Why Does Trypsin Cleave BPTI so Slowly?

Mikael Peräkylä[†] and Peter A. Kollman*[,‡]

Contribution from the Department of Chemistry, University of Kuopio, P.O. Box 1627, FIN-70211, Kuopio, Finland, and Department of Pharmaceutical Chemistry, University of California, San Francisco, California, 94143-0446

Received May 13, 1999. Revised Manuscript Received September 13, 1999

Abstract: Bovine pancreatic trypsin inhibitor (BPTI) binds to trypsin with high affinity. In the trypsin−BPTI complex the reactive peptide sequence of BPTI is positioned ideally for peptide bond hydrolysis. However, the peptide bond hydrolysis is $\sim 10^{11}$ slower for BPTI than for peptide substrates. We have applied the quantum mechanical-free energy (QM-FE) approach and molecular dynamics simulations to investigate peptide bond cleavage in the trypsin−BPTI and the trypsin−substrate systems. The activation free energy calculated for BPTI as substrate suggests a decrease in catalytic rate by 10^6-10^9, in qualitative agreement with the $\sim 10^{11}$ found experimentally. We provide evidence to support the hypothesis that the remaining 10^2-10^5 decrease in rate arises from the fact that the peptide bond in BPTI, once cleaved, is much more likely to re-ligate than to be further hydrolyzed to form products which will be released from trypsin. We have also demonstrated that the QM-FE approach, in combination with the free energy component analysis, should be a useful method to investigate enzyme catalysis and to derive information on the roles of different enzyme groups on the reaction energies.

Introduction

Protein inhibitors of proteases are ubiquitous in living systems. They play a vital role in many physiological processes by regulating the proteolytic activity of their target proteases. The majority of protein inhibitors of proteases known and characterized so far are directed toward serine proteases. These inhibitors can be classified into at least 16 families based on sequence similarity, topological similarity, and mechanism of binding.[1,2] The mechanism of inhibition can be presented in its simplest form by the equation

$$E + I \rightleftharpoons EI \rightleftharpoons EI^* \rightleftharpoons E + I^* \qquad (1)$$

where E is protease, I and I* are the virgin (reactive site peptide bond intact) and modified (reactive site peptide bond hydrolyzed) inhibitors, respectively, and EI is an enzyme−inhibitor noncovalent complex. Protease inhibitors form very stable complexes with proteases having K_i values in the $10^{-7}-10^{-13}$ M range.[1]

The inhibition mechanism of bovine pancreatic trypsin inhibitor (BPTI), which belongs to the pancreatic trypsin inhibitor (Kunitz) family is investigated in this study. BPTI, like most of the serine-protease-directed inhibitors, reacts with the target enzyme in a substrate-like manner. Such inhibitors all possess an exposed binding loop, which combines with the enzyme in a substrate-like manner and serves as a substrate for the enzyme (Figure 1). Although k_{cat}/K_m for this interaction is large, both k_{cat} and K_m are very small and, therefore, inhibitors are hydrolyzed slowly.[1,2] BPTI binds to trypsin with high affinity ($K_i = 10^{-13}$ M), forming a rigid and stable complex ($k_{cat} = 8.7 \times 10^{-10}$ s^{-1}).[3] The X-ray structures of several trypsin−BPTI complexes have revealed that there exist extensive contacts with the protease and inhibitor. LysI15 (I = Inhibitor) of the inhibitor is located in the specificity pocket (S1) of trypsin and AlaI16 is in the S1′ position (Figure 2).[1]

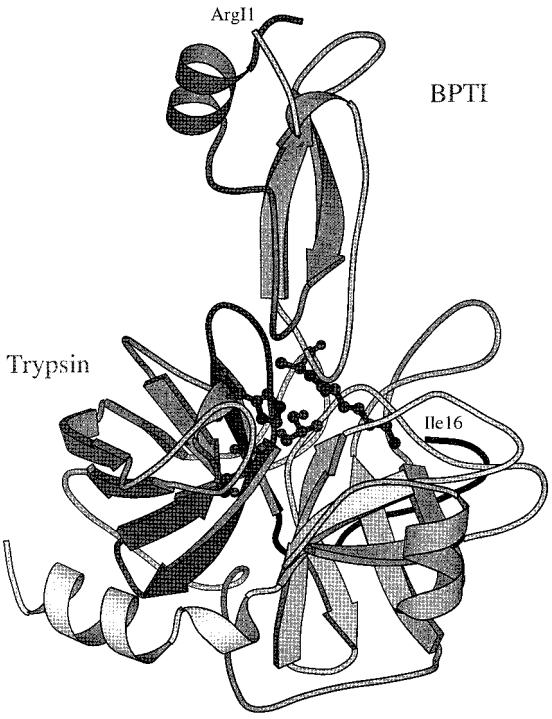

Figure 1. Structure of the trypsin−BPTI complex. The catalytic amino acid residues, His57, Asp102, and Ser195, of trypsin, and the reactive residues, LysI15 and AlaI16, of BPTI are shown in detail. This figure was prepared with MOLSCRIPT.[48]

[†] University of Kuopio.
[‡] University of California.

(1) Laskowski, M., Jr.; Kato, I. *Annu. Rev. Biochem.* **1980**, *49*, 593−626.
(2) Bode, W.; Huber, R. *Eur. J. Biochem.* **1992**, *204*, 433−451.
(3) Estell, D. A.; Wilson, K. A.; Laskowski, M., Jr. *Biochemistry* **1980**, *19*, 131−137.

10.1021/ja991602p CCC: $19.00 © 2000 American Chemical Society
Published on Web 03/24/2000

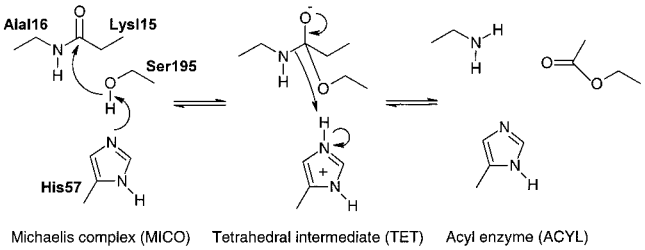

Figure 2. Reaction mechanism of the acylation step of the trypsin-catalyzed peptide hydrolysis and the structures of Michaelis complex (MICO), tetrahedral intermediate-like transition state (TET), and acyl-enzyme intermediate (ACYL) studied in this work. In the trypsin−BPTI system the scissile bond is between LysI15 and AlaI16.

The stable structures are close to Michaelis complexes (MICO) with little or no distortion from planarity in the LysI15−AlaI16 peptide bond, which is positioned ideally for peptide bond cleavage. The complex formation has occurred with only small conformational changes in the structure of inhibitor or protease. Thus, the structures show that inhibitors bind as if they were good substrates and form structures which closely represent frozen Michaelis complexes. Whereas the tight binding of protease inhibitors to proteases is nowadays well understood,[2,4] it is not clear why the peptide bonds remain intact in the complexes and are hydrolyzed so slowly. Two main reasons have been proposed to explain the slow hydrolysis of protease inhibitors.[3] First, the enzyme may bind to the inhibitor in such a way that favorable noncovalent bonds in the enzyme−inhibitor complex have to be disturbed on the formation of an enzyme−transition-state complex. In this way the large binding energy of the Michaelis-like complex is used to increase the energy of the transition state for bond cleavage. Second, following the cleavage of the active-site peptide bond, the newly formed termini are held in close proximity in the modified inhibitor−protease acyl-enzyme complex, greatly favoring reformation of the peptide bond. The tight binding inhibits the hydrolysis of the covalent acyl-enzyme complex (ACYL) by preventing a water molecule from entering the active site.[5−7]

