenberg, O. Teleman, and J. Stenflo, "Structure of the Ca^{2+}-free GLA domain sheds light on membrane binding of blood coagulation proteins," *Nat. Struct. Biol.*, vol. 2, pp. 504–509, 1995.

[51] C. H. Park and A. Tulinsky, "Three-dimensional structure of the kringle sequence: structure of prothrombin fragment 1," *Biochemistry*, vol. 25, pp. 7980–7987, 1986.

[52] M. Huang, A. C. Rigby, X. Morelli, M. A. Grant, G. Huang, B. Furie, B. Seaton, and B. C. Furie, "Structural basis of membrane binding by Gla domains of vitamin K-dependent proteins," *Nat. Struct. Biol.*, vol. 10, pp. 751–756, 2003.

[53] Y. Z. Ohkubo and E. Tajkhorshid, "Distinct structural and adhesive roles of Ca^{2+} in membrane binding of blood coagulation factors," *Structure*, vol. 16, pp. 72–81, 2008.

[54] L. A. Falls, B. C. Furie, M. Jacobs, B. Furie, and A. C. Rigby, "The ω-loop region of the human prothrombin γ-carboxyglutamic acid domain penetrates anionic phospholipid membranes," *J. Biol. Chem.*, vol. 276, pp. 23 895–23 902, 2001.

[55] D. Paul and J. H. Morrissey, "Stoichiometric analysis reveals a unique phosphatidylserine binding site in coagulation factor X," *J. Thromb. Haem.*, vol. 20, pp. 600–604, 2022.

[56] H. Mizuno, Z. Fujimoto, H. Atoda, and T. Morita, "Crystal structure of an anticoagulant protein in complex with the Gla domain of factor X," *Proc. Natl. Acad. Sci. USA*, vol. 98, pp. 7230–7234, 2001.

[57] Y. Shikamoto, T. Morita, Z. Fujimoto, and H. Mizuno, "Crystal structure of Mg^{2+}- and Ca^{2+}-bound Gla domain of factor IX complexed with binding protein," *J. Biol. Chem.*, vol. 278, pp. 24 090–24 094, 2003.

[58] V. Oganesyan, N. Oganesyan, S. Terzyan, D. Qu, Z. Dauter, N. L. Esmon, and C. T. Esmon, "The crystal structure of the endothelial protein C receptor and a bound phospholipid," *J. Biol. Chem.*, vol. 277, pp. 24 851–24 854, 2002.

[59] D. W. Banner, A. D'Arcy, C. Chène, F. K. Winkler, A. Guha, W. H. Konigsberg, and Y. Nemerson, "The crystal structure of the complex of blood coagulation factor VIIa with soluble tissue factor," *Nature*, vol. 380, pp. 41–46, 1996.

[60] D. Venkateswarlu, R. Duke, L. Perera, T. A. Darden, and L. G. Pedersen, "An all-atom solution-equilibrated model for human extrinsic blood coagulation complex (sTF-VIIa-Xa): a protein-protein docking and molecular dynamics refinement study," *J. Thromb. Haem.*, vol. 1, pp. 2577–2588, 2003.

[61] B. V. Norledge, R. J. Petrovan, W. Ruf, and A. J. Olson, "The tissue factor/factor VIIa/factor Xa complex: A model built by docking and site-directed mutagenesis," *Proteins*, vol. 53, pp. 640–648, 2003.

[62] A. G. Lee, "Lipid-protein interactions in biological membranes: a structural perspective," *Biochim. Biophys. Acta*, vol. 1612, pp. 1–40, 2003.

[63] S. J. Opella and F. M. Marassi, "Applications of NMR to membrane proteins," *Arch. Biochem. Biophys.*, vol. 628, pp. 92–101, aug 2017.

[64] Y. Gao, E. Cao, D. Julius, and Y. Cheng, "TRPV1 structures in nanodiscs reveal mechanisms of ligand and lipid action," *Nature*, vol. 534, pp. 347–351, 2016.

[65] P. Jin, D. Bulkley, Y. Guo, W. Zhang, Z. Guo, W. Huynh, S. Wu, S. Meltzer, T. Cheng, L. Y. Jan, Y.-N. Jan, and Y. Cheng, "Electron cryo-microscopy structure of the mechanotransduction channel NOMPC," *Nature*, vol. 547, pp. 118–122, jun 2017.

[66] N. Thonghin, V. Kargas, J. Clews, and R. C. Ford, "Cryo-electron microscopy of membrane proteins," *Methods*, vol. 147, pp. 176 – 186, 2018.

[67] K. Abe and Y. Fujiyoshi, "Cryo-electron microscopy for structure analyses of membrane proteins in the lipid bilayer," *Curr. Opin. Struct. Biol.*, vol. 39, pp. 71 – 78, 2016.

[68] R. O. Dror, R. M. Dirks, J. Grossman, H. Xu, and D. E. Shaw, "Biomolecular simulation: A computational microscope for molecular biology," *Annu. Rev. Biophys.*, vol. 41, pp. 429–452, jun 2012.

[69] E. H. Lee, J. Hsin, M. Sotomayor, G. Comellas, and K. Schulten, "Discovery through the computational microscope," *Structure*, vol. 17, pp. 1295–1306, 2009.

[70] C. D. Wassman, R. Baronio, Ö. Demir, B. D. Wallentine, C. Chen, L. V. Hall, F. Salehi, D. W. Lin, B. P. Chung, G. W. Hatfield, A. R. Chamberlin, H. Luecke, R. H. Lathrop, P. Kaiser, and R. E. Amaro, "Computational identification of a transiently open l1/s3 pocket for reactivation of mutant p53," *Nat. Commun.*, vol. 4, p. 1407, 2013.

[71] C. E. Clancy, G. An, W. Cannon, Y. Liu, E. May, P. Ortoleva, A. Popel, J. Sluka, and J. Su, "Multiscale modeling in the clinic: Drug design and development," *Ann. Biomed. Eng.*, vol. 44, pp. 2591–2610, 2016.

[72] J. R. Perilla, J. A. Hadden, B. C. Goh, C. G. Mayne, and K. Schulten, "All-atom molecular dynamics of virus capsids as drug targets," *J. Phys. Chem. Lett.*, vol. 7, pp. 1836–1844, 2016.

[73] J. D. Durrant and J. A. McCammon, "Molecular dynamics simulations and drug discovery," *BMC Biology*, vol. 9, p. 71, 2011.

[74] J. V. Vermaas, J. L. Baylon, M. J. Arcario, M. P. Muller, Z. Wu, T. V. Pogorelov, and E. Tajkhorshid, "Efficient exploration of membrane-associated phenomena at atomic resolution," *J. Membr. Biol.*, vol. 248, pp. 563–582, 2015.

[75] J. L. Baylon, J. V. Vermaas, M. P. Muller, M. J. Arcario, T. V. Pogorelov, and E. Tajkhorshid, "Atomic-level description of protein-lipid interactions using an accelerated membrane model," *Biochim. Biophys. Acta, Biomembr.*, vol. 1858, pp. 1573–1583, 2016.

[76] J. V. Vermaas, N. Trebesch, C. Mayne, S. Thangapandian, M. Shekhar, P. Mahinthichaichan, J. L. Baylon, T. Jiang, Y. Wang, M. P. Muller, E. Shinn, Z. Zhao, P. Wen, and E. Tajkhorshid, "Microscopic characterization of membrane transporter function by *in Silico* modeling and simulation," in *Computational Approaches for Studying Enzyme Mechanism Part B*, G. A. Voth, Ed. Academic Press, 2016, vol. 578, ch. 16, pp. 373–428.

