Dynamic Properties of the First Steps of the Enzymatic Reaction of Porcine Pancreatic Elastase (PPE). 2. Molecular Dynamics Simulation of a Michaelis Complex: PPE and the Hexapeptide Thr-Pro-nVal-Leu-Tyr-Thr

M. Geller,† S. M. Swanson, and E. F. Meyer, Jr.*

Contribution from the Department of Biochemistry and Biophysics, Texas A&M University, College Station, Texas 77843. *Received February 21, 1990*

Abstract: A hydrated Michaelis complex of a hexapeptide (Thr-Pro-nVal-Leu-Tyr-Thr) bound to porcine pancreatic elasase was simulated for 70 ps by use of molecular dynamics (AMBER 3.0). Input coordinates came from a recent crystallographic study of a related complex of elastase and an inhibitor. The dynamic properties of active-site residues (including four residues important for catalysis: Ser-195, His-57, Asp-102, and Ser-214) are analyzed and compared to several previous studies. One hydrogen bond, between N_ϵ of His-57 and the HO_γ groups of Ser-195, is thought to be crucial in the charge relay model of serine protease catalysis. This bond is not observed in simulations of the native enzyme, but is seen when the substrate replaces waters in the active site of the complex. Good hydrogen bonding requires modification of the standard AMBER side-chain charges of His-57 by taking into account the influence of Asp-102. This study reports a significantly higher frequency of attractive interactions and H-bond formation (during 75% of 70 ps) seen with the modified charges. At six times during the simulation, very short contact distances (2.56-2.59 Å) were seen in this bond. Such close donor-acceptor contacts could facilitate proton transfer from Ser-195 to His-57 and might be a crucial component in the catalytic mechanism.

The most extensively studied enzyme family to date, the serine proteinases, shares a catalytic mechanism that is only approximately understood. The well-known "catalytic triad" (Asp-102;His-57:Ser-195)[1], revealed in the late 1960s by Blow and co-workers,[2-4] forms the basis of a simple structure-function understanding of the catalytic mechanism. This sequence of events, leading to the hydrolysis of a peptide bond, is taken as a typical example by most biochemistry textbooks. Many experimental and theoretical efforts have been made to understand the chemical and physical nature of this enzymatic process in greater detail. In mechanistic terms, it is generally accepted[5] that the first microreversible step along the catalytic pathway involves formation of a *recognition* or *Michaelis* complex. Next, it is postulated that an H bond between the hydroxyl group (Ser-195) $O_\gamma H$ and the (His-57) N_ϵ atom is necessary to proceed and mediate proton transfer from the (Ser-195) O_γ to (His-57) N_ϵ atom. How a relatively inert alcohol (Ser) can be activated to become a powerful oxy anion is a central conundrum in catalytic studies.

The proper identification of the active-site residue responsible for the optimal pH (ca. 7) and pH dependence of catalysis in these extensively studied enzymes[6-10] is closely related with the answer to the question whether or not further proton transfer can occur from the (His-57) N_δ atom to Asp-102 (charge relay system).[11-14] Elucidation of the catayltic role of Asp-102 has been the subject of site-directed mutagenesis studies[15,16] and theoretical investigations.[17,18] Closer investigation reveals that there is a virtually conserved catalytic *tetrad* (including Ser-214),[19,20] which then suggests a mechanistic role for the adjacent water channel connecting the active site with the remote (15 Å) surface of the enzyme[19] by means of an unbroken chain of H bonds. A recent paper describes the stereoelectronic, structural, and kinetic factors contributing to the reaction pathway.[21] Recent theoretical studies of Warshel et al.[14,22] and Weiner et al.[23] used molecular dynamics and quantum mechanical methods to explore the chemical and physical nature of selected steps of catalysis. In spite of the fact that theoretical calculations are usually based upon different simplified models, which sometime make the results hard to compare, it is a desirable characteristic of modern scientific efforts that experimental and theoretical studies be mutually complementary, each thus challenging the other to provide deeper insight into the complex processes of nature. Although much progress has been made, our understanding of catalysis in the serine proteinases is still far from complete.

We have chosen to focus our attention on initial and subsequent states of the active site of the serine proteinase porcine pancreatic elastase (PPE). Some 15 crystallographic analyses of native and complexed PPE have been reported[24] or are in progress in this

† On leave from University of Warsaw, Department of Biophysics, Warsaw, Poland.

(1) Chymotrypsinogen numbering system has been adopted. Abbreviations used: PPE, porcine pancreatic elastase (EC 2.4.21.11); nVal, norvaline; NMeLeu, *N*-methylleucine; MD, molecular dynamics; ESP, molecular electrostatic potential charges; rms, root-mean-square.

(2) Matthews, B. W.; Sigler, P. B.; Henderson, R.; Blow, D. M. *Nature* **1967**, *214*, 652.
(3) Blow, D. M.; Birktoft, J. J.; Hartley, B. S. *Nature* **1969**, *221*, 337-340.
(4) Henderson, R.; Wright, C. S.; Hess, G. P.; Blow, D. M. *Cold Spring Harbor Symp. Quant. Biol.* **1971**, *36*, 63-70.
(5) Walsh, Ch. *Enzymatic Reaction Mechanism*; Freeman, W. H. and Co.: San Francisco, CA, 1979; pp 94-97.
(6) Hunkapillar, M. W.; Smallcombe, S.; Whitaker, D.; Richards, J. H. *Biochemistry* **1973**, *12*, 4732-4743.
(7) Koeppe, R.; Stroud, R. M. *Biochemistry* **1976**, *15*, 3458-3464.
(8) Robbillard, G.; Shulman, R. *J. Mol. Biol.* **1972**, *71*, 507-511.
(9) Markley, J. L.; Ibanez, I. B. *Biochemistry* **1978**, *17*, 4627-4640.
(10) Bachovchin, W.; Roberts, J. D. *J. Am. Chem. Soc.* **1978**, *100*, 8041-8047.
(11) Dewar, M. J. S. *Enzymes* **1986**, *36*, 8-20.
(12) Kossiakoff, A. A.; Spencer, S. A. *Biochemistry* **1981**, *20*, 6462-6473.
(13) Kollman, P. A.; Hayes, D. M. *J. Am. Chem. Soc.* **1981**, *103*, 2955-2961.
(14) Warshel, A.; Russell, S. T. *J. Am. Chem. Soc.* **1986**, *108*, 6569-6579.
(15) Sprang, S.; Standing, T.; Flettrick, R. J.; Finer-Moore, J.; Stroud, R. M.; Xuong, N.-H.; Hamlin, R.; Rutter, W. J.; Craik, C. S. *Science* **1987**, *237*, 905-909.
(16) Craik, C. S.; Roczniak, S.; Largeman, C.; Rutter, W. J. *Science* **1987**, *237*, 909-913.
(17) Warshel, A.; Sussman, F. *Proc. Natl. Acad. Sci. U.S.A.* **1986**, *83*, 3806-3810.
(18) Lesyng, B.; Meyer, E. F., Jr. *J. Comput.-Aided Mol. Des.* **1987**, *1*, 211-217.
(19) Meyer, E. F., Jr.; Cole, G. M.; Radhakrishnan, R.; Epp, O. *Acta Crystallogr.* **1988**, *B44*, 26-38.
(20) The only known exceptions to this conserved sequence homology are found in two granzymes: D and E (Ala) and F(Thr): Jenne, D. E.; Masson, D.; Zimmer, M.; Haefliger, J.-A.; Li, W.-H.; Tschopp, J. *Biochemistry* **1989**, *28*, 7953-7961.
(21) Dutler, H.; Bizzozero, S. A. *Acc. Chem. Res.* **1989**, *22*, 322-327.
(22) Warshel, A.; Naray-Szabo, G.; Sussman, F.; Hwang, J.-K. *Biochemistry* **1989**, *28*, 3629-3637.
(23) Weiner, S. J.; Siebel, G. L.; Kollman, P. A. *Proc. Natl. Acad. Sci. U.S.A* **1986**, *83*, 649-653.
(24) Bode, W.; Meyer, E. F., Jr.; Powers, J. C. *Biochemistry* **1988**, *28*, 1951-1963.

