PROTEINS
STRUCTURE ■ FUNCTION ■ BIOINFORMATICS

Pepsinogen-like activation intermediate of plasmepsin II revealed by molecular dynamics analysis

Ran Friedman* and Amedeo Caflisch*

Department of Biochemistry, University of Zürich, CH-8057 Zürich, Switzerland

ABSTRACT

Plasmepsins are pharmaceutically relevant aspartic proteases involved in haemoglobin degradation by the malaria causing parasites *Plasmodium* spp. They are translated as inactive proenzymes, with an elongated prosegment. On prosegment cleavage, plasmepsins undergo a series of hitherto unresolved conformational changes before becoming active. Here, the flexibility of plasmepsin and proplasmepsin and the activation process are investigated by multiple explicit water molecular dynamics simulations. The large N-terminal displacement and the interdomain shift from the proenzyme structure to active plasmepsin are promoted by essential dynamics sampling. An intermediate, stabilized by electrostatic interactions between the catalytic dyad and the N-terminus of mature plasmepsin, is observed along all activation trajectories. Notably, the stabilizing interactions in the activation intermediate of plasmepsin are similar to those in the X-ray structure of pepsinogen. In particular, the catalytic aspartates act as hydrogen bond acceptors for the N-terminal amino group and the Ser2 hydroxyl in plasmepsin, and the side chains of Lys36pro and Tyr9 in pepsinogen. The simulation results are used to suggest in vitro experiments to test the conformational transitions involved in the maturation of plasmepsin, and design small-molecule inhibitors.

Proteins 2008; 73:814–827.

Key words: plasmepsin; proplasmepsin; plasmodium; protease inhibitor design; essential dynamics sampling; molecular dynamics simulation; malaria.

INTRODUCTION

Malaria is a fatal mosquito-born disease, affecting 500 million people each year, of which 2.5 millions die.[1] Malaria is caused by parasites of the Plasmodium genus, which infect red blood cells and degrade haemoglobin as a nutritional source. There are four Plasmodia species which cause Malaria in human, of which *P. falciparum* is the most virulent. The existing drug therapy is toxic in the long term and is not always efficient, due to emerging drug resistance.[1,2] For this reason, there are ongoing efforts to find new drug targets in the proteome of *P. falciparum*. Proteins of the plasmepsin family seem promising in this aspect.[3] Plasmepsins are pepsin-like aspartic proteases, which are utilized by the Plasmodia for the initiation of haemoglobin breakdown, and are not found in other genera. Ten different proteins of the plasmepsin family have been identified in the genome of *P. falciparum*. Four of these enzymes are found in an acidic lysosome-like organelle, the *food vacuole*, in which haemoglobin degradation takes place, and are therefore the primary targets for drug development among plasmepsins. Although recent evidence shows that *P. falciparum* can survive without vacuolar plasmepsins in culture medium,[4] many plasmepsin inhibitors are fatal to the parasite.[3] Moreover, current development of anti-malarial drugs is based on the use of several compounds which bind different targets. Accordingly, plasmepsin inhibitors linked to Primaquine, a known anti-malarial drug, were developed recently, and showed high anti-plasmodial activity (IC_{50} of 0.1 μM).[5] It has also been suggested to use plasmepsins as drug targets together with falcipains, which are cysteine proteases unique to Plasmodia.[6]

Plasmepsin II (abbreviated PM; to simplify the notation, the Roman number II is dropped henceforth) is the most studied protein of the plasmepsin family. Its structure was solved by X-ray crystallography in the presence of different inhibitors and in the unbound form. The PM fold is typical of eukaryotic aspartic proteases. It consists of two homologous domains (the N- and C-domains), whose interface forms the binding cleft. The catalytic residues are D34 and D214. The binding site is covered by a β-hairpin structure (the flap) which is flexible to facilitate the attachment of the substrate. The flap is tethered to the vicinity of the catalytic site by a hydrogen bond between the side chains of Y77 and W41. A linear chain of hydrogen bonds, involving the binding site, the flap, and two structurally conserved

Additional Supporting Information may be found in the online version of this article.
Grant sponsor: Forschungskredit of the Unviversity of Zürich; Swiss National Science Foundation grant.
*Correspondence to: R. Friedman. E-mail: r.friedman@bioc.uzh.ch or A. Caflisch, Department of Biochemistry, University of Zürich, CH-8057 Zürich, Switzerland. E-mail: caflisch@bioc.uzh.ch
Received 5 December 2007; Revised 27 February 2008; Accepted 3 April 2008
Published online 22 May 2008 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/prot.22105

water molecules (T217–D214–water1–D34–S37–water2–Y77–W41), is present in the crystal structure of the apoenzyme; water 1 may be replaced by an oxygen atom of the inhibitor. This hydrogen-bonding arrangement is characteristic of pepsin-like aspartic proteases.[7] The catalytic dyad is surrounded by a symmetric hydrogen-bonding configuration, termed "the fireman's grip," which is also typical of aspartic proteases. The two aspartates, flap residues V78 and S79 and residues Y192 and S218 are hydrogen bonded to the inhibitors, as shown in crystal structures of PM.[8,9] There is a small variation in the size of the substrate binding cavity,[9] displaying the ability of the protease to accommodate itself to the target.

A different arrangement of the binding site was observed when PM was crystallized in the presence of a potent achiral inhibitor which binds it in the ratio of 2:1 (PDB: 2BJU,[10]) and in related structures (PDB: 2IGX, 2IGY,[11]). In this conformation, the flap is open and does not cap the binding site directly, and the W41–Y77 side chain hydrogen bond is missing. While the flap is more commonly found in a closed conformation, structures with an open flap were detected in another pepsin-like aspartic protease, β-secretase (BACE),[12,13] and in a retroviral aspartic protease.[14] In both cases, flap flexibility was also evident in molecular dynamics (MD) simulations.[15–19]

PM is translated as an inactive proenzyme, which is activated by cleavage of its 124-residue long prosegment. The conformational change of PM upon cleavage is hitherto unknown. The activation is not blocked by known aspartic protease inhibitors *in vivo*, suggesting the involvement of another enzyme rather than an autocatalytic mechanism.[20] *In vitro*, on the other hand, the enzyme undergoes autocleavage 12 residues upstream of the wild type N terminus at acidic pH (pH $\leq$ 4.7),[21] and the resulting protein is enzymatically active. The structure of truncated proPM (with residues 79pro-121pro) was solved by X-ray crystallography (PDB: 1PFZ,[22]). This conformation is inactive, as the two catalytic aspartates are too far apart to perform the water-assisted catalytic mechanism. Such conformation is unique to plasmepsins. In pepsinogen, the active site is preformed and is very similar to that of the active pepsin. Pepsinogen is inactive due to an interaction between a positive residue (K36pro) and the active site.[23] The mechanism of activation, as suggested for pepsin, cannot be applied to PM due to the difference in the structures of the proenzymes.

Because of the intrinsic importance of PM and the availability of crystal structures, several simulation studies have been published recently. Åqvist and coworkers have used molecular dynamics (MD) simulations to test the binding mode of PM inhibitors, estimate their free energy of binding,[24,25] and calculate the free energy profile for the catalytic mechanism.[26] We have recently reported explicit solvent MD simulations of complexed and uncomplexed PM, proPM and proPM cleaved of its prosegment, where the protonation state of the catalytic aspartates was assessed.[27] These simulations showed that the protonation of D214 (but not of D34) favors a stable active site structure.

To shed light on the conformational transition of the activation process, two MD simulation techniques are used in this study. First, multiple explicit water MD simulations are carried out to study the plasticity of proPM as well as complexed and uncomplexed PM. Uncomplexed PM is simulated with open and closed flap and with the H318Q substitution, which was identified in a genetic study of plasmepsins.[28] ProPM is also investigated in the absence of its prosegment and an N-domain shift is observed within tens of nanoseconds. Second, the maturation from an inactive, cleaved proenzyme to an active protease is studied by Essential Dynamics Sampling (EDS),[29,30] a simulation technique that effectively accelerates conformational transitions by selecting only those MD steps which do not increase the root mean square deviation (RMSD) between the actual conformation (cleaved proPM at the start of the EDS runs) and the target structure (mature PM). No deterministic force is added to the force field energy gradient in EDS (unlike other biased MD simulation techniques, for example, targeted[31,32] or steered MD[33]). Instead, when the target structure is not approached by a regular MD step, the coordinates and velocities are projected onto the essential subspace, extracted from an unbiased simulation of the target. As a consequence, it is guaranteed that the RMSD between the simulated and target structures will never increase but (unlike targeted MD) it is not guaranteed to decrease.

An essentially unique path for the conversion from proPM to PM is observed in *all* EDS simulations. Remarkably, the same intermediate structure, with the catalytic dyad stabilized by electrostatic interactions with the N-terminus, is identified along all EDS trajectories. The overall structure of the intermediate, as well as the salt bridge between the N-terminus and the catalytic dyad are preserved in unbiased MD simulations. Interestingly, the interactions stabilizing the catalytic site of the activation intermediate of PM are similar to those present in the active site of pepsinogen.

