Lone-Pair Directionality in Hydrogen Bond Potential Functions for Molecular Mechanics Calculations: The Inhibition of Human Carbonic Anhydrase II by Sulfonamides

Angelo Vedani and Jack D. Dunitz*

Contribution from the Organic Chemistry Laboratory, Swiss Federal Institute of Technology, CH-8092 Zürich, Switzerland. Received April 11, 1985

Abstract: Geometrical analysis of hydrogen bonds observed in crystal structure data retrieved from the Cambridge Crystallographic Data File reveals lone-pair directionality and leads to an extended hydrogen bond potential function for use in molecular mechanics calculations. This potential function includes as variables the hydrogen–acceptor separation, the angle subtended at the hydrogen atom, the angle at the acceptor atom, and the displacement of the hydrogen atom from a defined plane containing the lone-pair orbitals of the acceptor atom. With such directional terms included in the force field, electrostatic contributions to the force-field energy can be reduced or even entirely eliminated without adverse consequences, thus removing many problems associated with the assignment of atomic partial charges and an effective dielectric constant. The new hydrogen bond potential function has been incorporated in the molecular mechanics program YETI and used to refine the conformation of four complexes of human carbonic anhydrase II with small molecules: the natural substrate bicarbonate and three heterocyclic sulfonamide inhibitors.

The Cambridge Crystallographic Data File (CCDF)[1] presently contains the crystal structures of over 40 000 organic compounds and hence provides an excellent opportunity to identify common intra- and intermolecular bonding patterns and derive therefrom features of corresponding potential functions, complementary to theoretical calculations.

In recent publications, using structural data retrieved from the CCDF, Murray-Rust and Glusker[2] and Taylor and Kennard[3] have investigated the stereochemistry of hydrogen bonds in small molecule crystal structures with particular reference to preferred directions at the acceptor molecule. In a similar manner, Baker and Hubbard[4] have analyzed hydrogen-bonding patterns in well-refined protein structures. These studies all clearly showed the importance of two geometrical factors besides the donor–acceptor separation and the angle subtended at the H atom: the angle at the acceptor atom and the displacement of the donor atom from the plane defined by the lone-pair orbitals at the particular acceptor atom. For example, hydrogen bonds involving a ketone O atom as the acceptor (X—H···O=C<, X = O, N) prefer an H···O=C< angle of about 135° with the donor atom X located in the plane of the oxygen lone-pair orbitals.

This paper describes an analogous study of the geometry of hydrogen bonds involving as acceptors aromatic nitrogen and hydroxyl oxygen atoms embedded in several biologically important molecules or molecular fragments. In addition, we have analyzed the stereochemistry of the sulfonamide moiety with respect to hydrogen bonding. The results, including those found in the earlier studies,[2-4] are then used to design an extended hydrogen-bond potential function for use in molecular mechanics programs. The new hydrogen-bond potential function has been incorporated in the molecular mechanics program YETI and used to model details of the binding between the enzyme human carbonic anhydrase II and the natural substrate bicarbonate as well as three potent sulfonamide inhibitors: acetazolamide (2-acetamido-1,3,4-thiadiazole-5-sulfonamide), metazolamide (2-acetimido-3-methyl-1,3,4-thiadiazole-5-sulfonamide), and 2-nitro-thiophene-5-sulfonamide.

Analysis of Hydrogen Bond Geometry

With the information gathered from the CCDF, we have analyzed that the geometry of hydrogen bonds in which imidazole,

serine, threonine, tyrosine, adenine, cytosine, water, and sulfonamide fragments act as hydrogen-bond acceptors. The criteria for accepting an entry from the CCDF were that (1) hydrogen positions should be reported, (2) no atom heavier than potassium should be present in the structure so that C, N, O, and H atoms could be well located, (3) no disorder should be present, and (4) the R factor ($R = \sum ||F_o| - |F_c||/\sum |F_o|$) for the crystal structure determination should be less than 0.08. In our work, we take X–H···Y (X, Y = O, N) arrangements with d(X···Y) < 3.25 Å and (X–H···Y) > 90° as possible hydrogen bonds.

The search in CCDF was performed by defining a specific molecular fragment (program CONNSER) or by searching for a particular class of compounds (program BIBSER). The crystal data of the entries were retrieved with the program RETRIEVE, and those matching the aforementioned requirements were then selected with numeric data screening (program GEOM78). If a particular crystal structure had more than one entry in the database, only the best and/or most recent analysis was accepted. Polymorphic forms of the same compound were treated independently.

