

Article

Conformational Plasticity-Rigidity Axis of the Coagulation Factor VII Zymogen Elucidated by Atomistic Simulations of the N-Terminally Truncated Factor VIIa Protease Domain

Jesper J. Madsen ^{1,*}  and Ole H. Olsen ^{2,*}¹ Global and Planetary Health, College of Public Health, University of South Florida, Tampa, FL 33612, USA² Novo Nordisk Foundation Center for Basic Metabolic Research, Section for Metabolic Receptology, University of Copenhagen, Blegdamsvej 3b, DK-2200 Copenhagen, Denmark

* Correspondence: jespemadsen@usf.edu (J.J.M.); oho@sund.ku.dk (O.H.O.)

Abstract: The vast majority of coagulation factor VII (FVII), a trypsin-like protease, circulates as the inactive zymogen. Activated FVII (FVIIa) is formed upon proteolytic activation of FVII, where it remains in a zymogen-like state and it is fully activated only when bound to tissue factor (TF). The catalytic domains of trypsin-like proteases adopt strikingly similar structures in their fully active forms. However, the dynamics and structures of the available corresponding zymogens reveal remarkable conformational plasticity of the protease domain prior to activation in many cases. Exactly how ligands and cofactors modulate the conformational dynamics and function of these proteases is not entirely understood. Here, we employ atomistic simulations of FVIIa (and variants hereof, including a TF-independent variant and N-terminally truncated variants) to provide fundamental insights with atomistic resolution into the plasticity-rigidity interplay of the protease domain conformations that appears to govern the functional response to proteolytic and allosteric activation. We argue that these findings are relevant to the FVII zymogen, whose structure has remained elusive despite substantial efforts. Our results shed light on the nature of FVII and demonstrate how conformational dynamics has played a crucial role in the evolutionary adaptation of regulatory mechanisms that were not present in the ancestral trypsin. Exploiting this knowledge could lead to engineering of protease variants for use as next-generation hemostatic therapeutics.

Keywords: zymogen; structure-function relationship; allostery; tissue factor; molecular dynamics; replica-exchange



Citation: Madsen, J.J.; Olsen, O.H. Conformational Plasticity-Rigidity Axis of the Coagulation Factor VII Zymogen Elucidated by Atomistic Simulations of the N-Terminally Truncated Factor VIIa Protease Domain. *Biomolecules* **2021**, *11*, 549. <https://doi.org/10.3390/biom11040549>

Academic Editor: Umesh Desai

Received: 22 February 2021

Accepted: 6 April 2021

Published: 8 April 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Coagulation factor VIIa (FVIIa) is a protease that initiates the enzymatic cascade of reactions leading to a burst of thrombin activation and blood clot formation [1–4]. FVIIa consists of an N-terminal membrane-binding domain rich in γ -carboxyglutamic acid residues (Gla domain), two epidermal growth factor-like domains (EGF1 and EGF2), and a trypsin-like serine protease (SP) domain, which is responsible for performing the enzymatic function [3]. Like trypsin and most other trypsin-like proteases, FVIIa is produced as a single-chain zymogen, FVII, that requires proteolytic cleavage for activation. However, unlike trypsin, FVIIa is a catalytically incompetent protease on its own, relying on the cofactor, Tissue Factor (TF), as an allosteric activator. While numerous distinct structural aspects of FVIIa affect its biochemical and physiological functions [5–16], the focal point is the SP domain when it comes to the activation transition. The activation process is thoroughly studied and largely understood in the trypsinogen-to-trypsin transition [17–21]: proteolytic cleavage before position I16 (chymotrypsin numbering) produces a new N-terminal tail that spontaneously inserts into the activation pocket (or “Ile cleft”) and forms a salt bridge with D194, leading to a maturation of the S1 pocket and catalytic competency. In addition, three loop segments, activation loop 1, 2, and 3 (AL1–3), undergo

stabilization [19,22]. The N-terminal insertion of trypsinogen activation has even been observed previously using accelerated molecular dynamics simulations [23]. While the activation transition of FVII is known to follow the same hallmark steps as trypsinogen-trypsin for it become catalytically active, many important details concerning FVII and the zymogen-like FVIIa are lacking.

The structural and dynamic details of FVII before, during, and after each of the two activation transitions (proteolytic and allosteric) are of considerable interest. Structures of FVIIa [24] and the FVIIa-TF complex [25,26] have been tremendously helpful in elucidating the molecular basis of hemostasis. Unfortunately, key forms of this important protease (including FVII and the free, uninhibited FVIIa in the zymogen-like form) have remained elusive despite heroic efforts by expert crystallographers over the last several decades. Modeling and simulation efforts have been employed to compensate for the lack of these structures with some success [27,28].

The native trypsin does not undergo conformational or dynamic changes upon binding of active-site inhibitors [29]. However, other trypsin-like proteases that are regulated by protein cofactors, e.g., FVIIa or thrombin [30], rely critically on a dynamic apo state, presumably because the inherent conformational flexibility allows for a mechanism of allosteric activation by the cofactor(s). Extensive experimental and computational investigations [28,31–33] have resulted in the hypothesis that two primary allosteric pathways mediate TF-induced enhancement on FVIIa activity [34,35]. One pathway tethers the 170-loop to the SP domain by stabilizing the $\alpha 1$ helix (a.k.a. the TF-binding helix) and enables G223 to form a hydrogen bond with the mainchain nitrogen of R170C [28,36], which restrains the flexibility of W215 in the substrate pocket S2-S3. The other pathway, on the other hand, travels from L163 to F225, stabilizing AL3 and, via, S185, AL2. This facilitates the insertion of the N-terminus and the proper salt bridge formation between I16 and D194. Knowledge about the activation mechanism and allostery has already yielded successful engineering of FVIIa variants with enhanced activity independent of stimulation by endogenous TF [37–39]. Of particular interest is a recently engineered FVIIa-trypsin variant, FVIIa_{VYT}, which is completely TF-independent due to its grafted trypsin 170-loop (as previously described [40–42]) in addition to another mutation, L163V [39]. The grafted trypsin 170-loop and the single mutation seem to stabilize the first pathway and second pathway, respectively.

In this work, we employ a computational modeling and simulation approach to generate a faithful realization of FVII by N-terminal truncation of FVIIa. The method is validated by demonstrating that it can be used meaningfully for trypsin and that the N-terminally truncated variant, trypsin-desIVG, adopts the structural and dynamic characteristics of trypsinogen. To our knowledge, this is the first report of successful inactivation of a trypsin-like protease by truncation of the inserted N-terminal tail.

2. Materials and Methods

2.1. Molecular Dynamics (MD) Simulations

2.1.1. Temperature-Replica-Exchange Molecular Dynamics (T-REMD) Simulations

T-REMD [43] simulations of the trypsin-like SP domains were performed using Gromacs 5.0.4 [44] with the CHARMM36m force field [45]. Ca^{2+} bound to the calcium-binding loop was kept while crystal waters, inhibitors, and other co-crystallized moieties were removed in the protein preparation step. Explicit TIP3P water [46] with 150 mM KCl was used for solvation. A 12-Å cut-off for the van der Waals forces were used. Electrostatic forces were computed using the particle mesh Ewald method [47]. The Verlet cut-off scheme was used. A total of 10 replicas were swapped in a temperature range of 310.15 K (physiological temperature) to 331.21 K (lower than the protein denaturing temperature) in steps of $\Delta T = 2.34$ K. Systems were initialized with random velocities at the reference temperatures. The temperature and pressure were controlled using the Nosé-Hoover [48–50] and Parrinello-Rahman [51,52] methods, respectively, to sample the *NPT* ensemble at 1 bar and the reference temperature for the replica. The integration time

step was 2 fs, enabled by using H-bond restraints [53]. Exchange of adjacent replicas was attempted every 500 steps (i.e., every 1 ps) and the average exchange acceptance probability was ~0.20 for all systems. The T-REMD simulations were carried out for 100–200 ns per replica (200 ns for FVIIa and FVIIa-desIVG; 100 ns for the rest), resulting in a combined simulation time of 1–2 μ s/system. Only the physiological temperature (310.15 K) replicas were used for analysis of structural and dynamic properties of the proteins. An overview of the systems simulated can be seen in Table 1. These systems were chosen for the following reasons: trypsin(ogen) would allow benchmarking our approach using a system that was extensively studied and where the zymogen structure was known. This gave confidence to applying the computational method to the study of FVII(a) (and a key variant hereof), where the zymogen remained elusive despite significant experimental efforts. The FVIIa_{VYT} variant was chosen in addition because it was a soluble tissue factor-independent variant of FVIIa that was functionally the closest to that of trypsin. On the spectrum from FVIIa to trypsin, the FVIIa_{VYT} variant was therefore presumed to be the most trypsin-like construct among engineered variants of FVIIa in terms of dynamics and functional characteristics.

Table 1. Summary of the structures used, and the systems simulated in this study.

| Name | Description | PDB ID | Simulation Protocol | Simulation Time (μ s) |
|------------------------------|---|-------------|---------------------|----------------------------|
| FVIIa | Human FVIIa SP domain. | 1dan_H [25] | T-REMD | 2 |
| FVIIa-desIVG | Human FVIIa SP domain. In silico zymogenized. | 1dan_H [25] | T-REMD | 2 |
| FVIIa _{VYT} | Human FVIIa _{VYT} SP domain, a TF-independent variant of FVIIa. | 6r2w_H [39] | T-REMD | 1 |
| FVIIa _{VYT} -desIVG | Human FVIIa _{VYT} SP domain, a TF-independent variant of FVIIa. In silico zymogenized. | 6r2w_H [39] | T-REMD | 1 |
| Trypsin | Bovine trypsin. | 1j8a [54] | T-REMD | 1 |
| Trypsin-desIVG | Bovine trypsin. In silico zymogenized. | 1j8a [54] | T-REMD | 1 |
| Trypsinogen | Bovine trypsinogen. | 2tgt [55] | T-REMD | 1 |

FVIIa, coagulation factor VIIa; desIVG, N-terminally truncated construct; SP, serine protease; TF, tissue factor; T-REMD, Temperature-replica-exchange molecular dynamics.