MATERIALS AND METHODS

Molecular dynamics simulations

MD simulations of PM were carried out using six different structures as starting points (Table I): (1) PM with a closed flap, attached to the binding site through a hydrogen bond between the side chains of W41 and Y77,[34] (2) PM with the single amino acid substitution H318Q, which was identified in a genetic study of plasmepsins,[28] (3) PM with the flap region in an unusual

Table I
Summary of the MD Simulations Performed

PDB code	Name of simulation	Number of simulations[a]	Duration (ns)	Maximal backbone RMSD (nm)	Number of clusters[b]
1LF4	unliganded PM	2	20	0.25	5
1LF4	H318Q mutant of PM	1	20	0.31	3
2BJU	open flap PM	4	20	0.28	5
1LF2	PM / inhibitor complex	4	10, 10, 20, 20	0.28	4
1PFZ	proPM[c]	2	40	0.24	3
1PFZ	cleaved proPM	3	20, 20, 88	0.27	4

This table lists only the conventional MD runs started from the X-ray structures. EDS runs and MD simulations started from snapshots collected in EDS are listed in Table II.
[a]Duplicate runs were started using different random distributions of the initial velocities.
[b]One thousand structures, collected every 0.01 ns during the final 10 ns of the simulations, were clustered according to their structural similarities (see Methods). Clusters with less than 10 members were excluded. The number of clusters is an average over all simulations performed for a given system.
[c]Simulations of the proenzyme were performed with both catalytic aspartates negatively charged, while other runs were carried out with protonated D214, according to a previous study.[27]

open conformation,[10] simulated without an inhibitor, (4) proPM,[22] (5) proenzyme upon removal of the N-terminal prosegment (cleaved proPM), and (6) PM complexed with an inhibitor containing a hydroxypropylamine core.[9] The total simulation time, summed over all runs in Tables I and II, is 0.5 μs. The simulations were performed by use of the Gromacs program,[35,36] version 3.3.1, with the OPLS-AA force field.[37] Simulations of uncomplexed PM with open and closed flap conformations, inhibitor-bound PM (with a transition state analogue), proPM and cleaved proPM were carried out as reported in.[27] D214 was protonated in the simulations of PM and cleaved proPM, while both catalytic aspartates were negatively charged in the simulations of proPM. Lys, Arg, His, Glu, and all other Asp residues were simulated in their standard protonation states. All simulations were run in explicit solvent using periodic boundary conditions and particle mesh Ewald, and at constant temperature (300 K) and pressure (1 bar). Na^+ and Cl^- ions were added randomly by replacing non-crystallographic water molecules in order to neutralize the charge of the system and maintain a salt concentration of 0.1M.

Essential dynamics sampling

Here, EDS[29,30] was used to speed up the conformational transition from cleaved proPM to mature PM. Because of the time-scale of this process, it cannot be studied by conventional explicit solvent MD simulations, which are limited to 10–100 ns (the time-scale of the conformational changes upon cleavage of proPM is not known, but auto-activation of pepsinogen is in the order of 10^1 s).[38] To this end, we employed EDS in "contraction mode",[39] where the algorithm is implemented as follows:

1. Perform a regular MD step and calculate the heavy atom RMSD between the simulated structure and the reference (target) structure, $d_{\text{current,target}}$.
2. If $d_{\text{current,target}} \leq d_{\text{previous,target}}$, accept the step.
3. Otherwise, the coordinates and velocities are projected radially into the chosen subspace (ED subspace from a run of the target) centered on the target structure, with a contracting radius equal to $d_{\text{previous,target}}$, that is, the RMSD to the target will still be equal to $d_{\text{previous,target}}$ (for details, see[39]).

Five EDS simulations were started from three structures extracted from the three runs of cleaved proPM (Table II). Backbone heavy atom RMSD between these structures were between 0.15 nm and 0.23nm. Four of the five EDS runs used a target mature PM after MD equilibration (i.e., at production time $t = 0$). The fifth EDS run was carried out using a different target, taken from a second simulation of mature PM. The essential subspace[40] was calculated over the last 10 ns of the sim-

Table II
Summary of the EDS Runs and MD Simulations Started From EDS Snapshots

Name of simulation	Initial structure[a]	Duration (ns)	Protonation of D214	Backbone RMSD (nm)
EDS 1[b]	Cleaved 1	4	Protonated	0.08[c]
EDS 2	Cleaved 1	2	Protonated	0.15
EDS 3	Cleaved 2	5	Protonated	0.06
EDS 4	Cleaved 2	1	Protonated	0.17
EDS 5	Cleaved 3	4	Protonated	0.07
Int 1	EDS 3	20	Protonated	0.24[d]
Int 2	EDS 3	14	Not protonated	0.19
Int 3	EDS 4	16	Protonated	0.21
Int 4	EDS 4	10	Not protonated	0.18
Int 5	EDS 5	14	Protonated	0.23
Int 6	EDS 5	14	Not protonated	0.27

[a]The term "cleaved" refers to the simulations called "cleaved proPM" in Table I.
[b]Runs EDS 1 and EDS 2 differ in the target structure (extracted from an unliganded PM run), while in runs EDS 3 and EDS 4 different starting structures were used.
[c]RMSD from the target structure, calculated at the last frame of the simulation.
[d]Values of the RMSD from the starting structure, averaged over the last 10 ns of the simulation.

ulations of mature PM, using the g_covar utility in Gromacs, and taking all heavy atoms into account. The 782 eigenvectors with the highest eigenvalues (10% of the total number of eigenvectors) were used for EDS. The motion along these eigenvectors captures 71–74% of the heavy atom mean square fluctuations (MSF) in the *unbiased* simulations of cleaved proPM and only 59–68% in the biased simulations (see Supplementary Table III). This shows that there is a significant contribution from the unbiased modes in EDS. RMSD below 0.2 nm between the reference and target structure were reached in all simulations.

Trial EDS simulations were also carried out by using only the Cα carbons or backbone heavy atoms for the calculation of the essential subspace, but the EDS trajectories could not approach the target structure within few ns. The eigenvalues of the 782 aforementioned eigenvectors sum up to >99.5% of the MSF of all heavy atoms of the protein, while the 987 eigenvalues of the Cα MSF sum up only to 9% (note that it is not possible to compare these values with those given in Supplementary Table III where the eigenvalues were projected on the trajectory). Apparently, the information about the motions of the side chains is crucial for the success of the EDS as performed here.

Molecular dynamics simulations of the activation intermediate

An activation intermediate with an interaction between the catalytic dyad and the N-terminus of mature PM was identified in all five EDS simulations. Activation intermediate conformations were extracted from three EDS runs (see Table II) and used as starting structures for conventional MD simulations. These simulations were performed in duplicates, which differed in the protonation state of D214.

Clustering analysis

To identify the number of unique conformations in the trajectory, structures within each trajectory were clustered together if their structural similarity, as calculated by Cα RMSD, was smaller than 0.1 nm. A structure is added to a cluster when it is located less than a cutoff distance from a member of the cluster, and both structures have at least P common *neighbors*. The neighbors of a structure are the M other structures which are the most similar to it (as calculated by their mutual RMSD). Here, $P = 3$ and $M = 10$. These values reduce the sensitivity of the analysis to the cutoff distance. Clustering analysis was performed on the final 10 ns of the simulation, at 10, ps intervals, that is, 1000 structures were clustered. The Jarvis-Patrick algorithm,[41] as implemented in the g_cluster Gromacs utility was used for clustering. Clusters with less than 10 members were not regarded as relevant for the analysis, as they represent rare conformations.

Structural alignment and graphical presentation

Structural alignment of different proteins (PM and pepsinogen) was performed using the computer program Multiprot.[42] All protein figures were displayed by VMD.[43]

RESULTS

Overall stability and flexibility

Tables I and II list the simulations performed and the values of root mean square deviations (RMSD) from the initial (i.e., X-ray) structure for conventional MD simulations or from the target structure for EDS runs. The low RMSD values indicate that the former are stable while the latter reached the target. Furthermore, there is a good correlation between the Cα atomic fluctuations along the MD simulations and the crystallographic temperature factors (see Supplementary Fig. 1), except for loop residues 237–243, which are involved in crystal contacts in the complex with the inhibitor RS370 (PDB: 1LF2). Further structural analysis, including hydrogen bonds and salt bridges between the prosegment and the mature section of proPM is presented in the Supplementary Material.

Flap opening in the simulations

The flap of PM caps the substrate binding site in the crystal structures of the apo and inhibitor-bound enzymes.[8,34] Moreover, a hydrogen bond between the hydroxyl group of Y77 in the flap and W41 is a part of a linear chain of hydrogen bonds which is necessary to keep the integrity of the binding site.[7] In a previous computational study of β-secretase (BACE), an aspartic protease with a structural similarity to PM, it was shown that the flap tyrosine's side chain can flip its orientation and lose this hydrogen bond.[16] The flipping of the tyrosine makes the flap more mobile, and may be needed for substrate binding. Interestingly, structural evidence shows that some PM inhibitors bind to an open-flap conformation (PDB: 2BJU, 2IGX, 2IGY), where the hydrogen bond between W41 and Y77 is missing.[10,11]

The extent of the opening of the flap during the simulations can be measured by the distance between the nitrogen atom of V78 at the tip of the flap and the catalytic residues (Fig. 1). The corresponding distances in the X-ray structures, are 0.85, 0.96, and 1.40 nm for the apo enzyme, complexed enzyme with closed flap (PDB: 1LF2) and complexed open-flap structure (PDB: 2BJU), respectively (Fig. 1, vertical lines). In the absence of an inhibitor, the flap shows a considerable flexibility. The distance

