Noncovalent Structural Models for the Asp-His Dyad in the Active Site of Serine Proteases and for Solid-State Switching of Protonation States: Crystal Structure of the Associates of 1,1′-Binaphthyl-2,2′-dicarboxylic Acid with Imidazole in Dihydrated and in Anhydrous Forms[1]

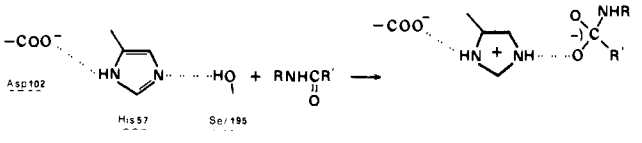
Mátyás Czugler,*[2a] János G. Ángyán,[2b] Gábor Náray-Szabó,[2b] and Edwin Weber[2c]

Contribution from the Central Research Institute of Chemistry, Hungarian Academy of Sciences, H-1525 Budapest, POB 17, Hungary, Chinoin Pharmaceutical and Chemical Works, H-1325 Budapest, POB 110, Hungary, and the Institut für Organische Chemie und Biochemie der Universität Bonn, D-5300 Bonn 1, West Germany. *Received June 24, 1985*

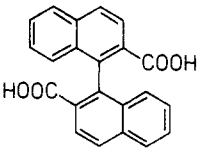
Abstract: Crystals with the stoichiometry 1:1:2 of 1,1′-binaphthyl-2,2′-dicarboxylic acid/imidazole/water (**2**) are formed in aqueous ethyl acetate. X-ray structure analysis [a = 8.955 (2) Å, b = 10.906 (3) Å, c = 12.018 (3) Å, α = 110.24 (2)°, β = 92.05 (3)°, γ = 97.56 (2)°, $P\bar{1}$, Z = 2, R = 0.028] reveals a proton transferred from one of the –COOH groups to imidazole, thereby establishing a strong intermolecular intraadduct hydrogen bond between the imidazolium and carboxylate ions. Two water molecules assist in stabilizing this ion pair via an extensive sheet of hydrogen bonding. The spatial arrangement of corresponding atomic sites of the ion pair and water molecules resembles closely that in a native enzyme *Streptomyces Griseus Protease A* (SGPA) [$\bar{\Delta}$ < 0.27 Å for seven fitted atoms]. Theoretical calculations accentuate the role of water molecules in this proton-transfer event and highlight a similar electrostatic pattern between **2** and native SGPA. On the contrary, an 1:1 adduct of 1,1′-binaphthyl-2,2′-dicarboxylic acid with imidazole (**3**), formed in water-free medium, seems to have less well-defined proton-transfer characteristics [a = 14.705 (3) Å, b = 9.396 (3) Å, c = 14.979 (3) Å, β = 102.06 (1)°, $P2_1/c$, Z = 4, R = 0.096] and more equivocal energies for different protonation states. Formal similarity of this 1:1 adduct **3** to topological requirements, forecasted for model building blocks for logic circuits at the molecular level, offers a way for studying structural relationships in such systems as well.

Modeling of active sites in enzymes is one of the challenging tasks in recent chemistry.[3] Efforts are largely based on recent knowledge on the structure of biopolymers[4] and model systems[3,5] and are driven by the hope to arrive at successful mimicking of certain enzymatic properties that may be useful in practical applications (selectivity, specificity, turnover rates, etc.).[6] Some of the successful models known hitherto have their functional groups attached to relatively complex artificial molecules with eminent inclusion properties.[5] This is understandable since inclusion compounds and enzymes exhibit many common features, e.g., taking advantage of nonbonding interactions in establishing a given structure like an enzyme–substrate or a host–guest complex. As the realm of inclusion compounds may roughly be divided into the families of molecular (capsule) and lattice (clathrate) inclusions,[7] we attempted to extend our studies into the latter direction. Such an approach seems to be verified also by the formidable difficulties when trying to construct a covalently bound model enzyme.[8] Access to simple lattice models is thought to clarify, as far as we are concerned, some structural conditions of a promising enzyme mimicry and is hoped to offer advantages

Scheme I



in the early steps of a design procedure. Therefore, we carried out complexation and crystallization experiments with 1,1′-binaphthyl-2,2′-dicarboxylic acid (**1**) (hereinafter BNDA), a



1

2: **1**·Imidazole·$2H_2O$

3: **1**·Imidazole

clathrate host that has recently been proved to possess eminent inclusion ability[9] and imidazole which due to its properties[10] plays an important role in many biochemical systems. The carboxylic groups essential in the complexing power of BNDA may play a role as in the serine protease enzymes. This family of enzymes cleaves peptide and ester bonds[11] by forming a tetrahedral intermediate at the peptide (ester) carbonyl with the use of (– + –) charge distribution of the so-called catalytic triad(hydroxyl of Ser 195, imidazole of His 57, and carboxylate of Asp 102; see Scheme I).

Thus, it seemed attractive to examine details of the crystal structure of an imidazole complex of BNDA both by X-ray

(1) Part of this work has been presented at the XIIIth IUCr Congress, Collective Abstract pp C-63, Hamburg, FRG, 1984.

(2) (a) Research Institute of Budapest. (b) CHINOIN. (c) University of Bonn.

(3) (a) Dugas, H., Penney, C., Eds. "Bioorganic Chemistry"; Springer: New York, 1981. (b) Green, B. S., Ashani, Y., Chipman, D., Eds. "Chemical Approaches to Understanding Enzyme Catalysis"; Elsevier: Amsterdam–New York, 1982. (c) Eggerer, H., Huber, R., Eds "Structural and Functional Aspects of Enzyme Catalysis"; Springer: New York, 1981. (d) Bartmann, W., Snatzke, G., Eds., "Structure of Complexes between Biopolymers and Low Molecular Weight Molecules"; Wiley: New York, 1982. (e) Nakanishi, K., D. "Bioorganic Studies on Receptor Sites"; Wiley: New York, 1984; Tetrahedron Symp. print, No. 14, p 455. (f) Cram, D. J. In "Chemistry for the Future"; Grünewald, H., Ed.; Pergamon Press: Oxford–New York, 1984.

(4) Bernstein, F. C.; Koetzle, T. F.; Williams, G. J. B.; Mayer, E. F.; Brice, M. D.; Rogers, J. R.; Kennard, O.; Shimanouchi, T.; Tasumi, M. *J. Mol. Biol.* **1977**, *112*, 535.

(5) (a) Breslow, R.; Trainor, G.; Ueno, A. *J. Am. Chem. Soc.* **1983**, *105*, 2739 and references therein. (b) le Noble, W. J.; Srivastava, S.; Breslow, R.; Trainor, G. *Ibid.* **1983**, *105*, 2748. (c) "Cyclodextrins"; Szejtli, J., Ed.; Reidel: Dordrecht, 1982. (d) Tabushi, I. *Acc. Chem. Res.* **1982**, *15*, 66.

(6) Todd, A. *Chem. Ind.* **1981**, 317.

