Computational Study of Conformational Transitions in the Active Site of Tosyl-α-chymotrypsin

Themis Lazaridis[†] and Michael E. Paulaitis*

Contribution from the Center for Molecular and Engineering Thermodynamics, Department of Chemical Engineering, University of Delaware, Newark, Delaware 19716

Received September 18, 1992[⊕]

Abstract: The motion of the tosyl group covalently bound in the active site of α-chymotrypsin has been studied using an empirical force field. A previously developed method employing an adiabatic projection of the potential energy hypersurface onto a small number of essential degrees of freedom has been used to determine multidimensional reaction paths for the rotation of the tosyl aromatic ring within and its movement out of the specificity pocket of the enzyme. The conformation of the specificity pocket was found to change under stochastic boundary, molecular dynamics simulations. The shape of the paths and the height of the calculated energy barriers were affected substantially by this conformational change. The free-energy profile for tosyl-ring rotation was also determined by free-energy perturbation in the presence and absence of explicit solvent water and in both cases was found to be lower than the energy barrier, indicating a positive activation entropy. On the basis of the calculated energy and free-energy barriers, the mobility of the bound inhibitor was found to be much lower than that observed in NMR experiments on the same system. Our calculations also indicate that the time scales for movement of the tosyl ring out of the pocket are much slower than that for rotation in the pocket. This result implies that the measured NMR correlation times for ring rotation are an average of correlation times for rotation in the pocket and in solution.

Introduction

The dynamic properties of biological macromolecules have attracted considerable attention over the past two decades (for reviews, see refs 1–6). Several experimental techniques, such as hydrogen exchange,[7] fluorescence quenching,[8] NMR,[9] laser flash photolysis,[10] and even X-ray crystallography,[11,12] have been used to firmly establish the existence of significant conformational fluctuations in the native, folded state of globular proteins. Several theoretical models emphasize the potential importance of such conformational fluctuations in enzyme function,[13–16] but experimental evidence for the direct implication of these fluctuations in catalytic activity is still lacking. One example where their importance is obvious, however, is myoglobin, which must undergo substantial conformational fluctuations to bind ligands.[17,18]

Motivation for this study stems from experiments carried out in our laboratories[19] and by others[20] to determine the effect of hydration on the catalytic activity of enzyme powders. In this work, a variety of experimental techniques have been used to show that both protein conformational dynamics and catalytic activity are linked to enzyme hydration. For example, the rate of hydrogen exchange and the rate of rotation of ESR spin probes in dry protein powders increase with the level of hydration, the latter also showing a high correlation with enzyme activity.[21] The level of hydration can also dramatically influence glass-transition-like phenomena in protein powders as measured, for example, by the mobility of Mössbauer labels embedded in protein molecules.[22,23] Dielectric relaxation measurements[24–27] have shown increased mobility of the protein as hydration level is increased. Finally, measurements of the catalytic activity of enzyme powders suspended in organic solvents in many cases show a rapid increase in activity with hydration level.[19,28]

The influence of hydration and dynamics on enzyme activity would best be elucidated by studies on proteins with well-characterized chemical reaction mechanisms, such as α-chymotrypsin.[29] This serine protease catalyzes the hydrolysis of amide and ester bonds preferably at sites bearing aromatic groups.[30,31]

(1) Gurd, F. R. N.; Rothgeb, T. M. *Adv. Protein Chem.* **1979**, *33*, 74.
(2) Debrunner, P. G.; Frauenfelder, H. *Annu. Rev. Phys. Chem.* **1982**, *33*, 283.
(3) Levitt, M. *Annu. Rev. Biophys. Bioeng.* **1982**, *11*, 251.
(4) Karplus, M.; McCammon, J. A. *Annu. Rev. Biochem.* **1983**, *53*, 263.
(5) McCammon, J. A.; Harvey, S. C. *Dynamics of Proteins and Nucleic Acids*; Cambridge University Press: Cambridge, 1987.
(6) Brooks, C. L., III; Karplus, M.; Pettitt, B. M. *Proteins: A theoretical perspective of dynamics, structure, and thermodynamics; Adv. Chem. Phys.* **1988**, *71*.
(7) Woodward, C. K.; Hilton, B. D. *Annu. Rev. Biophys. Bioeng.* **1979**, *8*, 99.
(8) Lakowicz, J. R.; Weber, G. *Biochemistry* **1973**, *12*, 4171.
(9) Wagner, G. *Q. Rev. Biophys.* **1983**, *16*, 1.
(10) Frauenfelder, H.; Parak, F.; Young, R. D. *Annu. Rev. Biophys. Biophys. Chem.* **1988**, *17*, 451.
(11) Frauenfelder, H.; Petsko, G. A.; Tsernoglou, D. *Nature* **1979**, *280*, 558.
(12) Artymiuk, P. J.; Blake, C. C. F.; Grace, D. E. P.; Oatley, S. J.; Phillips, D. C.; Sternberg, M. J. E. *Nature* **1979**, *280*, 563.
(13) Welch, G. R., Ed. *The Fluctuating Enzyme*; Wiley: New York, 1986.
(14) Cooper, A. *Prog. Biophys. Mol. Biol.* **1984**, *44*, 181.
(15) Lumry, R.; Rosenberg, A. *Coll. Int. CNRS* **1975**, *246*, 53.
(16) Careri, G.; Fasella, P.; Gratton, E. *Annu. Rev. Biophys. Bioeng.* **1979**, *8*, 69.
(17) Perutz, M. F.; Mathews, F. S. *J. Mol. Biol.* **1966**, *21*, 199.
(18) Case, D. A.; McCammon, J. A. *Ann. N. Y. Acad. Sci.* **1986**, *482*, 222.

(19) Paulaitis, M. E.; Sowa, M. J.; McMinn, J. H. *Ann. N. Y. Acad. Sci.* **1992**, *672*, 278.
(20) Rupley, J. A.; Careri, G. *Adv. Protein Chem.* **1991**, *41*, 37.
(21) Rupley, J. A.; Gratton, E.; Careri, G. *Trends Biochem. Sci.* **1983**, *8*, 18.
(22) Belonogova, O. V.; Frolov, E. N.; Illyustrov, N. V.; Likhtenshtein, G. I. *Mol. Biol. (Engl. Transl.)* **1979**, *13*, 428.
(23) Parak, F.; Frolov, E. N.; Mossbauer, R. L.; Goldanskii, V. I. *J. Mol. Biol.* **1981**, *145*, 825.
(24) Bone, S.; Pethig, R. *J. Mol. Biol.* **1982**, *157*, 571.
(25) Bone, S.; Pethig, R. *J. Mol. Biol.* **1985**, *181*, 323.
(26) Bone, S. *Biochim. Biophys. Acta* **1987**, *916*, 128.
(27) Singh, G. P.; Parak, F.; Hunklinger, S.; Dransfeld, K. *Phys. Rev. Lett.* **1981**, *47*, 685.
(28) Zaks, A.; Klibanov, A. M. *J. Biol. Chem.* **1988**, *263*, 8017.
(29) Schowen, R. L. In *Mechanistic Principles of Enzyme Activity*, Molecular Structure and Energetics; Liebman, J. F., Greenberg, A. Eds.; VCH: New York, 1988; Vol. 9.
(30) Fersht, A. *Enzyme structure and mechanism*, 2nd ed.; Freeman: New York, 1985.
(31) Fink, A. L. In *Enzyme Mechanisms*; Page, M. I., Williams, A. Eds.; Royal Society of Chemistry: London, 1987; chapter 10.

Tosyl-α-chymotrypsin

J. Am. Chem. Soc., Vol. 116, No. 4, 1994 1547

The origin of this specificity is evident in the crystal structures of the enzyme with inhibitors, where the aromatic groups are observed to bind in a hydrophobic "cleft" near the active site of the enzyme.[32–34] However, the mechanism of substrate specificity exhibited by the serine proteases is complex and still poorly understood.[35,36]

One interesting set of NMR experiments to probe the dynamics in the active site of α-chymotrypsin has shown that the tosyl ring in tosyl-chymotrypsin rotates rapidly in solution,[37] despite the fact that the tosyl group appears to bind tightly in the hydrophobic pocket in the crystal structure.[32] The correlation time for rotation of the ring was estimated via a diffusive model for ring motion to be about 1.1 ns at neutral pH and 0.14 ns at pH 4. It was also concluded from chemical shift measurements that the environment around the ring is predominantly polar. These observations could be explained by two possibilities: either the hydrophobic pocket relaxes in solution and allows solvent to enter and the ring to rotate rapidly or the ring does not reside in the pocket at all times but equilibrates between the conformation in the pocket and one or more conformations out of the pocket. This work raises the question of how flexible the active site is and how this flexibility might affect the catalytic activity. Since proper orientation of the substrate is required for catalytic reaction,[30] the degree of immobilization of the substrate in the specificity pocket of the enzyme will presumably influence the rate of hydrolysis of this substrate. Also, the dynamics of substrate binding may be another important factor in the expression of enzyme specificity.

The catalytic reaction pathway of serine proteases has been the subject of several quantum mechanical calculations[38–47] and of an empirical valence-bond theory.[48,49] The coupling of the dynamics of active site residues with the catalytic reaction has been addressed in a recent quantum mechanical model.[50] In addition, binding energies of substrates to chymotrypsin have been calculated by molecular mechanics methods.[51–53] Molecular dynamics simulations with empirical force fields have been performed on a number of serine proteases, including trypsin,[54–56] elastase,[57–61] and *Streptomyces griseus* protease A (SGPA).[62] A molecular dynamics study of chymotrypsin complexed with

enantiomeric substrates has been recently performed to determine the molecular origin of the relative stability of two types of acyl-enzymes.[63] These studies, however, have not addressed the dynamic aspects of serine protease catalysis, such as binding in the hydrophobic pocket and the mobility of the catalytic groups.

