solution of bicarbonate (4.0×10^{-2} M) was prepared by dissolving 400 mg of $KHCO_3$ in 100 mL of carefully degassed water.

Preparation of a Catalyst Solution. A catalyst solution was prepared by dissolving an appropriate amount of a catalyst into 5 mL of a carefully degassed buffer solution containing 4×10^{-4} M of neutral red or 2×10^{-4} M of *p*-nitrophenol as an indicator. After the solution was diluted to 10 mL with degassed water or an aqueous $ZnCl_2$ solution (2×10^{-2} M), the pH of the solution was adjusted to a certain desired value by careful addition of 10 N HCl or 10 N NaOH.

Determination of Buffer Factor. For the determination of buffer factor, Q, 1.0 mL of the catalyst solution was mixed with the same volume of carefully degassed water, and the diluted solution was titrated with 0.01 N HCl. The absorbance of *p*-nitrophenolate absorption at 400 nm or that of neutral red absorption at 540 nm was recored for every 5-10-μL addition of 0.1 N aqueous HCl. The absorbance of *p*-nitrophenolate absorption at 400 nm or that of neutral red absorption at 540 nm was recorded for every 5-10-μL addition of 0.1 N aqueous HCl. The absorbance changes observed (ΔA) were plotted against the total concentration of the added acid, Δ[HA], corrected for the volume change during the titration. The buffer factor, Q, was determined from the slope of the titration plot, $\Delta[HA]/\Delta A$.

Rate Measurements and Rate Analysis. All the kinetic measurements reported here were carried out with a Union RA-401 stopped-flow spectrometer thermostated at 25.0 °C. A solution of a catalyst and a CO_2 solution were driven together by 4 kg/cm^2 of nitrogen pressure and both solutions of equal volume were mixed within 2 ms. After mixing, the rapid absorbance change at 400 nm for *p*-nitrophenolate anion or at 540 nm for neutral red was stored in a Union RA-405 kinetic data processor and read out on a X-Y recorder after several traces were integrated. Usually, a reaction trace within 0.5-5 s was used for analysis. The absorbance change during the reaction was converted to the concentration change of proton (or bicarbonate) on multiplication by the buffer factor. Since the absorbance change observed for the reaction in an imidazole or imidazole-zinc buffer solution showed excellent linearity with time at the early stages of the reaction, the first-order rate constants were directly obtained from the slopes of the linear portions according to eq 2. For all buffer-indicator systems, first-order dependency of the

$$k_{obsd} = \frac{(\Delta A/\Delta t)_{initial} Q}{[CO_2]_0 \text{ or } [HCO_3^-]_0} \quad (2)$$

values of $(dA/dt)_{initial}$ on the initial concentration of CO_2 ((1.7×10^{-2})-(5.7×10^{-3}) M) was ascertained. Observed catalytic activity decreased during the course of the CO_2 hydration catalyzed by **3**-zinc complex. Therefore, the rate constants were estimated by the following methods:

(1) The observed kinetic curves were corrected by subtracting the absorbance changes due to the spontaneous CO_2 hydration and the deactivation reaction of $(His)_2CD$ under the corresponding buffer conditions from the observed overall absorbance change. The deactivation term was estimated from the rate constants of the reaction between free $(His)_2CD$ and CO_2 and the actual concentration of free $(His)_2CD$ calculated from $K_{Zn^{2+}}$.[22]

(2) On the basis of the corrected kinetic curves thus obtained, the first recycling point was so defined that 1 mol of CO_2 per 1 equiv of histamino grouping was hydrated. It was calculated by using the observed absorbance change and the buffer factor.

(3) Pseudo-zero-order rate constants were obtained from slopes of the corrected kinetic curves at the initial and first or second "recycling" points.

(4) The zero-order rate constants were converted to first-order rate constants (k_1, s^{-1}) by dividing by the concentration of CO_2 still remaining at each point.

(5) The catalytic constants (k_{cat}, M^{-1} s^{-1}) were obtained from $k_1/[(His)_2CD{\cdot}Zn^{2+}{\cdot}Imd]$, where the concentration of $(His)_2CD{\cdot}Zn^{2+}{\cdot}Imd$ was calculated from $K_Z{}^{n_{2+}}$.

Acknowledgment. We are grateful to Professor R. Breslow, Columbia University, for his thoughtful and helpful discussion on the catalysis.

Registry No. 1, 76700-69-1; **2**, 76700-71-5; **3**·4HCl, 90604-16-3; CO_2, 124-38-9; β-cyclodextrin, 7585-39-9; benzophenone-3,3′-disulfonyl chloride, 17619-15-7; histamine dihydrochloride, 56-92-8; carbonic anhydrase, 9001-03-0.

(21) $Ba(OH)_2$ titration of the CO_2 solution thus obtained showed that the concentration of CO_2 was 3.2×10^{-2} M. The observed value agreed with the reported value of the saturated concentration of CO_2 in water (3.38×10^{-2} M, see: Pocker, Y.; Bjorkquist, D. W. *Biochemistry* **1977**, *16*, 5698). Total concentration of carbonate species (HCO_3^- and H_2CO_3) under the present conditions was estimated to be less than 2×10^{-4} M on the basis of the observed pH value of 4.2 and the equilibrium constants (see: Faurholt, C. *J. Chim. Phys.* **1924**, *21*, 400). Therefore, the initial concentration of total carbonate, HCO_3^- plus $H_2CO_3^-$, immediately after 1:1 vol/vol mixing with catalyst solution, was 1×10^{-4} M. On the basis of these measurements, the pH drop caused by initially present HCO_3^- and H_2CO_3 must be negligibly small (0.004 pH unit) under the present condition.

