Backbone Flexibility Controls the Activity and Specificity of a Protein−Protein Interface: Specificity in Snake Venom Metalloproteases

Hannes G. Wallnoefer,[†] Torsten Lingott,[‡] José María Gutiérrez,[§] Irmgard Merfort,[‡] and Klaus R. Liedl*,[†]

Institute of General, Inorganic and Theoretical Chemistry, Faculty of Chemistry and Pharmacy, University of Innsbruck, Innrain 52a, A-6020 Innsbruck, Austria, Institute of Pharmaceutical Sciences, Department of Pharmaceutical Biology and Biotechnology, University of Freiburg, Stefan-Meier-Str. 19 (VF), D-79104 Freiburg, Germany, and Instituto Clodomiro Picado, Facultad de Microbiología, Universidad de Costa Rica, San José, Costa Rica

Received November 26, 2009; E-mail: klaus.liedl@uibk.ac.at

Abstract: Protein−protein interfaces have crucial functions in many biological processes. The large interaction areas of such interfaces show complex interaction motifs. Even more challenging is the understanding of (multi)specificity in protein−protein binding. Many proteins can bind several partners to mediate their function. A perfect paradigm to study such multispecific protein−protein interfaces are snake venom metalloproteases (SVMPs). Inherently, they bind to a variety of basement membrane proteins of capillaries, hydrolyze them, and induce profuse bleeding. However, despite having a high sequence homology, some SVMPs show a strong hemorrhagic activity, while others are (almost) inactive. We present computer simulations indicating that the activity to induce hemorrhage, and thus the capability to bind the potential reaction partners, is related to the backbone flexibility in a certain surface region. A subtle interplay between flexibility and rigidity of two loops seems to be the prerequisite for the proteins to carry out their damaging function. Presumably, a significant alteration in the backbone dynamics makes the difference between SVMPs that induce hemorrhage and the inactive ones.

Introduction

The ability of proteins to bind other biomolecules is crucial for many biological processes in living organisms. Investigating these interactions is a prerequisite for the design of new proteins with defined activity and selectivity that could be used as therapeutic agents as well as tools for cellular and molecular research and engineering.[1−6] Consequently, a comprehensive knowledge of the relations between sequence, structure, and the associated dynamics is necessary to understand their function and to design novel functional proteins.[7] However, there are still significant limitations on this topic, because there is no reasonable description of protein flexibility in most of the current protein design approaches.[1,8−18] It is generally accepted that proteins are inherently flexible,[19−27] and it has already been shown that this dynamic behavior is fundamental to their function.[28−31]

(8) Karanicolas, J.; Kuhlman, B. *Curr. Opin. Struct. Biol.* **2009**, *19* (4), 458–463.
(9) Li, L.; Liang, S.; Pilcher, M. M.; Meroueh, S. O. *Protein Eng.* **2009**, *22* (9), 575–586.
(10) Liang, S.; Li, L.; Hsu, W. L.; Pilcher, M. N.; Uversky, V.; Zhou, Y.; Dunker, A. K.; Meroueh, S. O. *Biochemistry* **2009**, *48* (2), 399–414.
(11) Lippow, S. M.; Tidor, B. *Curr. Opin. Biotechnol.* **2007**, *18* (4), 305–311.
(12) Rosenberg, M.; Goldblum, A. *Curr. Pharm. Des.* **2006**, *12*, 3973–3997.
(13) Poole, A. M.; Ranganathan, R. *Curr. Opin. Struct. Biol.* **2006**, *16* (4), 508–513.
(14) Ambroggio, X. I.; Kuhlman, B. *Curr. Opin. Struct. Biol.* **2006**, *16* (4), 525–530.
(15) Koder, R. L.; Dutton, P. L. *Dalton Trans.* **2006**, (25), 3045–3051.
(16) Butterfoss, G. L.; Kuhlman, B. *Annu. Rev. Biophys. Biomol. Struct.* **2006**, *35* (1), 49–65.
(17) Vizcarra, C. L.; Mayo, S. L. *Curr. Opin. Chem. Biol.* **2005**, *9* (6), 622–626.
(18) Chica, R. A.; Doucet, N.; Pelletier, J. N. *Curr. Opin. Biotechnol.* **2005**, *16* (4), 378–384.
(19) Lee, A. L.; Kinnear, S. A.; Wand, A. J. *Nat. Struct. Mol. Biol.* **2000**, *7* (1), 72–77.
(20) Fuentes, E. J.; Der, C. J.; Lee, A. L. *J. Mol. Biol.* **2004**, *335* (4), 1105–1115.
(21) Lange, O. F.; Lakomek, N. A.; Fares, C.; Schroder, G. F.; Walter, K. F. A.; Becker, S.; Meiler, J.; Grubmuller, H.; Griesinger, C.; de Groot, B. L. *Science* **2008**, *320* (5882), 1471–1475.
(22) Eisenmesser, E. Z.; Bosco, D. A.; Akke, M.; Kern, D. *Science* **2002**, *295* (5559), 1520–1523.

† University of Innsbruck.
‡ University of Freiburg.
§ Universidad de Costa Rica.

(1) Mandell, D. J.; Kortemme, T. *Nat. Chem. Biol.* **2009**, *5* (11), 797–807.
(2) Nooren, I. M. A.; Thornton, J. M. *EMBO J.* **2003**, *22* (14), 3486–3492.
(3) Jones, S.; Thornton, J. M. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93* (1), 13–20.
(4) Archakov, A. I.; Govorun, V. M.; Dubanov, A. V.; Ivanov, Y. D.; Veselovsky, A. V.; Lewi, P.; Janssen, P. *Proteomics* **2003**, *3* (4), 380–391.
(5) Keskin, O.; Gursoy, A.; Ma, B.; Nussinov, R. *Chem. Rev.* **2008**, *108* (4), 1225–1244.
(6) Reichmann, D.; Rahat, O.; Cohen, M.; Neuvirth, H.; Schreiber, G. *Curr. Opin. Struct. Biol.* **2007**, *17* (1), 67–76.
(7) Schueler-Furman, O.; Wang, C.; Bradley, P.; Misura, K.; Baker, D. *Science* **2005**, *310* (5748), 638–642.

In contrast to the historic "lock−key" model of Fischer,[32] which assumes the binding process to be a static insertion of a low-energy conformation of the ligand into a rigid binding site, a large number of possible conformations of the receptor and the binding proteins exists.[33,34] Moreover, virtually all binding processes go hand in hand with conformational changes. It is supposed that conformations of the bound state are sampled in the free state and conformations of the free state are accessible in the bound state. Experimentally, evidence was found for protein−substrate,[35−40] protein−protein,[21,41−45] and protein−nucleic acid complexes.[46−48] Consequently, proteins exist in a multitude of conformations in solutions.[49] This is the basis for the binding mechanism model known as "conformational selection".[50−52] In principle, the "conformational selection model" assumes that a protein already exists in multiple conformations, which represent its properties. A ligand chooses its most favored preexisting protein conformation upon binding, shifting the conformational ensemble of the receptor toward the bound conformation.

