J|A|C|S
ARTICLES
Published on Web 11/15/2002

Ab Initio QM/MM Dynamics Simulation of the Tetrahedral Intermediate of Serine Proteases: Insights into the Active Site Hydrogen-Bonding Network

Maya Topf,* Peter Várnai,† and W. Graham Richards*

Contribution from the Physical and Theoretical Chemistry Laboratory, University of Oxford, South Parks Road, OX1 3QZ, United Kingdom and Laboratoire de Biochimie Theorique, Institut de Biologie Physico-Chimique, 13 rue Pierre et Marie Curie, Paris 75005, France

Received March 18, 2002

Abstract: Ab initio QM/MM dynamics simulation is employed to examine the stability of the tetrahedral intermediate during the deacylation step in elastase-catalyzed hydrolysis of a simple peptide. An extended quantum region includes the catalytic triad, the tetrahedral structure, and the oxyanion hole. The calculations indicate that the tetrahedral intermediate of serine proteases is a stable species on the picosecond time scale. On the basis of geometrical and dynamical properties, and in agreement with many experimental and theoretical studies, it is suggested that the crucial hydrogen bonds involved in stabilizing this intermediate are between Asp-102 and His-57 and between the charged oxygen of the intermediate and the backbone N−H group of Gly-193 in the oxyanion hole. The mobility of the imidazolium ring between O_w and O^γ, two of the oxygens of the tetrahedral structure, shows how the intermediate could proceed toward the product state without a "ring-flip mechanism", proposed earlier on the basis of NMR data. In addition to the proposed $C^{\epsilon 1}$−H···O hydrogen bond between the imidazolium ring and the backbone carbonyl of Ser-214, we observe an alternative $C^{\epsilon 1}$−H···O hydrogen bond with the backbone carbonyl of Thr-213, that can stabilize the intermediate during the imidazolium movement. Proton hopping occurs between Asp-102 and His-57 during the simulation. The proton is, however, largely localized on the nitrogen, and hence it does not participate in a low-barrier hydrogen bond. The study also suggests factors that may be implicated in product release: breaking the hydrogen bond of the charged oxygen with the backbone of Ser-195 in the oxyanion hole and a loop opening between residues 216−225 that enables the breaking of a hydrogen bond in subsite S_3.

Introduction

Hydrogen-bonding networks often play a crucial role in the mechanism of enzymes. They can stabilize the conformation of the active sites of enzymes with the substrate oriented optimally for catalysis. They can also contribute to transition state stabilization by lowering the free energy of activation by several kcal mol^{-1} and by stabilizing intermediates on the potential energy surface.[1−3] An excellent example for a hydrogen-bonding network is the one formed with the catalytic triad of the serine protease (SP) family consisting of Asp-102, His-57, and Ser-195.[2,4−7]

The accumulated results of both experimental and theoretical work almost all support a common general mechanism for SP mediated amide hydrolysis. According to this mechanism, the reaction proceeds via an acylation step, during which a nucleophilic attack of the hydroxyl group of Ser-195 on the carbonyl carbon of the substrate leads to the formation of an acyl-enzyme (ester) intermediate (EA). The latter is hydrolyzed with the assistance of a water molecule during the deacylation step (Scheme 1), to give a noncovalent enzyme−product (EP) complex. Both the acylation and the deacylation steps of the reaction mechanism are believed to proceed via a negatively charged tetrahedral intermediate (TI1 and TI2), although there is no direct evidence for that.[2] The deprotonated side chain of His-57 (imidazole) acts as a general base, catalyzing the deprotonation of Ser-195 in the acylation step and of the hydrolytic water in the deacylation step. The protonated side chain (imidazolium) acts as a general acid, donating a proton to the leaving group in the acylation step and to Ser-195 in the deacylation step.

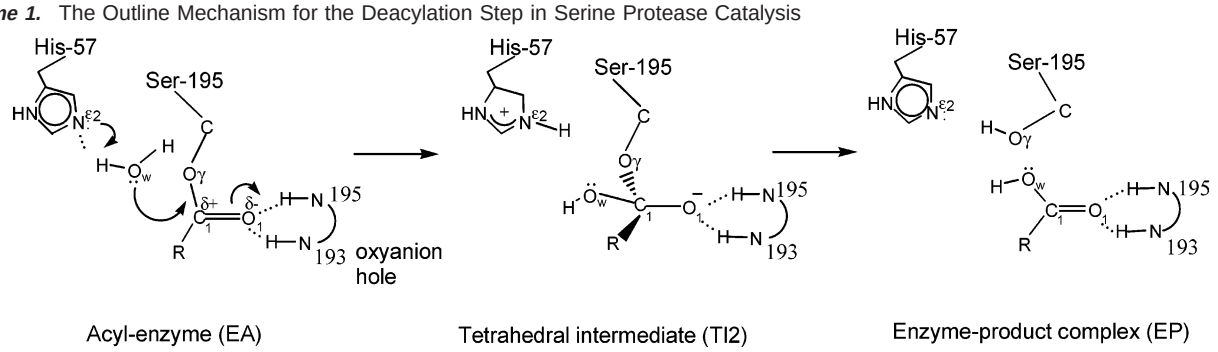
The two high-energy tetrahedral intermediates do not accumulate during peptide hydrolysis,[8,9] and, therefore, the evidence for their existence is indirect and is mainly based on analogy with the proposed mechanism of amide bond hydrolysis in small molecules[1] and from the use of "transition-state analogue" inhibitors, based upon Pauling's theory of enzyme catalysis. This included a crystal structure of bovine trypsin

* To whom correspondence should be addressed. Tel: +44-1865-275-908. Fax: +44-1865-275-905. E-mail: graham.richards@chem.ox.ac.uk.
† Institut de Biologie Physico-Chimique.

(1) Jencks, W. P. *Catalysis in Chemistry and Enzymology*; Dover: New York, 1987.
(2) Fersht, A. *Structure and Mechanism in Protein Science*; Freeman: New York, 1998.
(3) Warshel, A. *J. Biol. Chem.* **1998**, *273*, 27035−27038.
(4) Blow, D. M.; Birktoft, J. J.; Hartley, B. S. *Nature* **1969**, *221*, 337−340.
(5) Blow, D. M. *Acc. Chem. Res.* **1976**, *9*, 145−152.
(6) Kraut, J. *Annu. Rev. Biochem.* **1977**, *46*, 331−358.
(7) Goldblum, A. In *Computational Approaches to Biochemical Reactivity*; A, W., Ed.; Kluwer: Amsterdam, 1997; pp 295−340.
(8) Fastrez, J.; Fersht, A. R. *Biochemistry* **1973**, *12*, 1067−1074.
(9) Fersht, A. R.; Renard, M. *Biochemistry* **1974**, *13*, 1416−1426.

10.1021/ja026219q CCC: $22.00

Scheme 1. The Outline Mechanism for the Deacylation Step in Serine Protease Catalysis

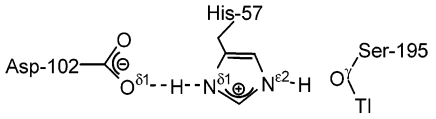


complexed with bovine pancreatic trypsin inhibitor,[10] and a crystal structure of monoisopropylphosphoryl (MIP)-trypsin with one of the unesterified oxygens hydrogen bonded to $N^{\epsilon 2}$ of His-57 and the other hydrogen bonded in the oxyanion hole,[11] followed by its high-resolution neutron structure.[12] Among other crystal structures were *Streptomyces griseus* A protease with Ac-Pro-Ala-Pro-phenylalaninal and with chymostatin, in which an aldehyde group replaces the peptide or ester moiety,[13,14] and a high-resolution structure of trypsin complexed with synthetic inhibitor *p*-midinophenylpyruvic acid, in which the geometry of the carbonyl carbon was between trigonal and tetrahedral.[15,16]

Recently, combined time-resolved/pH jump studies have been carried out on the PPE:BCM7 acyl-enzyme complex.[17] In this work, crystals of the PPE:BCM7 complex at pH 5 were exposed to short pH jumps to initiate reaction and then quenched by freezing with liquid nitrogen. This work led to the observation of electron density changes consistent with the formation of a tetrahedral intermediate (TI2) during hydrolysis of the ester, although this species only accumulated under specific conditions in a crystalline state. The geometry of the tetrahedral intermediate captured in the "pH 9 one minute structure" was similar to those in crystal structures of inhibitors such as trifluoromethyl ketones.[18] However, examination of this structure using molecular dynamics (MD) suggested that the experimental structure may not precisely represent an optimal arrangement for catalysis in solution.[19]

