Early Unfolding Response of a Stable Protein Domain to Environmental Changes†

Lalith Perera,*,‡ Thomas A. Darden,§ Robert E. Duke,|| Divi Venkateswarlu,|| and Lee G. Pedersen*,§,||

High Performance Computing Group, Information Technology Services, CB#3455, and Department of Chemistry, CB#3290, University of North Carolina, Chapel Hill, North Carolina 27599, and Laboratory of Quantitative and Computational Biology, National Institute of Environmental Health Science, Research Triangle Park, North Carolina 27709

Received: April 13, 2004

The X-ray crystal structure of γ-carboxyglutamic acid (Gla) domains is well-established. These domains are stable to long all-atom simulations in explicit solvent. Here we extend a prior simulation on the Gla domain of factor VIIa, an essential vitamin K-dependent protein involved in the initiation of blood coagulation, to ~20 ns in order to establish a reference point. We also subject this domain to a set of rational environmental changes using molecular dynamics techniques that accommodate long-range electrostatics accurately: (a) we move the seven bound calcium ions to >17.5 Å from any Gla residue and then simulate for 25 ns, and (b) in a separate calculation, we change all of the calcium ions to sodium ions and simulate for 20 ns. For both perturbed systems, the N-terminus chelation complex is initially greatly weakened, leading to increased motion of the ω loop (residues 1−11). In technique a, most calcium ions return to the preperturbation coordinating units within the time scale of the simulation. We track and display the sequence of calcium ion rebinding. The response of this complex, nonstandard system to the perturbations as estimated by accurate all-atom dynamics gives new details on the degree of sampling in early refolding events.

Introduction

A burst of recent simulation papers[1−8] has described in silico folding experiments on small peptides. These studies have focused on deriving relatively correctly folded structures from relatively random starting positions. It is currently a common practice to simplify the force fields for computational efficiency. In most cases, solvent has been included as a continuum model.[9−12] On the other hand, we have concentrated our efforts on providing improvements to force fields[13] and applications that begin near experimental structures.[14−15] Explicit solvent along with full treatment of electrostatics has been included. There are strong indications that for the simulation times we have studied, the methodology captures the essence of correct protein folding behavior: for instance, a partially folded NMR structure for the membrane-binding domain of factor IX was refolded properly with correct placement of metal ions,[16] a homology model for a human P450[17] was constructed from a rabbit X-ray crystal structure that evolved during the simulation a bent I-helix, and a unique water complex (both later seen experimentally)[19] and structures seen transiently in time-resolved X-ray photodissociation of CO in myoglobin[20] have been verified.[21] In addition, correct DNA behavior with respect to structure[22] and ion distributions[23] about DNA are also demonstrated.

The simulation times of our prior studies have typically been in the 0.5−4ns range. This is a very short time, given measured folding time scales:[1] 10 ns for formation of interresidue contacts,[24] 200 ns for α-helices[25] folding, 1−10 μs for β-hairpins[26] formation, and 10−1000 μs for single domain proteins to fold from extended structures.[27] The major question of the current work is, given an established structure of a crucial protein domain and given a known refolding response to an environmental change, are current simple simulation methods capable of manifesting the correct response when extended substantially in time scale? If not the full experimentally observed response, is the response in the correct direction? In this regard, it is useful to quote Anfinsen's hypothesis about protein folding:[28] "This hypothesis states that the three-dimensional structure of a native protein in its normal physiological milieu (solvent, pH, ionic strength, presence of other components such as metal ions or prosthetic groups, temperature, and other) is the one in which the Gibbs free energy of the whole system is lowest; that is, that the native conformation is determined by the totality of interatomic interaction and hence by the amino acid sequence, in a given environment". The role of environment is obviously emphasized in the full hypothesis.

γ-Carboxyglutamate (Gla) domains appear in vitamin K-dependent proteins of the blood coagulation cascade (II, VII, IX, X, C, S, Z).[29] These substantially homologous (~50% identity) domains are thought to be responsible for endowing vitamin K-dependent proteins with their membrane binding capability. Recently, Gla domains have also been discovered in other biological systems: Gas6,[20] PR-1,2,[31] TMG-3-4,[32] conus snails,[33] bone proteins,[34] and tunicates.[35]

The first Gla domain structure in the presence of calcium ions was determined for prothrombin fragment 1[36] in 1992 by the Tulinsky group. Previously, this group had determined the structure of the same protein fragment in the absence of calcium ions.[37] The Gla domain, however, in the non-calcium ion case

† Part of the special issue "Tomas Baer Festschrift".

* Corresponding authors. E-mail: Lalith_Perera@unc.edu (L.P.), Lee_Pedersen@unc.edu (L.G.P.).

‡ High Performance Computing Group, Information Technology Services, University of North Carolina.

§ National Institute of Environmental Health Science.

|| Department of Chemistry, University of North Carolina.

10.1021/jp048385l CCC: $27.50
Published on Web 07/29/2004

did not diffract sufficiently to determine a structure. Several years later, a group headed by Banner at Roche determined the structure of tissue factor (TF)/factor VIIa.[38] The Gla domain of VIIa (in the presence of calcium ions) was found to be very similar to that of prothrombin. An NMR structure (3D) of the Gla domain of factor IX[39] in the presence of calcium ions and moderate levels of denaturing agents was subjected to a subsequent genetic algorithm refinement/molecular dynamics improvement in water. This refined structure[16] also proved to be similar to the Tulinsky Gla domain. However, a NMR structure of the Gla domain of factor X in the absence of calcium ions[40] and other divalent ions led to the inference that calcium ions cause folding of the ω loop that is thought to interact with negatively charged phospholipids of membrane (platelet) surfaces. This structure was later compared to the structure of the Gla-EGF domain of factor X in the presence of calcium ions.[41] Recently, the Esmon group[42] determined the structure of a truncated Gla domain interacting with the endothelial cell receptor in the presence of calcium ions. Also Morita's group has determined the structure[43] of a heterodimer of factor X and a snake protein in the presence of calcium ions with and without magnesium ions. Finally, the structure of a Gla domain of human factor X with mixed divalent ions (calcium and magnesium) has very recently been determined.[44] It has been known since 1977[45] that blood does not coagulate in the presence of magnesium ions alone. This observation has been attributed to the concomitant lack of binding of Gla domains to negatively charged phospholipids when only magnesium ions are present. The recent structure with mixed ions[44] shows that the ω loop, thought to be a key determinant in binding of Gla domains to membranes, has a disrupted structure when magnesium ion is present.