2.1.2. The In Silico Zymogenized Variants

The trypsin-like proteases were zymogenized in silico to give constructs without the inserted N-terminal tail, thus mimicking the zymogen (denoted by the term “desIVG”). This was done by truncation of the SP domain N-terminal by removing the leading three amino acid residues (¹⁶IVG¹⁸). The new N-terminal residue, G19, was neutralized by amidation to avoid local artifacts caused by the presence of the positive charge. The truncation approach was found preferable when compared with applying a force (e.g., using steered MD) to pull the N-terminal tail out of the activation pocket because of the salt bridge between I16 and D194, which tended to persevere and cause the structure to unravel as the I16-D194 pair was being pulled.

2.2. Analyses

2.2.1. Root-Mean-Squared Deviation (RMSD) and Root-Mean-Squared Fluctuation (RMSF)

RMSD calculations were done with PyMOL’s align command without outlier rejection (cycles = 0). RMSF calculations were done using gmx rmsf (taking contributions from each atom in the residue selection) and the simulated B-factors (Debye-Waller factors) were calculated as [56]:

$$B_i = \frac{8\pi^2}{3} \langle MSF_i \rangle_{MD} \quad (1)$$

2.2.2. E/Z-E*/Z* Distributions

E/Z vs. E*/Z* are distinguished by a specific structural segment (the 215–217 segment) in the SP of the trypsin-like proteases: One that is fully accessible to the substrate (E/Z) and the other occluded by collapse of the 215–217 segment. The conformations were categorized with a distance-based cut-off criterion. For E/Z, the defining distances are W215/CG to H57/CG > 6.5 Å and W215/CG to S195/OG > 6 Å; otherwise the configuration is designated E*/Z*.

2.2.3. Principal Components Analysis (PCA)

PCA is a useful technique for reducing the dimensionality of MD simulation data and extracting larger amplitude motions. The method works by computing the eigenvectors of the mass-weighted covariance matrix of the atomic positional fluctuations (here we used all heavy atoms) [57]. To perform the PCA analysis, we used pytraj [58] and scikit-learn [59].

2.2.4. Hierarchical Clustering Analysis

Structural clustering analysis allowed us to assign different structures observed during the MD simulations into distinct groups. The distance metric used was the pairwise RMSD and we performed the clustering on the C α subset of atoms. Agglomeration was performed using the average/UPGMA linkage approach [60]. Pytraj [58] and SciPy [61]/NumPy [62] were used to perform the clustering analysis.

2.2.5. Dynamic Cross-Correlations

Dynamic cross-correlations between the atoms within the molecule is a measure of the extent to which they move together in a correlated manner. The values in the resulting N \times N heatmap (N = the number of C α atoms) correspond to the dynamic cross-correlation between each (i,j) atom and a correlation of 1(−1) indicates complete (anti)correlation. A value of 0 indicates no correlation exists. To perform the dynamic cross-correlation analysis, we used MD-TASK [63]/MDTraj [64].

2.2.6. Plotting and Visualization

Inspection of structures and trajectories was done using VMD [65] and PyMOL (Schrödinger LLC, Open-Source version 2.3.0). Plotting was done using matplotlib [66] and Xmgrace (Grace version 5.1.25). Figure layout was done with Inkscape (<http://www.inkscape.org/>, 1 February 2021, version 1.0).

3. Results and Discussion

3.1. The N-Terminally Truncated Trypsin Gradually Adopts Structural and Dynamic Features Characteristic to Trypsinogen

The results and conclusions presented in this work rely on the application of an *in silico* protocol for interrogating the conformational structure and dynamics of the inactive zymogen precursor of a trypsin-like protease. Specifically, the protocol involves the truncating the protease domain N-terminal by removing the leading three amino acid residues (¹⁶IVG¹⁸) and subjecting the inactivated construct to MD simulations. Here, we adopted a replica-exchange scheme to enhance sampling of the MD simulation, a popular approach in the study of biomolecular systems where an appropriate collective variable (CV, or reaction coordinate) is not readily defined. Defining a CV for biasing a trypsin-like protease toward its zymogen structure presupposes deep knowledge about the system, locally and globally. Intuitively, one could in principle expel the N-terminal tail from the activation pocket by applying a local bias e.g., along a vector projection on the exit path. However, this would not necessarily allow the resulting structure and dynamics of the protein to approach the zymogen-like state faster (in a global sense) than by performing regular MD simulations of the N-terminally truncated construct, as non-local slow degrees of freedom would still hamper the transition. To best handle these challenges, we choose

to combine N-terminally truncation with an enhanced sampling technique that facilitates global configurational exploration, namely T-REMD.

To validate our protocol, we first applied it to the evolutionary ancestor of FVIIa, trypsin, and the trypsinogen-trypsin activation transition. A straightforward way to assess whether the structural characteristics of the simulated trypsin-desIVG leans more toward trypsin or trypsinogen is to compute the root-mean-square deviation (RMSD). We observed that while trypsin-desIVG initially retained its trypsin-like characteristics on a global RMSD-based scale, it quickly (in ~20 ns) transitioned into, and remained in, a conformation more closely resembling the simulated average structure of trypsinogen (Figure 1), albeit with significant structural fluctuations. This was remarkable because it happened despite trypsin-desIVG being initialized from the actual trypsin structure (except the truncated N-terminus) and was compared with the simulated average structures of its structural reference and another crystal structure entirely (namely trypsinogen). However, recall that the MD-averaged structures are generated by positional averaging and does therefore not correspond to a structural realization of the conformational ensemble. This is because positional averaging will not, in general, preserve all bond lengths, angles, etc. A matrix for computed RMSD values among all simulated average structures and reference crystal structures used in this study is available in Table S1 (Supplementary Materials).

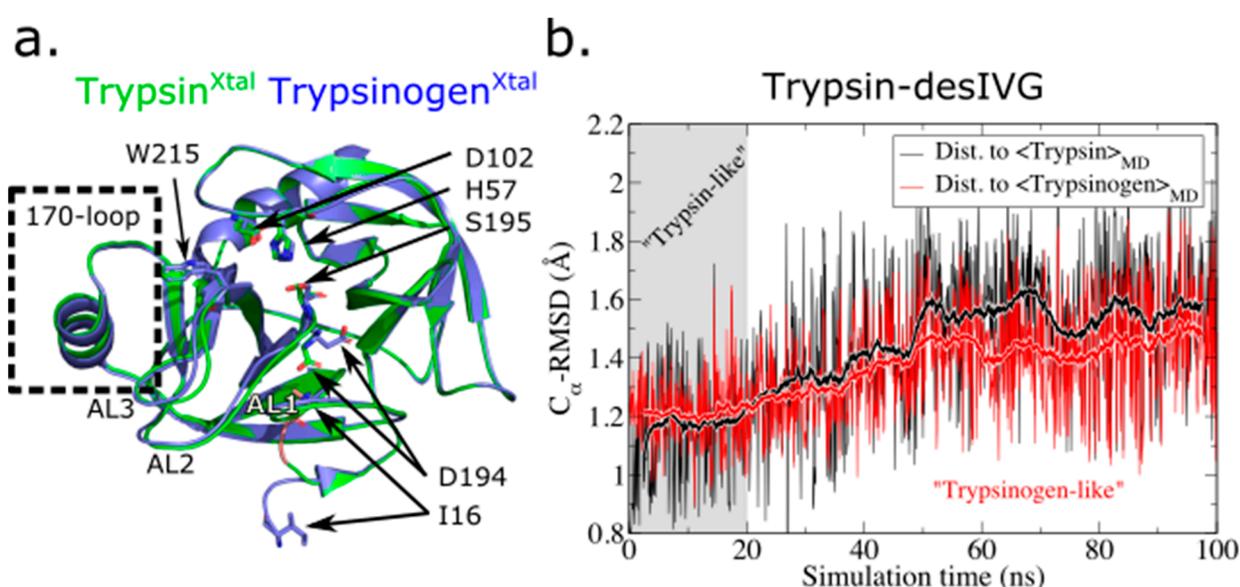


Figure 1. (a) Structural comparison of the trypsin and trypsinogen crystal structures. Select amino acid residues and spatial regions of special interest are labeled. Structure PDB IDs are indicated in Table 1. (b) The alpha-Carbon root-mean-squared deviation (C_{α} -Root-mean-squared deviation (RMSD)) from the simulated average structures of trypsin ($\langle \text{Trypsin} \rangle_{MD}$) and trypsinogen ($\langle \text{Trypsinogen} \rangle_{MD}$) is plotted for the trypsin-desIVG trajectory. The thicker lines with white edges indicate the running-average value over 50 steps.

An analysis of the per-residue root-mean-square fluctuation (RMSF) of the three simulated trypsin(ogen) constructs reveals that trypsin is very rigid with only minor flexible segments distributed throughout (e.g., at the tip of certain loops). Trypsin-desIVG and trypsinogen, on the other hand, exhibit a practically identical extended spatial region of high flexibility (Figure 2a,b). This flexible region encompasses activation loops 1 (residues 144–152), 2 (residues 184–193), and 3 (residues 221A–226) (AL1–3). Notice that the flexibility of the 170-loop is comparably low for all three trypsin(ogen) constructs. It is useful to further compare this with the experimentally determined B-factors from crystallography, which follow the same trends, albeit with much lower flexibility altogether due to the cryogenic temperatures and lattice conditions subjected to a protein in a crystal (Figure 2c).

