Inspection of the Binding Sites of Proteinase3 for the Design of a Highly Specific Substrate

Eric Hajjar,[†] Brice Korkmaz,[‡] Francis Gauthier,[‡] Bjørn Olav Brandsdal,[§] Véronique Witko-Sarsat,[⊥] and Nathalie Reuter*,[†]

Computational Biology Unit, BCCS, University of Bergen, N-5008 Bergen, Norway, INSERM U618, Université François Rabelais, 10 boulevard Tonnellé, F-37000 Tours, France, The Norwegian Structural Biology Centre, Faculty of Science, University of Tromsø, N-9037 Tromsø, Norway, and INSERM U507, Université René Descartes Paris 5, Hôpital Necker, 161 rue de Sèvres, F-75015 Paris, France

Received October 10, 2005

Proteinase3 (PR3) and human neutrophil elastase (HNE) are homologous proteases from the polymorphonuclear neutrophils and have been thought for a long time to have close enzymatic specificity. We have used molecular dynamics simulations to investigate and compare the interactions between different peptides and the two enzymes. The important role played especially by the C-terminal part of the peptides is confirmed. We provide a map of the subsites of PR3 and a description of the interaction scheme for six ligands. The main difference between HNE and PR3 concerns S2, S1′, S2′, and S3′. The recognition subsites in PR3 are interconnected; in particular, Lys99 participates to a hydrophobic (S4) and a polar (S2) pocket. On the basis of the simulations, we suggest that VADVKDR is a highly specific sequence for PR3; enzymatic assays confirm that it is cleaved by PR3 with a high specificity constant ($k_{cat}/K_m = 3\,400\,000\ M^{-1}\ s^{-1}$) and not by HNE.

I. Introduction

Serine proteinases of polymorphonuclear neutrophils (PMN), including proteinase 3 (PR3), human neutrophil elastase (HNE), and cathepsin G, have been considered for a long time as deleterious mediators able to degrade connective tissues.[1] These proteinases are translocated to the surface of the plasma membrane and/or released from the activated neutrophils at an inflammatory site.[2,3] These proteases are thus involved in proteolytic events associated with inflammation and are associated with several diseases such as emphysema,[4] cystic fibrosis, rheumatoid arthritis, or vasculitis.[5] Until now, the characterization of the roles of each of the proteases in the development of the lung diseases pathologies has been impaired by the lack of specific sensitive substrates. Proteases from the PMNs are homologous in sequence and have been thought for a long time to have close enzymatic specificity. However, they might have very different physiologic roles, especially in their intracellular targets. Recent work has pointed out that PR3 can have intracellular specific protein substrates, thus resulting in the regulation of intracellular functions such as proliferation[6] or apoptosis.[7] These activities were not observed for HNE. We have recently described that PR3 cleaved the cyclin-dependent kinase inhibitor, p21/waf1/CIP1, and the P1 cleavage site has been identified at the Ala45.[8] To analyze the physiological consequences of this cleavage, we have generated a p21 resistant to PR3 cleavage by mutating the Ala45 into Arg45. Most interestingly, the ectopic expression of the p21A45R into a monocytic cell line triggered its differentiation, thus demonstrating that PR3, by cleaving p21, prevented the differentiation.[8] PR3 can be considered as a relevant drug target, because (i) it is expressed only in the myeloid lineage and (ii) its specific inhibition—by potentiating differentiation—might be useful in the treatment of leukemia.[9,10]

PR3 and HNE share a high sequence identity (57%) (cf. Figure 1A), and PR3 is often classified as an elastase-like protease.[11,12] Their three-dimensional structure has the two six-stranded β-barrels (cf. Figure 1B) typical of chymotrypsin-like serine proteases. The active site is located between the two β-barrels and is composed of three amino acids: a histidine (His57), a serine (Ser195), and an aspartic acid (Asp102), forming the well-known catalytic triad. The correct positioning of the cleavable amide bond in the active site is ensured by interactions between the enzyme and the peptide that extend on both sides of the cleavable bond. These recognition sites determine the protease specificity (for reviews, see references 11−14) and are denominated using the Schechter and Berger nomenclature:[15] Sn, ..., S2, S1, S1′, S2′, ...Sn′ for the recognition sites on the enzyme and Pn, ..., P2, P1, P1′, P2′, ...Pn′ for the corresponding amino acids of the peptide, where the P1−P1′ bond is the cleavable bond (Figure 1C).

Compared to other serine proteases, there are relatively few available data on the specificity of PR3. The earliest studies revealed the preference of PR3 for "small hydrophobic amino acids" such as alanine, serine, and valine.[16] Brubaker et al.[17] emphasized the fact that norvaline (Nval) is preferred to a valine, which itself is better than an alanine. Substrates with a methionine at P1 are also efficiently cleaved by PR3.[18] As for HNE, it can tolerate larger hydrophobic side chains such as isoleucine, leucine, and norleucine (Nleu). The basis for the difference in P1 specificity between PR3 and HNE has been explained by comparing the X-ray structures of HNE and PR3.[19] The replacement of Val190 (HNE) by Ile (PR3) and of Ala213 (HNE) by Asp (PR3) would reduce the size of the S1 pocket and accounts for the preference of PR3 for smaller amino acids than HNE. The X-ray structure of PR3 provided a description of the other binding sites. The structure shows that S2 is a deep pocket in PR3, more polar than in HNE due to the Leu/Lys99 substitution. The catalytic Asp102 would be more solvent accessible than in HNE. The residue at position P3 would be mostly solvent exposed, and thus S3 would play no role in specificity. S4 is more polar and smaller for PR3 than for HNE (Leu99 substituted by Lys in PR3).[18] P1′ and P3′ specificities should be influenced by the loop containing Asp61. S2′ is more polar for PR3 than for HNE because of the Leu143 to Arg substitution.[18,19] Although Kam et al.[18] mentioned in 1991 the difference in the S′ subsites between PR3 and HNE, it is only

* To whom correspondence should be addressed. Tel: (+47) 555 84040. Fax: (+47) 555 84295. E-mail: Nathalie.Reuter@cbu.uib.no.
[†] University of Bergen.
[‡] Université François Rabelais.
[§] University of Tromsø.
[⊥] Université René Descartes Paris 5.

10.1021/jm051018t CCC: $33.50
Published on Web 01/26/2006

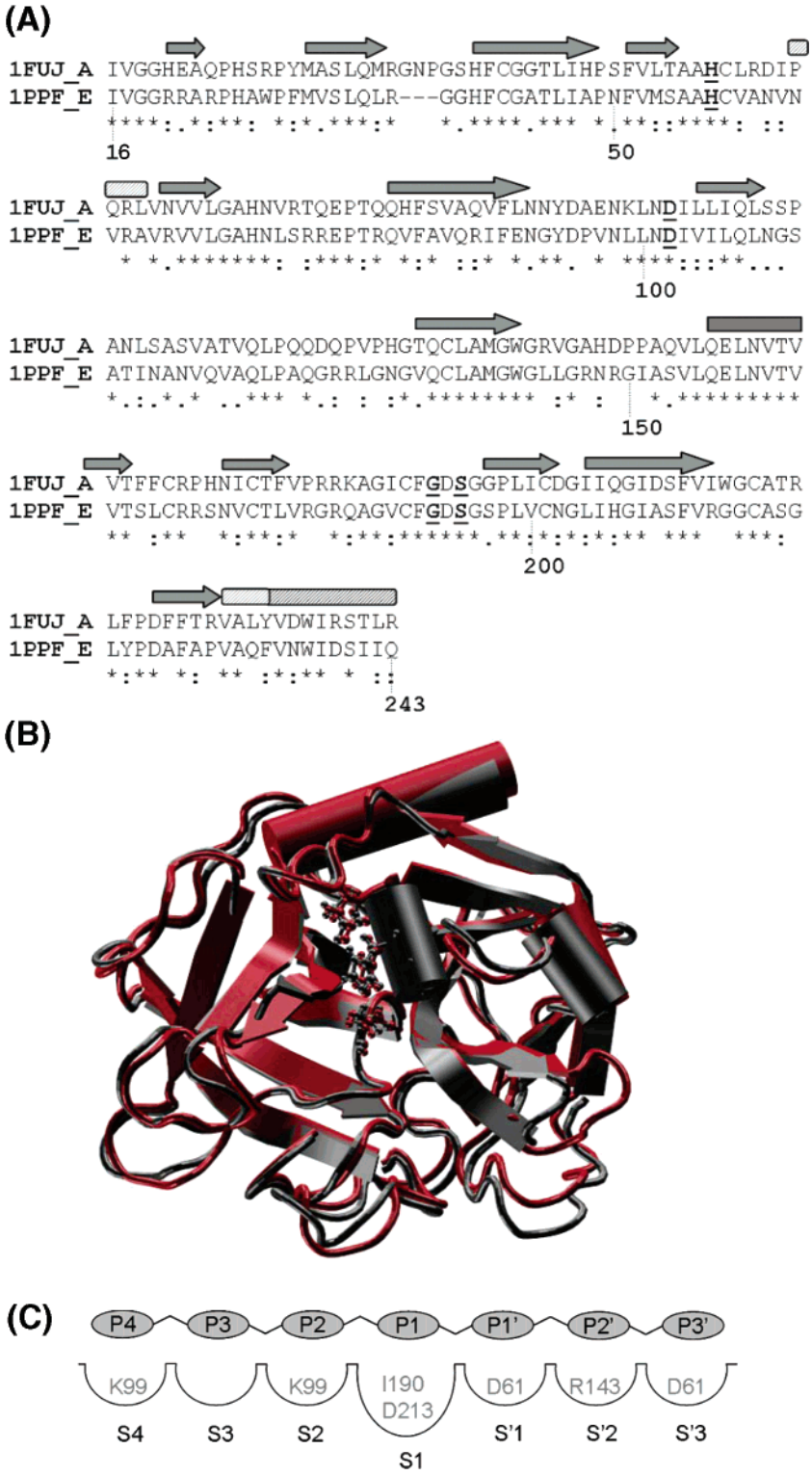

Figure 1. Similarities between HNE and PR3. (A) Sequence alignment of HNE and PR3; bold underlined letters are used for active site residues (H57, D102, G193, S195). Secondary structure elements are shown: plain gray arrows for extended sheets and dashed rectangles for helices (light gray and black for 3−10 and α-helices, respectively). (B) Structural alignment (rms deviation backbones: 2.43 Å); the enzymes (PR3 in red, HNE in gray) are drawn using the cartoon representation and the three residues of the active site using balls and sticks. (C) Schechter and Berger convention for the nomenclature of the enzyme−peptide interaction sites and residues postulated to constitute the S subsites in the case of PR3.[19]

recently that the importance of the S′−P′ interactions has been clearly established;[20,21] substrates extending beyond P1′ have a favorable effect on PR3 catalysis but not on HNE catalysis. Specific PR3 substrates are now available that allow quantification of the protease even in a complex biological medium.[20−22] Recently, Korkmaz et al.[22] reported PR3-specific peptides derived from serpin sequences and extending on the P′ sites. Interestingly these peptides contain a cysteine as P1 residue. The atomic/molecular basis of the S′−P′ interactions have not yet been fully characterized, and their description in the literature is based either on mutation studies or derived from the structures of HNE:inhibitors complexes.