Our group has for many years been engaged in the study of the structures of vitamin K-dependent proteins, including their Gla domains. This work has led to the development of simulation molecular dynamics (PME) techniques[12,13] that make possible stable simulations for ionic macromolecular systems. We have shown that the critical aspects of the Ca-bound Gla domain are stable[46] with PME for long simulations and that the method can be used for refinement of NMR data.[16]

Several interesting questions emerge. For example, how much of the structural folding event induced by calcium ions (ref 36 vs ref 37, ref 40 vs ref 41) can we see on simulation time scales of the order of 20 ns, a time scale compatible with current computer technology? To get at this question, we have performed three PME simulations: (1) FVII(1−49) with bound calcium, (2) FVII(1−49) with calcium ions moved from the positions found in X-ray crystal structure to bulk, and (3) FVII-(1−49) calcium ions removed, sodium ions substituted, electroneutral. The first simulation provides the reference for comparison as well as serving as a test of the simulation technique (i.e. does the methodology hold up for 20 ns on a system that we know maintains its X-ray structure in solution?). The second simulation allows us to test whether, on the 20 ns time scale, the calcium ions placed far away in solution can find their way back to reasonable binding locations. Finally, the third simulation provides a test of how much unfolding occurs when no divalent metal ion is present. Implicit in the latter two simulations is the question as to how much of the Gla domain and its divalent metal ion binding sites is preformed; i.e., does the protein structure unfold significantly after the perturbation of the ion environment?

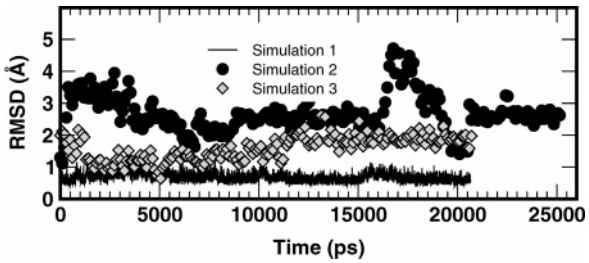

Figure 1. Root-mean-square deviations (rmsds) of backbone atoms calculated along the trajectories. The backbone atoms of first 45 residues in each configuration are used for optimal alignment with those of the X-ray crystal structure of calcium-bound, TF-bound FVIIa (the reference) and for the rmsd calculation. Solid line, solution FVII with calcium ions (simulation 1); filled circles, solution FVII with calcium ions are placed >17.5 Å from the protein for the initial structure (simulation 2); diamonds, solution FVII with sodium ions (simulation 3).

Methods

We have chosen the Gla domain (residues 1−49) from the TF/FVIIa/Ca^{2+} crystal structure[38] as our reference model (simulation 1). Water was added to create a minimum 12.5 Å shell about the protein to provide for periodic boundary conditions. Monovalent ions were added to ensure electroneutrality. The system in simulation 1 consisted of 809 peptide atoms, 7299 water molecules, seven calcium ions, and one Na-(I) counterion. For simulation 2, the calcium ions were removed from their bound positions from the initial reference model and placed randomly at least 17.5 Å from the closest protein atom. In this system, the total number of water molecules defined by an initial layer of at least 26 Å was increased to 11 401 to accommodate the solvation of ions placed significantly away from the protein. Finally, for simulation 3, the calcium ions of the initial reference system were replaced by sodium ions with sufficient sodium ions added randomly to the bulk for electroneutrality. A total of 15 sodium ions and 11 410 water molecules defined the environment in this system. Simulations 1 and 3 were performed for 20 ns, while simulation 2 was additionally extended by 5 ns to yield a total of 25 ns trajectory time.

The particle mesh Ewald (PME) option of a modified AMBER molecular dynamics program, PMEMD 3.1, was employed.[47] This version of AMBER is especially tailored for parallel computation. All standard procedures (constant pressure, shake only on hydrogen atoms, 1 fs time step) follow those given in ref 15.

Results and Discussion

In Figure 1, we present the root-mean-square deviations (rmsds) of the backbone atoms of the Gla domain of FVII. The rmsds were calculated by using the best alignment of the first 45 residues of the X-ray crystal structure[38] with snapshots selected every 10 ps of the trajectory. The last few residues (residue 46−49) have large fluctuations due to the detachment from the EGF1 domain that was necessary to prepare a separate Gla domain and, hence, are neglected in this calculation. As can be seen from the figure, the rmsds are small (∼0.7 Å) for the simulation of the Gla domain of FVII with calcium ions in which the X-ray crystal positions are used at the starting configuration (simulation 1). This simulation provides an excellent measure of the stability of the system during the 20 plus nanosecond simulation. It should also be taken as a general validation for the force field used in the simulation. On the other hand, the Gla domain systems initially prepared with the calcium ions placed away from the proteins (simulation 2) and

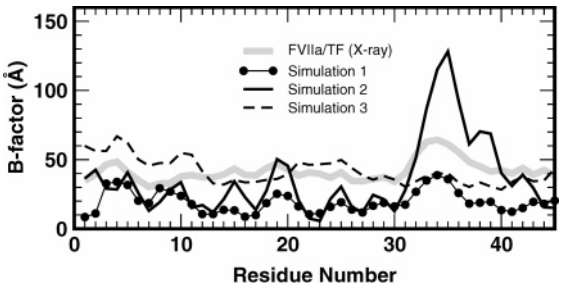

Figure 2. B-Factors ($8\pi^2\langle\Delta r^2\rangle/3$) calculated for the backbone atoms during the last 300 ps of dynamics (300 samples at a picosecond apart). Backbone atoms of residues 11−45 were used for the alignment of snapshots before calculating B-factors.

