A Theoretical Investigation of Tight-Binding Thermolysin Inhibitors

Jian Shen

Hoechst Marion Roussel, Inc., 2110 East Galbraith Road, P.O. Box 156300, Cincinnati, Ohio 45215-6300

Received October 7, 1996[⊗]

A tight-binding thermolysin inhibitor, Cbz-Phe-ψ[PO_2NH]-Leu-Ala (ZF^{p}LA, K_i = 0.068 nM), and its analogs, ZR^p(O)LA (R = Ala, Leu or Phe) have been studied using the finite difference solution to the linearized Poisson−Boltzmann equation (FDPB) and solvation entropy correction (SEC). The binding energy difference between conformationally different thermolysin inhibitors ZF^{p}LA and ZG^{p}LL is estimated using three approaches. Two of approaches use the X-ray structures of ZF^{p}LA-thermolysin and ZG^{p}LL-thermolysin structures. The third one uses both X-ray structures to calculate binding energy differences from ZF^{p}LA and ZG^{p}LL to a hypothetical intermediate Me^{p}LA. All the results are qualitatively correct with one closely reproducing the experimental value. The enhancement of the ZF^{p}LA binding is attributed largely to the solvation entropy or "hydrophobic force". The binding mode of the ZG^{p}LR N-terminal moiety appears to be electrostatically unfavorable. Reducing the polarity of that moiety is predicted to enhance binding affinity. The binding trends due to the hydrophobic variation of ZR^p(O)LA are calculated within 1 kcal/mol of the experimental values. Increasing lipophilicity of a ligand favors the binding due to the difference of surface area change between the free state and the bound state. The analysis of energetic components shows that these trends are not specific for the binding of phosphorus-containing inhibitors but are generally true for protein−ligand interactions. The electrostatic calculation does not support the involvement of the second protonation of ZF^{p}LA in binding. Therefore, reexamining the second protonation of ZF^{p}LA or seeking further experimental support seems appropriate. The structural sensitivity of the FDPB calculation was assessed by using ligand and receptor structures from different X-ray studies of thermolysin. The small deviations (<0.3 Å) in the receptor structures do not cause significant changes in electrostatic binding energy if there is no structural change in modified regions.

Introduction

The zinc endopeptidase thermolysin, with a variety of phosphorus-containing inhibitors, has been extensively studied by kinetic and thermodynamic experiments as well as X-ray crystallography.[1−6] Its structure and catalytic mechanism have served as models for studying enzyme inhibition and for designing inhibitors against angiotensin-converting enzyme and matrix metalloproteinases.[7−9] The abundant binding data and multiple high-resolution X-ray crystal structures have also attracted many theoretical studies for the fundamental understanding of molecular recognition.[10−15]

One series of the inhibitors,[1] Cbz-R-ψ[PO_2X]-Leu-Ala (where X = NH, NH_2, or O; R = Leu, Ala, or Phe) or ZR^p(X)LA, has not been subjected to detailed theoretical studies. Their unique slow-binding character was attributed to a single residue (R) modification at the P1 position. The X-ray study[2] revealed that the bound conformation of N-terminal moiety of ZF^{p}LA differs from that of ZG^{p}LL, one of another series of phosphorus-containing peptide inhibitors, ZG^p(X)L(Y)R (X = NH, O, or CH_2; Y = NH or O; R = Leu, Ala, Gly, Phe, H, or CH_3), shown in Figure 1. It has been suggested that all ZR^p(X)LA (R ≠ Gly) bind to thermolysin like ZF^{p}LA. The cause of slow binding can then be attributed to the bulky N-terminal moiety of ZR^{p}LA, which tends to trap, rather than to displace, a bound water molecule.

Surprisingly, ZF^{p}LA is the most potent (K_i = 0.068 nM) among all reported phosphorus-containing peptide inhibitors. The tight-binding character and the bound conformation of ZF^{p}LA must be controlled by thermodynamic factors. Holden et al.[2] suggested that the hydrophobic contribution from the P1 phenylalanine side chain and a second proton of the phosphonamide nitrogen in ZF^{p}LA stabilize the complex. In the absence of a thermodynamic analysis, however, it is difficult to verify this mechanism and quantify the various binding contributions.

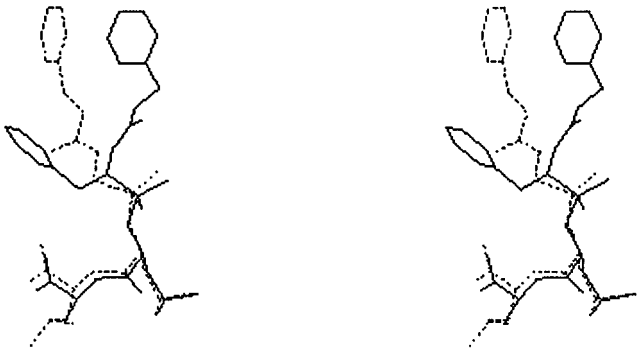


Figure 1. Stereoviews of superimposed bound ZF^{p}LA (solid line) and ZG^{p}LL (dash line). The structures were determined by X-ray. For ZF^{p}LA, the carbonyl of the carbobenzoxy and amide of Phe are hydrogen bonded to the thermolysin (not shown). The corresponding moiety of ZG^{p}LL is rotated 117°, forming a water mediated H-bond with the enzyme.