With existing computational power, conventional molecular dynamics simulations can probe the dynamics of protein molecules over a relatively short (subnanosecond) time scale. However, most functionally significant protein motions are thought to occur on a much longer time scale.[4] Thus, the calculation of rates of these processes cannot be accomplished by conventional methods. One promising alternative route for doing so is based on transition-state theory[64–67] and involves calculating the probability of finding the system at the transition state or, equivalently, the free-energy barrier. The reactive flux from the transition state (TS) can also be calculated to account for dynamic effects present in a condensed phase.[68] This approach was first applied to protein dynamics by Northrup et al.,[69] who studied the rate of rotation of tyrosine aromatic rings in BPTI. In that study, the free-energy profile was calculated by umbrella sampling. The rate of rotation, corrected for dynamic effects, was several orders of magnitude higher than the rate determined experimentally. Closer agreement with experiment was attained in later work with further improvements in the methodology and by including explicit solvent in the simulation.[70,71] Still, the calculated rate constant was 2 orders of magnitude higher than the experimental value.

One of the key requirements for the application of TS theory is the knowledge of the reaction coordinate or reaction path (RC). In the aforementioned work on tyrosine-ring rotations, the RC was empirically chosen as the χ_2 dihedral angle of tyrosine[72] or the difference between this and a virtual dihedral angle reflecting the major interactions of the ring with the surrounding protein.[69] The RC or intrinsic reaction coordinate[73] is defined as a path, tangent to the gradient, on the potential energy hypersurface that links the reactant and product states through a first-order saddle point (the TS). Several methods for determining RCs are now available. One approach is to calculate the optimal path between reactant and product states by energy minimizations,[74] Monte Carlo,[75] or variational methods.[76] Another approach is to first determine the saddle point or TS, usually by second-derivative methods (e.g., ref 77), and then follow the gradient downhill in energy to the reactant and product states. For small reactive systems, the calculation of reaction coordinates via second-derivative methods can be routinely performed with most currently

(32) Birktoft, J. J.; Blow, D. M. *J. Mol. Biol.* **1972**, *68*, 187.

(33) Fujinaga, M.; Sielecki, A. R.; Read, R. J.; Ardelt, W.; Laskowski, M., Jr.; James, M. N. G. *J. Mol. Biol.* **1987**, *195*, 397.

(34) Tulinsky, A.; Blevins, R. A. *J. Biol. Chem.* **1987**, *262*, 7737.

(35) Craik, C. S.; Largman, C.; Fletcher, T.; Roczniak, S.; Barr, P. J.; Fletterick, R.; Rutter, W. J. *Science* **1985**, *228*, 291.

(36) Hedstrom, L.; Szilagyi, L.; Rutter, W. J. *Science* **1992**, *255*, 1249.

(37) Ando, M. E.; Gerig, J. T.; Weigand, E. F. *J. Am. Chem. Soc.* **1982**, *104*, 3172.

(38) Scheiner, S.; Kleier, D. A.; Lipscomb, W. N. *Proc. Natl. Acad. Sci. U.S.A.* **1975**, *72*, 2606.

(39) Schneiner, S.; Lipscomb, W. N. *Proc. Natl. Acad. Sci. U.S.A.* **1976**, *73*, 432.

(40) Dewar, M. J. S.; Storch, D. M. *Proc. Natl. Acad. Sci. U.S.A.* **1985**, *82*, 2225.

(41) Kollman, P. A.; Hayes, D. M. *J. Am. Chem. Soc.* **1981**, *103*, 2955.

(42) Daggett, V.; Schroder, S.; Kollman, P. A. *J. Am. Chem. Soc.* **1991**, *113*, 8926.

(43) Weiner, S. J.; Seibel, G. L.; Kollman, P. A. *Proc. Natl. Acad. Sci. U.S.A.* **1986**, *83*, 649.

(44) Wipff, G.; Dearing, A.; Weiner, P. K.; Blaney, J. M.; Kollman, P. A. *J. Am. Chem. Soc.* **1983**, *105*, 997.

(45) Umeyama, H.; Imamura, A.; Nagata, C. *J. Theor. Biol.* **1973**, *41*, 485.

(46) Stamato, F. L. M. G.; Tapia, O. *Int. J. Quantum Chem.* **1988**, *33*, 187.

(47) Lamotte-Brasseur, J.; Dive, G.; Dehareng, D.; Ghuysen, J.-M. *J. Theor. Biol.* **1990**, *145*, 183.

(48) Warshel, A.; Russell, S. *J. Am. Chem. Soc.* **1986**, *108*, 6569.

(49) Warshel, A.; Sussman, F. *Proc. Natl. Acad. Sci. U.S.A.* **1986**, *83*, 3806.

(50) Sumi, H.; Ulstrup, J. *Biochim. Biophys. Acta* **1988**, *955*, 26.

(51) Platzer, K. E. B.; Momany, F. A.; Scheraga, H. A. *Int. J. Pept. Protein Res.* **1972**, *4*, 201.

(52) DeTar, D. F. *Biochemistry* **1981**, *20*, 1730.

(53) DeTar, D. F. *J. Am. Chem. Soc.* **1981**, *103*, 107.

(54) Wong, C. F.; McCammon, J. A. *Isr. J. Chem.* **1986**, *27*, 211.

(55) Wong, C. F.; McCammon, J. A. *J. Am. Chem. Soc.* **1986**, *108*, 3830.

(56) Brünger, A. T.; Huber, R.; Karplus, M. *Biochemistry* **1987**, *26*, 5153.

(57) Geller, M.; Carlson-Golab, G.; Lesyng, B.; Swanson, S. M.; Meyer, E. F., Jr. *Biopolymers* **1990**, *30*, 781.

(58) Geller, M.; Swanson, S. M.; Meyer, E. F., Jr. *J. Am. Chem. Soc.* **1990**, *112*, 8925.

(59) Lesyng, B.; Meyer, E. F., Jr. *J. Comput.-Aided Mol. Des.* **1987**, *1*, 211.

(60) Lesyng, B.; Meyer, E. F., Jr. *Biopolymers* **1990**, *30*, 773.

(61) Fujita, T.; Meyer, E. F., Jr. *J. Comput. Chem.* **1987**, *8*, 801.

(62) Avbelj, F.; Moult, J.; Kitson, D. H.; James, M. N. G.; Hagler, A. T. *Biochemistry* **1990**, *29*, 8658.

(63) Bemis, G. W.; Carlson-Golab, G.; Katzenellenbogen, J. A. *J. Am. Chem. Soc.* **1992**, *114*, 570.

(64) Keck, J. C. *Discuss. Faraday Soc.* **1962**, *33*, 173.

(65) Anderson, J. B. *J. Chem. Phys.* **1973**, *58*, 4684.

(66) Bennett, C. H. In *Algorithms for chemical computations*; Christofferson, R. E., Ed.; American Chemical Society: Washington, DC, 1977; pp 63–97.

(67) Berne, B. J. In *Multiple Time Scales*; Brackbill, J. U., Cohen, B. I., Eds.; Academic Press: New York, 1985.

(68) Chandler, D. *J. Chem. Phys.* **1978**, *68*, 2959.

(69) Northrup, S. H.; Pear, M. R.; Lee, C.-Y.; McCammon, J. A.; Karplus, M. *Proc. Natl. Acad. Sci. U.S.A.* **1982**, *79*, 4035.

(70) Ghosh, I.; McCammon, J. A. *Biophys. J.* **1987**, *51*, 637.

(71) Ghosh, I.; McCammon, J. A. *J. Phys. Chem.* **1987**, *91*, 4878.

(72) McCammon, J. A.; Karplus, M. *Proc. Natl. Acad. Sci. U.S.A.* **1979**, *76*, 3585.

(73) Fukui, K. *Acc. Chem. Res.* **1981**, *14*, 363.

(74) Elber, R.; Karplus, M. *Chem. Phys. Lett.* **1987**, *139*, 375.

(75) Pratt, L. R. *J. Chem. Phys.* **1986**, *85*, 5045.

(76) Gillilan, R. E.; Wilson, K. R. *J. Chem. Phys.* **1992**, *97*, 1757.

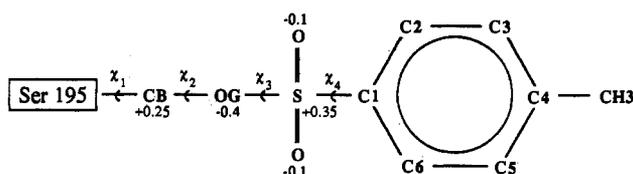(77) Cerjan, C. J.; Miller, W. H. *J. Chem. Phys.* **1981**, *75*, 2800.

Figure 1. Schematic of the tosyl group bound to the active site Ser 195 in α-chymotrypsin showing the partial charges on the atoms and the four explicit dihedral angles considered in the calculations.

available molecular orbital software packages. However, this approach rapidly becomes impractical as the number of atoms in the system increases.

Recently, methods for obtaining RCs for conformational transitions in larger systems have been developed by Elber and co-workers. These methods are based on the minimization of the sum of the energies of structures between the reactant and product states[74],[78] or on the energy minimizations orthogonal to a reference coordinate[79]–[81] and have been applied to transitions of myoglobin between local minima[74], α-helix formation of model tetrapeptides,[79],[80],[82] and diffusion of oxygen in myoglobin.[83] These and other studies[84] have been carried out in the full Cartesian coordinate space of the molecule which gives equal weight to all internal degrees of freedom.