We used molecular modeling and the structural data available on free PR3[19] and HNE (bound to OMTKY3[23]) to investigate

Table 1. Peptides: Sequences and Names[a]

peptide name	sequence							k_{cat}/K_m		ref
	P4	P3	P2	P1	P1	P2	P3	HNE	PR3	
Pns	Ala	Ala	Pro	Val	Lys	Gly	Asp	179000	62000	21
Ps	Val	Ala	Asp	Cys	Ala	Asp	Arg	nsh	651000	22
t-Pns	Ala	Ala	Pro	Val	Gly			77000	360	21
t-Ps	Val	Ala	Asp	Cys	Gly			24500[b]	10900[b]	22
Pb1	Ala	Ala	Ala	Ala	Ala			3460	9	21
Pb22	Ala	Ala	Ala	Ala	Lys	Gly	Asp	5000	2600	21

[a] Specificity constants (k_{cat}/K_m) reported in refs 21 and 22 for peptides of comparable sequences are given in M^{-1} s^{-1}. [b] Experimental values are reported for Val-Ala-Asp-Cys-*Ala*.

and compare the interactions between different peptides and the two enzymes at an atomic level of details. No structure of PR3−substrate complexes have been reported yet, unlike for HNE where several X-ray structures of enzyme−ligand complexes are available.[23−26] In what follows, free energies of binding of six different complexes of PR3 determined using the linear interaction energy (LIE) method are reported, and we present a thorough analysis of the interactions stabilizing the Michaelis complexes of PR3 and several peptides. Several computational methods can in principle be used to estimate the binding free energies,[27] and the LIE method has several appealing features in this respect. It allows determination of the binding free energy for ligands that are structurally and chemically different, which is not possible with the theoretically exact free energy perturbation method. It also provides energies that are more accurate than those obtained with simple but very fast scoring functions. We used specific and nonspecific peptides of PR3, based on molecular construction and molecular dynamics (MD) simulations of PR3−ligand complexes. In particular, peptides AAPVKGD[20,21] and VADCADR[22] were used, as well as alanine-rich substrates (AAAAA, AAAAKGD)[21] and short peptides (AAPVG, VADCG). Specificity constants of PR3 and HNE for these peptides[20−22] are summarized in Table 1. Results from MD simulations of HNE complexed with the same peptides are used as a reference to understand the factors governing the specificity of PR3. Using structural data provided by MD simulations, we draw an updated map of the enzyme interaction subsites S4 to S3′ and make an inventory of the interactions between PR3 and peptides. In addition we get information about the plasticity of the interaction sites. These results are then used to propose a consensus sequence made of seven amino acids and predicted to be specific for PR3. The ability of the proposed peptide to distinguish between HNE and PR3 is verified by enzymatic assays. This work provides new information with potential use for the design of peptidic substrates specific for PR3 but also for drug design and the identification of new intracellular targets of PR3.

II. Results

a. RMS Deviation along Simulation Time and B-Factors. We calculated RMS deviations between the backbone (N, Cα, C, O) of the initial structures of all complexes and the conformations obtained during the simulations. On the basis of the RMS variations we chose the sampling window to be from 500 to 2000 ps for all complexes. In particular, the RMS values are of the same range for the four complexes PR3:Ps, PR3:Pns, HNE:Ps, HNE:Pns; they vary from about 0.8 Å, at the beginning of the production runs, to 1.5 Å (plots available as Supporting Information). The structures of HNE and PR3 thus do not differ significantly from the X-ray structures of HNE:OMTKY3 and PR3 unbound, respectively. This is also the case for the other complexes.

Averaged atomic fluctuations were calculated from the simulations of all complexes, and compared to the ones derived from the experimental B-factors given in the PDB files (plots for Ps and Pns are given as Supporting Information). The trend is quantitatively well reproduced, where the regions of the proteins that fluctuate the most during the simulation have the highest experimental B-factors. In both cases, the magnitude of the theoretical fluctuations compares qualitatively well with the experimental fluctuations but are uniformly lower in the case of PR3. This difference is expected when comparing calculated and experimental fluctuations if the contribution from crystal disorders is not explicitly taken into consideration.[28] Residues of the catalytic triad (H57, D102, S195) are as expected at minima of the plots.

RMS deviations between simulated complexes and docked, solvated X-ray structures as well as comparisons between experimental and theoretical B-factors indicate that our simulations reproduce the available structural data well.

b. Catalytic Triad: Hydrogen Bond Network and Solvent Accessibility. In Table 2 we report the average distances between, on one hand, the N^{δ} of His57 and the carboxyl oxygens ($O^{\delta}1$ and $O^{\delta}2$) of Asp102 and, on the other hand, the Nϵ of His57 and O^{γ} of Ser195 (cf. Figure 2). Hydrogen bonds between these atoms are known to be crucial for the catalytic function of the protease.[11,13,19] For clarity, only distances for the complexes with Ps, Pns, t-Ps, and t-Pns are reported. The distances are slightly higher than the X-ray values but still reflect strong hydrogen bonds between the aspartate group and the imidazole ring. For example, $O^{\delta}1(O^{\delta}2)-N^{\delta}$ equals 3.16 (2.81), 3.37 (2.82), 3.25 (2.77), and 3.30 (2.91) Å, for HNE:Pns, PR3:Pns, HNE:Ps, and PR3:Ps, respectively. These distances agree well with the structural data; the hydrogen bond involving $O^{\delta}2$ (X-ray values: 2.56 and 2.54 Å, for HNE and PR3, respectively) is stronger than the one involving $O^{\delta}1$ (X-ray values: 3.25 and 3.50 Å, for HNE and PR3, respectively). We also observe strong hydrogen bonds between the hydroxyl group of Ser195 and the imidazole ring: 2.90, 2.92, 2.89, and 2.98 Å for HNE:Pns, PR3:Pns, HNE:Ps, and PR3:Ps, respectively. X-ray values are 2.48 and 2.36 Å for HNE and PR3, respectively. The angles calculated between acceptor, hydrogen, and donor atoms confirm the presence of these hydrogen bonds.

Interactions of Asp102 with Ala56 and Ser214 are important for a good orientation of the triad.[11] The same pattern as for the other hydrogen bonds is noticed, i.e., the average distances between donor and acceptor atoms and the corresponding angles are characteristic of strong hydrogen bonds for all complexes (results not shown). His−Asp distances in PR3:t-Ps are slightly different from the X-ray references; this is due to competing hydrogen bonds between the carboxyl group and water molecules.

An inventory of the water molecules in the vicinity of His57, Asp102, and Ser195 for all simulations showed that no water

Table 2. Catalytic Triad[a]

		HNE		PR3	
atom names	data from	distance	angle	distance	angle
$N^{\epsilon 2}{}_{His57}-O^{\gamma}{}_{Ser195}$	X-ray	2.48[23]		2.36[19]	
$N^{\epsilon 2}{}_{His57}-O^{\gamma}{}_{Ser195}$	MD Pns	2.90 ± 0.18	157 ± 12	2.92 ± 0.15	156 ± 12
$N^{\epsilon 2}{}_{His57}-O^{\gamma}{}_{Ser195}$	MD Ps	2.89 ± 0.13	154 ± 12	2.98 ± 0.19	155 ± 14
$N^{\epsilon 2}{}_{His57}-O^{\gamma}{}_{Ser195}$	MD t-Pns	2.87 ± 0.11	158 ± 10	2.88 ± 0.12	161 ± 11
$N^{\epsilon 2}{}_{His57}-O^{\gamma}{}_{Ser195}$	MD t-Ps	2.92 ± 0.18	160 ± 9	2.94 ± 0.18	158 ± 12
$O^{\delta 1}{}_{Asp102}-N^{\delta 1}{}_{His57}$	X-ray	3.25[23]		3.50[19]	
$O^{\delta 1}{}_{Asp102}-N^{\delta 1}{}_{His57}$	MD Pns	3.16 ± 0.22	143 ± 11	3.37 ± 0.22	123 ± 9
$O^{\delta 1}{}_{Asp102}-N^{\delta 1}{}_{His57}$	MD Ps	3.25 ± 0.22	141 ± 10	3.30 ± 0.27	145 ± 13
$O^{\delta 1}{}_{Asp102}-N^{\delta 1}{}_{His57}$	MD t-Pns	3.30 ± 0.22	139 ± 10	3.07 ± 0.26	143 ± 14
$O^{\delta 1}{}_{Asp102}-N^{\delta 1}{}_{His57}$	MD t-Ps	3.15 ± 0.22	147 ± 13	2.91 ± 0.22	158 ± 12
$O^{\delta 2}{}_{Asp102}-N^{\delta 1}{}_{His57}$	X-ray	2.56[23]		2.54[19]	
$O^{\delta 2}{}_{Asp102}-N^{\delta 1}{}_{His57}$	MD Pns	2.81 ± 0.13	148 ± 14	2.82 ± 0.13	165 ± 8
$O^{\delta 2}{}_{Asp102}-N^{\delta 1}{}_{His57}$	MD Ps	2.77 ± 0.11	153 ± 12	2.91 ± 0.19	150 ± 14
$O^{\delta 2}{}_{Asp102}-N^{\delta 1}{}_{His57}$	MD t-Pns	2.80 ± 0.16	156 ± 11	2.98 ± 0.22	151 ± 14
$O^{\delta 2}{}_{Asp102}-N^{\delta 1}{}_{His57}$	MD t-Ps	2.83 ± 0.13	148 ± 12	3.30 ± 0.37	135 ± 12

[a] Average distances (in Å) and angles (in deg) between principal hydrogen bond acceptors and donors within the catalytic triad (± standard deviations). Reference X-ray values are taken from the HNE:OMTKY3 complex [23] and unbound PR3.[19] Convention for atoms labeling is described on Figure 2.

Table 3. LIE Binding Energies[a]

	sequence							energy terms				
peptide name	P4	P3	P2	P1	P1′	P2′	P3′	vdW	Elec	ΔG_{bind}	k_{cat}/K_m	ref
Pb1	Ala	Ala	Ala	Ala	Ala	-	-	−26.2	−31.9	−26.8	9	21
t-Pns	Ala	Ala	Pro	Val	Gly	-	-	−30.3	−32.3	−29.0	360	21
Pb22	Ala	Ala	Ala	Ala	Lys	Gly	Asp	−44.5	−94.5	−62.9	2600	21
t-Ps[b]	Val	Ala	Asp	Cys	Gly	-	-	−25.6	−101.4	−63.5	10900[c]	22
Pns	Ala	Ala	Pro	Val	Lys	Gly	Asp	−44.9	−101.0	−65.9	62000	21
Ps[b]	Val	Ala	Asp	Cys	Ala	Asp	Arg	−37.6	−291.0	−164.3	651000	22
Popt	Val	Ala	Asp	Val	Lys	Asp	Arg	--49.1	−355.7	−177.5	3400000 ± 200	

[a] ΔG_{bind} values are given in kcal mol^{-1} as well as their van der Waals (vdW) and electrostatic (Elec) contributions (interaction energies) for each PR3:peptide complex. The corresponding values of k_{cat}/K_m (in M^{-1} s^{-1}) are also reported. [b] Peptides with a nonzero net charge ($\beta = 0.43$). [c] Experimental values are reported for Val-Ala-Asp-Cys-*Ala*.

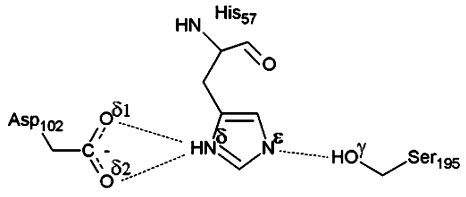

Figure 2. Hydrogen bond network between amino acids constituting the catalytic triad: His57, Asp102, and Ser195. Atom labels are the same as the ones used in Table 2.