Besides studying the tight-binding mechanisms, detailed calculation of binding energy differences is relevant to lead optimization in pharmaceutical research. The inhibitor modifications in this system are common in medicinal chemistry. It should be interesting to see whether a theoretical method can calculate the binding energies, in particular between conformationally differ-

[⊗] Abstract published in *Advance ACS Abstracts,* August 15, 1997.

S0022-2623(96)00695-4 CCC: $14.00

ent inhibitors ZF^{p}LA and ZG^{p}LL. Although the structural and conformational difference between these two inhibitors seems small compared with the diversity of compounds experimented with in drug discovery, accurate, and timely computations of their binding energies are still beyond the limit of simulation methods due to slow equalibration from one system to the other, a so called "sampling problem".

In this paper, a theoretical investigation of ZR^{p}LA binding using the finite difference solution to the linearized Poisson−Boltzmann equation and solvation entropy correction (FDPB+SEC) method[13] and a related binding mechanism study are described. The method includes the finite difference solution to the linearized Poisson−Boltzmann equation (FDPB)[16−19] for electrostatic interaction and molecular surface calculation for solvation entropy correction (SEC).[13] These two energetic components have been shown to account for most of the binding energy difference among ZG^p(X)L(Y)R inhibitors. Multiple X-ray structures of thermolysin complexs were used to access the binding differences of the conformationally different ZF^{p}LA, ZG^{p}LL, and related inhibitors. The consistent results obtained from several thermolysin structures not only indicate the predictive power of the calculations but also allow further exploration of the binding mechanism. The analysis of the calculated energy components and experimental data sheds light on the origin of the tight binding, hydrophobic interaction, and second protonation of ZF^{p}LA. These studies suggest that reducing the polarity of P1 amide moiety of ZG^{p}LR will increase binding affinities.

Method

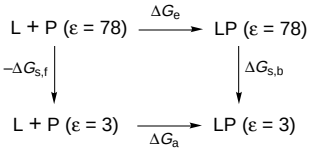
The calculation procedure for the FDPB + SEC method has been described previously.[13,20] Briefly, for a noncovalent binding process, L + P = LP, the binding free energy, ΔG, can be partitioned into electrostatic and nonelectrostatic components, ΔG_e and ΔG_n, respectively

$$\Delta G = \Delta G_e + \Delta G_n$$

Thus, the binding energy difference between a reference system and a modified system is

$$\Delta\Delta G = \Delta\Delta G_e + \Delta\Delta G_n$$

The calculation of ΔG_e contains both the solvation (ΔG_s) and assembly (ΔG_a) energies[16] as shown in the following scheme

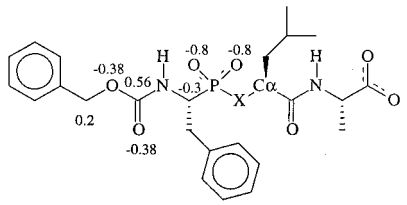


$$\Delta G_e = \Delta G_{s,b} - \Delta G_{s,f} + \Delta G_a$$
$$\Delta\Delta G_e = \Delta\Delta G_{s,b} - \Delta\Delta G_{s,f} + \Delta\Delta G_a$$

where ϵ is the dielectric of the environment. In this approach, two solvation energies are obtained from four electrostatic calculations using FDPB, and the assembly energy is analytically calculated using Coulomb's law. Due to the solute dielectric (i.e., 3 in our work), an identical dielectric ($\epsilon = 3$) is used for the "gas" phase to simplify the calculation of ΔG_a. Our $\Delta\Delta G_n$, or the solvation entropy correction (SEC), is derived as

$$\Delta\Delta G_n = 0.04(\Delta A_f - \Delta A_b)$$

where ΔA_f and ΔA_b are water accessible surface area changes of both the free state and the bound state in two binding complexes, respectively. The factor, 0.04 kcal/mol Å^2, is a statistical value derived from solvation entropies of a series of small molecules.[21]



X =	X(H)	P(O)	Cα
NH	-0.76(0.34)	1.32	0
O	-0.62	1.40	0.12
NH_2	-0.36(0.34)	1.22(-0.7)	0.16

Figure 2. Partial charges on ZF^p(X)LA used in the calculation, where X = NH, O, or NH_2. Charges on unspecified atoms are identical to those of similar atom groups in GROMOS.[24]

Newly released UHBD 5.1[22,23] was used for the FDPB calculation. Both the assembly energy and water accessible surface area calculations, which lead to SEC, were also performed within this program. The solvent region in the FDPB calculation was determined using the accessible surface of a probe of 1.1 Å radius and modeled with a dielectric of 78.[13] The solution was represented by an ionic strength of 0.01 M with an additional 1.4 Å of ion-excluding shell beyond the atomic radii of the molecules. A dielectric constant of 3.0 with a smooth dielectric boundary was used for the interior of ligands and proteins. The focusing option was applied to reduce the grid spacing to 0.25 Å around the binding site. GROMOS charges[24] were used for proteins. The partial charges for ZF^{p}LA and related analogs, depicted in Figure 2, are based on 6-31G* ab initio calculations of model compounds CH_3-PO_2-X-CH_3. Three X-ray thermolysin structures,[2,3] PDB4TMN (4tmn), PDB5TMN (5tmn), and PDB6TMN (6tmn) bound with ZF^{p}LA, ZG^{p}LL, and ZG^p(O)LL, respectively, were used as reference systems. Glu 143 and His 231 were protonated according to previous studies.[3,10,13] Polar hydrogens were added and energetically minimized using the program X-PLOR.[25] The coordinates of the modified inhibitors were edited on the basis of their reference structures. All calculations were done on SGI R4000. One complete binding energy calculation, including six FDPB calculations on a 100 × 100 × 100 grid, took about 1 h of CPU time.