In this study, we have used an ab initio quantum mechanics (QM) and molecular dynamics/free energy (FE) calculations based QM-FE approach[8] to study the hydrolysis of LysI15−AlaI16 peptide bond in trypsin−BPTI and trypsin−substrate complexes. Because acylation is the rate-limiting reaction step for peptide substrates,[9] we have calculated the activation free energies for the acylation step of peptide bond cleavage (from MICO to TET) in trypsin−BPTI complex and for trypsin-catalyzed hydrolysis of tetra (P3−P2−P1−P1′, SUB1) and penta peptide (P3−P2−P1−P1′-P2′, SUB2) substrates. The sequences of the two substrates were taken from the reactive-site sequence of BPTI. Comparison of the peptide bond hydrolysis of similar reactive-site sequences in trypsin−BPTI and trypsin−substrate complexes has allowed us to investigate factors responsible for the slow peptide bond hydrolysis of the protease inhibitor. MD simulations of 500 ps were also done for the trypsin−BPTI (TET/BPTI) and trypsin−SUB1 (TET/SUB1) tetrahedral intermediates and for the trypsin−BPTI (ACYL/BPTI) and trypsin−SUB1 (ACYL/SUB1) acyl-enzyme intermediates (Figure 2). Calculations for the tetrahedral intermediate were done to see how the catalytic groups are located in the active site for the cleavage of the peptide bond. Acyl-enzyme calculations were done to see how the newly formed termini in the BPTI and substrate and acyl group are located in the active site for the possible reformation of the broken peptide bond. Our calculations suggest that both of the previous explanations for slow hydrolysis of BPTI could play a role in explaining the $\sim 10^{11}$ slower hydrolysis of BPTI than that of substrate. Specifically, we find that the increased activation free energy for bond cleavage of BPTI than for substrate leads to a 10^6-10^9 reduction in rate. On the basis of the MD simulations on EI* (tetrahedral and acyl-enzyme intermediates), we suggest that the remaining reduction of 10^5 could be due to the fact that this is the ratio of reverse and forward rate constants for EI* going to either EI or E + I*.

Methods

The Quantum Mechanical-Free Energy (QM-FE) Approach. The quantum mechanical-free energy (QM-FE) approach is a general way to combine quantum mechanical calculations with classical mechanical free energy method to calculate the energetics of enzyme-catalyzed and solution reactions.[8] In the QM-FE approach, which is analogous to the approach introduced by Jorgensen,[10,11] the free energy difference, ΔG^*, between two structures is calculated as

$$\Delta G^* = \Delta E(\text{gas}) + \Delta G_{\text{int}} \qquad (2)$$

Here ΔE(gas) is the quantum mechanical energy difference between two QM fragments and ΔG_{int} is the difference in free energy of interaction between the QM fragments and the environment. The charge distributions and geometries of the QM fragments are from QM calculations, and standard free energy simulations are used to calculate the ΔG_{int} energies. In our approach we use the RESP[12,13] software and Lagrangian constraints to generate atomic point charges for the QM fragments.

Initial Structure Definition of Michaelis Complexes. The initial coordinate for simulations of trypsin−bovine pancreatic trypsin inhibitor (BPTI) and trypsin−substrate complexes were obtained from the X-ray crystal structure of bovine trypsin complexed with BPTI (Brookhaven Protein Data Bank code 2PTC) determined to 1.9 Å resolution.[14] The sequences of the two substrates studied, Acetyl-Pro-Cys-Lys-Ala-NH_2 (SUB1) and Acetyl-Pro-Cys-Lys-Ala-Arg-NH_2 (SUB2) were taken from the BPTI sequence.

The X-ray structure of trypsin−BPTI complex was solvated by adding a sphere of TIP3P[15] water molecules with a 20 Å radius from the Oγ of the catalytic Ser195 with the use of the "cap" option of the LEAP[16] program. Crystallographic waters were included in trypsin−BPTI as well as trypsin−substrate complexes. The trypsin−BPTI complex, which included a Ca^{2+} ion, was neutralized by placing 15 Cl^- ions in the positions of largest electrostatic potential as determined by the program CION of the AMBER[17] package. Counterions were fixed during all simulations. The solvated and neutralized system was

(4) Conte, L. L.; Chothia, C.; Janin, J. *J. Mol. Biol.* **1999**, *285*, 2177−2198.
(5) Longstaff, C.; Campbell, A. F.; Fersht, A. R. *Biochemistry* **1990**, *29*, 7339−7347.
(6) Perona, J. J.; Tsu, C. A.; Craik, C. S.; Fletterick, R. J. *J. Mol. Biol.* **1993**, *230*, 919−933.
(7) Castro, M. J. M.; Anderson, S. *Biochemistry* **1996**, *35*, 11435−11446.
(8) Stanton, R. V.; Peräkylä, M.; Bakowies, D.; Kollman, P. A. *J. Am. Chem. Soc.* **1998**, *120*, 3448−3457.
(9) Kraut, J. *Annu. Rev. Biochem.* **1977**, *46*, 331−358.
(10) Chandrasekhar, J.; Jorgensen, W. L. *J. Am. Chem. Soc.* **1985**, *107*, 2974−2975.
(11) Jorgensen, W. L. *Acc. Chem. Res.* **1989**, *22*, 184−189.
(12) Bayly, C. I.; Cieplak, P.; Cornell, W. D.; Kollman, P. A. *J. Phys. Chem.* **1993**, *97*, 10269−10280.
(13) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Kollman, P. A. *J. Am. Chem. Soc.* **1993**, *115*, 9620−9631.
(14) Marquart, M.; Walter, J.; Deisenhofer, J.; Bode, W.; Huber, R. *Acta Crystallogr.* **1983**, *B39*, 480−490.
(15) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J.; Impley, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926−935.
(16) Schafmeister, C. E. A. F.; Ross, W. S.; Romanovski, V. Leap: University of California, San Francisco, 1995.
(17) Pearlman, D. A.; Case, D. A.; Caldwell, J. W.; Ross, W. S.; Cheatham, T. E., III; Ferguson, D. M.; Seibel, G. L.; Singh, U. C.; Weiner, P. K.; Kollman, P. A. AMBER 4.1, University of California, San Francisco, 1995.

then energy-minimized for 1000 steps with use of the conjugate gradient algorithm. The minimized trypsin−BPTI system was used as starting point for molecular dynamics simulation performed at a constant temperature of 300 K with the use of the Berendsen algorithm.[18] Standard parameters of Cornell et al.[19] were used in all of the calculations unless otherwise noted. A time step of 1.5 fs and the SHAKE algorithm[20] to constrain bond distances were used along with a 12 Å nonbonded cutoff. In the simulations, we allowed movement only in residues within 14 Å of the Oγ of nucleophilic Ser195. MD equilibration of 75 ps was performed for the trypsin−BPTI complex before free energy calculations.

In the substrate simulations we used the solvated, neutralized, and equilibrated Acetyl-Ala-Phe-Arg-Ala-NH_2 (SUB3)-trypsin structure of our earlier work as an initial structure for trypsin.[8] The substrates studied in the present work were placed in the active site using the trypsin−BPTI X-ray structure as a template. The trypsin−SUB2 system was neutralized by adding an extra Cl^- with the program CION. After the substrate was placed in the active site, energy minimization of 1000 steps was performed. Molecular dynamics equilibration for 45 ps was done before starting the free energy simulations. On the basis of the evolution of energy of trypsin−BPTI and trypsin−substrate initial structures during the equilibration, the structures were considered to be stable and equilibrated for free energy simulations.

Geometries and Atomic Point Charges of QM Fragments. The part of the system considered quantum mechanically and the geometries and atomic partial charges of MICO and TET were taken from our earlier study. Since the geometry optimizations and charge determinations of the QM models are explained in detail there, they are only briefly described here. The geometries of MICO and TET were partly optimized by fixing the C_β and C_γ atoms of imidazole (His57), the C_β atoms of methanol (Ser195), and the carbonyl O and attached methyl C atoms of the substrate. The optimizations were done at the HF/6-31+G* level. The gas-phase reaction energies (ΔE(gas)) were calculated at the MP2/AUG-cc-pVDZ level using the HF/6-31+G* geometries. Atomic point charges were calculated with the RESP method at the HF/6-31+G* level. In the RESP fitting we used Lagrangian restraints to fix the net charge of the QM atoms to −0.07e. This ensures that the sum of the charges of the quantum mechanical atoms plus the charges of the molecular mechanical atoms His57, Ser195, and the substrate residues is exactly zero.