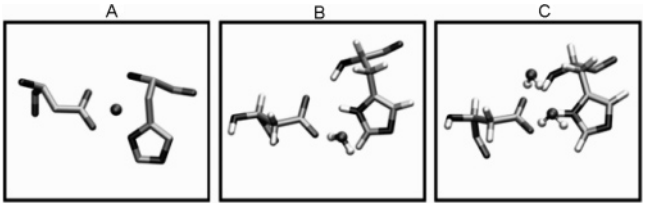

Figure 3. Solvent accessibility of the active site of PR3. X-ray structure of PR3 (A) and snapshots along the MD simulations of PR3:Ps (B) and PR3:t-Pns (C). For clarity, only His57, Asp102, and the water molecules interacting with the carboxyl group of Asp102 for more than 200 ps are represented.

molecules make close contact with the side chains of Ser195 or His57 during the simulation, unlike with Asp102. Asp102 is generally more solvent accessible in complexes with PR3 than with HNE (except in the complexes with Pns). Fujinaga et al. pointed out that Asp102 is more solvent accessible in unbound PR3 than in the complex HNE:OMTKY3; no water molecule is present around the carboxy group of Asp102 of HNE in the complex with OMTKY3 (1PPF), whereas one lies close to Asp102 of chain A in PR3 (cf. Figure 3A). This water molecule, numbered 196 in 1FUJ, is present around Asp102 in all MD simulations of complexes with PR3 (Figure 3B), except the one with Pns. The active sites of PR3 and HNE are more solvent accessible in the simulations with t-Ps and t-Pns than in the simulations with the long peptides. In particular, in PR3:t-Pns, hydrogen bonds between two water molecules and the carboxy group of Asp102 are observed (cf. Figure 3C).

c. LIE Binding Energies. The LIE method was used to evaluate the free energy of binding (ΔG^{LIE}_{bind}) of the complexes between PR3 and the peptides listed in Table 1. ΔG^{LIE}_{bind} values are presented in Table 3 in increasing order of affinity, together with the experimental values of the specificity constants k_{cat}/K_m, reported for comparable peptides. The peptides can be divided in three groups of ligands having comparable values of ΔG^{LIE}_{bind}: (i) Pb1 (−26.8 kcal/mol) and t-Pns (−29.0 kcal/mol), (ii) Pb22 (−62.9 kcal/mol), t-Ps (−63.5 kcal/mol), and Pns (−65.9 kcal/mol), and (iii) Ps (−164.3 kcal/mol).

Ps and Pns have the lowest free energy of binding for PR3, with Ps (−164.30 kcal/mol) having a much higher affinity than Pns (−65.9 kcal/mol). The two peptides have sequences that differ in many aspects (cf. Table 1). Both contain P′ charged amino acids, but they are distributed differently in the sequence. The P1 residue in Ps is polar (Cys) while it is a small hydrophobic amino acid (Val) in Pns. An additional difference is the presence of a negatively charged P2 residue (Asp) in Ps where Pns contains a Pro. The large difference in free energy of binding indicates that at least one of these sites is extremely important for the affinity to PR3.

The truncated versions of Ps and Pns which lack the P′ sites, t-Ps and t-Pns, have much lower affinities than the long peptides. The values of ΔG^{LIE}_{bind} are indeed decreased in the absence of

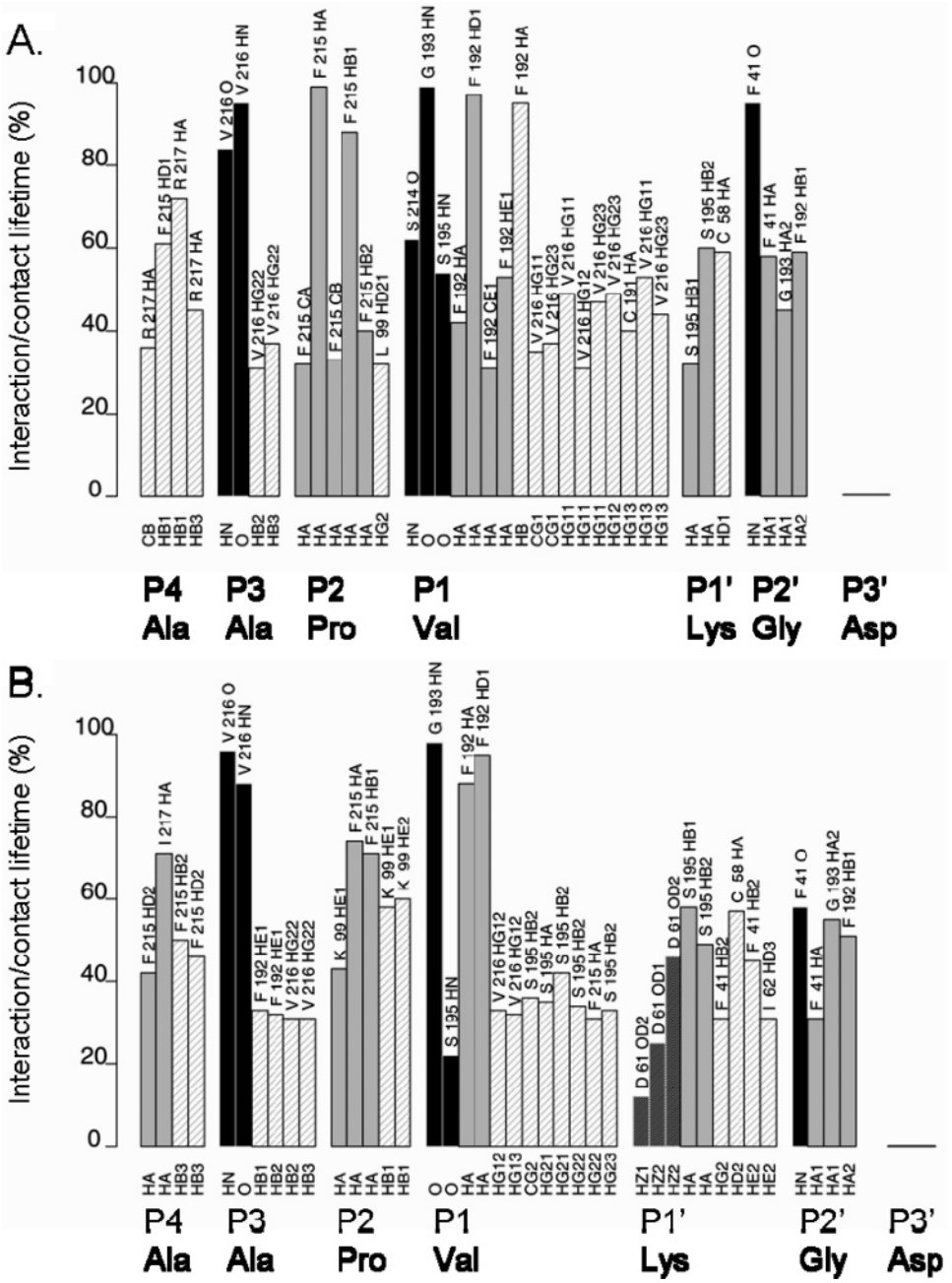

Figure 4. Interaction plots of HNE and PR3 with Pns. Hydrogen bonds and hydrophobic contacts between (A) HNE and Pns and (B) PR3 and Pns: atoms involved (on *X*-axis) and interactions lifetime (on *Y*-axis). Hydrogen bonds are represented by black bars; gray bars represent the lifetime of hydrophobic contacts. Hashed bars signify interactions involving side chains of the peptide, while nonhashed are used for interactions involving the backbone of the peptide.

the P′ tails (−63.5 and −29.0 kcal/mol, respectively). Even though our calculations of the binding affinities do not take into account an entropic penalty, the large difference between the truncated and the long peptides reveal the importance of the P′ sites in the binding affinity. The difference between t-Pns and t-Ps can be explained by the different nature of the P2 residue (Asp more favorable than Pro). This is also confirmed by the low affinity of Pb1, a sequence consisting only of alanines. Adding the same P′ tail to Pb1 as the one from Pns decreases the free energy of binding, confirming the importance of the role played by charged residues in the P′−S′ interactions between PR3 and its ligands.

The ranking of the free energy of binding agrees well with the ranking of catalytic efficiencies, suggesting an important role of the enzyme−peptide binding constants (*K*s) in the catalytic efficiency.

d. Peptide−Enzyme Interactions. We summarize below the interactions, hydrogen bonds, and hydrophobic 'contacts' between the two enzymes and four peptides observed during the MD simulations. Figure 4A,B, Figure 5A,B, and Figure 6A,B display the lifetime during the simulations of the ligand−enzyme interactions in six complexes: HNE:Pns, PR3:Pns, HNE:Ps, PR3:Ps, PR3:t-Pns, and PR3:t-Ps, respectively. The Schechter and Berger convention (cf. Figure 1C) is used to label the interaction sites. Parts A−D of Figure 7 show the structures obtained at the end of the simulations for PR3:Pns, PR3:Ps, HNE:Pns, and HNE:Ps.

(i) Pns Interacts Similarly with HNE and PR3. Except for P3′ (Asp), all residues of the peptide were found to have significantly long-lasting interactions with the enzymes (cf. Figure 4). Interactions between Pns and both enzymes show the same trend: they are dominated by hydrophobic interactions

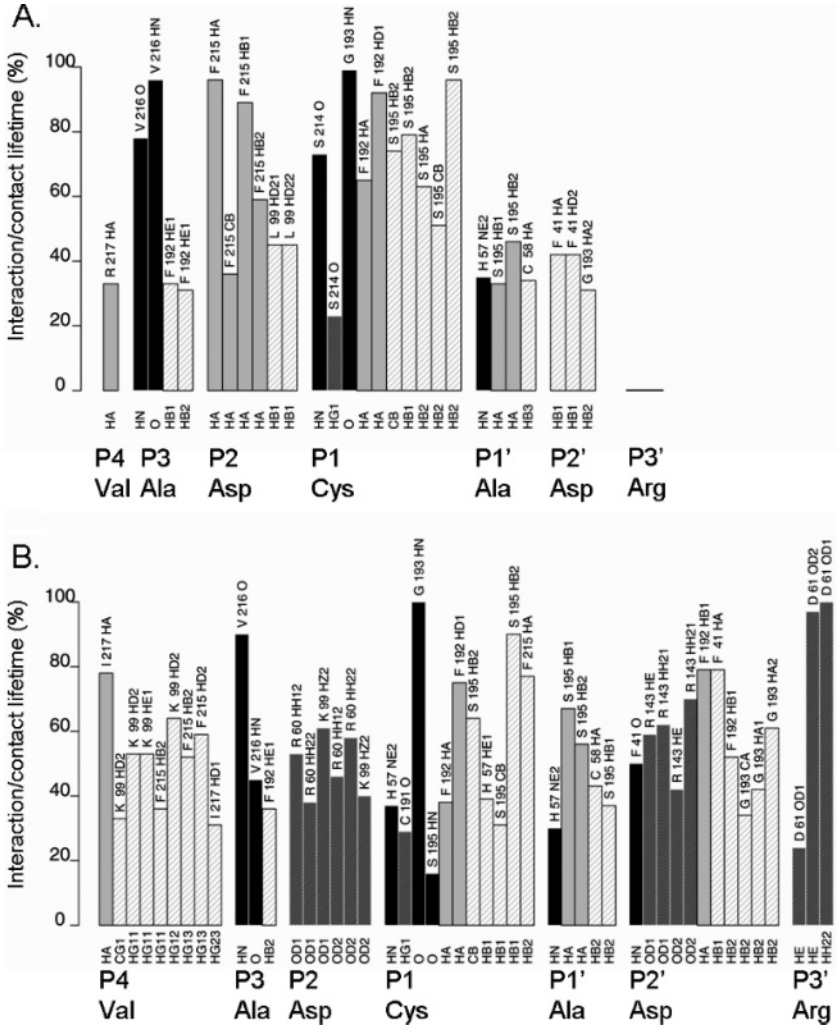

Figure 5. Interaction plots of HNE and PR3 with Ps. Hydrogen bonds and hydrophobic contacts between (A) HNE and Ps and (B) PR3 and Ps. See Figure 4 for legend.