[77] M. Chavent, A. L. Duncan, and M. S. Sansom, "Molecular dynamics simulations of membrane proteins and their interactions: From nanoscale to mesoscale," *Curr. Opin. Struct. Biol.*, vol. 40, pp. 8–16, 2016.

[78] H. Grubmüller, B. Heymann, and P. Tavan, "Ligand binding and molecular mechanics calculation of the streptavidin-biotin rupture force," *Science*, vol. 271, pp. 997–999, 1996.

[79] B. Isralewitz, M. Gao, and K. Schulten, "Steered molecular dynamics and mechanical functions of proteins," *Curr. Opin. Struct. Biol.*, vol. 11, pp. 224–230, 2001.

[80] L. G. Trabuco, E. Schreiner, J. Gumbart, J. Hsin, E. Villa, and K. Schulten, "Applications of the molecular dynamics flexible fitting method," *J. Struct. Biol.*, vol. 173, pp. 420–427, 2011.

[81] L. G. Trabuco, E. Villa, E. Schreiner, C. B. Harrison, and K. Schulten, "Molecular Dynamics Flexible Fitting: A practical guide to combine cryo-electron microscopy and X-ray crystallography," *Methods*, vol. 49, pp. 174–180, 2009.

[82] R. McGreevy, A. Singharoy, Q. Li, J. Zhang, D. Xu, E. Perozo, and K. Schulten, "xMDFF: Molecular dynamics flexible fitting of low-resolution X-Ray structures," *Acta Cryst. D*, vol. 70, pp. 2344–2355, 2014.

[83] B. J. Alder and T. E. Wainwright, "Studies in molecular dynamics. i. general method," *J. Chem. Phys.*, vol. 31, pp. 459–466, 1959.

[84] M. P. Muller, T. Jiang, C. Sun, M. Lihan, S. Pant, P. Mahinthichaichan, A. Trifan, and E. Tajkhorshid, "Characterization of lipid-protein interactions and lipid-mediated modulation of membrane protein function through molecular simulations," *Chem. Rev.*, vol. 119, pp. 6086–6161, 2019.

[85] S. J. Plimpton, "Fast parallel algorithms for short-range molecular dynamics," *J. Comp. Phys.*, vol. 117, pp. 1–19, 1995.

[86] J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kale, and K. Schulten, "Scalable molecular dynamics with NAMD," *J. Comp. Chem.*, vol. 26, pp. 1781–1802, 2005.

[87] H. J. C. Berendsen, W. F. van Gunsteren, E. Egberts, and J. de Vlieg, *Dynamic Simulation of Complex Molecular Systems*. ACS Publications, 1987, ch. 7, pp. 106–122.

[88] P. van der Ploeg and H. J. C. Berendsen, "Molecular dynamics simulation of a bilayer membrane," *J. Chem. Phys.*, vol. 76, pp. 3271–3276, Mar. 1982.

[89] P. van der Ploeg and H. J. C. Berendsen, "Molecular dynamics of a bilayer membrane," *Mol. Phys.*, vol. 49, pp. 233–248, 1983.

[90] H. Heller, M. Schaefer, and K. Schulten, "Molecular dynamics simulation of a bilayer of 200 lipids in the gel and in the liquid crystal-phases," *J. Phys. Chem.*, vol. 97, pp. 8343–8360, 1993.

[91] O. Edholm and J. Johansson, "Lipid bilayer polypeptide interactions studied by molecular dynamics simulation," *Eur. Biophys. J.*, vol. 14, pp. 203–209, 1987.

[92] T. Woolf and B. Roux, "Molecular dynamics simulation of the gramicidin channel in a phospholipid bilayer," *Proc. Natl. Acad. Sci. USA*, vol. 91, p. 11631, 1994.

[93] T. B. Woolf and B. Roux, "Structure, energetics, and dynamics of lipid-protein interactions: a molecular dynamics study of the gramicidin A channel in a DMPC bilayer," *Proteins: Struct., Func., Gen.*, vol. 24, pp. 92–114, 1996.

[94] J. W. Ponder and D. A. Case, "Force fields for protein simulations," in *Advances in Protein Chemistry*, F. M. Richards, D. S. Eisendberg, and J. Kuriyan, Eds. Elsevier Academic Press, 2003, vol. 66 Protein Simulations, pp. 27–85.

[95] A. D. MacKerell Jr., "Empirical force fields for biological macromolecules: Overview and issues," *J. Comp. Chem.*, vol. 25, pp. 1584–1604, 2004.

[96] O. Guvench and A. D. MacKerell, Jr, "Comparison of protein force fields for molecular dynamics simulations," *Methods Mol. Biol.*, vol. 443, pp. 63–88, 2008.

[97] L. Monticelli and D. P. Tieleman, "Force fields for classical molecular dynamics," in *Biomolecular Simulations: Methods and Protocols*, L. Monticelli and E. Salonen, Eds. Totowa, NJ: Humana Press, 2013, pp. 197–213.

[98] J. W. Ponder and D. A. Case, "Force fields for protein simulations," in *Protein Simulations*, ser. Advances in Protein Chemistry. Academic Press, 2003, vol. 66, no. Supplement C, pp. 27 – 85.

[99] B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, and M. Karplus, "CHARMM: A program for macromolecular energy, minimization, and dynamics calculations," *J. Comp. Chem.*, vol. 4, pp. 187–217, 1983.

[100] A. D. MacKerell, Jr., D. Bashford, M. Bellott, J. R. L. Dunbrack, J. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, B. Roux, M. Schlenkrich, J. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiorkiewicz-Kuczera, D. Yin, and M. Karplus, "Self-consistent parameterization of biomolecules for molecular modeling and condensed phase simulations," *FASEB J.*, vol. 6, pp. A143–A143, 1992.

[101] A. D. MacKerell, Jr., D. Bashford, M. Bellott, R. L. Dunbrack, Jr., J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, I. W. E. Reiher, B. Roux, M. Schlenkrich, J. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiorkiewicz-Kuczera, D. Yin, and M. Karplus, "All-atom empirical potential for molecular modeling and dynamics studies of proteins." *J. Phys. Chem. B*, vol. 102, pp. 3586–3616, 1998.

[102] H.-J. Weinmann, R. C. Brasch, W.-R. Press, and G. E. Wesbey, "Characteristics of Gd-DTPA complex: A potential NMR contrast agent," *Am. J. Roentgenology*, vol. 142, pp. 619–624, 1984.

[103] W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, Jr., D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, and P. A. Kollman, "A second generation force field for the simulation of proteins, nucleic acids, and organic molecules," *J. Am. Chem. Soc.*, vol. 117, pp. 5179–5197, 1995.

[104] W. F. van Gunsteren and H. J. C. Berendsen, *GROMOS Manual*, BIOMOS b. v., Lab. of Phys. Chem., Univ. of Groningen, 1987.

[105] W. Jorgensen and J. Tirado-Rives, "The OPLS potential functions for protein energy minimization for crystals of cyclic peptides and crambin," *J. Am. Chem. Soc.*, vol. 110, p. 3469, 1988.

[106] G. A. Kamisnki, R. A. Friesner, J. Tirado-Rives, and W. L. Jorgensen, "Evaluation and reparameterization of the OPLS-AA force field for proteins via comparison with accurate quantum chemical calculations on peptides," *J. Phys. Chem. B*, vol. 105, pp. 6476–6487, 2001.