0002-7863/90/1512-8925$02.50/0

laboratory. Careful comparisons of crystallographic data reveal that the *crucial H bond* (Ser-195···His-57) is frequently long and is reported with less than ideal angles, e.g., (Ser-195) C_β–O_γ–(His-57) N_ϵ of 87.7° for trypsin[25] or 85.1° for native PPE.[19] This angle ideally should be 109.5° in order to produce a linear O_γ–H···Nϵ H bond. Moreover, the Debye–Waller thermal (*B*) factor for O_γ (Ser-195) frequently is unusually high, thus indicating a high degree of mobility of the O_γ atom.[27] This mobility represents a finer mechanistic detail for subsequent theoretical calculations and may well provide fresh insight for the design of novel proteinase inhibitors. Better inhibitors are needed because of the role of these serine proteinase in degenerative diseases (pancreatitis, pulmonary emphysema, etc.).[28]

In our previous work[29] (part I), two different molecular dynamics simulations were performed to study whether there is any preference for formation of this catalytically crucial H bond in the native enzyme (PPE), i.e., in the absence of the substrate. During both simulations (total 150 ps) the H bond *was never formed*. In only two of the total 1500 snapshots of the trajectories could the formation of a significantly distorted (Φ = 44°; see below) H bond be observed.[29,30] The analysis showed that several water molecules form a well-defined dynamic network of H bonds inside the active site and are capable of preventing the formation of the (Ser-195) O_γ–H···N_ϵ (His-57) bond.

In our next study,[31] the first molecular dynamics simulation of the Michaelis complex of this enzyme was performed. A modified hexapeptide Thr-Pro-nVal-Leu-Tyr-Thr (I) was chosen as a model substrate because the crystallographic structure of an inhibitor complex (Thr-Pro-nVal-NMeLeu-Tyr-Thr, II) with PPE was known to 1.8-Å resolution.[32] It was found[32] that the *N*-methyl group responsible for the inhibitory properties of the latter compound is sterically in close proximity to active-site atoms; the currently understood geometric constraints in the active site are extremely tight.[33] Thus, the crystallographic geometry of this inhibitor should closely mimic the postulated geometry of a Michaelis complex containing the scissile substrate (I), which is hard to obtain crystallographically. For the sake of these calculations the *N*-methyl group of P'_1 (II) Leu[34] was replaced by an H atom in the simulation (I), making it identical with a normal, peptidic substrate.

Binding of substrate to the enzyme displaces the competitive water molecules from the active site; this was incorporated into our starting model. Hence, it could be assumed that the crucial H bond should be well formed; a "text-book" view would assume that it would be always formed. This first simulation showed,[31] however, that in close proximity of the His-57 acceptor, the carbonyl O atom of Ser-214 has a negative partial charge, which produces an attractive force on the (Ser-195) H(O_γ). This Coulombic attraction competes for formation of an H bond with His-57. Hence, the frequency of formation of this H bond critically depends on the proper calculation of the electrostatic interaction of the (Ser-195) O_γH group with His-57 as compared to its interaction with Ser-214. The stronger the former, the higher the frequency. The simulation,[31] in which the standard set of the electric net charges was employed, revealed that the (Ser-195) O_γH group still has a significant preference for a nonproductive interaction with the O=C carbonyl (Ser-214) rather than with the (His-57) N_ϵ atom.

Table I. Modified Net Atomic Charges (e) for N_δ-Protonated Histidine 57 Added to AMBER Library[a]

CA	CB	CG	CD	NE	CE	ND	HND
0.148	0.032	0.176	0.140	−0.610	0.344	−0.240	0.256
0.219	0.060	0.089	0.145	−0.527	0.384	−0.444	0.320

[a] Second line gives the standard values in the AMBER library. Net charges on atoms N, H, C, and O are the same as in the standard library.

According to the conclusions drawn in that paper we now present the next MD simulation of the same system, i.e., PPE with the hexapeptide Thr-Pro-nVal-Leu-Tyr-Thr (I), performed with an improved charge distribution for the side-chain atoms of His-57. The modification results from the inductive effect of Asp-102, a member of the catalytic tetrad, which forms a strong H bond with the (His-57) N_δH group.

Methods

The trajectory of the motion of the system was calculated by using the AMBER (version 3.0)[35] program package.

Starting Geometry. The 1.80-Å resolution X-ray structure of PPE complexed with a hexapeptide (II) was used as the starting geometry of the system (2275 atoms of PPE and 50 atoms of the substrate). As mentioned above, the *N*-methyl group of Leu in the hexapeptide was replaced by an H atom. All His residues were uncharged; the specific protonation form (N_δ or N_ϵ) was chosen in each case on the basis of favorable H-bonding arrangements found in the crystallographic structure. This and other parameters approximate conditions for pH 7. Hence His-210 was protonated at N_δ while N_ϵ protonation was chosen for His-40, −71, −91, and −200. To remain consistent with the accepted catalytic mechanism,[10,12,36] His-57 was protonated at N_δ, forming a strong H bond with the carboxylate group of Asp-102.

In order to include crystallographic water of hydration in the system, each water molecule either had to be buried inside the system or had to form at least two H bonds with the enzyme if it was located on the enzyme surface. This procedure led to retaining 54 molecules of crystallographic water in the system. For additional hydration, a box of Monte Carlo water, included in the AMBER database, was used. At first, a dome of water was created, centered on the (His-57) N_ϵ atom, containing those water molecules whose O atom positions were within 22.5 Å of the center. The closest distance between an O atom of water and the nearest heavy atom of the enzyme (or substrate) was 2.4 Å. This dome of water consisted of 736 molecules of water. By use of the same box of water, a second shell of water around the whole enzyme was also created. It consisted of those molecules of water for which O atom positions were closer than 7.0 Å to the nearest heavy atom of the enzyme and further than 2.4 Å. This shell of water consisted of 1325 water molecules. In total, 2115 water molecules were included. To prevent possible gradual expansion of the solvent into adjacent vacuum, O atoms of water molecules located in the outer 4.0-Å shell of water were held translationally fixed during the whole simulation.