(22) For example, in the case of Figure 3, the deactivation term was estimated from the reaction trace of Figure 3b and the concentration of free $(His)_2CD$ calculated from $K_{Zn^{2+}}$.

Quantum Chemical Calculation of the Enzyme-Ligand Interaction Energy for Trypsin Inhibition by Benzamidines

Gábor Náray-Szabó

Contribution from the CHINOIN Pharmaceutical and Chemical Works, H-1325 Budapest, Hungary. *Received September 1, 1983*

Abstract: Three-dimensional coordinates of the β-trypsin-benzamidine complex were used to construct a quantum chemical model for the enzyme-inhibitor interaction. The model included all enzyme and benzamidinium ($H_4N_2CC_6H_4X^+$) atoms and two bound water molecules (W702 and W710) located near the binding site. Hydration was also treated by using a simple model of the first hydration sphere including four water molecules. The enzyme-inhibitor interaction energy was approximated as $\sum V_{enz}(r_a)q_a$, where the first term is the enzyme electrostatic potential at the position of atom a and the second term is the charge on this atom. The potential was calculated from bond fragments as proposed previously while net charges were obtained from CNDO/2 calculations. The hydration energy was calculated by the same expression, where V_{enz} was replaced by the potential of the four water molecules. It was found that the experimental Gibbs free energy of association, ΔG_{exptl}, depends linearly on the calculated interaction energy. A similar linear relationship was observed between the hydration energy and ΔG_{exptl}, indicating that the enzyme may be treated as a "supersolvent". The "electrostatic lock" representing the active site of the enzyme is characterized, and the "key" (the charge pattern of the inhibitor) is found to fit into this lock. With use of this model as a guide, simple structure-activity relationships are derived which may be extended to other enzymes, like thrombin and plasmin.

In the search for potent bioactive compounds it is especially important to understand the nature of the interaction between biological macromolecules (enzymes, receptors) and small ligands (inhibitors, drugs). A promising way is combination of the complete three-dimensional structure of the biomacromolecule-ligand complex with theoretical energy calculations.[1-8] Owing

0002-7863/84/1506-4584$01.50/0

Table I. Substituent (XH_n or XO_2) Coordinates (pm) in the Coordinate System of Protein Data Bank

Benzamidine Coordinates

C1 (−167, 1442, 1669), C (−205, 1575, 1682), H2 (−260, 1620, 1609), C3 (−168, 1648, 1795), H3 (−196, 1745, 1904), C4 (−92, 1587, 1895), H4 (−66, 1639, 1976), C5 (−54, 1453, 1882), H5 (0, 1409, 1954), C6 (−92, 1381, 1769), H6 (−64, 1284, 1759), C7 (−207, 1364, 1547), N8 (−290, 1418, 1458), H8 (−317, 1366, 1376), H8′ (−326, 1511, 1472), N9 (−160, 1241, 1527), H9 (−187, 1189, 1444), H9′ (−98, 1200, 1596)

Hydrogen Coordinates of Bound Water Molecules

(a) W702 (−55, 1814, 2358), (−201, 1835, 2304), W710 (203, 1726, 2031), (157, 1879, 1967)

(b) W702 (−171, 1705, 2358), (−47, 1801, 2354), W710 (−129, 1876, 1985), (216, 1746, 2001)

XH_n or XO_2	X	H1 or O1	H2 (and H3) or O2
NH_2	(−53, 1662, 2013)	(7, 1747, 1983)	(−143, 1696, 2065)
OH	(−54, 1661, 2011)	(6, 1746, 1981)	
CH_3	(−52, 1665, 2017)	(8, 1750, 1987)	(−142, 1699, 2069)
F	(−55, 1660, 2009)		
Cl	(−45, 1678, 2038)		
NO_2	(−54, 1661, 2010)	(26, 1596, 2116)	(−95, 1812, 2024)

to formidable computational problems, theory cannot yield quantitative results which could, in principle, substitute experimental measurements, but, we feel, this is by no means a desirable goal. It is better to construct simple qualitative models which are based on a firm quantitative basis and may serve as an aid to designing new compounds possessing the desired activity. "Design" means here intuition combined with rigorous models in order to obtain useful working hypotheses for experiments. Enzyme–ligand interaction, especially competitive inhibition, is pictorially modeled by the "lock-and-key" or "hand-and-a-glove" analogies.[9] In the former the rigid macromolecule serves as a template where the ligand fits more or less precisely while in the latter both lock and key may relax through conformational changes to reach the optimum fit.

The illustrative lock-and-key analogy has no direct counterpart in a strict theoretical treatment of the enzyme–ligand interaction.[10] The energy of this latter can be partitioned into electrostatic exchange repulsion, inductive, charge-transfer, dispersion, and other terms and is only a fraction of the total Gibbs free energy of association measured in the biophase. Further, very important, contributions are the solvation energy change upon binding and the entropy term. On the other hand, a somewhat arbitrary classification can also be given: fitting the key into the lock requires roughly two conditions to be met, geometric and electrostatic. The former should account mainly for exchange repulsion and solvation (hydrophobic) effects while the latter is of purely classical origin. Other effects (especially entropy) are neglected. The geometry condition means that if the ligand fills the enzyme pocket better, its binding affinity gets larger. This principle, combined with elegant computer graphics, has been exploited recently to design potent thyroid hormone analogues.[7] The electrostatic condition means that the ligand charge distribution has to find its counterpart at the binding site in order to allow maximum interaction with the biomacromolecule.