It is widely accepted that two elements are central to the rate acceleration in SP catalysis: (a) the stabilization of the high-energy tetrahedral system by the two hydrogen bonds it forms with the backbone N−H groups of Ser-195 and Gly-193 (in mammalian isoenzymes, or with the N−H groups of Ser-221 and the side chain of Asn-155 in bacterial isoenzymes) in the so-called oxyanion hole, and (b) the electrostatic stabilization provided by Asp-102 to the protonated His-57.[2,6,20,21] Recently, however, two new hypotheses proposed other processes to explain the SP catalytic machinery. The first hypothesis, which

Scheme 2. The Proposed LBHB between His-57 and Asp-102 Upon Protonation of the Imidazole Base in the Tetrahedral Intermediate

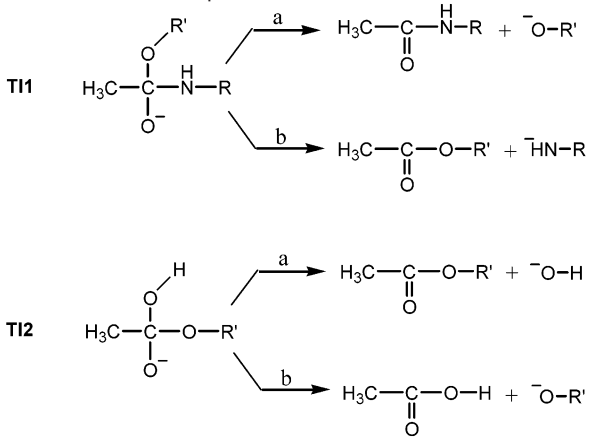


is based on ^{1}H NMR work,[22−25] claims that SP catalysis is partly explained by the transformation of the hydrogen bond between Asp-102 and His-57 into a "low-barrier" hydrogen bond (LBHB) upon protonation of the imidazole base in the transition state (Scheme 2). LBHBs are characterized by short distances between the heteroatoms involved (<2.65 Å), large low-field NMR chemical shifts exhibited by the shared proton (17−22 ppm), low fractionation factors, and positive values for the deuterium isotope effect on the NMR chemical shift of the LBHB. They are thought by some to supply 10−20 kcal mol^{-1} to the transition state by "resonance stabilization", whereas the energy supplied by a weak or conventional hydrogen bond is ∼2−12 kcal mol^{-1}.[26]

An unusual low-field signal was seen for the proton between His-57 and Asp-102 at low pH, in transition state analogues of chymotrypsin, trypsin, subtilisin, and α-lytic protease.[27−31] However, defining those hydrogen bonds as LBHBs is a question of interpretation.[32] According to the LBHB hypothesis, these hydrogen bonds require the absence of aqueous environment and a very small difference between the pK_a values of the heteroatoms involved (close to zero). This is already problematic, because NMR data clearly show that the proton between His-57 and Asp-102 is at least 85% localized on $N^{\delta 1}$ of His-57 and that the proton is not in a hydrophobic environment either.[32,33] Theoretical studies have also argued against the LBHB hypothesis, indicating that the new assumption, of a larger covalent character for hydrogen bonds between the enzyme and charged transition states of substrates than in

(10) Huber, R.; Kukla, D.; Bode, W.; Schwager, P.; Bartels, K.; Deisenhofer, J.; Steigemann, W. *J. Mol. Biol.* **1974**, *89*, 73−101.
(11) Chambers, J. L.; Straud, R. M. *Acta Crystallogr.* **1977**, *B33*, 1824−1837.
(12) Kossiakoff, A. A.; Spencer, S. A. *Biochemistry* **1981**, *20*, 6462−6474.
(13) Brayer, G. D.; Delbaere, L. T.; James, M. N.; Bauer, C. A.; Thompson, R. C. *Proc. Natl. Acad. Sci. U.S.A.* **1979**, *76*, 96−100.
(14) Delbaere, L. T.; Brayer, G. D. *J. Mol. Biol.* **1980**, *139*, 45−51.
(15) Marquart, M.; Walter, J.; Deisenhofer, J.; Bode, W.; Huber, R. *Acta Crystallogr.* **1983**, *B39*.
(16) Walter, J.; Bode, W. *Hoppe-Seyler's Z. Physiol. Chem.* **1983**, *364*, 949−959.
(17) Wilmouth, R. C.; Edman, K.; Neutze, R.; Wright, P. A.; Clifton, I. J.; Schneider, T. R.; Schofield, C. J.; Hajdu, J. *Nat. Struct. Biol.* **2001**, *8*, 689−694.
(18) Imperiali, B.; Abeles, R. H. *Biochemistry* **1986**, *25*, 3760−3767.
(19) Topf, M.; Várnai, P.; Schofield, C. J.; Richards, W. G. *Proteins* **2002**, *47*, 357−369.
(20) Bryan, P.; Pantoliano, W.; Quill, S. G.; Hsioa, H.-Y.; Poulos, T. *Proc. Natl. Acad. Sci. U.S.A.* **1986**, *83*, 3742−3745.
(21) Warshel, A.; Náray-Szabó, G.; Sussman, F.; Hwang, J.-K. *Biochemistry* **1989**, *28*, 3629−3637.
(22) Gerlt, J. A.; Gassman, P. G. *J. Am. Chem. Soc.* **1993**, *115*, 11552−11568.
(23) Gerlt, J. A.; Gassman, P. G. *Biochemistry* **1993**, *32*, 11943−11952.
(24) Frey, P. A.; Whitt, S. A.; Tobin, J. B. *Science* **1994**, *264*, 1927−1930.
(25) Cleland, W. W.; Kreevoy, M. M. *Science* **1994**, *264*, 1887−1890.
(26) Hibbert, F.; Emsley, J. *Adv. Phys. Org. Chem.* **1990**, *26*, 255−379.
(27) Halkides, C. J.; Wu, Y. Q.; Murray, C. J. *Biochemistry* **1996**, *35*, 15941−15948.
(28) Markley, J. L.; Westler, W. M. *Biochemistry* **1996**, *35*, 11092−11097.
(29) Cassidy, C. S.; Lin, J.; Frey, P. A. *Biochemistry* **1997**, *36*, 4576−4584.
(30) Lin, J.; Westler, W. M.; Cleland, W. W.; Markley, J. L.; Frey, P. A. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 14664−14668.
(31) Stratton, J. R.; Pelton, J. G.; Kirsch, J. F. *Biochemistry* **2001**, *40*, 10411−10416.
(32) Ash, E. L.; Sudmeier, J. L.; De Fabo, E. C.; Bachovchin, W. W. *Science* **1997**, *278*, 1128−1132.
(33) Bachovchin, W. W. *Magn. Reson. Chem.* **2001**, *39*, S199−S213.

Scheme 3. The Two Ways in Which the Tetrahedral Intermediates Can Collapse



solution, results in a reduction of the catalytic effect of these hydrogen bonds.[34,35] By using energy considerations they have shown that the pK_a of the catalytic His-57 and Asp-102 must be very different to stabilize the transition state (relative to water), and thus a LBHB in this system should not exist.

The second hypothesis proposed for the mechanism of SPs concerns the driving force for the collapse of the two tetrahedral intermediates. The first tetrahedral intermediate (TI1), once formed, can collapse in two ways, expelling the serine alkoxide ion or producing the acyl-enzyme intermediate by expelling the amine ion (Scheme 3). A potential problem, first recognized by Polgar,[36] arises because the alkoxide ion is a much better leaving group (p$K_b \approx 1$) than the amine anion (p$K_b \approx 16$), and, therefore, the first option, which is a back reaction, is favored. If the histidine imidazole is optimally aligned to accept a proton from the serine hydroxyl, it is also optimally aligned to donate the proton back to serine in the tetrahedral intermediate, giving a catalytic advantage to the back reaction, which is favored anyway. The same holds for the second tetrahedral intermediate (TI2), in which the imidazole is aligned to accept a proton from the hydrolytic water and is thus also aligned to donate it back to the hydroxyl ion, rather than to the serine alkoxide ion to complete the reaction.