The calculation of the geometric parameters (interatomic distances, angles, torsion angles, etc.) was performed with the program GEOM78, which was also used to obtain coordinates for the projection representations of the observed distributions (see Figures 1–4, 6, 8). The programs CONNSER, BIBSER, RETRIEVE, and GEOM78 are part of the CCDF package.[1] In what follows, "Don" signifies the hydrogen-bond donor and "Acc" the hydrogen-bond acceptor.

Aromatic Five-Membered Ring Nitrogen Atom. In this class, 64 examples (adenine, guanine, and imidazole) were found. As expected, they show a strong tendency towards a linear Don–H···Acc arrangement; only for 4 structures is the angle at the H atom less than 150°.

The distribution around the nitrogen acceptor atom is shown in Figure 1. There is a clear preference for the H atom to lie close to the C–N–C bisector, the direction in which the N lone pair should be concentrated. The slight asymmetry in the observed distribution is due mainly to the presence of ring substituents, as in the examples involving purine derivatives. Nevertheless, the deviation from the bisector is less than 30° for 95% of the entries.

Figure 2 shows the displacement of the H atom with respect to the aromatic plane containing the N lone-pair direction. The tendency of the H atom to lie close to this plane is evident. The angular deviation of the N···H vector from the aromatic plane is less than 30° for 88% of the entries.

Aromatic Six-Membered Ring Nitrogen Atom (Excluding Pyridines). In this class, 98 examples (adenine, guanine, and cytosine) were found. Again, the majority shows a tendency toward a linear Don–H···Acc arrangement; for 80 structures, the

(1) Allen, F. H.; Bellard, S.; Brice, M. D.; Cartwright, B. A.; Doubleday, A.; Higgs, H.; Hummelink, T.; Hummelink-Peters, B. G.; Kennard, O.; Motherwell, W. D. S.; Rodgers, J. R.; Watson, D. G. *Acta Crystallogr., Sect. B: Struct. Sci.* **1979**, *B35*, 2331–2339.
(2) Murray-Rust, P.; Glusker, J. *J. Am. Chem. Soc.* **1984**, *106*, 1018–1025.
(3) Taylor, R.; Kennard, O. *Acc. Chem. Res.* **1984**, *17*, 320–326.
(4) Baker, E. N.; Hubbard, R. E. *Prog. Biophys. Molec. Biol.* **1984**, *44*, 97–179.
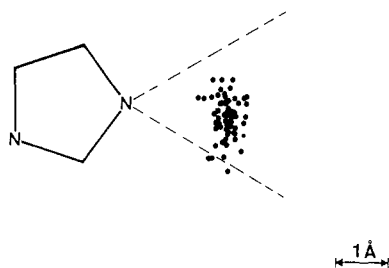
Figure 1. Distribution of donor H atoms around aromatic five-membered ring acceptor nitrogen atoms (top view). Dashed lines are drawn at ±30° from the C–N–C bisector.
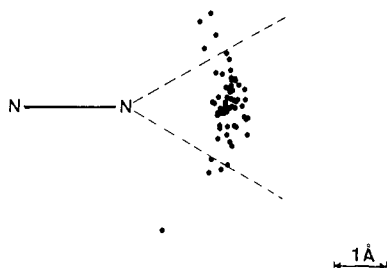

Figure 2. Distribution of donor H atoms around aromatic five-membered ring acceptor nitrogen atoms (side view). Dashed lines are drawn at ±30° from the ring plane, which has not been taken as a plane of symmetry.
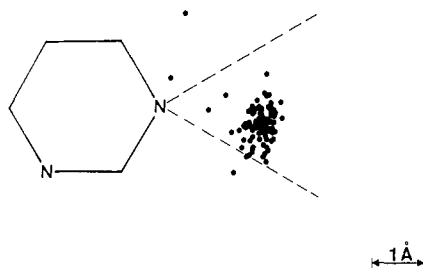

Figure 3. Distribution of donor H atoms around aromatic six-membered ring acceptor nitrogen atoms (top view). Dashed lines are drawn at ±30° from the C–N–C bisector.
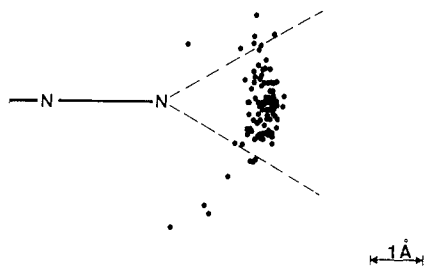

Figure 4. Distribution of donor H atoms around aromatic six-membered ring acceptor nitrogen atoms (side view). Dashed lines are drawn at ±30° from the ring plane, which has not been taken as a plane of symmetry.