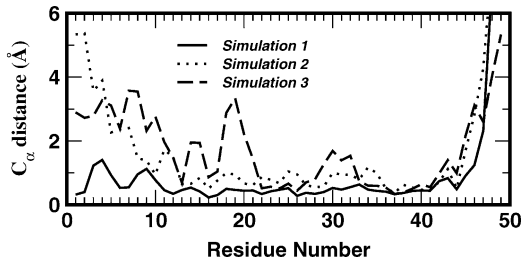

Figure 3. Cα differences [$|(r_{C\alpha}$(X-ray) − $r_{C\alpha}$(simulation))|] calculated when compared with the X-ray crystal structure. Differences were averaged over the last 300 ps of dynamics. Backbone atoms of residues 11−45 were used for the prior alignment.

with the sodium ions replacing calcium ions (simulation 3) show large fluctuations in the rmsd values compared to the wild type (simulation 1). In fact, for simulation 3 (crystallographic calcium ion positions are used to initially place half of the sodium ions), the perturbation is somewhat smaller than for simulation 3, for which calcium ions were initially placed at least 17.5 Å away from any of the protein atoms. For simulation 2, the major perturbation comes from the loss in the local balancing charge in the Gla domain. This fact is reflected in the larger values for the rmsd plot for simulation 2. The overall conclusion, however, is that even while there are significant deviations found in rmsd plots of simulations 2 and 3, the plateau in the rmsds toward the end of the simulations suggests that the peptides are in dynamic but stable conformations. The "blip" observed in the rmsd of simulation 2 at ∼17 ns will be considered in detail in a subsequent section.

The rmsds represent only the global dynamics of the peptide in that each configuration is represented only by a single number. However, thermal fluctuations of portions of the backbone atoms may also be evaluated. In Figure 2, the thermal fluctuations (or isotropic B-factors) calculated for the last 300 ps were compared with the experimental B-factors given in the X-ray crystal work.[38] The average B-factors for the backbone atoms are used for this plot. In the determination of B-factors, we used residues 11−45 for alignments of the snapshots before calculating the average fluctuations during the last 300 ps of dynamics. Since the B-factors were calculated using the coordinates at the last 300 ps of the trajectory, they provide a comparison of the average structural features at the end of the trajectories. As can be seen, both simulations 1 and 2 show quite comparable B-factors for the ω loop (residues 1−11). Large B-factors can be observed for simulation 3 compared to simulations 1 and 2. This may be due to the initially inadequate local charge provided by ions for counterbalancing the charges near the Gla residues and the mobility of some sodium ions that move in and out of coordination. The segment containing residues 30−35 falls into a turn region between two helices in the wild-type structure and displays a prominent peak in B-factors of almost all cases. For the calcium-outside simulation (simulation 2), this segment also has an interaction with the very mobile residues in the C-terminus leading to the large deviation seen in Figure 2.

Cα differences [$|(r_{C\alpha}$(X-ray) − $r_{C\alpha}$(simulation)| after optimal alignment] can be more valuable in assessing the true deviations of final structures when compared with the X-ray crystal structure. In Figure 3, the Cα differences are displayed after averaging over the final 300 ps time segment of each trajectory. Note here that the same residues (11−45) as used for previous B-factor calculation are used for the prior-to-calculation alignments. In simulation 1, these distances average ∼1 Å for the ω loop. However, displacements of the positions of the ω loop are large for both simulations 2 and 3. Interestingly, even though the N-terminus/ω loop perturbation is large for the simulation for which calcium ions were removed (simulation 2) from the X-ray positions, the final (20 ns) structures have undergone refolding due to reattachment of calcium ions to binding sites (not shown), and thus, the Cα difference plot for simulation 3 implies convergence toward the starting structure. The helical segments (residues 13−20, 25−31, and 35−44) on the other hand show the least deviation, indicating only marginal disruption to the structure in that region due to ion dynamics.

Consideration of the time dependence of the computed B-factors and Cα differences leads to a main conclusion that although there are some refolding events as a result of the initial ion placements in simulations 2 and 3, the overall secondary structure (particularly residues 11−45) remains near the original fold after 20 ns. To this point, however, we have examined only the overall structures of the Gla domain. The observations are largely consistent with the experimental (NMR) observations of the Stenflo group[40,41] for the Gla domain of FX (in different calcium ion environments, with and without). The unfolding/refolding events are tightly coupled to dynamic changes in the immediate environment. In simulation 1 (calcium ions were initially in the X-ray crystal positions), changes associated with the calcium−Gla coordinates are found to be small (see Table 1). The calcium ion coordinations were assessed during the last 100 ps of dynamics. We find some Ca−Gla exchange in the Gla−calcium network during that time window. However, we define a residue to be bound to a calcium ion if at least 75% of the time it makes contact with the ion. With this definition, only calcium ions 1, 2, 6, and 7 show conserved Gla coordination (i.e. bound) in the solution as compared to the crystal (though there are fluctuations in the coordinations), while calcium ions 3, 4, and 5 undergo changes in the Gla−calcium network by losing some of their contacts (crystal to solution). Both calcium ions 3 and 4 lose contacts with Gla-16, and calcium ion 5 loses its contact with Gla-14. Thus, even though the general features observed in the previous structure analysis do not point to local structural changes due to solvation of the protein, the Gla−calcium distribution does exhibit such local readjustments. However, it is significant that we do not see migration of any bound calcium ions to the solution during this 20 ns trajectory in simulation 1. In simulation 3, X-ray calcium ion positions were used to place seven sodium ions with extra (eight additional) sodium ions placed in the solution to provide a neutral unit cell. Interestingly for this case, we observed transitions to and from solvent by the sodium ions. In addition to the original sodium ions that participate in the Gla−sodium network, two additional sodium ions migrated into the region of the ion−Gla network. Thus, a somewhat rearranged ion−Gla network is observed for simulation 3. The most surprising observation occurs in simulation 2 (all calcium ions initially

TABLE 1: Calcium−Gla Coordination Network[a]