To resolve this perceived problem, it was first suggested that a proton transfer occurs along bent hydrogen bonds, when His-57 is rigidly held midway between the positions ideal for hydrogen bonding to Ser-195 and to the nitrogen atom of the scissile peptide bond.[37] Another proposed solution was that the mechanism is similar to that of the analogous mechanism for general base catalysis of aminolysis in solution, except that the diffusion process in solution is replaced by movement of the imidazole group relative to the substrate at the active site of the enzyme; that is, His-57 is moving between the ideal positions.[38] ^{15}N NMR work led to the proposal of the histidine moving mechanism,[39] suggesting that the most effective reaction profile in accelerating the rate of the formation of the acyl-enzyme might actually be when the Asp-His hydrogen bond breaks on the formation of the tetrahedral intermediate. Theoretical work, based on electron and atom transfer theory as well as dynamics of conformational nuclear modes, which was substantiated by structural and kinetic evidence,[40] showed that the imidazole of His-57 is mobile between the donor and acceptor groups. It was demonstrated that this mobility is crucial and is increased in the tetrahedral state due to the breaking of the histidine hydrogen bonds and the charge alignment of (Asp^- His^+ TI^-), which involves a much lower curvature of the potential surface than in both the initial enzyme−substrate state and the acyl-enzyme state. The activation energy liberated after the first proton transfer is not dissipated to the surrounding protein atoms or to the solution, but stored in a "conformational system" and used in the second proton transfer step (the collapse of the tetrahedral intermediate). A more recent ^{13}C NMR study, combined with inhibitor perturbation experiments, on α-lytic protease and subtilisin BPN′ suggested that the existence of a C−H···O hydrogen bond between His-57 and the backbone carbonyl of Ser-214 may enable a reaction-driven histidine "ring-flip mechanism".[41] According to the ring-flip mechanism, changes in positions between $N^{\epsilon 2}$−H and $C^{\epsilon 1}$−H and between $N^{\delta 1}$−H and $C^{\delta 1}$−H facilitate the transfer of the proton in the productive direction during both acylation and deacylation steps. This hypothesis is based on the assumption that the evolutionary conservation of the C−H···O hydrogen bond in all known serine SPs[42] indicates that it could play a crucial role in the enzyme mechanism. However, so far there is no experimental evidence for this mechanism.

To test the new hypotheses, the dynamics of the active site have to be examined at the quantum mechanical level (where bond-forming and -breaking processes can take place), while retaining the protein environment. This can be done by using a combined quantum mechanical/molecular mechanical (QM/MM) method.[43] The QM/MM potential is a combination of a QM potential, representing the part of the system (solute) where the most important electronic changes take place (e.g., the enzyme active site), and a MM force field, representing the rest of the system (e.g., the surrounding protein). The QM potential is perturbed by the coupled MM potential. The implementation of this method, first introduced by Warshel and Levitt,[44] is becoming increasingly popular for the study of chemical reactions in solution and in enzymes.[45−49] However, because of the high computational cost of representing some of the atoms quantum mechanically, the solute region needs to be chosen carefully.

In the present work, we present the results of a QM/MM dynamics simulation on the second tetrahedral intermediate (of the deacylation step, TI2) in elastase catalysis. The tetrahedral intermediate was chosen as a template because it is presumably similar to the transition state of the reaction and is thus ideal

(34) Warshel, A.; Papazyan, A.; Kollman, P. A. *Science* **1995**, *269*, 102−106.
(35) Warshel, A.; Papazyan, A. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 13665−13670.
(36) Polgar, L. *Acta Biochim. Biophys. Acad. Sci. Hung.* **1971**, *7*, 29−34.
(37) Wang, J. H. *Proc. Natl. Acad. Sci. U.S.A.* **1970**, *66*, 874−881.
(38) Satterthwait, A. C.; Jencks, W. P. *J. Am. Chem. Soc.* **1974**, *96*, 7018−7031.
(39) Bachovchin, W. W. *Biochemistry* **1986**, *25*, 7751−7759.
(40) Sumi, H.; Ulstrup, J. *Biochim. Biophys. Acta* **1988**, *955*, 26−42.
(41) Ash, E. L.; Sudmeier, J. L.; Day, R. M.; Vincent, M.; Torchilin, E. V.; Haddad, K. C.; Bradshaw, E. M.; Sanford, D. G.; Bachovchin, W. W. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 10371−10376.
(42) Derewenda, Z. S.; Derewenda, U.; Kobos, P. M. *J. Mol. Biol.* **1994**, *241*, 83−93.
(43) Gogonea, V. *Internet Electron. J. Mol. Des.* **2002**, *1*, 173−184.
(44) Warshel, A.; Levitt, M. *J. Mol. Biol.* **1976**, *103*, 227−249.
(45) Singh, U. C.; Kollman, P. A. *J. Comput. Chem.* **1986**, *7*, 718−730.
(46) Bash, P. A.; Field, M. J.; Davenport, R. C.; Petsko, G. A.; Ringe, D.; Karplus, M. *Biochemistry* **1991**, *30*, 5826−5832.
(47) Stanton, R. V.; Hartsough, D. S.; Merz, K. M., Jr. *J. Comput. Chem.* **1994**, *16*, 113−128.
(48) Thompson, M. A. *J. Phys. Chem.* **1995**, *99*, 4794−4804.
(49) Gao, J. In *Reviews in Computational Chemistry*; Boyd, D. R., Ed.; VCH: New York, 1996; Vol. 7, pp 119−185.

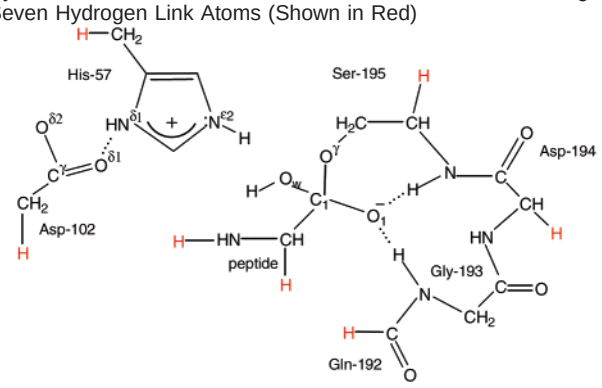
for examining the behavior of the hydrogen-bonding network at the transition state. In addition, it provided a way to test the stability of a tetrahedral intermediate, whose characteristics have been already examined in a previous work,[19] but only with the use of a MM force field that cannot describe bond-breaking and -forming processes. The computational methodology used is an improvement in comparison to previous MD studies in the area,[19,50,51] due to the use of a QM/MM potential with an extensive QM region. In addition, to give a reliable description of the intermediate, the quantum mechanical part was treated at the ab initio level, unlike most other studies using QM/MM potentials,[43−47] which have used semiempirical methods[52,53] for this part. This is computationally more demanding, but more desirable in the current system, due to the extensive hydrogen-bonding network, which would require parametrization work if semiempirical methods were used. Thus, the simulation is very challenging from the computational point of view, because a large quantum region (including both the catalytic triad and the oxyanion hole) is described at the ab initio level. Charge distributions, bond orders, and geometric parameters resulting from the analysis of the simulation provide a better understanding of how SPs work.

Methods

The structural model of the tetrahedral intermediate (TI2) is based on the acyl-enzyme crystal structure of the porcine pancreatic elastase (PPE) with its natural heptapeptide substrate, human β-casomorphin-7 (BCM7),[54] which was retrieved as entry 1QIX of the Brookhaven Protein Data Bank (PDB).[55] The system was set up following the procedure described in our earlier work.[19] The model contained 240 protein residues, 4 substrate residues, and 2800 added TIP3P water molecules,[56] giving a total of 12 053 atoms. The system was allowed to equilibrate for 50 ps using classical MD at 300 K, to relax the crystal structure, and was then minimized. This was performed using the program CHARMM (version 27).[57] The resulting minimized structure was the starting point for the QM/MM dynamics simulation, which was performed with CHARMM interfaced to GAMESS-US.[58]

The active site of the enzyme was described by a QM Hamiltonian, and the influence of the remainders of the protein and the solvent was incorporated via a coupled MM potential. To have an improved description of the highly polarized catalytic region of elastase, the QM region consisted of the catalytic triad and the oxyanion hole: the carboxylate of Asp-102, the imidazole ring of His-57, the tetrahedral fragment, a fragment of the peptide residue Ile-7, and fragments of Ser-195, Gln-192, Gly-193, and Asp-194 (Scheme 4). Hydrogen link atoms were added to satisfy the valence of the QM part.[45,59] These atoms, which do not interact with the neutral MM group consisting of

Scheme 4. The 56 Atom QM Region Used in the QM/MM Dynamics Simulation of the Tetrahedral Intermediate, Including Seven Hydrogen Link Atoms (Shown in Red)



the frontier atom and the atoms directly bonded to it,[60] were restrained on the bond between the QM and the MM atoms[61] with a harmonic potential (k = 1000 kcal mol^{-1} Å^{-2}). The system was also tested with link atoms, which interact with all MM charges. The Mulliken population[62] and the natural bond orbital (NBO) population[63] analyses of the QM wave function perturbed by the MM atoms, which were performed using the Gaussian 98 program,[64] showed no significant difference in the charges between the two cases. In total, 56 atoms were treated quantum mechanically, including 7 link atoms.