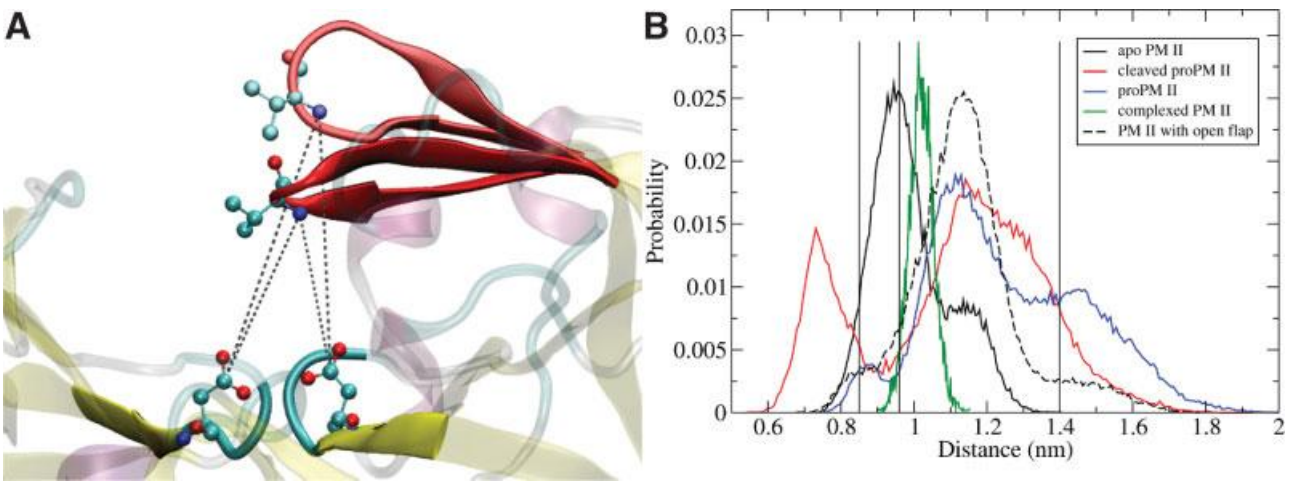

Figure 1

Flap motion. **A:** The distance between the Cγ atoms of the catalytic aspartates and the N atom of V78 at the tip of the flap (red) shown in the closed (PDB: 1LF4) and open (PDB: 2BJU) conformations. **B:** The distribution of the flap opening distance during the simulations. Vertical lines correspond to the distance in the crystal structures of apo PM with closed flap (0.85 nm, PDB:1LF4), PM/inhibitor complex with closed flap (0.96 nm, PDB: 1LF2) and open flap conformation (PDB:2BJU, 1.40 nm). [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

between the flap tip and the catalytic dyad is between 0.73 nm and 1.40 nm in the MD simulations of the apo protein with a maximum at 0.95 nm (Fig. 1, black solid histogram). The distribution is much wider, reaches a value of 1.78 nm and has a maximum at 1.14 nm for the simulations started from the structure with the open flap (Fig. 1, dashed histogram). This maximum is smaller than the 1.4 nm opening of the crystal conformation,[10] because of the presence of two inhibitor molecules in the substrate binding site in the X-ray structure (PDB: 2BJU), but not in the simulations. In the MD runs initiated from the closed flap conformation with a bound inhibitor, the presence of the inhibitor restrains the motions of the flap, and consequently the distribution is much narrower (Fig. 1, green histogram).

The flap is very mobile in the proenzyme, so much that its coordinates could not be resolved in the crystal structure.[22] The lack of electron density for the flap is a strong indicator for the absence of the W41–Y77 side chain hydrogen bond and hence the flap was modeled as open in the simulations of the proenzyme and cleaved proenzyme. During the simulations, the flap is indeed very mobile in the inactive proenzyme (Fig. 1, blue histogram), and the distribution has maxima at 1.11 nm and 1.45 nm, the latter being a value even larger than in the open-flap crystal structure of PM (PDB: 2BJU). Interestingly, when the prosegment is cleaved, the flap can assume a closed conformation (0.73 nm), although it has a higher tendency to be open [maximum of about 1.2 nm, red histogram in Fig. 1(B)] as in the proenzyme. Note that the closed flap conformation in the cleaved proenzyme is different with respect to that of active PM, as the W41–Y77 hydrogen bond is rarely present (not shown).

In computer simulations of BACE the side chain of flap residue Y71 spontaneously flipped its orientation within ~10 ns, causing a rupture of the Trp–Tyr hydrogen bond, which when formed stabilizes the closed conformation.[16] Such flipping of the side chain orientation of Y77 was not observed in the simulations of mature PM, regardless of the initial orientation of Y77 and W41. Interestingly, the hydrogen bond between Y77 and W41 formed spontaneously during one of the two MD runs of proPM and two of the three runs of cleaved proPM. In the latter, this hydrogen bond was stable until the end of the longest run. In the simulation of proPM, the formation of the flap hydrogen bond was transient and reversible, lasting only for 0.2% of the simulation time and hence not leading to flap closure. This indicates that the open flap conformation (without the W41–Y77 hydrogen bond) is the dominant one in the simulations, in contrast to the X-ray structure of *P. vivax* proPM.[44]

As example of spontaneous formation, the side chains of residues W41 and Y77 changed their orientation after 65ns of simulation time of cleaved proPM, forming a stable hydrogen bond (Fig. 2). Initially, the side chain of W41 was involved in a hydrogen bond with S37. After losing this hydrogen bond, the indole ring flipped its orientation and pointed towards Y77 with almost concomitant reorientation of the Y77 phenolic ring to form a stable hydrogen bond. It is not clear whether the reorientation of W41 and Y77 is a prerequisite for the formation of the active site, since the integrity of the active site was maintained in the absence of the W41–Y77 hydrogen

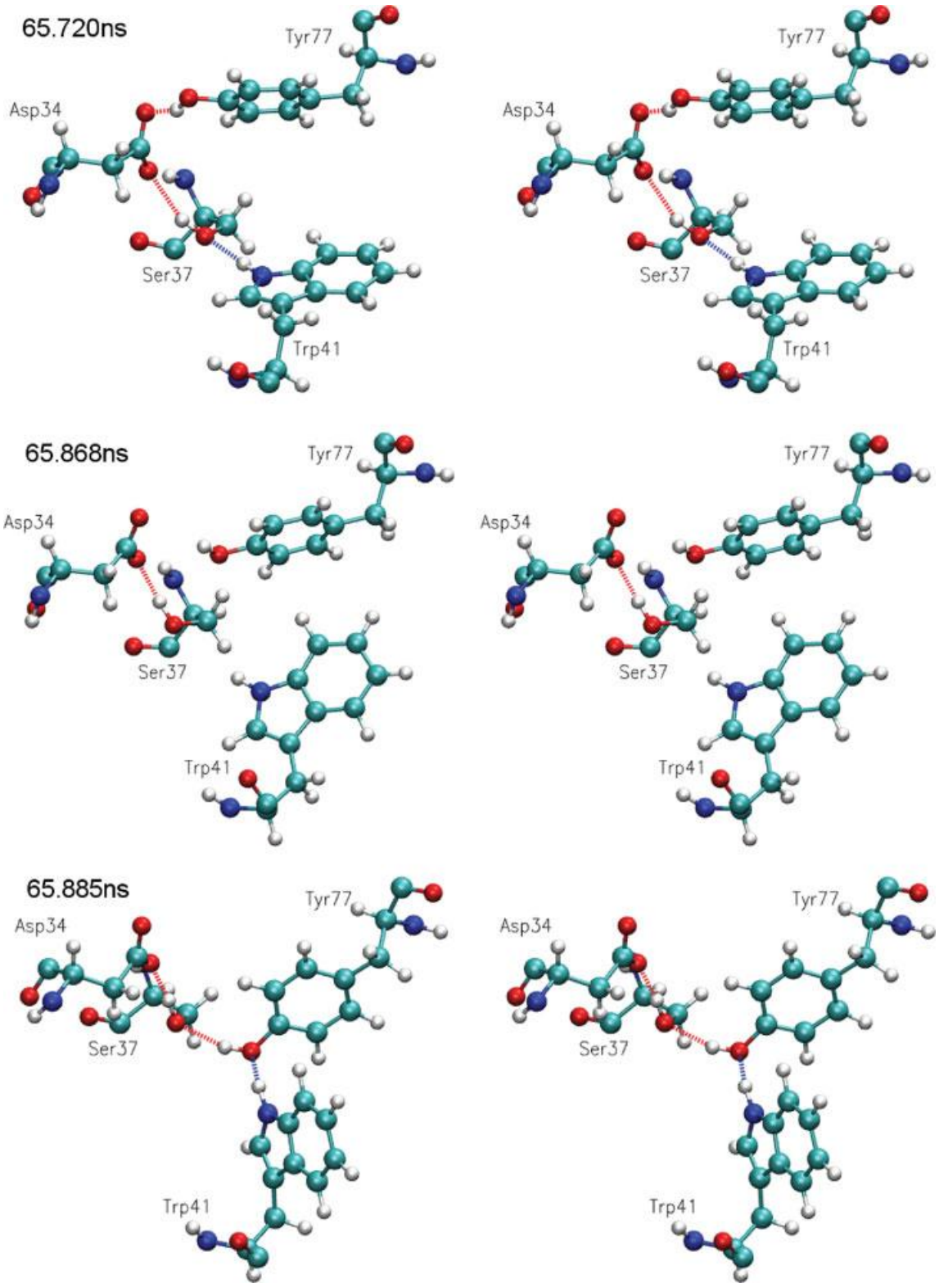

Figure 2

Spontaneous formation of the W41-Y77 side chain hydrogen bond. Stereoview of the formation of the hydrogen bond, which is characteristic of pepsin-like aspartic proteases,[7] during the 88 ns simulation of cleaved proPM. In the uppermost frame, W41 is hydrogen bonded to S37. On rupture of this hydrogen bond, the side chain of W41 flips (middle frame) and acts as a donor in a hydrogen bond to Y77 (bottom), as in the mature enzyme. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