Results

The calculated energies of the ligands in Table 1 are divided into four sets according to their reference systems. Because of hardware and software changes, recalculation of ZG^{p}LA in the second set serves as a control to ensure consistency. Compared with the previous calculation,[13] the $\Delta\Delta G_c$ of ZG^{p}LA only differs by 0.02 kcal/mol (1.34 of this study vs 1.32) showing excellent reproducibility. The small deviations in $\Delta\Delta G_e$ (−0.466 vs −0.357) and $\Delta\Delta G_n$ (1.80 vs 1.68) indicate that both the electrostatic and nonelectrostatic calculations in the current study are consistent with those reported earlier. The calculated $\Delta\Delta G_s$ can be verified qualitatively by their lipophilicity trends. A larger hydrophobic substituent or less polar modification increases the solvation energies. ZF^p(NH_2)LA has a much higher solvation energy because of its reduced negative charge (−1e). As we demonstrated previously,[20] the numerical error in $\Delta\Delta G_e$ should be less than 0.2 kcal/mol. The other uncertainty in $\Delta\Delta G_c$ comes from $\Delta\Delta G_n$, which may be as large as 15% of $\Delta\Delta G_n$.

Because the sensitivity of the calculated binding energy difference to the reference structure is unknown,

Table 1. Calculated Energy Differences for a Series of Thermolysin Inhibitors (kcal/mol)[a]

ligand	$\Delta\Delta G_{s,b}$	$\Delta\Delta G_{s,f}$	$\Delta\Delta G_a$	$\Delta\Delta G_e$	A_b	A_f	$\Delta\Delta G_n$	$\Delta\Delta G_c$	$\Delta\Delta G_{exp}$[b]
4tmn									
ZF^{p}LA	−1292.930	−74.751	−67.555	0	12401	785	0	0	(−14.05)
ZG^{p}LA	−1.406	−0.306	0	−1.100	12393	695	3.27(0.5)	2.17	3.29
Me^{p}LA	−4.362	1.390	2.106	−3.646	12429	479	13.34(2.0)	9.70	
ZF^p(O)LA	−0.178	1.373	6.552	5.001	12401	786	−0.01	4.99	3.90
ZL^p(O)LA	−0.264	1.366	6.552	4.920	12389	747	1.00(0.15)	5.93	5.53
ZA^p(O)LA	−1.398	1.263	6.552	3.891	12389	694	3.18(0.5)	7.07	6.11
ZG^{p}LL	−5.927	−0.275	5.74	0.088	12459	726	4.70(0.71)	4.79	2.94
ZF^p(NH_2)LA	25.452	29.367	18.479	14.564	12401	786	−0.021	14.54	
5tmn									
ZG^{p}LL	−1363.980	−75.026	−60.073	0	12469	726	0	0	(−11.11)
Me^{p}LA	−1.273	1.875	−1.105	−4.253	12450	469	9.55(1.4)	5.30	
ZG^{p}LA	−0.501	−0.035	0	−0.466	12471	684	1.80(0.3)	1.34	0.35
ZF^{p}LA	6.902	0.275	−4.92	1.707	12416	785	−4.45(0.67)	−2.74	−2.94
6tmn									
ZG^p(O)LL	−1385.327	−73.214	−52.031	0	12452	726	0	0	(−6.97)
ZG^{p}LL	0.283	−1.851	−5.882	−3.748	12452	726	−0.02	−3.77	−4.13
ZG^p(C)LL	0.271	0.485	−2.379	−2.593	12452	725	0.03	−2.56	−2.43
ZG^{p}L(O)L	0.211	−2.3	−3.365	−0.854	12452	727	−0.07	−0.93	−1.55
4tmn/5tmn									
ZG^{p}LL	−1298.857	−75.026	−61.815	0	12459	726	0	0	(−11.11)
ZG^p(O)LL	−0.359	1.836	6.099	3.904	12459	725	0.02	3.93	4.13
ZG^p(C)LL	−0.062	2.393	3.765	1.310	12459	725	0.06	1.37	1.70
ZG^{p}L(O)L	−0.18	−0.461	2.834	3.115	12459	727	−0.05	3.06	2.58

[a] Data are divided into four sets. The first three sets use 4tmn, 5tmn, and 6tmn as the reference structures for both the ligand and the receptor. The last reference system, 4tmn/5tmn, is a combination of thermolysin of 4tmn and ZG^{p}LL of 5tmn. The calculated energies of the first ligands in each set are absolute energies served as the reference systems. All other energies are relative to those of the first ligands in each set. Thus a reader can easily reproduce absolute solvation energies or assembly energies for each modified system and compare energy difference between any two inhibitors. The energy differences are as follows: $\Delta\Delta G_{s,b}$ and $\Delta\Delta G_{s,f}$, solvation energy differences in the bound state and in the free state, respectively; $\Delta\Delta G_a$, assembly energy differences; $\Delta\Delta G_e$, total electrostatic energy difference calculated by $\Delta\Delta G_{s,b} - \Delta\Delta G_{s,f} + \Delta\Delta G_a$; $\Delta\Delta G_n$, nonelectrostatic energy difference calculated by $0.04(\Delta A_f - \Delta A_b)$, where A_f and A_b are water accessible surface area for the free state and the bound state, respectively; the values in parentheses are estimated errors; $\Delta\Delta G_c$, calculated binding energy difference, i.e. $\Delta\Delta G_e + \Delta\Delta G_n$. [b] Experimental binding energy differences from refs 1, 5, and 6. The values in parentheses are absolute binding energies of the reference systems.