An alternative model is the induced fit. In contrast to the conformational selection, the conformation of the receptor in the complex is not formed until binding. The ligand binds and consequently induces conformational changes. However, even if those two mechanisms seem to be mutually exclusive, it was shown that the distinction is blurred.[53] The maltose binding protein is good evidence.[54] Several ground-state conformations that are significantly populated are "chosen" by ligands. However, the receptor conformation is altered upon binding. The result is a mixture of induced fit and conformational selection.

A fascinating feature of many important proteins is multi-specificity. Many proteins do not have a single interaction partner to mediate their function but instead can bind several partners with good affinity.[55,56] In terms of the "conformational selection model", this implies that the protein has a rugged energy hypersurface with numerous minima.[52] Every ligand chooses it a favorable counterpart as binding partner from the conformational ensemble of the receptor and binds to it.

For almost three decades, the hypothesis that structural ensembles are useful representations of the structure−dynamics relationship has been readily accepted.[57] This was shown for protein−ligand interactions,[58] protein−protein docking[59,60] and even protein design.[10,28,29,61−63] But it is very difficult to access such ensembles by experimental methods. The most suitable experimental technique is NMR, but even in this case, limitations in size and time resolution hamper the generation of a representative ensemble. Due to the comparably small size of SVMPs, NMR experiments could be feasible. However, no such results are presented in the literature so far. MD simulations offer a more viable approach. Starting from an X-ray structure, the time evolution of the system is observed, and thus an ensemble is generated.

Detailed analyses of unbound interfaces of protein−protein interactions by computer simulations are rather rare. The only analysis the authors are aware of was published by Grünberg et al.[64] In this study, only entropy calculations but no structural evaluations are presented. The main conclusion by Grünberg et al. is that protein−protein association does not lead to a general loss of conformational entropy. Also, an elaborate flexibility with complex (allosteric) rearrangements upon protein−protein association was found. This indicates the high complexity connected with protein−protein association.

(23) Henzler-Wildman, K. A.; Thai, V.; Lei, M.; Ott, M.; Wolf-Watz, M.; Fenn, T.; Pozharski, E.; Wilson, M. A.; Petsko, G. A.; Karplus, M.; Hubner, C. G.; Kern, D. *Nature* **2007**, *450* (7171), 838–844.

(24) Boehr, D. D.; McElheny, D.; Dyson, H. J.; Wright, P. E. *Science* **2006**, *313* (5793), 1638–1642.

(25) Schnell, J. R.; Dyson, H. J.; Wright, P. E. *Annu. Rev. Biophys. Biomol. Struct.* **2004**, *33* (1), 119–140.

(26) Wolf-Watz, M.; Thai, V.; Henzler-Wildman, K.; Hadjipavlou, G.; Eisenmesser, E. Z.; Kern, D. *Nat. Struct. Mol. Biol.* **2004**, *11* (10), 945–949.

(27) Eisenmesser, E. Z.; Millet, O.; Labeikovsky, W.; Korzhnev, D. M.; Wolf-Watz, M.; Bosco, D. A.; Skalicky, J. J.; Kay, L. E.; Kern, D. *Nature* **2005**, *438* (7064), 117–121.

(28) Ding, F.; Dokholyan, N. V. *PLoS Comput. Biol.* **2006**, *2* (7), e85.

(29) Larson, S. M.; England, J. L.; Desjarlais, J. R.; Pande, V. S. *Protein Sci.* **2002**, *11* (12), 2804–2813.

(30) Friedland, G. D.; Lakomek, N. A.; Griesinger, C.; Meiler, J.; Kortemme, T. *PLoS Comput. Biol.* **2009**, *5* (5), e1000393.

(31) Humphris, E. L.; Kortemme, T. *Structure* **2008**, *16* (12), 1777–1788.

(32) Fischer, E. *Ber. Dtsch. Chem. Ges.* **1894**, *27*, 2985–2991.

(33) Dill, K. A.; Chan, H. S. *Nat. Struct. Mol. Biol.* **1997**, *4* (1), 10–19.

(34) Hilser, V. J.; García-Moreno E., B.; Oas, T. G.; Kapp, G.; Whitten, S. T. *Chem. Rev.* **2006**, *106* (5), 1545–1558.

(35) Tang, C.; Schwieters, C. D.; Clore, G. M. *Nature* **2007**, *449* (7165), 1078–1082.

(36) Lu, Z. L.; Coetsee, M.; White, C. D.; Millar, R. P. *J. Biol. Chem.* **2007**, *282* (24), 17921–17929.

(37) Fenwick, R. B.; Prasannan, S.; Campbell, L. J.; Nietlispach, D.; Evetts, K. A.; Camonis, J.; Mott, H. R.; Owen, D. *Biochemistry* **2009**, *48* (10), 2192–2206.

(38) Saitoh, T.; Igura, M.; Obita, T.; Ose, T.; Kojima, R.; Maenaka, K.; Endo, T.; Kohda, D. *EMBO J.* **2007**, *26* (22), 4777–4787.

(39) Brath, U.; Akke, M. *J. Mol. Biol.* **2009**, *387* (1), 233–244.

(40) Keramisanou, D.; Biris, N.; Gelis, I.; Sianidis, G.; Karamanou, S.; Economou, A.; Kalodimos, C. G. *Nat. Struct. Mol. Biol.* **2006**, *13* (7), 594–602.

(41) Gsponer, J.; Christodoulou, J.; Cavalli, A.; Bui, J. M.; Richter, B.; Dobson, C. M.; Vendruscolo, M. *Structure* **2008**, *16* (5), 736–746.

(42) Junker, J. P.; Ziegler, F.; Rief, M. *Science* **2009**, *323* (5914), 633–637.

(43) Koglin, A.; Mofid, M. R.; Lohr, F.; Schafer, B.; Rogov, V. V.; Blum, M. M.; Mittag, T.; Marahiel, M. A.; Bernhard, F.; Dotsch, V. *Science* **2006**, *312* (5771), 273–276.

(44) Koglin, A.; Lohr, F.; Bernhard, F.; Rogov, V. V.; Frueh, D. P.; Strieter, E. R.; Mofid, M. R.; Guntert, P.; Wagner, G.; Walsh, C. T.; Marahiel, M. A.; Dotsch, V. *Nature* **2008**, *454* (7206), 907–911.

(45) Nevo, R.; Stroh, C.; Kienberger, F.; Kaftan, D.; Brumfeld, V.; Elbaum, M.; Reich, Z.; Hinterdorfer, P. *Nat. Struct. Mol. Biol.* **2003**, *10* (7), 553–557.

(46) Kalodimos, C. G.; Biris, N.; Bonvin, A. M. J. J.; Levandoski, M. M.; Guenneugues, M.; Boelens, R.; Kaptein, R. *Science* **2004**, *305* (5682), 386–389.

(47) Zhang, Q.; Stelzer, A. C.; Fisher, C. K.; Al-Hashimi, H. M. *Nature* **2007**, *450* (7173), 1263–1267.