Free Energy Simulations. In the free energy simulations we employed the thermodynamic integration (TI) method to calculate the relative free energies (ΔG_{int}) between MICO and TET in the enzyme environment. In the free energy calculations the QM model system is a function of perturbation parameter (λ). When $\lambda = 1$ the charge distribution and bond topology corresponds to MICO, and when $\lambda = 0$ the charges and topology are those of TET. In the free energy calculations we neglected any free energy changes within the perturbed group, since these are included in ΔE(gas). Thus, ΔG_{int} corresponds to the environmental effects on the ab initio QM fragments and when combined with the difference in the energies between the QM fragments an overall free energy (ΔG^*) between the states is calculated (eq 2).

In the TI method the total free energy change can be represented as a summation of the force field (V in eq 3) terms, such as bond, angle, dihedral angle, van der Waals, and electrostatic interactions (eq 3).

$$\Delta G = \int_0^1 \langle \partial V/\partial\lambda \rangle_\lambda \partial\lambda \approx \sum_i \langle \Delta V \rangle_{\lambda i} \Delta\lambda \qquad (3)$$

One can also represent the free energy as a sum of the different parts of the system. In this work we have decomposed the ΔG_{int} into components for each of the amino acids, solvent molecules, and counterions. The free energy components represent interactions between the structural components of the environment and the QM active site. It must be noted that although the total free energy is a state function, the components are not, and the decomposition depends on the pathway taken between the initial and final states. Therefore, the components should not be considered as quantitative predictions, but as representing the magnitude of the interactions and, therefore, able to give useful physical insight into the system studied.[21−25]

In the free energy perturbation calculations, we performed 30 ps equilibration runs prior to data collection. The perturbation runs ($\lambda = 1 \rightarrow \lambda = 0$) of 75 ps used 101 windows and 150 ps runs used 201 windows. Each window comprised 200 steps of equilibration and 300 steps of data collection. Each perturbation was run from both forward (from $\lambda = 0$ to 1) and reverse (from $\lambda = 1$ to 0) directions, to obtain an estimate of the error in the calculations from the hysteresis. The final ΔG_{int} values were calculated as averages from forward and reverse runs. We used a 12 Å residue-based cutoff for nonbonded interaction for 75 and 150 ps simulations. Additional 75 ps simulations were carried out using a dual cutoff method[26] with 12 Å primary and 22 Å secondary cutoff.

MD Simulations of Tetrahedral and Acyl-Enzyme Intermediates of Trypsin−BPTI and Trypsin−SUB1 Complexes. The neutralized, solvated, and energy-minimized Michaelis complex of the trypsin−BPTI system was used as starting point for tetrahedral (TET/BPTI) and acyl-enzyme intermediate (ACYL/BPTI) simulations of trypsin−BPTI. Similarly, the neutralized, solvated, and energy-minimized Michaelis complex of the trypsin−SUB1 system was used as starting point for tetrahedral (TET/SUB1) and acyl-enzyme intermediate (ACYL/SUB1) simulations of trypsin−SUB1. The LEAP program was used to modify the active-site structure to correspond to that of tetrahedral and acyl-enzyme intermediate. After minimization of 1000 steps, an equilibration run of 45 ps was performed. After that MD simulation of 500 ps was done at 300 K with 12 Å cutoff and a time step of 1.5 fs. During the simulations structures were saved every 100 steps (every 0.15 ps) for further analysis. In the tetrahedral intermediate simulations the parameters were the same as in free energy calculations. The torsion parameters for the ester group of acyl-enzyme intermediate were derived from the rotation barrier of methyl acetate calculated at the MP2/6-31+G*//HF/6-31G* level. The bond, angle, improper, and nonbonded parameters were taken from the work of Fox and Kollman.[21] The atomic partial charges for the ester group of the ACYL/BPTI and ACYL/SUB1 were obtained using RESP fit (HF/6-31G* level) and methyl acetate as a model. In the fit the net charge of two methyl end groups of methyl acetate was set to 0.082 e. It is equal but of opposite sign to the sum of the acyl-forming C═O and OH groups of LysI15 and Ser195. This constraint ensures that the total charge of serine and lysine residues in ACYL/BPTI and ACYL/SUB1 is zero.

Results

Reaction Activation (ΔG^*) and Interaction Free Energies (ΔG_{int}). Reaction activation free energies (ΔG^*) for hydrolysis of the peptide bond (MICO → TET) were calculated from the gas-phase energies of the QM models (ΔE(gas)) taken from ref 8 and the free energies of interaction (ΔG_{int}) between the QM models and the environment obtained from free energy perturbation calculations ($\Delta G^* = \Delta E(\text{gas}) + \Delta G_{int}$). For all reactions studied we used the same QM model. Thus, ΔE(gas) was 53.8 kcal/mol (at the MP2/AUG-cc-pVDZ//HF/6-31+G* level) for all the studied reactions and, therefore, ΔG_{int}'s are responsible for different ΔG^* values. The calculated ΔG_{int} and ΔG^* values are reported in Tables 1 and 2.

(18) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A. D.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684−3690.
(19) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M., Jr.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. *J. Am. Chem. Soc.* **1995**, *117*, 5179−5197.
(20) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327−341.
(21) Fox, T.; Scanlan, T. S.; Kollman, P. A. *J. Am. Chem. Soc.* **1997**, *119*, 11571−11577.
(22) Smith, P. E.; van Gunsteren, W. F. *J. Phys. Chem.* **1994**, *98*, 13735−13740.
(23) Brady, G. P.; Szabo, A.; Sharp, K. A. *J. Mol. Biol.* **1996**, *263*, 123−125.
(24) Boresch, S.; Karplus, M. *J. Mol. Biol.* **1995**, *254*, 801−807.
(25) Sun, Y.-C.; Veenstra, D. L.; Kollman, P. A. *Protein Eng.* **1996**, *9*, 273−281.
(26) van Gunsteren, W. F.; Berendsen, H. J. C. *Angew. Chem., Int. Ed. Engl.* **1990**, *29*, 992−1023.

Table 1. ΔG_{int} Values (kcal/mol) for MICO → TET

	75 ps 12 Å cutoff			150 ps 12 Å cutoff			75 ps 12 Å/22 Å cutoff[a]		
substrate	forward	reverse	average	forward	reverse	average	forward	reverse	average
BPTI	−29.4	−30.9	−27.1	−29.6	−28.8	−29.2	−27.1	−26.5	−26.8
SUB1	−40.3	−37.8	−39.1	−38.3	−36.7	−37.5	−39.6	−40.1	−39.9
SUB2	−38.2	−40.4	−39.3	−35.4	−35.9	−35.7			
SUB3	−39.7	−39.4	−39.6	−37.6	−38.6	−38.1			

[a] Dual cutoff method was used with 12 Å primary and 22 Å secondary cutoff.

Table 2. ΔG^* Values (kcal/mol) for MICO → TET[a]

	75 ps 12 Å cutoff	150 ps 12 Å cutoff	75 ps 12 Å/22 Å cutoff
substrate	ΔG^*	ΔG^*	ΔG^*
BPTI	23.6	24.6	27.0
SUB1	14.7	16.3	13.9
SUB2	14.5	18.1	−
SUB3	14.2	15.7	−

[a] $\Delta G^* = \Delta E + \Delta G_{int}$, averages of ΔG_{int} energies are used. ΔE = 53.8 kcal/mol, calculated at the MP2/AUG-cc-pVDZ//HF/6-31+G* level.