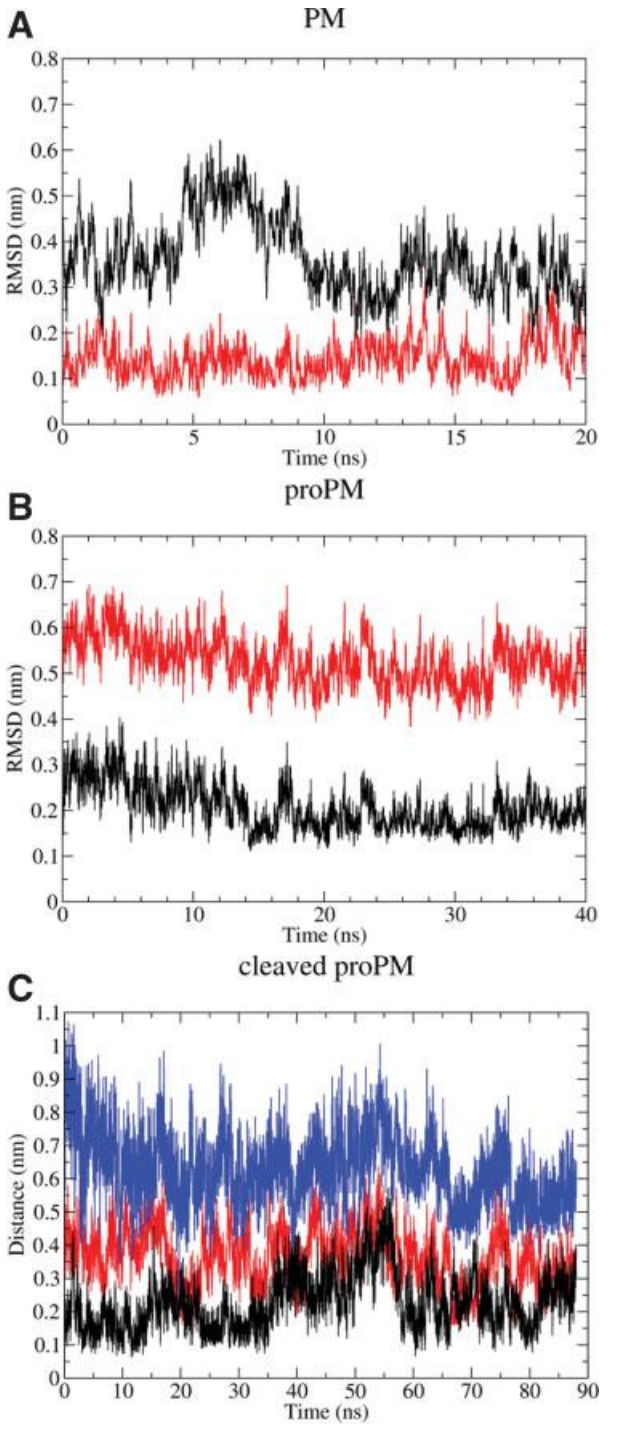

Figure 3

Orientation of the N-domain (residues 31–138). Time series of the Cα RMSD of the regular secondary structure elements of the N-domain of (**A**) PM, (**B**) proPM, and (**C**) cleaved proPM, using as reference structures PM (red) or proPM (black). Snapshots along the simulations were superimposed on the C-domain and central motif (residues 139–329), and RMSD of the N-domain were then calculated. The minimal distance between the Oδs of the catalytic aspartates is shown in blue (only for cleaved proPM). [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

bond.[27] The W41–Y77 side chain hydrogen bond, however, is present in the majority of the crystal structures of PM, including the apo-enzyme, indicating that in the mature protein the closed flap conformation is more likely to have the W41–Y77 side chain hydrogen bond formed.

Structural transitions in the simulations of the cleaved proenzyme

A comparison between the X-ray structures of PM and proPM reveals two main differences: a domain shift in the N-domain (residues 30–129) and a rearrangement of the mature N-terminus (residues 1–14).[22] The domain shift is visible when the structures of the enzyme and the proenzyme are superimposed on their C-domain and central motif. Neither the mechanism nor the time-scale of this conformational transition were elucidated by experiment.

The RMSD to the C-domain and central motif of the two conformations is small in all simulations (not shown). The orientation of the N-domain relative to the C-domain and central motif is maintained throughout the simulations of PM and proPM (Fig. 3). Remarkably, the cleavage of the prosegment, and the protonation of the catalytic dyad, promote an N-domain shift during the 88-ns simulation of cleaved proPM. At the start, the orientation of the N-domain is more similar to that in proPM (Cα RMSD = 0.17nm) than to PM (Cα RMSD = 0.37 nm), see Figure 4. However, as the simulation proceeds, the N-domain shifts its orientation, and approaches the PM conformation, so that at 70 ns the N-domain Cα RMSD from proPM and PM is 0.38 nm and 0.22 nm, respectively (see Figs. 3 and 4). The structural changes encompass the active site as well. The catalytic aspartates approach each other during the three cleaved proPM simulations, and reach the distance of a water-mediated hydrogen bond (5–6 Å), for example, in the 65–70 ns and 75–80 ns intervals of the longest run [Fig. 3(C)].

The integrity of the binding site in the mature enzyme is maintained through a linear chain of hydrogen bonds.[7] This arrangement is present in the simulations of mature PM, but not in those started from the cleaved proenzyme structure. The inability of the cleaved enzyme to fully form the active site as in mature PM during the three MD simulations calls for the application of advanced simulation techniques, as described below.

Simulations of the activation process by essential dynamics sampling

The structural changes on the cleavage of the proenzyme were promoted by EDS using the structure of the mature enzyme as a target. Five different simulations were performed (see Table II). In all runs, the cleaved proenzyme reached RMSD <0.2 nm from its target

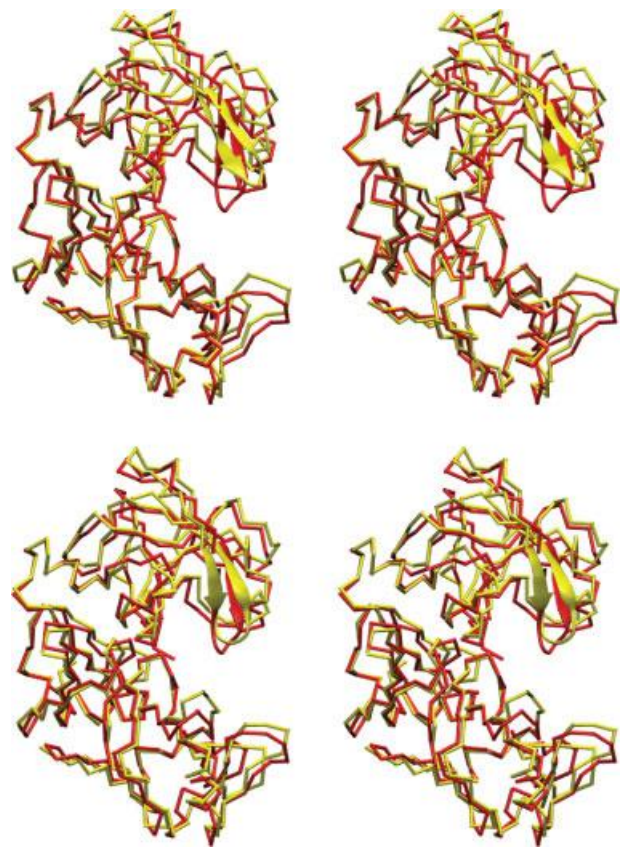

Figure 4

Changes in the structure of cleaved proPM. The structure of cleaved proPM (yellow) superimposed on the structure of PM (red), after energy minimization only (top) and after 70 ns of simulation (bottom) are shown in stereoview. The structures are superimposed on the C-domain and central motif (residues 139–329, which are at the bottom). Note that the 70 ns structure of cleaved proPM is more similar to PM than to proPM (Cα RMSD = 0.38 nm to proPM and 0.22 nm to PM). [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

within less than 2 ns (Fig. 5). It should be reminded that the time-scale of the EDS simulations is much shorter than that of the spontaneous process, as sampling is enhanced in the desired direction to allow the transition to complete within feasible time. Strikingly, although different initial structures (and essential components from two different MD simulations) were used in the simulations, the resulting trajectories were similar, for example, the protein always followed the same pathway from the proenzyme to the mature conformation (Fig. 6). This is shown for two simulations in Figure 7A, which depicts the backbone RMSD values of all structures from one EDS simulation with reference to all of the structures from a second EDS simulation. The RMSD of pairs of structures from similar times are low (~0.1 nm), which indicates that the pathways are similar. Furthermore, the complete transition is observed in all five EDS simulations, for example, conformations similar to the structure of the active enzyme were attained in the five EDS runs. As a counter example, the closed conformation of citrate synthase could not be reached in EDS simulations starting from an open conformation.[45]

The structural transition along the activation process is presented in Figure 6 (see also Supplementary Fig. 2). Before the EDS simulation, the N-terminus is a structureless coil, wiggling around the rest of the protein. At t = 40 ps, the first few residues already come closer to the mature segment of the protein. Interestingly, at 180 ps the N-terminus amino group interacts with the negatively charged binding site, while S2 forms hydrogen bonds with D34 and Y77 of the flap (Fig. 6, at 180 ps; see also the upper frame of Fig. 8). The interaction between the N-terminal residues and the catalytic site is identified in all simulations [Fig. 7(B)] and persists for 300–500 ps. The structure of the intermediate, with respect to the initial (cleaved pro-PM) and target (PM) conformations is shown in Figure 9. Note the different location of the N-terminus in each of these structures. To complete the transition, the N-terminus passes the catalytic site and starts interacting with β strands 4 (residues 153–157) and 5 (residues 166–170). Finally, the N-terminus forms an additional β strand using β strand 5 as a template.

