

Supplementary Material for:
**Pepsinogen-like activation intermediate of plasmeysin II revealed by
molecular dynamics analysis**

Ran Friedman and Amedeo Caffisch
*Department of Biochemistry, University of Zürich, Winterthurerstrasse 190,
CH-8057 Zürich, Switzerland*

The Supplementary Material include:

1. Supplementary section - structural stability and flexibility.
2. Supplementary section - hydrogen bonds and salt bridges between the prosegment and the mature portion of the protein.
3. Supplementary section - EDS simulations.
4. Supplementary figures 1-3

1 Structural stability and flexibility

The backbone heavy-atom root mean square deviations (RMSD) in the simulations are between 0.25 and 0.31 nm (main text, Table I). Clustering analysis (see Methods) identified 3-5 distinct protein structures over the last 10ns of the simulations. No structural or dynamic differences could be attributed to the H318Q mutation, and the simulations of PM and H318Q PM are therefore analyzed together.

The average root mean square fluctuations (RMSF) per residue, as calculated over all heavy atoms, are similar in all simulations. The most mobile residues are located on the termini, the tip of the flap (residues 78 and 79), and two loops which interconnect pairs of β -strands (residues 240, 241 and 278-283), see Supplementary Table 1 and Supplementary Figure 1. The enhanced mobility of these residues is in agreement with the temperature factors and is linked to the protein's function. Residues N78 and V79 are located on the tip of the flap and move upon substrate binding [1, 2], while residues P240 and F241 are part of a very mobile loop, which could not be located in the crystal structure of apo PM [2]. Kinetic evidence by Istvan and Goldberg shows that a mutation of residue F241 reduces the enzyme's k_{cat} for the cleavage of its natural substrate, haemoglobin [3]. The findings of Istvan and Goldberg were explained by interactions of the loop with substrate residues located distal from the active site. The enhanced flexibility of the loop, as evidenced in the simulations, can assist such interactions by allowing the distal substrate residues to accommodate to the protein after binding it, thereby anchoring the substrate to the enzyme and facilitating its cleavage.

The flexibility of proPM, as shown by its temperature factors, is smaller than the one of uncomplexed PM (see Supplementary Figure 1). The structural stability of proPM is also manifested by its low backbone RMSD and small number of structural clusters. The lack of plasticity of proPM confirms the proposed role of the prosegment as a 'harness', which prevents the formation of the mature active site by limiting the motions of the protein, as suggested by James and co-workers based on the structure of the proenzyme [4]. Cleaved proPM is more flexible than proPM. Its N-terminus is very mobile due to the

cleavage of the pro-mature junction, which breaks the pro-mature loop Y122pro–D4. Other very mobile regions of the cleaved proenzyme include the residues 48–52 and the flap tip (residues 77–81).

Table 1: Residues with RMSF > 0.25nm

Simulation	Residues with RMSF > 0.25nm
unliganded PM	50, 241, 278-283
unliganded H318Q mutant of PM	329
open flap PM	1, 241, 278-283
PM / inhibitor complex	79, 240, 241, 278-282, 329
proPM	p115, p116, 78, 329
cleaved proPM	1, 2, 78, 328, 329

2 Hydrogen bonds and salt bridges between the prosegment and the mature portion of the protein

The pro-mature junction in proPM is located on a tight loop between Y122pro and D4. The latter is a key residue in maintaining the conformation of the loop in the crystal structure [4], as it is hydrogen bonded to Y122pro and S1. In addition, D4 anchors the loop to the body of the protein by a salt bridge to K238 and a hydrogen bond to F241. This led to the conjuncture that D4 may play a key role in the cleavage of the proenzyme at low pH, as its protonation may result in disruption of these hydrogen bonds and the salt bridge to K238. In addition, the interactions between the prosegment and the C-domain are maintained by a hydrogen bond between D91pro and H164. This hydrogen bond and salt bridges in the prosegment are expected to be weak or absent at low pH [4]. A detailed analysis of the trajectories reveals that the network of hydrogen bonds that ensures the stability of the loop and anchors the prosegment to the C-domain is more complex than seen in the X-ray structure (Supplementary Table 2, Supplementary Figure 3). In the simulations, Y122pro is hydrogen bonded to D4 or S1, while H164 is hydrogen bonded to E87pro rather than D91pro. This redundancy of hydrogen bonding partners prevents the rapid release of the prosegment and might be necessary to tolerate mutations in the hydrogen bonding residues.