In our earlier QM-FE calculations we obtained ΔG^*'s of 14.2 kcal/mol (75 ps simulation) and 15.7 kcal/mol (150 ps simulation) for the trypsin-catalyzed reaction of Acetyl-Ala-Phe-Arg-Ala-NH_2 (SUB3, Table 2).[8] The calculated numbers are in good agreement with the experimental reaction activation free energy 15.1 kcal/mol for this substrate. This indicates that QM-FE approach can reproduce the essential interactions in the enzyme environment. Using the average ΔG_{int}'s from 150 ps simulations, we get ΔG^* of 16.3 kcal/mol for Acetyl-Pro-Cys-Lys-Ala-NH_2 (SUB1) and 18.1 kcal/mol for Acetyl-Pro-Cys-Lys-Ala-Arg-NH_2 (SUB2). In the case of trypsin−BPTI complex ΔG^* is 24.6 kcal/mol (150 ps), 8.3 kcal/mol larger than for SUB1 and 6.5 kcal/mol larger than for SUB2. We also carried out 75 ps free energy calculations for trypsin−SUB1 and trypsin−BPTI using a dual cutoff method with 12 Å primary and 22 Å secondary cutoff value. This had only a small effect (−0.8 kcal/mol, 75 ps simulations) on SUB1 system, but increased the ΔG_{int} of trypsin−BPTI system by 3.4 kcal/mol. When ΔG_{int} values calculated with the dual cutoff method are used, ΔG^* of trypsin−BPTI is 27.0 kcal/mol, 13.1 kcal/mol larger than for SUB1. All of these results show that formation of the acyl-enzyme intermediate has a significantly larger activation free energy in the trypsin−BPTI complex than in trypsin−substrate complexes.

Interaction Free Energy (ΔG_{int}) Components. The calculated interaction free energies (ΔG_{int}'s) of trypsin−BPTI and trypsin−SUB1 reactions were separated into components involving different parts of the systems. The free energy components of trypsin, substrate, "link atoms", all water molecules, BPTI, and counterions from the 75 ps and 75 ps dual cutoff simulations are reported in Table 3. In this analysis the substrate group includes substrate residues other than reactive LysI15 and AlaI16, and link atoms are atoms of LysI15, AlaI16, His57, and Ser195 not included in the QM model. ΔG_{int}'s from 12 Å cutoff and 12/22 Å dual cutoff simulation are close to each other for both trypsin−BPTI and trypsin−SUB1. Therefore, only components from the dual cutoff simulations are discussed here in more detail. The amino acid residues of trypsin have a contribution to ΔG_{int} (MICO → TET) of −47.6 kcal/mol for trypsin−SUB1 and a contribution of −44.4 kcal/mol for trypsin−BPTI system. The contributions of substrate and link atoms are 9.6 and 4.4 kcal/mol for trypsin−SUB1. These components are 1.8 and 0.6 kcal/mol less favorable for trypsin−BPTI than for trypsin−SUB1. Water molecules have a favorable ΔG_{int} contribution of −5.4 kcal/mol for trypsin−SUB1. In contrast, in trypsin−BPTI system water molecules have an unfavorable contribution of 1.0 kcal/mol to peptide cleavage. Thus, water molecules lower the relative activation free energy of trypsin−SUB1 6.4 kcal/mol compared to the trypsin−BPTI system. This is due to the fact that in trypsin−substrate system water molecules have access to the active site, whereas in the latter system, this penetration is prevented by BPTI (Figure 1). In addition to this effect, the amino acid residues of BPTI have an unfavorable contribution of 8.6 kcal/mol to the MICO → TET reaction. ΔG_{int} components of trypsin, substrate, link atoms, water, and BPTI are more positive ($\Delta\Delta G_{int}$, Table 3) in trypsin−SUB1 than in the trypsin−BPTI system. Thus, they all increase ΔG^* of trypsin−BPTI compared to trypsin−SUB1. In trypsin−SUB1 system, Cl^- counterions have a ΔG_{int} contribution of 0.1 kcal/mol, whereas counterions provide a relatively larger ΔG_{int} contribution of −9.0 kcal/mol in the trypsin−BPTI system. This contribution comes mainly from the interactions of two Cl^- ions, which are located ∼16.0 and ∼13.0 Å from the Nϵ atom of the catalytic His57 in positions where they are farther away from the oxyanion than His57. At these positions Cl^- ions neutralize positive charges of ArgI39 and LysI46 of BPTI and Lys60 of trypsin (Figure 3). ΔG_{int} components of these three residues in trypsin−BPTI are 2.0, 6.6, and 4.3 kcal/mol, respectively (see below, Table 4).

Table 3. Free Energy Components of ΔG_{Int} (kcal/mol) for MICO → TET in Trypsin−Acetyl-Pro-Cys-Lys-Ala-NH_2 (SUB1) and Trypsin−BPTI

	12 Å cutoff			12/22 Å dual cutoff		
group	ΔG_{int} (SUB1)	ΔG_{int} (BPTI)	$\Delta\Delta G_{int}$[a]	ΔG_{int} (SUB1)	ΔG_{int} (BPTI)	$\Delta\Delta G_{int}$[a]
trypsin	−47.4	−44.0	3.4	−47.6	−44.4	3.2
substrate[b]	9.2	10.6	1.4	9.6	11.4	1.8
link atoms[c]	5.0	5.4	0.4	4.4	5.0	0.6
water	−5.7	1.2	6.9	−5.4	1.0	6.4
BPTI		8.0	8.0		8.6	8.6
counterions	0.0	−10.6	−10.6	0.1	−9.0	−8.9

[a] $\Delta\Delta G_{int} = \Delta G_{int}(BPTI) - \Delta G_{int}(SUB1)$. [b] Amino acid residues of the substrate or the corresponding amino acids of BPTI excluding reactive Lys and Ala. In the case of SUB1 components of Acetyl, Pro, Cys, and NH_2 are included. In the case of BPTI components of Pro and Cys are included. [c] MM atoms of the split amino acid residues Ser, His, Ala, and Lys.

ΔG_{int}'s from 75 ps dual cutoff simulations were further divided into components of individual amino acid residues. Residues having the largest ΔG_{int} components in trypsin−SUB1 and trypsin−BPTI are listed in Table 4. Most of the amino acid residues with large components are charged ones, which interact with the catalytic groups by stabilizing/destabilizing electrostatic interactions. Asp102 has the largest ΔG_{int} component of −21.4 kcal/mol. Asp102 is part of the catalytic triad and stabilizes the protonated His57 of the tetrahedral intermediate. In BPTI there exist six positively charged residues, but no negatively charged residues which have large ΔG_{int} components. Figure 3 shows

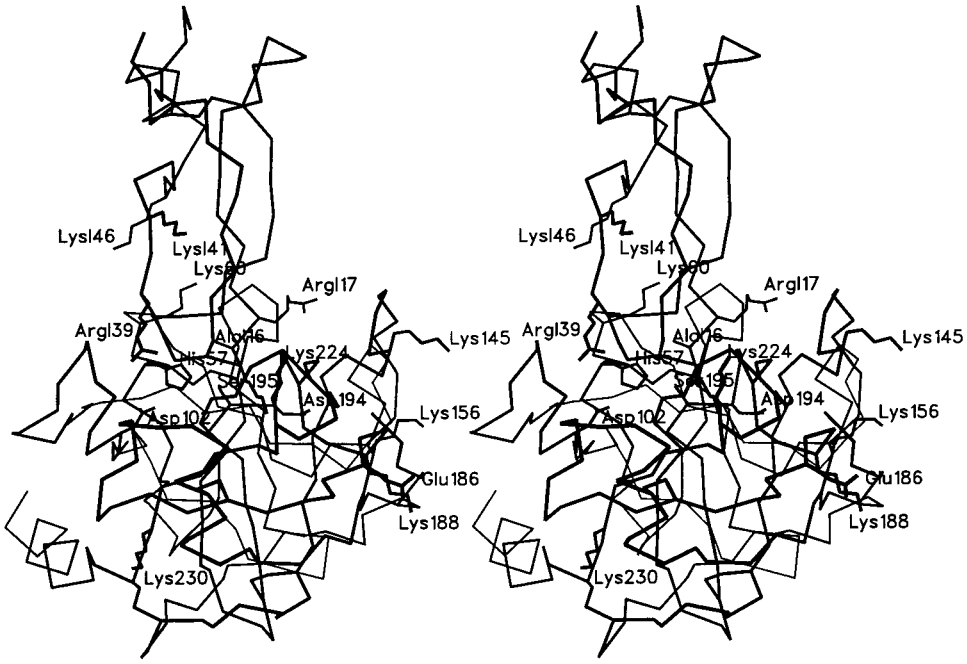

Figure 3. A stereoview of the trypsin−BPTI complex showing positions of selected charged amino acid residues with large free energy components of ΔG_{int} (Table 4). This figure was prepared with SETOR.[49]