(7) Weber, E.; Josel, H.-P. *J. Incl. Phenom.* **1983**, *1*, 79.

(8) Cram, D. J.; Katz, H. E. *J. Am. Chem. Soc.* **1983**, *105*, 135.

(9) Weber, E.; Csöregh, I.; Stensland, B.; Czugler, M. *J. Am. Chem. Soc.* **1984**, *106*, 3297 and references therein.

(10) Eigen, M. *Angew. Chem.* **1963**, *75*, 489.

(11) For references on structural aspects, see: (a) Kraut, J. *Annu. Rev. Biochem.* **1977**, *46*, 331. (b) Polgár, L.; Halász, P. *Biochem. J.* **1982**, *207*, 1.

0002-7863/86/1508-1275$01.50/0

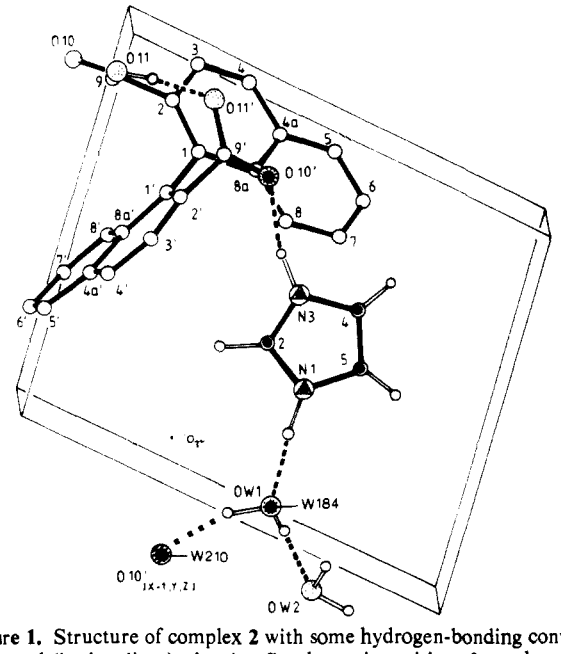


Figure 1. Structure of complex **2** with some hydrogen-bonding contacts indicated (broken lines), showing fitted atomic positions from the active site of native SGPA[21c] (black dots). A translated O10′ site is shown (hatched circle) and a transformed position of a would-be Oγ of 195 1965 is also noted. All but the relevant H atoms in H bonding are omitted for clarity.

crystallography and theoretical calculations. Further impetus was given to this analysis by the recognition of the importance of environment effects operating in such kinds of enzymes as well[12] and by the fact that most enzyme structures available up to now have somewhat limited resolution in X-ray experiments (rarely better than 2 Å).

Results and Discussion

Crystal Structure of 2. (1) Molecular Structure (Figure 1). The first crystallization from an aqueous media led to complex **2** which, as revealed by the subsequent structure determination, contained two molecules of water (cf. Experimental Section).

The basic skeleton for the BNDA host in **2** shows similarity to the dimensions of such compounds studied earlier[9] (Table IX in the supplementary material). However, some relevant changes associated with the geometry for the carboxyl groups do arise. First of all, we find one of the –COOH protons transferred to N(3). Thus, as it is also shown by the respective bonding dimensions (Figure 2), a carboxylate anion is formed with the simultaneous appearance of an imidazolium cation. The other carboxyl group displays also an unusual feature, having its –OH bond cis-positioned instead of the usual trans arrangement. Thereby, this H atom is directed toward the carboxylate anion and serves to complement its sphere of hydrogen bonding by an intramolecular interaction. Hydrogen-bonding characteristics of the anion in **2** seem to fulfill a kind of "saturation", paralleling recent observations.[13] This group is a recipient of four intra- and intermolecular/associate hydrogen bonds (cf. Table I and ref 14, 4a:0d pattern[14]). To arrive at such a molecular buildup, the orientations of the carboxyl(ate) groups are markedly changed from a more or less coplanar arrangement which is usually found for the BNDA host[9] to a more tilted one (Table II). This internal change of the structure may be assigned to the existence of ionized species in the crystal.

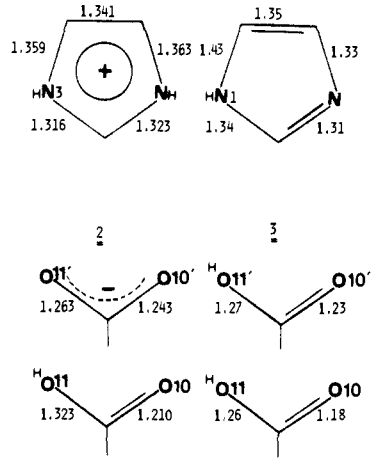


Figure 2. Comparison of the relevant selected dimensions of the –COOH and imidazolium/imidazole moieties in **2** and **3**. Esds are 0.001–3 and 0.010–14 Å for **2** and **3**, respectively.

Table I. Hydrogen Bonding Dimensions in **2** and **3**[a]

donor–H	D–H, Å	D–H···A, deg	A···H, Å	acceptor
		2		
O(11)–H(11)	1.01 (3)	172 (2)	1.56 (3)	$O(11')_i$[b]
O(W2)–H(1W2)	1.01 (3)	175 (2)	1.81 (3)	$O(11')_{ii}$
O(W1)–H(1W1)	0.95 (2)	157 (2)	1.95 (2)	$O(10')_{iii}$
N(3I)–H(3I)	0.98 (2)	163 (2)	1.80 (2)	$O(10')_i$
O(W1)–H(2W1)	0.94 (3)	170 (2)	1.87 (3)	$O(W2)_i$
C(2I)–H(2I)	1.05 (2)	140 (1)	2.35 (2)	$O(W2)_{iv}$
O(W2)–H(2W2)	1.04 (2)	171 (2)	1.98 (3)	$O(11)_v$
N(1I)–H(1I)	1.04 (3)	175 (2)	1.67 (3)	$O(W1)_i$
C(4I)–H(4I)	1.00 (2)	150 (2)	2.28 (2)	$O(10)_{vi}$
		3		
O(11′)–H(11′)	0.99	152	1.66	$O(11)_i$
O(11)–H(11)	1.12	171	1.66	$N(3I)_{ii}$
N(1I)–H(1I)	0.97	155	1.79	$O(10)_{iii}$
C(2I)–H(2I)	0.95	116	2.33	$O(10')_{iv}$
C(5I)–H(5I)	0.94	164	2.35	$O(11')_{iii}$

[a] Esd's are only given where appropriate. [b] Subscripts referring to symmetry operations relating H-bonded pairs of atoms are as follows: i, x, y, z; ii, $1-x, 1-y, 1-z$; iii, $x-1, y, z$; iv, $-x, 1-y, 1-z$; v, $x-1, 1+y, z$; vi, $x, 1+y, z$ for **2**; i, $1-x, -y, 1-z$; ii, $1-x, y-{}^1/_2, {}^1/_2-z$; iii, $-x, -y-{}^1/_2, z-{}^1/_2$; iv, $x, {}^1/_2-y, z-{}^1/_2$ for **3**.