Force Field Parameters. The AMBER united-atom potential model was used with a residue based cutoff distance of 8 Å for nonbonded interactions and unit dielectric constant. The mass of H was equal to 1 amu. The SPC water model was used in all calculations.[37] A partial charge set for norvaline was calculated previously.[31] An improved partial charge set for His-57 was calculated according to the AMBER procedure[38] with one modification. The ESP charges for the imidazole ring of His-57 were obtained from calculations performed for a complex of Asp-102 and His-57. The geometry of the complex was taken from the minimized structure of the enzyme calculated during the previous simulation.[31] To

(25) Marquart, M.; Walter, J.; Deisenhofer, J.; Bode, W.; Huber, R. *Acta Crystallogr.* **1983**, *B39*, 480–490.
(26) Hydrogen atom positions are not determined by macromolecular X-ray diffraction and therefore must be approximated geometrically from an analysis of neighboring atom positions.
(27) For example, the B for O_γ is ca. 22 Å^2 in PPE compared to neighboring side-chain atoms, which have relatively low *B* factor, 4–6 Å^2 (ref 19).
(28) Neurath, H. *J. Cell Biochem.* **1986**, *32*, 35–49.
(29) Geller, M.; Carlson-Golab, G.; Lesyng, B.; Swanson, S. M.; Meyer, E. F., Jr. *Biopolymers*, in press.
(30) Swanson, S. M.; Wesolowski, T.; Geller, M.; Meyer, E. F., Jr. *J. Mol. Graphics* **1989**, *7*, 240–242.
(31) Geller, M.; Swanson, S. M.; Meyer, E. F., Jr. *J. Biomol. Struct. Dyn.* **1990**, *7*, 1043–1052.
(32) Meyer, E. F., Jr.; Clore, G. M.; Gronenborn, A. M.; Hansen, H. A. S. *Biochemistry* **1988**, *27*, 725–730.
(33) Based on original observations by W. Bode (private communication), we find the current limit of displacement available to peptidic ligands to be of the order of 1.5 Å: Takahashi, L. H.; Radhakrishnan, R.; Rosenfield, R. E., Jr.; Meyer, E. F., Jr. *Biochemistry* **1989**, *28*, 7610–7617 (cf. Figure 3).
(34) Notation of: Schechter, I.; Berger, A. *Biochem. Biophys. Res. Commun.* **1967**, *27*, 157–162.
(35) *AMBER: Assisted Model Building With Energy Refinement*, Version 2. 1985; also Case, D.; Siebel, G.; Bush, P. Version 3.0, 1986, with update 1987, Department of Pharmaceutical Chemistry, University of California, San Francisco, CA.
(36) Bizzozero, S. A.; Dutler, H. *Bioorg. Chem.* **1981**, *10*, 46–62.
(37) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Hermans, J. In *Intermolecular Forces*; Pullman, B., Ed.; Reidel: Dordrecht, Holland, 1981; p 381.
(38) Weiner, S. J.; Kollman, P. A.; Case, D. A.; Singh, U. Ch.; Ghio, C.; Alagona, G.; Profeta, S., Jr.; Weiner, P. *J. Am. Chem. Soc.* **1984**, *106*, 764–784.

Table II. Cumulative rms Positional Fluctuations (in Angstroms) of Protein Atoms at Different Stages of Molecular Dynamics Simulation of a Complex of Porcine Pancreatic Elastase and Hexapeptide (I)[a]

		time						
no.[b]	subsystem	0–10 (0–10)	0–20 (11–20)	0–30 (21–30)	0–40 (31–40)	0–50 (41–50)	0–60 (51–60)	0–70 (61–70)
I	total	0.49	0.57	0.63	0.63	0.63	0.63	0.62
		(0.49)	(0.52)	(0.35)	(0.40)	(0.39)	(0.38)	(0.36)
	backbone	0.39	0.45	0.50	0.50	0.50	0.50	0.49
	side chain	0.56	0.65	0.71	0.72	0.71	0.71	0.70
II	total	0.62	0.72	0.81	0.83	0.84	0.83	0.82
		(0.62)	(0.59)	(0.59)	(0.61)	(0.61)	(0.60)	(0.58)
	backbone	0.44	0.56	0.64	0.65	0.64	0.64	0.63
	side chain	0.72	0.80	0.90	0.94	0.90	0.89	

[a] After least-squares fit of all non-hydrogen protein atoms relative to the average structure in each cumulative period of simulation (in Å); time in picoseconds. The numbers in parentheses are the corresponding subinterval values of rms for each 10-ps period of simulation. [b] I, enzyme; II, hexapeptide.

be consistent with the AMBER parametrization methods, the standard values of charges of backbone NH and C=O groups were unchanged (their sum equal to −0.246e). The bridge atoms (C_α, C_β) absorb the remaining charge distribution from the imidazole ring and the backbone segment to achieve neutrality of this residue.[38] This new set and, for comparison, the standard charges of histidine (also protonated on the N_δ atom; type HID[38]) are listed in Table I. As was expected, the negatively charged side chain of Asp-102, located near the $N_\delta H$ group of the imidazole ring, slightly shifts the electron density of the ring toward the N_ϵ and C_α atoms. For example, the electron density around the N_ϵ nucleus increases from −0.527e to −0.610e. For comparison's sake the same type of calculation has been performed for the pair Asp-102 and Ser-214 to check the possible influence of the former on the carbonyl group of the latter (which was important in further analysis). In this case, however, the corresponding changes were much smaller.

Minimization and Equilibration. After addition of H atoms (not included in the crystallographic analysis[26]), energy minimization was performed in three steps. First, the energy component from bulk solvent was minimized while the protein was held fixed in space. Second, the protein (with crystallographic water) was minimized while the nuclei of the crucial (His-57) N_ϵ, (Ser-195) O_γ, and C_β were held fixed in space. Finally, the energy of the entire system was minimized. Approximately 3000 minimization cycles were required, leading to the total energy of −29 970 kcal/mol and rms gradient of 0.1 kcal/(mol Å). Prior to MD data collection, the whole system was thermalized and equilibrated by the following procedure: with the protein held fixed, dynamical thermalization of the bulk water was performed at 285 K in the microcanonical ensemble (30 times for 0.2 ps; total time 6 ps), using a new set of random velocities each time. Then, the same thermalization was applied to the protein only; it was performed in four steps at 50, 100, 200, and 285 K (15 times for 0.2 ps at each temperature; total time 3 ps for each temperature), keeping the atoms (Ser-195) O_γ, (His-57) N_ϵ, C (nVal-3, substrate), and N (Leu-4, substrate) fixed in space. Next, the constraints applied to the above-mentioned atoms were removed, and with the velocities from the last step, an additional microcanonical simulation performed for 3 ps for the whole simulation showed that the temperature of the system was stable (T = 300 ± 3 K). After thermalization, the system was equilibrated by coupling it to a thermal bath (T = 300 K; tolerance of 5 K) with a temperature relaxation time t = 0.25 ps. Equilibration was started by using the restart file from the previous stage and was performed for 20 ps. The energy of the system was stable after 10 ps. Data were collecting during the subsequent 70-ps period at steps of 0.02 ps. During the whole simulation all bond lengths were constrained by use of the SHAKE algorithm,[39] with a small tolerance (0.0001 Å). A time step of 0.001 ps (1 fs) was used for integration of the equation of motion. After each stage, the structure of the system was checked by using graphics. During the whole thermalization and equilibration process 20 additional O atoms of bulk water (which have a tendency to move through the outer layer of the fixed water) were fixed in space.[40] No additional water molecules were fixed during the data collection period.