The QM region was treated with Hartree−Fock (HF) theory using the 3-21G basis set. This basis set was chosen taking into account the need for both computational accuracy and efficiency. The use of a larger basis set will increase the demand for computational resources and would not allow the application to as extensive a QM region as is described here, which is a necessity for the purpose of this study. Several studies have demonstrated that the 3-21G basis can give reasonable results for molecular geometries,[65] and the performance of the HF/3-21G:TIP3P potential was found to be in good agreement with full ab initio results[66] and has been used in condensed phase simulations.[67,68] Moreover, in our previous QM/MM study on the same system (but with a different size of the QM region), it was demonstrated that the geometries, charges, and bond orders obtained at the HF/3-21G level are similar to those obtained at the HF/6-31G* level.[69]

QM/MM dynamics were performed on the system for 26 ps, with stochastic boundary conditions as described earlier,[19,69] using an integration time step of 1 fs. This unusually long simulation took approximately 60 000 CPU h on 1 GHz Intel PIII processors. Frictional coefficients of 62 and 200 ps^{-1} for water oxygens and heavy atoms, respectively, were used in the protein buffer region.[70] For nonbonded

(50) Nakagawa, S.; Yu, H.-A.; Karplus, M.; Umeyama, H. *Proteins* **1993**, *16*, 172−194.
(51) De Santis, L.; Carloni, P. *Proteins: Struct., Funct., Genet.* **1999**, *37*, 611−618.
(52) Stewart, J. J. P. In *Reviews in Computational Chemistry*; Boyd, D. B., Ed.; VCH: New York, 1990; Vol. 1, pp 45−81.
(53) Zerner, M. C. In *Reviews in Computational Chemistry*; Boyd, D. B., Ed.; VCH: New York, 1991; Vol. 2, pp 313−365.
(54) Wilmouth, R. C.; Clifton, I. J.; Robinson, C. V.; Roach, P. L.; Aplin, R. T.; Westwood, N. J.; Hajdu, J.; Schofield, C. J. *Nat. Struct. Biol.* **1997**, *4*, 456−462.
(55) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. *Nucleic Acids Res.* **2000**, *28*, 235−242.
(56) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926−935.
(57) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187−217.
(58) Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S.; Windus, T. L.; Dupuis, M.; Montgomery, J. A., Jr. *J. Comput. Chem.* **1993**, *14*, 1347−1363.
(59) Field, M. J.; Bash, P. A.; Karplus, M. *J. Comput. Chem.* **1990**, *11*, 700−733.
(60) Bakowies, D.; Thiel, W. *J. Phys. Chem.* **1996**, *100*, 10580−10594.
(61) Reuter, N.; Dejaegere, A.; Maigret, B.; Karplus, M. *J. Phys. Chem. A* **2000**, *104*, 1720−1735.
(62) Mulliken, R. S. *J. Chem. Phys.* **1955**, *23*, 1833−1841.
(63) Reed, A. E.; Curtiss, L. A.; Weinhold, F. *Chem. Rev.* **1988**, *88*, 899−926.
(64) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Gonzalez, C.; Challacombe, M.; Gill, P. M. W.; Johnson, B. G.; Chen, W.; Wong, M. W.; Andres, J. L.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. *Gaussian 98*; Gaussian Inc.: Pittsburgh, PA, 1998.
(65) Hehre, W. J.; Radom, L.; Schleyer, P. v. R.; Pople, J. A. *Ab Initio Molecular Orbital Theory*; Wiley: New York, 1986.
(66) Freindorf, M.; Gao, J. *J. Comput. Chem.* **1996**, *17*, 386−395.
(67) Gao, J.; Freindorf, M. *J. Phys. Chem. A* **1997**, *101*, 3182−3188.
(68) Byun, K.; Mo, Y.; Gao, J. *J. Am. Chem. Soc.* **2001**, *123*, 3974−3979.
(69) Topf, M.; Várnai, P.; Richards, W. G. *Theor. Chem. Acc.* **2001**, *106*, 146−151.

Table 1. Selected Geometrical Parameters in the Active Site of the Simulated TI2 Averaged between 6 and 26 ps

		Distance (Å)			
O_1	Gly-193-N	2.64 ± 0.09	O_1	Gly-193-N−H	1.63 ± 0.10
O_1	Ser-195-N	3.07 ± 0.26	O_1	Ser-195-N−H	2.16 ± 0.30
His-57-$N^{\epsilon 2}$	C_1	3.53 ± 0.12	C_1	O_1	1.33 ± 0.03
His-57-$N^{\epsilon 2}$	O^w	3.23 ± 0.27	His-57-$N^{\epsilon 2}$−H	O^w	2.44 ± 0.31
His-57-$N^{\epsilon 2}$	Ser-195-O^{γ}	2.64 ± 0.10	His-57-$N^{\epsilon 2}$−H	Ser-195-O^{γ}	1.69 ± 0.19
His-57-$N^{\epsilon 2}$−H^1	His-57-$N^{\epsilon 2}$	1.03 ± 0.03	His-57-$N^{\delta 1}$−H	His-57-$N^{\delta 1}$	1.15 ± 0.14
His-57-$C^{\epsilon 1}$	Thr-213-O	3.67 ± 0.33	His-57-$C^{\epsilon 1}$−H	Thr-213-O	2.79 ± 0.41
His-57-$C^{\epsilon 1}$	Ser-214-O	2.98 ± 0.18	His-57-$C^{\epsilon 1}$−H	Ser-214-O	2.35 ± 0.30
His-57-$N^{\delta 1}$	Asp-102-$O^{\delta 1}$	2.54 ± 0.07	His-57-$N^{\delta 1}$−H	Asp-102-$O^{\delta 1}$	1.40 ± 0.16
Ser-214-O	P_1-N	3.48 ± 0.22	Ser-214-O	P_1-N−H	2.53 ± 0.24
Val-216-O	P_3-N	3.13 ± 0.22	Val-216-O	P_3-N−H	2.17 ± 0.23
Val-216-N	P_3-O	2.92 ± 0.12	Val-216-N−H	P_3-O	1.94 ± 0.13
W^1-O	O_w	2.85 ± 0.24	Arg-61-N−H	W_1-O	1.94 ± 0.20
W_1-H^1	O_w	2.53 ± 0.40	W_1-H^2	O_w	2.37 ± 0.51

	Angle (deg)		
O_w−C_1=O_1	111.4 ± 2.7	C_1···Ser-195-N···Gly-193-N···O_1	10.8 ± 5.1
O_w−C_1···P_1−C^{α}···O_1	124.4 ± 3.4	O_w−C_1−Ser-195-O^{γ}···O_1	120.5 ± 3.1

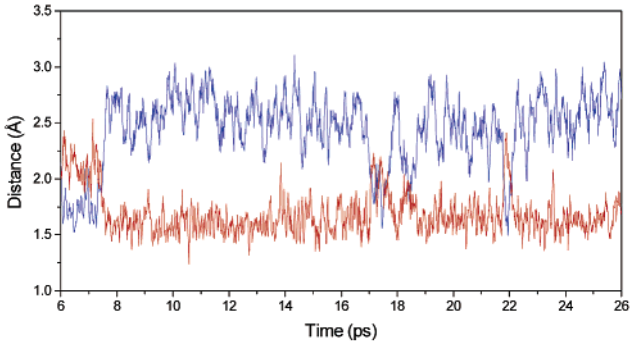


Figure 1. Distance between His-57-$N^{\epsilon 2}$−H and O^w blue, which both originated from the hydrolytic water in the acyl-enzyme state, in comparison with the distance between $N^{\epsilon 2}$−H and Ser-195-O^{γ} (red), between 6 and 26 ps of QM/MM dynamics simulation.

interactions between MM atoms, the electrostatic terms were truncated with a switch function between 10 and 14 Å, and the van der Waals terms were truncated with a shift function with a cutoff distance of 14 Å;[71] the dielectric constant was unity. The first 6 ps of the trajectory (before the system reached 300 K) were not analyzed, and coordinates along the remaining 20 ps were saved every 10 fs for analysis. Bond orders and charges were calculated every 0.5 ps from the converged QM wave function, perturbed by the MM atoms (to account for the influence of the environment), using the NBO analysis. For comparison purposes, the bond orders of a tetrahedral fragment [$CH_3C(O)^-OHOCH_3$] with two NH_3 molecules hydrogen bonded to its charged oxygen were calculated in the gas phase at the same level of theory (HF/3-21G).