Table 2: Hydrogen bond stability during the simulations of proPM

h-bond	Presence in the X-ray structure?	Occurrence %
E87pro-H164	no	35
D91pro-H164	yes	0
Y122pro-S1	no	21
Y122pro-D4	yes	19
S1-D4 ^a	yes	59
D4-K238	yes	89
D4-F241 ^b	yes	22

Hydrogen bonds are between side chain atoms unless otherwise stated. ^a Backbone hydrogen bond. ^b The side chain of D4 is hydrogen bonded to the backbone of F241.

3 EDS simulations

Table 3: Heavy-atom mean square fluctuations along the EDS input eigenvectors

Simulation	MSF ^a [nm ²]	Filtered MSF ^b [nm ²]	Residual MSF ^c %
Cleaved 1 ^d	25.05	17.86	29
Cleaved 2	33.30	24.31	27
Cleaved 3	35.55	26.23	26
EDS 1	75.16	47.61	37
EDS 2	77.34	52.85	32
EDS 3	68.85	40.85	41
EDS 4	106.53	72.45	32
EDS 5	76.55	48.90	36

This table gives the total MSF, MSF along the 782 EDS input eigenvectors and how much of the total MSF is captured when the motions along the EDS input eigenvectors are excluded. This shows that the residual MSF is larger for the EDS trajectories, i.e. the fluctuations which do not overlap with the input eigenvectors are substantial.

^a Total heavy atom mean square fluctuations were calculated by summing the eigenvalues of the (diagonal) MSF matrix, calculated over all heavy atoms in a trajectory.

^b To calculate the filtered MSF, the eigenvectors, which were used for the EDS simulations, were projected on the trajectory. This gives a new trajectory, where the motions are filtered, i.e., only the motions along the selected eigenvectors are represented. The MSF was then recalculated from the new matrix.

Filtering the trajectory is done as follows. First, a matrix of projections is obtained: $\mathbf{P} = \mathbf{X}\mathbf{T}$, where \mathbf{X} is the trajectory and \mathbf{T} is a matrix of eigenvectors. Then, a filtered trajectory \mathbf{X}^f is obtained as: $\mathbf{X}^f = \mathbf{P}\mathbf{T}^T$, where \mathbf{T}^T is the transpose of \mathbf{T} .

^c The residual trajectory is calculated as $1 - (\text{Filtered MSF}) / \text{MSF}$.

^d MSF were calculated over the last 10ns of the unbiased trajectories, or the whole trajectory for EDS.

References

1. Asojo, O. A, Afotina, A, Gulnik, S. V, Yu, B, Erickson, J. W, Randad, R, Medjahed, D, and Silva, A. M. Structures of Ser205 mutant plasmepsin II from *Plasmodium falciparum* at 1.8 angstrom in complex with the inhibitors rs367 and rs370. *Acta Crystallog. D* 2002; 58;2001–2008.
2. Asojo, O. A, Gulnik, S. V, Afonina, E, Yu, B, Ellman, J. A, Haque, T. S, and Silva, A. M. Novel uncomplexed and complexed structures of plasmepsin II, an aspartic protease from *Plasmodium falciparum*. *J Mol Biol* 2003; 327;173–181.
3. Istvan, E. S and Goldberg, D. E. Distal substrate interactions enhance plasmepsin activity. *J Biol Chem* Feb, 2005; 280(8);6890–6896.
4. Khazanovich Bernstein, N, Cherney, M. M, Loetscher, H, Ridley, R. G, and James, M. N. Crystal structure of the novel aspartic proteinase zymogen proplasmepsin II from *plasmodium falciparum*. *Nat Struct Biol* 1999; 6;32–37.