(2) Crystal Packing (Figure 3). The most interesting observations are related to certain electrostatic features inferred from the crystal structure and the characteristics of the participating entities.

An extensive network of hydrogen bonds in forms of closed loops and chains is seen in **2** (Figure 3, Table I). One of these motifs was already mentioned in context with the carboxylate anion. In contrast to this, the carboxyl group maintains only three H bonds (2a:1d). The H-bonding characteristics of the two water molecules shows also some deviations. One of them [O(W2)] is engaged, with its full binding capacity (2a:2d) involving even a short contact of the C–H···O type[15] with the imidazolium ring. The other one [O(W1)] maintains three hydrogen bonds (1a:2d), accepting only the hydrogen atom of N(1). These two water molecules, both relatively close to a center of symmetry, are placed in the crystal lattice so as to make up a *sheet* near $x \sim 0$ (Figure 3). This layer of water molecules may effectively shield the carboxylate charges facing each other in the crystal.[16] A further shielding may be exerted by the naphthyl groups protruding over the symmetry center related pairs of imidazolium cations. As requested by the symmetry, these cations are parallel to each other and at a distance

(12) (a) Náray-Szabó, G.; Polgár, L. *Int. J. Quantum Chem., Quantum Biol. Symp.* **1980**, *7*, 397. (b) Kollman, P. A.; Hayes, D. M. *J. Am. Chem. Soc.* **1981**, *103*, 2955. (c) Náray-Szabó, G.; Kapur, A.; Mezey, P. G.; Polgár, L. *THEOCHEM* **1982**, *90*, 137. (d) Umeyama, H., Nakagawa, S. *J. Theor. Biol.* **1982**, *99*, 759.

(13) Taylor, R., Kennard, O. *J. Am. Chem. Soc.* **1982**, *105*, 5063.

(14) In the following, we apply this shorthand notation for the description of acceptor/donor characteristics of a given entity in crystal structures **2** and **3**.

(15) (a) Taylor, R.; Kennard, O.; Versichel, W. *J. Am. Chem. Soc.* **1983**, *105*, 5761. (b) Taylor, R.; Kennard, O.; Versichel, W. *Acta Crystallogr., Sect. B* **1984**, *B40*, 280.

(16) It is worth noting that a similar layering of water molecules is to be observed in crystal structures which contain hydrated proton species, cf. ref 17.

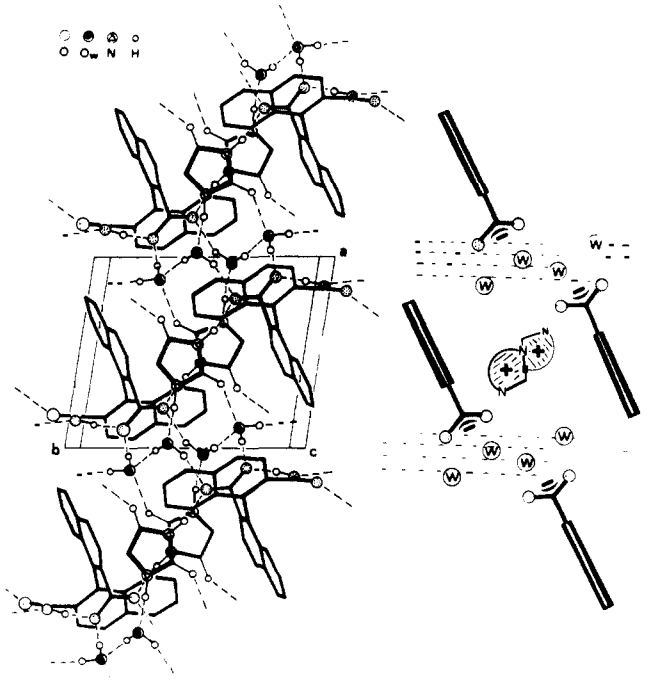


Figure 3. Hydrogen bonding and packing relations in **2**. The crystal consists of mirror-related homochiral strands of **1** fused by H bonding. Schematic representation highlights layering of water molecules, Im^+ cations, and carboxylate anions with their anchoring naphthyl moieties protruding over the cations.

of 3.4 Å. All but one H atom of the cation (0a:4d, cf. Table I) is involved in H bonds of either the N–H···O or C–H···O type.[13,15]

Possible Analogy of 2 to Serine Protease Enzymes. (1) Structural Similarity to SGPA (Figure 1). In the course of the structure determination, it became obvious that the presence of water molecules in **2** offers a possibility to extend a formal similarity to the active site of serine proteases. The first resemblance was through the number of water molecules (two) in the asymmetric unit of **2**. Interestingly, this number coincides with the number of internal water molecules associated with the active site of some serine protease enzymes (e.g., in α-chymotrypsin,[18] β-trypsin,[19] α-chymotrypsin dimer,[20] and *Streptomyces Griseus Protease A* and *B*[21,22]). Thus, it seemed worth examining steric relations between respective atomic sites in a serine protease enzyme and in **2**. The choice has fallen on two forms (the native enzyme and a peptide–aldehyde-inhibited one[21b]) of SGPA which have rather precise structures.[21c]

Atomic sites from the native enzyme representing the imidazole ring of His 57, Oδ1 of Asp 102, and a water W184 were subjected to a least-squares comparison[24] with the respective positions (imidazolium ring, O10′ and OW1) in **2**. The procedure reveals an appreciably good fit for these seven positions ($\bar{\Delta} \leq 0.27$ Å, Table III, Figure 1). To test the "predictive power" of the crystal

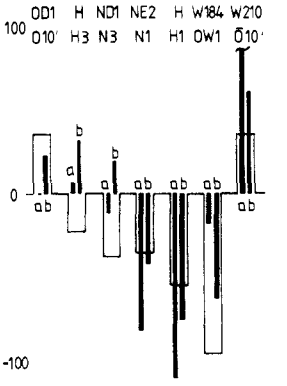
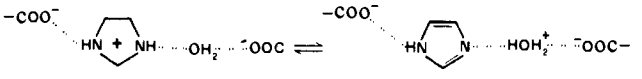


Figure 4. Comparison of the electrostatic potential patterns (in kJ/mol) at structurally corresponding atomic sites in **2** (empty bars) and in SGPA (full bars) including W326 in an optimized orientation. This water is hydrogen-bonded to Asp 102. Letters a and b denote potential values deriving from two different H-atom positions at Oγ of Ser 195, both of which may be regarded as being involved in H bonding with the simultaneous acceptor role of Oγ. For the sake of better comparison, we substracted 40 kJ/mol from each entry calculated for SGPA. Apart from the slight maximum character along Oδ1, H(Nδ1), and Nδ1, the minimum character of the patterns is retained. H-atom positions corresponding to a and b and for W326 are (−13.90, 19.50, 25.19), (−13.12, 20.03, 23.80), (−9.20, 21.07, 29.87), and (−8.11, 20.35, 19.09), respectively (given in Å, coordinates placed in the reference system of SGPA, ref 48). The value of the potential at W210 is +192 kJ/mol for a.