(Figure 4A, gray bars). Only a few hydrogen bonds (black bars) are observed, and all of them are observed between the backbone of the enzyme and the backbone of the peptide (plain black bars).

P1−S1. Except for the interactions stabilizing the oxyanion hole, P1−S1 interactions are predominantly hydrophobic in both complexes (gray bars). The backbone nitrogen atoms of both Ser195 and Gly193 make hydrogen bonds (plain black bars) with the backbone oxygen of P1−Val. The backbone carbonyl group of Ser214 makes a hydrogen bond with the amino group of P1−Val, only in the case of HNE. In HNE, the S1 pocket is mostly occupied by Val216 which has the highest number of contacts through its hydrophobic side chain with the side chain of P1−Val. The number of close contacts between Val216 and P1−Val is significantly lower in PR3:Pns; some of them are replaced by contacts between P1−Val and the C^{β} group of Ser195. In both PR3:Ps and PR3:Pns, the aromatic ring of Phe192 is close to the Cα of the P1 residue.

P2 to P4 (N-Terminus). Most interactions in this area are hydrophobic (gray bars), as expected from the nature of the sequence of Pns (P4−Ala, P3−Ala, P2−Pro). Only hydrophobic contacts were found between the P2−Pro and the enzymes; most of them through its $C^{\alpha}H^{\alpha}$ group with Phe215. Side chains of residue 99 ($C^{\delta}H^{\delta}{}_3$ Leu in HNE, $C^{\epsilon}H^{\epsilon}{}_2$ Lys in PR3) are also in the vicinity and contribute to the hydrophobicity of S2. The proximity of Lys99 suggests that S2 in PR3 could accommodate a polar residue. Val216 is present in S3 for both complexes. Its side chain is involved in hydrophobic interactions with the alanine side chain (gray hashed bars), while the backbone oxygen and nitrogen atoms make durable hydrogen bonds with the backbone nitrogen and oxygen atoms of P3−Ala (plain black bars), respectively. The hydrophobic S4 subsite consists of the side chains and H^{α} atoms of residues 217 (Arg in HNE, Ile in PR3) and 215 (Phe in both enzymes).

P1′ to P3′ (C-Terminus). Only PR3 can provide a polar partner to P1′−Lys, i.e., Asp 61 which through its carboxyl group makes durable hydrogen bonds with the lysine ammonium group. Residue 61 in HNE is a serine. The $C^{\beta}H^{\beta}{}_2$ group of the essential serine (195) contacts the aliphatic chain of P1′−Lys in both complexes. The P2′ residue in Pns is a glycine (Gly6), and we do not detect many interactions with HNE or with PR3. In both cases, only Phe41, Gly193, and Phe192 participate in the S2′ pocket. The most important feature is the hydrogen bond between the backbone amino group of Gly and the carbonyl of Phe41, which is conserved in both enzymes. Neither HNE nor PR3 provide strong interactions with the polar charged side chain of Asp7, which suggests that none of the enzymes can accommodate a negatively charged residue at the P3′ position.

(ii) Differential Interaction of Ps with HNE and PR3. Figure 5 summarizes enzyme−peptide interactions revealed by the simulations of HNE:Ps (A) and PR3:Ps (B). The number of hydrogen bonds (plain and dashed black bars) is significantly higher with PR3 than with HNE, and higher than in PR3:Pns.

P1. The oxyanion hole interactions are similar in PR3:Ps and PR3:Pns (Gly193−O lifetime 98%, Ser195−N < 20%). The hydrogen bond between Ser195−N and P1−O is not seen during the simulation of HNE:Ps. The interaction with the backbone oxygen atom of Ser214 is similar in HNE:Ps and HNE:Pns and is not observed between PR3 and Ps. In both cases, Val216 disappears from S1; instead the aliphatic groups of the cysteine side chain are in contact with S195 (CH_2) and Phe192 in both complexes. This can be explained by the fact that Cys makes the S1 pocket more polar than a valine; it illustrates the adaptability and the plasticity of the recognition sites. The sulfhydryl group makes a hydrogen bond with Ser214−O in HNE and Cys191−O in PR3. Interestingly, the polar Asp213 in PR3 is not seen to interact with the $S^{\gamma}H^{\gamma}$ group. His57−$N^{\epsilon}2$ is involved in a hydrogen bond with the amide group of P1−Cys, thus creating an interaction competing with the important His57−Ser195 hydrogen bond.

P2 to P4. The main difference between HNE:Ps and PR3:Ps lies in the S2 pocket. The carboxyl group of P2−Asp interacts through several hydrogen bonds with Arg60 and Lys99. The two corresponding residues in the sequence of HNE are Ala and Leu, respectively. As a consequence, no polar interaction is found between P2−Asp and HNE. P3−S3 interactions are similar in both complexes and resemble the ones found with Pns. S4 is again mostly hydrophobic as with Ps; of importance is the presence of the aliphatic groups ($C^{\delta}H^{\delta}{}_2$−$C^{\epsilon}H^{\epsilon}{}_2$) of Lys99 in PR3:Ps. The number of hydrophobic contacts between the enzyme and the side chain of P4−Val is higher in the complex with PR3 than with HNE.

P1′ to P3′. Although P1′−Ala is accommodated in the same way by both enzymes, the S2′ and S3′ pockets are clearly different in HNE and PR3. P2′−Asp finds a polar partner in PR3, i.e., Arg143, while P3′−Arg6 makes durable hydrogen bonds with the carbonyl group Asp61. No equivalent interactions are found in HNE where residues 143 and 61 are a leucine and an asparagine, respectively.

In HNE:Ps, the polar side chains of P2′ and P3′ are solvent exposed, and high fluctuations are observed during the simulations. After only 80 ps, P3′−Arg moves away from its initial conformation to interact with P2′−Asp, P1′−Ala, and strikingly even with the backbone and side chain of the P2−Asp as is illustrated by Figure 7D.

(iii) Simulation of Short Peptides with PR3, t-Pns, and t-Ps. Analysis of the contacts and hydrogen bonds in PR3:t-Pns and PR3:t-Ps is presented in parts A and B of Figure 6, respectively.

t-Pns. The pockets S4, S3, and S2 are similar in HNE:t-Pns and HNE:Pns. However the interaction between the enzyme and P1 is quite different; the hydrogen bond of Ser195HN with P1−O (oxyanion hole) disappears. Val216 and Cys191 move away from the vicinity of P1−Val, but the loss of these hydrophobic contacts is balanced by a gain of contacts with Ser195, Asp194, and Phe215. In the complex PR3:t-Pns, the number of hydrophobic contacts decreases slightly, but the same residues constitute the S4, S3, and S2 subsites as in PR3:Pns. Ile190 appears in S1, and its side chain contacts the side chain of P1−Val. The same hydrogen bond as the one found in all complexes with HNE is found between the backbones of P1 (NH) and Ser214 (O). In both complexes, the loss of the P′−S′ interactions influences almost exclusively S1 and not the other recognition subsites.

t-Ps. The number of interactions between PR3 and t-Ps is significantly lower than between PR3 and the corresponding residues of Ps (P4 to P1). Lys99 and Arg60 are absent from S4 and S2, respectively. The contribution of Lys99 to interactions with P2−Asp is increased while Arg60 is solvent exposed; its neighbor in the sequence, Asp61, has no interaction partner in the peptide, and both residues thus move away from the recognition sites. The P1−S1 pockets are slightly modified. A new hydrogen bond is observed between Ser214−O and Cys−N, the hydrogen bond between Ser195 and Cys−O (oxyanion hole) is not observed, while the hydrogen bond with His57−N^{ϵ} is not observed either. Phe215 is not present in S1. Clearly, the lack of strong and polar P′−S′ interactions results in significant modifications of the S recognition subsites.

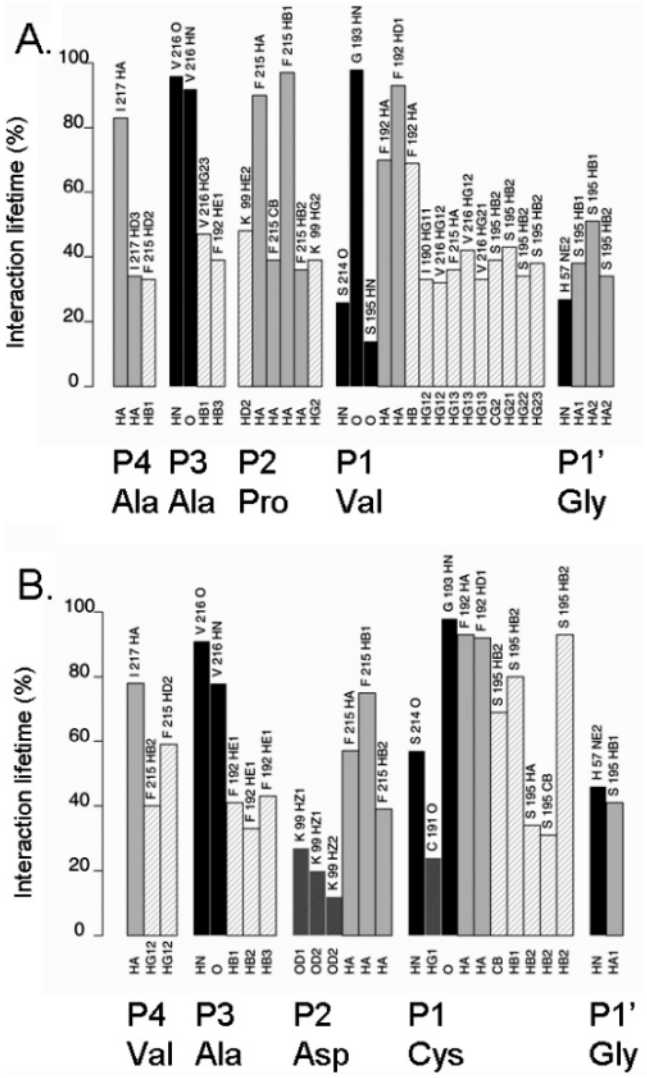

Figure 6. Interaction plots of PR3 with t-Pns and t-Ps. Hydrogen bonds and hydrophobic contacts between (A) PR3 and t-Pns and (B) PR3 and t-Ps. See Figure 4 for legend.

III. Discussion

The conformations of HNE collected during the simulations of the complexes were found to be structurally close to the reference X-ray structure (HNE:OMTKY3). The low RMS deviation and the qualitatively good agreement between theoretical and experimental B-factors establish the reliability of our simulation procedure. The structure of bound PR3 remains close to the X-ray structure of the unbound enzyme; the presence of the peptide does not cause major structural rearrangements in PR3.