[107] M. Schlenkrich, J. Brickmann, A. D. MacKerell, Jr., and M. Karplus, "Empirical potential energy function for phospholipids: Criteria for parameter optimization and applications," in *Biological Membranes: A Molecular Perspective from Computation and Experiment*, K. M. Merz and B. Roux, Eds. Boston: Birkhauser, 1996, pp. 31–81.

[108] R. W. Pastor and A. D. MacKerell Jr., "Development of the CHARMM force field for lipids," *J. Phys. Chem. Lett.*, vol. 2, pp. 1526–1532, 2011.

[109] S. E. Feller and A. D. MacKerell, Jr., "An improved empirical potential energy function for molecular simulations of phospholipids," *J. Phys. Chem. B*, vol. 104, pp. 7510–7515, 2000.

[110] J. B. Klauda, B. R. Brooks, A. D. MacKerell, Jr., R. M. Venable, and R. W. Pastor, "An ab initio study on the torsional surface of alkanes and its effect on molecular simulations of alkanes and a DPPC bilayer," *J. Phys. Chem. B*, vol. 109, pp. 5300–5311, 2005.

[111] 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., and R. W. Pastor, "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.

[112] R. M. Venable, A. J. Sodt, B. Rogaski, H. Rui, E. Hatcher, A. D. MacKerell, Jr., R. W. Pastor, and J. B. Klauda, "CHARMM all-atom additive force field for sphingomyelin: Elucidation of hydrogen bonding and of positive curvature," *Biophys. J.*, vol. 107, pp. 134–145, 2014.

[113] W. F. D. Bennett, C. K. Hong, Y. Wang, and D. P. Tieleman, "Antimicrobial peptide simulations and the influence of force field on the free energy for pore formation in lipid bilayers," *J. Chem. Theory Comput.*, vol. 12, pp. 4524–4533, 2016.

[114] Å. A. Skjevik, B. D. Made, R. C. Walker, and K. Teigen, "LIPID11: A modular framework for lipid simulations using amber," *J. Phys. Chem. B*, vol. 116, pp. 11 124–11 136, 2012.

[115] J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, and D. A. Case, "Developing and testing of a general amber force field," *J. Comp. Chem.*, vol. 25, pp. 1157–1174, 2004.

[116] C. J. Dickson, B. D. Madej, Å. A. Skjevik, R. M. Betz, K. Teigen, I. R. Gould, and R. C. Walker, "Lipid14: The amber lipid force field," *J. Chem. Theory Comput.*, vol. 10, pp. 865–879, feb 2014.

[117] B. D. Madej, I. R. Gould, and R. C. Walker, "A parameterization of cholesterol for mixed lipid bilayer simulation within the amber lipid14 force field," *J. Phys. Chem. B*, vol. 119, pp. 12 424–12 435, 2015.

[118] W. Kulig, M. Pasenkiewicz-Gierula, and T. Róg, "Topologies, structures and parameter files for lipid simulations in GROMACS with the OPLS-aa force field: DPPC, POPC, DOPC, PEPC, and cholesterol," *Dat. Brief*, vol. 5, pp. 333–336, 2015.

[119] A. Maciejewski, M. Pasenkiewicz-Gierula, O. Cramariuc, I. Vattulainen, and T. Ròg, "Refined OPLS all-atom force field for saturated phosphatidylcholine bilayers at full hydration," *J. Phys. Chem. B*, vol. 118, pp. 4571–4581, 2014.

[120] T. A. Halgren and W. Damm, "Polarizable force fields," *Curr. Opin. Struct. Biol.*, vol. 11, pp. 236–242, 2001.

[121] A. Warshel and M. Levitt, "Theoretical studies of enzymatic reactions: Dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme." *J. Mol. Biol.*, vol. 103, pp. 227–249, 1976.

[122] F. H. Stillinger and C. W. David, "Polarization model for water and its ionic dissociation products," *J. Chem. Phys.*, vol. 69, pp. 1473–1484, 1978.

[123] J. W. Ponder, C. Wu, P. Ren, V. S. Pande, J. D. Chodera, M. J. Schnieders, I. Haque, D. L. Mobley, D. S. Lambrecht, R. A. DiStasio, M. Head-Gordon, G. N. I. Clark, M. E. Johnson, and T. Head-Gordon, "Current status of the AMOEBA polarizable force field," *J. Phys. Chem. B*, vol. 114, pp. 2549–2564, 2010.

[124] Y. Shi, Z. Xia, J. Zhang, R. Best, C. Wu, J. W. Ponder, and P. Ren, "Polarizable atomic multipole-based AMOEBA force field for proteins," *J. Chem. Theory Comput.*, vol. 9, pp. 4046–4063, 2013.

[125] P. E. M. Lopes, J. Huang, J. Shim, Y. Luo, H. Li, B. Roux, and A. D. MacKerell, Jr., "Polarizable force field for peptides and proteins based on the classical Drude oscillator," *J. Chem. Theory Comput.*, vol. 9, pp. 5430–5449, 2013.

[126] J. Chowdhary, E. Harder, P. E. M. Lopes, L. Huang, A. D. MacKerell, Jr., and B. Roux, "A polarizable force field of dipalmitoylphosphatidylcholine based on the classical Drude model for molecular dynamics simulations of lipids," *J. Phys. Chem. B*, vol. 117, pp. 9142–9160, 2013.

[127] H. Li, J. Chowdhary, L. Huang, X. He, A. D. MacKerell Jr., and B. Roux, "Drude polarizable force field for molecular dynamics simulations of saturated and unsaturated zwitterionic lipids," *J. Chem. Theory Comput.*, vol. 13, pp. 4535–4552, 2017.

[128] J. Huang, P. E. Lopes, B. Roux, and A. D. MacKerell Jr, "Recent advances in polarizable force fields for macromolecules: microsecond simulations of proteins using the classical drude oscillator model," *J. Phys. Chem. Lett.*, vol. 5, pp. 3144–3150, 2014.

[129] I. Vorobyov, L. Li, and T. W. Allen, "Assessing atomistic and coarse-grained force fields for protein-lipid interactions: the formidable challenge of an ionizable side chain in a membrane," *J. Phys. Chem. B*, vol. 112, pp. 9588–9602, 2008.

[130] A. D. MacKerell, Jr., M. Feig, and C. L. Brooks, III, "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. Comp. Chem.*, vol. 25, pp. 1400–1415, 2004.

[131] R. B. Best, X. Zhu, J. Shim, P. E. M. Lopes, J. Mittal, M. Feig, and A. D. MacKerell, "Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone ϕ, ψ and side-chain χ_1 and χ_2 dihedral angles," *J. Chem. Theory Comput.*, vol. 8, pp. 3257–3273, 2012.

[132] K. Lindorff-Larsen, S. Piana, K. Palmo, P. Maragakis, J. L. Klepeis, R. O. Dror, and D. E. Shaw, "Improved side-chain torsion potentials for amber ff99SB protein force field," *Proteins*, vol. 78, pp. 1950–1958, 2010.

[133] D.-W. Li and R. Brüschweiler, "NMR-based protein potentials," *Angew. Chem. Int. Ed. Engl.*, vol. 49, pp. 6778–6780, 2010.