Scheme II



structure of **2**, a further point of identity was sought for a second water molecule (W210) of native SGPA, which is in hydrogen-bonding distance to W184. Both of them are supposed to mimic the C-terminal carboxylate–oxygen atomic positions of a would-be product (at the P_1 site) in the native enzyme.[21b] As revealed in Figure 1, the fitting procedure places W210 in the lattice of **2** distant (ca. 3.5 Å) to OW1. However, it is brought essentially to the same place ($\Delta \simeq 0.4$ Å) to an oxygen atom (O10′′) of the carboxylate anion translated by −1 along the *a* crystal axis (*a* = 8.955 Å). Thus, in the crystal structure of **2**, an atom with a partial negative charge assumes a steric position which is similar to that of an expected product in the enzyme.

(2) Electrostatical Similarity of 2 to the Active Site of SGPA (Figure 4). The structural analogy of selected portions in the structure of **2** to a part of the catalytic triad in serine proteases accentuates the importance of environmental effects and suggests appreciable electrostatic similarity as well. For example, in vacuo, a HCOOH– imidazole dyad exists in neutral form, and in the liquid (solution) state, it is just the environment which pushes the proton of the acid to the imidazole.[12c] The effect of the environment may be characterized by the electrostatic potentials created by the surrounding in a given motif. Figure 4 illustrates the distribution of the values of the electrostatic potential at atomic sites in the active site of SGPA and in **2**. A basically similar shape is shown for the two samples, reproducing the −+− distribution of charges characteristic for an enzyme with a tetrahedral intermediate. The value of the electrostatic potential at O10′ sites in the lattice of **2** is a decisive feature of this distribution. It is therefore somewhat unexpected that an even more positive potential is found at W210 in the enzyme. Thus, a potential which balances a partially negatively charged atom in the crystal structure of **2** is still a lower one than of the structurally similar W210 site of SGPA.[26] Therefore, on the basis of structural and electrostatic analogy between the two systems, one may conclude

(17) (a) Taesler, I. *Acta Univ. Upps.* **1981**, *S91*. (b) Taesler, I., Lundgren, J.-O. *Acta Crystallogr., Sect. B* **1978**, *B34*, 2424.
(18) Birktoft, J. J.; Blow, D. M. *J. Mol. Biol.* **1972**, *68*, 187.
(19) Bode, W.; Schwager, P. *J. Mol. Biol.* **1975**, *98*, 693.
(20) Blevins, R. A.; Tulinsky, A. *J. Biol. Chem.*, submitted for publication.
(21) (a) Sielecki, A. R.; Hendrickson, W. A.; Broughton, C. C.; Delbaere, L. T. J.; Brayer, G. D.; James, M. N. G. *J. Mol. Biol.* **1979**, *134*, 184. (b) James, M. N. G; Sielecki, A. R.; Brayer, G. D.; Delbaere, L. T. J. *J. Mol. Biol.* **1980**, *144*, 43. (c) James, M. N. G.; Sielecki, A. R., private communication. Coordinates of active site residues, His 57, Asp 102, Ser 195, and two internally bound water oxygen atoms (W184 and W210), for the native enzyme are from a further refinement with respect to the 1.8-Å structure.
(22) It is to be noted that no such water molecules were found in the "active site" of trypsinogen.[23] Also, according to recent illuminating studies of Meot-Ner, the most important changes in hydration enthalpies of bulky organic cations are brought about by the first two coordinating water molecules, cf. ref 25.
(23) Fehlhammer, H.; Bode, W.; Huber, R. *J. Mol. Biol.* **1977**, *111*, 415.
(24) Hoard, L. G. "ABCLS: Least-Squares Model-Fitting Program", University of Michigan: Ann Arbor, 1984.

(25) (a) Meot-Ner (Mautner), M. *J. Am. Chem. Soc.* **1984**, *106*, 1257, 1265. (b) Meot-Ner (Mautner), M. *Acc. Chem. Res.* **1984**, *17*, 186.
(26) It must be noted that W 210 in the enzyme is equidistant from O and N of Ser 195.

Table II. Least-Squares Planes and Their Dihedral Angles in **2** and **3**,[a] Relevant Dihedral Angles between Planes and Their Respective Equations. Equations of the Form $AX + BY + CZ - D = 0$ where X, Y, Z Refer to Orthogonal Coordinates

	2		3		
atoms	A	B	A	B	planes
C1	-22	43	23	-2	I
C2	37	32	-13	-29	
C3	15	-49	28	-4	
C4	-22	-38	-18	41	
C4a	-18	5	-19	31	
C5	2	56	17	-41	
C6	22	44	8	-36	
C7	25	-42	9	-1	
C8	-5	-50	-2	40	
C8a	-22	-1	-28	-4	
C9*	100	165	-107	-142	
C9*	-24	-12	-22	11	II
O10	0	0	0	0	
O11	0	0	0	0	
C2	0	0	0	0	
N1I	-2		-1		III
C2I	4		-15		
N3I	-2		12		
C4I	3		-21		
C5I	1		14		
O10*	141				
OW1*	-22				

plane	IA, deg	IIB, deg
IB		
2:	81.6 (1)	60.4 (1)
3:	87.7 (2)	34.7 (3)
IIA		
2	50.6 (1)	74.5 (1)
3	25.2 (3)	50.4 (5)
III		
2		39.2 (1)[b]
3		65.1 (4)[b]

	coefficients			
	A, deg	*B*, deg	*C*, deg	*D*, Å
IA				
2	0.697 (27)	0.530 (32)	-0.483	-1.085 (7)
3	-0.34 (14)	-0.79 (12)	-0.52	-7.29 (3)
IB				
2	0.506 (26)	-0.760 (34)	-0.408	-0.118 (4)
3	0.93 (15)	-0.31 (11)	-0.22	7.02 (2)
IIA				
2	0.987 (69)	0.045 (64)	0.157	8.301 (10)
3	0.17 (28)	0.53 (30)	0.83	9.01 (4)
IIB				
2	-0.106 (66)	-0.198 (65)	-0.974	-7.692 (7)
3	0.64 (24)	-0.24 (26)	-0.73	1.16 (5)
III				
2	0.120 (74)	0.429 (8)	-0.896	-4.906 (10)
3	-0.38 (28)	0.05 (33)	-0.92	-4.42 (3)

[a] Atoms marked by an asterisk are *not* defining the plane. [b] The respective values of these dihedral angles in SGPA (native and the peptide-aldehyde inhibited forms) are 50.1° and 87.5°, respectively.

that a $-+-$ charge distribution exists even in the native SGPA, the second negative charge being represented by W210.