In previous work,[85] we proposed an alternative method for determining RCs of conformational transitions in large systems. This method employs an adiabatic projection of the full internal coordinate space onto a smaller number of "essential" torsional degrees of freedom needed to describe the process under study. Here, we extend this work to study the dynamics of the tosyl ring in the active site of tosyl-α-chymotrypsin. Reaction paths for rotation of the tosyl ring within the pocket and its movement out of the pocket are determined, and free-energy profiles for ring rotation in vacuum and in explicit water are calculated by free-energy perturbation (FEP). The motions of catalytically active residues that occur during the simulations are also briefly discussed.

Methods

The program CHARMM,[86] augmented with code to perform the TS, RC, and FE calculations, was used for all simulations. The hydrogens that can participate in hydrogen bonds were included explicitly, while the nonpolar hydrogens were incorporated into the heavy atom to which they were bonded to form "united" atoms. The protonation state of all titratable groups was that corresponding to neutral pH. In particular, histidines were singly protonated (uncharged) and aspartates (including the catalytic Asp 102) were negatively charged. The catalytic His 57 was protonated at $N^{\delta 1}$, on the side of Asp 102. The tosyl group was built independently and attached to Ser 195 with the PATCH facility of CHARMM. The partial charges for the sulfur and oxygen atoms in the serine–tosyl complex were determined with Polygen's CHEMNOTE program and are shown in Figure 1. The partial charges of all other (united) atoms are zero.

The parameter set CHARMm 21A, distributed by Polygen, was used for the potential energy function. The potential parameters for the sulfonyl atoms in the tosyl group were not available and therefore were approximated by parameters for similar atom types in the CHARMm 21A parameter set. One particular parameter that did require special

(78) Czerminski, R.; Elber, R. *Int. J. Quantum Chem., Quantum Chem. Symp.* **1990**, *24*, 167.

(79) Czerminski, R.; Elber, R. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 6963.

(80) Czerminski, R.; Elber, R. *J. Chem. Phys.* **1990**, *92*, 5580.

(81) Ulitsky, A.; Elber, R. *J. Chem. Phys.* **1990**, *92*, 1510.

(82) Choi, C.; Elber, R. *J. Chem. Phys.* **1991**, *94*, 751.

(83) Nowak, W.; Czerminski, R.; Elber, R. *J. Am. Chem. Soc.* **1991**, *113*, 5627.

(84) Fischer, S.; Karplus, M. *Chem. Phys. Lett.* **1992**, *194*, 252.

(85) Lazaridis, T.; Tobias, D. J.; Brooks, C. L., III; Paulaitis, M. E. *J. Chem. Phys.* **1991**, *95*, 7612.

(86) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187.

consideration was the force constant (or intrinsic barrier) for rotation of the dihedral angle χ_4 (see Figure 1). The torsional potential given in CHARMm 21A has 3-fold periodicity with a force constant of 3.6 kcal/mol, which would result in a barrier of 7.2 kcal/mol. First, 3-fold periodicity is clearly incorrect for the rotation of an aromatic ring, and second, this barrier height appears to be much too high. Semiempirical calculations (CNDO/2) give values from 1.7 to 4.1 kcal/mol for the torsional barrier of benzenesulfonyl chloride, depending on the basis set used.[87],[88] In view of this uncertainty, all calculations reported herein have been performed without an intrinsic torsional barrier for this angle and reflect the effect of nonbonded interactions alone. However, when the calculations for ring rotation were repeated with an intrinsic torsional barrier of 3 kcal/mol, the same energy barrier and similar reaction paths were obtained. Apparently, the intrinsic torsional barrier does not simply add to the total barrier because the barrier is primarily determined by the strong nonbonded interactions between the ring and the surrounding protein groups.

The crystal structure of tosyl-α-chymotrypsin,[32] obtained from the Brookhaven Protein Data Bank,[89] was used as the starting structure. In this crystal structure, the positions of residues 9–13 are not shown and therefore these residues are omitted. Since this region is remote from the active site, omission of these residues should not affect the results of the present study. The explicit (polar) hydrogens were generated using the HBUILD facility of CHARMM.[90] The 50 crystallographically determined water molecules were also included at this point in the calculations to give a total atom count of 2292. Subsequently, all residues that did not have an atom within 18 Å from atom C1 of the tosyl group were deleted (138 residues, including 33 waters), bringing the total number of atoms down to 1242. The energy was then minimized with 100 steepest descent steps followed by 300 steps of adopted basis Newton–Raphson (ABNR) minimization.[86] In this minimization, only the atoms within 16 Å around tosyl C1 were allowed to move. An additional set of 500 ABNR steps were necessary for adequate convergence of the energy. After this initial stage of minimizations, only a slow relaxation of the energy was observed. The reproducibility of the adiabatic energy for a given set of the essential torsional angles was on the order of a few tenths of a kcal/mol, which was reasonable for a system this size and satisfactory for our purposes. A higher level of accuracy could be achieved by longer minimizations, at the expense of significantly more computer time. The energy-minimized crystal structure is shown in Figure 2.

The determination of TSs and RCs was performed on a reduced torsional adiabatic surface.[85] Four degrees of freedom were considered explicitly: the angles χ_1, χ_2, χ_3, and χ_4 shown in Figure 1. All remaining degrees of freedom were allowed to relax to their optimal values so as to minimize the energy. The use of more than one mapping degree of freedom is essential in this study; for example, if only χ_4 were varied while the energy was minimized, the other three angles could easily adjust to keep the ring in its most favorable orientation in the pocket. This approach assumes that the adiabatically projected structure for each set of the explicit degrees of freedom is unique. This assumption appeared reasonable, given previous observations that the protein matrix in the region of the active site of chymotrypsin remained in a single conformation during molecular mechanics energy minimizations.[52],[53] Another assumption is that a saddle point on this adiabatic surface corresponds to a saddle point on the full potential energy surface. The validity of this assumption for complex macromolecules is discussed elsewhere,[85] although a rigorous proof is lacking. No evidence for a breakdown in this assumption was obtained in this work. All TSs that were found by this method appeared to be true saddle points connecting local minima on the macromolecular energy hypersurface. The adiabatic surface was reached by ABNR minimizations with the four explicit dihedral angles constrained. In these calculations, only the atoms within 11 Å from tosyl C1 in the initial crystal structure (292 atoms) were allowed to move.

For the TS determination, two approaches were tested. The first was the Cerjan–Miller algorithm[77] implemented in the four-dimensional space of angles χ_1–χ_4. This method uses local information on the second derivatives of the adiabatic energy surface to walk uphill in the direction of the saddle point. The Hessian of the adiabatic surface was calculated numerically by performing small changes in the four dihedral angles and minimizing the energy with respect to all other degrees of freedom. The

(87) Boggia, L. K.; Filgueira, R. R.; Maranon, J.; Sorarrain, O. M. *Spectrosc. Lett.* **1978**, *11*, 143.

(88) van Eijck, B. P.; Hargittai, I.; Mayer, I. *J. Mol. Struct.* **1980**, *69*, 301.

(89) Bernstein, F. C. et al. *J. Mol. Biol.* **1977**, *112*, 535.

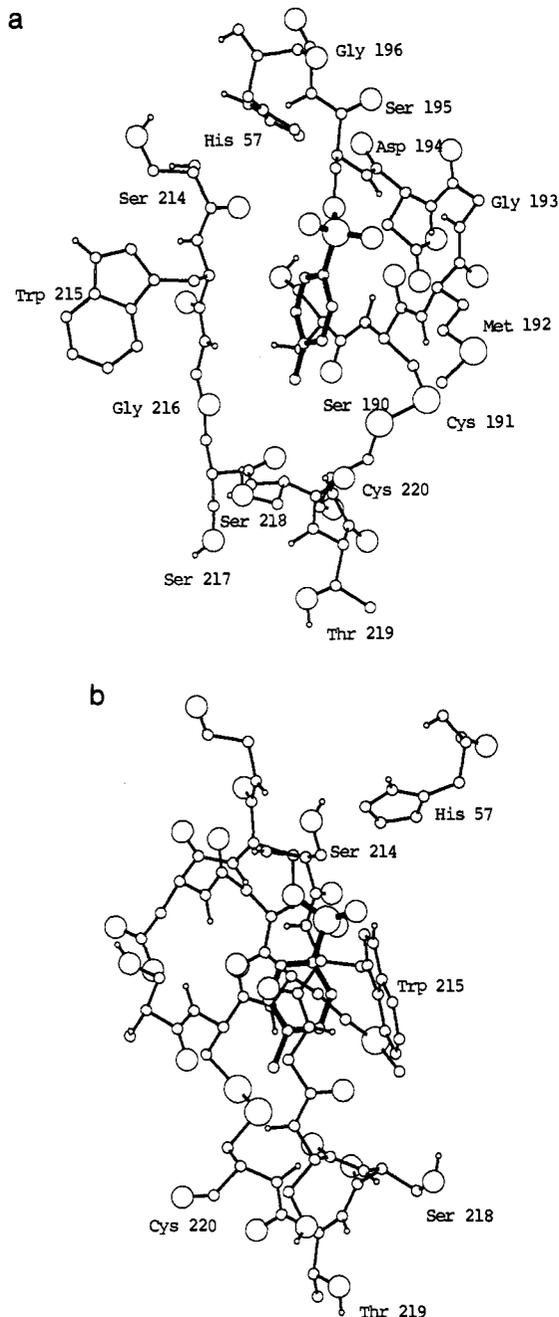(90) Brünger, A.; Karplus, M. *Proteins* **1988**, *4*, 148.