	X-ray crystal structure	FVII/Ca^{2+} (simulation 1)	FVII/Ca^{2+} (outside) (simulation 2)	FVII/Na^{+} (simulation 3)[b]
ion-1	Gla-7(2)	Gla-7	Gla-7(2)	**Gla-7**
	Gla-26	Gla-25	Gla-26(2)	**Gla-26(2)**
	Gla-29(2)	Gla-26	Gla-29(2)	**Gla-29(2)**
	Wat(2)	Wat(2)	Wat(2)	**Wat(2)**
ion-2	Gla-7(2)	Gla-7(2)	**Gla-6(2)**	**Gla-16**
	Gla-16	Gla-16	***Gla-7***	**Wat(3)**
	Gla-26	Gla-26	***Gla-26***	
	Gla-29	Gla-29	***Gla-29***	
	Wat(3)	Wat(3)		
ion-3	Asn-2	Asn-2	Gla-6(2)	**Ala-1**
	Gla-6	Gla-6	*Gla-7*	**Gla-16(2)**
	Gla-7	Gla-26	Gla-26	**Gla-26(2)**
	Gla-16(2)			**Wat**
	Gla-26(2)			
ion-4	Ala-1	Ala-1	Gla-6	**Gla-16**
	Gla-6(2)	Gla-6	Gla-26(2)	**Wat (4)**
	Gla-16(2)	Gla-20(2)		
	Gla-20(2)	Wat		
	Wat			
ion-5	Gla-14(2)	Gla-19(2)	Gla-19(2)	**Gla-19(2)**
	Gla-19(2)	Wat(3)	Wat(5)	**Wat(2)**
	Wat			
ion-6	Gla-20(2)	Gla-20(2)	Wat(6)	**Gla-29**
		Wat(5)		**Wat(3)**
ion-7	Gla-25	Gla-25	Gla-25	**Gla-25**
	Gla-29(2)	Gla-29(2)	Wat(6)	**Gla-26**
		Wat(3)	*Gla-29*	**Gla-29**
				Wat(5)
ion-8			Wat(6)	**Gla-7(2)**
				Gla-29(2)
				Wat(2)
ion-9				**Gla-16**
				Wat(4)

[a] Number of bonds with a particular residue is found in parentheses. "Wat" is the number of water molecules coordinated. Coordinates from the last 200 ps were used for this analysis. [b] Sodium ion coordination is represented in bold. In the calcium outside case, the sodium ion used for neutralization became strongly coordinated with Gla-6. The italic residues are currently in a weaker coordination with ions. The number within brackets denotes how many interaction sites per residue and the number of water molecules in the coordination with ions.

placed at least 17.5 Å away from the peptide). In this case, most of the calcium ions found their way back to coordinations with the Gla residues. Not only do the ions migrate long distances to initiate coordination with Gla residues, but additionally there appears to be a concerted response to restore the Gla−calcium network. The degree to which this happens over the time frame of the simulation was quite astounding. Five of the seven calcium ions in solution relocate to positions near the sites that are seen to be coordinated for the wild type calcium ion bound Gla domain (both X-ray and solution). Interestingly, a sodium ion used for the purpose of neutralizing the unit cell and originally far displaced from the peptide also establishes a coordination with several Gla residues that are located in a position that defines one of the " inside" calcium ion binding sites for the wild type.

It is rather informative for simulation 2 to display the timings of the approach of calcium ions leading to the re-formation of the Ca−Gla network. To accomplish this, we compute the *nearest* distances of *any* calcium ion to *any* carboxylate group of *any* Gla residue. In Figure 4, we present the time course of these minimum distances of calcium ions (or the sodium ion) to any carboxylate group. The nearest distances were calculated at 10 ps intervals. Initially all of the ions were at least 17.5 Å from the closest protein atom. While it took more than a

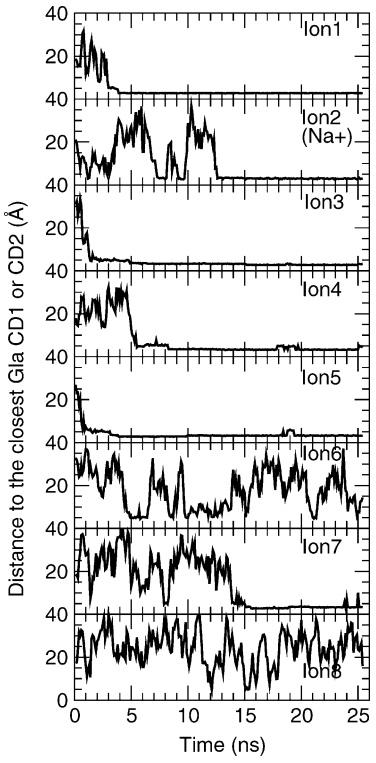

Figure 4. Distances (Å) of the ions in simulation 2 (involved in the ion−Gla network) to the *nearest* Gla CD1 or CD2 (carboxylate carbons) as a function of time. The numbering is the same as that in Table 1. The "blips" seen for ions 4, 5, 7 at > 15 ns are 5−10 Å fluctuations.

nanosecond for the first ion to migrate over 15 Å and to find a close contact with one of the carboxylate groups during the course of the next few nanoseconds, three other calcium ions found their way to reestablishing contacts with the protein. Interestingly, two calcium ions are found to make only occasional contacts with the protein but, in the 25 ns simulation, did not find a stable position for binding. After a few in and out transitions, the only sodium ion used in the simulation (for the purpose of neutralizing the system) found a stable position corresponding to a possible (prior) calcium binding site. Also, several of the calcium ions, even though having found relatively stable binding positions to the protein for a significant length of time, show transient 5−10 Å displacements from the Gla residues (see Figure 4, panels 2, 4, and 5 from the bottom). In fact, the peak seen for simulation 2 for the B-factors (in Figure 2) may be largely due to the part of the calcium−Gla network that remains mobile in simulation 2 after 20 ns. Also, the occupation of the sodium ion in a possible calcium binding site may partly be responsible for the fluctuations of the partially reestablished calcium−Gla network during the time scale of a few tens of nanoseconds.