Fujinaga et al.[19] reported that Asp102 is more solvent accessible in the unbound PR3 than in HNE because of the lack

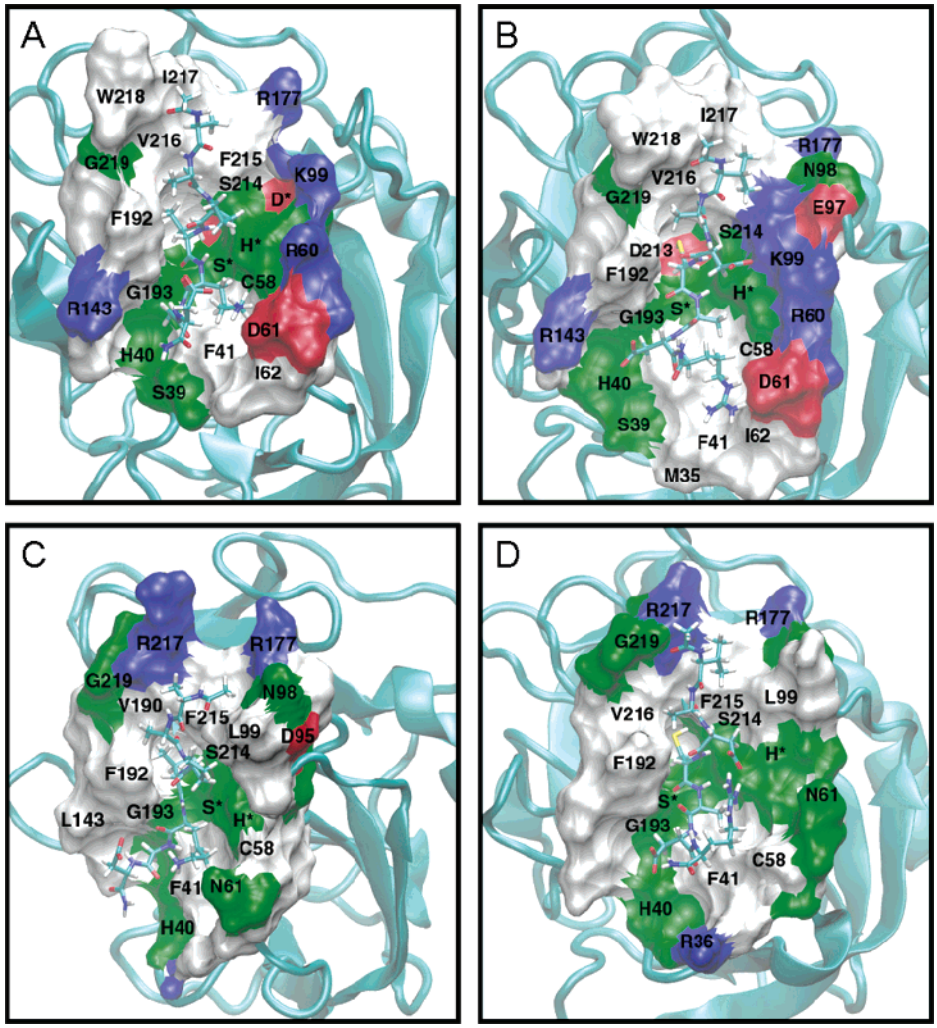

Figure 7. Structure of the enzyme−peptide complexes; details of the interaction surface (A) PR3−Pns, (B) PR3−Ps, (C) HNE-Pns, (D) HNE-Ps. The last conformation of each simulation is displayed. Atoms of the peptides are colored according to their chemical type (blue for nitrogen, red for oxygen, cyan for carbon, yellow for sulfur, and white for hydrogen), enzymes are colored in cyan, and the cartoon representation is used to highlight their secondary structures. The surface of the amino acids of the enzymes located within a distance of 8 Å from the peptide are displayed (representation of their solvent accessible surface calculated in the absence of the peptide) and colored according to their character: basic (blue), acidic (red), polar (green), or nonpolar (white). Residues are labeled following the numbering of chymotrypsin; residues of the catalytic triad are labeled H*, D*, and S*.

of branching at the γ position of residue 99 (Lys instead of Leu), thus opening access to Asp102. Our simulations show that this is also the case when PR3 is bound to a peptide in the Michaelis complex; more water molecules are found to be hydrogen bonded to Asp102 in complexes with PR3 than with HNE. The increased solvent accessibility of Asp102 in PR3 compared to HNE is also explained by the increased polar character of the surface residues surrounding the active site in PR3. The corresponding region is much more hydrophobic in the case of HNE (cf. Figure 7). This might have an effect on the catalytic efficiency of PR3 since solvent accessibility of the active site is important during the deacylation step. Indeed, this second step of the proteolysis involves the nucleophilic attack of a water molecule on the acylenzyme intermediate.

All complexes display a hydrogen bond network in good agreement with what is usually postulated for Michaelis complexes of chymotrypsin-like serine proteases.[11,13] Interestingly, even if Ps is not found to be efficiently cleaved by HNE,[22] the hydrogen bonds between amino acids of the catalytic triad of elastase in the HNE:Ps complex are strong and durable during the MD simulation. The reason for the lack of detectable hydrolysis of Ps by HNE (cf. Table 1) can thus not be explained by a perturbed geometry of the catalytic site that would impair its function. It instead suggests an important role for the recognition subsites.

The calculated free energies of binding reflect the importance of the P′−S′ interactions in the formation of the Michaelis complexes of PR3. The fact that the peptide with the highest k_{cat}/K_m constant also has the highest binding affinity to PR3 demonstrates the importance of the enzyme−ligand interactions, and especially in the S′ subsites, for the catalytic efficiency. It strongly suggests that these interactions play an important role in the first step of the reaction (the acylation is the rate-determining step) most probably by stabilizing the transition state leading to the formation of the tetrahedral intermediate. These interactions are absent in the deacylation step where the C-terminal end of the peptide has been released.

Although the importance of the S′ subsites in discriminating between PR3 and HNE has been revealed by experimental work, no structural study dedicated to the interactions between PR3 and peptidic substrates has been reported; there exists no X-ray structure of PR3 bound to a substrate. The information available is summarized in Figure 1C and stems mostly from the X-ray structure of PR3 unbound. The analysis of our MD simulations

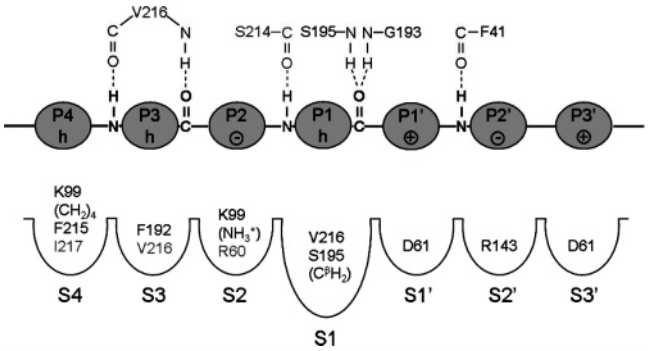

Figure 8. Consensus specific sequence of PR3 and map of the S subsites of PR3. Preferred types of amino acids at the P sites (h = hydrophobic, − = negatively charged, + = positively charged). Inventory of the principal amino acids constituting the recognition subsites of the enzyme; residues interacting with side chains of the peptide are listed below the peptide, and backbone−backbone interactions are sketched above it.

in term of hydrophobic contacts and hydrogen bonds between the enzymes and different peptides allows us to draw a detailed map of the recognition sites and highlight their plasticity (Figure 8). The comparison of the results obtained with HNE and PR3 leads to a better understanding of the basis for the specificity difference between the two enzymes. The cysteine residue in Ps (and t-Ps) interacts with the carbonyl group of residue 191 in PR3 and of Ser214 in HNE, but there is no evidence of an interaction with another polar side chain. In particular, Cys4 (P1) does not interact with Asp213, which is located in the S1 pocket. The simulations suggest that a polar amino acid in S1 is not something that PR3 can accommodate better than HNE, and relatedly Asp213 does not seem to have a role in specificity, in agreement with what was suggested by Fujinaga et al.[19] Ile190 of PR3 is not observed to be in close contact with the P1 side chain, neither when P1 is a valine (Pns, t-Pns) nor an alanine (Pb1, Pb22). The same applies to HNE and Val190. Instead, Val216, conserved in both enzymes, constitutes the main partner in the S1 site, together with the $C^{\beta}H_2$ group of Ser195 (cf. Figure 7). The S4 and S3 subsites are hydrophobic in PR3, as they are in HNE. In our simulations of PR3 complexes, depending on the peptide simulated, three different residues participate in S4: Phe215, Ile217, and the aliphatic part of the side chain of Lys99. Similarly, different residues are found in S3: Phe192 and Val216 but the latter only when the residue in P2 is hydrophobic such as a proline (Pns and t-Pns) or an alanine (Pb1 and Pb22) (not in the case of an Asp in P2 as in Ps, t-Ps or Popt). S2 in PR3 is clearly suited to accommodate a negatively charged amino acid, such as Asp, with Lys99 and Arg60 as potential hydrogen bond donors (cf. Figure 7B). Instead a leucine and an alanine are at position 99 and 60 in HNE. The P2 residue of the ligands is thus crucial in discriminating between PR3 and HNE. It is interesting to note that Lys99 contributes to both the hydrophobic S4 subsite and to the polar S2 subsite, playing two different roles. S1′ and S3′ are polar pockets made of, among others, Asp61. It participates to the two interaction subsites and interacts with lysine or arginine side chains. This is in line with the overlap of S1′ and S3′ described elsewhere.[11] S2′ is characterized by Arg143 which can interact with a negatively charged amino acid at the P2′ position of the sequence. In the sequence of HNE, residue 61 is an asparagine and residue 143 is a leucine. S1′, S2′, and S3′ subsites explain most of the specificity difference between HNE and PR3; this is in agreement with what has been reported by other groups.[19,22] Interestingly, a negatively charged residue in the P3′ position (like in Pns) does not find a polar partner in PR3 (neither in HNE) and remains solvent exposed. There is no overlap between S2′ and S1′ or S2′ and S3′. The P′ residues should be polar charged residues, but the sign of the charges and their position is important: P1′ should be positive, P2′ negative, and P3′ positive. Interestingly the lack of P′−S′ interactions also decreases the number of P−S interactions, as we observed when truncating Ps and Pns. Altering or replacing an amino acid at one end of the substrate is thus susceptible of modifying interactions with more than one recognition subsite and in some cases distant ones. The recognition sites are not independent of each other, and this can explain the complexity of interpreting site-directed mutagenesis experiments. Additionally, the simulations show that different amino acids of the enzyme can alternatively participate to a given recognition subsite depending on the nature of the corresponding amino acid in the ligand. This illustrates the adaptability of the enzyme; the recognition subsites are not rigid predefined pockets.

Based on the LIE binding affinities and our inventory of the amino acids constituting the enzyme recognition sites, we designed a peptide, named Popt, which should be highly specific for PR3. The sequence of Popt is Val-Ala-Asp-Val-Lys-Asp-Arg and was chosen to satisfy several criteria (cf. Figure 9): (i) hydrophobic amino acids at P4 and P3, (ii) negative group (e.g. Asp) at P2, (iii) small hydrophobic group at P1, (iv) positive groups (hydrogen bond donors) at P1′ and P3′, and (v) negative group at P2′. Simulations of complexes of PR3 and HNE with Popt were then performed following the same methodology as for the other complexes. The LIE binding energy was calculated and found equal to −177.47 kcal/mol. This reveals a slightly better affinity for the enzyme than Ps, without having a Cys at P1. The inventory of the interactions of Popt with PR3 and HNE is reported in Figure 9, the amino acids involved in the main interactions observed between PR3 and Popt are represented in Figure 10. It confirms what was expected from the previous simulations. In particular, P2−Asp interacts with Lys99 in PR3 but does not find any hydrogen bond partner in HNE. P1′−Lys and P3′−Arg make salt bridges with D61 of PR3; P3′−Arg is not found to interact with any residue of HNE, while P1′−Lys is hydrogen bonded to Asn61 (side chain and backbone oxygen) of HNE. The side chain of Arg143 of PR3 makes long lasting hydrogen bonds with the carboxyl group of P2′−Asp, which has no similar partner in HNE.

We verified our prediction with enzymatic assays. The consensus sequence VADVKDR was introduced as the peptidyl moiety of a fluorogenic substrate with intramolecularly quenched fluorescence. The corresponding substrate (Abz-VADVKDRQ-EDDnp) was rapidly cleaved by PR3, but not by HNE, and the cleavage site was identified by mass spectrometry as the VK bond. The specificity constant was 3400 ± 200 mM^{-1} s^{-1} (cf. Table 3) which compares with those obtained with the best substrates we have tested under the same experimental conditions (Korkmaz et al. unpublished data). It thus validates the prediction from the computations.