Results and Discussion

Hydrogen Bond between His-57-$N^{\epsilon 2}$−H and O_w or Ser-195-O^{γ}. The modeled tetrahedral intermediate (TI2) was found to be a stable species for the entire 26 ps of simulation time. This is consistent with recent time-resolved crystallographic data in which some form of a tetrahedral intermediate was detected.[17] At the beginning of the simulation, His-57-$N^{\epsilon 2}$−H formed a short hydrogen bond with O_w, the oxygen that originated from the nucleophilic water molecule in the acyl-enzyme state (Figure 1). This interaction was lost after ~7.6 ps due to a movement of the imidazolium ring toward Ser-195, and a new hydrogen bond was formed between $N^{\epsilon 2}$−H and Ser-195-O^{γ}. The imidazolium switched hydrogen bonds to O_w for short periods during the simulation, but spent most of the time interacting with O^{γ}. The $N^{\epsilon 2}$−H···O_w and $N^{\epsilon 2}$−H···O^{γ} distances, which were averaged between 6 and 26 ps, are 2.44 and 1.69 Å, respectively (Table 1). In addition, O_w was stabilized by a hydrogen-bonding interaction with a water molecule, which in turn was hydrogen bonded to the guanido group of Arg-61, throughout the simulation. The average O_w···C^1−O_1 angle is 111.4°, close to an ideal tetrahedral angle, and the average charge on O_w is −0.88*e* (Tables 1 and 2).

Table 2. Selected NBO Charges and Bond Orders in the QM Region Averaged between 6 and 26 ps

	atomic charges		bond orders
Gly-193-N	−0.82	O_1···Gly-193-N−H	0.13
Ser-195-N	−0.76	O_1···Ser-195-N−H	0.06
C_1	0.89	O_w−H	0.70
O_1	−0.96	O_1−C_1	1.14
O_w	−0.88	O_w−C_1	0.82
Ser-195-O^{γ}	−0.68	O^{γ}−C_1	0.79
His-57-$N^{\delta 1}$	−0.65	His-57-$N^{\epsilon 2}$−H···O_w	0.02
His-57-$N^{\epsilon 2}$	−0.64	His-57-$N^{\epsilon 2}$−H···Ser-195-O^{γ}	0.09
His-57-$C^{\epsilon 1}$	0.37	His-57-$N^{\epsilon 2}$−H	0.61
Asp-102-$O^{\delta 1}$	−0.87	His-57-$N^{\delta 1}$−H	0.48
Asp-102-$O^{\delta 2}$	−0.81	Asp-102-$O^{\delta 1}$···His-57-$N^{\delta 1}$−H	0.24

To complete the deacylation reaction in SP, a proton must transfer from His-57-$N^{\epsilon 2}$ to Ser-195-O^{γ}, which results in the collapse of the tetrahedral intermediate to the enzyme−product complex. The expected lifetime of the tetrahedral intermediate is not known, but, clearly, the simulation implies that the tetrahedral intermediate is stable on the picosecond time scale in solution. However, the mobility of the imidazolium ring of His-57 and the thermal fluctuations from O_w to O^{γ} probably permit the productive breakdown of the intermediate, which is in agreement with the results of the theoretical study mentioned above.[40] Thus, without a histidine flip, suggested in the ring-flip mechanism,[41] the tetrahedral intermediate could proceed to the product state. On the basis of preliminary calculations modeling the reaction, we expect the free energy of this step (the proton transfer from the histidine to O^{γ}) to be about 8 kcal mol^{-1}. However, it should be noted that although no ring flip was observed in the present simulation or in our previous 1 ns long MD simulation of the tetrahedral intermediate,[19] this possibility cannot be eliminated because those simulations are too short to observe such a movement, which is probably on the 10^{-5}−10^{-2} s time scale (based on tyrosine ring-flip rates[72]).

(70) Brunger, A. T.; Huber, R.; Karplus, M. *Biochemistry* **1987**, *26*, 5153−5162.
(71) Steinbach, P. J.; Brooks, B. R. *J. Comput. Chem.* **1994**, *15*, 667−683.

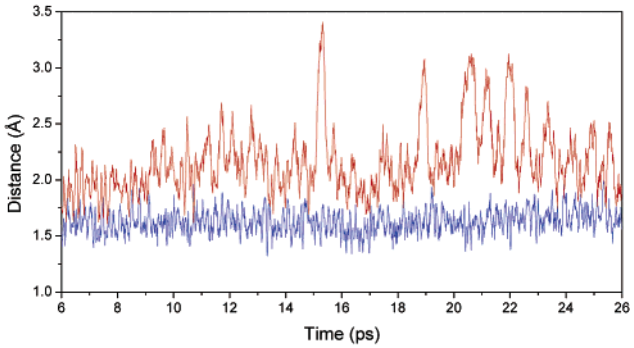


Figure 2. Hydrogen-bonding interactions in the oxyanion hole: the distances between O_1 and the backbone hydrogens (N−H) of Gly-193 (blue) and Ser-195 (red), between 6 and 26 ps of QM/MM dynamics simulation.

The Oxyanion Hole. As anticipated, the tetrahedral carbonyl oxygen (O_1) formed hydrogen bonds with the backbone N−H groups of Gly-193 and Ser-195 in the oxyanion hole and was located close to the plane defined by C_1 and the backbone N atoms. The O_1···Gly-193-N−H and O_1···Ser-195-N−H distances, which were averaged between 6 and 26 ps, are 1.63 and 2.16 Å, respectively, and the C_1···Ser-195-N···Gly-193-N···O_1 dihedral angle is 10.8° (Table 1). However, while the O_1···Gly-193-N−H hydrogen bond was continuously short and stable, the O_1···Ser-195-N−H hydrogen bond kept breaking and forming (Figure 2). The average charge on Gly-193-N−H is −0.82, and the average charge on Ser-195-N−H is −0.76. The average NBO bond order between O_1 and Gly-193-N−H is 0.13, while that between O_1 and Ser-195-N−H is 0.06 (Table 2). The bond orders of the corresponding hydrogen bonds between the charged oxygen of a tetrahedral fragment [$CH_3C(O)^-OHOCH_3$] and two NH_3 molecules are both 0.07. These results reveal an asymmetry between the stabilizing interactions in the oxyanion hole.

A number of theoretical[21,73] and experimental[20,74] studies on wild-type and mutants of a bacterial SP (subtilisin) have shown that the oxyanion hole residue Asn-155 is a key residue for the biological function, in that it provides a stabilization of the transition state relative to the ground state by approximately 5 kcal mol^{-1}. However, only one (theoretical) study on the mammalian isoenzymes has appeared to shed light on the crucial role of Gly-193,[51] which is presumably similar to Asn-155. The present results also indicate an important role of the Gly-193 hydrogen bond in the oxyanion hole, although they cannot provide the interaction energy. They are in contrast, however, with the time-resolved crystallographic data for the tetrahedral intermediate,[17] in which the O_1···Gly-193-N distance is longer than the O_1···Ser-195-N distance (2.95 and 2.65 Å, respectively). However, we believe that our results are more convincing because in the product release stage the inner O_1 and Ser-195-N−H bond probably breaks first, and hence the weakening of this hydrogen bond represents a step in the direction of the products. This assumption is consistent with our previous results,[69] in which O_1 was shown to be hydrogen bonded only to Gly-193-N−H in the enzyme−product complex (EP).

C−H···O Hydrogen Bond between His-57-C$^{\epsilon 1}$−H and the Backbone Carbonyl of Ser-214. The occurrence of a C−H···O hydrogen bond between the imidazolium ring of His-57 (via its C$^{\epsilon 1}$−H group) and the backbone carbonyl of Ser-214 in the active site of SPs has already been suggested in previous studies.[41,42] The present study not only supports this proposal at the TI2 stage, but also suggests that the C$^{\epsilon 1}$−H group might be involved in another interaction. According to our results, during the simulation, His-57-C$^{\epsilon 1}$−H was mostly hydrogen bonded to the backbone carbonyl of Ser-214. However, because of movements of the imidazolium ring, the C$^{\epsilon 1}$−H group also interacted with the backbone carbonyl of Thr-213. The interactions of the imidazolium C$^{\epsilon 1}$−H group with both Ser-214 and Thr-213 were also observed in our previous 1 ns long MD simulation of the TI2 model.[19] The distance between C$^{\epsilon 1}$−H and Thr-213-O had larger values than the distance between C$^{\epsilon 1}$−H and Ser-214-O: the averages between 6 and 26 ps are 2.79 and 2.35 Å, respectively (Table 1). However, on a few occasions during the simulation (∼1 ps in total), the C$^{\epsilon 1}$−H···Thr-213-O distance became shorter than the C$^{\epsilon 1}$−H···Ser-214-O distance (Figure 3A). An analysis of selected geometrical parameters related to the C$^{\epsilon 1}$−H···O═C interactions of Ser-214 and Thr-213 is presented in Table 3. These were averaged for each residue only when the C$^{\epsilon 1}$−H···O═C distance was shorter than the corresponding distance of the other residue. For an optimal C$^{\epsilon 1}$−H···O═C hydrogen bond, the H···O−C (ξ) and the H ···O−C−N (χ) angles should be close to 120° and 180°, respectively, as required by the idealized sp^2 hybridization of the oxygen atom.[75] The C$^{\epsilon 1}$−H···O distance should be less than 2.6 Å (the sum of the van der Waals radii