(3) Water-Mediated Proton Transfer. The analogy between structural and electrostatic features of **2** to some parts of the active site of serine protease enzymes calls for attention in regard of the role of such "conserved" water molecules observed independently in different structures. The topological similarity of water arrangement in **2** to those for H_3O^+ and $H_5O_2^+$ and higher hydrated proton structures,[17] which are thought to play an essential role in the extreme fast proton transfer in water,[10,27] suggests that these water molecules may also play a role in some kind of proton-transfer processes in an enzyme as well. Such a role would be, for example, to ensure the proper protonation states of the active site before an actual catalytic process begins.[28] The water molecules W184–W210 may provide an ideal pathway traversed by the proton leaving of Nϵ2 of His 57. Indeed, theoretical calculations for a proton transfer modeled by **2** and as depicted in Scheme II, reveal the following balance of energy.

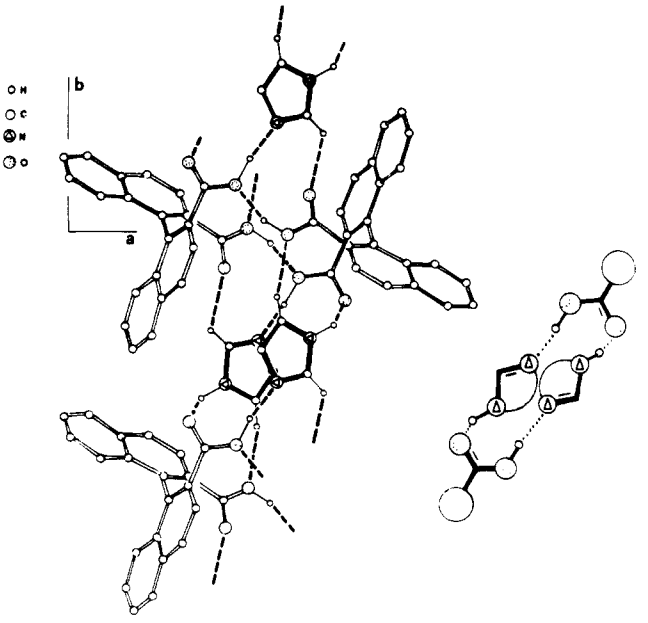


Figure 5. H bonding for **3** with a schematic representation of the central loop.

The change of the electrostatic lattice energy in **2** is -101 kJ/mol for the right-hand side of Scheme II, while the corresponding energy cost of breaking the N–H covalent bond may be approximated as $+115 \pm 5$ kJ/mol,[29] suggesting that the left-hand form of Scheme II is stabilized by 10–20 kJ/mol. This compares well with recent results of quantum chemical calculations for an amidine–water 1:2 model,[30] where such a mediator role for water molecules between different tautomeric forms was also proposed. A further support was lent to this assumption by a recent theoretical study for the 2-pyridone dihydrate system. Here, a somewhat similar H-bonding geometry as in **2** and essentially existence of an $H_5O_2^+$ ion as the mediating species in the transition state was established.[31] A $-+-$ charge (vide supra) pattern present in a native enzyme would directly favor such a H transfer. This event could be, for example, pivoted upon binding of the substrate molecule to the enzyme.

Crystal Structure of 3. The importance of water molecules assisting in a proton-transfer event in **2** made it plausible to grow crystals of BNDA with imidazole in a water-free medium. Thus, crystals **3** were grown with BNDA/imidazole 1:1 stoichiometry. This structure albeit with much worse accuracy than desired, revealed some differences as compared with **2**, e.g., the relative orientation of the –COOH groups with respect to their anchoring naphthyl moieties (cf. Table II). Though they are more tilted than found sofar in alcohol,[9,32] DMF,[33] and bromobenzene[33] in-

(27) (a) Lengyel, S.; Conway, B. E. In "Comprehensive Treatise of Electrochemistry"; Conway, B E., Bockhus, J. O. M., Yeager, E., Eds.; Plenum Press: New York, 1983; Vol. 5, Chapter 4. (b) Wang, J. H. *Proc. Natl. Acad. Sci.* U.S.A. **1970**, *66*, 874. (c) Ángyán, J.; Allavena, M.; Picard, M.; Potier, A.; Tapia, O. *J. Chem. Phys.* **1982**, *77*, 4723.

(28) I.e., creation of the ion-neutral form of the Asp 102-His 57 couple in Scheme I. This step is a prerequisite, creating of a good nucleophile from the –OH function of Ser 195.

(29) We approximated this quantity from the $H_3NH^+\cdots OH_2$ proton-transfer energy in vacuo; cf.; Scheiner, S. *Int. J. Quantum Chem.* **1983**, *23* 753.

(30) Zielinksi, T. J.; Poirier, R. A.; Peterson, M. R.; Csizmadia, I. G. *J. Comput. Chem.* **1983**, *4*, 419.

(31) Field, M. J.; Hillier, I. H.; Guest, M. F. *J. Chem. Soc., Chem. Commun.* **1984**, 1310.

(32) Czugler, M.; Weber, E. Unpublished X-ray analyses of BNDA inclusion compounds with *t*-BuOH and *n*-PrOH.

(33) Csöregh, I.; Czugler, M.; Sjögren, A.; Cserzö, M.; Weber, E. *J. Chem. Soc., Perkin Trans. 2*, in press.

Table III. Results of Least-Squares Fitting of Selected Atoms of **2** and of SGPA[a]

	assumed (model a)				fitted (model b)			
	X_0	Y_0	Z_0		X_0	Y_0	Z_0	Δ, Å
				Final Cartesian Coordinates				
N1I	1.8324	3.5782	7.4387	Nε2H[b]	1.8963	3.3575	7.6012	0.281
C2	2.6139	2.5799	7.0972	Cε1H	2.8901	2.8901	7.2721	0.358
N3I	3.8605	2.8649	7.3697	Nδ1H	4.0425	3.0281	7.4594	0.260
C4I	3.8919	4.0909	7.9552	CγH	3.8331	4.3042	7.9398	0.222
C5I	2.6268	4.5330	7.9992	Cδ2H	2.4839	4.5246	8.0155	0.144
O1O′	6.2311	1.5530	6.8997	Oδ1A	5.9966	1.4142	6.6655	0.359
OW1	−0.8588	3.6917	7.1548	OW184	−0.9448	3.7647	6.9230	0.258
				Other Atoms				
$O1O_T'$	−2.7239	1.5530	6.8997	OW210	−2.6085	1.8655	7.1277	0.404
				OγS	−0.1094	1.1092	7.0764	
				Oδ2A	6.8301	1.9153	8.6691	

Final Euler Angles[c] and Translations for Model b

A = 7.34°	B = 54.18°	C = 37.06°	
X_{tr} = 13.81 Å	Y_{tr} = −28.70 Å	Z_{tr} = 17.02 Å	

	$\Delta(X)^2$	$\Delta(Y)^2$	$\Delta(Z)^2$	Δ, Å
		Statistics for Seven Fitted Atoms		
av	0.0285	0.0217	0.0270	0.269
rms	0.0281	0.0195	0.0246	0.077
		Overall Statistics for Eight Atom Pairs		
av	0.0266	0.0312	0.0301	0.286
rms	0.0266	0.0324	0.0244	0.085

[a] Orthogonal coordinates for Models "a" and "b" representing selected atoms of **2** and SGPA; Euler angles A, B, and C and translations to obtain fitted coordinates. For the inverse, i.e., **2** fitted to SGPA, interchange A for C and apply all angles with opposite sign with the translations −18.8, 22.37, and 21.23 Å. [b] Suffixes H, S, and A denote atoms of residues His 57, Ser 195, and Asp 102 of SGPA. [c] C.f. ref 49.