In this paper we treat substituted benzamidinium inhibitors of β-trypsin. The three-dimensional structure of the β-trypsin–benzamidinium complex is known from X-ray crystallographic studies.[11] We selected seven benzamidine derivatives, 4-NH_2, 4-OH, 4-H, 4-Me, 4-F, 4-Cl, and 4-NO_2 in their protonated forms, since for these compounds extensive experimental work was done.[12–16] Data for the minimum inhibitory concentrations, coming from different laboratories, are roughly in accord; therefore we have a reliable activity data set. Conformational changes are negligible; therefore the application of the "lock-and-key" model seems to be allowed: to predict activity we do not need to consider conformational energy changes. We applied the electrostatic condition since geometry differences among various derivatives are so small that they are insufficient to explain the considerable change in the inhibitory power.

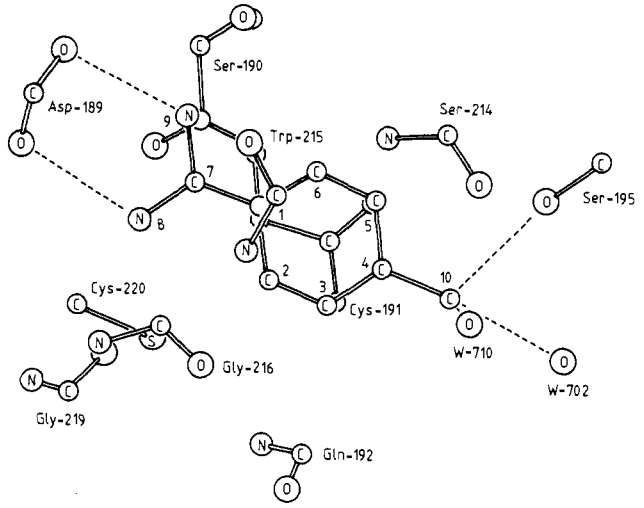

Figure 1. Schematic plot of the active site of β-trypsin with the bound inhibitor. Hydrogen atoms are omitted. Dotted lines represent hydrogen bonds.

In the following we apply a simple method to calculate the electrostatic energy between the enzyme and its inhibitor. On the basis of these calculations physically well-defined models will be proposed which reproduce activity trends quantitatively. Using these quantitative calculations as a basis we qualitatively characterize the "electrostatic lock" which is formed at the active site of β-trypsin. This mapping allows us to derive simple rules which may be used to classify various compounds into more active or less active groups.

(1) Platzer, K. E. B.; Momany, F. A.; Scheraga, H. A. *Int. J. Pept. Protein Res.* **1972**, 201.
(2) Warshel, A.; Levitt, M. *J. Mol. Biol.* **1976**, *103*, 227.
(3) Pincus, M. R.; Zimmerman, S. S.; Scheraga, H. A. *Proc. Natl. Acad. Sci. U.S.A.* **1977**, *74*, 4261.
(4) DeTar, D. F. *J. Am. Chem. Soc.* **1981**, *103*, 107.
(5) Höltje, H. D.; Simon, H. In "Steric Effects in Biomolecules"; Náray-Szabó, G., Ed.; Akadémiai Kiadó-Elsevier: Budapest–Amsterdam, 1982; p 359.
(6) Blaney, J. M.; Weiner, P. K.; Dearing, A.; Kollman, P. A.; Jorgensen, E. C.; Oatley, S. J.; Burridge, J. M.; Blake, C. C. F. *J. Am. Chem. Soc.* **1982**, *104*, 6424.
(7) Blaney, J. M.; Jorgensen, E. C.; Connolly, M. L.; Ferrin, T. E.; Langridge, R.; Oatley, S. J.; Burridge, J. M.; Blake, C. C. F. *J. Med. Chem.* **1982**, *25*, 785.
(8) Wipff, G.; Dearing, A.; Weiner, P. K.; Blaney, J. M.; Kollman, P. A. *J. Am. Chem. Soc.* **1983**, *105*, 997.
(9) Metzler, D. E. "Biochemistry"; Academic Press: New York–San Francisco–London, 1977.
(10) Zahradník, R.; Hobza, P.; Sauer, J. In "Steric Effects in Biomolecules"; Náray-Szabó, G., Ed.; Akadémiai Kiadó-Elsevier; Budapest–Amsterdam, 1982; p 327.
(11) Bode, W.; Schwager, P. *J. Mol. Biol.* **1975**, *98*, 693.
(12) Mares-Guia, M.; Nelson, D. L.; Rogana, E. *J. Am. Chem. Soc.* **1977**, *99*, 2331.
(13) Markwardt, F.; Landmann, H.; Walsmann, P. *Eur. J. Biochem.* **1968**, *6*, 502.
(14) Markwardt, F.; Walsmann, P.; Kazmirowski, H. G. *Pharmazie* **1969**, *24*, 400.
(15) Markwardt, F.; Landmann, H.; Walsmann, P. *Pharmazie* **1970**, *25*, 551.
(16) Markwardt, F.; Walsmann, P.; Stürzebecher, J.; Landmann, H.; Wagner, G. *Pharmazie* **1973**, *28*, 5.