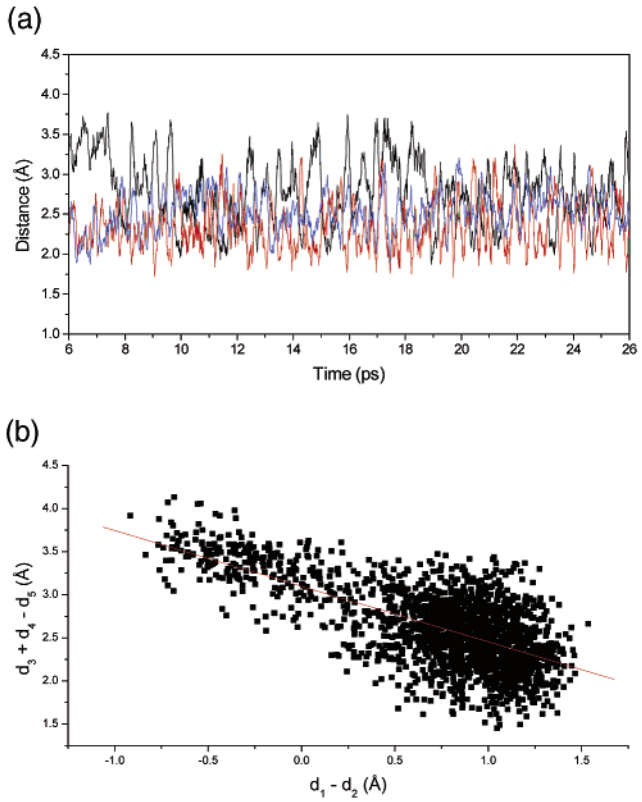


Figure 3. (A) Distances of Ser-214-O from P_1−N−H (blue) and from His-57-C$^{\epsilon 1}$−H (red) and of Thr-213-O from His-57-C$^{\epsilon 1}$−H (black), between 6 and 26 ps of QM/MM dynamics simulation. (B) Plot of the linear combination of the C$^{\epsilon 1}$−H···O(Ser-214) (d_3), C$^{\epsilon 1}$−H···O(Thr-213) (d_4), and Ser-214-O···P_1−N−H (d_5) distances versus the difference between the N$^{\epsilon 2}$−H^1···O_w (d_1) and N$^{\epsilon 2}$−H^1···O$^\gamma$ (d_2) distances. The correlation coefficient (R) is −0.7.

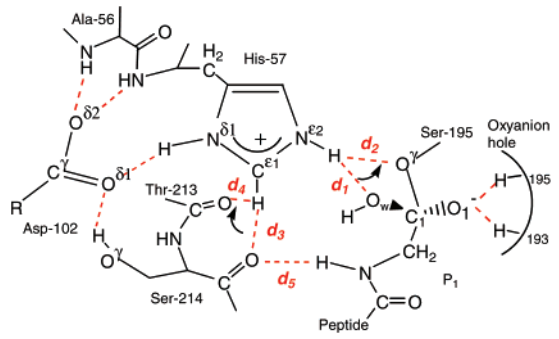
(72) Nall, B. T.; Zuniga, E. H. *Biochemistry* **1990**, *29*, 7576−7584.
(73) Hwang, J.-K.; Warshel, A. *Biochemistry* **1987**, *26*, 2669−2673.
(74) Wells, J. A.; Cunningham, B. C.; Graycar, T. P.; Estell, D. A. *Philos. Trans. R. Soc. London, Ser. A* **1986**, *A317*, 415−423.

Table 3. Selected Geometric Parameters of the $C^{\epsilon 1}-H\cdots O{=}C$ Hydrogen Bonds Formed between the Imidazolium Ring and Ser-214 and Thr-213[a]

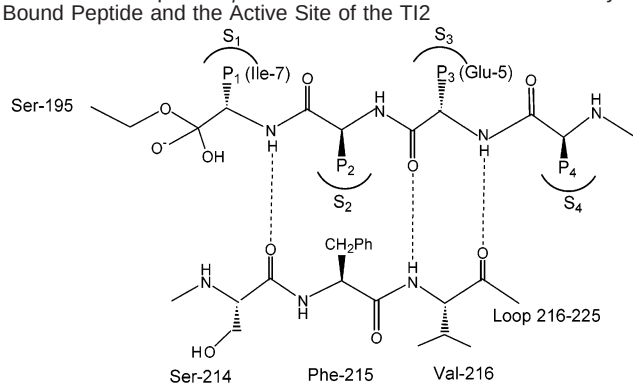
	Ser-214[a]	Thr-213[a]
	Distance (Å)	
$C^{\epsilon 1}$ O	2.76 ± 0.17	3.30 ± 0.07
$C^{\epsilon 1}-H$ O	2.16 ± 0.28	2.29 ± 0.09
	Angle (deg)	
$C^{\epsilon 1}-H\cdots O$	111 ± 14	158 ± 9
$H\cdots O-C$ (ξ)	101.4 ± 8.0	127.3 ± 5.1
$H\cdots O-C-N$ (χ)	157.2 ± 7.4	163.4 ± 7.2
θ ($\sin^{-1}(\sin\xi \sin\chi)$)	22	13

[a] The parameters were averaged for each residue over the simulation time only when its $C^{\epsilon 1}-H\cdots O{=}C$ distance was shorter than the corresponding distance for the other residue.

Scheme 5. The Hydrogen-Bonding Network Created in the Active Site of the Tetrahedral Intermediate



Scheme 6. Antiparallel β-Sheet Formed between a Productively Bound Peptide and the Active Site of the TI2



of oxygen and hydrogen atoms used for the simulation), and the $C^{\epsilon 1}\cdots O$ distance should be less than 3.5 Å (the sum of the van der Waals radii of oxygen and carbon atoms). The "elevation" angle, θ, which is the angle between the $C^{\epsilon 1}-H\cdots O{=}C$ hydrogen bond and the plane of the carbonyl carbon sp^2 electron orbitals,[76] should ideally be close to 0°. It is clear that the present results are consistent with all of these requirements; that is, the interaction between the imidazolium ring and Ser-214-O, or alternately that between the imidazolium ring and Thr-213-O, has all of the stereochemical characteristics of a proper hydrogen bond (Table 3).

In addition to the hydrogen bond it formed with His-57-$C^{\epsilon 1}-H$, Ser-214-O formed a hydrogen bond with the backbone N−H group of the P_1 residue, in subsite S_1 of the substrate binding pocket (Schemes 5 and 6). The Ser-214-O$\cdots P^1-N-H$ distance, which was averaged between 6 and 26 ps, is 2.53 Å (Table 1,

(75) Derewenda, Z. S.; Derewenda, U.; Kobos, P. M. *J. Mol. Biol.* **1995**, *252*, 248−262.
(76) Taylor, R.; Kennard, O. *J. Am. Chem. Soc.* **1982**, *104*, 5063−5070.

Table 4. Selected Geometrical Parameters of Two Minima Found for the TI2 Model Using an Extended QM Region

		Min1	Min2
	Distance (Å)		
His-57-$C^{\epsilon 1}-H$	Ser-214-O	2.20	2.68
His-57-$C^{\epsilon 1}-H$	Thr-213-O	3.69	2.44
His-57-$N^{\epsilon 2}-H$	O^w	1.75	2.38
His-57-$N^{\epsilon 2}-H$	Ser-195-O^{γ}	2.18	1.59
His-57-$N^{\delta 1}-H$	Asp-102-$O^{\delta 1}$	1.57	1.52
O_1	Gly-193-N−H	1.57	1.60
O_1	Ser-195-N−H	1.88	2.77
	Angle (deg)		
His-57-N−$C^{\alpha}-C^{\beta}-C^{\gamma}$		82.6	71.9
His-57-$C^{\alpha}-C^{\beta}-C^{\gamma}-N^{\delta 1}$		−117.4	−95.9

Figure 3A). As mentioned above, the movement of His-57 imidazolium between O_w and O^{γ} is related to fluctuations in the direction of the acyl-enzyme and the products, respectively. This motion can be represented by the difference between the $N^{\epsilon 2}-H\cdots O_w$ distance and the $N^{\epsilon 2}-H\cdots O^{\gamma}$ distance ($d_1 - d_2$) (Scheme 5, Figure 3B), with negative values showing backward, and positive values showing forward, movement along the reaction coordinate. The movement of the imidazolium ring affects the fluctuations in the $C^{\epsilon 1}-H\cdots O$ interactions between His-57 and the backbone carbonyl groups of Ser-214 and Thr-213. This movement can be therefore represented by the sum of these hydrogen-bond lengths ($d_3 + d_4$). However, movements of Ser-214-O are also likely to affect its hydrogen-bonding interaction with the P_1 substrate residue (Ile-7) via its N−H group (d_5). From an investigation of various combinations of the distances mentioned above, the best correlation was found between ($d_1 - d_2$) and ($d_3 + d_4$) − d_5, possibly reflecting to some extent reactant-like and product-like thermal fluctuations (Figure 3B). During those fluctuations, the hydrogen bond between His-57-$N^{\delta 1}-H$ and Asp-102-$O^{\delta 1}$ stayed short (<1.6 Å).