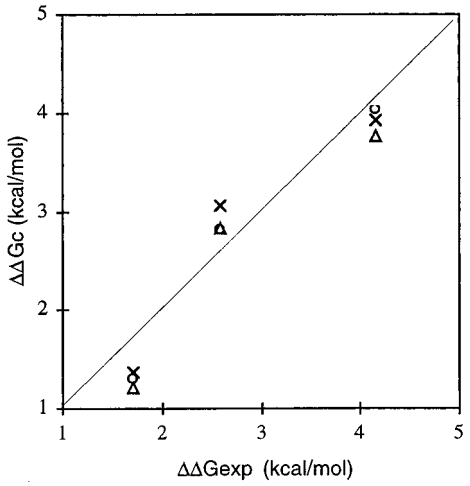


Figure 3. Calculated binding energy differences vs experimental values for ZG^p(O)LL, ZG^p(C)LL, and ZG^{p}L(O)L using 4tmn (×), 5tmn (○), and 6tmn (△) structures. $\Delta\Delta G$'s of 4tmn and 6tmn are converted from Table 1 and reference to the binding of ZG^{p}LL. The data of 5tmn are from previous calculations.[13]

a series of ZG^p(X)LL analogs (the third and forth sets in Table 1) were recalculated using 4tmn and 6tmn as reference systems. Together with 5tmn, these protein complex structures are similar in overall conformation with root mean square deviations ranging from just 0.1 to 0.3 Å. Because of the long-range nature of the electrostatic force, the accumulated structural deviations may affect calculated electrostatic energies. As shown in Figure 3, there is no significant change in $\Delta\Delta G_e$'s. The maximum deviation, including $\Delta\Delta G_s$ and $\Delta\Delta G_a$, is less than 0.2 kcal/mol. The small structural differences among the X-ray structures, however, alter the absolute solvation and assembly energies greatly.

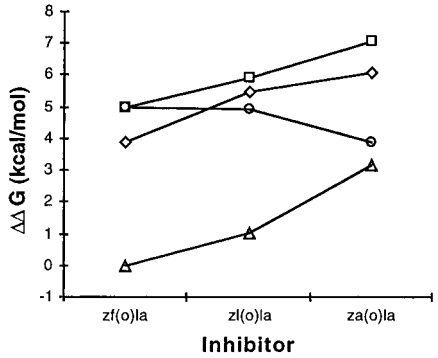


Figure 4. Calculated and experimental binding energy differences for ZR(O)LA, where ◇ is the experimental values; □, $\Delta\Delta G_c$; ○, $\Delta\Delta G_e$; and △, $\Delta\Delta G_n$.

For example, there is 65 kcal/mol of ΔG's difference (Table 1, column 2, rows 9 and 17) between ZG^{p}LL bound to either 4tmn or 5tmn. Clearly, the invariable part of the structures (thermolysin structure in this case) in the reference and modified systems have to be identical to minimize deviation in $\Delta\Delta G$'s. This guideline was used to devise the calculation schemes for the binding energy difference between ZF^{p}LA and ZG^{p}LL.

The binding energy differences of three ZR^p(O)LA analogs in the first set produced the correct binding affinity trends shown in Figure 4. The approximate 1 kcal/mol overestimation in binding energy for the three inhibitors is not from the R group but rather from the P1′ amide. Compared with the experimental 3.9 kcal/mol difference between ZF^{p}LA and ZF^p(O)LA, the calculated $\Delta\Delta G_e$ were overestimated by about 1 kcal/mol. Because the 4tmn protein structure produces consistent $\Delta\Delta G_e$ from ZG^{p}LL to ZG^p(O)LL (3.93 vs 4.13 of experiments), the overestimation in binding energy must be attributed to the ligand. The two ligand

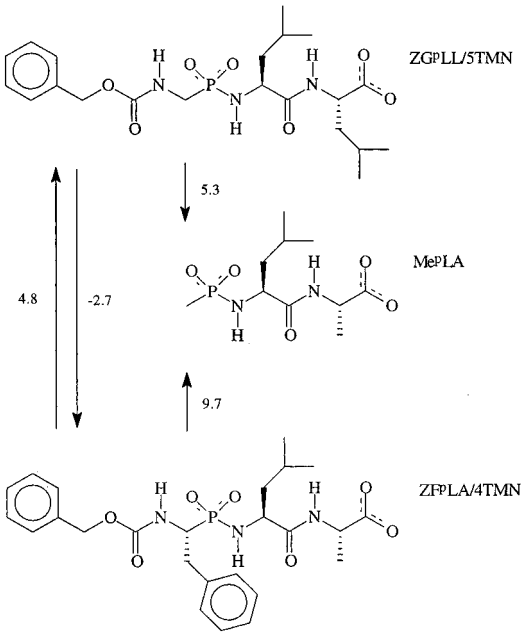


Figure 5. Structure of MeᵖLA and schemes to compute energy difference between ZFᵖLA and ZGᵖLL. The arrows indicate calculations from a reference system to a modified system. The values (taken from Table 1) along the arrows are calculated binding energy differences (in kcal/mol) associated with each process.