Graphics Utilities. Without computer graphics, it is virtually impossible to comprehend individual and composite interactions in such a study. By means of program MDKINO[30] it was possible to view in stereo on the Evans and Sutherland graphics system, like a movie, a sequence up to 3500 snapshots of the active site (~110 atoms) for the simulation. Careful visual examination led to the identification of specific simulation frames, which provided a starting point for subsequent statistical analyses.

H-Bond Criterion. During statistical analysis, a previously described procedure[29] was used to check the formation of a H bond between a donor group, D–H, and acceptor atom, A. Two conditions have to be satisfied for H bonding to occur. First, the distance between donor and acceptor atom should be shorter than R_{max}. The second condition depends on two angles. The first (α) is an angle between the lone-pair vector on the acceptor and the vector AD. The second (β) is an angle H–D–A. Then the sum (Φ) of absolute values of both angles, α and β, is calculated. The second condition requires that the value of Φ should be smaller than Φ_{max}. As in previous cases,[29,31] the R_{max} value of 3.3 Å and Φ_{max} equal to 45° were adopted for this calculation. We are conscious that this criterion, as well as any others, is somewhat arbitrary because it is impossible to differentiate precisely between H-bond formation and electrostatic attraction. Nevertheless, it is useful because it takes into account not only the distance between both heavy atoms (O–N) but also the relative orientation of a H–D(onor) group and the lone electron pair of an acceptor. Henceforth, existence of a H bond will be accepted only if this criterion is fulfilled. This criterion is stronger than, for example, that used by Koehler et al.,[41] where only a distance condition and one angle condition D–H–A > 135° were applied.

MD simulations were performed on the Cray X-MP and Y-MP at the Pittsburgh Supercomputer Center. Other calculations were performed on a VAX 11/750, and visual analysis was performed on Evans and Sutherland PS 330 display at our laboratory using the program FRODO (version 6.6). Stereo drawings were obtained with the program BALLSTICK.[42]

Results and Discussion

The temperature and energy of the system were quite stable during the whole simulation. The mean value and standard deviation of the temperature of the system was T = 300 ± 0.2 K. The mean of the total energy of the system was E_{tot} = −26 944 ± 63 kcal/mol.

Root-Mean-Square Analysis. An analysis of positional fluctuations is listed in Table II. Each column presents the fluctuation of coordinates relative to the mean conformation for each cumulative time span. After 30 ps the cumulative rms values are stable. Also, the rms values for each 10-ps time period (numbers in parentheses) are similar to each other. These, together with the stable value of the total energy, confirm that the system is in a state of equilibrium around which conformation space was sampled during the simulation. The larger rms fluctuations for the hexapeptide (I) are consistent with the experimental observations of higher thermal B factors and a K_M value in the millimolar range, suggesting that this substrate is weakly bound to the enzyme.[20] The rms analysis for chosen segments of the system is listed in Table III. Table III contains the pairwise rms differences between the X-ray, the minimized, and the time-averaged MD structures for the entire 70-ps simulation (mean MD). The values are typical for such a large hydrated system.[43] It is worth

(39) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327–341.

(40) To prevent this diffusion one should have the outer layer of the thickness of 5–6 Å. This would double the number of water molecules as well as time of the calculations. All 20 molecules were located further than 20 Å from the active site.

(41) Koehler, J. E. H.; Saenger, W.; van Gunsteren, W. F. *Eur. Biophys. J.* **1987**, *15*, 197–210.

(42) Radhakrishnan, R., unpublished.

Table III. rms Deviations in Molecular Dynamics Simulations of a Michaelis Complex of Porcine Pancreatic Elastase and Hexapeptide I[a]

subsystem	X-ray vs minimized	minimized vs mean MD	X-ray vs mean MD
enzyme			
total	0.62	1.00	1.20
backbone	0.56	0.78	0.88
side chains	0.71	1.40	1.41
Ser-195 O_γ	0.50	0.68	0.31
(side chain)	0.53	0.58	0.27
His-57 N_ϵ	0.79	0.70	0.26
(side chain)	0.65	0.55	0.42
Asp-102 C_γ	0.74	0.45	0.30
(side chain)	0.82	0.78	0.31
Ser-214 O_γ	0.50	0.61	0.38
(side chain)	0.54	0.66	0.28
hexapeptide			
total	0.80	1.32	1.02
backbone	0,56	1.04	0.82
side chains	0.93	1.41	1.10

[a] rms differences of atomic positions after a least-squares fit of all non-hydrogen protein atoms (in Å).

Table IV. Distances (Angstroms) between Chosen Atoms in Various Conformations of the Active Site of a Michaelis Complex; PPE + Hexapeptide I[a]

atoms				distances			
				X-ray	Min	no. 3006[b]	no. 2428[b]
Ser-195	O_γ	N_ϵ	His-57	3.10	2.91	2.58	3.24
Ser-195	O_γ	O	Ser-214	3.10	2.88	4.38	2.81
Ser-195	O_γ	O	Thr-213	3.60	3.52	4.05	3.44
Ser-195	O_γ	C	nVal-3	2.93	2.89	2.78	3.14
His-57	N_δ	$O_{\delta 1}$	Asp-102	2.54	3.17	3.12	2.81
His-57	N_δ	$O_{\delta 2}$	Asp-102	3.41	2.71	2.72	3.18
His-57	N	$O_{\delta 2}$	Asp-102	2.70	2.83	2.94	2.92
Ser-214	O_ϵ	$O_{\delta 1}$	Asp-102	2.84	2.70	2.66	2.59
nVal-3	N	O_γ	Ser-195	2.86	3.11	2.95	2.92
nVal-3	N	O_γ	Ser-214	3.35	3.18	4.27	3.69
nVal-3	O	N	Gly-193	2.64	2.76	2.79	2.85
nVal-3	O	N	Ser-195	3.04	3.23	3.16	3.02

[a] O denotes an oxygen atom of a carbonyl group, N denotes a nitrogen atom of the backbone. [b] Number of a snapshot chosen from 3500 sampled conformations in the MD simulation.

noting the rms values (X-ray vs mean MD) for the crucial residues in the active site are significantly lower than those for the whole enzyme. This result suggests that in spite of the conformational dynamics of the whole enzyme, the active site is more resistant to conformational changes.