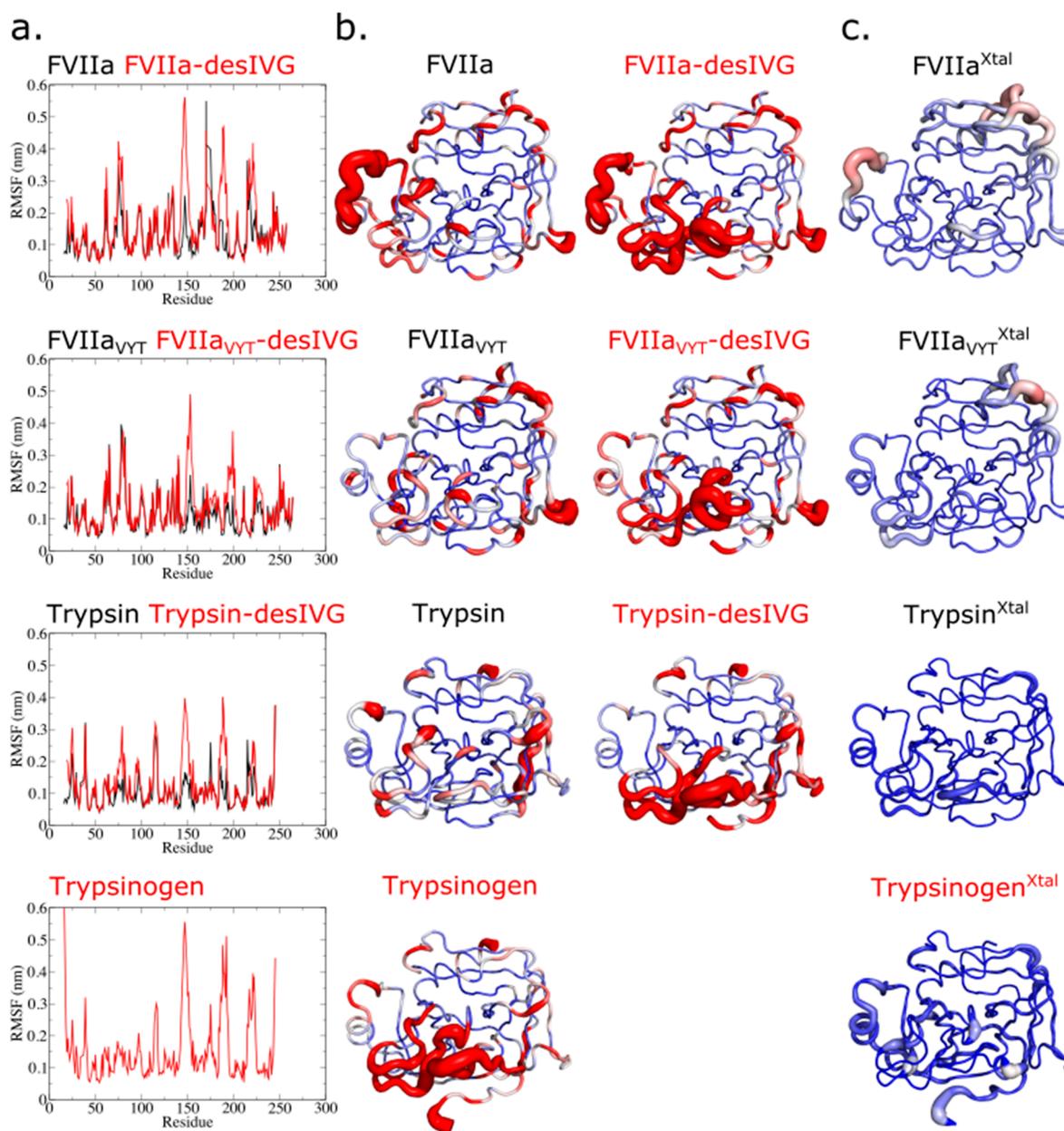


Figure 2. (a) Per-residue root-mean-square fluctuation (RMSF) of the simulated constructs. (b) Visualization of the B factors based on the computed RMSF values. Proteins are shown using a “sausage” cartoon representation where the thickness of the tube indicates the B-factor. The color scale goes from low (blue) over medium (white) to high (red). (c) Visualization of the crystallographic B-factors. Proteins are shown using a “sausage” cartoon representation where the thickness of the tube indicates the B-factor. The color scale goes from blue (low) over white (medium) to red (high). Structure PDB IDs are indicated in Table 1.

3.2. Conformational Flexibility of FVIIa, the TF-Independent Variant FVIIa_{VYT}, and Trypsin

The term flexibility refers here to the fast and reversible changes the system undergoes. The per-residue RMSF can be used as a measure to gauge the local conformational flexibility. A comparison between the three SP domains and their N-terminally truncated counterparts reveals that the activation pocket collapses (Figures S1 and S2), and that the surrounding segments (AL1–3 and S1 pocket) dramatically increase their structural flexibility as a result, when the inserted N-terminal tail is removed (Figure 2a,b). Such an increase is observed in all zymogen-enzyme pairs. It is thus expected to be shared by most trypsin-like zymogens based on the current results as well as available crystallographic evidence [67]. In

addition, we observed that the flexibility of the 170-loop was unaffected by the N-terminal insertion and the resulting structural maturation of the AL1–3 and S1 pocket, suggesting that a shortened 170-loop conferred intrinsic stability of the trypsin-like proteases but not vice-versa. It is informative to compare the flexibility of each structural element (e.g., 170-loop or AL1–3; see also Figure 1a and Figure S1) between the enzyme construct (FVIIa, FVIIa_{VYT}, trypsin) and the corresponding N-terminally truncated construct (-desIVG). The characteristics of, and differences between, the trypsinogen-trypsin and FVII-FVIIa conformational transitions are striking. Even though the percent sequence identity between trypsin/FVIIa and trypsin/FVIIa_{VYT}, are 41% and 42%, respectively [68], the carefully designed mutations of FVIIa_{VYT} (i.e., by grafting the trypsin 170-loop and L163V) manage to shift the conformational characteristics of the FVIIa_{VYT} SP domain to effectively mimic trypsin as gauged by local flexibility (Figure 2a,b). In particular, notice how the flexible regions (Figure 2b, red areas) throughout the FVIIa_{VYT} molecule matches up with the pattern observed in trypsin, not with that of FVIIa. This is clearly the case for the 170-loop, but closer inspection reveals that it also applies for the 215–217 segment that harbors W215. As we shall see in the following sections, various other structural and dynamic analyses of the conformational ensembles are consistent with this hypothesis.

In general, there is a similar distribution of flexible regions in the zymogen-like states and the active site triad conformations are preserved during the step of proteolytic activation, with FVIIa and FVIIa-desIVG exhibiting more variation in either measure. N-terminal tail insertion induces (directly or indirectly) the catalytically competent conformation by maturing the S1 pocket, AL1–3, and the oxyanion hole of the zymogens. This process is evidently accompanied by substantial conformational changes and changes in flexibility as well. The disorder-to-order transition is likely an essential regulatory mechanism for preventing premature proteolysis in more advanced organisms. The zymogen-like property of FVIIa may have its origins in a self-inhibitory property due, at least in part, to the outstanding conformational flexibility. It is unclear whether this regulatory framework is evolutionarily conserved among all species carrying FVIIa, something that has recently sparked some debate [69–71]. From the literature, we know that the energetic origins of the conformational changes induced by N-terminus insertion are both electrostatic and hydrophobic [23,72,73]. Thus, the zymogen-to-enzyme activation transition can be considered a cooperative process with distinct energetic and entropic contributions, making it difficult to delineate adequately on the basis of static structures. It is worth mentioning, however, that we were recently able to rationalize important aspects of FVIIa allostery by using the ensemble-refinement technique [74,75] that combines X-ray structure refinement with molecular dynamics in order to build an ensemble of structures consistent with the electron density map [76].

3.3. The Allosteric E-E* and Z-Z* Equilibria Are Influenced by the 170-Loop, but Independent of N-Terminus Insertion

Structural and biochemical investigations of thrombin, another trypsin-like protease, has suggested that a structural segment around residues 215–217 exhibits distinct conformations of direct functional relevance [30,77–82]. The terminology used here, E(Z) and E*(Z*), refer to the enzyme (E) or zymogen (Z) 215–217 segment being in the trypsin-like matured conformation (E/Z) or the substrate-occluding conformation (E*/Z*). Recent evidence supports that a similar conformational change is relevant for FVIIa [28,33,42]. In particular, the 215–217 segment has been implicated in substrate accessibility to the active site and cofactor-mediated allosteric regulation [28,42]. We have previously used MD simulations to successfully explore the conformational equilibrium of the 215–217 segment in variants of FVIIa [42].

Surprisingly, the current work suggests that proteolytic (in)activation and N-terminus insertion has little to no influence on E-E* and Z-Z* equilibria in any of the investigated protease variants (Figure 3a,b). FVIIa and FVIIa-desIVG have similarly substrate-occluding 215–217 segment conformations (Figure 3a). We have previously shown that FVIIa shifts its population from E* to E when complexed with TF [42] and we speculate that the

same might be true for FVII. The time evolution of the E/Z-E*/Z* states of the individual constructs show that transitions are intermittent and plentiful for FVIIa and FVIIa-desIVG, which are the two constructs that regularly visits both their respective E/Z and E*/Z* conformations (Figure S3). We find that the characteristic behavior of FVIIa_{VYT}(-desIVG) is effectively shifted toward that of trypsin(ogen), which is essentially locked in the E(Z) configuration (Figure 3b,c). These results suggest that the allosteric equilibria of the 215–217 segment is governed by the 170-loop, independently of N-terminus insertion, supporting the notion that W215 in the 215–217 segment is the endpoint of the proposed first allosteric pathway [35] alluded to the introduction.