Additionally, the pattern of the sequence proposed here corresponds to the sequence of the cleavage site of P21waf1 by PR3 that we have recently characterized.[8] The sequence around the cleavage site of P21waf1 is IQEARER. The protein is cleaved between A and R. With the exception of the P3 residue (Gln), the sequence complies with the criteria we found from our MD simulations: it contains a hydrophobic P4 (Ile), a negatively charged P2 (Glu), a small hydrophobic P1 (Ala), a positively charged P1′ (Arg), a negatively charged P2′ (Glu), and a positively charged P3′ (Arg).

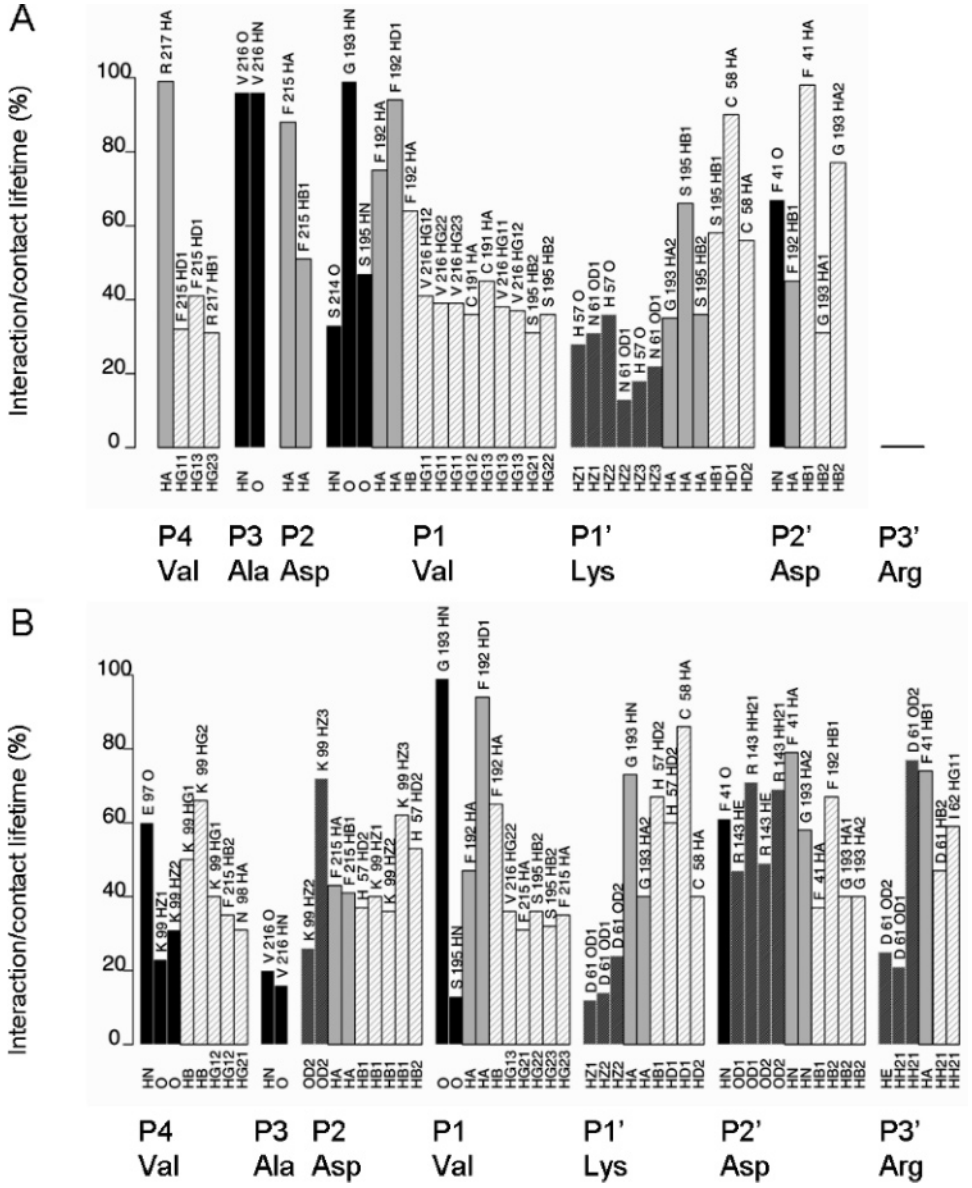

Figure 9. Hydrogen bonds and hydrophobic contacts between Popt (VADVKDR) and HNE (A) or PR3 (B). See Figure 4 for legend.

In conclusion, our simulations provide for the first time structures of complexes formed by PR3 and different peptides, specific or nonspecific for this enzyme. The network of hydrogen bonds between amino acids of the catalytic triad is characteristic of a Michaelis complex. The important role played by the interaction sites, and in particular by the C-terminal part of the peptides as suggested earlier,[22] is confirmed. We provide a dynamic map of the S subsites of the enzyme and a description of the interaction scheme for six different ligands. The main difference between HNE and PR3 concerns four different subsites: S2, S1′, S2′, and S3′ which all are more polar than in HNE (cf. Figure 7). The surface of PR3 around the active site is thus more polar than in HNE and explains the increased solvent accessibility of the aspartic acid (D102) of the catalytic triad in PR3 unbound and in Michaelis complexes. The seven pockets are shown to be interconnected; in particular, S4 and S2 share Lys99 which plays both a hydrophobic and a hydrophilic role, S1′ and S3′ overlap and share Asp61, and the lack of P′−S′ interaction modifies the network of S−P interactions. This is an important result in the context of drug design and investigation of enzyme−ligand binding sites by directed mutagenesis experiments. We also highlighted the plasticity of the pockets. Both the interdependence and the adaptability of the recognition pockets should be taken into consideration when designing new ligands.

The second important outcome of this work is the design of a peptide highly specific for PR3, containing a small hydrophobic amino acid at the P1 position. MD simulations of this peptide−VADVKDR−bound to PR3 and HNE confirmed the proposed interaction scheme, thus establishing the predictive power of our approach. Enzymatic assays confirmed its high specificity for PR3 vs HNE. Additionally, the pattern of amino acids in this sequence is similar to that of the cleavage region of P21waf1.[8] This work has therefore a potential in finding new intra- and extracellular targets of PR3, in addition to providing valuable information for the design of drug-like organic ligands. Indeed, the requirements we describe in terms of type and position of amino acids for a specific peptidic substrate can be expressed in terms of chemical characteristics of substituents for an organic ligand. In particular, it appears unlikely that a small molecule interacting only with the S1 site would be able to discriminate between PR3 and HNE. Optimally, it would need to have substituents containing a negatively charged group interacting with S2 and substituents mimicking the alternating

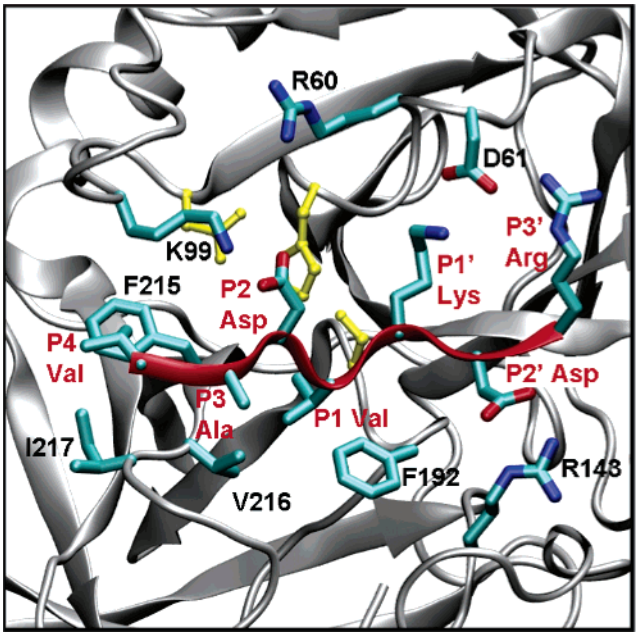

Figure 10. Structure of the complex between PR3 and Popt (VADVKDR): snapshot from the MD simulation. The secondary structure elements of the enzyme are represented using gray cartoons (arrows for extended strands, cylinders for helices). The backbone of the peptide is represented by a red ribbon. Side chains of the amino acids of the peptide (labeled with red characters) and side chains of amino acids constituting the subsites of the enzyme (labeled with black characters) are represented using sticks and colored following atom types (cf. legend Figure 7). The catalytic triad is represented using yellow balls and sticks.

charged pattern (+ − +) to optimize the interactions with the S1′ to S3′ sites.

IV. Methods

a. Peptide Sequences. Amino acid sequences of the simulated peptides were chosen from recent experimental studies reporting k_{cat}/K_m values for both HNE and PR3 with peptides of comparable sequences.[20−22] Our set of peptides included specific and nonspecific sequences of PR3 as well as sequences with and without P′ extension in order to investigate the influence of the S′−P′ interactions. Names and sequences of all peptides simulated are listed in Table 1, together with the k_{cat}/K_m values reported in references 20−22. Various terminal groups have been used in the enzymatic assays.[20−22] However, in our simulations we used simpler chemical groups—acetylated N-terminus and amidated C-terminal ends—but similar for all peptides in order to limit the sequence variation to the amino acids in P4 to P3′ positions.

Pns and t-Pns. The sequence AAPVKGD (referred to as 'nonspecific peptide' or Pns)[21] is not specific for PR3, with specificity constants of 179 000 and 62 000 for HNE and PR3, respectively. The use of a peptide with the same sequence truncated after P1, i.e., AAPVG (t-Pns), causes a 172-fold decrease of the specificity constant with PR3.

Ps and t-Ps. VADCADR (Ps)[22] is a sequence specific for PR3. It contains a Cys residue at the P1 position and is referred to as 'specific peptide' (or Ps) in this manuscript. We also used a peptide with the same sequence truncated after P1 (peptide referred to as 'truncated Ps' or t-Ps). Specificity constants for both HNE and PR3 with a peptide having a slightly different sequence (VADCA) have been reported by Korkmaz et al.[22]

Pb1 and Pb22. Two additional peptides were studied, AAAAA (Pb1) and AAAAKGD (Pb22), which have the same sequence as Pb1 for the P region and as Pns for the P′ region. The first one is more efficiently cleaved by HNE than by PR3, while the latter has a k_{cat}/K_m of the same range for both enzymes.

b. Starting Atomic Structures. Complexes. The initial set of coordinates used for the MD simulations was prepared from the X-ray structures of PR3 (PDB ID 1FUJ, resolution 2.2 Å)[19] and HNE complexed with OMTKY3 (third domain of the turkey ovomucoid inhibitor) (PDB ID 1PPF, resolution 1.8 Å).[23] OMTKY3 is a 56 residues long protein which interacts via its Leu18 with the S1 site of HNE. Three histidines of HNE were protonated on their δ nitrogen atom (25, 40, 57) and two on the ϵ nitrogen (71 and 210). The choice was based on visual inspection of the surrounding of each residue. The complexes between HNE and the different peptides were prepared by replacing the P4 to P3′ residues (P4 Ala15 − P3 Cys16 − P2 Thr17 − P1 Leu18 − P′1 Glu19 − P′2 Tyr20 − P′3 Arg21) of OMTKY3 by the sequences of our ligands. The remaining residues of OMTKY3 were then removed, and the N-terminus of each peptide was acetylated while their C-terminal end was amidated.