Interactions between the N-terminus and the active site during the activation process

In pepsinogen, the proenzyme is kept inactive due to interactions between a positively charged residue (K36pro in human and porcine pepsinogen) and the catalytic

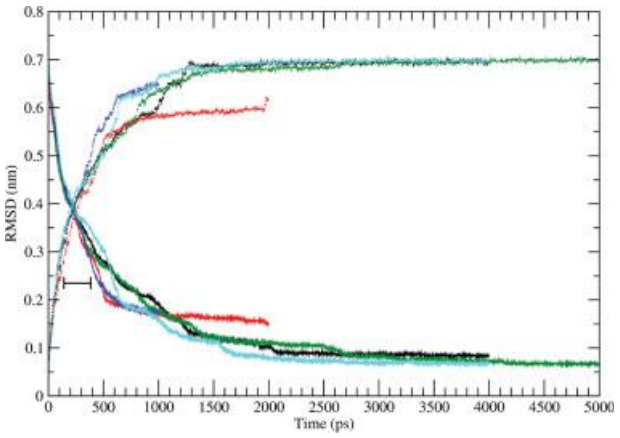

Figure 5

Backbone heavy atoms RMSD from the initial (cleaved proPM, dotted line) and target (mature PM, solid line) structures along the five EDS runs. The different colors refer to different simulations, which also differ in length. The black horizontal bar indicates the time interval during which the N-terminus is in contact with the catalytic dyad (see also Fig. 7). [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

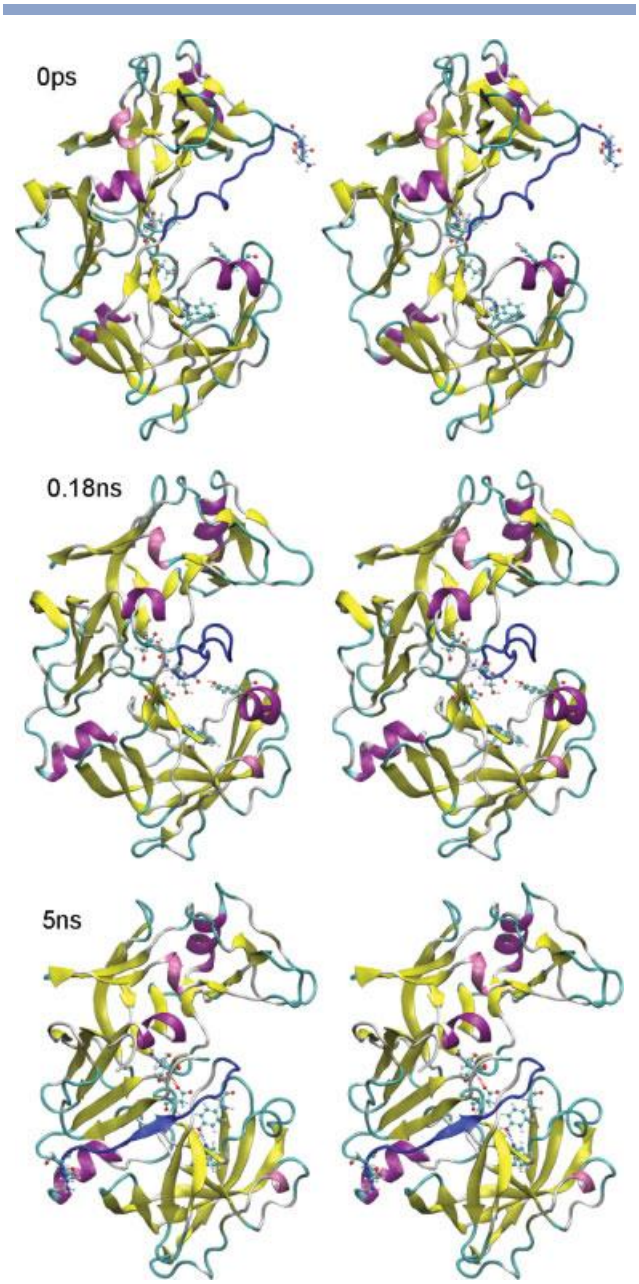

Figure 6

Stereoview of snapshots from one of the EDS simulations of the activation process of PM. The N-terminal segment (residues 1–14) is shown in blue regardless of its secondary structure. Residues 15–329 are color-coded according to their secondary structure (α helices in purple, 3–10 helices in pink, β strands in yellow, tight turns in cyan and coils in white). See also Supplementary Figure 2. Note that the orientation is rotated with respect to Figure 4, so that the N-domain is at the bottom to better visualize the motion of the N-terminal segment. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

aspartates.[23,46] Furthermore, the catalytic residue D32 acts as an acceptor for a hydrogen bond with the phenolic moiety of Y9. Such interactions were not observed in the crystal structure of proPM.[22]

In all EDS simulations, the N-terminus of PM interacts with the catalytic aspartates for up to 0.5 ns, either by a salt bridge with the amino group or by hydrogen bonds involving the side chain (and sporadically also the backbone NH) of S2 [Fig. 7(B)]. To test the stability of these interactions in the absence of any bias, six conventional MD simulations were started from EDS conformations of the protein, with the N-terminus close to the active site [see circles and arrows in Fig. 7(B) and Table II]. In the first three runs, PM was protonated at D214; the minimal distance between the two N-terminal residues (S1, S2) and the Oδ atoms of the catalytic aspartates is shown in Figure 10. In two of the three simulations, the N-terminus is either in a direct contact with the catalytic dyad (black time series in Fig. 10) or is separated by a single water molecule (blue and black time series in Fig. 10). In the third simulation, a transient separation is observed between 6 and 13 ns with contact reformation at 14 ns. It should be mentioned that three EDS simulations were run with protonated D214 because this is the protonation state in the mature protein,[27] but it is reasonable to assume that the protonation takes place during the activation process, after cleavage. For this reason, three additional simulations of the intermediate structure were performed with unprotonated D214. In these three simulations, the N-terminus amino group formed a stable salt bridge with the catalytic pair, and consequently the minimal distances were 0.2 nm or lower (results not shown). Notably, the average backbone RMSD from the initial conformation of all control runs is $\leq$0.27 nm. Such structural stability suggests that the activation intermediate is (meta)stable on time-scales longer than 10 ns. The catalytic site of this intermediate is structurally similar to that of pepsinogen (PDB: 2PSG[47]), as displayed in Figure 8. Note the similarity between the electrostatic interactions of the amino group in the N-terminus of PM or K36pro of pepsinogen with the catalytic dyad. In addition, in both cases one of the catalytic residues accepts a hydrogen bond from a hydroxyl moiety, where the hydrogen bond donors are Y9 and S2 in pepsinogen and PM, respectively. These interactions prevent the binding of a substrate and may also stabilize the unprotonated state of both catalytic aspartates.

DISCUSSION

The activation of pepsin-like aspartic proteases is a complex process. In pepsinogen, this process starts by a conformational change which takes milliseconds[48] and does not involve cleavage. The new conformation is termed *intermediate 1*. A cascade of events leads to the cleavage of the pro–mature junction and the formation of *intermediate 2*, where the prosegment is already cleaved but residues A1pro-F26pro still interact non-co-