Figure 5a show a representative configuration of FVII wild type selected near the end of simulation 1. The N-terminal Ala-1 forms an elaborate chelation complex with several calcium ions and Gla side chains. Three hydrophobic residues of the ω loop are directed away from the calcium−Gla network. This finely folded structure was also the case for the X-ray crystal structure. The backbone ribbon for this solution structure superimposes well with the X-ray crystal structure, and at least six calcium ions participate in the tightly knitted calcium−Gla network. A similar orientation from the end of simulation 3 is displayed in Figure 5b along with the superimposed backbone of the X-ray crystal structure. A clustering of sodium ions in an arrangement different from that of the calcium ion arrangement of the X-ray

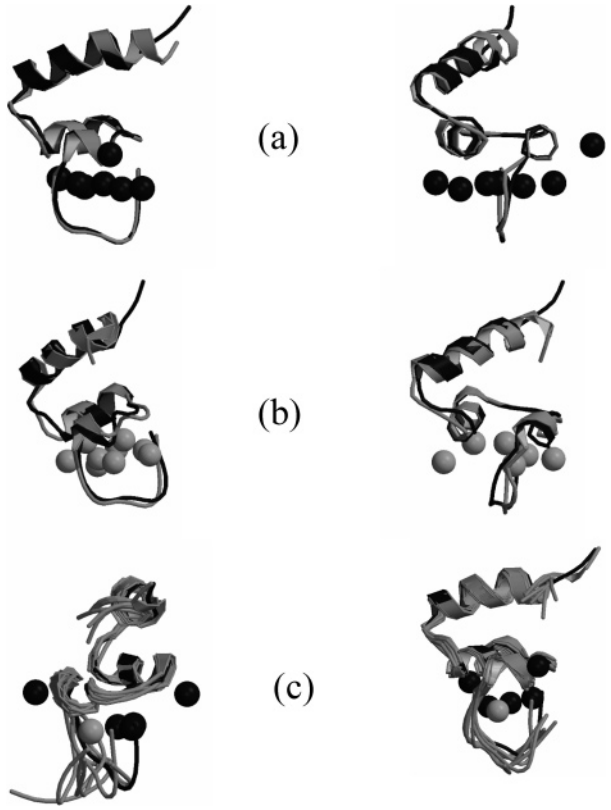

Figure 5. Representative snapshots near the end of the simulations. Backbone ribbons are displayed along with the backbone ribbons (light lines) of FVII in the X-ray crystal structure (dark line). Two views rotated by 90° are shown. Backbone atoms of residues 11−45 are optimally aligned with the X-ray crystal structure. All simulation ions involved in the Gla−ion network are also shown (calcium ions, dark spheres; sodium ions, light spheres). (a) wild-type FVII (from simulation 1 at $T = 20$ ns). (b) FVII with sodium ions (only) (from simulation 3 at 20 ns). (c) FVII with calcium ions moved away from the X-ray crystal positions at the beginning of the simulation (simulation 2). Backbone ribbons at various time intervals were displayed to emphasize the mobility of the ω loop during the dynamics.

crystal structure can be seen here; though the N-terminal Ala-1 is still found to be participating in the calcium−Gla network, the position of the ω loop is moved from the original X-ray crystal position. The most interesting structural changes are seen in representative configurations selected at 5 ns intervals from simulation 2 (see Figure 5c). The ω loop (residues 1−11) shows considerable motion during the dynamics period. What is the cause of the substantial ω loop motion? For one thing, some of the newly bound calcium ions are in fact found to be those that have returned to the interior of the peptide structure. Several of those ions are involved in the structure of the ω loop, in particular the N-terminus. That is, we find that the 25 ns time period is apparently long enough to observe repositioning of a majority of calcium ions to the original binding sites found in the wild type calcium bound FVII configuration. There are important precedents for this conjecture. The mobility of the ω loop has also been reported experimentally in the calcium-free NMR structures of factor X[41] and in the experimental magnesium ion bound structure of factor IX.[48]

A comparison of the hydrogen-bonding patterns for the stable structure (simulation 1) with the ion perturbed structures (simulations 2 and 3) gives more detailed evidence of structural changes (Table 2). We calculated the H-bonding for 100 configurations (spaced at 10 ps) during the last nanosecond of the simulations. A major set of H-bonds are associated with the N-terminal Ala-1. These H-bonds are involved in the Gla−calcium chelation complex at the N-terminus for the X-ray crystal structure. For the wild type case, three H-bonds involving Ala-1 are stable for essentially the entire nanosecond with a fourth H-bond observed at 59% of the time. This scenario is dramatically changed for both simulations 2 and 3. In both cases the above-mentioned H-bonds are observed less than 50% of the time, indicating a weaker N-terminal attachment to the main structure than is required for the ω loop positioning. However, the H-bonds beyond the ω loop are found to be mostly conserved. This observation is consistent with the structures shown in Figure 5. Also, the dynamic nature of the system is further confirmed by the less than 100% occupancy of the H bonds. As mentioned previously, the large peak in the B-factor (simulation 2) for the residue segment 31−35 can be rationalized via the H-bonding pattern. The backbone of Lys-32 is tightly bound with the backbone of Arg-28 in the X-ray crystal structure and in all three simulation structures. However, the neighboring residue Gla-29 is in a complex coordination state with several calcium ions (X-ray and simulation 1 structure). Removal (simulation 2) or change to sodium ions (simulation 3) of these calcium ions could be responsible for the mobility of Gla-29. This mobility would then be translated to the segment (residues 31−35) via strong H-bonds such as the Lys32−Arg28 bond that is conserved. In the wild-type simulation (simulation 1), there is no such drastic change associated with the Ca−Gla network, and the residue 31−35 segment is stable.

In the NMR study of Sunnerhagen et al.[41] on the calcium-free Gla domain structures of factor X, the hydrophobic residues in the ω loop (Phe-4, Leu-5, and Val-8) were found to be in the interior of the domain with the exposed Gla residues in the loop. That is, an inversion of structure occurred for the ω loop in the absence of calcium ions as compared to the X-ray crystal (or simulation 1) structure. The main conclusion of this study was that calcium binding leads to exposure of the hydrophobic residues of the ω loop so that they are available for insertion into the membrane surface. When the experimental system was compared with a calcium ion bound model system derived from the prothrombin X-ray crystal structure,[36] the rmsd for backbone atoms for residues 4−44 was reported to be 1.47 Å, whereas residues 13−18, 24−30, and 34−42 had rmsds less than 0.6 Å.[41] In the recent magnesium ion bound form of factor X (complexed with peptide M84R),[44] the X-ray crystal structure revealed a form for which the first 11 residue fragment (ω loop) is in a novel conformation. Wang et al.[44] argues that this conformation likely represents an intermediate folding state of the domain. The ω loop is found to be quite extended and does not carry the signature of an ω loop. The reported NMR structures for factor IX with magnesium ions[48] are found to have intermediate structures between the calcium-bound form (prothrombin and factor VII X-ray crystal structures) and the structure with magnesium and M84R. We cannot directly compare our simulated structures with the various experimental structures, since the conditions employed in the experimental studies are somewhat drastic compared to the simulation conditions. For example, we have attempted to replace the calcium ions by sodium ions (simulation 3), but the positions of the seven sodium ions remained at positions found in the bound-calcium ion structure. Even with the reduction in charge (calcium 2+ to sodium +1), the initial sodium ion positions are electrostatically favorable for attraction due to opposite charge interactions (Gla and sodium). However, additional sodium ions were found in the Gla−calcium networking region, and a disrupted Gla−ion network was observed. In the simula-