Table 4. Contribution of Selected Amino Acid Residues to Free Energies for MICO → TET in Trypsin−Acetyl-Pro-Cys-Lys-Ala-NH_2 (SUB1) and Trypsin−BPTI

	contribution		
group	trypsin− SUB1	trypsin− BPTI	notes
Ile16	−6.8	−6.0	*N*-terminus of trypsin
Phe41	2.4	2.4	C═O (main chain)---H−N (main chain) of P2′
Lys60	3.4	4.3	
Asp102	−21.4	−21.6	Asp of catalytic triad
Lys145	−9.4	−9.6	
Lys156	−5.0	−5.0	
Lys159	−3.8	−1.4	
Glu186	4.0	4.0	
Lys188	−5.0	−5.0	
Asp189	4.1	3.6	
Gln192	−5.4	−5.5	C_αH---oxyanion interaction
Gly193	−5.4	−5.4	N−H of oxyanion hole
Asp194	5.8	5.8	
Val213	−2.8	−2.6	C═O (main chain)---H−Cϵ of His57
Lys222	−3.8	−3.8	
Lys224	−5.4	−5.4	
Lys230	8.0	9.4	
CysI14	10.0	10.4	P2
LysI15	−11.0	−11.2	P1
AlaI16	14.8	15.6	P1′
ArgI17	−	−8.6	P2′
ArgI20	−	1.4	
GlyI36	−	−3.8	C═O (main chain)- - -H−Cγ of His57
ArgI39	−	2.0	
LysI41	−	7.4	
LysI46	−	6.6	

how the selected charged residues are located in the trypsin−BPTI complex. There are also several noncharged residues with large components. All of these residues have direct interactions with the catalytic residues (Figure 4). On the basis of location of the noncharged residues they can be divided into two groups: (i) residues which interact with the oxyanion of the tetrahedral intermediate and those (ii) which interact with the

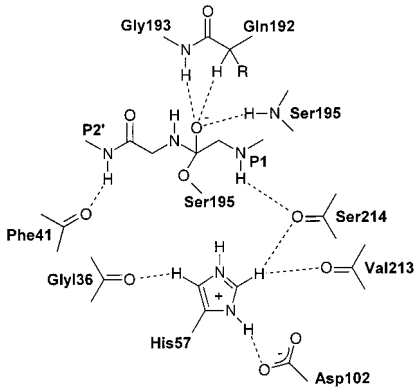

Figure 4. Interactions between the trypsin and BPTI residues and the active-site residues important for the catalytic reaction.

catalytic histidine. Gly193 is part of the oxyanion hole, and its main chain N−H group makes a hydrogen bond with the anionic oxygen. This interaction has ΔG_{int} component of −5.4 kcal/mol. The main chain N−H of Ser195 makes the second hydrogen bond of the oxyanion hole. Contribution of this interaction is not seen in Table 4, because part of Ser195 belongs to the QM model. In addition to these two hydrogen bonds, Gln192, which has ΔG_{int} of −5.4 kcal/mol in trypsin−SUB1 and −5.5 kcal/mol in trypsin−BPTI, interacts with the oxyanion through the hydrogen of its α-carbon. In addition to Asp102 of the catalytic triad, decomposition of ΔG_{int}'s revealed two other interactions, one from trypsin and one from BPTI, which considerably stabilize the protonated His57 of the tetrahedral intermediate. Val213 of trypsin has ΔG_{int} of −2.8 kcal/mol in trypsin−SUB1 and −2.6 kcal/mol in trypsin−BPTI. This favorable contribution is due to interactions between the main chain C═O of Val213 and Cϵ−H of His57. GlyI36 of BPTI, which has ΔG_{int} component of −3.8 kcal/mol, makes similar main chain interaction with Cγ−H of His57. In addition, Ser214 (ΔG_{int} = −0.7 kcal/mol) makes main chain C═O to histidine Cϵ−H interaction (Figure 4), which further contribute to the tight binding of His57. Binding of the reactive sequence of BPTI/substrate is facilitated by hydrogen bonds between the

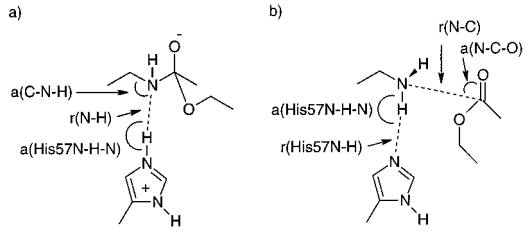

Figure 5. Definition of geometrical parameters for the (a) trypsin−BPTI (TET/BPTI) and trypsin−SUB1 (TET/SUB1) tetrahedral intermediate and (b) trypsin−BPTI (ACYL/BPTI) and trypsin−SUB1 (ACYL/SUB1) acyl-enzyme intermediate.

Table 5. Geometrical Parameters for Active-Site Groups in 500 ps MD Simulation of Tetrahedral Intermediate of Trypsin−BPTI (TET/BPTI) and Trypsin−SUB1 (TET/SUB1) Complexes (Figure 5a)[a]

parameter	average	max	min	SD[b]
		TET/BPTI		
r(N−H)	2.89	4.17	2.09	0.32
a(N−H−N)	145.5	177.5	108.9	10.6
a(C−N−H)	107.5	136.0	84.7	6.5
		TET/SUB1		
r(N−H)	2.72	4.24	1.99	0.33
a(N−H−N)	152.3	180.0	104.0	10.0
a(C−N−H)	104.6	132.1	63.4	8.6

[a] Distances in Å and angles in deg. Number of structures is 3400. [b] SD = standard deviation.

main chain C=O of Phe41 of trypsin (ΔG_{int} = 2.4 kcal/mol) and the main chain N−H of P2′ (Arg/NH_2) of BPTI/substrate and an analogous interaction between the C=O of Ser214 of trypsin and the N−H at P1 of BPTI/substrate (Lys).

Structures of Tetrahedral and Acyl-Enzyme Intermediates. MD simulations (500 ps) of tetrahedral intermediates of trypsin−BPTI (TET/BPTI) and trypsin−SUB1 (TET/SUB1) were carried out to see how the N−H of His57 and N of the scissile peptide bond are oriented (Figure 5a). In the trypsin-catalyzed reaction the tetrahedral intermediate forms an acyl-enzyme intermediate by proton transfer from the protonated His57 to the N of the scissile peptide bond (Figure 2). It can be assumed that this takes place readily if the transferred H of the His57 is close to the N of the scissile bond (N−H distance) and the His57−H bond points toward N of the scissile bond (i.e., the His57−H−N angle should be close to 180°). Also, the transferred H should approach the N of the scissile bond in an angle suitable for reaction (i.e., the C−N−H angle should be close to 120°). Results from the analyses of these geometrical parameters are reported in Table 5. In the TET/BPTI trajectory the N−H distance (average is 2.89 Å, standard deviation, SD, is 0.32 Å) and N−H−N (av 145.5°, SD = 10.6°) and C−N−H (av 107.5°, SD = 6.5°) angles are suitable for the peptide bond cleavage. In the TET/SUB1 trajectory the average N−H distance (2.72 Å) and the N−H−N angle (152.3°) are slightly more favorable for the reaction than in the case of TET/BPTI. The C−H−N angles are within 3° in the two cases. All in all, the geometrical parameters in the tetrahedral intermediate are slightly more favorable for cleavage of the peptide bond for TET/SUB1 than TET/BPTI.