structures, ZGᵖLL of 5tmn and ZFᵖLA of 4tmn from the X-ray data are indeed different around the P1 amide. Besides bidentate vs monodentate bindings of the phosphoramide oxygens to the Zn, the inhibitor atoms around P1 amide in ZFᵖLA are less well defined than those in ZGᵖLL. Approximations in potential function may also contribute to the overestimation. Excluding this 1 kcal/mol overestimated energy from $\Delta\Delta G_c$, the modification at P1 can be very well explained by the calculation. The relative affinities among the ZRᵖ(O)LA analogs are dominated by $\Delta\Delta G_n$, which is primarily due to the surface area changes of free ligands.

The binding energy difference between ZFᵖLA and ZGᵖLL was assessed by three calculation schemes shown in Figure 5. Two of the binding energy differences using 4tmn and 5tmn as reference protein structure are 4.79 and 2.74 kcal/mol, respectively, which yield an average binding energy difference of 3.8 ± 1.0 kcal/mol. The calculation with 5tmn closely reproduced the experimental 2.94 kcal/mol of binding energy difference. The discrepancies among the calculated values come mainly from $\Delta\Delta G_e$ not $\Delta\Delta G_n$. For example, the $\Delta\Delta G_e$ deviation between 4tmn and 5tmn systems is 1.8 kcal/mol (1.71 + 0.09 from Table 1, column 5, rows 12 and 7), which accounts for most of their $\Delta\Delta G_c$ deviation, 2.1 kcal/mol.

Due to substantial conformational changes from bound ZFᵖLA to bound ZGᵖLL, we speculated that either 4tmn or 5tmn alone can represent both the reference and the modified systems. Thus, a third approach was devised to incorporate both protein structures assuming that they are equally accurate. In this approach, a hypothetical inhibitor MeᵖLA is used as the modified inhibitors in two calculations, from ZFᵖLA and ZGᵖLL to MeᵖLA. There is no conformational change in each calculation because the coordinates of MeᵖLA are the same as the corresponding reference coordinates. The combination of the two hypothetical binding processes ($\Delta\Delta G_c$ = 9.70 and 5.30 kcal/mol in Table 1, column 9, rows 3 and 10) yields a binding energy difference of 4.4 kcal/mol between ZFᵖLA and ZGᵖLL. Although the value falls in between the first two results, it is not as close to experimental data as the average value. All three calculations consistently reproduced the experimental observation that ZFᵖLA has higher affinity than ZGᵖLL.

Discussion

ZFᵖLA Protonation. We have tried two models for calculating the binding energy of ZFᵖLA. One model is a singly protonated phosphonamide as is the case with ZGᵖLR. The other, denoted as ZFᵖ(NH_2)LA, is a doubly protonated model, suggested by the X-ray study.[2] Our calculation shows that ZFᵖ(NH_2)LA is unlikely to be involved in the energetic binding. First, using ZFᵖ-(NH_2)LA as both the free state and the bound state (assuming the second protonation occurs before binding) results in a dramatic loss of binding energy (~14 kcal/mol). The calculated binding energy difference between ZFᵖ(NH_2)LA and ZFᵖ(O)LA is 9.5 kcal/mol (14.51−4.99 from Table 1, column 9, rows 8 and 4), which is inconsistent with the experimental value of −3.9 kcal/mol.

Alternatively, we can assume that ZFᵖLA binds to the active site before the second protonation. The total binding should be the sum of ZFᵖLA binding and a second proton transfer. The calculated binding energy difference between ZFᵖLA and ZFᵖ(O)LA is −5.0 kcal/mol, which is very close to the experimental value. That leaves little, if any, binding contribution from the second protonation. It should be mentioned that the −3.9 kcal/mol of experimental binding energy difference between ZFᵖLA and ZFᵖ(O)LA is almost identical to −4.0 kcal/mol of averaged binding energy difference between other phosphonamidates and phosphonates in ZGᵖLR series, where the phosphonamide nitrogen is considered singly protonated.

The second protonation would have a significant impact on the solvation energy and assembly energy because of a unitary charge change on the inhibitor. According to our calculation for the bound state, it will cause a total of 43 kcal/mol (25 kcal/mol of solvation and 18 kcal/mol of assembly from Table 1, columns 2 and 4, row 8) loss of affinity compared to the reference system, or about 37 kcal/mol loss compared to ZFᵖ(O)LA binding. A similar amount of energy (negative) would be required to compensate this loss for stabilizing the complex. One neglected energetic process accompanying the second protonation of ZFᵖLA is deprotonation. It appears that the proton could come directly from the solvent, or more plausibly, be shuttled by Glu 134.[2] The net result is the same as a proton transfer from solvent to the complex. The associated energy can be decomposed into proton desolvation, proton assembly to the complex in gas phase, and resolvation of the bound proton. The energies associated with the later two steps are already included in the energy of the bound state ($\Delta\Delta G_{s,b}$ + $\Delta\Delta G_a$) for ZFᵖ(NH_2)LA. Because of the positive energy of proton desolvation, the deprotonation would not be able to stabilize the complex. If the second proton can be confirmed by an experiment, another proton donor, or some other theoretical basis, must be sought.