(48) Al-Hashimi, H. M.; Walter, N. G. *Curr. Opin. Struct. Biol.* **2008**, *18* (3), 321–329.

(49) Boehr, D. D.; Nussinov, R.; Wright, P. E. *Nat. Chem. Biol.* **2009**, *5* (11), 789–796.

(50) Tsai, C. J.; Nussinov, R. *Protein Sci.* **1997**, *6* (1), 24–42.

(51) Tsai, C. J.; Nussinov, R. *Protein Sci.* **1997**, *6* (7), 1426–1437.

(52) Ma, B.; Kumar, S.; Tsai, C. J.; Nussinov, R. *Protein Eng.* **1999**, *12* (9), 713–720.

(53) Wong, S.; Jacobson, M. P. *Proteins* **2008**, *71* (1), 153–164.

(54) Hammes, G. G.; Chang, Y. C.; Oas, T. G. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106* (33), 13737–13741.

(55) Fromer, M.; Yanover, C.; Linial, M. *Proteins* **2010**, *78* (3), 530–547.

(56) Han, J. D.; Bertin, N.; Hao, T.; Goldberg, D. S.; Berriz, G. F.; Zhang, L. V.; Dupuy, D.; Walhout, A. J. M.; Cusick, M. E.; Roth, F. P.; Vidal, M. *Nature* **2004**, *430* (6995), 88–93.

(57) Hartmann, H.; Parak, F.; Steigemann, W.; Petsko, G. A.; Ponzi, D. R.; Frauenfelder, H. *Proc. Natl. Acad. Sci. U.S.A.* **1982**, *79* (16), 4967–4971.

(58) Wei, B. Q.; Weaver, L. H.; Ferrari, A. M.; Matthews, B. W.; Shoichet, B. K. *J. Mol. Biol.* **2004**, *337* (5), 1161–1182.

(59) Chaudhury, S.; Gray, J. J. *J. Mol. Biol.* **2008**, *381* (4), 1068–1087.

(60) Prasad, J. C.; Goldstone, J. V.; Camacho, C. J.; Vajda, S.; Stegeman, J. J. *Biochemistry* **2007**, *46* (10), 2640–2654.

(61) Fu, X.; Apgar, J. R.; Keating, A. E. *J. Mol. Biol.* **2007**, *371* (4), 1099–1117.

(62) Kono, H.; Saven, J. G. *J. Mol. Biol.* **2001**, *306* (3), 607–628.

(63) Kraemer-Pecore, C. M.; Lecomte, J. T. J.; Desjarlais, J. R. *Protein Sci.* **2003**, *12* (10), 2194–2205.

(64) Grünberg, R.; Nilges, M.; Leckner, J. *Structure* **2006**, *14* (4), 683–693.

For the investigation of multispecific protein−protein interfaces, snake venom metalloproteases (SVMPs) are excellent prototypes provided by nature. Inherently, their function is to hydrolyze a variety of extracellular matrix proteins, including those of the basement membranes (BM) surrounding endothelial cells in the microvasculature, thus inducing hemorrhage.[65,66] This hemorrhagic activity can either occur locally at the site of venom injection or systemically, affecting multiple organs.[67] The mechanism by which viperid venoms induce hemorrhage is dependent on the proteolytic degradation of components of the BM in the wall of microvessels, especially capillaries.[67,68] It has been hypothesized that *in vivo* hemorrhage occurs by a two-step mechanism: first, proteins of the BM and of the endothelial cell membrane, involved in cell−BM adhesion, are cleaved, which leads to the weakening of the mechanical stability of capillaries; thereafter, in a second step, the hemodynamic forces normally operating in the microcirculation, like hydrostatic pressure and shear stress, contribute to the distension and rupture of the capillary wall, with the ensuing extravasation.[67] Thus, hydrolysis of BM components and its binding is a key step in the pathogenesis of venom-induced hemorrhage.

Interestingly, however, SVMPs differ significantly in their ability to induce hemorrhage. Despite having similar proteolytic activity toward diverse substrates *in vitro*, some of these enzymes are hemorrhagic, whereas others are not.[67,69] We assume that there are structural determinants in the metalloproteinase domain that enable them to induce hemorrhage. However, because no structures of protein−protein complexes are available, the characterization of such structural features remains elusive. Structural analyses of several hemorrhagic and nonhemorrhagic SVMPs revealed small variations in a loop region surrounding the active site.[70] This variable region contains about 25 amino acid residues and is interrupted by the structurally conserved three residue motif (Cys-Ile/Val-Met) of the Met-turn only.

Despite intensive investigations in this research area, the detailed structural determinants of hemorrhagic activity remain unknown, and there is still no possibility to predict the hemorrhagic potential of SVMPs on the basis of structure or sequence analysis. One study suggested a slight, but not linear, correlation of polar molecular surface area characteristics and hemorrhagic potential.[71] In a recently solved complex structure of BaP1, the first part of the above-mentioned loop region was found to exist in two conformations. This finding led to the assumption that flexibility of this loop could be responsible for hemorrhagic activity.[72]

To investigate the dynamics in this region, we have analyzed four different SVMPs with consistent activity data. They were measured either in our own laboratory or with the same experimental protocol.[73−75] Two of them exert hemorrhagic activity, while the other two do not induce this effect. The active

ones are BaP1 from *Bothrops asper*[72] and acutolysin-A from *Agkistrodon acutus*,[76] and the inactive ones are leucurolysin-A from *Bothrops leucurus*[77] and H2-proteinase from *Trimeresurus flavoviridis*,[78] respectively.

In a pairwise sequence alignment, BaP1 and leucurolysin-A exhibit the same type of amino acid side chains in 78% of all residues (Figure 1). Even more impressive is the fact that only 5% of all amino acid side chains are of different chemical properties. Compared with H2-proteinase, the sequential similarity of BaP1 is still rather high (53%), especially considering that the similarity between both active and both inactive proteins is equal (53%) and even lower (51%), respectively. Comparing all four SVMPs, the only major difference in sequence occurs at the loop region directly following the highly conserved active site, which surrounds the so-called Met-turn (green shaded area in Figure 1). Nevertheless, it is not possible to correlate the amino acid composition directly to the activity behavior (hemorrhagic actives and inactives) of the proteins. Indeed, there are some remarkable similarities within the group of actives and inactives. For example, the hemorrhagic active proteins (BaP1 and acutolysin-A) both present a Gly-Ser-Cys-Ser-Cys-Gly-Ala/Gly-Lys-Ser (residues 154−163) sequence in the part before the Met-turn, whereas the inactives do not show any identical residues in this section besides the two conserved cysteinyl residues. This is a further hint that flexibility might play a role to distinguish between active and inactive proteins, bearing in mind that these amino acid side chains are known to enable a broader main chain conformational flexibility.[80,81] In the part after the Met-turn (residues 167−177), no direct correlation between sequence and activity can be observed. Either active and inactive proteins possess the same amino acid side chain (positions 167, 169−172, and 174−177) or, like at position 173, they do not show the same properties within the actives (Val and Glu) and the inactives (Lys and Gly). Therefore and as mentioned above, sequence alignments can at most lead to the localization of the potential interaction area but do not explain in detail why some SVMPs are hemorrhagic active and some are not.