In the case of PR3, chains A to D in the PDB file are extremely similar (rmsd: 0.4 Å) with B-factors of comparable magnitude, and therefore the choice of the chain should not influence the outcome of the MD simulations. We chose to use chain A. The monomer contains 221 residues. There are four disulfide bridges, between cysteines 42−58, 136−201, 168−182, and 191−220. Histidine residues numbered 24, 48, 132 were protonated at their Nϵ nitrogen atom, and the others (20, 40, 57, 71, 82, 147, and 179) at their Nδ nitrogen atom. The complexes were built by superposing PR3 on HNE of the HNE:OMTKY3 complex (rmsd on backbone atoms: 2.43 Å), deleting HNE and then replacing OMTKY3 residues such as it had been done for the complexes with HNE.

Only water molecules from the X-ray data close to the enzymes (within 8 Å distance from the enzyme) and not overlapping with the enzyme or the peptide were kept (distance to the complex greater than 2.5 Å). The aspartic acid present in the S1 pocket of PR3, Asp213, was chosen to be protonated as suggested by Fujinaga et al.[19] It was confirmed by pK_a calculations on the enzyme with and without the peptides, using the UHBD program[29] and the PKALK package[30] (E.Hajjar, unpublished results). Otherwise, the coordinates of the missing hydrogen atoms were determined using the HBUILD[31] facility in Charmm, at pH 7. As a result, the total charge of PR3 was equal to +2, and the total charge of HNE was equal to +11. To avoid simulations of non-neutral systems, ions were added to the enzyme−peptide complexes that had a net charge different from zero. The ions were added in the boxes in a random manner but at a significant distance from the complexes (>8 Å); no ions came in contact with the enzymes or peptides during the course of the simulations.

Peptides in Water. To calculate LIE binding energies, simulations of the ligands in water were performed. The initial coordinates for the peptides were taken from the enzyme−peptide complexes. Counterions were added when necessary to neutralize the system.

c. MD Simulations. The complexes were refined by geometry optimization to remove unfavorable steric contacts caused by manual docking. The Charmm software (v30b1),[32] together with the Charmm27 force field,[33] were used. In the next step, each complex (protein, peptides, and crystal water molecules) was solvated in a cubic box (75 Å) of pre-equilibrated water molecules using VMD.[34] Peptides (for the LIE binding energies) were solvated in smaller boxes (60 Å). The resulting conformations were used as the starting set of coordinates for the MD simulations and have been used as reference structures during the analysis of the trajectories (referred to as initial conformations in the Results section).

All simulations were performed using version 2.5 of the NAMD program,[35] with version 27 of the Charmm force field.[33] The energy of the solvated systems was briefly minimized, and the systems were then brought to 300 K by a stepwise heating (10, 100, 200, and 300 K) using periodic reassignment of the velocities to control the temperature. The systems were examined in the (N,P,T) ensemble, where the equations of motion were integrated using a multiple-time step algorithm;[36,37] bonded forces were evaluated every femtosecond (fs), short-range nonbonded forces every 2 fs, and long-range electrostatics every 4 fs. Langevin dynamics was used to control the pressure (target pressure: 1 atm, oscillation period: 75 fs, damping time scale: 25 fs) and the temperature

(target temp: 300 K, coupling coefficient: 1.0 ps^{-1}) during the whole simulations. SHAKE[38,39] was used to constrain the bond lengths between hydrogen and oxygen atoms of the water molecules. The particle Mesh Ewald algorithm[40,41] was used to estimate the long-range electrostatic forces by interpolating charges on a grid with 80 Å edges for the complexes and 60 Å for the peptides in water. Once the systems had reached 300 K, they were stabilized during 100 ps. During that time, velocities were reassigned periodically (every 10 fs), and one conformation was stored every picosecond. After the equilibration phase, production runs were performed for 2000 ps.

d. Trajectory Analysis. Atomic Fluctuations. The experimental B-factors from X-ray diffraction can be converted to atomic rms fluctuations:[42]

$$\langle r_j^2 \rangle^{1/2} = \sqrt{\frac{3}{8\pi^2} B_j} \quad (1)$$

where $\langle \Delta r_j^2 \rangle^{1/2}$ and B_j represent the root-mean-square fluctuations and the temperature factor of atom j, respectively. The rms atomic fluctuations can also be calculated over an MD simulation. In general, comparison between rms fluctuations obtained from this equation and from MD can only be qualitative.[28,43] It may serve to check whether the MD simulation correctly reproduces the difference in mobility of the different parts of the molecule as given by the crystallographic data. We calculated $\langle \Delta r_j^2 \rangle$ for each heavy atom of the backbone with respect to its average position during the MD simulations. Atomic fluctuations were then averaged (mass-weighted) for each residue.

Hydrophobic Contacts and Hydrogen Bonds. To detect the presence of hydrogen bonds, Charmm makes use of the donor and acceptor definitions in the protein structure file (psf). In our analyses of the trajectories, an acceptor and a donor are considered to interact by a hydrogen bond if the distance between A and D is lower than or equal to 2.4 Å,[44] and if the value of the angle formed by D−H−A is greater than or equal to 130°. Charmm was also used to analyze the hydrophobic contact between amino acids of the enzymes and the substrates. A hydrophobic contact is assumed to exist when two candidate atoms are closer to each other than a cutoff distance (3 Å). We built a list of candidate atoms restricted to aliphatic groups of amino acids side chains (Charmm force field nomenclature: ca; cb; cg1; cg2; ha*; hb*; hg; hg2*; type cg except the one of hsd, hse, asn, tyr, asp; type hg1 except for cys, thr, ser; type cd1 except for tyr; type cd2 except for tyr, hsd, hse; type cay and type hy*). The inventory was restricted by considering the contacts having a total lifetime (sum of the lifetimes of all events) above 30% of the sampling time. The presence of water molecules in the vicinity of the catalytic triad was also investigated (cutoff distance: 3.5 Å and minimum lifetime of a contact: 3 ps).

LIE Binding Energies. The LIE method [45] was used to calculate free energies of binding for all PR3−peptide complexes. This method estimates the binding free energy through a polar and a nonpolar contribution obtained from structural ensembles generated using MD or Monte Carlo simulations. The polar or electrostatic contribution is obtained using linear response theory,[45,46] while the nonpolar contribution is estimated by optimization of a weight coefficient with respect to reproducing experimental data. The free energy of binding of the complex is derived by considering only two states: (1) free ligand in the solvent and (2) ligand bound to the solvated protein. The binding free energy is estimated according to the LIE equation:

$$\Delta G_{bind}^{LIE} = \alpha(\langle V_{l-s}^{vdw} \rangle_{bound} - \langle V_{l-s}^{vdw} \rangle_{free}) + \beta(\langle V_{l-s}^{el} \rangle_{bound} - \langle V_{l-s}^{el} \rangle_{free}) \quad (2)$$

where $\langle V_{l-s} \rangle_{bound}$ and $\langle V_{l-s} \rangle_{free}$ denote the ensemble average of the van der Waals (vdw) and electrostatic (el) ligand-surrounding (l-s) interaction energies when the ligand is in the complex (bound) and in water (free). Average interaction energies have been taken from the ensemble of configurations generated by our MD simulations (sampling windows, i.e., 0.5−2.0 ns). Interaction energies between each substrate (heptapeptide) and its surroundings (PR3, water molecules, ions) have been calculated in every step of the free and bound ligand MD simulations, using the 'pairInteraction' method in NAMD. We have used $\alpha = 0.5$ and $\beta = 0.5$ or 0.43, depending on whether the net charge of the ligand was equal to 0 or not, respectively. These parameters have been used successfully for calculating binding affinities of large peptidic substrates of serine proteases.[47−50]

e. Measurement of k_{cat}/K_m Value (Popt). Enzyme Assays. PR3 and HNE were from Athens Research & Technology (Athens, GA). PR3 activity was measured at 37 °C in 50 mM Hepes, pH 7.4, 0.75 M NaCl, supplemented with 0.05% Igepal CA-630 (v/v) to take into account the great propensity of PR3 to stick to plastic and glass surfaces when in dilute solution. PR3 was titrated with α1-PI as reported before.[20] The synthetic fluorogenic substrate Abz-VADVKDRQ-EDDnp was synthesized as described before,[20] and a stock solution (2 mM) was prepared in 30% (v/v) *N*,*N*-dimethylformamide and diluted to 0.5 mM with 50 mM Hepes buffer pH 7.4. Substrate hydrolysis was followed by measuring the fluorescence at $\lambda_{ex} = 320$ nm and $\lambda_{em} = 420$ nm in a Hitachi F-2000 spectrofluorometer. The specificity constants (k_{cat}/K_m) were determined under first-order conditions using a substrate concentration below the K_m (1 μM). The final PR3 concentration was 10 nM.

Acknowledgment. Funding for N.R and E.H. was provided by the National Program for Research in Functional Genomics in Norway (FUGE) in the Research Council of Norway. Parallab (High Performance Computing Laboratory at the University of Bergen) is thankfully acknowledged for provision of CPU time on its IBM Regatta supercomputer. This work was supported by EGIDE (N.R. and VWS) and by the Association Vaincre la Mucoviscidose (F.G. and V.W.S.) and the "Leg Poix" (V.W.S.). We thank BAXTER (Extramural Grant, V.W.S.) and AMGEN (V.W.S.) for their generous financial support. B.O.B. gratefully acknowledges financial support from the Research Council of Norway. The Norwegian Structural Biology Centre is supported by FUGE. Annick Dejaegere and David Liberles are thanked for careful reading of the manuscript, and Knut Børve is also thanked for fruitful discussions.

Supporting Information Available: Plot representing the rms deviations of the backbone of the enzymes (PR3 and HNE) during the MD simulations of the complexes with Ps and Pns, plot comparing the calculated and experimental averaged atomic fluctuations, and the sequence of OMTKY3. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Owen, C. A.; Campbell, E. J. The cell biology of leukocyte-mediated proteolysis. *J. Leukocyte Biol.* **1999**, *65*, 137−150.
(2) Csernok, E.; Ernst, M.; Schmitt, W.; Bainton, D. F.; Gross, W. L. Activated neutrophils express proteinase 3 on their plasma membrane in vitro and in vivo. *Clin. Exp. Immunol.* **1994**, *95*, 244−250.
(3) Owen, C. A.; Campbell, M. A.; Sannes, P. L.; Boukedes, S. S.; Campbell, E. J. Cell surface-bound elastase and cathepsin G on human neutrophils: a novel, non-oxidative mechanism by which neutrophils focus and preserve catalytic activity of serine proteinases. *J. Cell Biol.* **1995**, *131*, 775−789.
(4) Campbell, E. J.; Campbell, M. A.; Boukedes, S. S.; Owen, C. A. Quantum proteolysis by neutrophils: implications for pulmonary emphysema in alpha 1-antitrypsin deficiency. *J. Clin. Invest.* **1999**, *104*, 337−344.
(5) Witko-Sarsat, V.; Lesavre, P.; Lopez, S.; Bessou, G.; Hieblot, C.; Prum, B.; Noel, L. H.; Guillevin, L.; Ravaud, P.; Sermet-Gaudelus, I.; Timsit, J.; Grunfeld, J. P.; Halbwachs-Mecarelli, L. A large subset of neutrophils expressing membrane proteinase 3 is a risk factor for vasculitis and rheumatoid arthritis. *J. Am. Soc. Nephrol.* **1999**, *10*, 1224−1233.
(6) Witko-Sarsat, V.; Canteloup, S.; Durant, S.; Desdouets, C.; Chabernaud, R.; Lemarchand, P.; Descamps-Latscha, B. Cleavage of p21waf1 by proteinase-3, a myeloid-specific serine protease, potentiates cell proliferation. *J. Biol. Chem.* **2002**, *277*, 47338−47347.