Tosyl-α-chymotrypsin

J. Am. Chem. Soc., Vol. 116, No. 4, 1994 1549

a

b

Figure 2. Conformation of minimum energy in the active site of tosyl-α-chymotrypsin (energy-minimized crystal structure). The tosyl ring is shown with heavy, dark bonds: (a) face view and (b) side view looking from the 214–217 side after rotation around the vertical direction by 90°.

increment for the angles was approximately 4°, although higher values were used in regions where the adiabatic energy surface was relatively flat. This method was initially found to work but was computationally expensive due to the required minimizations; for a single evaluation of the 4 × 4 Hessian, a total of 20 minimizations are needed, each consisting of a minimum of 150 ABNR steps.

Therefore, we examined methods for the determination of the TS that use only first derivatives. The great advantage of first derivatives in the present internal coordinate approach is that they can be calculated by rigid changes of the dihedral angles instead of energy minimizations. This is because the relationship

$$\frac{\partial E_{adiab}}{\partial \chi_i} = \frac{\partial E}{\partial \chi_i} \qquad (1)$$

where E is the energy and E_{adiab} the adiabatic energy is always valid. However, the second derivative of the adiabatic energy cannot be calculated from rigid-body changes because the analogous expression for the second

derivative

$$\frac{\partial^2 E_{adiab}}{\partial \chi_i{}^2} = \frac{\partial^2 E}{\partial \chi_i{}^2} \qquad (2)$$

is only true in the limit of weak coupling between internal degrees of freedom, which cannot be assumed in this study. The method we used is similar to those developed by Elber and co-workers for application in Cartesian coordinate space.[80,81] This method involves the generation of a chain of structures between the initial and final conformation and the subsequent updating of this chain until, in the limit of convergence, it becomes tangent to the gradient at all points. In our case, each point on the chain is a set of the four explicit angles. The initial chain is simply a linear interpolation between the initial and final structures in internal coordinate (IC) space. The partial derivatives $\partial E/\partial \chi_i$ are calculated numerically by rigid moves of each of the four dihedral angles (an increment of 0.1° was used), and this vector is projected onto the direction perpendicular to the local direction of the chain. The latter is approximated by $\chi^{k+1} - \chi^{k-1}$, where k is the number of the current structure on the chain. Then, a step in the resultant direction is taken and the procedure repeated. Although this method can be used to obtain the whole path (with some appropriate modification accounting for the metric of the internal coordinate space), we only used it to determine the TS. After the general region of the TS had been located, an accurate determination of the TS was obtained by adding more structures to the chain in that region and repeating the above procedure.

After the TS was determined, the RC was calculated by following the gradient as prescribed elsewhere,[85] that is, at each point we follow the contravariant components of the gradient

$$v^i = g^{ij}v_j \qquad (3)$$

where

$$g^{ij} = \sum_{k=1}^{3N} \frac{\partial q^i}{\partial x^k} \frac{\partial q^j}{\partial x^k} \qquad (4)$$

is the contravariant metric tensor in internal coordinate space and the

$$v_i = \frac{\partial E}{\partial q^i} \qquad (5)$$

are the covariant components of the gradient. In the above equations, the q^i are internal and x^k are Cartesian coordinates. The values g^{ij} are calculated as a function of the internal coordinates by the expressions tabulated by Decius.[91] The steepest descent (SD) path is initiated at the TS either by moving in both directions corresponding to the negative eigenvector or, when the chain method is used, by starting from the two structures straddling the TS.

The FE profile for the conformational transitions was calculated by performing a number of simulations at regularly placed points along the RC and sampling a hyperplane orthogonal to the RC. The relevant constraint is a linear combination of the four dihedral angles χ_1–χ_4:

$$v_i q^i = \text{constant} \qquad (6)$$

This constraint was implemented as in ref 85. The values of the partial derivatives were first smoothed to remove noise that results from incomplete convergence of the energy minimizations and small fluctuations around the path inherent to gradient following. At each point in the simulation, the structure was "perturbed" to the previous and following hyperplane by vectors parallel to the RC. The difference in energy was calculated and stored, and the free-energy difference was then obtained from

$$\Delta A_{1\to2} = -k_B T \ln \left\langle \exp\left(-\frac{E_2 - E_1}{kT}\right)\right\rangle_1 \qquad (7)$$

For the simulations, we used stochastic boundary conditions.[92,93] The protein was partitioned into three regions. The first contained all residues that have at least one atom within 14 Å from tosyl C1 excluding main-chain atoms beyond 15 Å. This region (724 atoms) was subjected to deterministic molecular dynamics using the Verlet algorithm. The second region comprised all residues that have at least one atom within 16 Å

(91) Decius, J. C. *J. Chem. Phys.* **1948**, *16*, 1025.
(92) Brooks, C. L., III; Brunger, A.; Karplus, M. *Biopolymers* **1985**, *24*, 843.
(93) Brooks, C. L., III; Karplus, M. *J. Chem. Phys.* **1983**, *79*, 6312.

from tosyl C1 excluding all atoms that belong to region 1. This region (292 atoms) was subjected to Langevin dynamics. The remaining atoms (the reservoir region) were deleted.

The heavy atoms in the Langevin region were constrained by harmonic forces to mimic the effect of the omitted reservoir atoms. The force constants were obtained from the crystallographic B factors through

$$k = \frac{4\pi^2 k_B T}{B} \tag{8}$$

Since these were not available for the tosyl-α-chymotrypsin structure, we assumed the value 20 Å^2. The average value of B in crystals of native chymotrypsin is slightly less than this value.[94,95] The force constants were scaled such that atoms closer to the innermost or reaction region are constrained to a smaller extent than those closer to the reservoir region. The friction coefficient for the heavy protein atoms in the Langevin region was taken to be 250 ps^{-1}. This value was obtained from velocity autocorrelation functions in conventional molecular dynamics simulations of proteins[92] and increased somewhat for faster equilibration.

For the simulations in water, an 18-Å sphere of equilibrated TIP3P[96] water centered at tosyl C1 was overlayed on the protein and all water molecules closer than 2.8 Å from any protein heavy atom were deleted. The water was briefly equilibrated and the process repeated several times to ensure that all voids were filled. A total of 223 water molecules (in addition to 15 crystallographically determined water molecules) remained in the system. In addition to the two crystallographic water molecules in the back of the specificity pocket, two water molecules were accommodated in the back of the pocket close to Tyr 228. These water molecules, after equilibration, formed hydrogen bonds to the backbone atoms of Trp 215 and Ser 214, the side chains of Tyr 228 and Ser 190, and each other.

SHAKE[97] was used for all bonds with hydrogen. A cutoff of 8 Å was used for the nonbonded interactions. The electrostatic interactions were shifted to bring them to zero at the cutoff distance, while for the van der Waals interactions a switching function was applied. Each FEP simulation consisted of 4000 steps of equilibration and 16 000 steps of sampling, with the time step set equal to 1.5 fs. The temperature was set at 300 K and controlled to within 5 K during equilibration and to within 2 K during sampling by periodically scaling the velocities. Most simulations were performed on the Cray-YMP at the Pittsburgh Supercomputing Center while a few were performed on a SGI IRIS 4D-220S workstation.

Results

Paths Based on the Energy-Minimized Crystal Structure (EMCS). To determine the transition state for rotation of the aromatic ring in the pocket, the Cerjan–Miller method in the space of the four dihedral angles χ_1–χ_4 was used. A starting conformation was obtained from the EMCS by moving χ_4 while constraining χ_1, χ_2, and χ_3 so that the ring was perpendicular to the pocket. The transition state (TSROT) was determined after about 20 iterations, and its structure is shown in Figure 3. We see that in the TS, the ring is positioned at the rim of the pocket with an orientation perpendicular to it. Apparently, there is too much strain for rotation of the ring entirely within the pocket, and this path provides a lower energy barrier. One of the SD paths initiated from this TS led directly to the minimum-energy conformation in the pocket (POCK) and the other to a conformation identical to POCK but with the ring rotated by 180°. The energy profile and the changes in the dihedral angles χ_1–χ_4 along this path are shown in Figure 4a,b, respectively. Rotation is captured primarily through the change in χ_4 which exhibits an overshoot at one end of the path. Movement of the ring to the rim of the pocket is reflected by the change in χ_2. The dihedral angles χ_1 and χ_3 remain almost constant along the path.

The potential energy barrier for this interconversion is $\sim$10–11 kcal/mol as shown in Figure 4a. The different contributions to the energy barrier are also shown in this figure. Almost half

(94) Tsukada, H.; Blow, D. M. *J. Mol. Biol.* **1985**, *184*, 703.

(95) Blevins, R. A.; Tulinsky, A. *J. Biol. Chem.* **1985**, *260*, 4264.

(96) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926.