Table II. Electrostatic Potentials and Net Charges at Atoms of Various Benzamidinium Inhibitors

	−V, kJ/mol					net charge, millielectron						
		$2H_2O^a$		hydrate[b]								
atom	enzyme	a	b	c	d	NH_2	OH	H	CH_3	F	Cl	NO_2
H8	836	13	14	252	242	180	181	181	179	184	183	182
H8′	666	18	14	299	286	191	191	192	191	193	192	196
N8	668	17	16	209	197	−189	−189	−188	−189	−187	−188	−183
H9	773	13	12	254	243	180	181	181	179	184	183	182
H9′	561	17	14	301	287	191	191	192	191	193	192	196
N9	632	17	14	210	197	−189	−189	−188	−189	−187	−188	−183
C7	581	20	19	178	163	372	373	372	371	375	374	374
C1	468	31	29	129	107	−51	−50	−36	−51	−43	−44	−11
C2	450	37	37	110	86	51	55	41	48	56	52	37
H2	497	28	30	182	163	10	13	11	9	19	15	24
C3	401	60	61	69	32	−29	−31	10	−16	−23	−13	30
H3	376	56	61	49	13	40	48	37	34	62	51	67
C4	415	82	80	77	22	160	200	54	103	251	165	112
H4	428	113	111	95	5			38				
C5	422	51	45	74	33	−23	−26	10	−16	−23	−13	30
H5	437	43	33	58	15	44	51	37	35	62	51	67
C6	428	34	30	113	87	52	55	41	50	56	53	57
H6	433	26	21	187	165	10	13	11	9	19	15	24
X10	433	124	122	118	−3	−201	−228		−34	−192		299
Cl10	434	127	124	138	−9						−82	
H10	440	271	256	135	−7	101	158		38			
H10′	376	110	114	115	−7				38			
H10″	496	82	70	115	−7	101			22			
O1	495	69	48	112	−23							−246
O2	385	95	112	181	−23							−246

[a] (a) hydrogen locations as in (a), Table I, middle section; (b) hydrogen locations as in (b), middle section. [b] (c) hydration model as in Figure 2, position a; (d) hydration model as in Figure 2, position b.

Model and Method

We used three-dimensional coordinates of the β-trypsin–benzamidinium complex as obtained by X-ray crystallographic measurements[11] and given in the Protein Data Bank.[17] Coordinates of enzyme and benzamidinium hydrogen atoms were generated by the PROTPOT[18] and XANADU[19] programs by using standard C–H, N–H, and O–H bond lengths and orientations. Rotating amino and hydroxyl substituents were oriented in staggered positions with respect to the directly linked groups. The only exception is the hydroxyl group of the active Ser-195 side chain which we rotated to reach an optimum hydrogen-bonding position with $N^{\epsilon 2}$ of His-57. This allows at the same time the favorable interaction between the positively charged benzamidinium ion and oxygen lone pairs. Similarly, the OH, NH_2, and CH_3 substituents, considered as proton donors, were rotated into an optimum hydrogen-bonding position with the oxygen atom of Ser-195. Substituent coordinates are given in Table I.

X-ray diffraction measurements located two bound water molecules in the vicinity of the inhibitor, W702 and W710.[11] These were also considered in the present study. Since hydrogen positions are uncertain, we used two slightly different coordinate sets to have an impression on the influence of orientation. Both arrangements involve hydrogen bonds with the benzamidine substituent and neighboring amino acid residues. Coordinates are given in Table I. Graphical representation of atoms in the close environment of the inhibitor is given in Figure 1.

To account for the solvent effect we modeled the first hydration shell around the inhibitor (Figure 2). To ensure comparability of hydration energies for all substituted derivatives four water molecules were considered in the hydration sphere. The only change was in the orientation near the C4 substituent (X). Positions a and b (Figure 2) can be generated from each other by interchanging hydrogen atoms and lone pairs.

It is an extremely important problem how to consider ionization degree of dissociable side chains (Glu, Asp, Lys, Arg). Experimental[20] and theoretical[21,22] evidence indicates that these are very strongly shielded in the biophase, supposedly by counterions located in their vicinity.[22] To mimic this situation we considered terminal groups of all surface side chains in their un-ionized form (–COOH for Glu and Asp, $-NH_2$ for Lys, and $-NHC(NH)NH_2$ for Arg). On the other hand, we treated two buried aspartates (Asp-102 and Asp-189) as negatively charged, giving thus a total charge of 2– to the enzyme. Asp-102 is the member of the catalytic triad and is ionized at any pH.[23] It is the His-57 side chain which gets protonated upon acidification,[23] but the determination of inhibitory potencies was done at pH 8.0;[12-16] therefore His-57 has to be considered as neutral. Since the benzamidine and Asp-189 form an ion pair in the enzyme–inhibitor complex,[11] this side chain is also considered as fully charged. Potential values obtained from

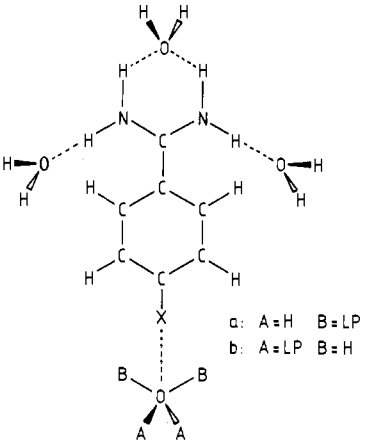

Figure 2. Model of the first hydration sphere around substituted benzamidines. For the water molecule, bound to X, lone pairs are indicated as hypothetical atoms.