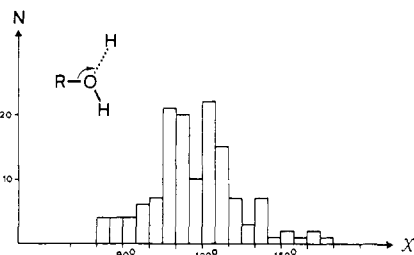**Figure 5.** Distribution of the angle χ (H···O–R) at acceptor hydroxyl O atoms.
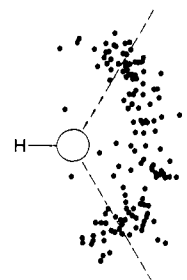

Figure 6. Distributions of donor H atoms around acceptor hydroxyl O atoms in a Newman projection down the C–O bond (H–O bond for H_2O). Dashed lines indicate the "lone-pair directions" at $\omega = \pm 2\pi/3$.
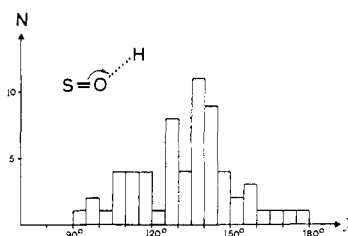

Figure 7. Distribution of the angle χ (H···O–S) at acceptor sulfonamide oxygen atoms.
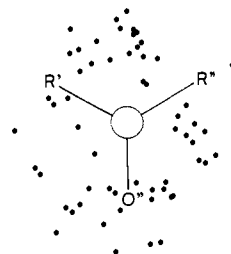

Figure 8. Distribution of donor H atoms around acceptor sulfonamide O atoms in a Newman projection down the S–O_{Acc} bond.

angle at the H atom is within 30° of linearity.

The distribution around the nitrogen acceptor atom is shown in Figure 3. As found for the five-membered rings, there is a clear preference for the H atom to lie close to the C–N–C bisector, the slight asymmetry in the distribution being again attributable mainly to disturbance by ring substituents. The deviation from the bisector is less than 30° for 97% of the entries.

The displacement of the H atom with respect to the aromatic plane is shown in Figure 4. As found for the five-membered rings, there is an evident tendency for the H atom to lie close to this plane. The angular deviation of the N···H vector from the aromatic plane is less than 30° for 86% of the entries.

Hydroxyl Oxygen Atoms as Acceptors. The hydroxyl group can function both as a hydrogen-bond donor and as a acceptor. Most previous studies have concentrated on the donor aspect. Here we examine the behavior as the acceptor. In this class, we found 138 entries (serine, threonine, tyrosine, and water molecules hydrogen-bonded to amino acids). Although no bifurcated (i.e., three-center) hydrogen bonds were found, the linearity of the Don–H···O arrangement is slightly less pronounced than for the nitrogen acceptor fragments: for 29 structures (21%), the angle at the H atom deviates by more than 30° from linearity.

Figure 5 shows the distribution of the angle at the acceptor O atom. It varies from 81° to 166°, with 79% of the entries in the range from 95° to 135°, the mean value being 117°. Figure 6 shows the distribution of the H atoms in a Newman projection down the C–O bond (H–O bond for H_2O). The observed distribution can be viewed as a superposition of two types of hydrogen-bonding pattern at the hydroxyl oxygen: (1) with H roughly in the direction of a tetrahedral sp³ lone-pair lobe (torsion angle ∼ +120° or −120°) and (2) with H roughly in the direction of a trigonal sp² lone-pair lobe (torsion angle ∼180°).