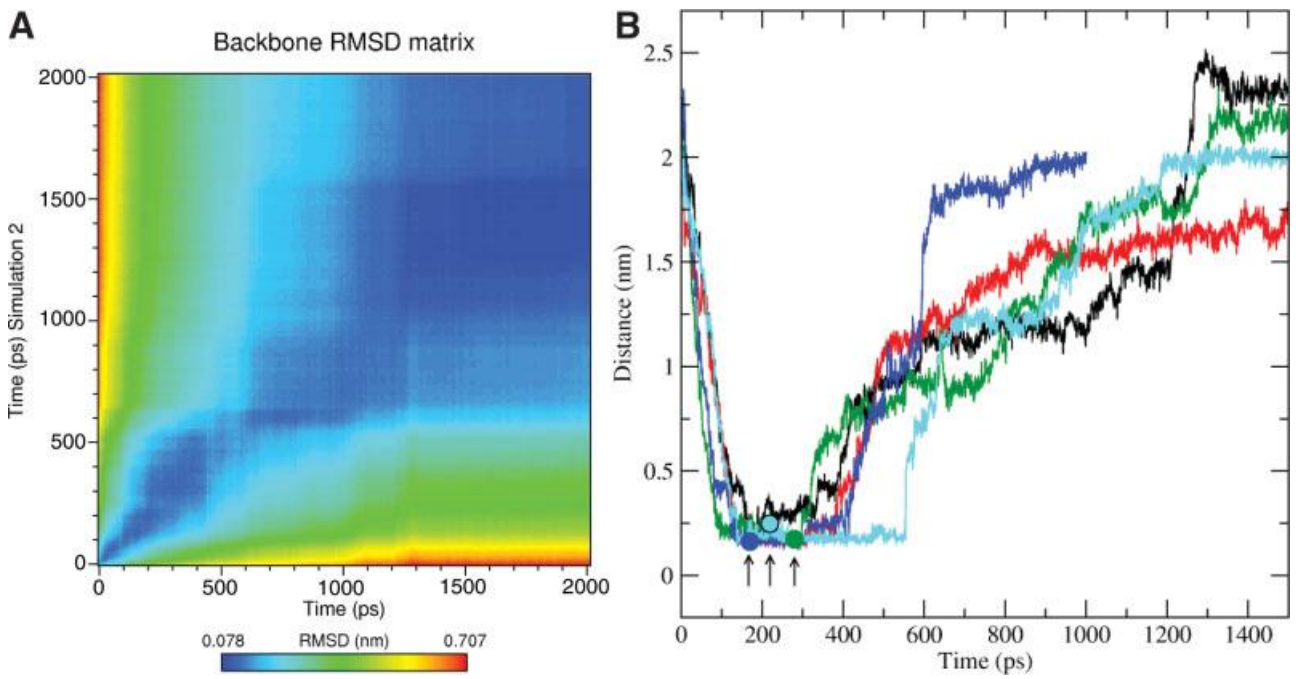

Figure 7

Structural transitions in the EDS simulations. **A:** Backbone RMSD matrix of all snapshots from one EDS simulation relative to all snapshots from a second EDS simulation, displayed for the first 2 ns of the simulations. The blue region between 100 ps and 500 ps is the activation intermediate. **B**: Time series of the minimal distances between S1 and S2 of the N-terminus and the Oδ atoms of D34 and D214 (shown only for t $\leq$1500 ps, truncated after the distances reached a plateau). The different colors refer to different EDS simulations. Starting points for the simulations of the activation intermediate are indicated by circles (see the arrows at the bottom of the figure).

valently with the mature portion.[49] Intermediate 2 can be stabilized by transfer to neutral pH, and its structure was solved by X-ray crystallography.[50]

The kinetics of the activation process of PM are known to a lesser extent. The structure of proPM differs in many aspects from the one of pepsinogen. In proPM, there are no interactions between the active site of the proenzyme and the prosegment.[22] On the contrary, in pepsinogen the prosegment's K36pro side chain is involved in a salt bridge with the catalytic dyad and there is a hydrogen bond between the phenolic side chain of Y9 (as a donor) and D32. Another major difference is that in pepsinogen the catalytic site is preformed (including the flap hydrogen bond between the side chains of W39 and Y75),[47] while in the structure of proPM the two catalytic aspartates are too far apart to support the enzymatic mechanism, and the W41–Y77 side chain hydrogen bond is apparently not present.[22] Interestingly, this hydrogen bond formed spontaneously in two of the three simulations of cleaved proPM.

The simulations of proPM, cleaved proPM and the activation process shed some light on the maturation of PM. In proPM, the prosegment, and mature portion of the protein interact via 22 hydrogen bonds (average in the simulations). The transient protonation of aspartate and glutamate residues at acidic pH is likely to disrupt the hydrogen bond connections between the prosegment's two α helices and the mature portion of the protein, eventually leading to destabilization of the helices.[22] The β strand of the proenzyme is anchored to the body of the protein by six hydrogen bonds, and is not expected to be affected directly by low pH, but will be dispatched due to the release of the helices. Once the prosegment is cleaved, the mature N-terminus can detach from the rest of the protein and moves towards the catalytic site, forming the activation intermediate. This intermediate is observed in all EDS runs and is stable (in the 10–20 ns time scale) in six control simulations started from three different snapshots of the EDS intermediate.

The activation intermediate is not sampled during the unbiased simulations of the cleaved proenzyme suggesting that there is a free energy barrier between the cleaved proenzyme and the intermediate, which cannot be overcome in the timescale of the simulations. In the cleaved proenzyme, the 14 N-terminal residues are not anchored to the body of the protein and their conformational freedom prevents a fast transition. When applying EDS, this mainly entropic barrier is easily overcome, and the intermediate is formed within less than 200 ps. The second

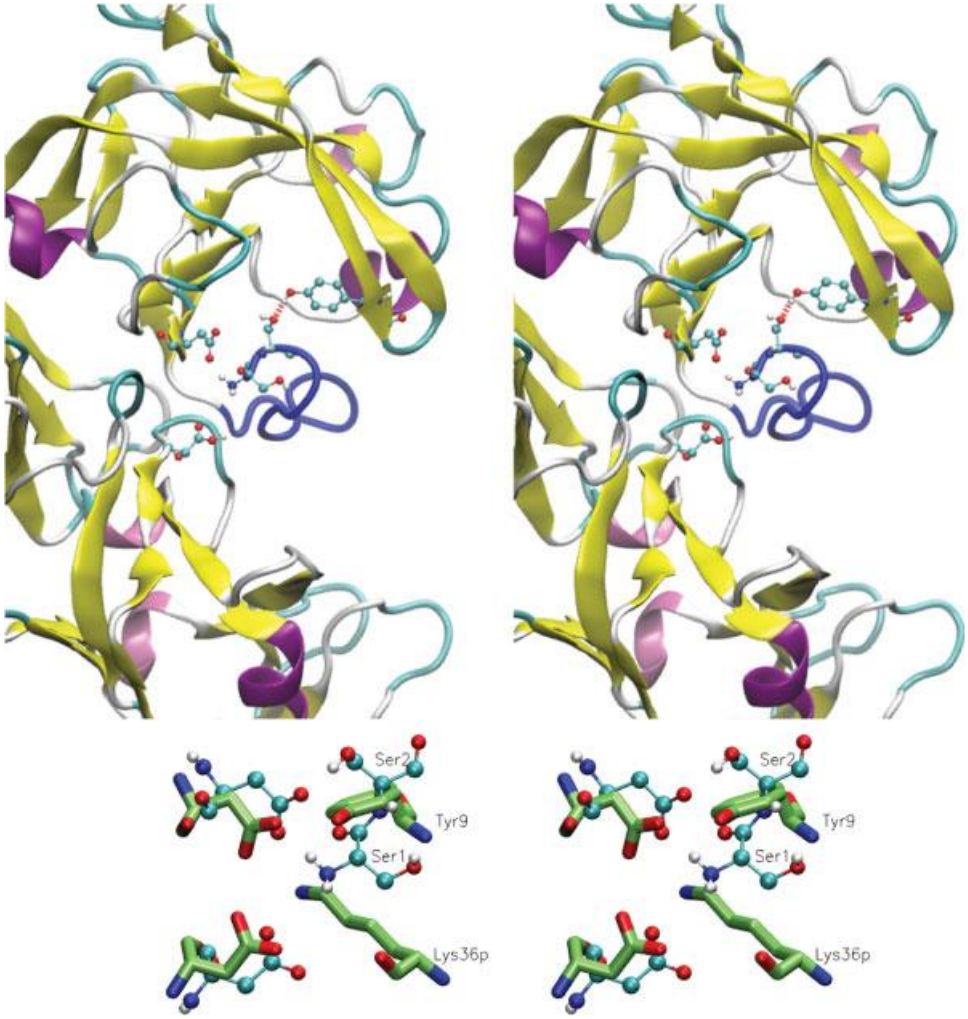

Figure 8

Similarity between the activation intermediate of PM and pepsinogen. (Top) A snapshot of the interaction between Y77 of the flap and S2, while the N-terminal amino group interacts with the catalytic dyad, taken from a simulation of the intermediate with protonated D214. The hydrogen bond between the side chains of Y77 and S2 was identified in about 35% of the simulation time in two MD simulations of the intermediate (with D214 protonated and unprotonated). The protein is colored as in Figure 6. Residues S1, S2, D34, Y77, and D214 are shown in ball and stick representation. (Bottom) Stereoview of the active site interactions in PM and pepsinogen. PM residues are shown in ball and stick representation, while pepsinogen residues are displayed by cylinders. The overlap was obtained by automatic alignment and involved 242 (out of 329) PM residues.

barrier, between the intermediate and the final state involves a significant enthalpic contribution due to the electrostatic interactions between the N-terminus and the catalytic site. This barrier is also not overcome in the control simulations, started from the intermediate conformations sampled by EDS, but it seems to be smaller when one of the catalytic residues is protonated, as suggested for pepsinogen.[50] Protonation of the catalytic dyad weakens the electrostatic attraction, promoting the activation process.