Enhanced Binding. The dominant $\Delta\Delta G_n$ between ZFᵖLA and ZGᵖLA is in good agreement with the experimental observation that the hydrophobic charac-

ter of the P1 Phe side chain is a major contributor to the enhanced binding of ZF^{p}LA. The analysis of the decomposed energy also indicates that the orientation of P1 amide may contribute significantly to the binding difference. ZF^{p}LA has two H-bonds from P1 amide to the back bone N of Trp 115 and OH of Tyr 157. The corresponding amide of ZG^{p}LL rotates to the opposite direction forming a water mediated H-bond network between the nitrogen of P1 Gly^{p} and the N of Trp 115. Although both orientations are permissible in conformational space, the associated $\Delta\Delta G_a$ shows that the two H-bonds enhance binding better than the water mediated H-bond does. The rotation of the amide moiety causes ZG^{p}LL to loss 4.92 to 5.74 kcal/mol in $\Delta\Delta G_a$. Thus, in terms of electrostatic binding, the P1 amide moiety seems to be a potential target for ZG^{p}LR binding improvement.

This idea is supported by an experimental compound, *n*-hexyl-PO2-Leu-Trp-NHMe, which has improved affinity ($\Delta G_{exp} = -12.2$ kcal/mol, $\Delta\Delta G_{exp} = -1.1$ kcal/mol).[5] To calculate its binding energy requires a bound structure of the P2′ Trp side chain and the C-terminal NHMe, which is not available. While the modeling of the structure is beyond the scope of this paper, the role of hydrophobic hexyl group at N-terminus can be investigated by the FDPB method assuming the hexyl chain aligned with the N-terminal moiety of ZG^{p}LL. A preliminary calculation of a model compound, *n*-hexyl-PO2-Leu-Leu, results in a $\Delta\Delta G_e$ of -0.9 kcal/mol. Thus, it confirms that reducing the polarity at the N-terminus can improve the electrostatic binding affinity for ZG^{p}LR.

Hydrophobic Modification. Hydrophobic modification has long been recognized as a way to improve binding affinity and other properties.[26−29] The calculations of hydrophobic modifications, such as ZR^{p}(O)LA and the previous ZG^{p}(X)LR, indicate some underlying trends in energetic components and the corresponding nature of the binding. Interestingly, electrostatic binding (or electrostatic solvation because assembly energy is zero) prefers a small hydrophobic group, while the nonelectrostatic $\Delta\Delta G_n$ is in favor of a large hydrophobic group, see Figure 4. For example, $\Delta\Delta G_e$ and $\Delta\Delta G_n$ between ZF^{p}(O)LA and ZA^{p}(O)LA are -1.1 and 3.2 kcal/mol, respectively. In terms of the thermodynamics of our model, the former is more solvation enthalpy in nature and the later solvation entropy.

Further examination of $\Delta\Delta G_e$ components reveals that the absolute values of $\Delta\Delta G_{s,f}$ are much smaller than that of $\Delta\Delta G_{s,b}$. For the above pair of inhibitors, the absolute value of $\Delta\Delta G_{s,f}$ and $\Delta\Delta G_{s,b}$ are 0.1 and 1.2 kcal/mol, respectively. In this study, a hydrophobic modification in the free state influences only the inhibitor solvation. In contrast, it affects the solvation of both inhibitor and protein in bound state. This difference leads to negative $\Delta\Delta G_e$ for a smaller hydrophobic modification, i.e. from ZF^{p}(O)LA to ZA^{p}(O)LA. Therefore, the enhanced electrostatic binding for smaller hydrophobic modifications is due to extended solvation of the binding complexes.

The $\Delta\Delta G_n$ calculation is directly related to the solvent assessable surface areas of the free state and the bound state. The same hydrophobic change (Ala to Phe) results in 12 Å^2 of ΔA_b and 92 Å^2 of ΔA_f. The smaller ΔA_b is due to less solvation of the bound inhibitor. Unless a modified hydrophobic group is fully surrounded by solvent in bound state (then $\Delta A_b = \Delta A_f$), ΔA_f is always greater than ΔA_b with increasing hydrophobicity. According to this definition, the corresponding $\Delta\Delta G_n$ is always positive in favor of greater hydrophobic modification.

Based on this model, the interplay between $\Delta\Delta G_e$ and $\Delta\Delta G_n$ determines the net binding effect on a hydrophobic modification. The calculations of modifications at P1 and P2′ of these inhibitors show that the free state component of $\Delta\Delta G_n$, $0.04\Delta A_f$, is often the leading term in the binding energy differences. Thus, one may use $0.04\Delta A_f$ as a rough estimate for the effect of a hydrophobic modification when a receptor structure is unknown. This approximation seems to have the same basis as widely used statistical methods and empirical methods that correlate either lipophilicity (log *P*) or surface area of ligands to binding energies with variable coefficients.[27,28] In general, however, surface area or properties of the bound state as well as electrostatic solvation should be included.

Binding Water and Other Interactions. The binding water in 4tmn appears to be kinetically important and probably relevant to binding thermodynamics as well. However, an explicit representation of binding waters in FDPB binding energy difference calculation has several unsolved problems. First, the coordinates of water protons are not available. The reliability of modeled proton coordinates are unknown. Second, any ligand or protein molecular modification near a binding water is likely to change the coordinates of the water thereby increasing the uncertainty in the calculation. Third, a receptor−ligand system with an explicit binding water becomes a trimolecular system. There is a concern about how to count for the configurational entropy change when a binding water is replaced by a molecular modification on a ligand or a protein. It appears that more fundamental research is needed to solve this issue. On the other hand, the similar high dielectric character of ice supports the use of the continuum model for stationary binding waters as an approximation.