Lone-Pair Directionality in Potential Functions

J. Am. Chem. Soc., Vol. 107, No. 25, 1985 7655

Table I. Parameters for the Hydrogen Bond Potential Function Used in YETI

$$E_{HB} = \left(\frac{A'}{r_{H \cdots A}{}^i} - \frac{C'}{r_{H \cdots A}} \right) \cos^k (\theta_{D-H \cdots A}) \cos^m (\chi_{H \cdots A-AA} - \chi_0) \cos^n (\omega_{H \cdots A-AA-AB} - \omega_0)$$

acceptor type	definition range[a] of θ function	definition range[a] of χ function	optimal angle at acceptor: χ_0	displacement factor ?
carbonyl O		90° < X < 180°	135°	yes
carbonyl O		90° < X < 180°	135°	yes
hydroxyl O	90° < θ < 180°	60° < X < 180°	109.5° (sp³), 120° (sp²)	yes[b]
sulfonamide O		90° < X < 180°	135°	no
aromatic five-membered ring N		72° < X < 180°	126°	yes
aromatic six-membered ring N		60° < X < 180°	120°	yes

[a] If the angle in question lies outside the defined range, the E_{HB} term is set equal to zero. Electrostatic and van der Waals interactions are still computed. [b] sp²: $\omega_0 = 0.0$. sp³: $\omega_0 = \pm 2\pi/3$.

In agreement with results of the analysis by Taylor and Kennard[3] of 196 O–H···O< hydrogen bonds, the distribution of the observed torsion angles in the range $105° < |\omega| < 180°$ is fairly homogenous and thus does not indicate any strong preference of the hydrogen-bond donor for tetrahedral (i.e., sp³) lone-pair directions. The distribution for phenolic hydroxyl groups (i.e., tyrosyl-OH) is not markedly different from that of the class as a whole.

Sulfonamide Oxygen Atom. In this class, we found 49 entries with 62 matching fragments. For 43 entries (69%), the angle subtended at the hydrogen atom is within 30° of linearity.

The distribution of the H···O=S angle at the O acceptor atom (illustrated in Figure 7) varies from 92° to 176°, with 43 structures (69%) in the range from 115° to 155° (mean value, 134°). These results are similar to those obtained by Murray-Rust and Glusker for hydrogen bonding to carbonyl oxygens.[2]

Figure 8 shows the distribution of donor H atoms in a Newman projection down the S–O$_{Acc}$ bond. The apparent absence of any preferred orientation of the H atoms in this projection would seem to be in conflict with the proposal by Platt and Robson[5] of localized lone pairs at sulfonamide oxygens.

Potential Function for Hydrogen Bonds

In molecular mechanics calculations, nonbonded interactions between pairs of atoms are generally represented by a van der Waals and an electrostatic (Coulomb) term. The van der Waals term (e.g., a 12/6 Lennard-Jones potential function) optimizes the separation of two atoms to a preassigned value, the van der Waals distance, usually taken as the sum of some standard set of van der Waals radii such as those of Pauling or Bondi.[6] Since the distance between a pair of atoms involved in a hydrogen bond is often considerably shorter than the van der Waals distance, Gelin and Karplus in 1979 introduced a new term designed to model the interaction between a pair of atoms (Don···Acc) involved in a hydrogen bond[7]

$$E_{HB} = A'/r^{12} - C'/r^{10} \qquad (1)$$

where r is the Don···Acc separation and the coefficients A' and C' depend on various factors such as the donor and acceptor atom type, the equilibrium Don···Acc separation, and the well depth.[8] Although no explicitly directional terms are present, some directionality at the acceptor is introduced by the net effect of the electrostatic and van der Waals interactions of neighboring atoms.

In the molecular mechanics program AMBER,[9] the same type of potential function for hydrogen bonds is used, except that the 12/10 term is assigned to the H···Acc interaction[10] rather than to the Don···Acc one, which is described by a 12/6 Lennard-Jones potential function. This change seems justified, since it is the H···Acc separation that is usually much shorter than the van der

Waals distance ($d_{vdw} \sim 2.60$ Å) and thus calls for special treatment. However, this requires the positions of the H atoms partaking in hydrogen bonding to be known explicitly and also involves a significant increase in the number of variables.

In a slightly different approach, Levitt[11,12] represents the H···Acc interaction by an appropriately parametrized 12/6 term together with a modifying function depending on the angle at the H atom.

Directionality at the acceptor molecule was specifically allowed for by Brooks et al. in the molecular mechanics program CHARMM:[13]

$$E_{HB} = (A'/r^i - C'/r^j) \cos^m (\theta_{D-H \cdots A}) \cos^n (\chi_{H \cdots A-AA}) \qquad (2)$$

The first factor corresponds to the earlier potential function (1) (where $i = 12$, $j = 10$), the second takes into account the nonlinearity at the H atom (Don–H···Acc), and the third involves the direction of the H atom with respect to the molecule (or molecular fragment) in which the acceptor atom resides, specifically the angle χ between the H···Acc direction and the vector from the acceptor to an atom (called the antecedant atom, AA) to which it is covalently bonded. The exponent m (=0, 2, 4) depends on the nature of the donor, while the exponent n (=0, 2) depends on the acceptor.