[134] C. Oostenbrink, A. Villa, A. E. Mark, and W. F. Van Gunsteren, "A biomolecular force field based on the free enthalpy of hydration and solvation: The GROMOS force-field parameter sets 53A5 and 53A6," *J. Comp. Chem.*, vol. 25, pp. 1656–1676, 2004.

[135] A. D. MacKerell Jr. and N. K. Banavali, "All-atom empirical force field for nucleic acids: II. Application to molecular dynamics simulations of DNA and RNA in solution," *J. Comp. Chem.*, vol. 21, pp. 105–120, 2000.

[136] E. J. Denning, U. D. Priyakumar, L. Nilsson, and A. D. Mackerell, Jr., "Impact of 2'-hydroxyl smapling on the conformational properties of RNA: Update of the CHARMM all-atom additive force field for RNA," *J. Comp. Chem.*, vol. 32, pp. 1929–1943, 2011.

[137] K. Hart, N. Foloppe, C. M. Baker, E. J. Denning, L. Nilsson, and A. D. Mackerell, Jr., "Optimization of the CHARMM additive force field for DNA: Improved treatment of the BI/BII conformational equilibrium," *J. Chem. Theory Comput.*, vol. 8, pp. 348–362, 2012.

[138] A. Pérez, I. Marchán, D. Svozil, J. Sponer, T. E. Cheatham, C. A. Laughton, and M. Orozco, "Refinement of the AMBER Force Field for Nucleic Acids: Improving the Description of α/γ Conformers," *Biophys. J.*, vol. 92, pp. 3817–3829, 2007.

[139] R. Aduri, B. T. Psciuk, P. Saro, H. Taniga, H. B. Schlegel, and J. SantaLucia, "AMBER force field parameters for the naturally occurring modified nucleosides in RNA," *J. Chem. Theory Comput.*, vol. 3, pp. 1464–1475, 2007.

[140] T. A. Soares, P. H. Hünenberger, M. A. Kastenholz, V. Kräutler, T. Lenz, R. D. Lins, C. Oostenbrink, and W. F. Van Gunsteren, "An improved nucleic acid parameter set for the GROMOS force field," *J. Comp. Chem.*, vol. 26, pp. 725–737, 2005.

[141] O. Guvench, E. Hatcher, R. M. Venable, R. W. Pastor, and A. D. Mackerell, Jr., "CHARMM additive all-atom force field for glycosidic linkages between hexopyranoses," *J. Chem. Theory Comput.*, vol. 5, pp. 2353–2370, 2009.

[142] E. P. Raman, O. Guvench, and A. D. Mackerell, Jr., "CHARMM additive all-atom force field for glycosidic linkages in carbohydrates involving furanoses," *J. Phys. Chem. B*, vol. 40, pp. 12 981–12 994, 2010.

[143] O. Guvench, S. S. Mallajosyula, E. P. Raman, E. Hatcher, K. Vanommeslaeghe, T. J. Foster, F. W. Jamison, and A. D. MacKerell, "CHARMM additive all-atom force field for carbohydrate derivatives and its utility in polysaccharide and carbohydrate protein modeling," *J. Chem. Theory Comput.*, vol. 7, pp. 3162–3180, 2011.

[144] T. M. Glennon and K. M. Merz, Jr., "A carbohydrate force field for AMBER and its application to the study of saccharide to surface adsorption," *J. Mol. Struc.*, vol. 395–396, pp. 157–171, 1997.

[145] K. N. Kirschner, A. B. Yongye, S. M. Tschampel, and J. González-Outeiriño, "GLYCAM06: A generalizable biomolecular force field. carbohydrates," *J. Comp. Chem.*, vol. 29, pp. 622–655, 2008.

[146] W. Damm, A. Frontera, J. Tirado-Rives, and W. L. Jorgensen, "OPLS all-atom force field for carbohydrates," *J. Comp. Chem.*, vol. 18, pp. 1955–1970, 1997.

[147] D. Kony, W. Damm, S. Stoll, and W. F. Van Gunsteren, "An improved OPLS force field for carbohydrates," *J. Comp. Chem.*, vol. 23, pp. 1416–1429, 2002.

[148] S. E. Feller, D. Yin, R. W. Pastor, and A. D. MacKerell Jr., "Molecular dynamics simulation of unsaturated lipids at low hydration: Parametrization and comparison with diffraction studies," *Biophys. J.*, vol. 73, pp. 2269–2279, 1997.

[149] K. Vanommeslaeghe and A. D. MacKerell, Jr., "Automation of the CHARMM General Force Field (CGenFF) I: Bond perception and atom typing," *J. Chem. Inf. Model.*, vol. 52, pp. 3144–3154, 2012.

[150] K. Vanommeslaeghe, E. P. Raman, and A. D. MacKerell, Jr., "Automation of the CHARMM General Force Field (CGenFF) II: Assignment of bonded parameters and partial atomic charges," *J. Chem. Inf. Model.*, vol. 52, pp. 3155–3168, 2012.

[151] J. D. Yesselman, D. J. Price, J. L. Knight, and C. L. Brooks, III, "MATCH: An atom–typing toolset for molecular mechanics force fields," *J. Comp. Chem.*, vol. 33, pp. 189–202, 2012.

[152] C. G. Mayne, J. Saam, K. Schulten, E. Tajkhorshid, and J. C. Gumbart, "Rapid parameterization of small molecules using the Force Field Toolkit," *J. Comp. Chem.*, vol. 34, pp. 2757–2770, 2013.

[153] L. Verlet, "Computer 'experiments' on classical fluids: I. Thermodynamical properties of Lennard-Jones molecules," *Physical Review*, vol. 159, pp. 98–103, 1967.

[154] B. Hess, "P-LINCS: A parallel linear constraint solver for molecular simulation," *J. Chem. Theory Comput.*, vol. 4, pp. 116–122, 2008.

[155] W. van Gunsteren and H. Berendsen, "Algorithms for macromolecular dynamics and constraint dynamics," *Mol. Phys.*, vol. 34, pp. 1311–1327, nov 1977.

[156] J.-P. Ryckaert, G. Ciccotti, and H. J. C. Berendsen, "Numerical integration of the Cartesian equations of motion of a system with constraints: Molecular dynamics of *n*-alkanes," *J. Comp. Phys.*, vol. 23, pp. 327–341, 1977.

[157] H. C. Andersen, "Rattle: A "velocity" version of the shake algorithm for molecular dynamics calculations," *J. Chem. Phys.*, vol. 52, pp. 24–34, 1983.

[158] S. Miyamoto and P. A. Kollman, "SETTLE: An analytical version of the SHAKE and RATTLE algorithm for rigid water molecules," *J. Comp. Chem.*, vol. 13, pp. 952–962, 1992.

[159] B. Hess, H. Bekker, H. J. C. Berendsen, and J. G. E. M. Fraaije, "LINCS: a linear constraint solver for molecular simulations," *J. Comp. Chem.*, vol. 18, pp. 1463–1472, sep 1997.

[160] M. Paterlini and D. M. Ferguson, "Constant temperature simulations using the langevin equation with velocity verlet integration," *Chem. Phys.*, vol. 236, pp. 243–252, sep 1998.

[161] N. Goga, A. J. Rzepiela, A. H. de Vries, S. J. Marrink, and H. J. C. Berendsen, "Efficient algorithms for langevin and DPD dynamics," *J. Chem. Theory Comput.*, vol. 8, pp. 3637–3649, oct 2012.