Conformation Analysis. To facilitate further discussion, Figure 1 presents four stereoviews (Figure 1b (minimized) in the supplementary material) of the active site of porcine pancreatic elastase with bound hexapeptide (four middle residues (2–5) of the latter are shown). The crystallographic and minimized structures as well as conformations of two chosen MD snapshots are shown. Distances between selected atoms for all these conformations are listed in Table IV. The general features of these conformations are similar; the interaction regions of His-57, Asp-102, and Ser-214, the location of the carbonyl O (nVal-3) atom in the oxy anion hole, as well as the overall shape of the substrate (residues 2–5). There are, however, some important dissimilarities connected mainly with the manner of interaction of the (Ser-195) $O_\gamma H$ group. Examination of the conformations shows the presence of four possible donor/acceptor atoms (for H-bond formation) that may strongly interact with this group electrostatically. These are as follows: the (His-57) N_ϵ atom, two adjacent carbonyl groups (Ser-214) C=O and (Thr-213) C=O (not shown on Figure 1), and the (nVal-3) NH backbone group. The enzymatically important H-bond acceptor for the HO_γ proton is, of course, (His-57) N_ϵ. However, the two negatively charged carbonyl O atoms may, as it will be shown later, disturb the formation of this crucial H bond.

Table V. Identification of Selected Hydrogen Bonds in the Active-Site Region of Porcine Pancreatic Elastase with Bound Hexapeptide[a]

no.	donor		acceptor	
	residue	atom	residue	atom
1	Ser-195	O_γ	His-57	N_ϵ
2	nVal-3	N	Ser-195	O_γ
3	nVal-3	N	Ser-214	O
4	Ser-195	O_γ	Ser-214	O
5	His-57	N_δ	Asp-102	$O_{\delta 1}/O_{\delta 2}$
6	His-57	N	Asp-102	$O_{\delta 2}$
7	Ser-214	O_γ	Asp-102	$O_{\delta 1}$
8	Gly-193	N	nVal-3	O
9	Ser-195	N	nVal-3	O
10	Gln-192	N_ϵ	Pro-2	O

[a] The time evolution of these bonds is shown in Figure 4. Bonds 2, 3, and 8–10 are between the protein and the substrate.

Examination of the crystallographic and minimized structures (Figure 1a,b and Table IV) shows that in both structures the H bond between the (His-57) N_ϵ atom and the (Ser-195) HO_γ group is formed, with more distortion found in the crystal structure. Residues 2–5 of the substrate (I) may form three H bonds with the enzyme in the active-site region. The first is an H bond between the O=C group of nVal-3 and the HN backbone group of Gly-193 or Ser-195 in the oxy anion hole. The second forms between the carbonyl O atom of Pro-2 and the amino group of Gln-192. The third involves the NH backbone group of nVal-3 and either the O=C oxygen (Ser-214) or the (Ser-195) O_γ atom. In the minimized structure the NH group of nVal-3 forms a weak H bond with the O=C group of Ser-214 (3.2 Å; Table IV).

The first chosen MD snapshot (3006; Figure 1c) corresponds to the case when the crucial H bond between the (Ser-195) $O_\gamma H$ group and the (His-57) N_ϵ acceptor is well formed. Moreover, in this conformation, and in six others, the distance between the donor and the acceptor atoms is *less than 2.6* Å (in this case, 2.58 Å) with $\Phi = 21°$. From the data presented in Figure 2a (supplementary material), it is seen that the (Ser-195) O_γ donor atom is located near the (His-57) N_ϵ acceptor in ~75% of the snapshots. The average distance for those conformations is 2.85 Å, which is favorable for H-bond formation. However, as will be discussed below, values of Φ for many of these cases are larger than 45°, which indicates significant distortion of this bond (see Figure 3). The rest (~25%) of the sampled conformations correspond to two families. The first one, much more prevalent, is presented in Figure 1d. In this case, the (Ser-195) $O_\gamma H$ group is directed toward the (Ser-214) O=C group. The distance between the O_γ atom and the (Ser-214) O(=C) atoms, 2.81 Å, is favorable for H-bond formation. The Φ value (56°) is always higher than the criterion threshold (45°) value, indicating electrostatic attraction rather than H-bond formation. The second-type of conformation, which occurs rather rarely, corresponds to the cases when the $O_\gamma H$ group is pointing toward the substrate (the low value of distances presented in broken lines in Figure 2b (supplementary material).

H-Bond Network. Visual analysis of the collected snapshots, using the program MDKINO[30] made it possible to observe dynamic behavior of the active site as well as H-bond networks in the extended active-site region. The most interesting donor and acceptor groups, identified in Table V, were selected for further analysis. The dynamic behavior of chosen H bonds is presented in Figure 4. Data presented in Figure 4 exhibit the existence of a specific H-bond network in the active site. Many of these H bonds are rather stable during the entire simulation. This is true especially for the residues Asp-102, His-57, and Ser-214. The $CO_{\delta 1}$ and $CO_{\delta 2}$ groups of Asp-102 are acceptors; (His-57) $N_\delta H$, (Ser-214) $O_\gamma H$ and (His-57) NH are donors. Some of the H bonds are formed temporarily. These include the following: (Ser-195) O_γ—H···N_ϵ (His-57), (nVal-3) N—H···O_γ (Ser-195), (nVal-3) C=O···H—N (Ser-195 or Gly 193) in the oxy anion

(43) van Gunsteren, W. F.; Berendsen, H. J. C. In *Molecular Dynamics and Protein Structure*; Hermans, J., Ed.; Polycrystal Book Service: Dayton, OH, 1985; pp 13, 91.