Methods

The X-ray structures of BaP1 (PDB code: 2w15), acutolysin-A (1bud), and H2-proteinase (1wni) were extracted from the Protein Data Bank (PDB).[82] The structure of leucurolysin-A is unreleased

(65) Bjarnason, J. B.; Fox, J. W. *Pharmacol. Ther.* **1994**, *62*, 325–372.
(66) Hati, R.; Mitra, P.; Sarker, S.; Bhattacharyya, K. K. *Crit. Rev. Toxicol.* **1999**, *29* (1), 1–19.
(67) Gutiérrez, J. M.; Rucavado, A.; Escalante, T.; Díaz, C. *Toxicon* **2005**, *45* (8), 997–1011.
(68) Markland, F. S. *Toxicon* **1998**, *36* (12), 1749–1800.
(69) Fox, J. W.; Serrano, S. M. T. *J. Proteomics* **2009**, *72* (2), 200–209.
(70) Watanabe, L.; Shannon, J. D.; Valente, R. H.; Rucavado, A.; Alape-Girón, A.; Kamiguti, A. S.; Theakston, D. G.; Fox, J. W.; Gutiérrez, J. M.; Arni, R. K. *Protein Sci.* **2003**, *12* (10), 2273–2281.
(71) Ramos, O. H. P.; Selistre-De-Araujo, H. S. *Toxicon* **2004**, *44* (5), 529–538.
(72) Lingott, T.; Schleberger, C.; Gutierrez, J. M.; Merfort, I. *Biochemistry* **2009**, *48* (26), 6166–6174.

(73) Takahashi, T.; Ohsaka, A. *Biochim. Biophys. Acta* **1970**, *198*, 293–307.
(74) Xu, X.; Wang, C.; Liu, J.; Lu, Z. *Toxicon* **1981**, *19* (5), 633–644.
(75) Wallnoefer, H. G.; Lingott, T.; Escalante, T.; Ferreira, R. N.; Nagem, R. A. P.; Gutiérrez, J. M.; Merfort, I.; Liedl Klaus R. manuscript in preparation, 2010.
(76) Gong, W.; Zhu, X.; Liu, S.; Teng, M.; Niu, L. *J. Mol. Biol.* **1998**, *283* (3), 657–668.
(77) Ferreira, R. N.; Rates, B.; Richardson, M.; Guimaraes, B. G.; Sanchez, E. O. F.; Pimenta, A. M. d. C.; Nagem, R. A. P. *Acta Crystallogr., Sect. F: Struct. Biol. Cryst. Commun.* **2009**, *65* (8), 798–801.
(78) Kumasaka, T.; Yamamoto, M.; Moriyama, H.; Tanaka, N.; Sato, M.; Katsube, Y.; Yamakawa, Y.; Omori-Satoh, T.; Iwanaga, S.; Ueki, T. *J. Biochem.* **1996**, *119* (1), 49–57.
(79) Larkin, M. A.; Blackshields, G.; Brown, N. P.; Chenna, R.; McGettigan, P. A.; McWilliam, H.; Valentin, F.; Wallace, I. M.; Wilm, A.; Lopez, R.; Thompson, J. D.; Gibson, T. J.; Higgins, D. G. *Bioinformatics* **2007**, *23* (21), 2947–2948.
(80) Dahl, D. B.; Bohannan, Z.; Mo, Q.; Vannucci, M.; Tsai, J. *J. Mol. Biol.* **2008**, *378* (3), 749–758.
(81) Anderson, R. J.; Weng, Z.; Campbell, R. K.; Jiang, X. *Proteins* **2005**, *60* (4), 679–689.
(82) Berman, H.; Henrick, K.; Nakamura, H. *Nat. Struct. Mol. Biol.* **2003**, *10* (12), 980.
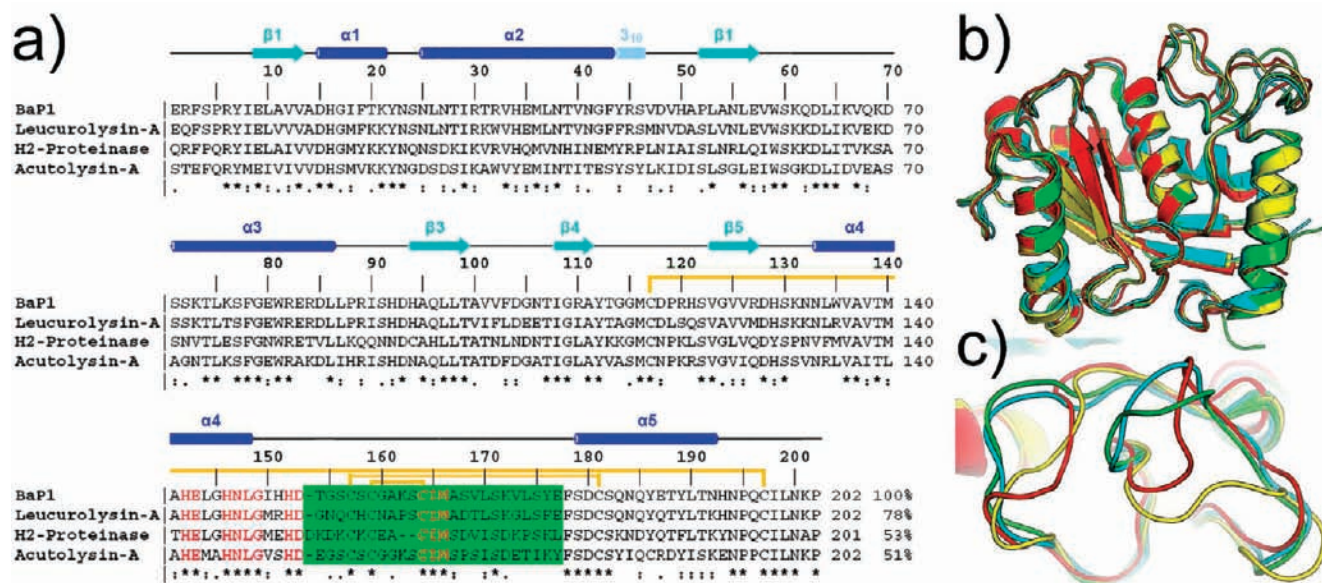
Figure 1. (a) Sequence alignment of the four SVMPs performed with ClustalW2.[79] Residue numbering and secondary structure elements (blue, α-helices; cyan, β-sheets; pale blue, 3₁₀-helix) correspond to those of BaP1 (2w15). Highly conserved residues of the active site and the Met-turn are depicted in red letters, and residues of the highly dynamic loop area are green shaded. The quantity of residues and sequence identities compared with BaP1 are given at the end of the sequences. Pairwise alignments of the other three proteins resulted in following sequence identities: leucurolysin-A and H2-proteinase 53%; leucurolysin-A and acutolysin-A 50%; H2-proteinase and acutolysin-A 51%. Consensus symbol definitions: "∗" denotes identical residues; ":" conserved residue substitution; "." semiconserved residue substitution. (b) The hemorrhagic actives (BaP1 (green) and acutolysin-A (red)) and the inactives (H2-proteinase (yellow) and leucurolysin-A (cyan)) show a high structural similarity. Differences arise from residues 156−175. (c) An enlarged depiction of the residues 156−175. Structural deviations in the X-ray structures are not correlated with the hemorrhagic activity. All four protein structures show deviations in the backbone structure of residues 156−175; however, no separation from actives (red and green) and inactives (cyan and yellow) by means of structural analyses is possible.