[162] J. V. Vermaas, T. V. Pogorelov, and E. Tajkhorshid, "Extension of the highly mobile membrane mimetic to transmembrane systems through customized *in Silico* solvents," *J. Phys. Chem. B*, vol. 121, pp. 3764–3776, 2017.

[163] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola, and J. R. Haak, "Molecular dynamics with coupling to an external bath," *J. Chem. Phys.*, vol. 81, pp. 3684–3690, 1984.

[164] G. J. Martyna, D. J. Tobias, and M. L. Klein, "Constant pressure molecular dynamics algorithms," *J. Chem. Phys.*, vol. 101, pp. 4177–4189, 1994.

[165] S. E. Feller, Y. Zhang, R. W. Pastor, and B. R. Brooks, "Constant pressure molecular dynamics simulation: The Langevin piston method," *J. Chem. Phys.*, vol. 103, pp. 4613–4621, 1995.

[166] C. L. Brooks III, "Methodological advances in molecular dynamics simulations of biological systems," *Curr. Opin. Struct. Biol.*, vol. 5, pp. 211–215, 1995.

[167] T. Darden, D. York, and L. G. Pedersen, "Particle mesh Ewald: An $N\cdot\log(N)$ method for Ewald sums in large systems," *J. Chem. Phys.*, vol. 98, pp. 10 089–10 092, 1993.

[168] U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, and L. G. Pedersen, "A smooth particle mesh Ewald method." *J. Chem. Phys.*, vol. 103, pp. 8577–8593, 1995.

[169] Y. Z. Ohkubo, T. V. Pogorelov, M. J. Arcario, G. A. Christensen, and E. Tajkhorshid, "Accelerating membrane insertion of peripheral proteins with a novel membrane mimetic model," *Biophys. J.*, vol. 102, pp. 2130–2139, 2012.

[170] J. L. Baylon, I. L. Lenov, S. G. Sligar, and E. Tajkhorshid, "Characterizing the membrane-bound state of cytochrome P450 3A4: Structure, depth of insertion, and orientation," *J. Am. Chem. Soc.*, vol. 135, pp. 8542–8551, 2013.

[171] J. L. Baylon and E. Tajkhorshid, "Capturing spontaneous membrane insertion of the influenza virus hemagglutinin fusion peptide," *J. Phys. Chem. B*, vol. 119, pp. 7882–7893, 2015.

[172] J. V. Vermaas and E. Tajkhorshid, "A microscopic view of phospholipid insertion into biological membranes," *J. Phys. Chem. B*, vol. 118, pp. 1754–1764, 2014.

[173] T. V. Pogorelov, J. V. Vermaas, M. J. Arcario, and E. Tajkhorshid, "Partitioning of amino acids into a model membrane: Capturing the interface," *J. Phys. Chem. B*, vol. 118, pp. 1481–1492, 2014.

[174] M. J. Arcario and E. Tajkhorshid, "Membrane-induced structural rearrangement and identification of a novel membrane anchor in Talin F2F3," *Biophys. J.*, vol. 107, pp. 2059–2069, 2014.

[175] A. E. Blanchard, M. J. Arcario, K. Schulten, and E. Tajkhorshid, "A highly tilted membrane configuration for the pre-fusion state of synaptobrevin," *Biophys. J.*, vol. 107, pp. 2112–2121, 2014.

[176] J. V. Vermaas and E. Tajkhorshid, "Conformational heterogeneity of α-synuclein in membrane," *Biochim. Biophys. Acta, Biomembr.*, vol. 1838, pp. 3107–3117, 2014.

[177] J.-F. Rhéault, E. Gagné, M. Guertin, G. Lamoureux, M. Auger, and P. Lagüe, "Molecular model of hemoglobin N from *Mycobacterium tuberculosis* bound to lipid bilayers: A combined spectroscopic and computational study," *Biochemistry*, vol. 54, pp. 2073–2084, 2015.

[178] J. J. Madsen, Y. Z. Ohkubo, G. H. Peters, J. H. Faber, E. Tajkhorshid, and O. H. Olsen, "Membrane interaction of the factor VIIIa discoidin domains in atomistic detail," *Biochemistry*, vol. 54, pp. 6123–6131, 2015.

[179] M. P. Muller, Y. Wang, J. H. Morrissey, and E. Tajkhorshid, "Lipid specificity of the membrane binding domain of coagulation factor X," *J. Thromb. Haem.*, vol. 15, pp. 2005–2016, 2017.

[180] Y. Qi, X. Cheng, J. Lee, J. V. Vermaas, T. V. Pogorelov, E. Tajkhorshid, S. Park, J. B. Klauda, and W. Im, "CHARMM-GUI HMMM builder for membrane simulations with the highly mobile membrane-mimetic model," *Biophys. J.*, vol. 109, pp. 2012–2022, 2015.

[181] M. C. Gregory, M. A. McLean, and S. G. Sligar, "Interaction of KRas4b with anionic membranes: A special role for PIP_2," *Biochem. Biophys. Res. Commun.*, vol. 487, pp. 351–355, 2017.

[182] J. V. Vermaas and E. Tajkhorshid, "Differential membrane binding mechanics of synaptotagmin isoforms observed at atomic detail," *Biochemistry*, vol. 56, pp. 281–293, 2017.

[183] D. R. McDougle, J. L. Baylon, D. D. Meling, A. Kambalyal, Y. V. Grinkova, J. Hammernik, E. Tajkhorshid, and A. Das, "Incorporation of charged residues in the CYP2J2 FG loop disrupts CYP2J2–lipid bilayer interactions," *Biochim. Biophys. Acta, Biomembr.*, vol. 1848, pp. 2460–2470, 2015.

[184] L. Zhang, M. Rajendram, D. B. Weibel, A. Yethiraj, and Q. Cui, "Ionic hydrogen bonds and lipid packing defects determine the binding orientation and insertion depth of reca on multicomponent lipid bilayers," *J. Phys. Chem. B*, vol. 120, pp. 8424–8437, 2016.

[185] K. K. Skeby, O. J. Andersen, T. V. Pogorelov, E. Tajkhorshid, and B. Schiott, "Conformational dynamics of the human islet amyloid polypeptide in a membrane environment: Toward the aggregation prone form," *Biochemistry*, vol. 55, pp. 2031–2042, 2016.

[186] M. Christensen, K. K. Skeby, and B. Schiøtt, "Identification of key interactions in the initial self-assembly of amylin in a membrane environment," *Biochemistry*, vol. 56, pp. 4884–4894, 2017.

[187] B. Roux, "The calculation of the potential of mean force using computer simulations," *Comput. Phys. Commun.*, vol. 91, pp. 275–282, 1995.

[188] S. Park and K. Schulten, "Calculating potentials of mean force from steered molecular dynamics simulations," *J. Chem. Phys.*, vol. 120, pp. 5946–5961, 2004.

[189] M. Moradi and E. Tajkhorshid, "Mechanistic picture for conformational transition of a membrane transporter at atomic resolution." *Proc. Natl. Acad. Sci. USA*, vol. 110, pp. 18 916–18 921, 2013.

[190] M. Moradi and E. Tajkhorshid, "Computational recipe for efficient description of large-scale conformational changes in biomolecular systems," *J. Chem. Theory Comput.*, vol. 10, pp. 2866–2880, 2014.

[191] M. Moradi, G. Enkavi, and E. Tajkhorshid, "Atomic-level characterization of transport cycle thermodynamics in the glycerol-3-phosphate:phosphate transporter," *Nat. Commun.*, vol. 6, p. 8393, 2015.