(97) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327.
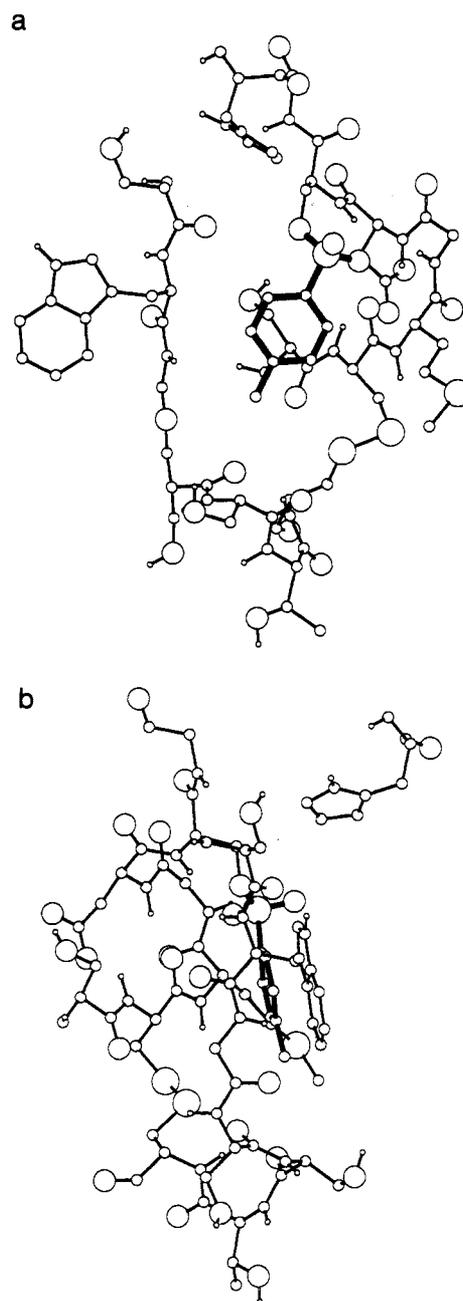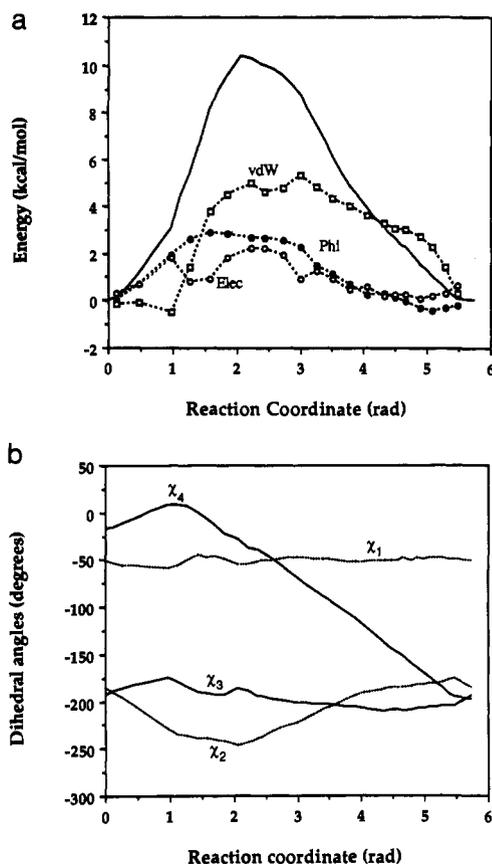
a

b

Figure 3. Transition state for rotation in the pocket, based on the energy-minimized crystal structure: (a) and (b) as in Figure 2.

of this energy is due to loss of van der Waals interactions, and the rest is divided equally between electrostatic and torsional energy. We note again that this result does not include any intrinsic torsional potential for χ_4. One can observe an anomaly in the relative contributions of van der Waals and electrostatic energy in the region of the RC where the overshoot in χ_4 occurs.

To determine the path out of the pocket, a local energy minimum was first located for the ring outside the pocket. This was accomplished visually by interactive graphical manipulation of the tosyl ring and subsequent energy minimization. This local minimum, called LM1, is shown in Figure 5. In this conformation, the aromatic ring points away from the active site and is free to rotate without steric hindrance from the surrounding protein. Transition states between LM1 and POCK were found by applying the "chain" method described above. A TS (TS21) was located which was verified to have one and only one negative eigenvalue on the adiabatic torsional space of χ_1–χ_4. One of the SD paths initiated from this TS led to LM1 and the other to a new local minimum, LM2. Repeating this procedure between LM2 and

Tosyl-α-chymotrypsin

J. Am. Chem. Soc., Vol. 116, No. 4, 1994 1551

Figure 4. Profiles for ring rotation in the pocket, based on the energy-minimized crystal structure. (a) Energy profile: solid line = total energy; open squares = van der Waals interactions; dark circles = torsional energy; and open circles = electrostatic energy. (b) Dihedral angle profile.

POCK, a second TS, TS32, was determined which was found to connect LM2 with yet another local minimum, LM3. In LM3, the ring is actually in the pocket, in an orientation similar to that in POCK but with different values for the torsional angles $\chi_1-\chi_4$. The same procedure was then applied to the interconversion between LM3 and POCK to give the third TS, TSPOCK3. The energy and the dihedral angle profiles for the overall path, POCK to LM1, is shown in Figure 6a,b, respectively. The energy barrier for the LM3–POCK interconversion is the highest along the path ($\sim$10 kcal/mol) and makes this transition the rate-limiting step. The magnitude of this barrier is approximately equal to that for rotation in the pocket as shown in Figure 4a.

Paths Based on the Dynamically Equilibrated Structure (DES). On the basis of the RC determind in the previous section, calculation of the FE profile for ring rotation was attempted. However, the dynamics trajectories exhibited such a large drift from the calculated path that inordinately large "perturbations" in dihedral angles were needed to move from one hyperplane to the next, resulting in poor convergence of the relative free energies. Thus, it became impossible to calculate the free-energy difference between subsequent points on the path.

Examination of the dynamic trajectories revealed a systematic conformational change in the specificity pocket. In all simulations (i.e., all points along the reaction path), one side of the pocket (left side in Figure 2a, residues 214–220, hereafter called the "specificity loop") moved upward relative to the other side and the tosyl ring, and the bottom of the pocket (residues 217–219) closed behind the tosyl ring decreasing the opening of the pocket (Figure 7). This region showed the largest displacement (>2 Å), while the rest of the specificity loop was displaced by 1–1.5 Å. The movement of the specificity loop was accompanied by a similar movement of residues 222–228, which form a β-sheet with the specificity loop. Thus, the observed movement was more or less
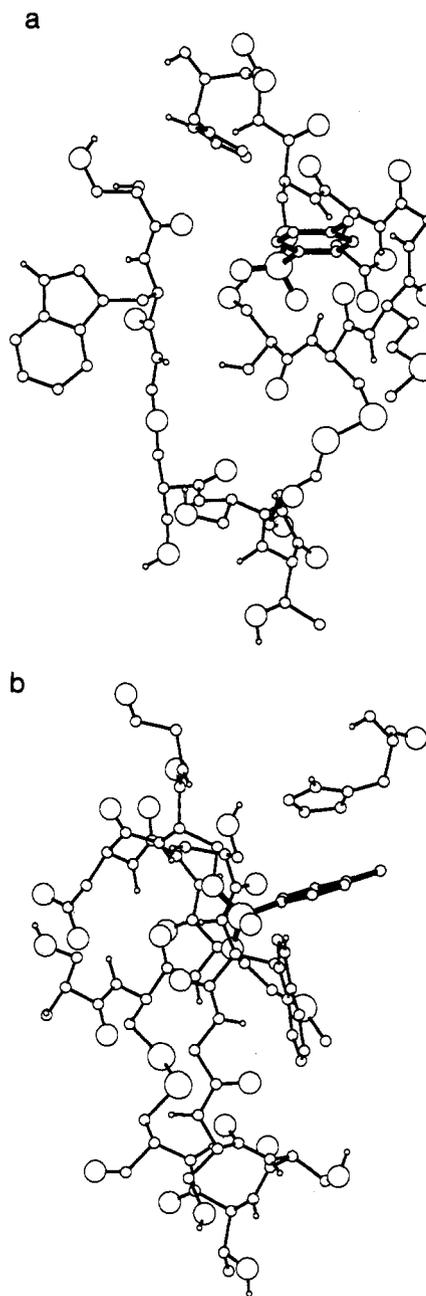

Figure 5. First local energy minimum out of the pocket (LM1): (a) and (b) as in Figure 2.

a rigid-body movement of this β-sheet with respect to the rest of the protein. The reasons for this conformational change are not obvious. It cannot be due to omission of the solvent, since exactly the same conformational change was found to occur in simulations with explicit water molecules. Furthermore, the assignment of initial velocities at the beginning of the simulation was found to play no role. In addition, we observed this conformational change in simulations of the whole protein, and therefore, it cannot be an artifact of the stochastic boundary setup. In all cases, the conformational change was rapid, occurring within the equilibration period, so it would not be difficult to test whether it will happen in simulations of the crystallographic dimer in the region of the active site or with a different parameter set. Preliminary simulations with the new all-atom CHARMM22 parameters exhibited smaller deviations from the active site crystal structure.