Table IV. Summary of Crystal and Experimental Data on **2** and **3**

	2	**3**
	$C_{22}H_{13}O_4^-/C_3H_5N_2^+/2H_2O$	$C_{22}H_{14}O_4/C_3H_4O_2$
fw	446.5	410.4
a, Å	8.955 (2) Å	14.705 (3)
b, Å	10.906 (3)	9.396 (3)
c, Å	12.018 (3)	14.979 (3)
α, deg	110.24 (2)	90
β, deg	92.05 (2)	102.06 (1)
γ, deg	97.56 (2)	90
space group	$P\bar{1}$	$P2_1/c$
Z	2	4
D_{calcd}	1.363 Mg·m^{-3}	1.347 Mg·m^{-3}
λ	Mo Kᾱ (0.71073 Å)	Mo Kᾱ
crystal size, mm	0.3 × 0.2 × 0.3	0.03 × 0.08 × 0.35
θ range	1.5–25.0°	1.5–25.0°
N_{tot}	3825	3427
N_R	1975 [$>3\sigma(I)$]	924 [$>1.9\sigma(I)$]
N_{par}	299 + 88	279
R_R	0.028	0.096
R_w	0.041	0.056
R_{tot}	0.073	0.096

clusion compounds, they are by far not perpendicular to their naphthyl rings as in **2**.[34] Moreover, such an arrangement would exclude a possible assistance of one of the carboxyl groups in proton transfer as observed for **2**.

Packing Features and Energy Relations of Different Tautomeric Forms in 3 (Figure 5). There seems to be a genuine difference in the hydrogen-bonding pattern of **3** (Figure 5) compared with **2** (Figure 3). In **3**, the most characteristic feature is a central loop of hydrogen bonds which in this respect is resembling the pattern found for various alcohol clathrates of BNDA.[9] Though in imidazole the donor–acceptor functions are more distant than in an alcoholic –OH, it may still act analogously to alcohols. Therefore, the arrangement in **3** may be indicative that at least a reasonable portion of the molecular associates in **3** do exist in the neutral form. Analysis of possible H-bonding contacts also reveals that the imidazole ring is involved in two C–H···O interactions as classified by Kennard and her co-workers.[13] Another interesting aspect of the central H-bonding loop is seen in Figure 5. It has principally the same topology as it was forecasted for a model molecule containing bis(imidazolylhemiquinolyl) functions.[35] This kind of structure is expected to ensure proper steric and energetic pathways for switched proton transfer which are thought of being used in constructing logic circuits at the molecular level.[35] Such a switching phenomena in **3** could be recognized as a kind of structural disorder. Indeed, the lower accuracy of the structure determination of **2** could not only be due to unfavorable experimental conditions (insufficient single-crystal size) but due to a coupled positional disorder affecting O- and H-atomic positions.

Assessment of Protonation States in 2 and 3: The Stabilizing Role of Water. The structure determination of complex **3** did not yield all the important atomic positions unambiguously.

To circumvene this problem, a theoretical approach was adopted. The calculation of the energy difference between forms characterized by different protonation states essentially involves the determination of the proton-transfer energy corrected by the effect of the crystalline environment.

The first question concerns the problem, whether the structures **2** and **3** are essentially ionic or neutral. In **2**, the electrostatic lattice energy was found to favor the ionic form by about 67 kJ/mol, while the in vacuo proton-transfer energy could be estimated to be in the range of 60–100 kJ/mol, favoring the neutral form. Taking into account the correction of the systematic errors owing to the use of CNDO/2 charges,[36] we find that for the crystal **2**, the loss of covalent energy upon ion-pair formation is overbalanced by the stabilization in the electrostatic lattice energy.

The calculations offer also an easy way to elucidate the role of water molecules in stabilizing the ionic form of **2**. The electrostatic lattice energy difference between the ionic and neutral structures with dropping the water molecules gives 47 kJ/mol (uncorrected value). That is, this hypothetical anhydrous structure made from **2** is less favored in the ionic form: about 30% of the

(34) A coordinato-clathrate of BNDA with Me_2SO,[33] however, has a rather similar dihedral angle (33.9°) for one of the –COOH groups as compared with **3**. A value in this range seems to reflect rather the steric than the electronic requirements of the guest molecules.[9]

(35) Aviram, A.; Seiden, P. E.; Ratner, M. A. In "Molecular Electronic Devices"; Carter, F. L., Ed.; Marcel Dekker: New York and Basel, 1983.

(36) Voogd, J.; Derissen, J. L.; van Duijneveldt, F. B. *J. Am. Chem. Soc.* **1981**, *103*, 7702.

total stabilization effect can be attributed to the presence of water molecules in **2**.

In the anhydrous crystal **3**, only a small stabilization (11 kJ/mol) of the ionic form is ensured by electrostatic lattice energy, which is by far not enough to overweight the loss in covalent energy in the course of the proton transfer from the carboxyl to the imidazole, leading to the ionic form of the crystal. Therefore, it is quite probable that the protonation state corresponds to a neutral structure of **3**.

This means as well that it is just water which induces proton transfer. Such an event may be thought of similar to ionic hydration. In aqueous solutions, ions or ion pairs are favored over neutral structures, which in general are more stable in vacuo.[37]

The neutral form of **3** has two possible proton positions on the imidazole ring, N1 and N3. These two positions are energetically equivalent and they may be occupied with the same weight. Therefore, it may happen that the failure of locating the proton along the imidazole–carboxyl H bond is due to structural disorder (cf. preceding section).

Conclusions

We have shown by X-ray analysis and theoretical calculations that some structural and electrostatic aspects of rather complex molecules are successfully modeled via building of crystal lattices. By this method, we also have access to model systems that are not amenable either to direct studies or those which are not precise enough. Thus, the crystal structure of a simple model which shows striking structural and electrostatic similarity to serine protease enzymes is now offered for further spectroscopic and theoretical determinations. Such a model seems to advance our knowledge about specific hydration properties and possible functions of water in this type of enzymes.

The proton-transfer mediating role as put forward seems to be ideally suited for water in accord with earlier[10,27] and more recent studies.[25,30,31] Invariably present "fixed" water molecules in active sites of serine protease enzymes, therefore, may also exert their influence even in a catalytic step.[38] Also, relatively little attention was paid until now to the possible role of hydrogen bonds of the C–H···O type in this class of enzymes. Both, in the hydrated and anhydrous crystals, imidazole rings are intimately involved in these kinds of interactions.[40]

Another aspect of this study is concerned with the energetic and steric arrangement of the associated molecules in the crystal structure of **3**. Similarities to theoretically constructed molecules[35] seems also to verify efforts that are aimed at engineering of crystal structures.