In this study, the solvation process, including both the enthalpic and entropic contributions in ligand−protein binding, is considered. It is commonly agreed that the former can be accurately and practically calculated using FDPB. The later is simply a statistically derived factor relating the surface area of a molecule to its solvation entropy. Other widely used molecular interactions[30] such as conformational energies and van der Waals force were completely neglected. Yet, the results qualitatively agree with our current knowledge about molecular solvation and interaction and are quantitatively close to experimental observations. A reasonable explanation is that the differences in omitted interactions are probably small in the systems we studied. Despite the imperfections, the combination of accuracy, speed, and mechanistic insight provided by FDPB + SEC is unique among many binding energy calculation methods.

Structural Variations. The underlying rational of using one static structure (an X-ray structure) for binding energy deference calculation assumes that the structure represents the ensemble averages of both reference and modified systems. From the results of ZG^{p}(X)LL binding using 4tmn, 5tmn, and 6tmn, small structural deviations (<0.3 Å) in receptor structures (with no modification) can be tolerated in the calculated

binding energy difference of small molecular modification such as NH to O on ligands. This is probably due to the error cancellation in computing binding energy differences. However, when a large modification or conformational change occurs at ligands such as ZF^{p}LA and ZG^{p}LL, the uncertainties in $\Delta\Delta G_e$ become large and may exceed the $\Delta\Delta G_e$ itself. This dependency may indicate that one reference structure is no longer a representative of both ensemble averages of reference and modified systems. Errors in X-ray structures could be another factor. On the other hand, $\Delta\Delta G_n$ seems not very sensitive to reference structures. Thus, good qualitative results may still be obtained if $\Delta\Delta G_n$ dominates the binding difference.

While this work concentrated on multiple X-ray complexes, it was also intended to examine the feasibility of applications to modeled structures such as docked inhibitors and homologous proteins. This study implies that the use of correctly modeled structures may be able to produce qualitative binding energy trends, which are useful in structure-based design. The errors in the calculations of different molecular modifications can serve as the limit of quantitative binding energy calculations. No doubt, more work needs to be done before the method can be routinely used in pharmaceutical research.

Conclusion

The binding energy difference between conformationally different thermolysin inhibitors ZF^{p}LA and ZG^{p}LL is estimated using FDPB + SEC method. The results of all three approaches, from ZF^{p}LA to ZG^{p}LL, from ZG^{p}LL to ZF^{p}LA and from both ZF^{p}LA and ZG^{p}LL to a hypothetical intermediate Me^{p}LA, are qualitatively correct with one closely reproducing the experimental value. The enhancement of the ZF^{p}LA binding is attributed largely to the solvation entropy or "hydrophobic force". The binding mode of the ZG^{p}LR N-terminal moiety appears to be electrostatically unfavorable. Reducing the polarity of that moiety is predicted to enhance binding affinity. The binding trends due to the hydrophobic variation of ZR^p(O)LA are calculated within 1 kcal/mol of the experimental values. Increasing lipophilicity of a ligand favors the binding due to the difference of surface area change between the free state and the bound state. The analysis of energetic components shows that these trends are not specific for the binding of the phosphorus-containing inhibitors, but are generally true for protein−ligand interactions. The electrostatic calculation does not support the involvement of the second protonation of ZF^{p}LA in binding. Therefore, reexamining the second protonation of ZF^{p}LA or seeking further experimental support seems appropriate.

Acknowledgment. Author is particularly grateful to Drs. T. Pelton, D. Kominos, and R. Vaz for their comments and help in revising the manuscript.

References

(1) (a) Bartlett, P. A.; Marlowe, C. K. Evaluation of intrinsic binding energy from a hydrogen bonding group in an enzyme inhibitor. *Science* **1987**, *235*, 569−571. (b) Bartlett, P. A.; Marlowe, C. K. Possible role for water dissociation in the slow binding of phosphorus-containing transition-state-analogue inhibitors of thermolysin. *Biochemistry* **1987**, *26*, 8553−8561.

(2) Holden, H. M.; Tronrud, D. E.; Monzingo, A. F.; Weaver, L. H.; Matthews, B. W. Slow- and fast-binding inhibitors of thermolysin display different modes of binding: Crystallographic analysis of extended phosphonamidate transition-state analogues. *Biochemistry* **1987**, *26*, 8542−8552.

(3) Tronrud, D. E.; Holden, H. M.; Matthews, B. W. Structures of two thermolysin-inhibitor complexes that differ by a single hydrogen bond. *Science* **1987**, *235*, 571−574.

(4) Matthews, B. W. Structural Basis of the action of thermolysin and related zinc peptidases. *Acc. Chem. Res.* **1988**, *21*, 333−340.

(5) Grohelny, D.; Goli, U. B.; Galardy, R. E. Binding of phosphorus-containing inhibitors of thermolysin. *Biochemistry* **1989**, *28*, 4948−4951.

(6) Morgan, B. P,; Scholtz, J. M.; Ballinger, M. D.; Zipkin, I. D.; Bartlett, P. A. Differential Binding Energy, A detailed evaluation of the influence of hydrogen-bonding and hydrophobic groups on the inhibition of thermolysin phosphorus-containing inhibitors. *J. Am. Chem. Soc.* **1991**, *113*, 297−307.