MD simulations of trypsin−BPTI (ACYL/BPTI) and trypsin−SUB1 (ACYL/SUB1) acyl-enzyme intermediates were carried out to find out how the newly formed NH_2 terminus of AlaI15 and C=O group of the acyl-enzyme's ester group are oriented (Figure 5b). It can be assumed that reformation of the peptide bond and the Michaelis complex (via tetrahedral intermediate) takes place readily if the nitrogen lone pair of the terminal NH_2 is close to the carbonyl carbon of the acyl group (N−C distance) and approaches it perpendicular to the plane of C=O group (N−C−O angle). In addition, the leaving hydrogen of NH_2 group should be close to the N of His57, which can act as a general base in the reformation reaction (His57N−H distance). Also, the His57N−H bond should point toward the histidine's N (i.e., the His57N−H−N angle should be close to 180°). Results from the analyses of the two distances and angles are reported in Table 6. In the ACYL/BPTI trajectory the N of the NH_2 group stays for the most of the time less than 3.2 Å from the carbonyl carbon and within 20° from the ideal attack angle 90°. Average values are 3.02 Å (SD = 0.14 Å) for the N−O distance and 86.1° (SD = 7.6°) for the N−C−O angle. Thus, the approach of the terminal N is particularly suitable for reformation of the broken peptide bond. The distance (His57N−H) and angle (His57N−H−N) of a hydrogen of the NH_2 group are also suitable for reformation of the peptide bond. The average distance is 2.90 Å (SD = 0.52 Å), and angle, 150.52° (SD = 15.7°). The parameters for the latter interaction (His57N−H−N) show significantly more fluctuation than the ones for the attack of N. During the first 150 ps of the ACYL/SUB1 trajectory, the active-site structure remained suitable, although to a lesser extend than for ACYL/BPTI, for reformation of the peptide bond (Table 6, ACYL/SUB1, 0−150 ps). After 150 ps the noncovalently bound part of the substrate changed conformation. As a result of this the NH_2 group of the newly formed terminus became more exposed to solvent and would probably be protonated in physiological conditions. This conformational change took place in about 70 ps after which the structure stayed stable for the rest of the 500 ps simulation (Table 6, ACYL/SUB1, 220−500 ps). The geometrical parameters show that from 220 ps to 500 ps the C=O of the acyl group and N of His57 are far apart (4.9−5.4 Å) from the NH_2 terminus and unsuitably positioned for reformation of the peptide bond. Simulations of ACYL/BPTI and ACYL/SUB1 indicate that the terminal N of the broken peptide bond of BPTI remains suitably positioned for reformation of the peptide bond, whereas in the case of substrate the cleaved part of the peptide is ready to leave the active site.

Table 6. Geometrical Parameters for Active-Site Groups in 500 ps MD Simulation of Acyl-Enzyme Intermediate of Trypsin−BPTI (ACYL/BPTI) and Trypsin−SUB1 (ACYL/SUB1) Complexes (Figure 5b)[a]

parameter	average	max	min	SD[b]
		ACYL/BPTI		
r(N−C)	3.02	3.66	2.64	0.14
r(H−NE2)	2.90	4.60	1.80	0.54
a(N−C−O)	85.7	111.1	50.6	7.6
a(N−H−N)	150.5	179.7	96.5	15.9
		ACYL/SUB1 (0−150 ps)[c]		
r(N−C)	3.53	5.23	2.78	0.47
r(H−NE2)	2.71	5.01	3.21	0.80
a(N−C−O)	89.6	112.7	46.8	8.2
a(N−H−N)	140.4	179.6	31.0	35.2
		ACYL/SUB1 (220−500 ps)[d]		
r(N−C)	4.94	6.98	3.73	0.47
r(H−NE2)	5.44	7.73	3.47	0.63
a(N−C−O)	85.3	111.4	61.0	7.2
a(N−H−N)	78.1	146.3	7.4	26.3

[a] Distances in Å and angles in deg. Number of structures is 3400. [b] SD = standard deviation. [c] Data were collected from 0 to 150 ps of the 500 ps simulation. [d] Data were collected from 220 to 500 ps of the 500 ps simulation.

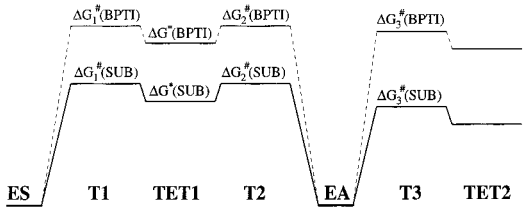

Figure 6. Schematic free energy diagram showing trypsin reaction with substrates (SUB, solid line) and BPTI (dashed line). T1 and T2 are the transition states for the formation and breakdown of the tetrahedral intermediate for acylation (TET1); T3 is the transition state for formation of the tetrahedral intermediate for deacylation (TET2). The corresponding free energies are indicated, with $\Delta G^{\#}$ used for the transition states and ΔG^{*} for the intermediate.

Discussion

Free Energies of BPTI and Substrate Hydrolysis by Trypsin. We can put our calculations in perspective by considering Figure 6, which shows a schematic free energy diagram for amide hydrolysis by serine proteases. It is clear both from calculations and experiments that $\Delta G_1^{\#}$ and $\Delta G_2^{\#}$, the activation free energies for formation and breakdown of the tetrahedral intermediate for acylation are comparable in free energy to each other and to the free energy for the formation of tetrahedral intermediate, ΔG^{*}, calculated here and in ref 8. Furthermore, for normal amide substrates, it is known that $\Delta G_3^{\#}$ is lower than $\Delta G_1^{\#}$ and $\Delta G_2^{\#}$, since acylation is rate-limiting. Our calculations have shown that $\Delta\Delta G^{*} = \Delta G^{*}(\mathrm{BPTI}) - \Delta G^{*}(\mathrm{SUB})$ is 8−13 kcal/mol, in the correct range but smaller than the 15 kcal/mol difference observed experimentally for $\Delta\Delta G^{\#}$. Given that this is the only free energy we can calculate directly with the available state-of-the-art methodology, we have analyzed some aspects of $\Delta G_2^{\#}$ and $\Delta G_3^{\#}$ by means of qualitative MD calculations. In the case of $\Delta\Delta G_1^{\#} = \Delta G_1^{\#}(\mathrm{BPTI}) - \Delta G_1^{\#}(\mathrm{SUB})$ or $\Delta\Delta G_2^{\#} = \Delta G_2^{\#}(\mathrm{BPTI}) - \Delta G_2^{\#}(\mathrm{SUB})$, there is no reason to expect them to be the same as $\Delta\Delta G^{*}$. In fact, given the more precise geometrical requirements in chemical reactions compared to noncovalent interactions, it would certainly be reasonable for $\Delta\Delta G_1^{\#}$ and $\Delta\Delta G_2^{\#}$ to be a few kcal/mol greater than $\Delta\Delta G^{*}$, which, if $\Delta\Delta G^{*}$ were at the upper end of 8−13 kcal/mol, could resolve the discrepancy with experiment. The geometrical analysis we have done on the MD simulations on the TET structures of BPTI and SUB do show that the structure is somewhat more favorable for His to substrate proton transfer in SUB, supporting the possibility that $\Delta\Delta G_2^{\#}$ is somewhat larger than $\Delta\Delta G^{*}$. However, an alternative possibility is that $\Delta\Delta G_3^{\#}$ is even larger than $\Delta\Delta G_1^{\#}$ or $\Delta\Delta G_2^{\#}$. Our molecular dynamics simulations on the acyl-enzyme intermediates (EA) for both SUB and BPTI show a significant difference between the qualitative structures in only 500 ps. In BPTI, the structure is stable, whereas in SUB, the cleaved part of the substrate moves substantially away from being able to re-ligate. This movement is SUB will allow a water to approach and lead to a low $\Delta G_3^{\#}(\mathrm{SUB})$. On the other hand, the stability of EA for BPTI significantly disfavors water entry and leads to $\Delta G_3^{\#}(\mathrm{BPTI})$ being larger than the re-ligation of the bond ($\Delta G_2^{\#}(\mathrm{BPTI})$). Nonetheless, we are not in a position to quantitatively calculate $\Delta G_3^{\#}(\mathrm{BPTI})$ and show whether it is indeed larger than $\Delta G_2^{\#}(\mathrm{BPTI})$. $\Delta G_3^{\#}(\mathrm{BPTI})$ can be much larger than $\Delta G_3^{\#}(\mathrm{SUB})$ and still be smaller than $\Delta G_2^{\#}(\mathrm{BPTI})$ and thus, not be the rate-limiting step in the reaction. It is likely that whether $\Delta G_1^{\#}(\mathrm{BPTI})/\Delta G_2^{\#}(\mathrm{BPTI})$ are larger or smaller than $\Delta G_3^{\#}$ can be determined experimentally by spectroscopic or mass spectroscopy approaches, given that the reactions are quite slow.