Experimental Section

(1) Crystal Structure Determination of 2. Crystals of complex **2** were obtained by carefully mixing ethyl acetate solutions of BNDA and imidazole (0.0092 g (0.027 mmol) and 0.0037 g (0.054 mmol)) with traces of water (5 μL) in the latter. Well-formed irregular prism-shaped crystals appear in a few hours.[41] One such specimen was sealed in a thin-walled glass capillary for X-ray data collection. Data were collected on an automated four-circle diffractometer (CAD-4). The initial structure model was developed by direct method[42] (MULTAN 78) applied for 260 E values and included all non-H atoms of BNDA except for the two water and imidazole molecules which were in turn found in difference electron-density (e.d.) synthesis.

Full-matrix treatment[43] of the adjustable parameters led rapidly to convergence and H atoms belonging to the water molecules, and a carboxyl group showed up even at this isotropic stage already (R = 0.075). It was also then noted that the carboxyl group having more uniform C–O distances seems to have its proton lost to the neighboring imidazole which in turn possesses an extra protonic site at N(3). Apart from conventional difference e.d. syntheses, thus, the distribution around the nontrivial H-atom sites was examined by calculating slant-plane slices of the e.d. in the respective atomic planes.[43]

Though all H atoms sought could be seen, they were only added when their positions were confirmed by the outlined procedure after some anisotropic refinements of the non-H atoms. Near to the termination of the refinement procedure, positional and isotropic thermal parameters of all H atoms were also adjusted in separate steps using a restricted data set with reflections of the lower scattering angle range ($\sin \vartheta/\lambda_{max} = 0.4$). All H atoms showed remarkable stability and bonding dimensions expected for a room-temperature X-ray study. It is to be noted that both water oxygen atoms and also the imidazolium cation had rather low and isotropic thermal parameters, indicating that this crystal lattice is built up via strong intermolecular interactions. Before the final refinement step (R = 0.028), an empirical extinction coefficient was also introduced and refined. The final ΔF map had its maximum value (0.08 e/Å^3) near (0.94 Å) to OW1, while its minimum value (–0.10 e/Å^3) is in the vicinity of O11 and O11′ atoms (1.5 and 1.7 Å, respectively).

(2) Crystal Structure Determination of 3. Crystals of **3** were obtained by the same procedure as for **3** with the exception that no water was added to either solutions. On mixing, "schlieren" develops and crystals grow much more slowly and remain less developed in spite of efforts varying growth conditions (temperature, concentration from 1:2 to 1:4 mole ratio of BNDA/imidazole, etc.).

Only a small crystal of 0.05 × 0.08 × 0.30 mm in size could be chosen as the best for measuring purposes. Data were collected and processed[43] much the same way as for **2**, including structure solution and refinement. However, due to the unfavorable observations/parameter ratio and nonpositive definite problems at the anisotropic stage of refinement, constrained group-temperature factors were introduced for the non-hydrogen atoms. Those H atoms which could be placed by geometry evidence were also added with ca. 20% higher isotropic thermal parameters than derived from the B_{eq} value of their parent atoms. The final model could be accessed only in a rather slow convergence. Peaks from a final difference e.d. synthesis were at this stage taken as indicative of putative H-atom positions at the –COOH groups and added resulting in the final R of 0.096 for 927 observations. These positions, however, do not resolve the ambiguity deriving from the bonding dimensions of the respective non-H atoms of the –COOH and imidazole groups. The final ΔF map contained features ranging from 0.6 to –0.7 e/Å^3, of which no physical significance could be attributed.

(3) Energy Calculations. To calculate relative energies of different protonation states in **2** and **3**, we used a model where BNDA is replaced by two HCOOH (or $HCOO^-$) molecules in a spatial position corresponding to the substituents of the binaphthyl group. The total energy change was estimated by the formula

$$\Delta E = \Delta E_c + \Delta E_e \qquad (1)$$

where subscripts c and e stand for covalent and environmental contributions, respectively.

For the HCOOH + imidazole $\rightleftarrows$ $HCOO^-$ + imidazolium process, E_c is estimated by ab initio quantum chemical methods using a split valence-shell (4-31G) basis set.[12b-d] In a previous work,[12c] we applied empirical corrections to estimate ΔE_c, and a range from 15 to 101 kJ/mol was given. Kollman[12b] and Umeyama[12d] obtained 80 and 67 kJ/mol, respectively. Summing up these results, one would most probably suggest 80 ± 20 kJ/mol for ΔE_c.

We applied an electrostatic approach to estimate ΔE_e

$$\Delta E_e = \sum_i V_i^{IP}(r_i) q_i^{IP} - \sum_j V_j^{N}(r_j) q_j^{N} \qquad (2)$$

where the first and second terms represent the interaction energy between the environment and the ion-pair (IP) and neutral forms (N), respectively. $V^{IP}(r_i)$ and $V^N(r_i)$ are the electrostatic potentials produced by the environment of ion-pair and neutral structures at atom i, bearing a net charge of q_i. Though polarization, charge transfer, and other effects are neglected if eq 2 is applied, similar approximation was found to work

(37) Warshel, A. *Acc. Chem. Res.* **1981**, *14*, 284.

(38) Stein, R. L. *J. Am. Chem. Soc.* **1983**, *105*, 5111 and references therein. In spite of numerous efforts, the controversy concerning the true nature of the catalytic mechanism of serine proteases (cf. ref 11b and 39) seems to be unsettled until now.

(39) Kossiakoff, A. A.; Spencer, S. A. *Biochemistry* **1981**, *20*, 6462.

(40) Note that atom C2 of the imidazole rings in both cases is involved in C–H···O hydrogen bonds, which may affect the interpretation of NMR data related to this carbon atom; c.f.; Hunkapiller, M. W.; Forgac, M. D.; Whitaker, D. R.; Richards, J. H. *Biochemistry* **1973**, *12*, 4732.

(41) Formation of crystals **2** has been detected under other circumstances as well. E.g., adding equimolar amounts of NaOH and *N*-methylacetamide, imidazole, and excess water yields invariably crystals of **2**. This may be taken indicative that the steric and electrostatic arrangement in **2** is a preferred one over a pH range ≥7, which, actually, lies close to the optimum operating pH for serine proteases.

(42) Main, P.; Hull, S. E.; Germain, G.; Declercq, P. J.; Woolfson, M. M. "MULTAN 78"; Universities of York: and Louvain: England and Belgium, 1978.