In crystallographic studies of the various native enzyme and enzyme–inhibitor structures, the specificity pocket is reproduced quite well, although small differences have been observed in the conformation of Ser 218 between the dimeric α-chymotrypsin

1552 *J. Am. Chem. Soc., Vol. 116, No. 4, 1994*
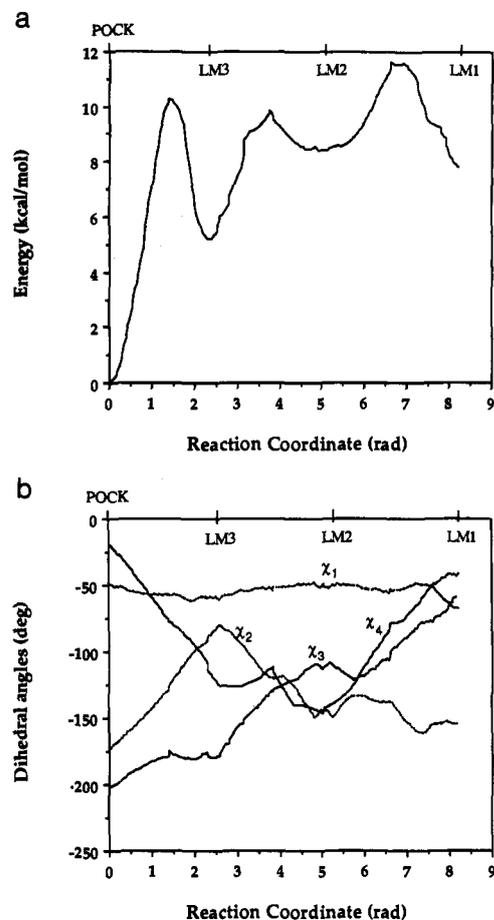
Lazaridis and Paulaitis

a

b

Figure 6. Profiles for movement of the ring out of the pocket, based on the energy-minimized crystal structure: (a) energy profile and (b) dihedral angle profile.

and the monomeric γ-chymotrypsin. These deviations are probably due to interaction of the two protein molecules in the asymmetric unit of the α-chymotrypsin crystals[98] and, more importantly, are much smaller than those observed in the dynamic trajectories. However, a high-resolution (1.5 Å) crystallographic study of a chymotrypsin–inhibitor complex showed significant differences in the specificity pocket for the free and bound enzyme.[99] These experimental observations suggest that this part of the active site is quite flexible, in agreement with our simulations, although the actual changes are different. The crystallographic *B* factors associated with the atoms in the specificity loop are also relatively high,[94,95] indicating high mobility. The specificity loop was also found to be mobile and to deviate considerably from the crystal structure in a simulation of SGPA,[62] which possesses a three-dimensional structure similar to that of chymotrypsin. In terms of RMS differences, the observed deviations from the crystal structure are relatively moderate. After we energy minimized the final structure from the dynamics trajectory, the RMS deviation of all atoms between this and the EMCS was 1.35 Å. This is comparable to what was observed in other simulation studies,[62,100,101] but it is large enough to change the RC for rotation expressed in terms of dihedral angles.

Therefore, it became necessary to determine new reaction paths based on the dynamically equilibrated protein structure. In doing

(98) Cohen, G. H.; Silverton, E. W.; Davies, D. R. *J. Mol. Biol.* **1981**, *148*, 449.

(99) Ringe, D.; Seaton, B. A.; Gelb, M. H.; Abeles, R. H. *Biochemistry* **1985**, *24*, 64.

(100) Berendsen, H. J. C.; van Gunsteren, W. F.; Zwinderman, H. R. J.; Geurtsen, R. G. *Ann. N. Y. Acad. Sci.* **1986**, *482*, 269.

(101) Ichiye, T.; Olafson, B. D.; Swaminathan, S.; Karplus, M. *Biopolymers* **1986**, *25*, 1909.
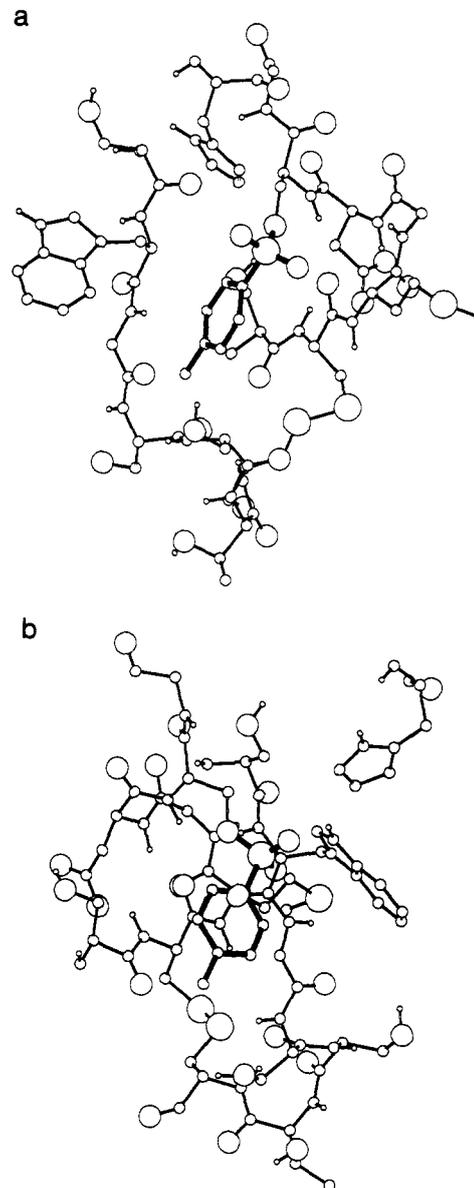
a

b

Figure 7. Conformation of minimum energy in the active site of tosyl-α-chymotrypsin (dynamically equilibrated structure): (a) and (b) as in Figure 2.

so, we observed significant differences between the new paths and those previously obtained on the basis of the energy-minimized crystal structure, as shown by the energy and dihedral angle profiles in Figure 8a,b, respectively. Now, the ring rotates fully within the pocket. The only angle that changes substantially is χ_4. In addition, the energy barrier is higher (~15 kcal/mol), apparently due to a lower energy in the ground state as the pocket relaxes around the inhibitor. The path calculated without any metric correction (eq 3) was also qualitatively similar to that of Figure 8b, suggesting that this correction can be omitted without significant error.

In attempting to determine the path out of the pocket, it was found that this transition is now severely hindered by the lower part of the specificity pocket which has moved closer to the ring. In particular, atoms 217 O and, primarily, 220 SG of the disulfide bridge impose steric constraints to the motion of the ring in and out of the pocket by nonbonded repulsions with the *p*-methyl group of the ring. This results in a breakdown of our adiabatic approximation in four-angle space, since the same set of values for the dihedral angles will correspond to different conformations depending on which side of the barrier one initiates the RC calculations. That is, because of the strength of the nonbonded
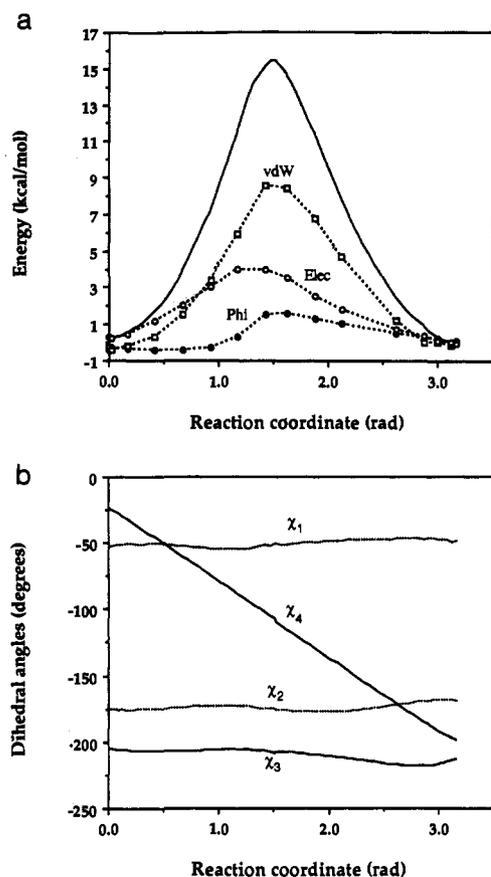
Tosyl-α-chymotrypsin

J. Am. Chem. Soc., Vol. 116, No. 4, 1994 1553

a

b

Figure 8. Profiles for rotation of the ring within the pocket, based on the dynamically equilibrated structure: (a) energy profile and (b) dihedral angle profile.

Table 1. Values of the Four Dihedral Angles and the Energy for the Local Minima and the Transition States Determined in This Work for Tosyl-α-chymotrypsin

	χ_1[b]	χ_2	χ_3	χ_4	energy (kcal/mol)[a]
X-ray[c]	−64	106	−161	26	
POCK	−51	174	166	−18	0.0
LM1	−44	−167	−48	−71	7.2
LM2	−51	−146	−110	−146	8.3
LM3	−59	−89	−177	−116	5.1
TSROT	−54	113	173	−28	10.4
TSPOCK3	−58	−130	−179	−76	10.1
TS32	−52	−121	−131	−116	9.4
TS21	−56	−138	−107	−83	11.3
POCKD	−51	−172	152	−20	0.0
LMD1	−44	−95	−96	−25	9.7
LMD2	−51	−104	−152	−116	9.1
LMD2′	−49	−71	−122	−142	4.0
LMD3	−65	−92	159	−99	8.3
TSROTD	−51	−176	152	−110	15.4
TSDPOCK3	−61	−108	152	−76	10.4
TSDPOCK2	−60	−118	173	−81	23.9
TSD32	−61	−84	168	−120	17.1
TSD2′2	−48	−88	−146	−131	9.6
TSD21	−46	−86	−108	−56	11.0
POCKWA	−56	172	174	−32	0.0
POCKWB	−58	−112	157	−105	0.6
TSB	−61	−111	157	174	13.8
TSBTOA	−59	−132	157	−178	14.1

[a] The energies are relative to the lowest local minimum for a given set of structures. [b] The dihedral angles are defined as follows: χ_1, 195 N–195 CA–195 CB–195 OG; χ_2, 195 CA–195 CB–195 OG–TOS S; χ_3, 195 CB–195 OG–TOS S–TOS C1; and χ_4, 195 OG–TOS S–TOS C1–TOS C2. [c] Birktoft and Blow.[31]
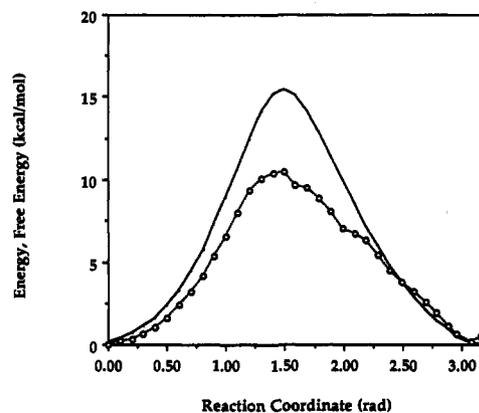

Figure 9. Free-energy profile (open circles) for rotation of the tosyl ring within the specificity pocket of α-chymotrypsin in the absence of explicit solvent. The solid line is the adiabatic energy profile.

interactions between the ring and the pocket and the large distance between the *p*-methyl group and the anchor of the tosyl group (Ser 195 CA), slight changes in the backbone or in the bond angles around Ser 195 can move the ring to either side of the barrier posed by the specificity loop for a given set of χ_1–χ_4 values.