What we have shown in this work is that $\Delta\Delta G^{*}$ (and by implication $\Delta\Delta G_1^{\#}$ and $\Delta\Delta G_2^{\#}$) is of the correct magnitude (8−13 kcal/mol) to explain the effective experimental $\Delta\Delta G^{\#}$ (∼15 kcal/mol) and, based on more qualitative MD simulations, $\Delta G_3^{\#}(\mathrm{BPTI})$ would be expected to be substantially larger than $\Delta G_3^{\#}(\mathrm{SUB})$.

Peptide Bond Hydrolysis in the Trypsin−Substrate System. Serine proteases share similar structural parts important for their catalytic power and specificity. They all possess a catalytic triad (serine, histidine, and aspartate residues), the oxyanion hole, an extended site for nonspecific binding, and often a pocket for specific binding.[27] The roles of different parts of the catalytic machinery in the catalytic activity and specificity have been elucidated by experimental and computational methods and are nowadays well characterized.[27−30] Here we have applied the QM-FE approach together with the free energy component analysis to investigate the roles of the different parts of the catalytic center in the catalytic power of trypsin. These calculations provide interesting insight into the factors affecting the catalytic effect on peptide bond hydrolysis by trypsin and demonstrate the usefulness of the QM-FE approach in studying enzyme catalysis. Free energy component analysis has been earlier used to analyze for example the stability and activity of mutant proteins[31−33] and (differential) binding energies of ligands to protein receptors.[21,34,35] It must be noted that free energies are sensitive to the pathway chosen in the free energy calculation, and therefore, care must be taken to interpret the components.[21−25] However, we share the opinion that if properly used, free energy component analysis is a valuable computational tool, which gives useful physical information on the system studied.

The free energy component analysis on trypsin-catalyzed reaction resulted in a number of amino acid residues with large components (Table 4). There are several charged amino acids with large components. Most of them are located far away from the active site. Several of those residues are close to oppositely charge amino acid which compensate for their effect on reaction or, as in the case of trypsin−BPTI, this compensation is done by counterions. The salt bridge of Ile16-Asp194 is formed when trypsinogen is cleaved to make trypsin. Its formation creates a well-formed active-site pocket, but its contribution to the ΔG^{*} for MICO → TET (−1.0 kcal/mol for trypsin−substrate and −0.2 kcal/mol for trypsin−BPTI) is modest. The component analysis revealed several interesting active-site interactions of neutral amino acid residues (Table 4, Figure 4). In addition to the generally recognized oxyanion hole interactions and the importance of Asp102 of the catalytic triad, there are several interactions with His57 and the substrates' P1 and P2′ residues which stabilize the tetrahedral transition state and are important for tight binding[36] of the catalytic histidine and the substrate/inhibitor. GlyI36, which is calculated to stabilize TET by 3.8

(27) Branden, C.; Tooze, J. *Introduction to Protein Structure*; Garland Publishing: New York, 1991.
(28) Daggett, V.; Schröder, S.; Kollman, P. A. *J. Am. Chem. Soc.* **1991**, *113*, 8926−8935.
(29) Warshel, A.; Russell, S. *J. Am. Chem. Soc.* **1986**, *108*, 6569−6579.
(30) Warshel, A.; Naray-Szabo, G.; Sussman, F.; Hwang, J.-K. *Biochemistry* **1989**, *28*, 3629−3637.
(31) Daggett, V.; Brown, F.; Kollman, P. *J. Am. Chem. Soc.* **1989**, *111*, 8247−8256.
(32) Tidor, B.; Karplus, M. *Biochemistry* **1991**, *30*, 3217−3228.
(33) Prevost, M.; Wodak, S. J.; Tidor, B.; Karplus, M. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 10880−10884.
(34) Miyamoto, S.; Kollman, P. A. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 8402−8406.
(35) Kollman, P. *Chem. Rev.* **1993**, *93*, 2395−2417.
(36) Lu, W.; Qasim, M. A.; Laskowski, M., Jr.; Kent, S. B. H. *Biochemistry* **1997**, *36*, 673−679.

kcal/mol, has been found to be important for the stability of trypsin−BPTI complex. A GlyI36 → Ala mutation decreases K_i of the trypsin−BPTI complex by ∼5 kcal/mol.[7] GlyI36 is part of the conserved interface sequence and is important in specific binding interactions. As suggested by Figure 4, many of the neutral residues contributing to a stabilizing ΔG_{int} (MICO → TET) do so using their backbone atoms. Nonetheless, it would be interesting to mutate some of them to see if their side chains are critical to align the backbone to help stabilize TET. One could also use backbone-engineered protein or substrate to probe the energetic contribution of backbone hydrogen bonds to binding and catalysis.[36,37] The free energy component of Asp102 of the catalytic triad and Gly192 of the oxyanion hole can be compared with the experimental data on site-specific mutagenesis. Asp102 → Asn mutation decreases k_{cat} by ∼2.5 × 10^4 (6 kcal/mol) in trypsin,[38] and Asn155 → Leu decreases k_{cat} by ∼2.5 × 10^2 (3.3 kcal/mol) in subtilisin.[39] The corresponding estimates from ΔG_{int} component analysis for Asp102 (21.4 kcal/mol) and Gln192 (5.4 kcal/mol) are significant overestimates but provide qualitative values for site-directed mutants. The calculated components are overestimates partially because in mutagenesis experiments removal of an amino acid residue is compensated for by other interactions. Also, structural changes caused by mutations are not accounted for in the free energy component analysis. If more accurate estimates for the effects of mutagenesis are needed, one can do the actual mutations by means of standard free energy perturbation methods,[35,40−42] or other computational approaches can be applied.[43] When free energy component analysis is done in conjunction with QM-FE approach, one gets in a single simulation qualitative estimates for the roles of individual amino acid residues in enzyme catalysis and even for site-directed mutagenesis. This approach should be generally applicable to study the energetics and catalysis of any enzymatic reaction.

Why Does Trypsin Hydrolyze BPTI so Slowly? Trypsin hydrolyses the LysI15−AlaI16 peptide bond of BPTI in a rate that is ∼10^{11} slower than for peptide substrates. Our calculations suggest that a 10^6−10^9 reduction in rate can be accounted for by an increased activation free energy for bond cleavage of BPTI than of substrate in the acylation step. The calculated free energy components can be used to explain the origin of this reduction in structural terms (Table 3). Amino acid residues of trypsin, substrate, and link atoms, which are structural groups common to both trypsin−BPTI and trypsin−substrate, all increase the free energy of forming the transition state in the trypsin−BPTI reaction compared to this process in trypsin−substrate. About half of the total increase in the activation free energy is due to these three groups of residues. This indicates that a part of the favorable protein−protein interactions in the trypsin−BPTI complex are lost upon going from the Michaelis complex to the transition state. Here the ground-state interactions are used to decrease the reaction rate. The second half of the increased activation free energy for bond cleavage in the trypsin−BPTI is due to medium effects. Namely, water molecules, which enter the active site in the trypsin−substrate system, stabilize the transition state of substrate by 5−6 kcal/mol. In the trypsin−BPTI system, BPTI prevents the water molecules from penetrating the active site, and consequently, the stabilization of the transition state by water molecules is lost. The role of water molecules in stabilization of serine protease-catalyzed reaction[29] and in other enzyme reactions has been noted earlier.[44,45] BPTI also affects the reaction rate by electrostatic and direct interactions with the active-site residues. These interactions were calculated to decrease the rate of peptide bond hydrolysis by 9 kcal/mol. Outside the reactive active-site sequence there are charged residues which, by electrostatic interactions, contribute unfavorably to the reaction. However, these residues are on the surface of the protein and are neutralized by the counterions, whose free energy contributions almost exactly compensate for the unfavorable effects of the charged residues. This indicates that the charged surface residues of BPTI are not important for the inhibition of the peptide bond hydrolysis. In the case of trypsin, Soman et al.[43] made a similar observation based on electrical potential calculations using the finite difference Poisson−Boltzmann method. They found that active sites of cow and rat trypsin are effectively shielded from surface charges. Surface charges may be important for the activity of protease in physiological environment and for specificity and tight binding between protease inhibitor and protease.[2,43] We note that more rigorous ways (such as reaction field, Ewald, or multipole methods) to include long-range electrostatic effects could be employed to get more accurate estimates of the effects of charged residues and counterions on the calculated free energy components and activation energies. However, the fact that we used a large cutoff value (22 Å) for nonbonded interactions, showed that the free energies were relatively insensitive to whether a 12 or 22 Å cutoff was used, and the fact that we made comparisons between similar systems suggests that the methods used are adequately accurate for our purposes.