(43) All programs used in these computations are of those of the Enraf-Nonius SDP package (V. 18.0) together with some auxiliary programs written in house. Program SLPF (Slant Plane Fourier) is a local PDP 11/34 (64k) adaptation (Dr. L. Párkányi, CRIC) of a van der Waal's program.

sufficiently for the hydrated $FH\cdots NH_3$ prototype H-bonded system.[44]

The environment was modeled by adding successive layers of unit cells to the central (000) cell. This procedure leads to crystal blocks of similar shape to that of the unit cell. In the present study for both structures, two additional layers (corresponding to a crystal block containing 125 unit cells) proved to be sufficient to simulate the potential of the infinite crystal.

V of eq 2 was calculated by the bond increment method[45] which has already been proven to be useful in the calculation of the electrostatic potential maps of small molecules[45] and proteins.[12a,12c,46] A particular advantage of this method is that very large systems can be handled since the computational work (including the construction of the wave function) is proportional to the first power of the number of bonds in the system. For unit cells which lie far from the (0,0,0) one, the electrostatic potential was replaced by a simple monopole expression which made the calculations quite cheap even for such a large system as treated here. Atomic net charges were obtained from CNDO/2 calculations.[47]

According to a recent work by Voogd et al.,[36] the CNDO/2 charges used in our calculations generally underestimate the electrostatic lattice energies of amino acid and peptide crystals. Different systematic errors were found in predicting the energy of ion-pair (error of about 100 kJ/mol) and neutral (error of about 60 kJ/mol) structures. It means that the ΔE_e values obtained by our calculations with CNDO/2-based charges are underestimated by about 40 kJ/mol.

The protein potential of SGPA was calculated by using the PROTPOT program[46] with three-dimensional coordinates from Protein Data Bank.[48] Surface ionizable side chains were considered as fully shielded.[46]

Acknowledgment. M. C. is indebted to Prof. A. Kálmán for his support in this work. E. W. acknowledges the financial support of the Deutsche Forschungemeinschaft. We thank Prof. M. N. G. James and Dr. A. Sielecki (Edmonton) for some updated coordinates of active site residues of SGPA and Prof. A. Tulinsky (East Lansing, MI) for sending a manuscript prior to publication.

Registry No. 2, 99827-47-1; **3**, 99827-48-2; SGPA, 55326-50-6; serine protease, 37259-58-8.

Supplementary Material Available: Relative atomic coordinates, anisotropic thermal parameters, selected bond lengths and angles, and structure factor tables for **2** and **3** (13 pages). Ordering information is given on any current masthead page.

(44) (a) Ángyán, J.; Náray-Szabó, G. *Theoret. Chim. Acta* **1983**, *64*, 27. (b) Ángyán, J.; Náray-Szabó, G. *Acta Chim. Acad. Sci. Hung.* **1984**, *116*, 141.

(45) (a) Náray-Szabó, G. *Int. J. Quantum Chem.* **1979**, *16*, 265. (b) Náray-Szabó, G.; Grofcsik, A.; Kösa, K.; Kubinyi, M.; Martin, A. *J. Comput. Chem.* **1981**, *2*, 58.

(46) Ángyán, J.; Náray-Szabó, G. *J. Theor. Biol.* **1983**, *103*, 349.

(47) Pople J. A., Beveridge, D. L., Eds. "Approximate Molecular Orbital Theory"; McGraw-Hill: New York, **1970**.

(48) James, M. N. G.; Sielecki, A. R. Protein Data Bank File 1983, 85SB13, 73.

(49) The premultiplier matrix of a vector is of the form

cB	sB·sC	−sB·cC
sA·sB	cA·cC − sA·cB·sC	sA·cB·cC + cA·sC
cA·sB	−cA·cB·sC − sA·cC	cA·cB·cC − sA·sC

where "s" and "c" denote sine and cosine functions of the Euler angles A, B, and C, cf.: Schilling, J. W. Ph.D. Dissertation, University of Michigan, Ann Arbor, 1968.

Ultraviolet Raman Hypochromism of the Tropomyosin Amide Modes: A New Method for Estimating α-Helical Content in Proteins

Robert A. Copeland and Thomas G. Spiro*

Contribution from the Department of Chemistry, Princeton University, Princeton, New Jersey 08544. *Received July 23, 1985*

Abstract: Ultraviolet resonance Raman spectra, obtained with 200-nm excitation from a H_2 Raman-shifted Nd:YAG laser, are recorded for tropomyosin. The intensities of the amide II and III modes, and of the amide II′ mode in D_2O, increase with increasing pH, as the α-helical protein unfolds. A linear relationship between intensity and α-helical content has been established. This relationship has been used to estimate the helix content of myoglobin, bovine serum albumin, and cytochrome *c*, good agreement being found with previous determinations from X-ray crystallography or optical rotatory dispersion. The intensity increase associated with the loss of the helical structure is attributed to the UV absorption hypochromism of α-helical peptides. The molar Raman scattering factors for the amide II, II′, and III modes in helical and nonhelical structures are proportional to the squares of the molar absorptivities at 190 nm, the center of the amide absorption band, establishing local resonance for these amide modes with the π–π^* transition. The amide I (C=O stretching) band intensity is nearly independent of α-helical content, suggesting that it is enhanced primarily via higher lying electronic transitions.

The advent of reliable high-power pulsed lasers, whose wavelengths can readily be shifted over a wide range of the electromagnetic spectrum via nonlinear optical techniques (crystal mixers and stimulated Raman shifters) has made possible the routine acquisition of Raman spectra with ultraviolet excitation. This development has opened the way to systematic exploration of the resonance Raman characteristics of UV chromophores, including the purine and pyrimidine bases of nucleic acids[1-4] and the aromatic side chains of proteins,[5-10] as well as the amide bonds of the polypeptide backbone.[8-12] Resonance Raman spectroscopy is a useful structure probe due to its high sensitivity (easily extending into the micromolar range) and the chromophore specificity associated with wavelength matching of the laser to a

* Author to whom correspondence should be addressed.

(1) Fodor, S. P. A.; Rava, R. P.; Hays, T. R.; Spiro, T. G. *J. Am. Chem. Soc.* **1985**, *105*, 1520–1529.

(2) Ziegler, L. D.; Hudson, B.; Strommen, D. P.; Peticolas, W. L. *Biopolymers* **1984**, *23*, 2067–2081.

(3) Kubasek, W. L.; Hudson, B.; Peticolas, W. L. *Proc. Natl. Acad. Sci. U.S.A.* **1985**, *82*, 2369–2373.

(4) Fodor, S. P. A.; Spiro, T. G. *J. Am. Chem. Soc.*, in press.

(5) Rava, R. P.; Spiro, T. G. *J. Am. Chem. Soc.* **1984**, *106*, 4062–4064.

(6) Johnson, C. R.; Ludwig, M.; O'Donnell, S.; Asher, S. A. *J. Am. Chem. Soc.* **1984**, *106*, 5008–5010.

(7) Rava, R. P.; Spiro, T. G. *J. Phys. Chem.* **1985**, *89*, 1856–1861.

(8) Rava, R. P.; Spiro, T. G. *Biochemistry* **1985**, *24*, 1861–1865.

0002-7863/86/1508-1281$01.50/0