(17) 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.
(18) Ángyán, J.; Náray-Szabó, G., Program PROTPOT, Budapest, 1981.
(19) Sheldrick, G., Program XANADU, Cambridge, 1975.
(20) Rees, D. C. *J. Mol. Biol.* **1980**, *141*, 323.
(21) Churg, A. K.; Weiss, R. M.; Warshel, A.; Takano, T. *J. Phys. Chem.* **1983**, *87*, 1683.
(22) Ángyán, J.; Náray-Szabó, G. *J. Theor. Biol.* **1983**, *103*, 349.
(23) Polgár, L.; Halász, P. *Biochem. J.* **1982**, *207*, 1.

Table III. Interaction Energies (kJ/mol) As Calculated by Eq 1 between the Inhibitor and Different Constituents of the Biological System[a]

X	enzyme (1)	W702 + W710[b] (2)	hydration sphere (3)	ΔE_t (1) + (2) − (3)	σ^c	ΔG_{exptl}[d]	ΔG_{exptl}[e]
NH_2	611.9	36.2–40.1	213.9	434.2–438.1	−0.66	28.0	28.1
OH	604.3[f]	44.6–48.1[f]	211.9	437.0–440.5[f]	−0.37	25.0–28.5	24.6
	613.2[g]	14.7–18.2[g]		416.0–419.5[g]			
H	604.9	26.9–28.4	215.9	415.9–417.4	0.0	26.1–26.7	25.4
CH_3	604.6	33.8–36.4	216.7	421.7–424.3	−0.17	24.7–26.0	25.8
F	606.9	15.2–16.7	211.3	410.8–412.3	0.06	23.3–25.4	
Cl	606.0	21.1–22.5	212.2	414.9–416.4	0.23	22.4–26.0	25.1
NO_2	601.0	29.9–32.3	219.2	411.7–414.1	0.78	19.6–20.6	19.8

[a] Position of rotating OH and NH_2 groups is optimized. [b] For two reasonable orientations of water molecules (cf. Table I). [c] Hammett substituent constant (ref 31). [d] Experimental Gibbs free energy of association (ref 12). [e] Experimental Gibbs free energy of association (ref 13–16). [f] H in H10 position (cf. Table II). [g] H in H10″ position (cf. Table II).

this model are given in Table II.

The enzyme–inhibitor interaction energy was approximated as

$$E_{int} = \sum_a V_{enz}(r_a) q_a \qquad (1)$$

where $V_{enz}(r_a)$ is the enzyme electrostatic potential as calculated by the PROTPOT program[18] and r_a and q_a are the position vector and CNDO/2 net charge of atom a of the inhibitor. The same expression was used to calculate the interaction energy with the hydration shell. In this case V_{enz} was replaced by V_{hydr} which was calculated by the ELPO program.[24] PROTPOT and ELPO calculate the electrostatic potential from transferable bond fragments as proposed previously,[25,26] thus the computational work becomes proportional only to the first power of the number of bonds in the molecule. This allows treatment of such a large system as β-trypsin in full.

Though our aim in this paper is not to obtain quantitative interaction energies for the enzyme–substrate interaction we discuss the applicability of eq 1 in some detail. It is known that an exact theoretical calculation of nonbonded interaction energies between molecular species is an extremely difficult task.[10] Even at the ab initio level a large basis set has to be used which makes calculations impracticable for larger systems. Semiempirical zero differential overlap methods, like CNDO/2, are known to overestimate association energies and to underestimate equilibrium hydrogen-bond distances.[27,28] However, dipole moments for molecules with first-row atoms are correctly reproduced.[29] From this we can conclude that, at fixed distances, trends in electrostatic interaction energies, like for the ion pair in our case, will be adequately described. Furthermore, we have shown in a recent paper that eq 1 correctly reflects trends in exactly calculated CNDO/2 interaction energies for the $(H_3N\cdots HF)\cdot 6H_2O$ system.[30] On the basis of this argument we may state that if a linear relation between interaction energies, calculated by eq 1 and obtained by experiment, will be found, it should not be considered as an artifact without theoretical justification.

Results

Interaction energies of eq 1, as calculated for different parts of the biological environment, are given in Table III. The range for the interaction energy with the bound water molecules, W702 and W710, represents the two orientations specified in Table I. The corresponding potential values are given in columns a and b in Table II. The hydration energy was calculated by using the data set of column c or d in Table II. We selected always the larger of the two values; this is given in Table III.

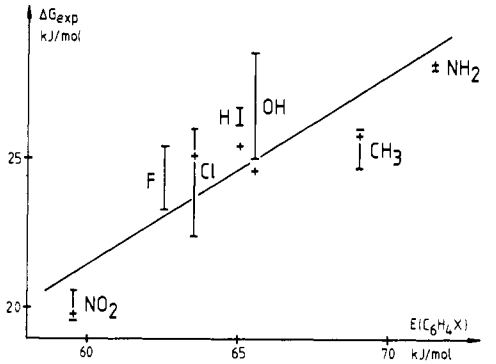

Figure 3. Plot of the experimental Gibbs free energy of association, ΔG_{exptl}, as a function of the calculated electrostatic interaction energy of eq 1 (cf. Table III).