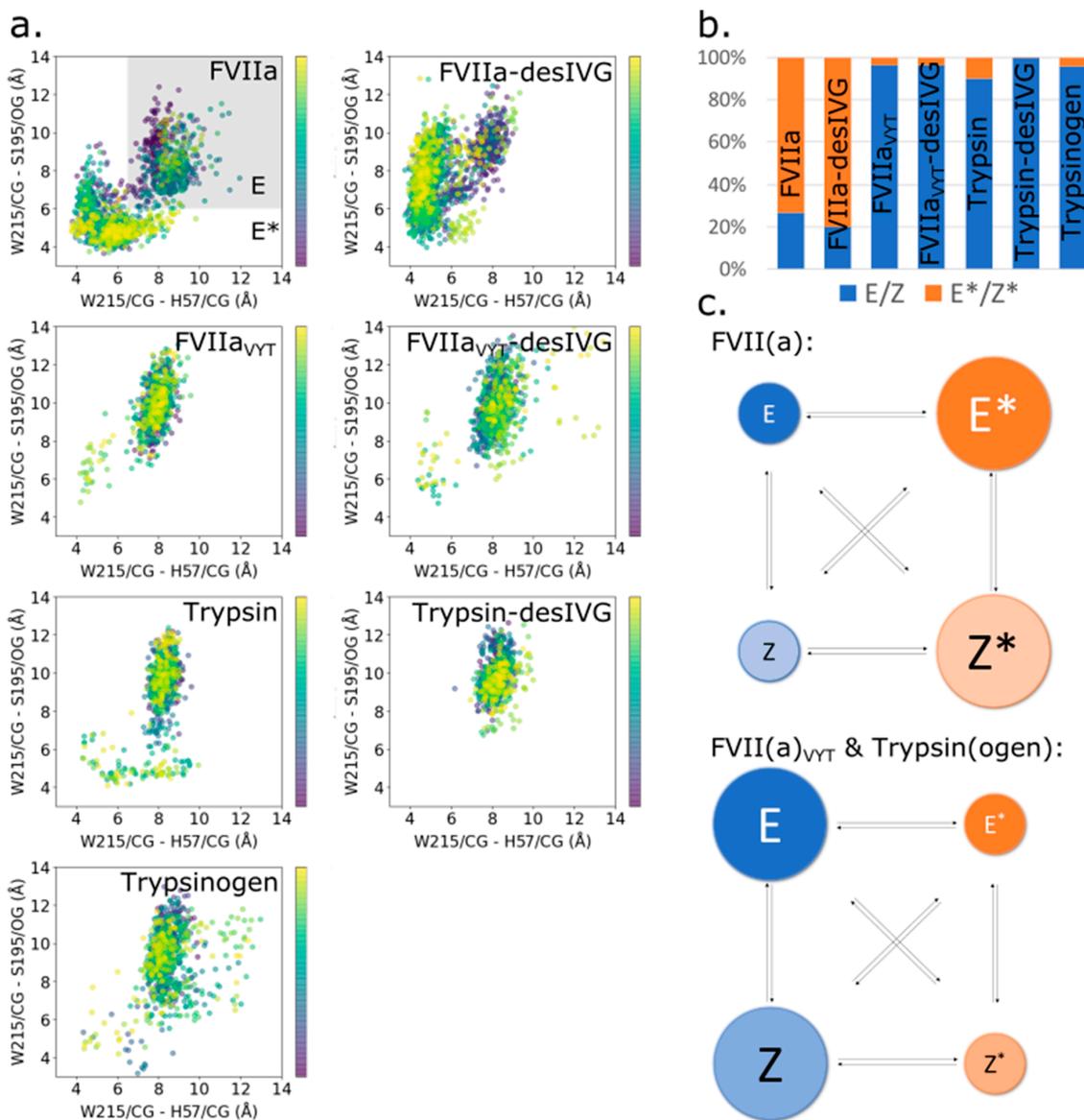


Figure 3. (a) Scatter plots of the W215 location relative to the active site residues H57 and S195 of the simulated constructs. The color bar scale indicates the simulation time starting from 0 ns (purple dots) to the end of the trajectory (yellow dots). (b) Quantification of E-E*/Z-Z* equilibria from (a) based on the cut-off criterion. Errors in the estimates based on bootstrapping with 1000 samples were <1% and errorbars showing the bootstrap standard error are therefore too tiny to be seen. (c) Schematic depiction of relative populations of E-E*/Z-Z*. FVII(a) is found predominantly in Z*/E* configurations, while FVII(a)_{VYT} and trypsin(ogen) are found to almost entirely occupy Z/E configurations. Size ratios between the spheres are indicative only and not to scale; see the histogram in (b) for accurate quantification.

3.4. Conformational Plasticity-Rigidity Axis Intrinsically Governed by the 170-Loop with the N-Terminus Insertion as Extrinsic Modulator

There are distinct differences in both structure and dynamics of FVIIa, FVIIa_{VYT}, and trypsin. In the above (Section 3.2), we address the conformational flexibility of the simulated proteins and variants by assessing structural fluctuations. Here, we look into notions of plasticity and rigidity, which are increasingly appreciated as essential for elucidating structural and functional characteristics of proteins [21,83–88]. Though, the terms are ill-defined despite being commonly used. We adopt the definition by Csermely, who propose that plasticity and rigidity are opposing characteristics of the system due to external or internal changes [89]. Structural plasticity and rigidity are defined in terms of structural properties and they need not necessarily be synonymous with their functional counterparts.

We performed PCA of the simulations of the three proteases and their corresponding zymogen counterparts to complement the structural analyses (Figure 4). PCA is useful technique for reducing the dimensionality of MD simulation data and extract larger amplitude motions and, as such, allows us to concretely compare the dynamics exhibited by the proteins in the essential subspace [57,90]. The analysis reveals that the simulated conformational ensembles for FVIIa and FVIIa-desIVG exhibit much larger subspace variation; even more so than that of trypsinogen, which is the least compact of the rest. Here, we use the term “compactness” to visually assess the PCA analyses. Used this way, (lack of) “compactness” is consistent with conformational rigidity (plasticity) as defined and used in the current work. Generally, we find that the constructs can be ordered in terms of their subspace compactness as follows (from most to least): trypsin, FVIIa_{VYT}, trypsin-desIVG, FVIIa_{VYT}-desIVG, trypsinogen, FVIIa, and FVIIa-desIVG (Figure 4). The unambiguous trends that emerge by inspection of the distributions of the first two principal components are that (1) truncating the N-terminal tail increases subspace variability (with trypsin-desIVG approximating trypsinogen), (2) FVIIa_{VYT} effectively adopts the dynamical characteristics of trypsin, and (3) FVIIa/FVIIa-desIVG exhibits the most subspace variability by far. Furthermore, the gaussian and unimodal nature of the PC1-PC2 scatter plots for FVIIa indicates the presence of substantial conformational plasticity of FVII/FVIIa. The compact, non-gaussian and multimodal nature of FVIIa_{VYT} and trypsin, on the other hand, indicates conformational rigidity. Recall that major conformational changes between distinct conformational states are expected to be captured reasonably well by the first few PCs while the dynamics of a rigid protein in a narrow energetic well may involve independent motions of many different parts of the protein and, as such, exhibit higher dimensionality in the essential dynamics space as well as local frustration [57,90–92].

These observations demonstrate that the *in silico* zymogenization method we apply enables us to transform the SP domains into their corresponding zymogens, not just measured by calculating the RMSD to a reference structure (as shown in Figure 1), but also in the essential dynamics subspace (Figure 4). Taken together with considerations of molecular flexibility (Figure 2), these analyses are effective in elucidating the structural and dynamic features of the conformational ensembles. In addition, there is some evidence from the current work that inactivation by N-terminus truncation gives rise to a bifurcated conformational ensemble for trypsin and FVIIa_{VYT}, but for not for FVIIa. This finding comes about by performing RMSD-based hierarchical clustering on the simulated ensemble of structures and characterizing the number of clusters using a cut-off or a dendrogram (Figures S4–S10). The increase in the number of distinct clusters after N-terminal truncation for trypsin and FVIIa_{VYT}, and lack hereof for FVIIa, is consistent with the rigidity of trypsin and FVIIa_{VYT}, and the plasticity of FVIIa. However, the choice of clustering method has some influence on the results. To further support the idea of different occurrence of structural clusters in the simulated ensembles between trypsin/FVIIa_{VYT} and FVIIa, we computed per-residue dynamic cross-correlations over the individual trajectories. The dynamic cross-correlations were curiously similar for all cases, suggesting that the dynamic footprints of a bifurcated conformational ensemble may be averaged out over the sampled time scale (Figures S11–S17). It is conceivable that the similarities among the

cross-correlation heatmaps are at least in part due to the frequent and constant exchanges between replicas.