[192] E. Darve, D. Rodríguez-Gómez, and A. Pohorille, "Adaptive biasing force method for scalar and vector free energy calculations," *J. Chem. Phys.*, vol. 128, p. 144120, 2008.

[193] J. Schlitter, M. Engels, and P. Krüger, "Targeted molecular dynamics: A new approach for searching pathways of conformational transitions," *J. Mol. Graphics*, vol. 12, pp. 84–89, 1994.

[194] G. Fiorin, M. L. Klein, and J. Hénin, "Using collective variables to drive molecular dynamics simulations," *Mol. Phys.*, vol. 111, pp. 3345–3362, 2013.

[195] H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, and P. E. Bourne, "The protein data bank," *Nucleic Acids Res.*, vol. 28, pp. 235–242, 2000.

[196] H. Berman, K. Henrick, and H. Nakamura, "Announcing the worldwide Protein Data Bank," *Nat. Struct. Biol.*, vol. 10, p. 980, 2003.

[197] K. Kelly, "3D bioinformatics and comparative protein modeling in MOE," *Chem Comput Group*, pp. 1–10, 2013.

[198] C. R. Søndergaard, M. H. Olsson, M. Rostkowski, and J. H. Jensen, "Improved treatment of ligands and coupling effects in empirical calculation and rationalization of pKa values," *J. Chem. Theory Comput.*, vol. 7, pp. 2284–2295, 2011.

[199] M. H. Olsson, C. R. Søndergaard, M. Rostkowski, and J. H. Jensen, "Propka3: consistent treatment of internal and surface residues in empirical pKa predictions," *J. Chem. Theory Comput.*, vol. 7, pp. 525–537, 2011.

[200] I. Halperin, B. Ma, H. Wolfson, and R. Nussinov, "Principles of docking: An overview of search algorithms and a guide to scoring functions," *Proteins: Struct., Func., Bioinf.*, vol. 47, pp. 409–443, 2002.

[201] I. T. Desta, K. A. Porter, B. Xia, D. Kozakov, and S. Vajda, "Performance and its limits in rigid body protein-protein docking," *Structure*, vol. 28, pp. 1071–1081.e3, 2020.

[202] A. M. Bonvin, "Flexible protein–protein docking," *Curr. Opin. Struct. Biol.*, vol. 16, pp. 194–200, 2006.

[203] E. Katchalski-Katzir, I. Shariv, M. Eisenstein, A. A. Friesem, C. Aflalo, and I. A. Vakser, "Molecular surface recognition: determination of geometric fit between proteins and their ligands by correlation techniques," *Proc. Natl. Acad. Sci. USA*, vol. 89, pp. 2195–2199, 1992.

[204] V. A. Roberts, E. E. Thompson, M. E. Pique, M. S. Perez, and L. F. Ten Eyck, "DOT2: Macromolecular docking with improved biophysical models," *J. Comp. Chem.*, vol. 34, pp. 1743–1758, 2013.

[205] N. A. Baker, D. Sept, S. Joseph, M. J. Holst, and J. A. McCammon, "Electrostatics of nanosystems: Application to microtubules and the ribosome," *Proc. Natl. Acad. Sci. USA*, vol. 98, pp. 10 037–10 041, 2001.

[206] M. F. Sanner, A. J. Olson, and J.-C. Spehner, "Fast and robust computation of molecular surfaces," in *Proceedings of 11th ACM Symposium on Computational Geometry.* New York, NY: ACM press, 1995, pp. 331–340.

[207] E. W. Davie, K. Fujikawa, and W. Kisiel, "The coagulation cascade: Initiation, maintenance, and regulation," *Biochemistry*, vol. 30, pp. 10 363–10 370, 1991.

[208] S. H. Qureshi, L. Yang, C. Manithody, and A. R. Rezaie, "Membrane-dependent interaction of factor Xa and prothrombin with factor V in the prothrombinase complex," *Biochemistry*, vol. 48, pp. 5034–5041, 2009.

[209] D. Venkateswarlu, L. Perera, T. Darden, and L. G. Pedersen, "Structure and dynamics of zymogen human blood coagulation factor X," *Biophys. J.*, vol. 82, pp. 1190–1206, 2002.

[210] N. Tavoosi and J. H. Morrissey, "Influence of membrane composition on the enhancement of factor VII/tissue factor activity by magnesium ions," *Thromb. Haemost.*, vol. 111, pp. 565–779, 2014.

[211] K. Vadival, S. Agah, A. S. Messer, D. Cascio, M. S. Bajaj, S. Krishnaswamy, C. T. Esmon, K. Padmanabhan, and S. P. Bajaj, "Structural and functional studies of γ-carboxyglutamic acid domains of factor VIIa and activated protein C: role of magnesium at physiological calcium," *J. Mol. Biol.*, vol. 425, pp. 1961–1981, 2013.

[212] S. X. Wang, E. Hur, C. A. Sousa, L. Brinen, E. J. Slivka, and R. J. Fletterick, "The extended interactions and Gla domain of blood coagulation factor Xa," *Biochemistry*, vol. 42, pp. 7959–7966, 2003.

[213] W. Humphrey, A. Dalke, and K. Schulten, "VMD – Visual Molecular Dynamics," *J. Mol. Graphics*, vol. 14, pp. 33–38, 1996.

[214] M. Hoffman and D. M. Monroe III, "A cell-based model of hemostasis," *Thromb. Haemost.*, vol. 85, pp. 958–965, 2001.

[215] D. M. Monroe and N. S. Key, "The tissue factor–factor VIIa complex: procoagulant activity, regulation, and multitasking," *J. Thromb. Haem.*, vol. 5, pp. 1097–1105, 2007.

[216] K. Hansson and J. Stenflo, "Post-translational modifications in proteins involved in blood coagulation," *J. Thromb. Haem.*, vol. 3, pp. 2633–2648, 2005.

[217] B. Dahlbäck and B. O. Villoutreix, "Regulation of blood coagulation by the protein C anticoagulant pathway," *Arter. Thromb. Vasc. Biol.*, vol. 25, pp. 1311–1320, 2005.

[218] B. Dahlbäck, "Blood coagulation," *Lancet*, vol. 355, pp. 1627–1632, 2000.

[219] G. L. Nelsestuen, W. Kisiel, and R. G. Di Scipio, "Interaction of vitamin K dependent proteins with membranes," *Biochemistry*, vol. 17, pp. 2134–2138, 1978.

[220] Y. Rodríguez, M. Mezei, and R. Osman, "The PT1-Ca^{2+} Gla domain binds to a membrane through two dipalmitoylphosphatidylserines. a computational study," *Biochemistry*, vol. 47, pp. 13 267–13 278, 2008.

[221] G. E. Gilbert and A. A. Arena, "Activation of the factor viiia-factor ixa enzyme complex of blood coagulation by membranes containing phosphatidyl-l-serine (*)," *Journal of Biological Chemistry*, vol. 271, pp. 11 120–11 125, 1996.

[222] J. E. Vance and R. Steenbergen, "Metabolism and functions of phosphatidylserine," *Progress in Lipid Research*, vol. 44, pp. 207–234, 2005.

[223] M. A. Grant, R. F. Baikeev, G. E. Gilbert, and A. C. Rigby, "Lysine 5 and phenylalanine 9 of the factor IX ω-loop interact with phosphatidylserine in a membrane-mimetic environment," *Biochemistry*, vol. 43, pp. 15 367–15 378, 2004.