and was kindly provided by R. N. Ferreira and A. M. de Pimenta, Universidade Federal de Minas Gerais, Brasil. These four proteins were chosen because their activity was determined consistently with the same experimental protocol.[73−75] For BaP1, both conformations of the 160−162 loop were used (referred as "BaP1 conformer 1" and "BaP1 conformer 2" in the following, leading to a total of three simulations for the hemorrhagic active proteins). BaP1 is ligated with a peptidomimetic inhibitor in the proteolytic binding site. This ligand is retained for the simulations.

Because the structure of leucurolysin-A contains several 1,2-dihydroxyethane (in the following abbreviated with "edo") molecules, which are crystallization agents, two simulations were set up. One with edo molecules and another one without them. This provides two independent starting points (referred as "leucurolysin-A with edo" and "leucurolysin-A without edo" in the following). Edo is consequently treated in the same way as the ligand in BaP1.

The structures were imported into MOE 2010.09[83] and protonated with the protonate3D tool.[84] Afterwards, the protonation state of the amino acids, especially of the histidines, was revised manually. This was especially necessary for the Zn²⁺ coordinating histidines, because MOE proposed in some cases deprotonated species, which were changed into neutral histidines coordinating with the lone pair.

The structures were further prepared in the leap tool of amber10.[85] For the proteins, the amber force field 99SB[86] was employed. In BaP1, a nonstandard residue is present. This pyroglutamate unit was constructed from a glutamate and the atom types

were assigned manually. The partial charges were calculated with RESP HF-6/31G* using antechamber[85] and Gaussian 03.[87] The same procedure for partial charges and the amber generalized atom force field (GAFF)[88] were applied for the ligand in BaP1 and edo in leucurolysin-A. For Zn²⁺ ion, the parameters according to Ryde[89] were used. All water molecules present in the X-rays were retained. The structures were then soaked in a box of TIP3P water molecules.

The equilibration for the MD simulation was carried out using the sander tool of amber10. The protocol consists of 15 steps. All minimizations comprised 2000 steps of steepest decent followed by 2000 steps of conjugate gradient. Firstly only hydrogens were minimized to gain a reasonable hydrogen bond network. Afterwards a minimization of the water molecules was performed with a positional restraint of 1000 kcal mol⁻¹ Å⁻² on the protein heavy atoms. For the next three simulation steps, the restraint was kept on the protein heavy atoms. Firstly the temperature was raised from 100 to 300 K over 100 ps in a *NVT* ensemble followed by 200 ps simulation at 300 K with a *NPT* ensemble to adjust the box size. Afterwards a cooling back to 100 K in a *NVT* ensemble was performed. Then a series of nine minimizations with decreasing positional restraint on the protein heavy atoms (1000, 500, 100, 50, 20, 10, 5, 2, 0 kcal mol⁻¹ Å⁻²) was performed. Thus, steric clashes and similar problems should be removed without significantly disturbing the structure. Finally, the system was again heated to the simulation temperature of 300 K over a 100 ps *NVT* simulation.

Production runs were carried out with the pmemd tool of amber10. All simulations were performed in a *NPT* ensemble with a nonbonded cutoff of 8 Å, particle mesh ewald (PME with Amber defaults values; ewald coefficient 0.34864 Å⁻¹; cutoff 8 Å; tolerance

(83) MOE (The Molecular Operating Environment), version 2010.01, software available from Chemical Computing Group Inc., 1010 Sherbrooke Street West, Suite 910, Montreal, Canada H3A 2R7, http://www.chemcomp.com, 2009.

(84) Labute, P. *Proteins* **2009**, *75* (1), 187–205.

(85) *AMBER 10*; University of California: San Francisco: 2008.

(86) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. *Proteins* **2006**, *65* (3), 712–725.

(87) Frisch, M. J. et al. *Gaussian 03*, revision C.02; Gaussian, Inc.: Wallingford CT, 2004.

(88) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollmann, P. A.; Case, D. A. *J. Comput. Chem.* **2004**, *25* (9), 1157–1174.

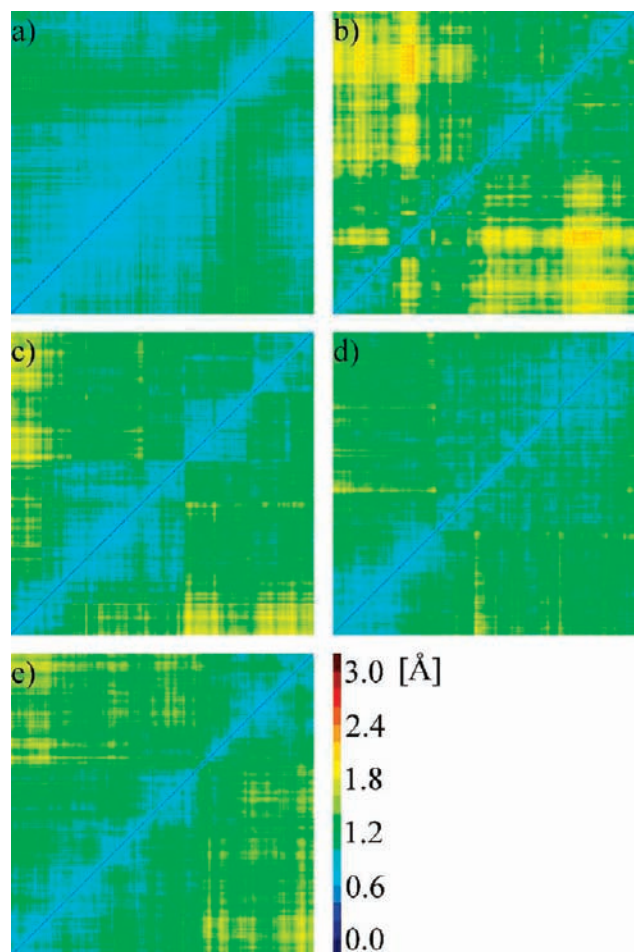(89) Ryde, U. *Proteins* **1995**, *21* (1), 40–56.

Figure 2. Two-dimensional rmsd plots for the C_α atoms for (a) acutolysin-A, (b) BaP1 comformer 2, (c) H2-proteinase, (d) leucurolysin-A with edo, and (e) leucurolysin-A without edo. Only in the BaP1 conformer 1 simulation a conformational change occurs (see Figure 3).