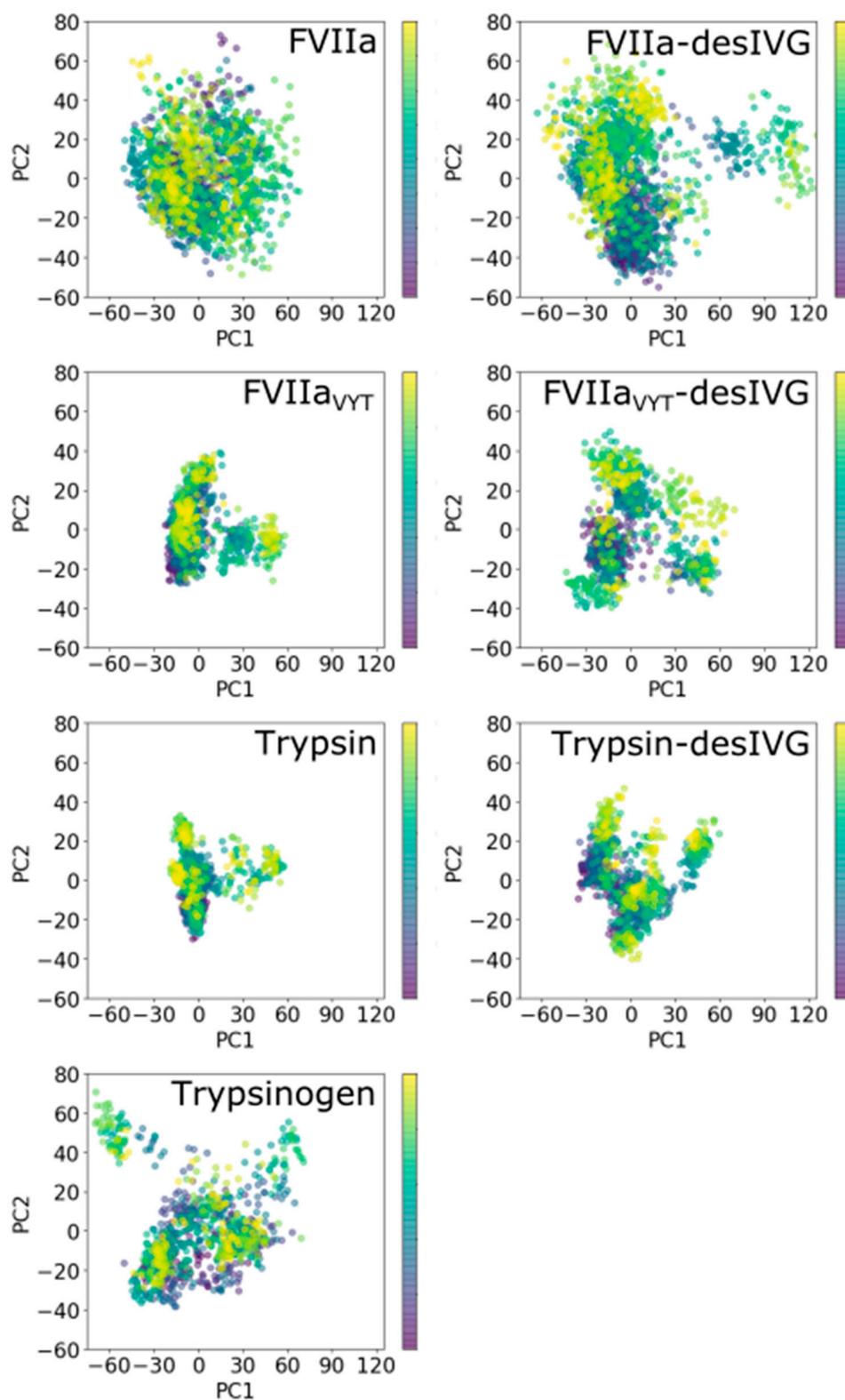


Figure 4. Principal component analysis of the simulated constructs. The plots show the first (PC1) and second (PC2) principal component against each other. The color bar scale indicates the simulation time starting from 0 ns (purple dots) to the end of the trajectory (yellow dots).

It's remarkable that the engineered variant, FVIIa_{VYT}, clearly exhibits the characteristics of trypsin in excellent agreement with experimental measurements, reassuring us that this variant effectively “beats TF at its own game” as proclaimed [39]. Concretely, this means that FVIIa_{VYT} exhibited 60-fold higher amidolytic activity than FVIIa, and displayed similar factor X activation and antithrombin inhibition kinetics to the FVIIa-sTF complex [39]. It is perplexing, however, that the chemical reactivity of the N-terminal tail of FVIIa_{VYT} (gauged via the carbamylation assay [93]) falls strictly in-between that of FVIIa and FVIIa-sTF, with a three-fold higher KNCO $t_{1/2}$ than the former and a four-fold lower KNCO $t_{1/2}$ than the latter. Our results do not resolve this apparent paradox and further investigations are needed. In particular, it remains unclear from a structural point of view exactly how KNCO chemically reacts with the N-terminal primary amine of I16. The trypsinogen-to-trypsin transition would suggest that the N-terminal tail is available for chemical modification only once expelled from the activation pocket and extruding outward from the SP domain body (in a trypsinogen-like configuration [55]). However, it might be possible for water and KCNO to access the N-terminal residue between the loop segment that forms the rim of the S1 pocket (residues 190–195) and AL1. MD simulations have previously indicated that solvation and destabilization of the I16-D194 salt bridge can occur by this mechanism [28] and it would be interesting to look further into this aspect using specialized analysis methods [94–96] or collective-variable (CV)-based enhanced sampling methods such as metadynamics [97]. This would be particularly pertinent in relation to assessing the solvent accessibility of I16 in different conformational states of FVIIa.

3.5. A Glimpse of the Elusive FVII Zymogen Exposed at Last?

The structural basis for the functional and biochemical similarities between the zymogen-like FVIIa and FVII is poorly understood. The main reasons are that key forms of this important protease (including FVII and the free, uninhibited FVIIa in the zymogen-like form) have remained elusive despite heroic efforts by expert crystallographers. An eccentric activation mechanism of FVII involving reconfiguration of β -strands of the protein structural scaffold has been proposed based on a crystal structure [9]. However, detailed experimental and theoretical investigations [31,98,99] do not support the reregistration hypothesis. Remarkably, a structure of the free FVIIa without an inhibitor has been obtained by soaking out the inhibitor from benzamidinium-inhibited FVIIa crystals [24]. While this structure alludes to added flexibility of FVIIa in the absence of inhibitor, it does not resolve the issues relating to the zymogen-like character of the free and uninhibited FVIIa in solution.

It is a reasonable assumption that the functional similarities between the zymogen-like FVIIa and FVII arise due to their structural and dynamic similarities. This work sheds some light on the properties of both FVIIa and FVII insofar that FVIIa-desIVG is a reasonable surrogate for FVII, which we have argued in the above. It must be emphasized, however, that it is not known how many distinct conformational basins (or dominant conformational intermediates) there are in the activation process from FVII via FVIIa to FVIIa-TF. Our results demonstrate the vivid flexibility of parts of the FVIIa-desIVG molecule (Figures 2 and 3) and the inherent plastic character of its conformational ensemble (Figure 4). This is consistent with the view that FVII may not have a dominant, well-defined structure in large parts of the molecule.

It is worth noting that even in cases where the zymogen structures of advanced trypsin-like proteases are available (such as profactor D [100] and factor XI [101]; see the review by Gohara and Di Cera for an overview of available zymogen structures of trypsin-like proteases [67]), a recurring theme is that the parts of the structures that we are most interested in are too flexible and divulge insufficient electron density for reliable model building.

4. Conclusions

We report MD simulations of FVIIa, FVIIa_{VYT} (a TF-independent variant), trypsin(ogen) and their N-terminally truncated counterparts. By employing the replica exchange method for improved sampling, we have explored the conformational dynamics of these protease constructs at effective time scales beyond that what the corresponding conventional MD simulations would have explored.

Our results demonstrate the feasibility of a *in silico* zymogenization protocol that involves N-terminal truncation of the three leading amino acid residues of the activated protease to produce a construct that takes on the structural and dynamic characteristics of the zymogen. The method is used to shed light on the structure and dynamics of the elusive FVII zymogen. By comparing and contrasting with trypsin and the TF-independent FVIIa variant, FVIIa_{VYT}, we show that the conformational plasticity-rigidity axis is primarily governed by the 170-loop and modulated by N-terminal tail insertion. Conformational flexibility is affected mostly by N-terminal tail insertion. Finally, the allosteric E-E* and Z-Z* equilibria are influenced by the 170-loop, but independent of N-terminal tail insertion.

That a relatively minor modification in the 170-loop can make FVIIa TF-independent likely means that TF's role in supporting the zymogen-like FVIIa is compensating for an effect that the trypsin 170-loop supplies to the SP domain in terms of flexibility and rigidity (but that the FVIIa 170-loop fails to). Considering that N-terminal tail insertion is a mechanistic step conserved among other trypsin-like proteases, this naturally restricts the space of evolutionary possibility to making adaptations that utilize the conformational plasticity-rigidity axis.