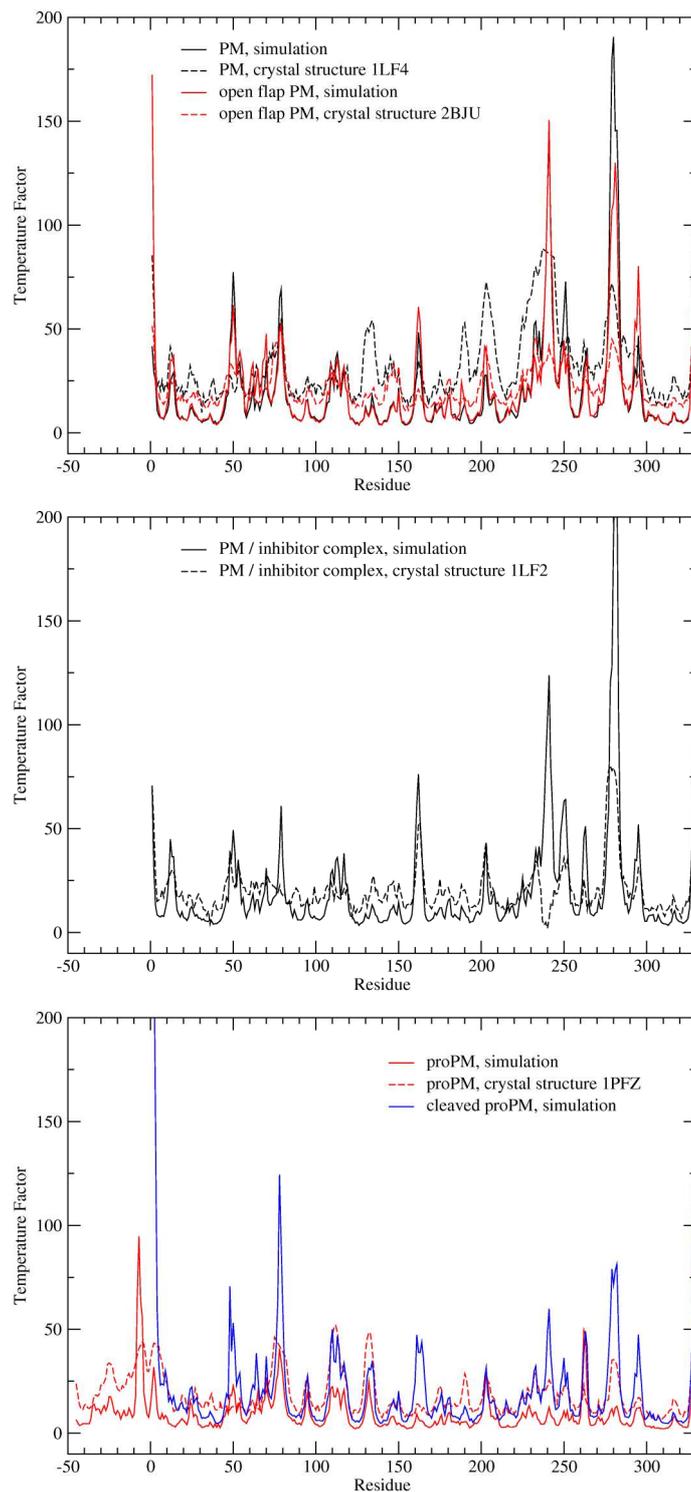


Figure 1: Temperature factors of $C\alpha$ atoms of PM and proPM. The temperature factors (in \AA^2) are computed from the RMSF as $B = 8\pi^2 \text{RMSF}^2 / 3$. The values are averaged over the last 10ns of each simulation and truncated at $B=200$ to keep the scale. The largest discrepancy between experimental temperature factors and those calculated from the RMSF along the MD run is located at loop residues 237-243, which are involved in crystal contacts with a neighbouring molecule in the X-ray structure of PM/inhibitor complex (PDB 1LF2).