0.00001[90] for long-ranging electrostatics and SHAKE for the bonds involving hydrogens. As a temperature regulation system, the Langevin algorithm[91] was applied (time constant 1.0 ps; the seed for the pseudorandom number generator was set to 71277; collision frequency, $\gamma = 3.0$ ps^{-1}).

Calculations of root-mean-square deviations (rmsd, positional root-mean-square deviations of the C_α atoms), B-factors, 2D-rmsd plots (a rmsd value for the backbone of every snapshot compared with every other snapshot is displayed as a heatmap), and representative structures[92] were done with ptraj. Representative structures were calculated on 1200 snapshots, extracted every 25 ps, applying the means algorithm. B-factors were calculated as the mean square positional fluctuations of an atom multiplied with (8/3)π^2.

Pictures of structures and structural analysis were performed in PyMOL.[93]

Results

MD Simulations. After the equilibration, for all six complexes, 30 ns of MD simulation was performed. Stable trajectories with mean backbone rmsd's well below 2 Å could be generated for further analysis (Figure 2).
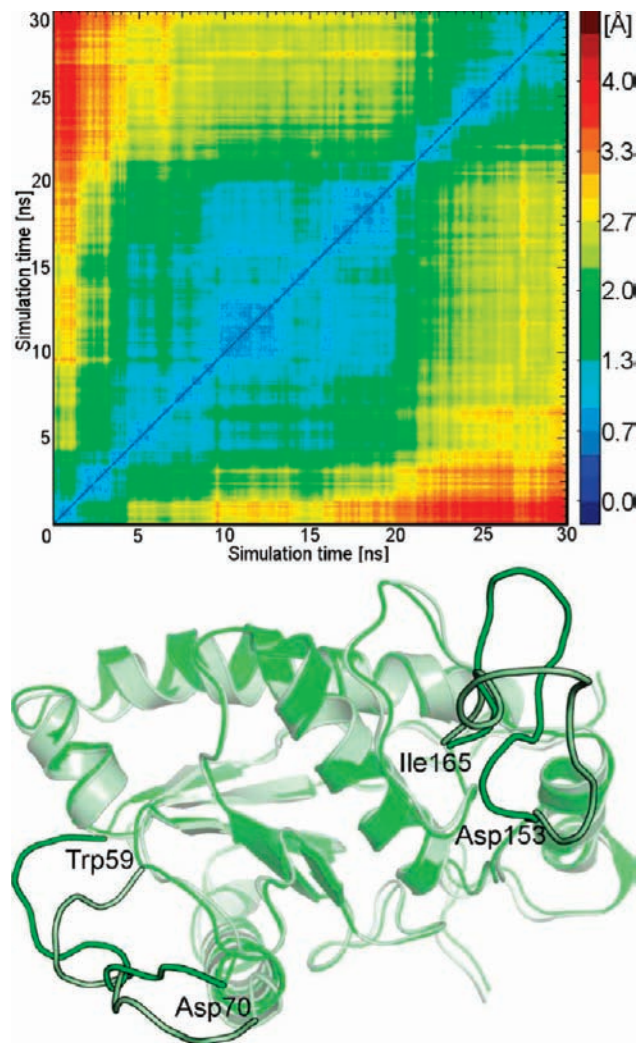
(90) Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98* (12), 10089–10092.
(91) Wu, X.; Brooks, B. R. *Chem. Phys. Lett.* **2003**, *381* (3−4), 512–518.
(92) Shao, J.; Tanner, S. W.; Thompson, N.; Cheatham, T. E. *J. Chem. Theory Comput.* **2007**, *3* (6), 2312–2334.
(93) *PyMOL*; DeLano Scientific: Palo Alto, CA, 2002.

Figure 3. Top, BaP1 conformer 1 undergoes a conformational change after about 20 ns. The 2D-rmsd plot shows two stable conformations. The first is sampled from about 7 to 20 ns, and the second from 20 to 30 ns. Within these simulation times, the conformations stay stable. Besides the flexibility of residues 156−167, which is the main focus of this study, residues 59−70 rearrange. Presumably, this does not affect the former loop. Bottom, Alignment of a snapshot after 15 ns (green) overlaid with a snapshot after 25 ns (pale green). The two loops that undergo the structural transition are highlighted. The first loop lasts from residues 59 to 70, while the second (residues 153−163) is in the protein−protein-interface area.

Conformer 1 of BaP1 undergoes a conformational change in two regions, which is connected with a larger rmsd shift (Figure 3).

Besides a transition in the loop of interest for the protein−protein interface (residues 153−165; all residue numbers mentioned in the following correspond to BaP1 (2W15); the other proteins were aligned on BaP1 for analyses), also a region from residue 59 to 70 rearranges itself. However, the latter loop is on the opposite site of the protein and therefore should not disturb the flexibility of the former loop. As Figure 3 illustrates, the rest of the protein perfectly stays in its structural shape. A similar event is not observed in any of the other simulations (see Figure 2).