This problem was solved by including one more explicit degree of freedom in the RC calculations: the virtual dihedral angle 220 SG–215 CA–195 CA–TOS CH3. The advantage of choosing this particular virtual angle is that it can be changed by rigid-body displacements of the tosyl ring, which is useful in calculating the partial derivatives $\partial E/\partial\chi_i$. If rigid-body displacements were not possible, these derivatives would have to be calculated by energy minimizations, thus increasing the computational cost of the calculations by about 10-fold.

Two paths out of the pocket were ultimately found, a direct path (POCKD–LMD2–LMD2′–LMD1) with an energy barrier of 23.9 kcal/mol and an indirect path (POCKD–LMD3–LMD2–LMD2′–LMD1) with a somewhat lower energy barrier of 17.1 kcal/mol. The χ_1–χ_4 dihedral angles for these local minima and their connecting transition states are given in Table 1. In this table, only relative energies are reported between structures that have resulted from one another by energy minimization; i.e., structures that belong to the same adiabatic energy surface. Both paths give a large energy barrier for the movement of the ring out of the pocket relative to rotation in the pocket. We believe this is due to the distortion of the pocket that occurred after the dynamics run, since the movement of the specificity loop upward, closer to the ring, is primarily responsible for these large barriers. Other paths were also explored; in particular, a crankshaft motion of χ_1 and χ_3 to move the ring out of the pocket without bumping into the bottom of the pocket. This path also failed to provide a lower energy barrier.

Calculation of Free-Energy Profiles. A total of 16 constrained molecular dynamics simulations were performed along the RC describing ring rotation in the pocket for the DES (Figure 8). The fluctuations of each of the four angles during these simulations were approximately ±5° to ±30°. The calculated FE profile is shown in Figure 9 with the energy profile from Figure 8. The FE barrier is lower than the energy barrier by about 5 kcal/mol. The statistical uncertainty in the free energy has been estimated from the standard deviation of values calculated from different bins of the data. The error bars thus calculated are ±0.27 kcal/mol at the top of the barrier and ±0.41 kcal/mol at the ends of the profile. The sample-size hysteresis error was also estimated[102] and found to be less than or equal to the random error at all points. If we assume that the adiabatic energy barrier for this motion (which corresponds to the true energy barrier at 0 K) is equal to the energy barrier at 300 K, the activation entropy, calculated from the difference, is about 16 cal/(mol K). The activation entropy was also calculated by the finite-difference

(102) Wood, R. H.; Mühlbauer, W. C. F.; Thompson, P. T. *J. Phys. Chem.* **1991**, *95*, 6670.

1554 *J. Am. Chem. Soc., Vol. 116, No. 4, 1994*
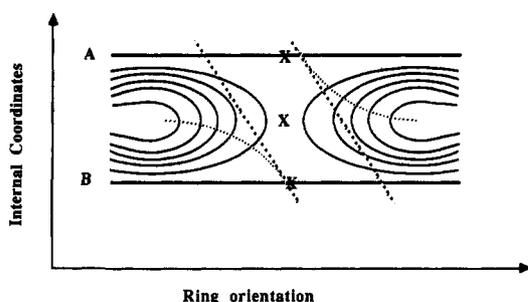
Lazaridis and Paulaitis

Figure 10. Schematic diagram for the two paths that effect rotation of the tosyl aromatic ring (straight horizontal lines) on a contour plot of the potential energy surface. The dashed lines are the linear orthogonal constraints implemented in the free-energy calculations. The dotted lines are the ideal curvilinear constraints. X denotes transition states.

temperature derivative method,[103] but this method gave a value with a very large statistical uncertainty in the present case.

The FE calculation was repeated in the presence of explicit solvent to examine possible effects of hydration on the mobility of the inhibitor. However, it was found that the ring jumped to another path for the rotation of the LMD3 conformation. In this conformation, as in LM3, the ring is in the pocket, but the torsional angles $\chi_1-\chi_4$ have values different from those in POCKD. This was verified by calculating the new path using the chain method. This path (referred to as path B) is very much like that shown in Figure 8 (referred to as path A), only the lines for χ_2 and χ_4 are translated, the first upward by about 60° and the second downward by about 80°. An attempt to perform a FE calculation on this new path led to exactly the opposite occurrence near the top of the barrier: a jump from path B to path A. The situation can be visualized on the schematic diagram in Figure 10. There exist two paths (A and B) that effect a flip of the aromatic ring. The barrier between the paths is high near the ground states but becomes progressively lower and approaches k_BT close to the TS. Therefore, near the TS, the ring can jump easily from one path to the other. It is likely that our choice of orthogonal hyperplane (denoted by the dashed lines in Figure 10) for the constraints is the source of this problem. In principle, the constraint should be along a curved, geodesic line (dotted line in Figure 10) orthogonal to the contours of constant energy. With that constraint, jumps between the two paths would not be possible. However, it would be difficult to implement a complex, curvilinear constraint like this in our current methods.

The jump between the two paths does not occur in the absence of water solvent. Apparently, conformation LMD3 is stabilized by water relative to conformation POCK, and also, the barrier for the interconversion between LMD3 and POCKD is lowered. From graphics observations, it is apparent that this might be due to enhanced interactions between water and the sulfonyl oxygens in conformation LM3. In POCKD, the two oxygens point sideways, toward the two sides of the pocket, whereas in conformation LMD3, one oxygen points directly into the solvent and can form hydrogen bonds with water molecules. The additional two water molecules in the back of the pocket that were inserted during the set up of the simulation with water do not seem to be involved in this stabilization of LMD3, since the same jump occurs when these are deleted.

Figure 10 suggests that there might be another TS linking the paths A and B, and this was confirmed by applying the chain method between two points on the A and B paths. (The chain method can be applied even when the two end points are not local minima.) The path initiated from this TS (TSBTOA in Table 1) was found to link the "reactant" state of path A with the "product" (ring-flipped) state of path B.

To obtain the FE profile along one of the paths for rotation in the presence of the solvent, we implemented a rather *ad hoc*

procedure. Since the movement of χ_2 is responsible for the jump between paths, we imposed a weak harmonic force constraining χ_2 near the values appropriate for path B. The FE barrier obtained by this approximation is somewhat lower than the corresponding barrier in vacuum, but the curves are similar. The error bars are approximately ±0.3 kcal/mol at the top of the barrier and ±0.6 kcal/mol at the ends. The fact that water does not have much effect on either the conformation of the pocket or the value and the nature of the free-energy barrier for ring rotation in the pocket should perhaps be expected, since the specificity pocket is primarily hydrophobic in character and does not interact strongly with water. Different observations would most likely be made in a situation where intraprotein hydrogen bonds or salt bridges are present.[104]

On the basis of calculated free-energy barriers of 9–10 kcal/mol for rotation within the pocket, with or without explicit waters, and the elementary TS theory expression

$$k = \kappa \frac{k_B T}{h} \exp\left(-\frac{\Delta G}{RT}\right),$$

an estimate of the rate of ring rotation in the pocket would be (3.5×10^5)–(1.9×10^6) sec^{-1}, which corresponds to a correlation time of 500 ns–2.8 μs. A typical value of 0.1–0.5 for the transmission coefficient κ would further reduce the estimated rate of rotation, since we have set κ equal to 1.0 in these calculations. The experimental correlation time for rotation of the tosyl ring is 0.14–1.1 ns.[37] There is, however, evidence that this rapid rotation may be due to the fact that the tosyl ring also resides, at least part of the time, outside the specificity pocket.[105] The much longer correlation times we obtained for rotation in the pocket support this notion. Much longer correlation times were also measured for a fluorinated aromatic inhibitor that more closely resembles an actual substrate for the enzyme (>10 μs) and for a fluorinated biphenyl ring (260 ns) where it is believed that one phenyl ring resides in the pocket and is relatively immobile, while the other points toward the surface of the enzyme and rotates rapidly.[106,107] In the latter case, the measured value of the correlation time would be an average for the two ring conformations and thus underestimates the correlation time for rotation in the pocket alone. Collectively, these NMR results indicate that the present computational result for the correlation time ($\geq$500 ns) should be reasonable.