In addition to the Michaelis complex to the tetrahedral intermediate reaction (Figure 6, $\Delta G_1^{\#}$), the step from the tetrahedral intermediate to the acyl-enzyme intermediate ($\Delta G_2^{\#}$) and the step leading to the entering of an acyl group-hydrolyzing water molecule to the active site and hydrolysis of acyl-enzyme ($\Delta G_3^{\#}$) may both or either one of them contribute to the slower hydrolysis of BPTI than of substrate. MD simulations on trypsin−BPTI and trypsin−substrate tetrahedral intermediates suggested that, although the structure of the active site is slightly more favorable for peptide bond hydrolysis in the case of substrate than that of BPTI, it is likely that $\Delta\Delta G_2^{\#}$ would be at the most a few kcal/mol larger than $\Delta\Delta G^*$. On the other hand, simulations on trypsin−BPTI and trypsin−substrate acyl-enzyme intermediates showed clearly different behavior for the two systems. In the substrate simulation, the cleaved part of the substrate changed conformation after 150 ps of simulation leading to a structure in which reformation of the peptide bond is not likely and which is on the path to diffusion of the cleaved peptide from the active site, allowing a water molecule to enter the active site. In contrast to this, MD simulation on the trypsin−BPTI acyl-enzyme intermediate suggested that the NH_2 terminus stays close to the ester group of acyl-enzyme in an orientation where the reformation of the peptide bond takes place readily. That the NH_2 group released on cleavage of the scissile bond remains in its original position is supported by an NMR study on chymotrypsin inhibitor 2 (CI2) in which the active-site

(37) Shin, I.; Ting, A. Y.; Schultz, P. G. *J. Am. Chem. Soc.* **1997**, *119*, 12667−12668.
(38) Craik, C. S.; Roczniak, S.; Largman, C.; Rutter, W. J. *Science* **1987**, *237*, 909−913.
(39) Bryan, P.; Pantoliano, M. W.; Quill, S. G.; Hsiao, H.-Y.; Poulos, T. *Proc. Natl. Acad. Sci. U.S.A.* **1986**, *83*, 3743−3745.
(40) Rao, S. N.; Singh, U. C.; Bash, P. A.; Kollman, P. A. *Nature* **1987**, *328*, 551−554.
(41) Warshel, A.; Sussman, F.; Hwang, J.-K. *J. Mol. Biol.* **1988**, *201*, 139−159.
(42) Simonson, T.; Brünger, A. T. *Biochemistry* **1992**, *31*, 8661−8674.
(43) Soman, K.; Yang, A.-S.; Honig, B.; Fletterick, R. *Biochemistry* **1989**, *28*, 9918−9926.
(44) Peräkylä, M.; Kollman, P. A. *J. Am. Chem. Soc.* **1997**, *119*, 1189−1196.
(45) Jensen, G. M.; Warshel, A.; Stephens, P. J. *Biochemistry* **1994**, *33*, 10911−10924.

scissile bond was cleaved.[46] CI2 is a small (83 residues) inhibitor of serine proteases having a mechanism of inhibition similar to that of BPTI. The NMR data on CI2 indicated that the loop region of CI2, where the cleaved NH_2 terminus is located, stays relatively ordered even in an uncomplexed inhibitor. It is probable that the relative rigidity of the active-site loop would be enhanced by interactions with the protease favoring the uncleaved complex. These data suggest that the rest of the rate reduction in trypsin−BPTI compared to that in the trypsin−substrate system, 10^2-10^5, could come from a relative rate for reformation of the peptide bond that is larger than the rate for formation of product and free enzyme. Also, the high stability and rigid structure of the trypsin−BPTI complex are beneficial for the inhibitory action of the protease inhibitor. Thus, larger dissociation constants of the protein inhibitor−protease complexes are predicted to increase the hydrolysis of peptide bonds by decreasing the activation free energy barrier for peptide bond cleavage and by increasing the rate of formation of free enzyme and modified inhibitor from the acyl-enzyme intermediate. Support for this suggestion is found from the K_m and k_{cat} values of protease inhibitors (Kunitz family) of serine proteases.[3,47] For the inhibitors the increase in K_m generally leads to increase in k_{cat} values. Also, in some cases there are similar changes in K_m and k_{cat} values, resulting in quite similar k_{cat}/K_m values despite large changes in individual numbers. These individual values can be explained by the importance of binding energy in the inhibitory function of serine protease inhibitors.

Conclusions

We have presented QM-FE calculations on trypsin−substrate and trypsin−BPTI systems. Because we are studying the same process as those previously presented by Stanton et al.,[8] we do not need to do any additional quantum mechanical calculations, merely calculate the classical free energy difference between the Michaelis complexes (MICO) and the tetrahedral intermediate (TET) model for the transition state. The two additional substrates studied here yield $\Delta G^{\#}$ estimates in good agreement with that calculated previously[8] and with the value found experimentally for specific amide substrates. The value calculated for BPTI as substrate suggests a decrease in catalytic rate by 10^6-10^9, in qualitative agreement with the 10^{11} found experimentally. We speculate that the remaining 10^2-10^5 decrease in rate arises from the fact that the peptide bond, once cleaved, is much more likely to re-ligate to form the intact inhibitor than to form the acyl-enzyme and have the NH_2 terminal part of the inhibitor diffuse away. We support this speculation with molecular dynamics simulations of the tetrahedral intermediate of BPTI and substrate and simulations of the cleaved BPTI and substrate attached to trypsin. These simulations show that through much of the 500 ps trajectory, the NH_2 end of cleaved BPTI is in excellent position to attack the acyl C═O and regenerate intact BPTI. In contrast to this, the NH_2 end of the cleaved peptide substrate moved away (in 200 ps) from the vicinity of the acyl C═O favoring hydrolysis of acyl-enzyme intermediate.

We have also shown the power of free energy component analysis to give insight into the nature of enzyme catalysis. We find that the free energy of stabilization of TET over MICO comes substantially from Asp102 and the oxyanion hole, as expected, but there are surprisingly large contributions from some other backbone interactions. We suggest that mutating these might be interesting to examine induced effects on k_{cat} for trypsin. Free energy components are also useful for understanding why the $\Delta G^{\#}$ for MICO → TET is ∼8−13 kcal/mol less favorable for BPTI than for a normal substrate. We suggest that this comes from the stronger interactions in MICO that must be broken to go to TET and the fact BPTI, unlike a small substrate, does not allow water to approach as closely to and stabilize the transition state (tetrahedral intermediate) for amide hydrolysis. Thus, these calculations further demonstrate the power of QM-FE calculations to connect enzyme structure to catalytic action.

Acknowledgment. P.A.K. acknowledges the grant support of the NIH (GM-29072). M.P. is grateful to the Academy of Finland for support.

JA991602P

(46) Shaw, G. L.; Davis, B.; Keeler, J.; Fersht, A. R. *Biochemistry* **1995**, *34*, 2225−2233.
(47) Otlewski, J.; Zbyryt, T. *Biochemistry* **1994**, *33*, 200−207.
(48) Kraulis, P. *J. Appl. Crystallogr.* **1991**, *24*, 924−950.
(49) Evans, S. *J. Mol. Graphics* **1993**, *11*, 134−138.