In comparison of the two conformations of BaP1, neither a conversion from one conformation into the other nor a common end point conformation is found. Both are rather stable during the whole time of 30 ns of computer simulations.
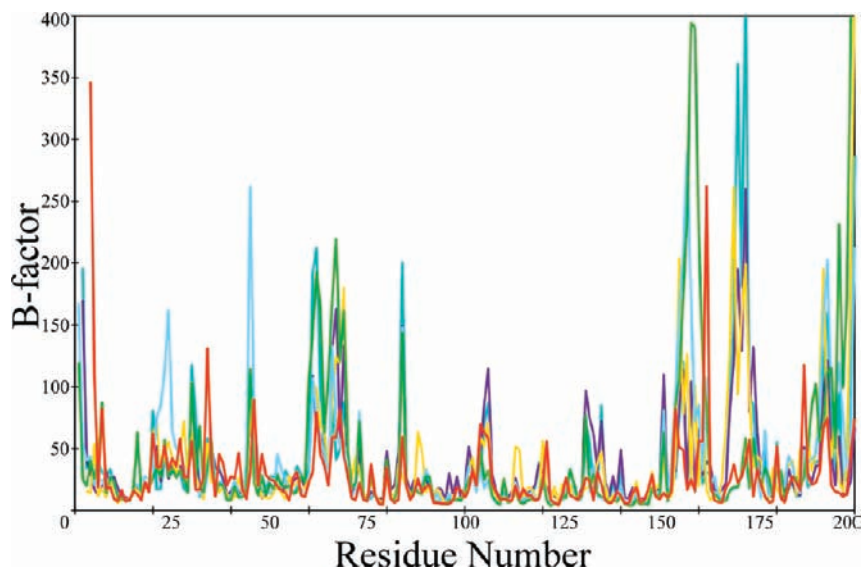
Figure 4. Residue-wise *B* factor plots for all six complexes over the entire simulation (30 ns). While all six proteins show similar dynamic behavior in most parts of the sequences, the protein interface region 156−175 provides remarkable differences. Figure 5 gives a close up view of this interesting region. Color code: red, acutolysin-A; green, BaP1 conformer 1; blue, BaP1 conformer 2; yellow, H2-proteinase; turquoise, leucurolysin-A with edo; magenta, leucurolysin-A without edo.
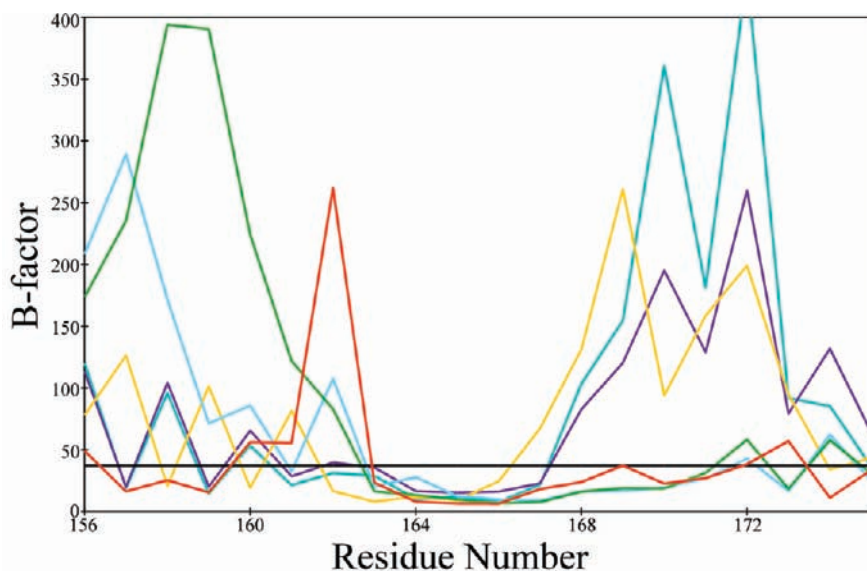

Figure 5. Detailed residue-wise *B* factor plot for the flexible loops (residues 156−175). While in the rest of the proteins a similar dynamic behavior is found, a separation of hemorrhagic active and inactive proteins is found here. The actives undergo greater conformational changes in the region 156−163, and the inactives are more flexible in the 167−175 part, respectively. The ability to bind the target proteins is likely to be related to the dynamics of this loop. Color code: red, acutolysin-A; green, BaP1 conformer 1; blue, BaP1 conformer 2; yellow, H2-proteinase; turquoise, leucurolysin-A with edo; purple, leucurolysin-A without edo; black, mean *B* factor over all residues and simulations. To illustrate the convergence of the results, the Supporting Information provides the plot just for the last 15 ns and just for the backbone atoms in the last 15 ns.

Analyses of the flexibility of the complexes in the simulations are performed with calculation of *B* factors. Plots for residue-wise (mean *B* factor for every residue) *B* factors for all six complexes are shown in Figure 4.

Despite small differences in intensity, the overall motions are very similar. The only exception is the region from residue 156 to 175. A close up view (Figure 5) shows that two flexibility groups can be observed.

The hemorrhagic active proteins (acutolysin-A and the two conformers of BaP1) exhibit a significantly higher flexibility in the first part of the loop (before the Met-turn; residues 156−163), while the inactives have much higher flexibility in the loop region after the Met-turn (residue numbers 166−175). Acutolysin-A has a peak of dynamics at the end of the loop at

Met163, the residues before have no significantly higher flexibility than the inactives at the same positions. BaP1 conformer 1 is by far the most flexible example and hence reaches the highest *B* factors. The whole loop area from Ser156 to Lys162 shows highly dynamic behavior. In contrast to acutolysin-A, BaP1 conformer 2 shows its peak of flexibility in the loop from residue 157 to 159.

Hemorrhagic inactive proteins (H2-proteinase and leucurolysin-A with and without edo) show a different behavior. Slightly above-average flexibility in the 156−165 region is observed, but they have significantly lower *B* factors (about three times) than the actives. However, a consistent picture is observed in the loop area after the Met-turn from residue 168 to 175. In all three cases, a double peak at almost the same residue numbers
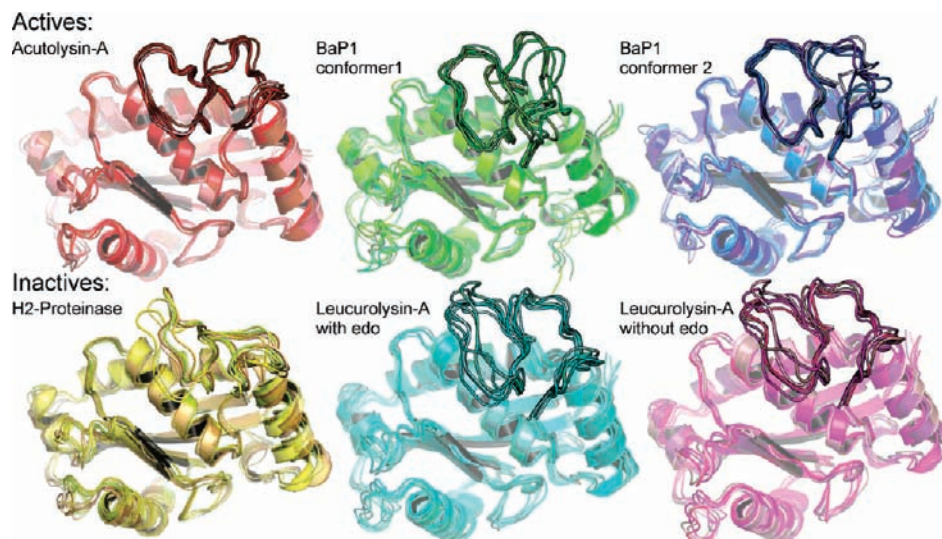
Figure 6. Alignment of the X-ray structure and five representative structural clusters observed during the simulations. Residues 156−175 are highlighted. For the hemorrhagic active proteins (first line), the right part of the highlighted loops (residues 156−163) performs significant conformational changes, while the left part (residues 167−175) keeps its conformation during the entire 30 ns of MD simulation. For the inactive proteins (second line), the dynamics of the loops are the opposite way around.

is found. Residues 170 (in case of H2-proteinase residue 169) and 172 are the most flexible amino acid side chains. The active proteins do not show a similar behavior. They show almost no dynamics with *B* factors far below average after the Met-turn region.

Splitting the *B* factors into backbone and side chain contributions, the inactives show analogous composition of the overall *B* factors in the 168−175 region. *B* factors calculated from the side chain motions are twice as large as the backbone atom ones. However, backbone *B* factors are by far (a factor of 5−7) higher than those for hemorrhagic active proteins. The same applies to the side chain atoms. The actives have their highest flexibility in the area before the Met-turn (residues 156−165), where the *B* factor composition is not homogeneous. While acutolysin-A and conformer 2 of BaP1 behave similarly to the inactives in the region after the Met-turn (a ratio of 2:1 of side chain and backbone *B* factors), BaP1 conformer 1 shows a much larger backbone contribution (almost 1:1 ratio). Thus, the backbone significantly contributes to the dynamics and flexibility differences between hemorrhagic actives and inactives.