[224] S. Jo, T. Kim, V. G. Iyer, and W. Im, "CHARMM-GUI: A web-based graphical user interface for CHARMM," *J. Comp. Chem.*, vol. 29, pp. 1859–1865, 2008.

[225] S. Jo, J. B. Lim, J. B. Klauda, and W. Im, "CHARMM-GUI membrane builder for mixed bilayers and its application to yeast membranes," *Biophys. J.*, vol. 97, pp. 50–58, 2009.

[226] J. C. Phillips, D. J. Hardy, J. D. C. Maia, J. E. Stone, J. V. Ribeiro, R. C. Bernardi, R. Buch, G. Fiorin, J. Hénin, W. Jiang, R. McGreevy, M. C. R. Melo, B. Radak, R. D. Skeel, A. Singharoy, Y. Wang, B. Roux, A. Aksimentiev, Z. Luthey-Schulten, L. V. Kalé, K. Schulten, C. Chipot, and E. Tajkhorshid, "Scalable molecular dynamics on CPU and GPU architectures with NAMD," *J. Chem. Phys.*, vol. 153, p. 044130, 2020.

[227] E. E. Kooijman, K. M. Carter, E. G. Van Laar, V. Chupin, K. N. Burger, and B. De Kruijff, "What makes the bioactive lipids phosphatidic acid and lysophosphatidic acid so special?" *Biochemistry*, vol. 44, pp. 17 007–17 015, 2005.

[228] J. J. Shin and C. J. Loewen, "Putting the pH into phosphatidic acid signaling," *BMC biology*, vol. 9, pp. 1–11, 2011.

[229] Y. Nemerson, "The reaction between bovine brain tissue factor and factors VII and X," *Biochemistry*, vol. 5, pp. 601–608, 1966.

[230] M. Kalafatis, N. A. Swords, M. D. Rand, and K. G. Mann, "Membrane-dependent reactions in blood coagulation: role of the vitamin K-dependent enzyme complexes," *Biochim. Biophys. Acta, Mol. Basis of Dis.*, vol. 1227, pp. 113 – 129, 1994.

[231] N. Mackman, R. E. Tilley, and N. S. Key, "Role of the extrinsic pathway of blood coagulation in hemostasis and thrombosis," *Arter. Thromb. Vasc. Biol.*, vol. 27, pp. 1687–1693, 2007.

[232] K. Kamata, H. Kawamoto, T. Honma, T. Iwama, and S.-H. Kim, "Structural basis for chemical inhibition of human blood coagulation factor Xa," *Proc. Natl. Acad. Sci. USA*, vol. 95, pp. 6630–6635, 1998.

[233] C. D. Dickinson, C. R. Kelly, and W. Ruf, "Identification of surface residues mediating tissue factor binding and catalytic function of the serine protease factor VIIa," *Proc. Natl. Acad. Sci. USA*, vol. 93, pp. 14 379–14 384, 1996.

[234] W. Ruf, "Factor VIIa residue Arg 290 is required for efficient activation of the macromolecular substrate factor X," *Biochemistry*, vol. 33, pp. 11 631–11 636, 1994.

[235] D. Kirchhofer, M. T. Lipari, P. Moran, C. Eigenbrot, and R. F. Kelley, "The tissue factor region that interacts with substrates factor IX and factor X," *Biochemistry*, vol. 39, pp. 7380–7387, 2000.

[236] Q. Huang, P. F. Neuenschwander, A. R. Rezaie, and J. H. Morrissey, "Substrate recognition by tissue factor-factor VIIa – evidence for interaction of residues Lys^{165} and Lys^{166} of tissue factor with the 4-carboxyglutamate-rich domain of factor X," *J. Biol. Chem.*, vol. 271, pp. 21 752–21 757, 1996.

[237] S. Roy, P. E. Hass, J. H. Bourell, W. J. Henzel, and G. A. Vehar, "Lysine residues 165 and 166 are essential for the cofactor function of tissue factor," *J. Biol. Chem.*, vol. 266, pp. 22 063–22 066, 1991.

[238] K. Ke, J. Yuan, and J. H. Morrissey, "Tissue factor residues that putatively interact with membrane phospholipids," *PLoS One*, vol. 9, p. e88675, 2014.

[239] M. Huang, R. Syed, E. A. Stura, M. J. Stone, R. S. Stefanko, W. Ruf, T. S. Edgington, and I. A. Wilson, "The mechanism of an inhibitory antibody on TF-initiated blood coagulation revealed by the crystal structures of human tissue factor, Fab 5G9 and TF·5G9 complex," *J. Mol. Biol.*, vol. 275, pp. 873 – 894, 1998.

[240] W. Ruf, J. Shobe, S. M. Rao, C. D. Dickinson, A. Olsen, and T. S. Edginton, "Importance of factor VIIa gla-domain residue Arg-36 for recognition of the macromolecular substrate factor X gla-domain," *Biochemistry*, vol. 38, pp. 1957–1966, 1999.

[241] A. R. Rezaie, P. F. Neuenschwander, J. H. Morrisseyy, and C. T. Esmon, "Analysis of the functions of the first epidermal growth factor-like domain of factor X," *J. Biol. Chem.*, vol. 268, pp. 8176–8180, 1993.

[242] M. Ndonwi, G. J. Broze, Jr., S. Agah, A. E. Schmidt, and S. P. Bajaj, "Substitution of the gla domain in factor X with that of protein C impairs its interaction with factor VIIa/tissue factor — lack of comparable effect by similar substitution in factor IX," *J. Biol. Chem.*, vol. 282, pp. 15 632–15 644, 2007.

[243] Y. Z. Ohkubo, J. H. Morrissey, and E. Tajkhorshid, "Dynamical view of membrane binding and complex formation of human factor VIIa and tissue factor," *J. Thromb. Haem.*, vol. 8, pp. 1044–1053, 2010.

[244] Z. Rao, P. Handford, M. Mayhew, V. Knott, G. G. Brownlee, and D. StuartZ, "The structure of a Ca^{2+}-binding epidermal growth factor-like domain: Its role in protein-protein interactions," *Cell*, vol. 82, pp. 131 – 141, 1995.

[245] D. Zhong, M. S. Bajaj, A. E. Schmidt, and S. P. Daja, "The n-terminal epidermal growth factor-like domain in factor IX and factor X represents and important recognition motif for binding to tissue factor," *J. Biol. Chem.*, vol. 277, pp. 3622–3631, 2002.

[246] J. Du, K. Wichapong, T. M. Hackeng, and G. A. F. Nicolaes, "Molecular simulation studies of human coagulation factor VIII C domain-mediated membrane binding," *Thromb. Haemost.*, vol. 113, pp. 373–384, 2015.

[247] T. Jiang, P.-C. Wen, N. Trebesch, Z. Zhao, S. Pant, K. Kapoor, M. Shekhar, and E. Tajkhorshid, "Computational dissection of membrane transport at a microscopic level," *Trends Biochem. Sci.*, vol. 45, pp. 202–216, 2020.

[248] C. Mayne, W. Toy, K. Carlson, T. Bhatt, S. W. Fanning, G. L. Greene, B. S. Katzenellenbogen, S. Chandarlapaty, J. A. Katzenellenbogen, and E. Tajkhorshid, "Defining the energetic basis for a conformational switch mediating ligand-independent activation of mutant estrogen receptors in breast cancer," *Mol. Cancer Res.*, vol. 19, pp. 1559–1570, 2021.