Because the two $C^{\epsilon 1}-H\cdots O$ hydrogen bonds in the current study involve interactions between MM and QM atoms, to ensure that they are correct, the system was also tested with all of the atoms involved included in the QM region. To do so, two structures obtained from the trajectory, each representing one of those interactions, were minimized using an extended QM region, which also included the full Ser-214 residue and the backbone of Thr-213 (76 atoms in total). In the first structure (obtained at 11.5 ps), the $C^{\epsilon 1}-H\cdots O$ hydrogen bond was formed with Ser-214-O, and in the second (obtained at 12.9 ps), it was formed with Thr-213-O. The minimizations were performed until the average root-mean-square (RMS) gradient was less than 0.1 kcal mol^{-1} $Å^{-1}$. It was found that the two minima exist for the tetrahedral intermediate, with similar energies: one which is closer to the reactant and the other which is closer to the product (Table 4, Figure 4). In the first minimum (Min1), the $C^{\epsilon 1}-H$ is hydrogen bonded to Ser-214, with distances of 2.20 and 3.69 Å between $C^{\epsilon 1}-H$ and the carbonyl oxygens of Ser-214 and Thr-213, respectively. In the second minimum (Min2), the $C^{\epsilon 1}-H$ is closer to Thr-213, with distances of 2.68 and 2.44 Å between $C^{\epsilon 1}-H$ and Ser-214 and Thr-213, respectively. The hydrogen bond formed by the imidazolium via $N^{\epsilon 2}-H$ is with O_w in Min1 and with O^{γ} in Min2, and the hydrogen bond between His-57-$N^{\delta 1}-H$ and Asp-102-$O^{\delta 1}$ is short in both minima. To change from Min1 to Min2, the side chain of His-57 has to rotate its χ_1 (N−$C^{\alpha}-C^{\beta}-C^{\gamma}$ dihedral angle) by ∼11°

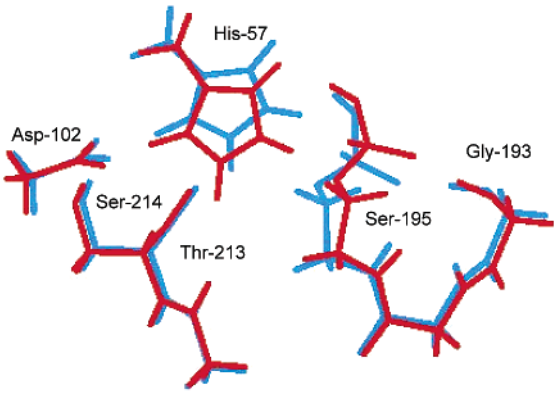


Figure 4. A superposition of the active site residues of the two minima found for the TI2 model using an extended QM region, Min1 (blue) and Min2 (red).

and its χ_2 ($C^{\alpha}-C^{\beta}-C^{\gamma}-N^{\delta 1}$ dihedral angle) by ~22°. Interestingly, the oxyanion hole hydrogen bonds are different in the two minima; the $O_1\cdots$Gly-193-N−H and $O_1\cdots$Ser-195-N−H distances are 1.57 and 1.88 Å in Min1, respectively, and 1.60 and 2.77 Å in Min2, respectively. This result gives more justification to the proposal stated above, that the closer the intermediate gets to the product, the weaker the $O_1\cdots$Ser-195-N−H hydrogen bond in the oxyanion hole becomes.

As stated in the Introduction, the histidine ring-flip process is based on the assumption that the evolutionary conservation of the $C^{\epsilon 1}-H\cdots O$ hydrogen bond between His-57 and the backbone carbonyl of Ser-214 in all known SPs indicates a crucial role in the mechanism.[41] Although supporting the existence of the $C^{\epsilon 1}-H\cdots O$ hydrogen bond, the current results indicate that the existence of this bond does not necessarily require an explanation involving a new mechanism, because this interaction can be partially maintained during the movement of the imidazolium of His-57 toward Ser-195-O^{γ}. Although it is correct that this movement causes the $C^{\epsilon 1}-H\cdots O$ hydrogen bond with Ser-214 to elongate, this loss of stabilization is presumably compensated for by an alternative $C^{\epsilon 1}-H\cdots O$ hydrogen bond that is formed with Thr-213. Not only that, but according to the ring-flip mechanism, during the flip of the histidine ring the hydrogen bond between Asp-102 and His-57 is broken, while in the present study (and in our previous QM/MM[69] and MD[19] studies), the movement suggested for the ring does not require the breaking of this hydrogen bond. In a recent ab initio study on C−H···X hydrogen bonds,[77] the binding energy of $H_2C{=}CH_2\cdots OH_2$ was found to be 1.5 kJ mol^{-1} (~0.4 kcal mol^{-1}) in the gas phase, with an H···O bond length is 2.5 Å. In comparison, a weak or conventional hydrogen bond can vary between ~2 and 12 kcal mol^{-1}.[26] Thus, it is more likely that only the hydrogen-bonding interaction of the histidine ring with Ser-214 will be easily lost, and not the "conventional" hydrogen bond between Asp-102 and His-57.

Hydrogen Bond between His-57-$N^{\delta 1}$−H and Asp-102-$O^{\delta 1}$. A short hydrogen bond was found between $N^{\delta 1}$ of His-57 imidazolium and $O^{\delta 1}$ of Asp-102 carboxylate, in agreement with the MD study on the tetrahedral intermediate[19] and with the time-resolved crystal structure.[17] In addition, during the course of the simulation, proton hopping occurred on several occasions between $N^{\delta 1}$ and $O^{\delta 1}$. The $N^{\delta 1}\cdots O^{\delta 1}$, $N^{\delta 1}-H\cdots O^{\delta 1}$, and $N^{\delta 1}\cdot\cdot H$ distances, averaged between 6 and 26 ps, are 2.54, 1.40, and 1.15 Å, respectively (Table 1). The average bond order between the proton and $O^{\delta 1}$ is 0.24, and the average bond order between the proton and $N^{\delta 1}$ is 0.48 (Table 2). The average charge on $N^{\delta 1}$ is $-0.65e$, and the average charge on $O^{\delta 1}$ is $-0.87e$.

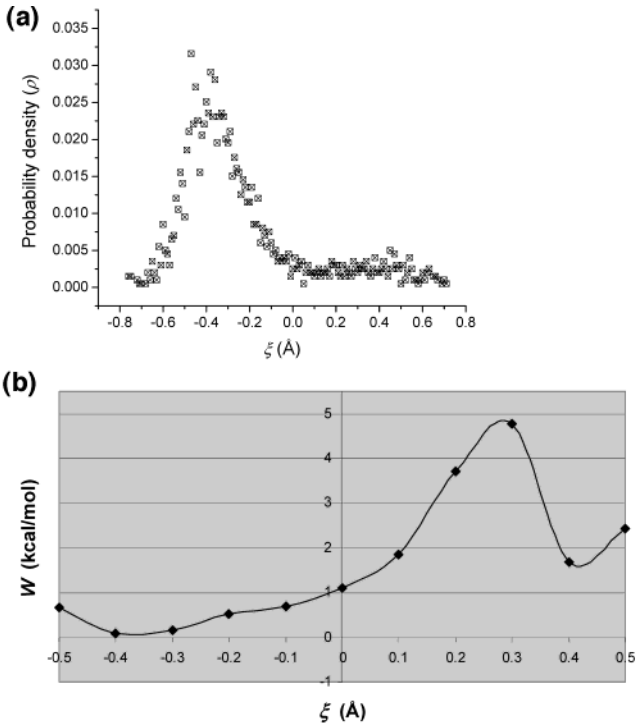


Figure 5. (A) The probability density distribution, ρ, of the proton hopping from His-57-$N^{\delta 1}$ to Asp-102-$O^{\delta 1}$ as a function of the reaction coordinate ξ ($= d(N^{\delta 1}\cdots H) - d(O^{\delta 1}\cdots H)$), calculated from the trajectory at HF/3-21G//CHARMM level. (B) PMF of the proton transfer from His-57-$N^{\delta 1}$ to Asp-102-$O^{\delta 1}$, calculated from the probability distribution, ρ, and including a correction obtained at the B3LYP/6-31G*//CHARMM level.