Consequently, the trajectories were clustered and representative structures were calculated for every cluster in each trajectory. Hence, every single structure represents a cluster of many snapshots, which are related to the representative one (Figure 6).

Except for residues 156−175, all parts of the proteins stay in perfect alignment (the only exception is the above-mentioned second mobile loop in BaP1 conformer 1). Hence, the clustering was done on residues 156−175 and therefore the structural changes of the presumed protein−protein interface are investigated. Flexibility, as represented by the B-factors, is not due to a higher local fluctuation, but rather a slow flip of the backbone. This results in a fan-like shaped conformational ensemble as shown in Figure 6. Although, hemorrhagic active and inactive proteins show this behavior, they differ in the region of its occurrence. While in case of the actives the already as flexible identified loop (156−163) before the Met-turn is confirmed flexible, the inactives stay rigid here. However, after

the Met-turn (167−175) the actives exhibit a rigid behavior and the inactives show the fan-shape conformations.

Discussion

We present MD simulations of SVMPs, proteins that have to bind membrane proteins to induce their hemorrhagic effect. However, these enzymes have drastically differing hemorrhagic activity, although the proteins possess a high sequential and structural homology. The first hint that dynamics may play a crucial role for this protein−protein interaction site already emerged from the X-ray structure of BaP1.[72] Parts of a loop (residues 159−163) were found to exist in two clearly distinguishable conformations.

Consequently, we suggest that this enables this area to be the multispecific protein−protein interface, which allows SVMPs to induce hemorrhage. To test this assumption, we have set up six simulations of four SVMPs. Under consistent experimental conditions, two of them are hemorrhagic active, while the other two are not. Our simulations show converged behavior as evidenced by rather stationary dynamics. Only the surface area, composed of two loops (residues 156−165 and 167−175), exhibits a remarkable movement pattern. Both loops can be observed in many different conformations. This is not restricted to side chain atoms, but the whole backbone flips. The large motions are highly noteworthy, because residue 157 and residue 159 are cross-linked disulfide cysteines usually implying a higher rigidity.

Even more remarkably, we find a separation of hemorrhagic active and inactive proteins. While the actives exhibit the enhanced motions in the first loop (residues 156−165), the inactives have a dynamic loop at residues 167−175. This supports our premise that this particular region is responsible for the different hemorrhagic potency and, thus, the ability to bind the target proteins in the BM.

According to the "conformational selection" theory, we suppose that a sufficient adaptability of the loop 156−165 is a prerequisite for the docking of different target proteins. In this model, the main task of the loop seems to be providing the "correct" conformation for every binding partner. Therefore, the loop 156−165 may control the target specificity of SVMPs.
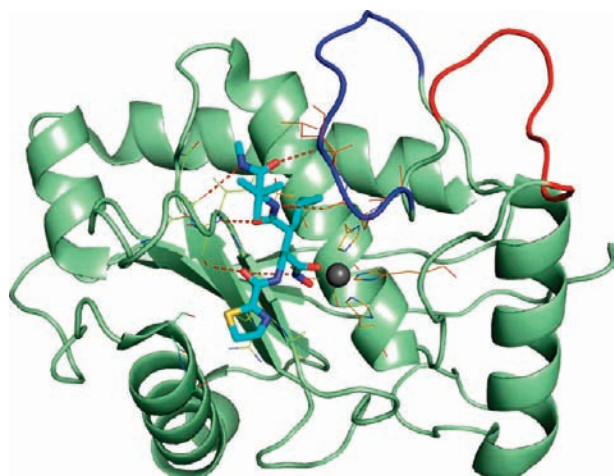
Figure 7. The loop that is flexible in the actives is shown in red (residues 156−165), whereas the one that is flexible in the inactives is shown in blue (residues 167−175). The gray sphere represents the zinc ion, which is close to the loops. The structure of BaP1 (green; pdb code 2W15) and the ligand (cyan) in the binding site illustrate the close vicinity. Dynamics of the two loops and a rigidization of them upon target protein binding are expected to influence the conformation of the proteolytic binding site.

Further, we surmise that the location of the interface introduces the possibility to induce conformational changes in the proteolytic binding site (Figure 7).

The dynamic interactions over several hydrogen bonds allow direct conformational communication. Hence, if the loops become rigid upon binding, then propagation into the binding site may be the next step. However, only an experimentally determined structure (or several of them) of the protein−protein complex could confirm this assumption. To illustrate the alteration of the binding site dimensions during the simulations, the Supporting Information provides distance vs simulation time plots of relevant distances for all six simulations. Again, a different behavior for hemorrhagic active and inactive proteins is found.

The fact that the whole backbone moves is of great interest for the field of protein engineering. Most current protein design methods assume a fixed backbone. In consideration of this, the sequence is changed in order to generate new properties of the target protein. However, in the case of multispecificity of SVMPs, this would not lead to success. From our results, we argue that not sequence deviations in the 157−175 region but a considerably enhanced flexibility resulting from these small sequence deviations alter the ability to bind the target proteins.

Consequently we surmise that backbone flexibility could be also essential for other protein−protein interactions and their specificity.

Conclusions

SVMPs have different abilities to bind target proteins. Until now, no explanation of the different hemorrhagic potential, which is related to binding ability, was found. We suggest that this is a problem completely related to dynamics. Despite a high sequential homology, the loop region 156−175 shows totally different flexibility for actives and inactives. This area is suspected to be the multispecific protein−protein interface.

In contrast to the assumptions in most current protein design approaches, the specificity of the interface is connected with large backbone shifts. A fan-shaped conformational ensemble for the protein−protein interface region is found in our simulations. A certain combination of flexibility (residues 156−165) and rigidity of the neighboring loop (residues 167−175) provides an appropriate association domain for every target protein. Hemorrhagic inactive proteins do not offer this particular flexibility. Hence, a separation in context of activity can be achieved by *B* factor calculation.

A deeper understanding of the functionality of the mechanism of hemorrhage is fundamental to create treatment possibilities for the victims of snake bites. However, SVMPs are even more interesting, due to their high structural analogy to human proteins like ADAMs (a disintegrin and metalloproteinase domain containing enzymes) and MMPs (matrix metalloproteinases).

Finally, we emphatically recommend considering backbone flexibility as a source of multispecificity in protein−protein interfaces. It is likely that the mechanism suggested in this study is not limited to SVMPs and their substrates but can be transferred to other protein−protein interfaces.

Supporting Information Available: *B* factor plots for the last 15 ns of the simulation and *B* factor plots just for the backbone atoms in the last 15 ns to illustrate convergence, distance vs simulation time plots of relevant distances of the binding site for all six simulations, a mol2-file for the pyroglutamate unit, and complete ref 87. This material is available free of charge via the Internet at http://pubs.acs.org.

JA909908Y