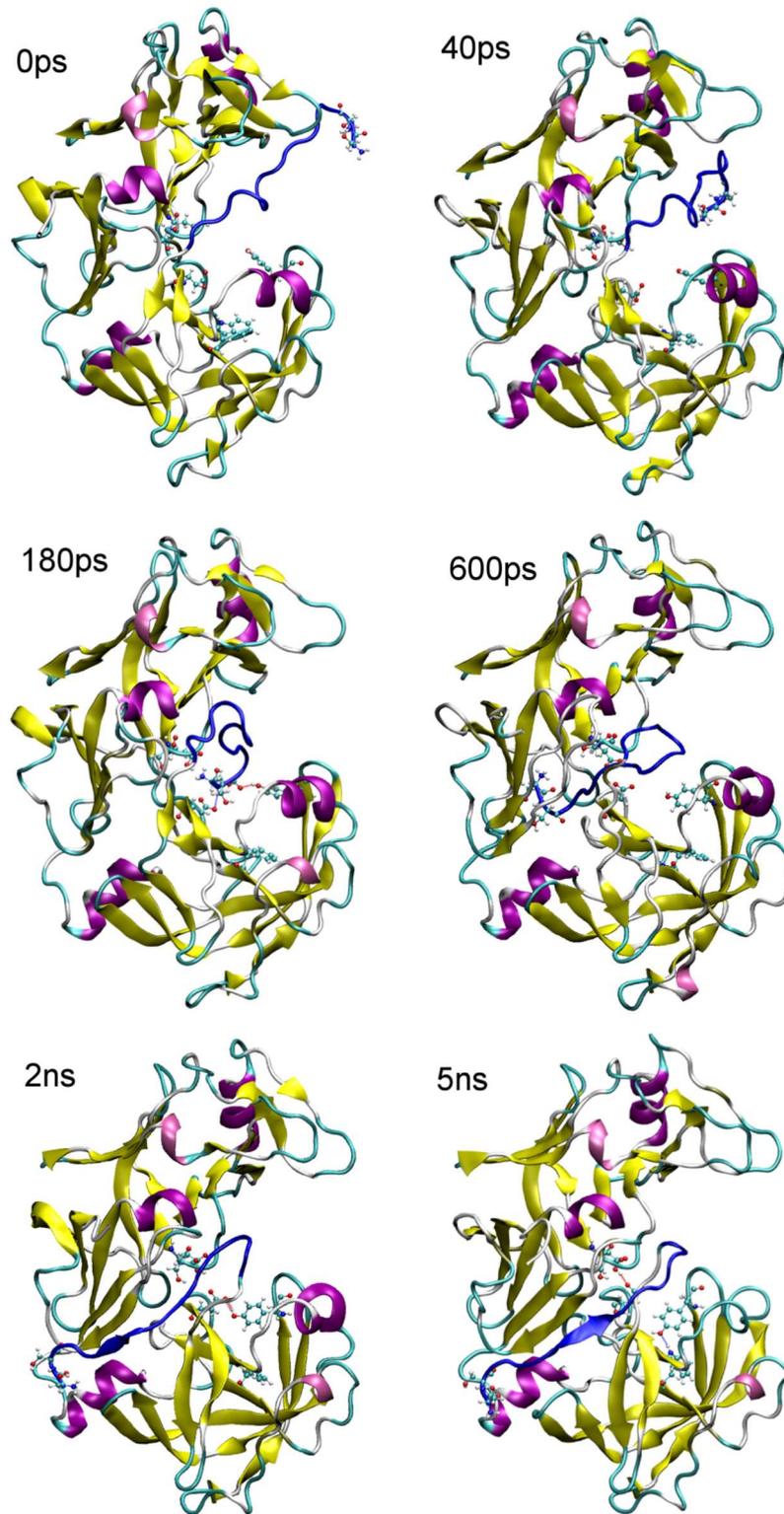


Figure 2: Snapshots from one of the EDS simulations of the activation process of PM. The N-terminal segment (residues 1-14) is shown in blue regardless of its secondary structure. Residues 15-329 are color-coded according to their secondary structure (α helices in purple, 3-10 helices in pink, β strands in yellow, tight turns in cyan and coils in white).