[249] T. Jiang, W. Han, M. Maduke, and E. Tajkhorshid, "Molecular basis for differential anion binding and proton coupling in the Cl^-/H^+ exchanger ClC-ec1," *J. Am. Chem. Soc.*, vol. 138, pp. 3066–3075, 2016.

[250] M. H. Cheng and I. Bahar, "Molecular mechanism of dopamine transport by human dopamine transporter," *Structure*, vol. 23, pp. 2171–2181, nov 2015.

[251] S. Pant and E. Tajkhorshid, "Microscopic characterization of GRP1 PH domain interaction with anionic membranes," *J. Comp. Chem.*, vol. 41, pp. 489–499, 2020.

[252] D. Gorgun, M. Lihan, K. Kapoor, and E. Tajkhorshid, "Binding mode of SARS-CoV2 fusion peptide to human cellular membrane," *Biophys. J.*, vol. 120, p. 191a, 2021.

[253] M. Selander-Sunnerhagen, M. Ullner, E. Persson, O. Teleman, J. Stenflo, and T. Drakenberg, "How an epidermal growth factor (EGF)-like domain binds calcium. high resolution NMR structure of the calcium form of the NH2-terminal EGF-like domain in coagulation factor X," *J. Biol. Chem.*, vol. 267, pp. 19 642–19 649, 1992.

[254] J. Åqvist, "Ion-water interaction potentials derived from free energy perturbation simulations," *J. Phys. Chem.*, vol. 94, pp. 8021–8024, 1990.

[255] R. J. Baugh and S. Krishnaswamy, "Role of the activation peptide domain in human factor X activation by the extrinsic Xase complex," *J. Biol. Chem.*, vol. 271, pp. 16 126–16 134, 1996.

[256] C. Eigenbrot, D. Kirchhofer, M. S. Dennis, L. Santell, R. A. Lazarus, J. Stamos, and M. H. Ultsch, "The factor VII zymogen structure reveals reregistration of β strands during activation," *Structure*, vol. 9, pp. 627–636, 2001.

[257] E. Papagrigoriou, P. A. McEwan, P. N. Walsh, and J. Emsley, "Crystal structure of the factor XI zymogen reveals a pathway for transactivation," *Nat. Struct. Mol. Biol.*, vol. 13, p. 557, 2006.

[258] M. Budayova-Spano, M. Lacroix, N. M. Thielens, G. J. Arlaud, J. C. Fontecilla-Camps, and C. Gaboriaud, "The crystal structure of the zymogen catalytic domain of complement protease C1r reveals that a disruptive mechanical stress is required to trigger activation of the C1 complex," *EMBO J.*, vol. 21, pp. 231–239, 2002.

[259] M. Megyeri, V. Harmat, B. Major, A. Végh, J. Balczer, D. Héja, K. Szilágyi, D. Datz, G. Pál, and P. Závodszky, "Quantitative characterization of the activation steps of mannan-binding lectin (MBL)-associated serine proteases MASPs points to the central role of MASP-1 in the initiation of the complement lectin pathwa," *J. Biol. Chem.*, vol. 288, pp. 8922–8934, 2013.

[260] L. Kalé, R. Skeel, M. Bhandarkar, R. Brunner, A. Gursoy, N. Krawetz, J. Phillips, A. Shinozaki, K. Varadarajan, and K. Schulten, "NAMD2: Greater scalability for parallel molecular dynamics," *J. Comp. Phys.*, vol. 151, pp. 283–312, 1999.

[261] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, "Comparison of simple potential functions for simulating liquid water," *J. Chem. Phys.*, vol. 79, pp. 926–935, 1983.

[262] T. Zama, M. Murata, R. Watanabe, K. Yokoyama, T. Moriki, H. Ambo, H. Murakami, M. Kikuchi, and Y. Ikeda, "A family with hereditary factor X deficiency with a point mutation gla32 to gln in the gla domain (factor X tokyo)," *Brit. J. Haem.*, vol. 106, pp. 809–811, 1999.

[263] F. Birkle and J. H. Morrissey, "A serine loop in tissue factor mediates substrate selectivity by the tissue factor–factor VIIa complex," *J. Thromb. Haem.*, vol. 19, pp. 75–84, 2021.

[264] D. Kirchhofer, C. Eigenbrot, M. T. Lipari, P. Moran, M. Peek, and R. F. Kelley, "The tissue factor region that interacts with factor Xa in the activation of factor VII," *Biochemistry*, vol. 40, pp. 675–682, 2001.

[265] H. Brandstetter, M. Bauer, R. Huber, P. Lollar, and W. Bode, "X-ray structure of clotting factor IXa: Active site and module structure related to Xase activity and hemophilia B," *Proc. Natl. Acad. Sci. USA*, vol. 92, pp. 9796–9800, 1995.

[266] M. P. Bicocchi, M. Pasino, C. Rosano, A. C. Molinari, E. D. Valle, T. Lanza, F. Bottini, and M. Acquila, "Insight into molecular changes of the FIX protein in a series of Italian patients with haemophilia B," *Haemophilia*, vol. 12, pp. 263–270, 2006.

[267] S. M. Medfisch, E. M. Muehl, J. H. Morrissey, and R. C. Bailey, "Phosphatidylethanolamine-phosphatidylserine binding synergy of seven coagulation factors revealed using nanodisc arrays on silicon photonic sensors," *Sci. Rep.*, vol. 10, pp. 1–7, 2020.

[268] M. Dockal, R. Hartmann, M. Fries, M. C. L. Thomassen, A. Heinzmann, H. Ehrlich, J. Rosing, F. Osterkamp, T. Polakowski, U. Reineke et al., "Small peptides blocking inhibition of factor Xa and tissue factor-factor VIIa by tissue factor pathway inhibitor (TFPI)," *J. Biol. Chem.*, vol. 289, pp. 1732–1741, 2014.

[269] Z.-F. Huang, T. Wun, and G. Broze Jr, "Kinetics of factor Xa inhibition by tissue factor pathway inhibitor," *J. Biol. Chem.*, vol. 268, pp. 26 950–26 955, 1993.

[270] G. J. Broze, L. A. Warren, W. F. Novotny, D. A. Higuchi, J. J. Girard, and J. P. Miletich, "The lipoprotein-associated coagulation inhibitor that inhibits the factor VII-tissue factor complex also inhibits factor Xa: insight into its possible mechanism of action," *Blood*, vol. 71, 1988.

[271] T. J. Girard, L. A. Warren, W. F. Novotny, K. M. Likert, S. G. Brown, J. P. Miletich, and G. J. Broze, "Functional significance of the kunitz-type inhibitory domains of lipoprotein-associated coagulation inhibitor," *Nature*, vol. 338, pp. 518–520, 1989.

[272] L. C. Petersen, C. A. Sprecher, D. C. Foster, H. Blumberg, T. Hamamoto, and W. Kisiel, "Inhibitory properties of a novel human kunitz-type protease inhibitor homologous to tissue factor pathway inhibitor," *Biochemistry*, vol. 35, pp. 266–272, 1996.

[273] M. J. Burgering, L. P. Orbons, A. van der Doelen, J. Mulders, H. J. Theunissen, P. D. Grootenhuis, W. Bode, R. Huber, and M. T. Stubbs, "The second kunitz domain of human tissue factor pathway inhibitor: cloning, structure determination and interaction with factor Xa," *J. Mol. Biol.*, vol. 269, pp. 395–407, 1997.