TABLE 2: H-Bonding Calculated during the Last Nanosecond of Dynamics

FVII/Ca^{2+} (simulation 1)	FVII/Ca^{2+} (outside) (simulation 2)	FVII/Na^{+} (simulation 3)
Ala-1:Gln-21 (N−H2−O), 97[a]	Ala-1:Gla-6 (N−H1−OE4), 45	Ala-1:Gla-20 (N−H1−OE4), 37
Ala-1:Gla-26 (N−H3−OE3), 59	Ala-1:Gla-20 (N−H2−OE3), 29	Ala-1:Gla-26 (N−H3−OE2), 41
Ala-1:Gla-26 (N−H3−OE4), 99	Ala-1:Gla-20 (N−H2−OE4), 39	Ala-1:Gla-20 (N−H3−OE4), 38
Ala-1:Gla-20 (N−H1−OE4), 97	Ala-1:Gla-20 (N−H1−OE4), 30	Ala-1:Gla-26 (N−H2−OE2), 40
Leu-8:Leu-5 (N−H−O), 34	Ala-1:Gla-6 (N−H3−OE4), 30	Gla-6:Phe-4 (N−H−O), 30
Arg-9:Gla-6 (N−H−O), 64	Leu-5:Asn-2 (N−H−O), 40	Leu-8:Leu-5 (N−H−O), 27
	Leu-5:Asn-2 (N−H−OD1), 27	
	Gla-6:Asn-2 (N−H−O), 60	
	Leu-8:Leu-5 (N−H−O), 45	
	Arg-9:Leu-5 (N−H−O), 59	
Ser-12:Gla-16 (N−H−OE4), 93	Ser-12:Gla-16 (N−H−OE3), 88	Ser-12:Pro-10 (N−H−O), 45
Arg-15:Gla-16 (NE−HE−OE2), 100	Ser-12:Gla-16 (OG−HG−OE3), 73	Arg-15:Ser-12 (N−H−OG), 30
Gla-16:Ser-12 (N−H−O), 77	Leu-13:Gla-16 (N−H−OE3), 86	Cyx-17:Leu-13 (N−H−O), 77
Cyx-17:Leu-13 (N−H−O), 71	Gla-14:Ser-12 (N−H−O), 64	Lys-18:Arg-15 (N−H−O), 73
Lys-18:Leu-13 (N−H−O), 28	Arg-15:Gla-19 (NE−HE−OE3), 39	Gla-19:Arg-15 (N−H−O), 95
Gla-19:Gla-14 (N−H−O), 85	Arg-15:Ser-12 (N−H−OG), 38	Gla-20:Gla-16 (N−H−O), 32
Cyx-22:Tyr-44 (N−H−OH), 33	Gla-16:Gla-16 (N−H−OE3), 68	Ser-23:Gla-26 (N−H−OE1), 40
Ser-23:Gla-26 (N−H−OE3), 35	Cyx-17:Leu-13 (N−H−O), 82	Ser-23:Gla-26 (OG−HG−OE1), 94
Ala-27:Ser-23 (N−H−O), 80	Lys-18:Gla-14 (NZ−HZ3−OE1), 44	Ala-27:Ser-23 (N−H−O), 53
Arg-28:Phe-24 (N−H−O), 66	Gla-19:Gla-14 (N−H−O), 58	Arg-28:Gla-25 (NE−HE−OE4), 32
Gla-29:Gla-25 (N−H−O), 62	Gla-20:Arg-15 (N−H−O), 39	Arg-28:Phe-24 (N−H−O), 91
Arg-28:Lys-32 (NE−HE−O), 67	Gln-21:Gla-20 (N−H−OE1), 81	Ile-3Ala-27 (N−H−O), 39
Ile-30:Ala-27 (N−H−O), 51	Ala-27:Ser-23 (N−H−O), 73	Phe-31:Arg-28 (N−H−O), 30
Phe-31:Ala-27 (N−H−O), 71	Arg-28:Phe-24 (N−H−O), 95	Lys-32:Arg-28 (N−H−O), 92
Lys-32:Arg-28 (N−H−O), 84	Ile-30:Ala-27 (N−H−O), 73	Asp-33:Phe-31 (N−H−O), 90
Asp-33:Phe-31 (N−H−O), 97	Phe-31:Ala-27 (N−H−O), 32	Arg-36:Asp-33 (N−H−OD2), 30
Gla-35:Gla-35 (N−H−OE4), 63	Lys-32:Arg-28 (N−H−O), 91	Thr-37:Asp-33 (N−H−O), 60
Arg-36:Asp-33 (N−H−OD2), 44	Asp-33:Phe-31 (N−H−O), 83	Thr-37:Asp-33 (OG1−HG1−O), 99
Thr-37:Asp-33 (N−H−O), 87	Arg-36:Asp-33 (N−H−OD2), 46	Lys-38:Ala-34 (N−H−O), 97
Thr-37:Asp-33 (OG1−HG1−O), 98	Thr-37:Asp-33 (OG1−HG1−O), 100	Lys-38:Gla-35 (NZ−HZ2−OE2), 37
Lys-38:Ala-34 (N−H−O), 90	Thr-37:Asp-33 (N−H−O), 83	Leu-39:Gla-35 (N−H−O), 46
Leu-39:Gla-35 (N−H−O), 81	Lys-38:Ala-34 (N−H−O), 87	Trp-41:Thr-37 (N−H−O), 92
Phe-40:Arg-36 (N−H−O), 48	Lys-38:Gla-35 (NZ−HZ1−OE1), 32	Ile-42:Lys-38 (N−H−O), 75
Trp-41:Thr-37 (N−H−O), 86	Lys-38:Gla-35 (NZ−HZ3−OE1), 28	Ser-43:Leu-39 (OG−HG−O), 46
Ile-42:Lys-38 (N−H−O), 66	Leu-39:Gla-35 (N−H−O), 59	Ser-43:Phe-40 (N−H−O), 47
Ser-43:Phe-40 (N−H−O), 37	Phe-40:Arg-36 (N−H−O), 44	Tyr-44:Phe-40 (N−H−O), 97
Tyr-44:Phe-40 (N−H−O), 68	Trp-41:Thr-37 (N−H−O), 90	Ser-45:Trp-41 (N−H−O), 95
Tyr-44:Cyx-17 (OH−HH−O), 65	Ile-42:Lys-38 (N−H−O), 60	Ser-45:Trp-41 (OG−HG−O), 90
Ser-45:Trp-41 (N−H−O), 84	Ser-43:Phe-40 (N−H−O), 41	Asp-46:Ile-42 (N−H−O), 51
Ser-45:Trp-41 (OG−HG−O), 98	Tyr-44:Phe-40 (N−H−O), 92	Gly-47:Ser-43 (N−H−O), 35
Asp-48:Ser-45 (N−H−O), 48	Tyr-44:Cyx-17 (OH−HH−O), 60	Gly-47:Tyr-44 (N−H−O), 43
Gln-49:Ser-45 (N−H−O), 51	Ser-45:Trp-41 (N−H−O), 97	Asp-48:Ser-43 (N−H−O), 74
	Ser-45:Trp-41 (OG−HG−O), 85	
	Asp-46:Ile-42 (N−H−O), 52	