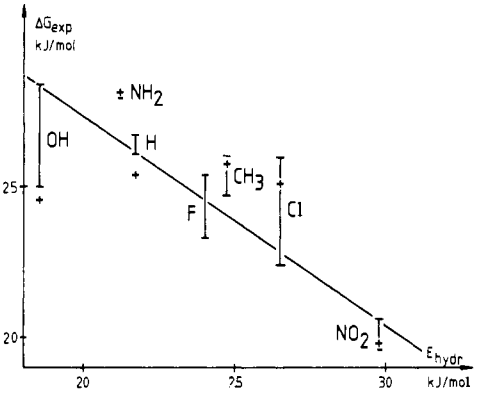

Figure 4. Plot of the experimental Gibbs free energy of association as a function of the calculated hydration energy (cf. Table III).

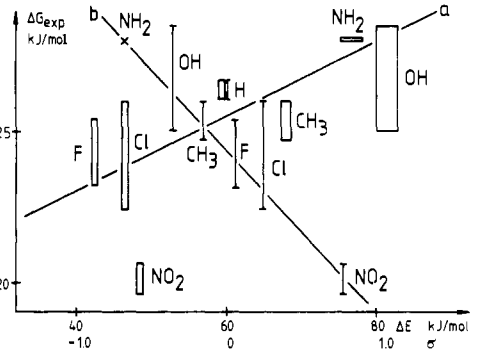

Figure 5. Plot of the experimental Gibbs free energy of association (a) as a function of the net binding energy, $E_t = E(\text{enzyme}) + E(\text{W702} + \text{W710}) - E(\text{hydration})$, and (b) as a function of the Hammett substituent constant. The width of rectangles represents the range given in Table III, column 2.

Experimental Gibbs free energies of association are plotted vs. enzyme–inhibitor interaction energies, hydration energies, and

(24) Náray-Szabó, G. *QCPE* **1980**, *13*, 396.
(25) Náray-Szabó, G. *Int. J. Quant. Chem.* **1979**, *16*, 265.
(26) Náray-Szabó, G.; Grofcsik, A.; Kósa, K.; Kubinyi, M.; Martin, A. *J. Comput. Chem.* **1981**, *2*, 58.
(27) Kollman, P. A.; Allen, L. C. *J. Am. Chem. Soc.* **1970**, *92*, 753.
(28) Gregory, A. R.; Paddon-Row, M. N. *J. Am. Chem. Soc.* **1976**, *98*, 7521.
(29) Pople, J. A.; Beveridge, D. L. "Approximate Molecular Orbital Theory"; McGraw-Hill: New York, 1970.
(30) Ángyán, J.; Náray-Szabó, G. *Theoret. Chim. Acta* **1983**, *64*, 27.
(31) McDaniel, D. H.; Brown, H. C. *J. Org. Chem.* **1958**, *23*, 420.

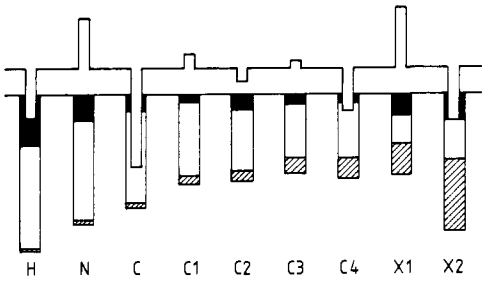

Figure 6. Graphic representation of the "electrostatic lock" along the substituted benzamidine molecule. Notice how the upper charge pattern (the "key" for the OH derivative) fits into the lower electrostatic one. Columns, corresponding to negative potential and positive charge values, point downward. Full areas correspond to the enzyme potential if the Asp-189 and Ser-195 side chains are excluded. Empty areas represent the contribution of the latter while the fraction, originating from W702 and W710, is shaded.

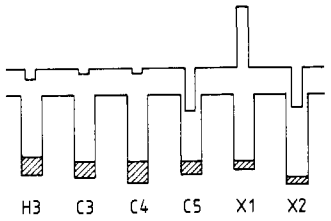

Figure 7. Same as in Figure 6 for the 3 substituent. The enzyme potential without Asp-189 and Ser-195 contribution is not given.

the calculated ne energy of association in Figures 3, 4, and 5, respectively. Also included in Figure 5 is the ΔG_{exptl}–Hammet σ plot.

Discussion

It is seen from Table III that there is a good correlation between calculated and experimental interaction energies (Figure 3).

A linear dependence is observed between ΔG_{exptl} and the hydration energy of the substituted phenyl ring (Figure 4) which may be possible only for closely related molecules like our benzamidine series. This is probably a manifestation of the idea of Warshel,[32] who stated that enzymes act as "supersolvents" for substrates and inhibitors. If the analogy is complete, enzyme binding and hydration energies should run parallel as in Figure 4.

Figure 5 shows the relation between $\Delta E_t = E(\text{enzyme}) + E(\text{W702,W710}) - E(\text{hydration})$ for the substituted phenyl moiety and ΔG_{exptl}. The correlation is fair except for the nitro derivative. This is due to the closeness of W710 to the O2 atom of the nitro group (245 pm). The electrostatic approximation of eq 1 does not take exchange repulsion into account; therefore the interaction energy between W710 and the inhibitor molecule is too large for the 4-NO_2 derivative. As is seen from Figure 5, an excellent linear dependence exists between Hammett substituent constants and experimental Gibbs free energies. This is in accord with previous results.[33,34]

It is clear from the above discussion and from the linear dependence of ΔG_{exptl} on the calculated association energy that the electrostatic model works well in our case. The next step is to map the binding pocket. Figures 6 and 7 depict the change of the electrostatic potential along the benzamidine molecule in two, approximately perpendicular, directions. *These figures represent the electrostatic lock.* It is seen how the charge pattern of the inhibitor molecule fits into the lock ensuring optimum interaction. Structural water molecules, W702 and W710, produce a potential which increases the bonding ability of the whole system; the effect is especially strong at substituent atoms. Figure 7 depicts the potential change in the region of the substituent in the 3 position. This is similar to Figure 6: a decrease along the substituent atoms is observed, though it is much less marked than for the molecule substituted in the 4 position. The effect of W702 and W710 is not very important.