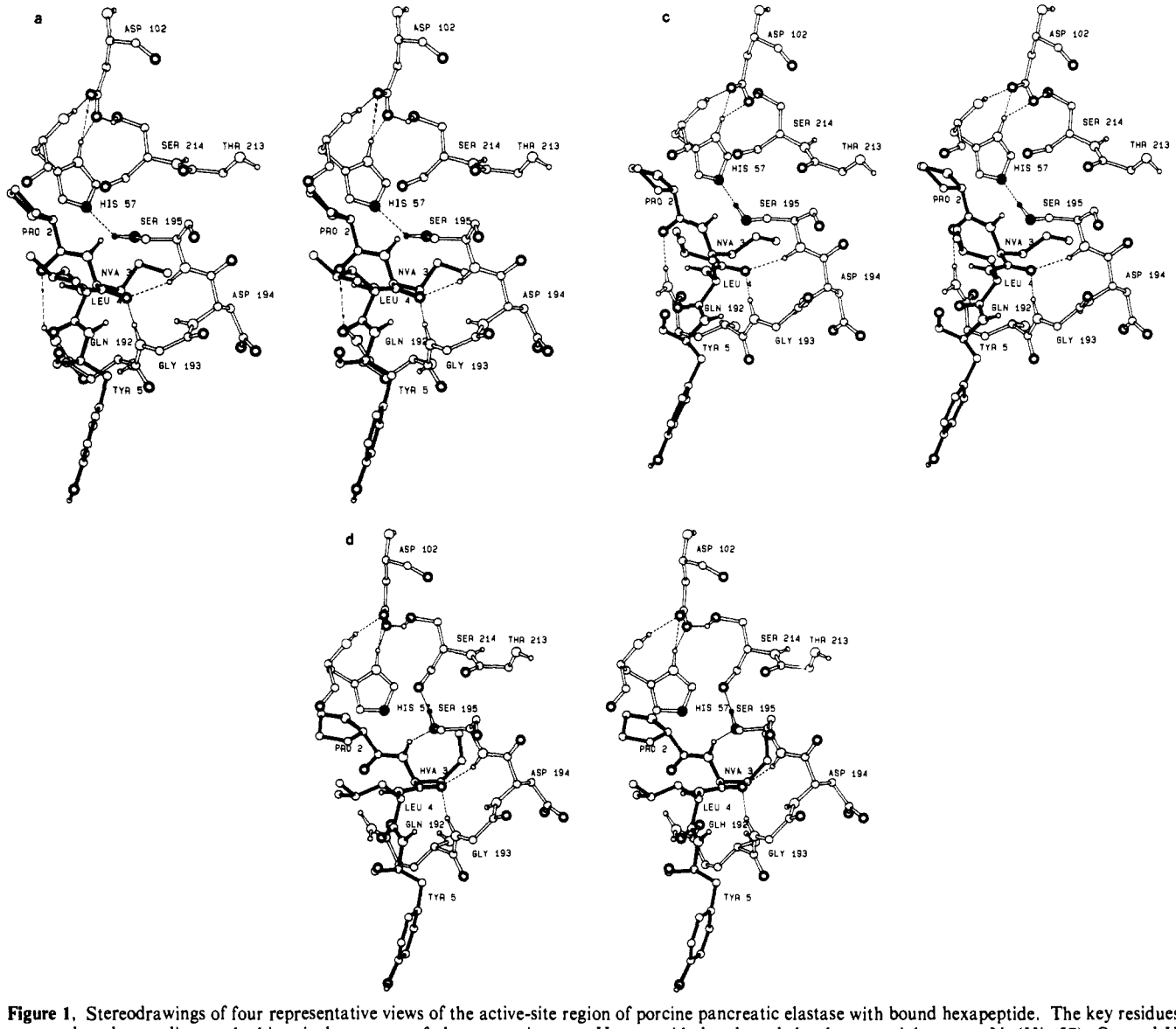


Figure 1. Stereodrawings of four representative views of the active-site region of porcine pancreatic elastase with bound hexapeptide. The key residues are numbered according to the historical sequence of chymotrypsinogen. Hexapeptide bonds and the three crucial atoms: N_ϵ (His-57), O_γ, and H (Ser-195) are drawn totally filled. (a) Crystallographic; (b) minimized (in supplementary material); (c) snapshot 3006, an example of a conformation with a very close approach (r = 2.58 Å) between donor and acceptor atoms in the crucial H bond between (Ser-195) $O_\gamma H$ and N_ϵ (His-57); (d) snapshot 2428, an example of a conformation in which the crucial H bond is not formed: instead, there is an electrostatic attraction between the donor group and the carbonyl oxygen of Ser-214 and also a new H bond between the (nVal-3; substrate) NH group and the O_γ atom (Ser-195).

hole, (Gln 192 (N_ϵ—H1 or N_ϵ—H2···O=C (Pro-2; substrate). The first row in Figure 4 shows the dynamic properties of the crucial H bond (1) between (Ser-195) $O_\gamma H$ and (His-57) N_ϵ. It is now ~10 times more readily formed than in our previous simulation[31] in which the standard, default set of charges was used for His-57. The mode of formation of the H bond (2) between the NH group of nVal-3 (substrate) and the (Ser-195) O_γ atom, presented in the second row, points to the possible role of the substrate in restoration of the H bond between Ser-195 and His-57 (in addition to exclusion of competing, active-site water molecules). Rotation of the (Ser-195) $O_\gamma H$ group toward the carbonyl group of Ser-214 leads to the breaking of this H bond. However, the rotation simultaneously changes orientation of the two lone electron pairs located on the O_γ nucleus favorably for H-bond (3) formation with the NH group of nVal-3. This interaction, being an attraction of the O_γ nucleus, leads to the restoration of the previous H bond (2) (compare the pairwise complementarity of bars in the first two rows in Figure 4). An empty fourth row shows that the above-mentioned complementarity between interaction the $O_\gamma H$ group and the O=C group of Ser-214 is due to electrostatic attraction rather than to H-bond (4) formation. Data presented in rows 5–7 demonstrate that the catalytic tetrad H-bonding network between His-57, Asp-102, and Ser-214 is rather stable in time. The next two rows (8, 9) show the evolution of the two H bonds that help to stabilize the carbonyl group of nVal-3 in the oxy anion hole. It is seen that when one of them is well formed, the other is usually broken, and vice versa. For comparison, Figure 4b (supplementary material) shows the same analysis when the angular threshold for the H bond was changed to $\Phi = 55°$. It is seen that many more H bonds are now formed, especially for cases 1, 2, and 7 and 8 (oxy anion hole); this exhibits the existence of dynamic distortions of these bonds.

The appearance of the short H bonds between Ser-195 and His-57 during the simulation was the subject of further study to determine whether such a short contact could be an artifact resulting from the force field potential used in the simulation.[38] Energy analysis performed for snapshot 3006 and the minimized structure showed that the AMBER energy of interaction between Ser-195 and His-57 for the former is higher than the latter by ~4.2 kcal/mol. By use of the ab initio Gaussian 86 program, the energy of a model system that should mimic the interaction between both residues was calculated. This system consisted of an imidazole ring with methylated C_γ atom (model for the side chain of His) and methyl alcohol (model for the side chain of Ser). All energy calculations and optimizations were performed at the HF/6-31G* level. First, the geometry of both subunits was

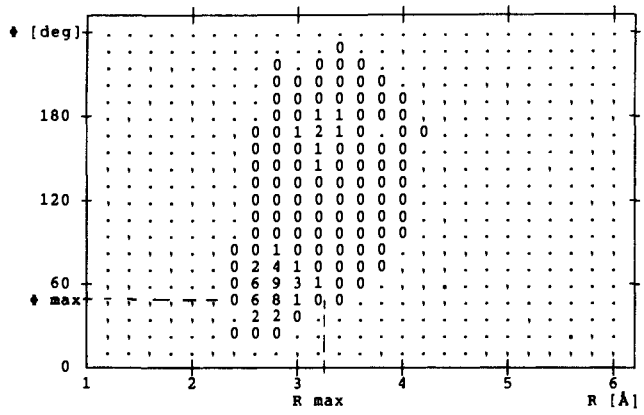


Figure 3. Frequency of occurrence of the crucial H bond (His-57 N_ϵ ···HO_γ Ser-195) determined by two parameters R and Φ during the molecular dynamics simulation of porcine pancreatic elastase with bound hexapeptide. Frequencies are scaled to be inside the interval (0, 10) and truncated to the integer value 0 ... 9. • indicates an exact zero. The hydrogen-bond region is located in the left lower corner. For explanation of the meaning of the parameters R and Φ, see Methods.