The probability density distribution of the proton to hop from His-57-$N^{\delta 1}$ to Asp-102-$O^{\delta 1}$, ρ, calculated as a function of the reaction coordinate, ξ, which was chosen as the difference between the distances of the proton from those two atoms ($\xi = d(N^{\delta 1}\cdots H) - d(O^{\delta 1}\cdots H)$), is presented in Figure 5A. The distances were obtained from the simulation at the HF/3-21G//CHARMM level, sampling the trajectory every 10 fs (starting at 6 ps). According to this, the proton was not equally shared between the $O^{\delta 1}$ and $N^{\delta 1}$, but was localized on $N^{\delta 1}$. Subsequently, the potential of mean force (PMF) for the proton transfer was calculated from the probability distribution ($\rho(\xi)$) in the following way:

$$W(\xi) = -RT \ln \rho(\xi)$$

where R is the gas constant, and T is the temperature of the simulation (300 K). In addition, in order to improve the results, a correction from a higher level of theory was added to the PMF. To do so, the potential energy surface (PES) of exactly the same system was twice mapped along the reaction coordinate, with the QM region treated at a low level (LL) and a high level (HL) of theory. The difference in the energy as a function of the reaction coordinate, $E(\xi)$, at the B3LYP/6-31G*//CHARMM level (HL) and at the HF/3-21G//CHARMM level (LL) of theory was added to $W(\xi)^{LL}$ in the following way:

(77) Hartmann, M.; Wetmore, S. D.; Radom, L. *J. Phys. Chem.* **2001**, *105*, 4470−4479.

$$W(\xi)^{HL} = W(\xi)^{LL} + E(\xi)^{HL} - E(\xi)^{LL}$$

assuming that the entropic contribution is similar at both levels, which is reasonable given that the bulk of the system entropy is accounted for by the MM region, and this is largely independent of the level of QM theory. The second assumption is that the minimized structure is representative of an average structure at 300 K.

The PES was mapped using an adiabatic mapping procedure; structures along the reaction coordinate (r) were minimized using the RESD mode[78] in CHARMM27, with a force constant of 1000 kcal mol^{-1} Å^{-2} for the harmonic constraint. Because of the extended QM region used in this study (Scheme 4), the computations required for the QM/MM minimizations were extremely expensive, and, thus, while the low-level energies ($E(\xi)^{LL}$) were determined after minimization at the HF/3-21G//CHARMM level, the high-level energies ($E(\xi)^{HL}$) were determined by a single point calculation at the B3LYP/6-31G*//CHARMM level on the structures minimized along the reaction coordinate at a lower level of theory (HF/6-31G*//CHARMM). The minimizations were performed until the average RMS gradient was less than 0.05 kcal mol^{-1} Å^{-1}.

The PMF of the proton to hop from His-57-N$^{\delta 1}$ to Asp-102-O$^{\delta 1}$ calculated from the probability density including the correction along the reaction coordinate, $W(\xi)^{HL}$, is presented in Figure 5B. The calculated increase in free energy when the proton transfers to Asp-102 is ~1.6 kcal mol^{-1}, with an activation barrier of ~4.8 kcal mol^{-1}. This increase corresponds to a difference of ~1.2 pK_a units and to a 94:6 division of the proton between His-57 and Asp-102, respectively. Without including the correction, the calculated increase in free energy is 1.2 kcal mol^{-1}, and the activation barrier is 1.5 kcal mol^{-1}, which corresponds to a difference of 0.9 pK_a unit and to a 89:11 division. Thus, the calculated increase in free energy is similar at both levels of theory (HF/3-21G//CHARMM and B3LYP/6-31G*//CHARMM), although the activation barrier is underestimated at the HF/3-21G//CHARMM level. In any case (with or without including the correction to the free energy), the results are consistent with recent NMR data which show that the proton is at least 85% localized on the N$^{\delta}$,[1,32,33] and they indicate even more strongly that this hydrogen bond does not fall under the definition of a LBHB, which requires equal pK_a values for the heteroatoms involved (i.e., that the proton is equally shared between them). Furthermore, these results support the claim that the proton hopping destabilizes the intermediate and, thus, produces an anticatalytic effect.[34]

The Binding Pocket and Loop 216−225. In the MD study described in our previous paper,[19] a loop opening was observed for the TI2 model (but not for the model of the acyl-enzyme) between residues Val-216 and Pro-225, below the binding pocket of the substrate. Although the present simulation was too short to observe a full loop opening, the constant increase in the RMS deviation of the location of the loop from its original location along the trajectory (reaching a mean value of 0.6 Å in the last 3 ps) could possibly indicate such a change. This is in agreement with a recent study on over 500 sequences of "trypsin-like" homologous enzymes, which investigated the structure−function relationship in trypsins[79] and suggested that the same loop (216−225) can modulate either or both of the S_1 sites or the subsite preferences on the N-terminal side of the scissile bond. In the time-resolved crystallographic work, this loop was seen to have the largest structural changes.[17]

In addition, the P_3 residue (Glu-5) fluctuated more than the other peptide residues in the binding pocket, as seen in the MD simulation. The mean RMS deviation of the location of P_3 from its original location is 0.24 Å, whereas those of P_1, P_2, and P_4 are 0.16, 0.13, and 0.15 Å, respectively. This result could be associated with the movement of loop 216−225, because the Val-216-N−H···P_3-O hydrogen bond in subsite S_3, which is further away from the center of the loop, stayed stable along the trajectory, while the P_3-N−H···Val-216-O kept forming and breaking. The P_3-O···Val-216-N−H and P_3-N−H···Val-216-O distances, which were averaged between 6 and 26 ps, are 1.94 ± 0.13 and 2.17 ± 0.23 Å, respectively (Table 1). As suggested in the MD study, the movement of the loop and the P_3 residue could be implicated in product release.

Conclusion

Ab initio QM/MM dynamics simulation of the tetrahedral intermediate shows that the intermediate is a stable species on the picosecond time scale. On the basis of geometrical and dynamical properties, and in agreement with many experimental and theoretical studies, it is suggested that the crucial hydrogen bonds involved in stabilizing this intermediate are between Asp-102 and the protonated His-57 and between the charged oxygen of the intermediate and the N−H group of Gly-193 in the oxyanion hole. Proton hopping occurred between Asp-102 and His-57 during the simulation; however, the proton was mostly localized on the nitrogen, and hence it is not participating in a low-barrier hydrogen bond. When the proton hopping occurs, it has a destabilizing effect on the intermediate.

The mobility of the imidazolium of His-57 between O_w and O^{γ}, two of the oxygens of the tetrahedral structure, implies that the intermediate can proceed to the product state without a ring-flip mechanism, although the simulation was not long enough to observe this. During this movement, the hydrogen bond with Asp-102 is maintained as well as a weak C$^{\epsilon 1}$−H···O hydrogen bond with the backbone carbonyl of Ser-214. The loss in the latter interaction is compensated for by an alternative C$^{\epsilon 1}$−H···O hydrogen bond between the imidazolium and the backbone carbonyl of Thr-213.

The study also indicates two factors that may be implicated in product release: breaking the hydrogen bond of the charged oxygen with the backbone N−H group of Ser-195 in the oxyanion hole, and the beginning of a loop opening between residues 216−225 that possibly induces a break of the hydrogen bond in subsite S_3.

Acknowledgment. We are grateful to Ben Webb and Christopher J. Schofield for helpful discussions. M.T. is funded by the Overseas Research Students (ORS) Awards Scheme and the Wingate Scholarships. P.V. thanks the Wellcome Trust for an International Prize Travelling Research Fellowship (grant reference number 060078). The work was in part supported by a grant from the National Foundation for Cancer Research. We thank the Oxford Supercomputing Centre for the generous allocation of computer time.

JA026219Q

(78) Eurenius, K. P.; Chatfield, D. C.; Brooks, B. R.; Hodoscek, M. *Int. J. Quantum Chem.* **1996**, *60*, 1189−1200.

(79) Perona, J. J.; Craik, C. S. *J. Biol. Chem.* **1997**, *272*, 29987−29990.