The positive entropy contribution to the barrier, suggested by the qualitative difference between the adiabatic energy and the free energy in Figure 9, can be conceptualized in terms of the "width of the reaction valley" along the RC. In the ground state, the ring is tightly held by the pocket and this valley is narrow. As the ring rotates, the pocket opens up to accommodate it, and this allows the ring more freedom along the direction orthogonal to the path; i.e., the reaction valley becomes wider near the top of the barrier. The experimental determination of activation enthalpies and entropies from the temperature dependence of measured rates, such as NMR correlation times, is subject to considerable uncertainty, and the activation parameters thus determined may not reflect intrinsic properties for the process under study. Northrup et al.[69] suggest, for example, that calculated activation parameters may include changes in the viscosity of the solvent or changes in the protein conformation with temperature. In addition, interpretation of NMR results is usually based on adopting simple dynamic models for the process studied and is contingent upon the validity of these models. Nevertheless, a survey of available experimental results is still profitable. Experimental activation entropies for the dynamics

(103) Brooks, C. L., III. *J. Phys. Chem.* **1986**, *90*, 6680.

(104) MacKerell, A. D., Jr.; Nilsson, L.; Rigler, R.; Saenger, W. *Biochemistry* **1988**, *27*, 4547.

(105) O'Connell, T. M.; Gerig, J. T.; Williams, P. G. *J. Am. Chem. Soc.*, submitted for publication.

(106) Gerig, J. T.; Hammond, S. J. *J. Am. Chem. Soc.* **1984**, *106*, 8244.

(107) Kairi, M.; Gerig, J. T. *Magn. Res. Chem.* **1990**, *28*, 47.

Table 2. Average Values and Fluctuations of Key Interatomic Distances and Residues in the Active Site of Tosyl-α-chymotrypsin in the Presence and Absence of Water

	vacuum		water	
	average	fluctuations	average	fluctuations
His 57 Ne2–Ser 195 Og	5.03	0.36	4.55	0.27
His 57 Nd1–Asp 102 Od1	3.05	0.16	3.14	0.19
His 57 Nd1–Asp 102 Od2	3.14	0.23	4.27	0.27
Asp 194 Od1–Ile 16 N	2.71	0.11	2.74	0.11
Asp 194 Od2–Ile 16 N	2.82	0.11	3.04	0.24
His 57[a]		0.193		0.193
Ser 195		0.200		0.209
Asp 102		0.145		0.150
Met 192		0.358		0.244
specificity loop[b]		0.214		0.163

[a] Only heavy atoms are included in the averages for the residues His 57, Ser 195, Asp 102, and Met 192. [b] Average for the backbone atoms of Ser 214 through Thr 219.

of the tosyl ring have not been reported. For the rotation of a tyrosine ring in BPTI, large positive activation entropies were measured experimentally but were not reproduced by umbrella-sampling calculations.[69–71] Kairi and Gerig[107] also report positive activation entropies for the motion of fluorinated biphenyl groups in the active site of α-chymotrypsin. Interestingly, a negative activation entropy is reported for the same process in the denatured enzyme.[107]

Free-energy calculations on the paths out of the pocket were not attempted, in part because solvent effects must be included for these paths. As will be discussed below, these calculations are significantly more complicated, particularly in the presence of explicit solvent. The NMR results do not give directly the rate of ring motion in and out of the pocket but do establish that this interconversion is fast compared to NMR time scales. This gives an estimated upper bound for the FE barrier of 16 kcal/mol.[108] Thus, the energy barriers calculated here are consistent with this upper bound if we assume a modest positive value for the activation entropy. We believe, however, that these energy barriers for movement out of the pocket may be somewhat high due to the conformational change that constricts the opening of the pocket in the molecular dynamics simulations.

Overall Dynamics in the Active Site. Table 2 contains the results of a brief analysis of the dynamics trajectories in vacuum and in explicit water. The values listed are averages from the constrained simulations used to obtained the FE results reported in the previous section. Several observations are worth noting. The average distance between His 57 N^{e2} and Ser 195 O$^\gamma$ is quite large. The shorter value in water is due to the fact that path B is followed in these simulations, which brings O$^\gamma$ closer to the histidine. His 57 is quite mobile, but the hydrogen bond with Asp 102 is maintained. Also, the ion pair Asp 194–Ile 16 is maintained throughout the simulations. In the presence of water, the distances between Asp 102 and His 57 and between Asp 194 and Ile 16 are somewhat longer. Since no direct interaction occurs between water and these pairs, this could be due to the dielectric effect of water. Met 192 exhibits the highest fluctuations, due to its long side chain which can move around freely. The fluctuations of the backbone atoms of the specificity loop are seen to be lower in water.

Discussion

The results presented herein show that multidimensional reaction paths for conformational rearrangements in the active sites of enzymes can be determined on an adiabatic potential energy hypersurface in reduced internal coordinate space. The adiabatic approximation, by which certain degrees of freedom (DOF) are considered explicitly and the remaining DOF are allowed to relax, has been found to work well even for a complex

molecule the size of a protein. More importantly, its inevitable failures are obvious to the user and can be remedied by increasing the number of explicit DOF, appropriately taking into account the interactions that cause the breakdown of the approximation. RC determination methods that require only first derivatives of the potential energy can handle quite a large number of explicit DOF. Methods that use second derivatives, such as the Cerjan–Miller method, seem to be rather expensive, and their computational requirements will scale much faster with the number of degrees of freedom than first derivative methods. In any case, close interaction between the program and the user and the aid of computer graphics will be necessary, until "smarter" algorithms can be devised.

Some obvious extensions of the present methodology would be useful. One is the accommodation of Cartesian coordinates as well as internal coordinates in the set of "essential" DOF. That would enable one, for example, to study pathways for aromatic ring insertion into the specificity pocket *prior* to the formation of the covalent bond between the inhibitor and the active site serine, which more likely corresponds to the real sequence of events in enzyme–ligand recognition. In such a calculation, the three translational and three rotational degrees of freedom of the substrate should be treated explicitly. There is, of course, the question of what the metric should be in such a space, but this seems to be a matter of secondary importance, judging from these results and also from our previous work (Figure 1).[85] These extended techniques could also be used to study the putative coupling of conformational motions to the chemical events of enzymatic catalysis using combined quantum mechanics–molecular mechanics Hamiltonians.[109] Reaction paths can potentially be more informative than MD simulations, since they reveal the intrinsic dynamics of a process without being obscured by thermal noise.

FEP appears to be an accurate method for obtaining free-energy profiles for conformational transitions, as long as the system follows a single, well-defined RC. The quality of the calculations can be seen in Figure 9 in the small overall hysteresis; i.e., in the fact that the curves return to zero after a full ring flip, as they should. The statistical noise, as measured by the method of subaverages, appears to be low. Compared to umbrella sampling, the present implementation of the FEP method has the advantage that any difficulties or failures become evident to the researcher through large deviations from the RC and very large estimated sample-size hysteresis errors.[102] Therefore, we believe that when FEP can be applied, it will give accurate results. In contrast, it appears to us that it is more difficult to tell when something has gone awry in umbrella-sampling calculations. The latter method also suffers from uncertainties in matching the different windows. However, FEP is expensive and requires performing simulations along the entire path.

One goal of this work was to address the question posed in the Introduction: Does the hydrophobic pocket relax in solution thus allowing solvent to enter and the ring to rotate rapidly, or does the ring equilibrate between its conformation in the pocket and one or more conformations out of the pocket? Our results support the latter in the sense that the calculated rotational correlation time for the ring in the pocket is much slower than those measured for this system. Our results also suggest that the ring will rotate in the pocket, as well as in solution, since the time scales for movement out of the pocket are much slower than that for rotation in the pocket. This implies that measured NMR correlation times are an average of the correlation times for rotation in the pocket and in solution weighted by the relative populations of these two conformational states. Although questions also remain about the authenticity of the conformational change observed in the molecular dynamics simulations, the facility by which this change

(108) Gerig, J. T. Private communication.

(109) Alagona, G.; Desmeules, P.; Ghio, C.; Kollman, P. A. *J. Am. Chem. Soc.* **1984**, *106*, 3623.

occurs leads one to suggest that it may actually occur as a fluctuation, perhaps on a longer time scale, with an evident functional role. A slight movement of the left side of the chain upward locks the substrate in a position favorable to catalysis, and a slight movement downward facilitates its release.

There is another factor that should necessarily enter into this discussion when we consider the movement of the ring out of the pocket, namely solvation and the hydrophobic interaction. When the ring moves out of the pocket, some water molecules will take its place in the specificity pocket, as was observed in crystal structures of the native enzyme.[94,95] The presence of solvent around the enzyme can have two opposing effects. One is to provide van der Waals interactions to the ring, which tend to stabilize conformations out of the pocket. The second would be the effectively unfavorable (hydrophobic) interaction of the ring with water, which would tend to stabilize the conformation in the pocket. It is this interaction that is thought to drive the ring into the pocket in the first place. Another question to be answered is why two water molecules are not displaced upon substrate binding and whether they play a biological role.[95]

One might hope that a FEP simulation of the ring moving in and out of the pocket in the presence of explicit water molecules would account for all these interactions and would give, in principle, the correct relative free energy. However, there are practical reasons for which this approach is not likely to work. Performing "perturbations" of the ring toward a pocket full of water molecules does not appear to be a simulation that will produce convergent, meaningful results. In reality, the exact sequence of events that leads to the displacement of water molecules and insertion of the aromatic ring in the pocket is of fundamental interest and may have a significant effect on the kinetics of binding. It has been suggested that the displacement of a water molecule is the rate-limiting step for inhibitor binding.[110,111] This problem could be the subject of more advanced computations, in which the coordinates of the ring and the displaced water molecules are considered explicitly.

(110) Holden, H. M.; Tronrud, D. E.; Monzingo, A. F.; Weaver, L. H.; Matthews, B. W. *Biochemistry* **1987**, *26*, 8542.
(111) Bartlett, P. A.; Marlowe, C. K. *Biochemistry* **1987**, *26*, 8553.