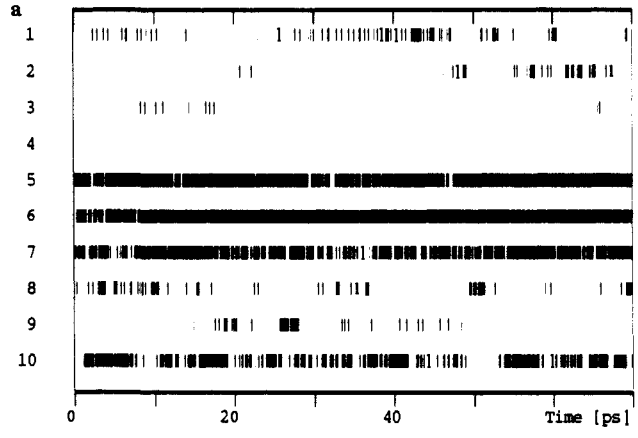


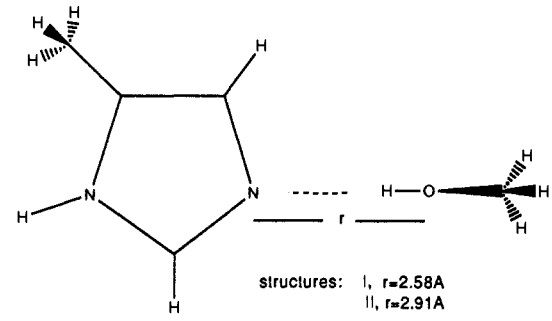
Figure 4. Time evolution of a hydrogen-bonding network in the active site of a Michaelis complex: PPE + hexapeptide. Each small bar denotes the occurrence of the hydrogen bond in a given snapshot. Molecular dynamics simulation, time = 70 ps. Numbers on the left side denote the hydrogen bond (see Table V). (a) $\Phi = 45°$, (b) $\Phi = 55°$ (supplementary material).

optimized (planar imidazole ring). The starting conformation of the whole system was taken from the simulation (snapshot 3006). Then the relative orientation was optimized by assuming that the OH group of the alcohol is coplanar with the imidazole ring and keeping the distance (r) between the O atom (methanol) and the N_ϵ (imidazole) fixed and equal to the dynamically observed distance, 2.58 Å. The relative orientation of the optimized system was essentially unchanged in comparison with the simulation data. The change of the angle between the O atom of methanol, the N_γ, and the C_δ was less than 2°. Final optimization of the system, keeping the relative orientation fixed, led to structure I (Chart I). Next, with the relative angular orientation fixed, the relative distance was optimized, leading to $r = 2.91$ Å. With this orientation fixed, the whole system was once again optimized, leading to the structure II. The energy difference between system I and system II is 4.1 kcal/mol. Although this is only a rough estimation of the difference of the interaction energy between Ser-195 and His-57 for the two above-mentioned conformations, it suggests that such a short H bond may be formed under the conditions of the simulation, which plausibly can be related to a natural conformational state of the active site.

Conclusions

There is a specific H-bond network in the active-site region of PPE; it strongly resembles one found in the native case studied in our previous work.[29] The residue involved are Asp-102, His-57, and Ser-214; both the $CO_\delta 1$ and $CO_\delta 2$ atoms of ionized Asp-102

Chart I



are acceptors, while (His-57) $N_\delta H$, (Ser-214) $O_\gamma H$, and (His-57) NH are donors. These H bonds are quite stable during the entire simulation. It was previously shown that replacement of Asp-102 by Asn leads to a significantly different pattern of H bonds in this region and destabilizes conformations of the imidazole ring of His-57.[18] Hence, one of the possible roles of this stable network of H bonds is to maintain the proper orientation of the imidazole ring relative to the Ser-195. Location of the carbonyl group of the substrate (nVal-3) in the oxy anion hole is also stable in time. A Coulombic attraction between this group and NH backbone groups of Ser-195 and Gly 193 leads frequently to H-bond formation (oxy anion hole), usually with only one of the latter (Figure 4).

Improvement of the description of the electrostatic interaction in the active site, presented in this study, takes into account the inductive effect of the Asp-102 carboxylic acid group on the electron charge distribution of His-57. Including the new set of net charges for His-57 results in significant changes in the conformation of the (Ser-195) $O_\gamma H$ group. Now, in ~75% of the 70-ps MD simulation, the group is located in the proximity of the (His-57) N_ϵ atom, which is necessary for initializing the first step of the enzymatic reaction. Previously, only 5% of the collected snapshots exhibited this type of conformation. The result of this improvement has two major implications. First, it points out that one possible role of Asp-102 is to facilitate, via an inductive effect, formation of the catalytically crucial H bond between (His-57) N_ϵ and the (Ser-195) $O_\gamma H$ group. Second, as a cautionary note for "black-box" MD simulations, it points out that a more precise description of the electrostatic interaction is sometimes necessary for a proper description of events. Usually, the potential parameters employed in MD simulations calculate the energy of the electrostatic interactions on the basis of net charges obtained for each of the components (residues in the case of proteins) independently. It is perhaps noteworthy here that our conclusions on nonoccurrence of this H bond in the native case should be still valid. This is due to the fact that the changes involve only the side chain of His-57. Hence, the forces between the $O_\gamma H$ group of Ser-195 or H_2O molecules and the side chain would be changed proportionally if this new set of charges were used. The result also indicates the unexpectedly important role of the backbone NH group of nVal-3 (substrate), which is located in the proximity of the (Ser-195) $O_\gamma H$ group. When the group points toward the O=C group of Ser-214, the two lone electron pairs of the O_γ atom are directed to the NH group, resulting in the electrostatic attraction or even formation of an H bond between them. This interaction helps to restore the crucial H bond between Ser-195 and His-57 (compare Figure 1 c,d and Figure 4). Moreover, the location of the NH group of nVal-3 might have significant influence upon proton transfer along the H bond.

One important point that needs further analysis is the occurrence of short H bonds between (Ser-195) O_γ and (His-57) N_ϵ. In six different steps during the simulation, *this distance was shorter than 2.6* Å. The unfavorable increase in the interaction energy from both residues was compensated by dynamic interactions with the rest of the system. This suggests that occurrences of these short contacts have an important biochemical consequence. On the basis of ab initio calculations, performed for small H-bond systems,[44] we can expect that the barrier for proton transfer in

such a case should be lower than the barrier for distances that are closer to equilibrium (2.8-2.9 Å).

A persistent question, the mode of activation of Ser to become a powerful nucleophile, thus may be viewed in a new light as a result of these simulations. The effect of protein-mediated close contact of donor and acceptor may be one of the factors that increases the speed and efficiency of this enzymatic reaction. Quantum chemical methods should be employed to elucidate the problem. These calculations are now in progress in this laboratory.