Supplementary Materials: The following are available online at <https://www.mdpi.com/article/10.3390/biom11040549/s1>, Table S1: RMSD values from MD trajectories are reported with reference to the averaged structure over the whole trajectory, Figure S1: FVIIa (PDB ID: 1dan_H) shown as ribbon diagram with certain key loops (170-loop, AL1-3) and amino acids marked with colored labels, Figure S2: Structure of FVIIa (top panels) and a closeup view of the region determining the E/Z-E*/Z* conformation (bottom panels) at different stages during the simulation, Figure S3: Time evolution of E/Z-E*/Z* conformational transitions during the simulations, Figure S4: Dendrogram of RMSD-based hierarchical clusters from the simulated ensemble of FVIIa, Figure S5: Dendrogram of RMSD-based hierarchical clusters from the simulated ensemble of FVIIa-desIVG, Figure S6: Dendrogram of RMSD-based hierarchical clusters from the simulated ensemble of FVIIaVYT, Figure S7: Dendrogram of RMSD-based hierarchical clusters from the simulated ensemble of FVIIaVYT-desIVG, Figure S8: Dendrogram of RMSD-based hierarchical clusters from the simulated ensemble of trypsin, Figure S9: Dendrogram of RMSD-based hierarchical clusters from the simulated ensemble of trypsin-desIVG, Figure S10: Dendrogram of RMSD-based hierarchical clusters from the simulated ensemble of trypsinogen, Figure S11: Heatmap of per-residue dynamic cross-correlations over the simulated trajectory of FVIIa, Figure S12: Heatmap of per-residue dynamic cross-correlations over the simulated trajectory of FVIIa-desIVG, Figure S13: Heatmap of per-residue dynamic cross-correlations over the simulated trajectory of FVIIaVYT, Figure S14: Heatmap of per-residue dynamic cross-correlations over the simulated trajectory of FVIIaVYT-desIVG, Figure S15: Heatmap of per-residue dynamic cross-correlations over the simulated trajectory of trypsin, Figure S16: Heatmap of per-residue dynamic cross-correlations over the simulated trajectory of trypsin-desIVG, Figure S17: Heatmap of per-residue dynamic cross-correlations over the simulated trajectory of trypsinogen.

Author Contributions: Conceptualization, J.J.M. and O.H.O.; methodology, J.J.M.; software, J.J.M.; validation, J.J.M.; formal analysis, J.J.M.; investigation, J.J.M.; resources, J.J.M.; data curation, J.J.M.; writing—original draft preparation, J.J.M.; writing—review and editing, J.J.M. and O.H.O.; visualization, J.J.M.; supervision, O.H.O. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. MacFarlane, R. An enzyme cascade in the blood clotting mechanism, and its function as a biochemical amplifier. *Nature* **1964**, *202*, 498–499. [[CrossRef](#)]
2. Davie, E.W.; Ratnoff, O.D. Waterfall Sequence for Intrinsic Blood Clotting. *Science* **1964**, *145*, 1310–1312. [[CrossRef](#)]
3. Furie, B.; Furie, B.C. The molecular basis of blood coagulation. *Cell* **1988**, *53*, 505–518. [[CrossRef](#)]
4. Davie, E.W.; Fujikawa, K.; Kisiel, W. The coagulation cascade: Initiation, maintenance, and regulation. *Biochemistry* **1991**, *30*, 10363–10370. [[CrossRef](#)] [[PubMed](#)]
5. Neuenschwander, P.F.; Morrissey, J.H. 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.* **1994**, *269*, 8007–8013. [[CrossRef](#)]
6. Leonard, B.J.N.; Clarke, B.J.; Sridhara, S.; Kelley, R.; Ofosu, F.A.; Blajchman, M.A. Activation and active site occupation alter conformation in the region of the first epidermal growth factor-like domain of human factor VII. *J. Biol. Chem.* **2000**, *275*, 34894–34900. [[CrossRef](#)]
7. Prasad, R.; Sen, P. Structural modulation of factor VIIa by full-length tissue factor (TF1-263): Implication of novel interactions between EGF2 domain and TF. *J. Biomol. Struct. Dyn.* **2018**, *36*, 621–633. [[CrossRef](#)]
8. Ohkubo, Y.Z.; Madsen, J.J. Uncovering Membrane-Bound Models of Coagulation Factors by Combined Experimental and Computational Approaches. *Thromb. Haemost.* **2021**, in press. [[CrossRef](#)]
9. Eigenbrot, C.; Kirchhofer, D.; Dennis, M. The factor VII zymogen structure reveals reregistration of β strands during activation. *Structure* **2001**, *9*, 627–636. [[CrossRef](#)]
10. Lazarus, R.A.; Olivero, A.G.; Eigenbrot, C.; Kirchhofer, D. Inhibitors of Tissue Factor—Factor VIIa for Anticoagulant Therapy. *Curr. Med. Chem.* **2004**, *11*, 2275. [[CrossRef](#)] [[PubMed](#)]
11. Colina, C.M.; Venkateswarlu, D.; Duke, R.; Perera, L.; Pedersen, L.G. What causes the enhancement of activity of factor VIIa by tissue factor? *J. Thromb. Haemost.* **2006**, *4*, 2726–2729. [[CrossRef](#)]
12. Ohkubo, Y.Z.; Tajkhorshid, E. Distinct structural and adhesive roles of Ca²⁺ in membrane binding of blood coagulation factors. *Structure* **2008**, *16*, 72–81. [[CrossRef](#)]
13. Ohkubo, Y.Z.; Morrissey, J.H.; Tajkhorshid, E. Dynamical view of membrane binding and complex formation of human factor VIIa and tissue factor. *J. Thromb. Haemost.* **2010**, *8*, 1044–1053. [[CrossRef](#)] [[PubMed](#)]
14. Lee, C.J.; Chandrasekaran, V.; Wu, S.; Duke, R.E.; Pedersen, L.G. Recent Estimates of the Structure of the Factor VIIa (FVIIa)/Tissue Factor (TF) and Factor Xa (FXa) Ternary Complex. *Thromb. Res.* **2010**, *125*, S7–S10. [[CrossRef](#)] [[PubMed](#)]
15. Lee, C.J.; Wu, S.; Bartolotti, L.J.; Pedersen, L.G. Molecular dynamics simulations of the binary complex of human tissue factor (TF(1-242)) and factor VIIa (TF(1-242) /fVIIa) on a 4:1 POPC/POPS lipid bilayer. *J. Thromb. Haemost.* **2012**, *10*, 2402–2405. [[CrossRef](#)] [[PubMed](#)]
16. Persson, E.; Madsen, J.J.; Olsen, O.H. The length of the linker between the epidermal growth factor-like domains in factor VIIa is critical for a productive interaction with tissue factor. *Protein Sci.* **2014**, *23*, 1717–1727. [[CrossRef](#)]
17. Bode, W.; Huber, R. Induction of the bovine trypsinogen—Trypsin transition by peptides sequentially similar to the N-terminus of trypsin. *FEBS Lett.* **1976**, *68*, 231–236. [[CrossRef](#)]
18. Bode, W.; Schwager, P.; Huber, R. The transition of bovine trypsinogen to a trypsin-like state upon strong ligand binding. *J. Mol. Biol.* **1978**, *118*, 99–112. [[CrossRef](#)]
19. Huber, R.; Bode, W. Structural basis of the activation and action of trypsin. *Acc. Chem. Res.* **1978**, *266*, 114–122. [[CrossRef](#)]
20. Bode, W. The transition of bovine trypsinogen to a trypsin-like state upon strong ligand binding: II. The binding of the pancreatic trypsin inhibitor and of isoleucine-valine and of sequentially related peptides to trypsinogen and to p-guanidinobenzoate-trypsinogen. *J. Mol. Biol.* **1979**, *127*, 357–374. [[CrossRef](#)]
21. Plattner, N.; Noé, F. Protein conformational plasticity and complex ligand-binding kinetics explored by atomistic simulations and Markov models. *Nat. Commun.* **2015**, *6*, 7653. [[CrossRef](#)]
22. Fehlhammer, H.; Bode, W.; Huber, R. Crystal structure of bovine trypsinogen at 1.8 Å resolution: II. Crystallographic refinement, refined crystal structure and comparison with bovine trypsin. *J. Mol. Biol.* **1977**, *111*, 415–438. [[CrossRef](#)]
23. Boechi, L.; Pierce, L.; Komives, E.A.; McCammon, J.A. Trypsinogen activation as observed in accelerated molecular dynamics simulations. *Protein Sci.* **2014**, *23*, 1550–1558. [[CrossRef](#)] [[PubMed](#)]
24. Sichler, K.; Banner, D.W.; D’Arcy, A.; Hopfner, K.-P.; Huber, R.; Bode, W.; Kresse, G.-B.; Kopetzki, E.; Brandstetter, H. Crystal Structures of Uninhibited Factor VIIa Link its Cofactor and Substrate-assisted Activation to Specific Interactions. *J. Mol. Biol.* **2002**, *322*, 591–603. [[CrossRef](#)]
25. Banner, D.W.; D’Arcy, A.; Chène, C.; Winkler, F.K.; Guha, A.; Konigsberg, W.H.; Nemerson, Y.; Kirchhofer, D. The crystal structure of the complex of blood coagulation factor VIIa with soluble tissue factor. *Nature* **1996**, *380*, 41–46. [[CrossRef](#)]
26. Pike, A.C.W.; Brzozowski, A.M.; Roberts, S.M.; Olsen, O.H.; Persson, E. Structure of human factor VIIa and its implications for the triggering of blood coagulation. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 8925–8930. [[CrossRef](#)]
27. Perera, L.; Darden, T.A.; Pedersen, L.G. Predicted solution structure of zymogen human coagulation FVII. *J. Comput. Chem.* **2002**, *23*, 35–47. [[CrossRef](#)]
28. Madsen, J.J.; Persson, E.; Olsen, O.H. Tissue factor activates allosteric networks in factor VIIa through structural and dynamic changes. *J. Thromb. Haemost.* **2015**, *13*, 262–267. [[CrossRef](#)]