The validity of (2) rests on the assumption that a linear arrangement is preferred for both the Don–H···Acc angle and (for certain acceptor types) the angle subtended at the acceptor atom. However, it now seems clear that the second assumption is incorrect for most acceptor types. Even for ketone O, the arc-like distribution of χ angles observed in small-molecule crystal structures shows a pair of broad maxima at approximately $|\chi| = 135°$ separated by a shallow saddle point at $\chi = 180°$ (see ref 2, p 1021). For protein structures, the corresponding distributions at the amide carbonyl O are different for α-helices, β-turns, and parallel and antiparallel β-sheets, but none show a maximum at $\chi = 180°$ (see ref 4, p 147). In fact, the preferred angle seems to depend on the type of acceptor atom and on the nature of the molecule in which it is embedded.

While it is true that the hydrogen-bond geometry in any given structure is a compromise among many factors, the new results underline the need for a more versatile expression to allow for variation in the angle at the acceptor atom for different kinds of acceptor molecules. Moreover, there is apparently also a need for an additional factor to allow for the displacement of the H atom from a plane defined by the lone-pair orbitals at the acceptor molecule. To meet all these requirements, a modified and extended version of (2) can be written as

$$E_{HB} = (A'/r^i{}_{H \cdots A} - C'/r^j{}_{H \cdots A}) \cos^k (\theta_{D-H \cdots A}) \times$$
$$\cos^m (\chi_{H \cdots A-AA} - \chi_0) \cos^n(\omega_{H \cdots A-AA-AB} - \omega_0) \qquad (3)$$

The first and second factors correspond to those in (2). The third factor implies that there is an optimal angle χ_0 at the acceptor atom, to be assigned for any given acceptor type. The analysis

(5) Platt, E.; Robson, B. *J. Theor. Biol.* **1982**, *96*, 381–395.

(6) Bondi, A. *J. Phys. Chem.* **1964**, *68*, 441–451.

(7) Gelin, B. R.; Karplus, M. *Biochemistry* **1979**, *18*, 1256–1268.

(8) The total energy of the hydrogen bond is obtained by adding the electrostatic term to (1).

(9) Weiner, P. K.; Kollman, P. A. *J. Comput. Chem.* **1981**, *2*, 287–303.

(10) Wipff, G.; Dearing, A.; Weiner, P. K.; Blaney, J. M.; Kollman, P. A. *J. Am. Chem. Soc.* **1983**, *105*, 997–1005.

(11) Levitt, M. In "Protein Folding"; Jaenicke, R., Ed.; Elsevier-North Holland: Amsterdam, 1980; p 17.

(12) Hall, D.; Pavitt, N. *J. Comput. Chem.* **1984**, *5*, 441–450.

(13) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187–217.

7656 *J. Am. Chem. Soc., Vol. 107, No. 25, 1985*

Vedani and Dunitz

Table II. Force-Field Parameters Used in the Program YETI

atom type	polarizability α, Å^{-3}	effective no. of outer-shell electrons	van der Waals radius, Å
Nonbonded Parameters[a]			
C	1.65	5	1.80
*CH$_{aliphatic}$	1.35	6	1.85
*CH$_{aromatic}$	2.07	6	1.90
*CH$_2$	1.77	7	1.90
*CH$_3$	2.17	8	1.95
H$_{acidic}$	0.04	1	0.80
OH	0.59	7	1.60
O$_{carbonyl}$	0.84	7	1.60
O$_{carbonyl}$	2.14	7	1.60
N	1.15	6	1.65
S	2.00	15	1.90
Zn	0.10	27	0.74

Don$\cdots$Acc	well depth E_{min}, kcal/mol	equilib dist (H$\cdots$Acc), Å
Hydrogen Bond Parameters[b]		
O–H$\cdots$O	4.25	1.79
O–H$\cdots$N	4.00	1.89
N–H$\cdots$O	3.50	1.89
N–H$\cdots$N	3.00	1.99

[a] Nonbonded parameters for calculating the coefficients A and C in the 12/6 Lennard-Jones potential function; adapted from ref 9 and 13. Asterisks indicate extended atom types. [b] Hydrogen bond parameters for calculating the coefficients A' and C' in (3); adapted from ref 13.

of small-molecule structures retrieved from the CCDF suggests that χ_0 is about 135° for carbonyl, carboxyl, and sulfonamide oxygen atoms, 109.5° (sp^3) or 120° (sp^2) for hydroxyl oxygen atoms, 126° for aromatic five-membered ring nitrogen atoms (imidazole), and 120° for aromatic six-membered ring nitrogen atoms (pyrimidine).