[a] The percentage of time the H-bonds were observed is also shown.

tion for which calcium ions were initially placed far outside the vicinity of the protein (simulation 2), there exists the opportunity for ions to diffuse toward the protein, since the distances are still near enough for ions to feel the strength of the long-range electrostatic interactions. However, this situation is more interesting than simulation 3, since the calcium ions avidly located the coordination sites while the protein was in a local structure near that suitable for lipid binding. One might have predicted that early in simulation 2, repulsive forces would distort the protein significantly before the ions could return. This mostly did not happen. Although the limitations in the time scales used for the simulations play a role in obscuring a complete picture of dynamical events, we were able to observe some of the early changes that the protein undergoes during the unfolding and refolding due to the incorrect positioning of ions (as in simulation 2) or finding the incorrect type of ions at the sites occupying them (as in simulation 3).

Summary

The structures of the Gla domain bound with divalent metal ions of several homologous vitamin K-dependent proteins is well-established. The calcium ion bound form is critical for proper membrane binding and subsequent proteolytic events that maintain normal hemostasis. Thrombotic disease states are directly related to the membrane binding provided by this domain. In this work we have simulated the stability of the Gla domain with respect to drastic changes in the ion environment. The gross changes have been studied experimentally. We find by unconstrained molecular dynamics simulations for ∼20 ns that (a) the X-ray crystal structure is largely maintained in solution (with calcium ions bound); (b) when sodium ions replace the calcium ions, relatively small global structure changes are seen (Figure 5b) locally; particularly at the critical N-terminus, the hydrogen-bond network is greatly weakened (Table 2); (c) when all calcium ions are removed to distant solution locations in the initial structure, significant global changes occur in the N-terminus ω loop region (residues 1−11), but much of the structure is unchanged. Several of the displaced calcium ions diffuse back, in the electrostatic field of the ion-stripped protein, to locations similar to those of the X-ray crystal/simulation 1 structures (Figures 4 and 5C and Table 1). These

simulations provide us with an unusual glimpse of the relative speeds of detailed ion movements to global protein folding/unfolding.

Acknowledgment. This work was supported by the National Institute of Health (HL-06350) and NSF (ITR/AP 1-110-3265). We acknowledge the use of computational resources provided by Academic Technology and Networks at UNC—Chapel Hill and the Biomed. Unit of the Pittsburgh Supercomputing Center.