29. Peterson, F.C.; Gordon, N.C.; Gettins, P.G.W. High-level bacterial expression and ¹⁵N-alanine-labeling of bovine trypsin. Application to the study of trypsin–Inhibitor complexes and trypsinogen activation by NMR spectroscopy. *Biochemistry* **2001**, *40*, 6275–6283. [[CrossRef](#)]
30. Lechtenberg, B.C.; Johnson, D.J.D.; Freund, S.M.V.; Huntington, J.A. NMR resonance assignments of thrombin reveal the conformational and dynamic effects of ligation. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 14087–14092. [[CrossRef](#)] [[PubMed](#)]
31. Rand, K.D.; Jørgensen, T.J.D.; Olsen, O.H.; Persson, E.; Jensen, O.N.; Stennicke, H.R.; Andersen, M.D. Allosteric activation of coagulation factor VIIa visualized by hydrogen exchange. *J. Biol. Chem.* **2006**, *281*, 23018–23024. [[CrossRef](#)]
32. Rand, K.D.; Andersen, M.D.; Olsen, O.H.; Jørgensen, T.J.D.; Ostergaard, H.; Jensen, O.N.; Stennicke, H.R.; Persson, E. The origins of enhanced activity in factor VIIa analogs and the interplay between key allosteric sites revealed by hydrogen exchange mass spectrometry. *J. Biol. Chem.* **2008**, *283*, 13378–13387. [[CrossRef](#)]
33. Song, H.; Olsen, O.H.; Persson, E.; Rand, K.D. Sites involved in intra- and interdomain allostery associated with the activation of factor VIIa pinpointed by hydrogen-deuterium exchange and electron transfer dissociation mass spectrometry. *J. Biol. Chem.* **2014**, *289*, 35388–35396. [[CrossRef](#)] [[PubMed](#)]
34. Olsen, O.H.; Persson, E. Cofactor-induced and mutational activity enhancement of coagulation factor VIIa. *Cell. Mol. Life Sci.* **2008**, *65*, 953–963. [[CrossRef](#)] [[PubMed](#)]
35. Persson, E.; Olsen, O.H. Allosteric activation of coagulation factor VIIa. *Front. Biosci.* **2011**, *16*, 3156–3163. [[CrossRef](#)] [[PubMed](#)]
36. Persson, E.; Olsen, O.H. Activation loop 3 and the 170 loop interact in the active conformation of coagulation factor VIIa. *FEBS J.* **2009**, *276*, 3099–3109. [[CrossRef](#)] [[PubMed](#)]
37. Persson, E.; Kjalke, M.; Olsen, O.H. Rational design of coagulation factor VIIa variants with substantially increased intrinsic activity. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 13583–13588. [[CrossRef](#)]
38. Nielsen, A.L.; Sorensen, A.B.; Holmberg, H.L.; Gandhi, P.S.; Karlsson, J.; Buchardt, J.; Lamberth, K.; Kjelgaard-Hansen, M.; Ley, C.D.; Sørensen, B.B.; et al. Engineering of a membrane-triggered activity switch in coagulation factor VIIa. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 12454–12459. [[CrossRef](#)]
39. Sorensen, A.B.; Tuneew, I.; Svensson, L.A.; Persson, E.; Østergaard, H.; Overgaard, M.T.; Olsen, O.H.; Gandhi, P.S. Beating tissue factor at its own game: Design and properties of a soluble tissue factor-independent coagulation factor VIIa. *J. Biol. Chem.* **2020**, *295*, 517–528. [[CrossRef](#)]
40. Soejima, K.; Mizuguchi, J.; Yuguchi, M.; Nakagaki, T.; Higashi, S.; Iwanaga, S. Factor VIIa modified in the 170 loop shows enhanced catalytic activity but does not change the zymogen-like property. *J. Biol. Chem.* **2001**, *276*, 17229–17235. [[CrossRef](#)]
41. Soejima, K.; Yuguchi, M.; Mizuguchi, J.; Tomokiyo, K.; Nakashima, T.; Nakagaki, T.; Iwanaga, S. The 99 and 170 loop-modified factor VIIa mutants show enhanced catalytic activity without tissue factor. *J. Biol. Chem.* **2002**, *277*, 49027–49035. [[CrossRef](#)]
42. Sorensen, A.B.; Madsen, J.J.; Svensson, L.A.; Pedersen, A.A.; Østergaard, H.; Overgaard, M.T.; Olsen, O.H.; Gandhi, P.S. Molecular basis of enhanced activity in factor VIIa-trypsin variants conveys insights into tissue factor-mediated allosteric regulation of factor VIIa activity. *J. Biol. Chem.* **2016**, *291*, 4671–4683. [[CrossRef](#)] [[PubMed](#)]
43. Sugita, Y.; Okamoto, Y. Replica-exchange molecular dynamics method for protein folding. *Chem. Phys. Lett.* **1999**, *314*, 141–151. [[CrossRef](#)]
44. Abraham, M.J.; Murtola, T.; Schulz, R.; Pall, S.; Smith, J.C.; Hess, B.; Lindah, E. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **2015**, *1–2*, 19–25. [[CrossRef](#)]
45. Huang, J.; Rauscher, S.; Nawrocki, G.; Ran, T.; Feig, M.; De Groot, B.L.; Grubmüller, H.; Mackerell, A.D. Charmm36M: An Improved Force Field for Folded and Intrinsically Disordered Proteins. *Nat. Methods* **2017**, *14*, 71–73. [[CrossRef](#)] [[PubMed](#)]
46. Jørgensen, W.L.; Chandrasekhar, J.; Madura, J.D.; Impey, R.W.; Klein, M.L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926–935. [[CrossRef](#)]
47. Darden, T.; York, D.; Pedersen, L. Particle mesh Ewald: An $N^2 \log(N)$ method for Ewald sums in large systems. *J. Chem. Phys.* **1993**, *98*, 10089–10092. [[CrossRef](#)]
48. Nose, S. A unified formulation of the constant temperature molecular dynamics methods. *J. Chem. Phys.* **1984**, *81*, 511. [[CrossRef](#)]
49. Nose, S. A molecular dynamics method for simulations in the canonical ensemble. *Mol. Phys.* **1984**, *52*, 255–268. [[CrossRef](#)]
50. Hoover, W.G. Canonical dynamics: Equilibrium phase-space distributions. *Phys. Rev. A* **1985**, *31*, 1695. [[CrossRef](#)]
51. Parrinello, M.; Rahman, A. Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys.* **1981**, *52*, 7182. [[CrossRef](#)]
52. Nose, S.; Klein, M.L. Constant pressure molecular dynamics for molecular systems. *Mol. Phys.* **1983**, *50*, 1055. [[CrossRef](#)]
53. Hess, B.; Bekker, H.; Berendsen, H.J.C.; Fraaije, J.G.E.M. LINCS: A Linear Constraint Solver for molecular simulations. *J. Comput. Chem.* **1997**, *18*, 1463–1472. [[CrossRef](#)]
54. Cuesta-Seijo, J.A.; Garcia-Granda, S. Trypsin as a model for high resolution x-ray diffraction in proteins. *Bol. R. Soc. Hist. Nat. Sec. Geol.* **2002**, *97*, 123–129.
55. Walter, J.; Steigemann, W.; Singh, T.P.; Bartunik, H.; Bode, W.; Huber, R. On the disordered activation domain in trypsinogen: Chemical labelling and low-temperature crystallography. *Acta Crystallogr. Sect. B* **1982**, 1462–1472. [[CrossRef](#)]
56. Yang, L.; Eyal, E.; Chennubhotla, C.; Jee, J.; Gronenborn, A.M.; Bahar, I. Insights into equilibrium dynamics of proteins from comparison of NMR and X-ray data with computational predictions. *Structure* **2007**, *15*, 741–749. [[CrossRef](#)] [[PubMed](#)]
57. Amadei, A.; Linssen, A.B.; Berendsen, H.J. Essential dynamics of proteins. *Proteins* **1993**, *17*, 412–425. [[CrossRef](#)] [[PubMed](#)]