The fourth factor allows for the displacement of the H atom from a defined plane. This is done by computing a torsion angle ω involving the H atom and three atoms of the required plane (A, AA, and AB, where AA and AB are selected atoms in the acceptor molecule). The phase-shift parameter ω_0 is set equal to $\pm 2\pi/3$ if tetrahedral sp^3 lone-pair orbitals at the acceptor are believed to be involved (e.g., R–OH, H$_2$O). Where no definite lone-pair directions can be assumed, e.g., sulfonamide O atoms, this factor is taken as unity (i.e., $n = 0$).

The exponents i and j determine the width of the potential valley. In agreement with Weiner and Kollman,[9] we use $i = 12$ and $j = 10$. The exponents k, m, and n allow for acceptor-specific slopes of the penalty function and can be derived from corresponding distribution patterns observed in crystal structures retrieved from the CCDF.[14]

This hydrogen-bond potential function (3) has been incorporated in the molecular mechanics program YETI designed for substrate docking to an extended active site region of a protein.[15] This program searches for a minimum-energy conformation of a small-molecule substrate–protein complex by performing torsional rotations of segments of amino acid side chains of the protein, combined with translation, rotation, and internal torsional motions of the substrate molecule. It uses a steepest-descent minimizer with quadratic-step interpolation.

Apart from the new hydrogen-bond potential function, the YETI force field is largely taken over from those in the programs AMBER and CHARMM. A switching function (from ref 13) for smooth cutoff is assigned to nonbonded interaction terms. It is usually set at 9.5/10.0 Å (switch on/switch off) for electrostatic terms, 7.5/8.0 Å for Van der Waals interactions, and 4.5/5.0 Å (H$\cdots$Acc separation) for hydrogen bonds. The force-field parameters are given in Table II.

(14) For our refinements, we used $k = 2$, $m = 2$, and $n = 2$ (but $n = 0$ for sulfonamide O acceptors). Energy calculations based on [k,m,n] values derived from actual distributions observed in the CCDF (carbonyl and carboxyl O [4,2,2], hydroxyl O [2,2,1], sulfonamide O [2,2,0], and aromatic ring N [4,2,2]) did not lead to significantly different results.

(15) YETI is written in FORTRAN 77, contains about 5500 lines of code, and is currently dimensioned for 400 atoms.

Methods

Our interest has been focused so far on the inhibition of human carbonic anhydrase II by sulfonamide drugs. Carbonic anhydrase, a monomeric metalloenzyme (Zn), consists of 260 amino acids and catalyzes the reversible hydration of carbon dioxide to bicarbonate. The crystal structure of human carbonic anhydrase II (HCA II, the high active isoenzyme) has been determined to a resolution of 2.1 Å by Kannan et al.[16-18] Accurate binding constants for about 30 sulfonamide inhibitors are known,[19-21] thus providing an opportunity for comparison with values based on estimated substrate–protein binding energies.

In a previous study,[22] the conformation of the native enzyme had been obtained from the Brookhaven Protein Data Bank[23] and relaxed to a convergence energy of 0.010 kcal/Mol by using the molecular mechanics program AMBER.[9,24] Atomic coordinates for the small molecules were either retrieved directly from the CCDF or constructed from appropriate molecular fragments. Small molecule substrates were then fitted visually to the receptor site by interactive computer graphics techniques.