The stability of the activation intermediate in the unbiased simulations and its similarity to pepsinogen indicate that the EDS simulations generate a probable activation trajectory. Yet, more evidence is needed to better elucidate the maturation process. PM activation involves cleavage, conformational changes and protonation of a catalytic residue. This makes it difficult to study the whole process by using bulk experiments, since the kinetics of one process may not be well separated from the kinetics of other processes (e.g., detachment of the N-terminus from the catalytic site and protonation of D214). In principle, single molecule fluorescence resonance energy transfer (FRET) experiments could be used to overcome this kinetic obstacle (separation of timescales), but the currently used fluorophores are not sensitive enough to distinguish between a salt bridge distance

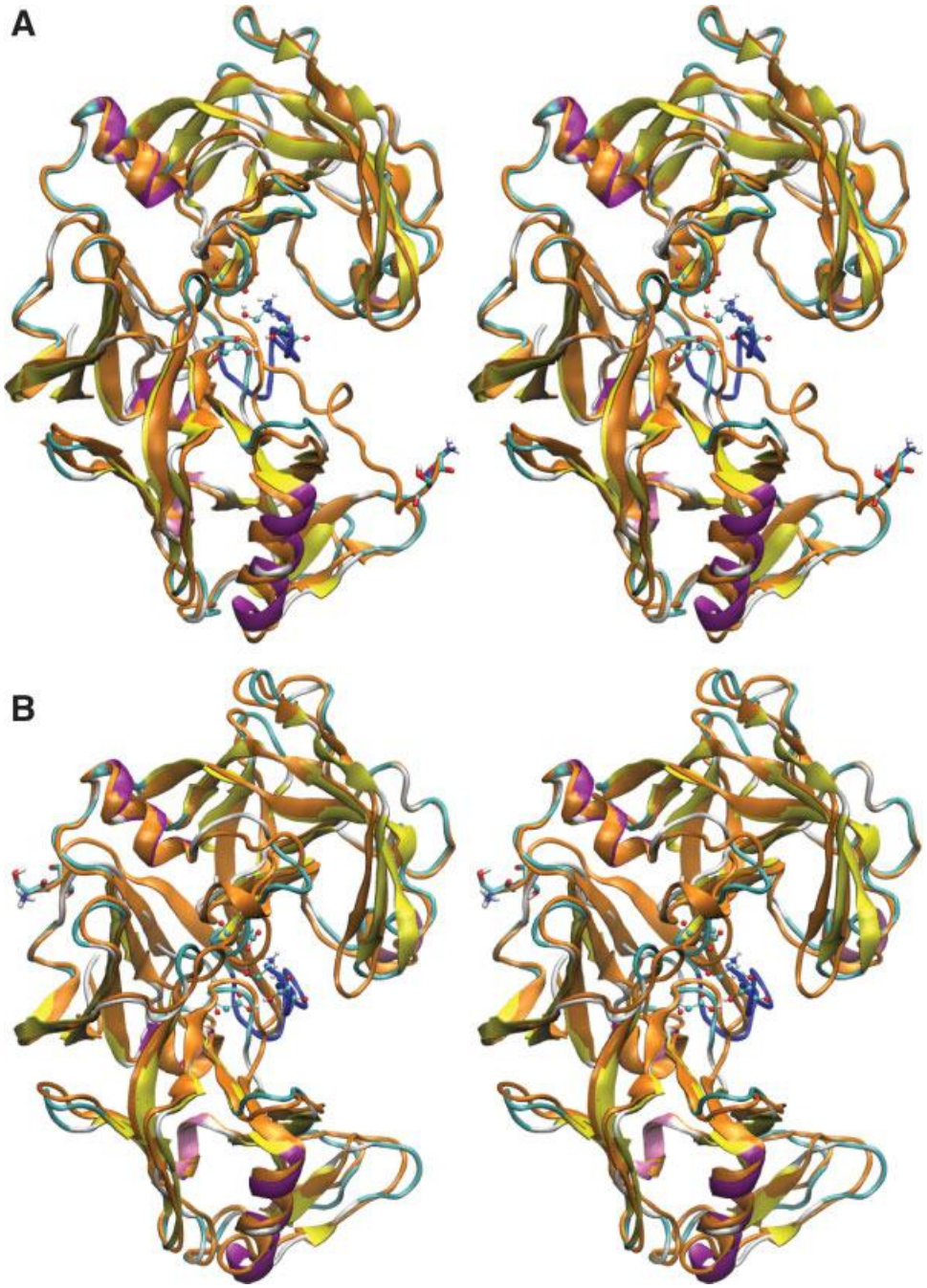

Figure 9

The activation intermediate. Stereoview of the activation intermediate (taken from the simulation EDS 4 at t = 170 ps, that is, the initial conformation for simulations Int 3 and Int 4) structurally aligned (**A**) on the initial conformation, prior to EDS and (**B**) on the target structure, that is, active PM. The activation intermediate is colored as in Figure 6, and the orientation is as in Figure 8. The reference structure is colored orange. Residues S1, S2, D34, and D214 are represented as balls-and-sticks in full color for the intermediate, while the two N-terminal residues of the reference structures are shown by sticks only. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

and an extended conformation of the N-terminus. It may therefore be helpful to use cross-linking experiments followed by mass spectroscopy.[51,52] Interactions between the amino group and carboxylates in the protein could be identified by carbodiimide cross-linkers, which link primary amines to carboxylates.

The activation trajectory and the atomic details of the intermediate are useful for the design of PM inhibitors

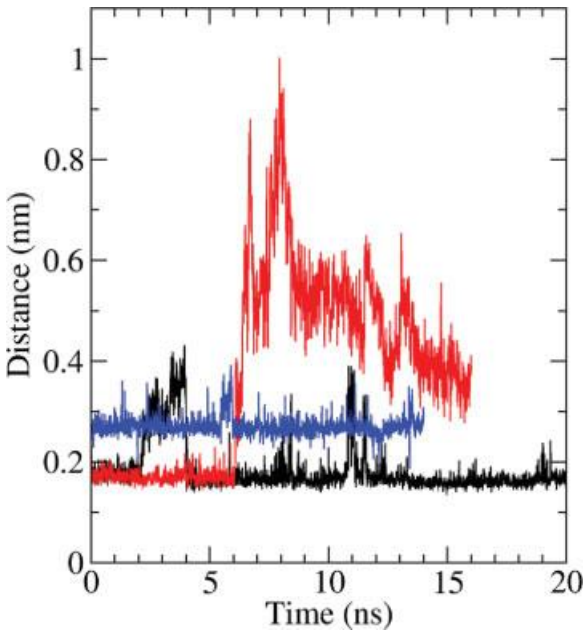

Figure 10

Interactions between the two N-terminal residues and the catalytic dyad. The minimal distance between the two N-terminal residues (S1, S2) and the Oδatoms of the catalytic aspartates is shown for unbiased runs with D214 protonated, started from the activation intermediate obtained by EDS (simulations Int 1, Int 3, and Int 5 in Table II). [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

that bind the premature active site when the catalytic aspartates are already adjacent, but the linear chain of hydrogen bonded residues is not fully formed. It was previously shown that globin and pepstatin can bind an activation intermediate of pepsinogen and act as inhibitors.[53] A possible strategy for the development of novel PM inhibitors involves high-throughput docking. In this approach, a large library of compounds could be screened against multiple conformations of PM with different values of distance and relative orientation of the catalytic aspartates, sampled along the activation trajectory. Inspired by the stable arrangement of the N-terminal serine, which is in close contact with the catalytic dyad in the activation intermediate, a more focused strategy could involve the docking of a tailored library consisting of compounds with a primary (or secondary) amino group and a hydroxyl group. The stability of the activation intermediate and the similarity between the X-ray structures of *P. falciparum* proPM[22] and *P. vivax* proPM[44] suggest that inhibitors of the activation process of PM might be useful to fight plasmodia in general.

ACKNOWLEDGMENTS

The authors thank Dr. Ben Schuler for interesting discussions about the experimental validation of the results. The simulations were performed on the Matterhorn cluster of the University of Zurich and we gratefully acknowledge the support of C. Bolliger, T. Steenbock and A. Godknecht.