(7) Rich, D. H.; Northrop, D. B. Enzyme kinetics in drug design: Implications of multiple forms of enzyme on substrate and inhibitor structure-activity correlations. In *Computer-aided Drug Design*; Perun, T. J., Propst, C. L., Eds.; Marcel Dekker: New York, 1989; pp 185−250.

(8) Hangauer, D. G. Computer-aided design and evaluation of angiotensin converting enzyme inhibitor. In *Computer-aided Drug Design*; Perun, T. J., Propst, C. L., Eds.; Marcel Dekker, New York, 1989; pp 253−326.

(9) Henderson, B.; Docherty, A. J. P.; Beeley, N. R. A. Design of inhibitors of articular cartilage destruction. *Drug Future* **1990**, *15*, 495−507.

(10) Bash, P. A; Singh, U. C.; Brown, F. K.; Langridge, R.; Kollman, P. A. 1987. Calculation of the relative change in binding free energy of a protein-inhibitor complex. *Science* **1987**, *235*, 574−576.

(11) Merz, K. M.; Kollman, P. A. Free energy perturbation simulations of the inhibition of thermolysin: Prediction of the free energy of binding of a new inhibitors. *J. Am. Chem. Soc.* **1989**, *111*, 5649−5658.

(12) Wendoloski, J. J.; Shen, J.; Oliva, M. R.; Weber, P. C. Biophysical Tools for Structure-Based Drug Design. *Pharmacol. Ther.* **1993**, *60*, 169−183.

(13) Shen, J.; Wendoloski, J. J. 1995. Binding of phosphorus-containing inhibitors to thermolysin studied by the Poisson-Boltzmann Method. *Protein Sci.* **1995**, *4*, 373−381.

(14) Head, R. D.; Smythe, M. L.; Oprea, T. I.; Waller, C. L.; Green, S. M.; Marshall, G. R. Validate: A new method for the receptor-based prediction of binding affinities of novel ligands. *J. Am. Chem. Soc.* **1996**, *118*, 3959−3969.

(15) Wasserman, Z. R.; Hodge, C. N. Fitting an inhibitor into the active site of thermolysin: A molecular dynamics case study. *Protein* **1996**, *24*, 227−237.

(16) Gilson, M. K.; Honig, B. Calculation of the total electrostatic energy of a macromolecular system: Solvation energies, binding energies and conformational analysis. *Proteins* **1988**, *4*, 7−18.

(17) Harvey, S. C. Treatment of electrostatic effects in macromolecular modeling. *Proteins* **1989**, *5*, 78−92.

(18) Davis, M. E.; McCammon, J. A. Electrostatics in biomolecular structure and dynamics. *Chem. Rev.* **1990**, *90*, 509−521.

(19) Sharp, K. A.; Honig, B. Electrostatic interaction in macromolecules: Theory and applications. *Annu. Rev. Biophys. Chem.* **1990**, *19*, 301−332.

(20) Shen, J.; Wendoloski, J. J. Electrostatic binding energy Calculation Using the finite difference solution to the linearized Poisson-Boltzmann equation: Assessment of its accuracy. *J. Compt. Chem.* **1995**, *17*, 350−357.

(21) Ben-Naim, A.; Marcus, Y. Solvation thermodynamics of nonionic solutes. *J. Chem. Phys.* **1984**, *81*, 2016−2027

(22) Davis, M. E.; Madura, J. D.; Luty, B. A.; McCammon, J. A. Electrostatic diffusion of molecules in solution: simulation with the University of Houston Brownian Dynamics program. *Comp. Phys. Commun.* **1991**, *62*, 187−190.

(23) Briggs, J. M.; Madura, J. D.; Davis, M. E.; Gilson, M. K.; Antosiewicz, J.; Luty, B. A.; Wade, R. C.; Bagheri, B.; Ilin, A.; Tan, R. C.; McCammon, J. A. *University of Houston Brownian Dynamics Program User's Guide and Programmer's Manual Release 5.1.exp*, available at Molecular Simulations, Inc., San Diego, 1995.

(24) van Gunsteren, W. F.; Berendsen, H. J. C. *Groningen Molecular Simulation (GROMOS) Library Manual*; Biomos: Nijenborgh 16, The Netherlands, 1987.

(25) Brunger, A. T. *X-PLOR Manual v 2.0*; Yale University: New Haven, CT, 1991.

(26) Tanford, C. *The hydrophobic Effect: Formation of Micelles and Biological Membranes*; John Wiley & Sons: New York, 1973.

(27) Hansch, C.; Leo, A. *Exploring QSAR*; American Chemical Society: Washington, DC, 1995.

(28) Gao, J.; Qiao, S.; Whitesides, G. M. Increasing binding constants of ligands to carbonic anhydrase by using "greasy tails". *J. Med. Chem.* **1995**, *38*, 2292−2301.

(29) Testa, B.; Carrupt, P-A.; Gatrick, C.; Billois, F.; Weber, P. Lipophilicity in molecular modeling. *Pharm. Res.* **1996**, *13*, 335−343.

(30) Ajay; Murcko, M. A. Computational methods to predict binding free energy in ligand-receptor complex. *J. Med. Chem.* **1995**, *38*, 4953−4967.

JM9606958