(7) Pederzoli, M.; Kantari, C.; Gausson, V.; Moriceau, S.; Witko-Sarsat, V. Proteinase-3 induces procaspase-3 activation in the absence of apoptosis: potential role of this compartmentalized activation of membrane-associated procaspase-3 in neutrophils. *J. Immunol.* **2005**, *174*, 6381−6390.
(8) Dublet, B.; Ruello, A.; Pederzoli, M.; Hajjar, E.; Courbebaisse, M.; Canteloup, S.; Reuter, N.; Witko-Sarsat, V. Cleavage of p21/WAF1/CIP1 by Proteinase 3 modulates differentiation of a monocytic cell line: Molecular analysis of the cleavage site. *J. Biol. Chem.* **2005**, *280*, 30242−30253.
(9) Molldrem, J. J.; Clave, E.; Jiang, Y. Z.; Mavroudis, D.; Raptis, A.; Hensel, N.; Agarwala, V.; Barrett, A. J. Cytotoxic T lymphocytes specific for a nonpolymorphic proteinase 3 peptide preferentially inhibit chronic myeloid leukemia colony-forming units. *Blood* **1997**, *90*, 2529−2534.
(10) Molldrem, J.; Dermime, S.; Parker, K.; Jiang, Y. Z.; Mavroudis, D.; Hensel, N.; Fukushima, P.; Barrett, A. J. Targeted T-cell therapy for human leukemia: cytotoxic T lymphocytes specific for a peptide derived from proteinase 3 preferentially lyse human myeloid leukemia cells. *Blood* **1996**, 88, 2450−2457.
(11) Hedstrom, L. Serine protease mechanism and specificity. *Chem. Rev.* **2002**, *102*, 4501−4524.
(12) Rawlings, N.; Tolle, D.; AJ, B. MEROPS: the peptidase database. *Nucleic Acids Res.* **2004**, *32* (Database issue), D160−164.
(13) Perona, J. J.; Craik, C. S. Structural basis of substrate specificity in the serine proteases. *Protein Sci.* **1995**, *4*, 337−360.
(14) Czapinska, H.; Otlewski, J. Structural and energetic determinants of the S1-site specificity in serine proteases. *Eur. J. Biochem.* **1999**, *260*, 571−595.
(15) Schechter, I.; Berger, A. On the size of the active site in proteases. I. Papain. *Biochem. Biophys. Res. Commun.* **1967**, *27*, 157−162.
(16) Rao, N. V.; Wehner, N. G.; Marshall, B. C.; Gray, W. R.; Gray, B. H.; Hoidal, J. R. Characterization of proteinase-3 (PR-3), a neutrophil serine proteinase. Structural and functional properties. *J. Biol. Chem.* **1991**, *266*, 9540−9548.
(17) Brubaker, M. J.; Groutas, W. C.; Hoidal, J. R.; Rao, N. V. Human neutrophil proteinase 3: mapping of the substrate binding site using peptidyl thiobenzyl esters. *Biochem. Biophys. Res. Commun.* **1992**, *188*, 1318−1324.
(18) Kam, C. M.; Kerrigan, J. E.; Dolman, K. M.; Goldschmeding, R.; Von dem Borne, A. E.; Powers, J. C. Substrate and inhibitor studies on proteinase 3. *FEBS Lett.* **1992**, *297*, 119−123.
(19) Fujinaga, M.; Chernaia, M. M.; Halenbeck, R.; Koths, K.; James, M. N. The crystal structure of PR3, a neutrophil serine proteinase antigen of Wegener's granulomatosis antibodies. *J. Mol. Biol.* **1996**, *261*, 267−278.
(20) Korkmaz, B.; Attucci, S.; Hazouard, E.; Ferrandiere, M.; Jourdan, M. L.; Brillard-Bourdet, M.; Juliano, L.; Gauthier, F. Discriminating between the activities of human neutrophil elastase and proteinase 3 using serpin-derived fluorogenic substrates. *J. Biol. Chem.* **2002**, *277*, 39074−39081.
(21) Koehl, C.; Knight, G.; Bieth, J. Compared action of neutrophil proteinase 3 and elastase on model substrates. Favorable effect of S′−P′ interactions on proteinase 3 catalysts. *J. Biol. Chem.* **2003**, *278*, 12609−12612.
(22) Korkmaz, B.; Attucci, S.; Moreau, T.; Godat, E.; Juliano, L.; Gauthier, F. Design and use of highly specific substrates of neutrophil elastase and proteinase 3. *Am. J. Respir. Cell Mol. Biol.* **2004**, *30*, 801−807.
(23) Bode, W.; Wei, A.; Huber, R.; Meyer, E.; Travis, J.; Neumann, S. X-ray crystal structure of the complex of human leukocyte elastase (PMN elastase) and the third domain of the turkey ovomucoid inhibitor. *EMBO J.* **1986**, *5*, 2453−2458.
(24) Navia, M.; McKeever, B.; Springer, J.; Lin, T.; Williams, H.; Fluder, E.; Dorn, C.; Hoogsteen, K. Structure of human neutrophil elastase in complex with a peptide chloromethyl ketone inhibitor at 1.84 Å resolution. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 7−11.
(25) Cregge, R.; Durham, S.; Farr, R.; Gallion, S.; Hare, C.; Hoffman, R.; Janusz, M.; Kim, H.; Koehl, J.; Mehdi, S.; Metz, W.; Peet, N.; Pelton, J.; Schreuder, H.; Sunder, S.; Tardif, C. Inhibition of human neutrophil elastase. 4. Design, synthesis, X-ray crystallographic analysis, and structure−activity relationships for a series of P2-modified, orally active peptidyl pentafluoroethyl ketones. *J. Med. Chem.* **1998**, *41*, 2461−2480.
(26) Macdonald, S. J. F.; Dowle, M. D.; Harrison, L. A.; Clarke, G. D. E.; Inglis, G. G. A.; Johnson, M. R.; Shah, P.; Smith, R. A.; Amour, A.; Fleetwood, G.; Humphreys, D. C.; Molloy, C. R.; Dixon, M.; Godward, R. E.; Wonacott, A. J.; Singh, O. M. P.; Hodgson, S. T.; Hardy, G. W. Discovery of further pyrrolidine trans-lactams as inhibitors of human neutrophil elastase ({HNE}) with potential as development candidates and the crystal structure of HNE complexed with an inhibitor (GW475151). *J. Med. Chem.* **2002**, *45*, 3878−3890.
(27) Brandsdal, B. O.; Osterberg, F.; Almlof, M.; Feierberg, I.; Luzhkov, V. B.; Aqvist, J. Free energy calculations and ligand binding. *Adv. Protein Chem.* **2003**, *66*, 123−158.
(28) Brooks, C. L. I.; Karplus, M.; Pettitt, B. M. *Proteins: A Theoretical Perspective of Dynamics, Structure and Thermodynamics*; John Wiley & Sons: New York, 1988.
(29) Briggs, J. M.; Davis, M. E.; Desai, B. H.; Gilson, M. K.; Ilin, A.; Luty, B. A.; McCammon, J. A.; Madura, J. D.; Tan, R. C.; Wade, R. C. Program. *University of Houston Brownian Dynamics* (UHBD), 1993.
(30) Schaefer, M.; Sommer, M.; Karplus, M. pH-dependence of protein stability: Absolute electrostatic free energy differences between conformations. *J. Phys. Chem. B* **1997**, *101*, 1663−1683.
(31) Brunger, A. T.; Karplus, M. Polar hydrogen positions in proteins: Empirical energy placement and neutron diffraction comparison. *Proteins* **1988**, *4*, 148−156.
(32) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. CHARMM: A Program for Macromolecular Energy, Minimization, and Dynamics Calculations. *J. Comput. Chem.* **1983**, *4*, 187−217.
(33) Foloppe, N.; MacKerell, A. D. All-atom empirical force field for nucleic acids: I. Parameter optimization based on small molecule and condensed phase macromolecular target data. *J. Comput. Chem.* **2000**, *21*, 86−104.
(34) Humphrey, W.; Dalke, A.; Schulten, K. VMD − Visual Molecular Dynamics. *J. Mol. Graphics* **1996**, *14*, 33−38.
(35) Kale, L.; Skeel, R.; Bhandarkar, M.; Brunner, R.; Gursoy, A.; Krawetz, N.; Phillips, J.; Shinozaki, A.; Varadarajan, K.; Schulten, K. NAMD2: Greater scalability for parallel molecular dynamics. *J. Comput. Phys.* **1999**, *151*, 283−312.
(36) Martyna, G. J.; Tuckerman, M. E.; Tobias, D. J.; Klein, M. L. Explicit reversible integrators for extended systems dynamics. *Mol. Phys.* **1996**, *87*, 1117−1157.
(37) Izaguirre, J. A.; Reich, S.; Skeel, R. D. Longer time steps for molecular dynamics. *J. Chem. Phys.* **1999**, *110*, 9853−9864.
(38) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. Numerical integration of the Cartesian equations of motion of a system with constraints: Molecular dynamics of n-alkane. *J. Comput. Phys.* **1977**, *23*, 327−341.
(39) Andersen, H. C. Rattle − a Velocity Version of the Shake Algorithm for Molecular-Dynamics Calculations. *J. Comput. Phys.* **1983**, *52*, 24−34.
(40) Darden, T.; York, D.; Pedersen, L. Particle Mesh Ewald − an N.Log(N) Method for Ewald Sums in Large Systems. *J. Chem. Phys.* **1993**, *98*, 10089−10092.
(41) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A smooth particle mesh Ewald method. *J. Chem. Phys.* **1995**, *103*, 8577−8593.
(42) Willis, B. T. M.; Prior, A. W. *Thermal Vibrations in Crystallography*; Cambridge University Press: Cambridge, 1975; pp 1−280.
(43) Stote, R.; Dejaegere, A.; Karplus, M. *Computational Approaches to Biochemical Reactivity*; Kluwer Academic Publisher: 1997; pp 153−198.
(44) Deloof, H.; Nilsson, L.; Rigler, R. Molecular-dynamics simulation of galanin in aqueous and nonaqueous solution. *J. Am. Chem. Soc.* **1992**, *114*, 4028−4035.
(45) Aqvist, J.; Medina, C.; Samuelsson, J. E. A new method for predicting binding affinity in computer-aided drug design. *Protein Eng.* **1994**, *7*, 385−391.
(46) Lee, F. S.; Chu, Z. T.; Bolger, M. B.; Warshel, A. Calculations of antibody−antigen interactions: microscopic and semi-microscopic evaluation of the free energies of binding of phosphorylcholine analogs to McPC603. *Protein Eng.* **1992**, *5*, 215−228.
(47) Brandsdal, B. O.; Aqvist, J.; Smalas, A. O. Computational analysis of binding of P1 variants to trypsin. *Protein Sci.* **2001**, *10*, 1584−1595.
(48) Almlof, M.; Brandsdal, B. O.; Aqvist, J. Binding affinity prediction with different force fields: examination of the linear interaction energy method. *J. Comput. Chem.* **2004**, *25*, 1242−1254.
(49) Leiros, H. K.; Brandsdal, B. O.; Andersen, O. A.; Os, V.; Leiros, I.; Helland, R.; Otlewski, J.; Willassen, N. P.; Smalas, A. O. Trypsin specificity as elucidated by LIE calculations, X-ray structures, and association constant measurements. *Protein Sci.* **2004**, *13*, 1056−1070.
(50) Almlof, M.; Aqvist, J.; Smalas, A. O.; Brandsdal, B. O. Probing the effect of point mutations at protein−protein interfaces with free energy calculations. *Biophys. J.* In press.

JM051018T