In the new calculations with YETI, the refinement of the active site region included 368 atoms of the enzyme and up to 16 atoms of the substrate.[25] When the aforementioned cutoff distances were used, the list of nonbonded pairs of atoms numbered about 23 000. The automatic survey of the environments of the ~ 60 H atoms attached to N or O atoms revealed a total of about 230 possible hydrogen bonds. According to (3), the energy of each of these depends on the relative positions of five atoms (Don, H, Acc, and two acceptor antecedent atoms). Hydrogen atoms attached to C were incorporated within "extended atom types". The convergence criterion was set at 0.001 kcal/mol.[26]

We have performed parallel refinements with different microscopic dielectric constants, $\epsilon = 1.0, 2.0, 4.0$, and 8.0, as well as with distance-dependent dielectric constants $\epsilon = r, 2r$, and $4r$ as proposed by Warshel and Levitt[27] to allow for polarization effects (here r is the dimensionless number equal to the interatomic distance expressed in angstroms). With $\epsilon = 1.0$ and $\epsilon = r$, we obtain unrealistic coordination geometry at the metal center, e.g., Zn$\cdots$N distances of less than 1.75 Å; moreover $\epsilon = 1.0$ (and also 2.0) leads to an overweighting of the electrostatic terms and thus to much poorer hydrogen-bond geometries. Best consistency with known structural data (retrieved from the CCDF) was achieved by using $\epsilon = 4r$.

For comparison, refinements were made with all atomic charges set to zero. The only other change was that the van der Waals radius of Zn was reduced (from 0.74 Å) to 0.55 Å, corresponding to van der Waals' distances Zn$\cdots$O, 1.95 Å, and Zn$\cdots$N, 2.00 Å. Slightly different conformations were obtained for amino acid side chains involved in hydrogen bonding (i.e., His, Gln, Glu, Thr, Tyr, and the small molecule). Compared with the best previous refinements, using partial atomic charges with $\epsilon = 4r$, the hydrogen-bond geometries improved significantly. The coordination at the metal center could be reproduced surprisingly well by the van der Waals potential function alone: mean Zn$\cdots$N = 2.08 (3) Å compared with 2.04 (9) Å using the additional electrostatic term.

Thus, one way of avoiding the problems of assigning atomic partial charges and artificial dielectric constants might be to rely more on parameters directly obtainable from structural data, such as nonbonded distances, coordination geometries at metal centers, and detailed hydrogen-bond geometries. Since partial atomic charges are not in any sense

(16) Kannan, K. K. In "Biophysics and Physiology of Carbon Dioxide"; Bauer, C., Gros, H., Bartels, H., Eds.; Springer: Berlin, 1979; pp 184–205.

(17) "Biology and Chemistry of the Carbonic Anhydrases"; Tashian, R. E., Hewett-Emmet, D., Eds.; New York Academy of Science: New York, 1984.

(18) Lindskog, S.; Henerson, H. E.; Kannan, K. K.; Liljas, A.; Nyman, P. O.; Strandberg, B. In "The Enzymes"; Boyer, P., Ed.; New York Academy of Science: New York, 1971; Vol. V, pp 587–665.

(19) Taylor, P. W.; King, R. W.; Burgen, A. S. V. *Biochemistry* 1970, 9, 2638–2645.

(20) Sprague, J. M. In "Topics in Medicinal Chemistry"; Rabinowitz, J. L., Myerson, R. M., Eds.; Interscience: New York, 1968; pp 1–63.

(21) Kakeya, N.; Aoki, M.; Kamada, A.; Yata, N. *Chem. Pharm. Bull.* 1969, 17, 1010–1018.

(22) Vedani, A.; Meyer, E. F., Jr. *J. Pharm. Sci.* 1984, 73, 352–358.

(23) Bernstein, F.; Koetzle, T. F.; Williams, G. J. B.; Meyer, E. F., Jr.; Brice, M. D.; Rodgers, J. R.; Kennard, O.; Shimanouchi, T.; Tasumi, M. *J. Mol. Biol.* 1977, 112, 535.

(24) These calculations were performed on a VAX 11/780 computer of the Physics Department, Texas A&M University, College Station, TX.

(25) These calculations were performed on a DEC 10/99 and a CYBER 180-825 computer of the ETH Zürich.

(26) This low value was chosen because of the poor convergence behavior of the steepest-descent minimizer used in the current version of YETI. A Newton–Raphson procedure will be implemented in the near future.

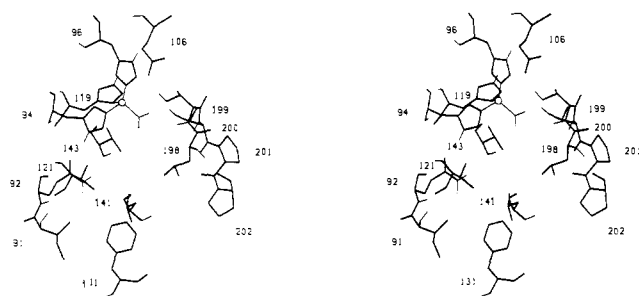(27) Warshel, A.; Levitt, M. *J. Mol. Biol.* 1976, 103, 227–249.