58. Roe, D.R.; Cheatham, T.E. PTRAJ and CPPTRAJ: Software for processing and analysis of molecular dynamics trajectory data. *J. Chem. Theory Comput.* **2013**, *9*, 3084–3095. [[CrossRef](#)]
59. Pedregosa, F.; Varoquax, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.
60. Sokal, R.R.; Michener, C.D. A statistical method for evaluating systematic relationships. *Univ. Kans Sci. Bull.* **1958**, *38*, 1409–1438.
61. Virtanen, P.; Gommers, R.; Oliphant, T.E.; Haberland, M.; Reddy, T.; Cournapeau, D.; Burovski, E.; Peterson, P.; Weckesser, W.; Bright, J.; et al. SciPy 1.0: Fundamental algorithms for scientific computing in Python. *Nat. Methods* **2020**, *17*, 261–272. [[CrossRef](#)]
62. Harris, C.R.; Millman, K.J.; van der Walt, S.J.; Gommers, R.; Virtanen, P.; Cournapeau, D.; Wieser, E.; Taylor, J.; Berg, S.; Smith, N.J.; et al. Array programming with NumPy. *Nature* **2020**, *585*, 357–362. [[CrossRef](#)] [[PubMed](#)]
63. Brown, D.K.; Penkler, D.L.; Amamuddy, O.S.; Ross, C.; Atilgan, A.R.; Atilgan, C.; Bishop, Ö.T. MD-TASK: A software suite for analyzing molecular dynamics trajectories. *Bioinformatics* **2017**, *33*, 2768–2771. [[CrossRef](#)]
64. McGibbon, R.T.; Beauchamp, K.A.; Harrigan, M.P.; Klein, C.; Swails, J.M.; Hernández, C.X.; Schwantes, C.R.; Wang, L.-P.; Lane, T.J.; Pande, V.S. MDTraj: A Modern Open Library for the Analysis of Molecular Dynamics Trajectories. *Biophys. J.* **2015**, *109*, 1528–1532. [[CrossRef](#)]
65. Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual molecular dynamics. *J. Mol. Graph.* **1996**, *14*, 33–38. [[CrossRef](#)]
66. Hunter, J.D. Matplotlib: A 2D Graphics Environment. *Comput. Sci. Eng.* **2007**, *9*, 90–95. [[CrossRef](#)]
67. Gohara, D.W.; Di Cera, E. Allostery in trypsin-like proteases suggests new therapeutic strategies. *Trends Biotechnol.* **2011**, *29*, 577–585. [[CrossRef](#)]
68. Sievers, F.; Wilm, A.; Dineen, D.; Gibson, T.J.; Karplus, K.; Li, W.; Lopez, R.; McWilliam, H.; Remmert, M.; Söding, J.; et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **2011**, *7*, 539. [[CrossRef](#)]
69. Beeler, D.L.; Aird, W.C.; Grant, M.A. Evolutionary conservation of the allosteric activation of factor VIIa by tissue factor in lamprey. *J. Thromb. Haemost.* **2018**, *16*, 734–748. [[CrossRef](#)]
70. Madsen, J.J.; Persson, E.; Olsen, O.H. Evolutionary conservation of the allosteric activation of factor VIIa by tissue factor in lamprey: Comment. *J. Thromb. Haemost.* **2018**, *16*, 1450–1454. [[CrossRef](#)]
71. Beeler, D.L.; Aird, W.C.; Grant, M.A. Evolutionary conservation of the allosteric activation of factor VIIa by tissue factor in lamprey: Reply. *J. Thromb. Haemost.* **2018**, *16*, 1454–1456. [[CrossRef](#)]
72. Hedstrom, L.; Lin, T.Y.; Fast, W. Hydrophobic interactions control zymogen activation in the trypsin family of serine proteases. *Biochemistry* **1996**, *35*, 4515–4523. [[CrossRef](#)] [[PubMed](#)]
73. Pasternak, A.; White, A.; Jeffery, C.J.; Medina, N.; Cahoon, M.; Ringe, D.; Hedstrom, L. The energetic cost of induced fit catalysis: Crystal structures of trypsinogen mutants with enhanced activity and inhibitor affinity. *Protein Sci.* **2001**, *10*, 1331–1342. [[CrossRef](#)]
74. Burnley, B.T.; Afonine, P.V.; Adams, P.D.; Gros, P. Modelling dynamics in protein crystal structures by ensemble refinement. *Elife* **2012**, *2012*, 1–29. [[CrossRef](#)]
75. Burnley, B.T.; Gros, P. phenix. ensemble_refinement: A test study of apo and holo BACE1. *Comput. Crystallogr. Newsl.* **2013**, *4*, 51–58.
76. Sorensen, A.B.; Madsen, J.J.; Frimurer, T.M.; Overgaard, M.T.; Gandhi, P.S.; Persson, E.; Olsen, O.H. Allostery in Coagulation Factor VIIa Revealed by Ensemble Refinement of Crystallographic Structures. *Biophys. J.* **2019**, *116*, 1823–1835. [[CrossRef](#)]
77. Huntington, J.A.; Esmon, C.T. The molecular basis of thrombin allostery revealed by a 1.8 Å structure of the “slow” form. *Structure* **2003**, *11*, 469–479. [[CrossRef](#)]
78. Bah, A.; Garvey, L.C.; Ge, J.; Di Cera, E. Rapid kinetics of Na⁺ binding to thrombin. *J. Biol. Chem.* **2006**, *281*, 40049–40056. [[CrossRef](#)] [[PubMed](#)]
79. Gandhi, P.S.; Chen, Z.; Mathews, F.S.; Di Cera, E. Structural identification of the pathway of long-range communication in an allosteric enzyme. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 1832–1837. [[CrossRef](#)] [[PubMed](#)]
80. Vogt, A.D.; Bah, A.; Di Cera, E. Evidence of the E*-E equilibrium from rapid kinetics of Na⁺ binding to activated protein C and factor Xa*. *J. Phys. Chem. B* **2010**, *114*, 16125–16130. [[CrossRef](#)]
81. Niu, W.; Chen, Z.; Gandhi, P.S.; Vogt, A.D.; Pozzi, N.; Pelc, L.A.; Zapata, F.; Di Cera, E. Crystallographic and kinetic evidence of allostery in a trypsin-like protease. *Biochemistry* **2011**, *50*, 6301–6307. [[CrossRef](#)]
82. Vogt, A.D.; Chakraborty, P.; Di Cera, E. Kinetic Dissection of the Pre-existing Conformational Equilibrium in the Trypsin Fold. *J. Biol. Chem.* **2015**, *290*, 22435–22445. [[CrossRef](#)] [[PubMed](#)]
83. Nagendra, H.G.; Sukumar, N.; Vijayan, M. Role of water in plasticity, stability, and action of proteins: The crystal structures of lysozyme at very low levels of hydration. *Proteins* **1998**, *32*, 229–240. [[CrossRef](#)]
84. Huse, M.; Kuriyan, J. The conformational plasticity of protein kinases. *Cell* **2002**, *109*, 275–282. [[CrossRef](#)]
85. Pochapsky, T.C.; Kazanis, S.; Dang, M. Conformational Plasticity and Structure = Function. *Antioxid. Redox Signal.* **2010**, *13*, 1273–1296. [[CrossRef](#)]
86. Tong, M.; Seeliger, M.A. Targeting conformational plasticity of protein kinases. *ACS Chem. Biol.* **2015**, *10*, 190–200. [[CrossRef](#)]
87. Spoerri, P.M.; Sapra, K.T.; Zhang, C.; Mari, S.A.; Kato, H.E.; Kobilka, B.K.; Müller, D.J. Conformational Plasticity of Human Protease-Activated Receptor 1 upon Antagonist- and Agonist-Binding. *Structure* **2019**, *27*, 1517–1526. [[CrossRef](#)]
88. Crean, R.M.; Gardner, J.M.; Kamerlin, S.C.L. Harnessing Conformational Plasticity to Generate Designer Enzymes. *J. Am. Chem. Soc.* **2020**, *142*, 11324–11342. [[CrossRef](#)]

89. Csermely, P. Plasticity-rigidity cycles: A general adaptation mechanism. *arXiv* **2015**, arXiv:1511.01239.
90. Martín-García, F.; Papaleo, E.; Gomez-Puertas, P.; Boomsma, W.; Lindorff-Larsen, K. Comparing Molecular Dynamics Force Fields in the Essential Subspace. *PLoS ONE* **2015**, *10*, e0121114. [[CrossRef](#)] [[PubMed](#)]
91. Gasper, P.M.; Fuglestad, B.; Komives, E.A.; Markwick, P.R.L.; McCammon, J.A. Allosteric networks in thrombin distinguish procoagulant vs. anticoagulant activities. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 21216–21222. [[CrossRef](#)]
92. Fuglestad, B.; Gasper, P.M.; McCammon, J.A.; Markwick, P.R.L.; Komives, E.A. Correlated motions and residual frustration in thrombin. *J. Phys. Chem. B* **2013**, *117*, 12857–12863. [[CrossRef](#)]
93. Higashi, S.; Nishimura, H.; Aita, K.; Iwanaga, S. Identification of regions of bovine factor VII essential for binding to tissue factor. *J. Biol. Chem.* **1994**, *269*, 18891–18898. [[CrossRef](#)]
94. Hummer, G.; Rasaiah, J.C.; Noworyta, J.P. Water conduction through the hydrophobic channel of a carbon nanotube. *Nature* **2001**, *414*, 188–190. [[CrossRef](#)]
95. Magdziarz, T.; Mitusińska, K.; Gołdowska, S.; Płuciennik, A.; Stolarczyk, M.; Lugowska, M.; Góra, A. AQUA-DUCT: A ligands tracking tool. *Bioinformatics* **2017**, *33*, 2045–2046. [[CrossRef](#)] [[PubMed](#)]
96. Tjørnelund, H.D.; Madsen, J.J.; Peters, G.H.J. Water-Intake and Water-Molecule Paths to the Active Site of Secretory Phospholipase A2 Studied Using MD Simulations and the Tracking Tool AQUA-DUCT. *J. Phys. Chem. B* **2020**, *124*, 1881–1891. [[CrossRef](#)] [[PubMed](#)]
97. Laio, A.; Parrinello, M. Escaping Free-Energy Minima. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 12562–12566. [[CrossRef](#)]
98. Olsen, O.H.; Nielsen, P.F.; Persson, E. Prevention of beta strand movement into a zymogen-like position does not confer higher activity to coagulation factor VIIa. *Biochemistry* **2004**, *43*, 14096–14103. [[CrossRef](#)]
99. Perera, L.; Pedersen, L.G. A reconsideration of the evidence for structural reorganization in FVII zymogen. *J. Thromb. Haemost.* **2005**, *3*, 1543–1545. [[CrossRef](#)]
100. Jing, H.; Macon, K.J.; Moore, D.; DeLucas, L.J.; Volanakis, J.E.; Narayana, S.V.L. Structural basis of profactor D activation: From a highly flexible zymogen to a novel self-inhibited serine protease, complement factor D. *EMBO J.* **1999**, *18*, 804–814. [[CrossRef](#)]
101. Papagrigoriou, E.; McEwan, P.A.; Walsh, P.N.; Emsley, J. Crystal structure of the factor XI zymogen reveals a pathway for transactivation. *Nat. Struct. Mol. Biol.* **2006**, *13*, 557–558. [[CrossRef](#)] [[PubMed](#)]