References and Notes

(1) Garcia, A. E.; Onuchic, J. N. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 13898−13903.
(2) Villa, J. A.; Ripoli, D. R.; Scheraga, H. A. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 14812−14816.
(3) Nymeyer, H.; Garcia, A. E. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 13934−13939.
(4) Ikeda, K.; Higo, J. *Protein Sci.* **2003**, *12*, 2542−2548.
(5) Chowdhury, S.; Lee, M. C.; Xiong, G.; Duan, Y. J. *Mol. Biol.* **2003**, *327*, 711−717.
(6) Zagrovic, B.; Snow, C.; Shirts, M.; Pande, V. J. *Mol. Biol.* **2002**, *323*, 927−937.
(7) Qiu, L.; Pabit, S. A.; Roitberg, A. E.; Hagen, S. J. *J. Am. Chem. Soc.* **2002**, *124*, 12952−12953.
(8) Snow, C. D.; Zagrovic, B.; Pande, V. S. **2002**, *124*, 14548−14549.
(9) Cotesta, S.; Tavernelli, I.; Di Iorio, E. E. *Biophys. J.* **2003**, *85*, 2633−2640.
(10) van der Spoel, D.; Lindahl, E. J. *Phys. Chem. B.* **2003**, *107*, 11178−11178.
(11) Daidone, I.; Amadei, A.; Roccatano, D.; Di Nola, A. *Biophys. J.* **2003**, *85*, 2865−2871.
(12) Gnanakaran, S.; Nymeyer, H.; Portman, J.; Sanbonmatsu, K. Y.; Garcia, A. E. *Curr. Opin. Struct. Biol.* **2003**, *13*, 168−174.
(13) Darden, T.; York, D.; Pedersen, L. J. *Chem. Phys.* **1993**, *98*, 10089−10092; Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. *J. Chem. Phys.* **1995**, *103*, 8577−8593.
(14) Perera, L.; Darden, T. A.; Pedersen, L. G. *Throm. Haemost.* **2001**, *85*, 596−603.
(15) Venkateswarlu, D.; Duke, R.; Perera, L.; Pedersen, L. G. J. *Throm. Haemost.* **2003**, *1*, 2577−2588.
(16) Li, L. P.; Darden, T. A.; Freedman, S. J.; Furie, B. C.; Furie, B.; Baleja, J. D.; Smith, H.; Hiskey, R. G.; Pedersen, L. G. *Biochemistry* **1997**, *36*, 2132−2138.
(17) Gorokhov, A.; Negishi, M.; Johnson, E. F.; Pedersen, L. C.; Perera, L.; Darden, T. A.; Pedersen, L. G. *Biophys. J.* **2003**, *84*, 57−68.
(18) William, P. A.; Cosme, J.; Sridhar, V.; Johnson, E. F.; McRee, D. E. *Mol. Cell* **2000**, *5*, 121−131.
(19) Wester, M. R.; Johnson, E. F.; Marques-Soares, C.; Dansette, P. M.; Mansuy, D.; Stout, C. D. *Biochemistry* **2003**, *42*, 6370−6379.
(20) Schotte, F.; Lim, M.; Jackson, T. A.; Smirnov, A. V.; Soman, J.; Olson, J. S.; Phillips, G. N.; Wulff, M.; Anfinrud, P. A. *Science* **2003**, *300*, 1944−1947.
(21) Anfinrud, P. A.; Schotte, F.; Hummer, G.; Wulff, M. Abstract of papers of the 227th National Meeting (Anaheim, CA) 2004, PHYS-478.
(22) Cheatham, T. E.; Miller, J. L.; Fox, T.; Darden, T. A.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 4193−4194; York, D. M.; Yang, W. T.; Lee, H.; Darden, T.; Pedersen, L. G. *J. Am. Chem. Soc.* **1995**, *117*, 5001−5002.
(23) Kombo, D. C.; Young, M. A.; Beveridge, D. L. *Proteins* **2000**, *39*, 212−215.
(24) Lapidus, L. J.; Eaton, W. A.; Hofrichter, J. *Proc. Nat. Acad. Sci.* **2000**, *97*, 7220−7225.
(25) Williams, S.; Causgrove, T. P.; Gilmanshin, R.; Fang, K. S.; Callender, R. H.; Woodruff, W. H.; Dyer, R. B. *Biochemistry* **1996**, *35*, 691−697.
(26) Munoz, V.; Thompson, P. A.; Hofrichter, J.; Eaton, W. A. *Nature* **1997**, *390*, 196−199.
(27) Mayor, U.; Johnson, C. M.; Daggett, V.; Fersht, A. R. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 13518−13522.
(28) Anfinsen, C. B. *Science* **1973**, *181*, 223−230.
(29) Stenflo, J. *Crit. Rev. Eukar. Gene Exp.* **1999**, *9*, 59−88.
(30) Manafioletti, G.; Brancolini, C.; Avanzi, G.; Schneider, C. *Mol. Cell Biol.* **1993**, *13*, 4976−4985.
(31) Kulman, J. D.; Harris, J. E.; Haldeman, B. A.; Davie, E. W. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 9058−9062.
(32) Kulman, J. D.; Harris, J. E.; Xie, L.; Davie, E. W. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 1370−1375.
(33) Bulaj, G.; Buczek, O.; Goodsell, I.; Jimenez, E. C.; Kranski, J.; Nielsen, J. S.; Garrett, J. E.; Olivera, B. M. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, 100-Supp. 2, 14562−14568.
(34) Simes, D. C.; Williamson, M. K.; Ortiz-Delgado, J. B.; Viegas, C. S. B.; Price, P. A.; Cancela, M. L. J. *Bone Min. Res.* **2003**, *18*, 244−259.
(35) Wang, C. P.; Yagi, K.; Lin, P. J.; Jin, D. Y.; Makabe, K. W.; Stafford, D. W. J. *Throm. Haemost.* **2002**, *1*, 118−123.
(36) Soriano-Garcia, M.; Padmanabhan, K.; Devos, A. M.; Tulinsky, A. *Biochemistry* **1992**, *31*, 2550−2566.
(37) Seshadri, T. P.; Tulinsky, A.; Skrzypczakjankun, E.; Park, C. H. *J. Mol. Biol.* **1991**, *220*, 481−494.
(38) Banner, D. W.; Darcy, A.; Chene, C.; Winkler, F. K.; Guha, A.; Konigsberg, W. H.; Nermerson, Y.; Kirchofer, D. *Nature* **1996**, *380*, 41−46.
(39) Freedman, S. J.; Furie, B. C.; Furie, B.; Baleja, J. D. *Biochemistry* **1995**, *34*, 12126−12137; *J. Biol. Chem.* **1995**, *270*, 7980−7987.
(40) Sunnerhagen, M.; Forsen, S.; Hoffren, A−M.; Drakenberg, T.; Teleman, O.; Stenflo, J. *Nat. Str. Biol.* **1995**, *2*, 504−509.
(41) Sunnerhagen, M.; Olah, G. A.; Stenflo, J.; Forsen, S.; Drakenberg, T.; Trewhella, J. *Biochemistry* **1996**, *35*, 11547−11559.
(42) Oganesyan, V.; Oganesyan, N.; Terzyan, Q. D.; Dauter, Z.; Esmon, N. L.; Esmon, C. T. *J. Biol. Chem.* **2002**, *277*, 24851−24854.
(43) Shikamoto, Y.; Morita, T.; Fujimoto, Z.; Mizuno, H. J. *Biol. Chem.* **2003**, *278*, 24090−24094.
(44) Wang, S. X.; Hur, E.; Sousa, C. A.; Brinen, L.; Slivka, E. J.; Fletterick, R. J. *Biochemistry* **2003**, *42*, 7959−7966.
(45) Prendergast, F. G.; Mann, F. G. *J. Biol. Chem.* **1977**, *252*, 840−850.
(46) Venkateswarlu, D.; Perera, L.; Darden, T.; Pedersen, L. G. *Biophys. J.* **2002**, *82*, 1190−1206.
(47) Duke, R. E.; Pedersen, L. G. *PMEMD 3.1* (version of AMBER), University of North Carolina-Chapel Hill, 2003.
(48) Freedman, S. J.; Blostein, M. D.; Baleja, J. D.; Jacob, M.; Furie, B. C.; Furie, B. *J. Biol. Chem.* **1996**, *271*, 116227−116236.