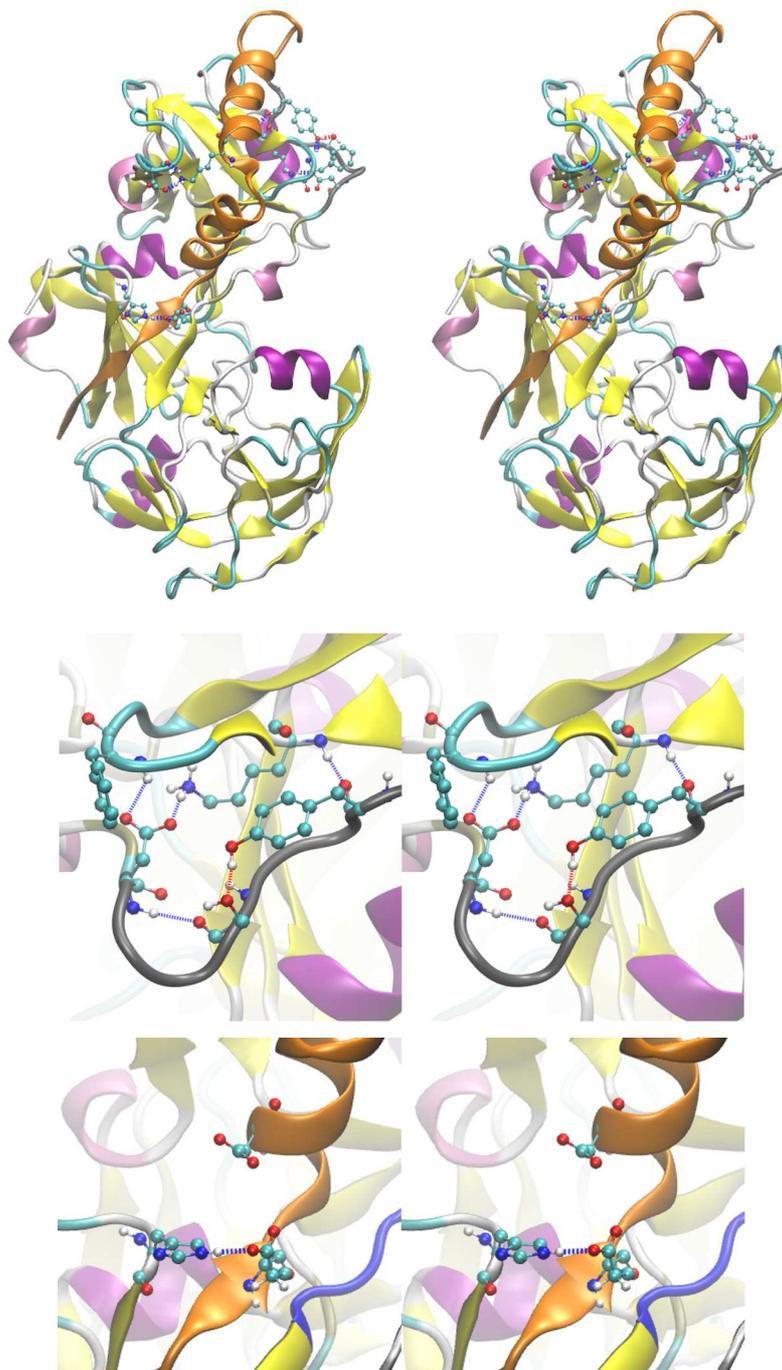


Figure 3: Interactions between the prosegment and mature portion in proPM. The prosegment is colored in orange, the pro-mature loop (Y122p-D4) in gray and the mature portion is color-coded according to the secondary structure (α helices in purple, 3-10 helices in pink, β sheets in yellow, turns in cyan and coils in white). Residues Y122pro, S1 and D4 of the pro-mature loop, and residues K238 and F241 which interact with them are shown with balls and sticks, and their interaction is highlighted in the middle frame. Residues K103pro and D279, which form a salt bridge during the simulations, as well as residues E87pro and H164 are also shown with balls and sticks. The frame on the bottom highlights residues H164, E87pro (which is hydrogen bonded to H164 in the simulations) and D91pro (which is hydrogen bonded to H164 in the crystal structure.)