Acknowledgment. We thank Mgr Antoni Łączkowski for help in displaying some of the statistical data. Financial support has been provided by the National Science Foundation (DMB - 8517286), Robert A. Welch Foundation (A-328) and the Texas Agriculture Experiment Station. Computational resources were provided by the Pittsburgh Supercomputing Center and Associate Provost for Computing, Dr. John Dinkel. M.G. wishes to acknowledge travel support from the Ministry of Science and Higher Education (Poland) within the project C.P.B.P. 01.06.

(44) Lesyng, B., private communication.

Supplementary Material Available: Minimized sterodrawing of the active-site region of PPE (Figure 1b), time evolution of selected distances between the side chain of Ser-195 and His-214 (Figure 2a) and Ser-214 and nVal-3 (Figure 2b) in MD simulation of a Michaelis complex, and time evolution of a H-bonding network in the active site of a Michaelis with $\Phi = 55°$ (Figure 4b) (4 pages). Ordering information is given on any current masthead page.

Sterically Encumbered Functional Groups: An Investigation of Endo versus Exo Phosphoryl Complexation Using ^1H and ^{31}P NMR

Bernard P. Friedrichsen, Douglas R. Powell,† and Howard W. Whitlock*

Contribution from the Samuel M. McElvain Laboratories of Organic Chemistry, Chemistry Department, University of Wisconsin—Madison, Madison, Wisconsin 53706.
Received May 21, 1990. Revised Manuscript Received July 23, 1990

Abstract: The synthesis of phosphine oxide bifunctional macrocycles **1-4** is reported. Additionally, the X-ray crystal structures for exo-exo diyne **1** and endo-exo hosts **2** and **4** are presented. Assignment of the two phosphorus signals in the ^{31}P NMR spectra of **2** and **4** and the aromatic proton signals in the ^1H NMR spectra of **2** and **4** is reported. The complexation behavior of macrocycles **1-4** and precyclophane **8** with a variety of neutral organic guests and Ph_2SnCl_2 is investigated by using ^1H and ^{31}P NMR as investigative instrumental probes. Initial endo complexation is the preferred mechanism in the 1:2 complexation of **2** with guests, while initial exo complexation is preferred for the complexation of **4** with guests. **2** forms 1:2 complexes with pentafluorophenol, 2,6-dimethyl-4-nitrophenol, and acetic acid via initial exo complexation. Association constants determined from these experiments reveal that the exo phosphoryl binding site in **4** is higher than those in the other reported phosphine oxides. An X-ray crystal structure of the 1:1 complex of **4** with diphenyltin dichloride was obtained to explore this anomaly, and it is reported.

Introduction

This paper describes the synthesis and complexation behavior of phosphine oxide bifunctional macrocycles **1-4** (Figure 1).[1] The design of host molecules capable of binding neutral organic guests is an area of rapidly expanding interest,[2] Cram,[3] Lehn,[4] Vögtle,[5] Diederich,[6] and others have made significant advances in the field of host-guest complexation.[7] Some of our past and continuing research has led to the construction of large preorganized macrocyclic cavities bearing concave functionalities.[8] Most of these have utilized naphthalenes and diyne bridges as their basic structural units while pyridines have served as the hydrogen-bond accepting sites. We were interested in expanding our arsenal of complexation functionality as well as our general structural framework. In 1985 Breslow et al. communicated the synthesis of an exo(C-Me)-exo(C-Me) host.[9] We were interested in the anatomical structure of this species since it closely resembles our own macrocycles with its diyne bridges and well-defined three-dimensional cavity. Phosphine oxides have been recognized as strong hydrogen-bond acceptors,[10] and we contemplated using these as our loci complexation. Etter has recently reported the successful cocrystallization of various hydrogen donors with triphenylphosphine oxide.[11] Most of the literature regarding phosphorus-containing macrocycles has focused on the development of macrocyclic phosphines and phosphites as ligands for

† X-ray crystallographer for the Chemistry Department at the University of Wisconsin—Madison.

(1) Friedrichsen, B. P.; Whitlock, H. W. *J. Am. Chem. Soc.* **1989**, *111*, 9132.

(2) For an excellent review, see: Diederich, F. *Angew. Chem., Int. Ed. Engl.* **1988**, *27*, 362-386.

(3) (a) Tucker, A. J.; Knobler, C. B.; Trueblood, K. N.; Cram, D. J. *J. Am. Chem. Soc.* **1989**, *111*, 3688. (b) Paek, K. S.; Cram, D. J. *Bull. Kor. Chem. Soc.* **1989**, *10*, 568. (c) Sherman, J. C.; Cram, D. J. *J. Am. Chem. Soc.* **1989**, *111*, 4257.

(4) (a) Fages, F.; Desvergne, J. P.; Kotzybahibert, F.; Lehn, J. M.; Marsau, P.; Albrechtgary, A. M.; Bouaslaurent, H.; Aljoubbeh, M. *J. Am. Chem. Soc.* **1989**, *111*, 8672. (b) Hosseini, M. W.; Lehn, J. M.; Comarmand, J. *Helv. Chim. Acta* **1989**, *72*, 1066. (c) Hosseini, M. W.; Kintzinger, J. P.; Lehn, J. M.; Zahidi, A. *Helv. Chim. Acta* **1989**, *72*, 1078.

(5) (a) Wallon, A.; Peter-Katalinic, J.; Werner, W. M.; Vögtle, F. *Chem. Ber.* **1990**, *123*, 375. (b) Peter-Katalinic, J.; Ebmeyer, F.; Seel, C.; Vögtle, F. *Chem. Ber.* **1989**, *122*, 2391. (c) Ebmeyer, F.; Vögtle, F. *Chem. Ber.* **1989**, *122*, 1725.

(6) Seward, E. M.; Hopkins, R. B.; Sauerer, W.; Tam, S. W.; Diederich, F. *J. Am. Chem. Soc.* **1990**, *112*, 1783. (b) Hester, M. R.; Uyecki, M. A.; Diederich, F. *Isr. J. Chem.* **1989**, *29*, 201. (c) Diederich, F.; Cutter, H. D. *J. Am. Chem. Soc.* **1989**, *111*, 8438.

(7) (a) Breslow, R.; Greenspoon, N.; Guo, T.; Zarzycki, R. *J. Am. Chem. Soc.* **1989**, *111*, 8296. (b) Petti, M. A.; Sheppod, T. J.; Dougherty, D. A. *Tetrahedron Lett.* **1986**, *27*, 5563. (c) Canceill, J.; Lacombe, L.; Collet, A. *J. Am. Chem. Soc.* **1986**, *108*, 4230. (d) Williams, K.; Askew, B.; Ballester, P.; Buhr, C.; Jeong K. S.; Jones, S.; Rebek, J., Jr. *J. Am. Chem. Soc.* **1989**, *111*, 1090.

0002-7863/90/1512-8931$02.50/0