REFERENCES

1. Linares GEG, Rodriguez JB. Current status and progresses made in malaria chemotherapy. Curr Med Chem 2007;14:289–314.
2. Noranate N, Durand R, Tall A, Marrama L, Spiegel A, Sokhna C, Pradines B, Cojean S, Guillotte M, Bischoff E, Ekala M-T, Bouchier C, Fandeur T, Ariey F, Patarapotikul J, Le Bras J, Trape JF, Rogier C, Mercereau-Puijalon O. Rapid dissemination of *Plasmodium falciparum* drug resistance despite strictly controlled antimalarial use. PLoS ONE 2007;2:e139.
3. Ersmark K, Samuelsson B, Hallberg A. Plasmepsins as potential targets for new antimalarial therapy. Med Res Rev 2006;26:626–666.
4. Bonilla JA, Bonilla TD, Yowell CA, Fujioka H, Dame JB. Critical roles for the digestive vacuole plasmepsins of *Plasmodium falciparum* in vacuolar function. Mol Microbiol 2007;65:64–75.
5. Dell'Agli M, Parapini S, Galli G, Vaiana N, Taramelli D, Sparatore A, Liu P, Dunn BM, Bosisio E, Romeo S. High antiplasmodial activity of novel plasmepsins I and II inhibitors. J Med Chem 2006; 49:7440–7449.
6. Liu J, Istvan ES, Gluzman IY, Gross J, Goldberg DE. *Plasmodium falciparum* ensures its amino acid supply with multiple acquisition pathways and redundant proteolytic enzyme systems. Proc Natl Acad Sci USA 2006;103:8840–8845.
7. Andreeva NS, Rumsh LD. Analysis of crystal structures of aspartic proteinases: on the role of amino acid residues adjacent to the catalytic site of pepsin-like enzymes. Protein Sci 2001;10:2439–2450.
8. Silva AM, Lee AY, Gulnik SV, Maier P, Collins J, Bhat TN, Collins PJ, Cachau RE, Luker KE, Gluzman IY, Francis SE, Oksman A, Goldberg DE, Erickson JW. Structure and inhibition of plasmepsin II: a hemoglobin-degrading enzyme from *Plasmodium falciparum.* Proc Natl Acad Sci USA 1996;93:10034–10039.
9. Asojo OA, Afotina A, Gulnik SV, Yu B, Erickson JW, Randad R, Medjahed D, Silva AM. Structures of Ser205 mutant plasmepsin II from *Plasmodium falciparum* at 1.8 angstrom in complex with the inhibitors rs367 and rs370. Acta Crystallograph D 2002;58:2001–2008.
10. Prade L, Jones AF, Boss C, Richard-Bildstein S, Meyer S, Binkert C, Bur D. X-ray structure of plasmepsin II complexed with a potent achiral inhibitor. J Biol Chem 2005;280:23837–23843.
11. Boss C, Corminboeuf O, Grisostomi C, Meyer S, Jones AF, Prade L, Binkert C, Fischli W, Weller T, Bur D. Achiral, cheap, and potent inhibitors of Plasmepsins I, II, and IV. ChemMedChem 2006;1: 1341–1345.
12. Hong L, Tang J. Flap position of free memapsin 2 (β-secretase): a model for flap opening in aspartic protease catalysis. Biochemistry 2004;43(16):4689–4695.
13. Patel S, Vuillard L, Cleasby A, Murray CW, Yon J. Apo and inhibitor complex structures of BACE (β-secretase). J Mol Biol 2004; 343(2):407–416.
14. Wlodawer A, Gustchina A. Structural and biochemical studies of retroviral proteases. Biochim Biophys Acta 2000;1477(1/2):16–34.
15. Cascella M, Micheletti C, Rothlisberger U, Carloni P. Evolutionarily conserved functional mechanics across pepsin-like and retroviral aspartic proteases. J Am Chem Soc 2005;127:3734–3742.
16. Gorfe AA, Caflisch A. Functional plasticity in the substrate binding site of β-secretase. Structure 2005;13:1487–1498.
17. Carnevale V, Raugei S, Micheletti C, Carloni P. Convergent dynamics in the protease enzymatic superfamily. J Am Chem Soc 2006; 128:9766–9772.

18. Chang CE, Shen T, Trylska J, Tozzini V, McCammon JA. Gated binding of ligands to HIV-1 protease: Brownian dynamics simulations in a coarse-grained model. Biophys J 2006;90:3880–3885.
19. Tozzini V, Trylska J, Chang CE, McCammon JA. Flap opening dynamics in HIV-1 protease explored with a coarse-grained model. J Struct Biol 2007;157:606–615.
20. Francis SE, Banerjee R, Goldberg DE. Biosynthesis and maturation of the malaria aspartic hemoglobinases plasmepsins I and II. J Biol Chem 1997;272:14961–14968.
21. Hill J, Tyas L, Phylip LH, Kay J, Dunn BM, Berry C. High level expression and characterisation of Plasmepsin II: an aspartic proteinase from *Plasmodium falciparum*. FEBS Lett 1994;352:155–158.
22. Khazanovich Bernstein N, Cherney MM, Loetscher H, Ridley RG, James MN. Crystal structure of the novel aspartic proteinase zymogen proplasmepsin II from *Plasmodium falciparum*. Nat Struct Biol 1999;6:32–37.
23. James MN, Sielecki AR. Molecular structure of an aspartic proteinase zymogen, porcine pepsinogen, at 1.8 A resolution. Nature 1986; 319:33–38.
24. Ersmark K, Feierberg I, Bjelic S, Hamelink E, Hackett F, Blackman MJ, Hulten J, Samuelsson B, Åqvist J, Hallberg A. Potent inhibitors of the *Plasmodium falciparum* enzymes plasmepsin I and II devoid of cathepsin D inhibitory activity. J Med Chem 2004;47(1):110–122.
25. Ersmark K, Nervall M, Hamelink E, Janka LK, Clemente JC, Dunn BM, Blackman MJ, Samuelsson B, Åqvist J, Hallberg A. Synthesis of malarial plasmepsin inhibitors and prediction of binding modes by molecular dynamics simulations. J Med Chem 2005;48:6090–6106.
26. Bjelic S, Åqvist J. Catalysis and linear free energy relationships in aspartic proteases. Biochemistry 2006;45:7709–7723.
27. Friedman R, Caflisch A. The protonation state of the catalytic aspartates in plasmepsin II. FEBS Lett 2007;581:4120–4124.
28. Barry AE, Leliwa-Sytek A, Man K, Kasper JM, Hartl DL, Day KP. Variable SNP density in aspartyl-protease genes of the malaria parasite *Plasmodium falciparum*. Gene 2006;376:163–173.
29. Amadei A, Linssen AB, de Groot BL, van Aalten DMF, Berendsen HJC. An efficient method for sampling the essential subspace of proteins. J Biomol Struct Dyn 1996;13:615–625.
30. de Groot BL, Amadei A, van Aalten DM, Berendsen HJ. Toward an exhaustive sampling of the configurational spaces of the two forms of the peptide hormone guanylin. J Biomol Struct Dyn 1996;13: 741–751.
31. Schlitter J, Engels M, Krüger P, Jacoby E, Wollmer A. Targeted molecular dynamics simulation of conformational change: application to the T ↔ R transition in insulin. Mol Simul 1993;10:291–309.
32. Ferrara P, Apostolakis J, Caflisch A. Targeted molecular dynamics simulations of protein unfolding. J Phys Chem 2000;104:4511–4518.
33. Isralewitz B, Gao M, Schulten K. Steered molecular dynamics and mechanical functions of proteins. Curr Opin Struct Biol. 2001;11: 224–230.
34. Asojo OA, Gulnik SV, Afonina E, Yu B, Ellman JA, Haque TS, Silva AM. Novel uncomplexed and complexed structures of plasmepsin II: an aspartic protease from *Plasmodium falciparum*. J Mol Biol 2003;327:173–181.
35. Berendsen HJC, van der Spoel D, Vandrunen R. Gromacs—a message-passing parallel molecular-dynamics implementation. Comput Phys Commun 1995;91:43–56.
36. van der Spoel D, Lindahl E, Hess B, Groenhof G, Mark AE, Berendsen HJC. GROMACS: fast, flexible, and free. J Comput Chem 2005; 26:1701–1718.
37. Jorgensen WL, Maxwell DS, Tirado-Rives J. Development and testing of the OPLS all-atom force Field on conformational energetics and properties of organic liquids. J Am Chem Soc 1996;118:11225–11236.
38. al Janabi J, Hartsuck JA, Tang J. Kinetics and mechanism of pepsinogen activation. J Biol Chem 1972;247:4628–4632.
39. Daidone I, Amadei A, Roccatano D, Nola AD. Molecular dynamics simulation of protein folding by essential dynamics sampling: folding landscape of horse heart cytochrome c. Biophys J 2003;85:2865–2871.
40. Amadei A, Linssen AB, Berendsen HJC. Essential dynamics of proteins. Proteins 1993;17:412–425.
41. Jarvis RA, Patrick EA. Clustering using a similarity measure based on shared nearest Neighbors. IEEE T Comput 1973;C22:1025–1034.
42. Shatsky M, Nussinov R, Wolfson HJ. A method for simultaneous alignment of multiple protein structures. Proteins 2004;56:143–156.
43. Humphrey W, Dalke A, Schulten, K. VMD: visual molecular dynamics. J Mol Graph 1996;14:33–38.
44. Khazanovich Bernstein N, Cherney MM, Yowell CA, Dame JB, James MN. Structural insights into the activation of *P. vivax* plasmepsin. J Mol Biol 2003;329:505–524.
45. Daidone I, Roccatano D, Hayward S. Investigating the accessibility of the closed domain conformation of citrate synthase using essential dynamics sampling. J Mol Biol 2004;339:515–525.
46. Moore SA, Sielecki AR, Chernaia MM, Tarasova NI, James MN. Crystal and molecular structures of human progastricsin at 1.62 A resolution. J Mol Biol 1995;247:466–485.
47. Sielecki AR, Fujinaga M, Read RJ, James MN. Refined structure of porcine pepsinogen at 1.8 A resolution. J Mol Biol 1991;219:671–692.
48. Auer HE, Glick DM. Early events of pepsinogen activation. Biochemistry 1984;23:2735–2739.
49. Foltmann B, Jensen AL. Human progastricsin. Analysis of intermediates during activation into gastricsin and determination of the amino acid sequence of the propart. Eur J Biochem 1982;128:63–70.
50. Khan AR, Cherney MM, Tarasova NI, James MN. Structural characterization of activation 'intermediate 2' on the pathway to human gastricsin. Nat Struct Biol 1997;4:1010–1015.
51. Trester-Zedlitz M, Kamada K, Burley SK, Fenyo D, Chait BT, Muir TW. A modular cross-linking approach for exploring protein interactions. J Am Chem Soc 2003;125:2416–2425.
52. Sinz A. Chemical cross-linking and mass spectrometry to map three-dimensional protein structures and protein-protein interactions. Mass Spectrom Rev 2006;25:663–682.
53. Marciniszyn JJ, Huang JS, Hartsuck JA, Tang J. Mechanism of intramolecular activation of pepsinogen. Evidence for an intermediate delta and the involvement of the active site of pepsin in the intramolecular activation of pepsinogen. J Biol Chem 1976;251:7095–7102.