Table IV. pK_i Values for Substituted Benzamidine Derivatives (K_i in mol/L)[34]

	β-trypsin		plasmin		thrombin	
X	3-X	4-X	3-X	4-X	3-X	4-X
	First-Order Substituents (S_I)					
CH_3	4.495	4.523	3.638	3.523	3.377	3.444
OH	4.481	4.301		3.301	4.092	3.658
OCH_3	4.602	4.495	3.959	3.639	3.569	3.377
OC_2H_5	4.569	4.000	3.770	3.347	3.745	3.347
	Second-Order Substituents (S_{II})					
NO_2	3.887	3.482	3.222	2.658	3.263	2.149
$COOCH_3$	4.056	3.523	3.284	3.301	3.215	
$COOC_2H_5$	4.187	3.699	3.301	3.097	3.377	2.666
$COCH_3$	4.620	3.495	3.721	2.658	3.377	2.569
$CONHCH_3$	3.824	3.854	3.125	2.921	2.726	2.482

Our PROTPOT program offers the possibility of analyzing the protein electrostatic potential in terms of residual or side-chain contributions. To study the effect of distant environment on the electrostatic lock we excluded first the Asp-189 and Ser-195 side chains from the model. The potential values at atomic positions, given in Figure 6, are −296, −215, −184, −174, −198, −191, −203, −254, −296, and −288 kJ/mol, respectively. In another model we excluded residues depicted in Figure 1, i.e., Asp-189, Ser-190, Cys-191, Gln-192, Ser-195, Ser-214, Trp-215, Gly-216, Gly-219, and Cys-220, from the model. The potential values for the above positions are −129, −134, −142, −159, −164, −186, −202, −242, −270, and −262 kJ/mol, respectively. Finally, the additional exclusion of Asp-102 yielded −20, −18, −13, −17, −26, −38, −35, −64, −74, and −69 kJ/mol, respectively. It is clear from these calculations not only that neighboring side chains determine the electrostatic lock but also that the effect of distant residues is important. Certainly, the charged Asp-102 side chain is very important: it endows the potential pattern with a decreasing character in the vicinity of C4, X10, and H10″. However, this character is conserved even if Asp-102 is excluded from the protein model. The present result is in accord with our previous studies indicating the important role of protein electrostatic potential in enzymatic processes.[36]

The graphic representation of the electrostatic lock in Figures 6 and 7 allows derivation of some simple and general rules which may be used to predict relative inhibitory potencies of different substituted derivatives. Table IV includes activity data for some selected substituents in the 4 and 3 position. Each substituent has at least two atoms: one, X1, linked directly to C4 of the phenyl ring and another, X2, bound to X1. Furthermore, the X1X2 bond is polar. Two classes are defined: S_I where X1 is the more negative (*CH*H_2, *OH*, *OC*H_3, *OC*H_2*C*H_3) and S_{II} where X1 is the more positive (*NO*O, *CO*O*C*H_3, *CO*O*C*$_2H_5$, *CO*CH_3, *CO*NH*C*H_3); X1 and X2 are represented by italic type. The classes correspond approximately to first- and second-order substituents well-known in organic chemistry, and they are located in Table IV in the top and bottom sections, respectively.

On the basis of Figures 6 and 7 it is easily seen that S_I type substituents bind stronger than S_{II} ones since the ∓ charge distribution is attracted by the descending potential pattern. For the opposite ± distribution repulsion occurs. Clearly, the very strong attraction of the benzamidinium substituent by Asp-189 determines binding, but substituent effects are primarily due to this difference. Furthermore, interaction of the substituent in the 3 position with the protein electrostatic potential is much weaker than in the 4 position. This difference is even more marked if bound water molecules are also considered. Therefore the in-

(32) Warshel, A. *Acc. Chem. Res.* **1981**, *14*, 284.
(33) Coats, E. A. *J. Med. Chem.* **1973**, *16*, 1102.
(34) Labes, D.; Hagen, V. *Pharmazie* **1979**, *34*, 649.
(35) Henkel, H. J.; Labes, D.; Hagen, V. *Pharmazie* **1983**, *38*, 342.
(36) Náray-Szabó, G.; Bleha, T. In "Molecular Structure and Conformations: Recent Advances"; Csizmadia, I. G., Ed.; Elsevier: Amsterdam, 1982; p 267.

Table V. Exceptions to Eq 2 As Compared to the Total Number of Possible Relations in Table IV

relation	β-trypsin	plasmin	thrombin
$S_I^4 \geq S_I^3$	12/16	11/12	12/16
$S_I^4 \geq S_{II}^3$	6/20	4/20	2/20
$S_I^4 \geq S_{II}^4$	0/20	0/20	0/16
$S_I^3 \geq S_{II}^3$	4/20	1/15	0/20
$S_I^3 \geq S_{II}^4$	0/20	0/15	0/16
$S_{II}^3 \geq S_{II}^4$	1/25	3/25	0/20
$S_I \geq S_{II}$	10/80	5/70	2/72

hibitory potency of various substituted derivatives can be ordered as follows:

$$S_I^4 \geq S_I^3 \geq S_{II}^3 \geq S_{II}^4 \quad (2)$$

Upper indices denote substituent positions. Table V illustrates the validity of eq 2. There is only one relation, $S_I^4 \geq S_I^3$, where it is not fulfilled. The less rigorous $S_I \geq S_{II}$ classification (with no respect to substituent position) is valid in 90% of the possible comparisons.

Table V includes data for plasmin and thrombin, too. The three-dimensional structure of these enzymes, also members of the serine proteinase family, is not known at present. However, using computer-generated models, Furie and co-workers have shown that the internal structure and active site of thrombin are similar to β-trypsin.[37] This should mean that the protein electrostatic potential near the active site is also similar in these enzymes;[22] i.e., eq 2 holds also for thrombin. Tables IV and V show that this is the case. Furthermore, eq 2 is fulfilled for plasmin, another important serine proteinase. This indicates that the plasmin active site and "electrostatic lock" should be similar to those of β-trypsin and thrombin.

(37) Furie, B.; Bing, D. H.; Feldmann, R. J.; Robinson, D. J.; Burnier, J. P.; Furie, B. C. *J. Biol. Chem.* **1982**, *257*, 3875.

Conclusions

The present paper describes the quantum chemical calculation of the enzyme–ligand interaction energy for a series of substituted benzamidine inhibitors of β-trypsin. To our knowledge this is the first study which considers the protein in full; others either use empirical methods or model the whole system by a limited number of amino acid residues.[1-8] Our main conclusions are the following.

(1) The experimental Gibbs free energy of association changes parallel with the calculated interaction energy. A similar, linear, dependence is observed between ΔG_{exptl} and the hydration energy.

(2) The "electrostatic lock", representing the active site, is characterized. The electrostatic potential of the enzyme is lower at the amidine and C4 substituents while it is larger along the phenyl ring. The charge pattern of the inhibitor fits into this lock as a key.

(3) By use of the electrostatic lock as a guide simple structure–activity relationships are derived. For X1X2Z type substituents, where X1 is directly attached to the ring and X2 to X1, it is the direction of the X1X2 polarity which determines activity. If X1 is the more negative (OH, OCH_3, etc.), the inhibitory activity is larger than in the case where X1 is the more positive (NO_2, $COCH_3$, etc.).

Registry No. 4-Aminobenzamidinium, 57867-44-4; 4-hydroxybenzamidinium, 57867-43-3; benzamidinium, 53356-58-4; 4-methylbenzamidinium, 57867-49-9; 4-fluorobenzamidinium, 57867-48-8; 4-chlorobenzamidinium, 57867-47-7; 4-nitrobenzamidinium, 57867-51-3; β-trypsin, 9002-07-7.

Flavin and 5-Deazaflavin Photosensitized Cleavage of Thymine Dimer: A Model of in Vivo Light-Requiring DNA Repair

Steven E. Rokita*[1] and Christopher T. Walsh

Contribution from the Departments of Chemistry and Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139. *Received December 29, 1983*

Abstract: This paper presents a model photochemical system for the cleavage of the cyclobutane ring of *cis,syn*-thymine dimer and the production of thymine. Many characteristics of this system mimic the in vivo thymine dimer repair phenomenon attributed to the action of photoreactivation enzymes. Lumiflavin, 5-deazariboflavin, and 8-methoxy-7,8-didemethyl-N^{10}-ethyl-5-deazaflavin will each sensitize dimer cleavage. This cleavage depends on the concentration of both thymine dimer and sensitizer, on the extent of irradiation, and on the wavelength of irradiation. Maximum dimer cleavage occurs when the wavelength of irradiation corresponds to the longest wavelength λ_{max} of the sensitizers. This model system also has a distinct pH dependence; dimer cleavage requires a pH of greater than 10. Initial characterization of the mechanism for dimer cleavage catalyzed by either flavin or 5-deazaflavin is also presented and compared to previously described model systems. A derivative of 8-hydroxy-5-deazaflavin, the chromophore of the *Streptomyces griseus* photoreactivation enzyme, is not able to sensitize thymine dimer cleavage under any conditions presented here. Other electron-rich flavins presented in this report are also unable to catalyze dimer cleavage under model conditions.

Intrastrand *cis,syn*-thymidine dimers are the predominant lesion in UV-irraiated DNA.[2,3] Two distinct mechanisms are used by organisms to repair this lesion. The well-characterized "dark" repair system depends on the action of a dimer-specific endonuclease followed by DNA polymerase.[4] The second system requires the presence of visible light to essentially reverse the adduct forming [2 + 2] cycloaddition of adjacent thymidines. Although this photoreactivation actvity was discovered over 30 years ago,[5] specific proteins (photoreactivation enzyme (PRE),

(1) Current address: Rockefeller University, New York, NY 10021.
(2) Hariharan, P. V.; Cerutti, P. A. *Biochemistry* **1977**, *16*, 2791.
(3) Wang, S. Y., Ed. "Photochemistry and Photobiology of Nucleic Acids"; Academic Press: New York, 1976; Vol. 1 and 2.
(4) Demple, B.; Linn, S. *Nature (London)* **1980**, *287*, 203.
(5) (a) Dulbecco, R. *J. Bacteriol.* **1950**, *59*, 329. (b) For recent review, see: Sutherland, B. M. *Enzymes* **1981**, *24*, 481.

0002-7863/84/1506-4589$01.50/0