Figure 9. Stereoscopic view of the active site cleft of native human carbonic anhydrase II. The zinc atom is represented by a sphere; bonds from O and N atoms to H or Zn atoms are drawn light.

observable quantities and since their usefulness for such calculations is at least questionable, it is difficult to see what would be lost if they were to disappear from the scene. A recent analysis by Pettitt and Karplus[28] of the role of electrostatics in the structure, energy, and dynamics of *N*-methyl alanylacetamide showed that most structural and dynamic properties are relatively insensitive to the choice of the atomic partial charge model, even for the zero-charge case.

Results

The results reported in the following section were obtained by using atomic partial charges for amino acids given by Blaney et al.[29] and for sulfonamides by Korolkovas et al.,[30] combined with a microscopic dielectric constant $\epsilon = 4r$. The cutoff distances were set at 9.5/10.0 Å (electrostatic), 7.5/8.0 Å (van der Waals), and 4.5/5.0 Å (H-bond; H···Acc separation).

1. Native Human Carbonic Anhydrase II (HCA II). Calculations for native HCA II with a water molecule as the fourth ligand of the zinc atom yielded a slightly distorted tetrahedral coordination at the metal center (see Figure 9): Zn···N 2.16 Å (His 94), 1.95 Å (His 96), 2.05 Å (His 119), Zn···O 1.91 Å (H$_2$O), X···Zn···Y angles 91°, 101°, 106°, 109°, 115°, and 132°. This is in good agreement with the geometries retrieved from the CCDF of small-molecule complexes containing tetra-coordinated zinc (1.89 Å < Zn···O < 2.05 Å and 1.91 Å < Zn···N < 2.14 Å).

The catalytically important water molecule is hydrogen-bonded to the hydroxyl group of Thr 199. An O···O distance of 2.74 Å and an almost linear O–H···O arrangement (167°) are found. Moreover, the H atom is located not too far from the direction of the (sp^3) lone-pair orbital at the hydroxyl O atom (angle at the acceptor = 81°, torsion angle 116°).

2. Natural Substrate Bicarbonate. From NMR and spectral absorbance evidence,[16,31] the hydroxyl O atom of the bicarbonate ion is believed to replace the water molecule as the fourth ligand of the zinc atom and to be hydrogen bonded to the hydroxyl O atom of Thr 199. A second O atom of HCO$_3$⁻ occupies the fifth coordination site of the metal. This general arrangement is assumed in the present work.

The accommodation of the natural substrate by the protein is found to involve some rearrangement of the active site from the native enzyme. The change from a four- to a five-coordinated metal center leads to marked displacements ($d_{RMS} > 0.2$ Å) of the residues His 94, His 96, Leu 198, and the zinc. Distances at the metal center are Zn···N, 2.48 Å (His 94), 1.99 Å (His 96), and 2.22 Å (His 119) Zn···O, 1.92 and 1.97 Å. Thus, compared with the native enzyme, the interactions of His 94 and His 119 with the metal center are weakened significantly. The resulting coordination is close to a tetragonal pyramid, with its base built by the N atoms of His 96 and His 119 and the two O atoms of the substrate. The axial ligand (His 94) appears to be the most weakly bonded. The angles between the axial and the equatorial
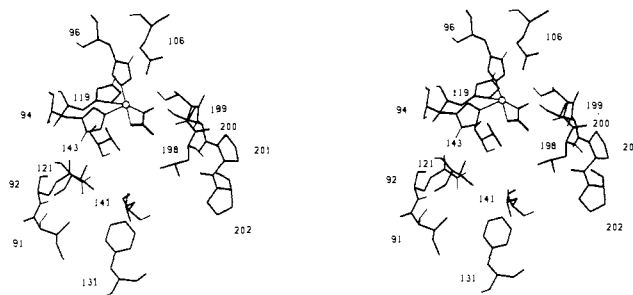
Figure 10. Stereoscopic view of the active site cleft of human carbonic anhydrase II with the natural substrate bicarbonate. The zinc atom is represented by a sphere; bonds from O and N atoms to H or Zn atoms are drawn light. The substrate molecule is drawn enhanced.
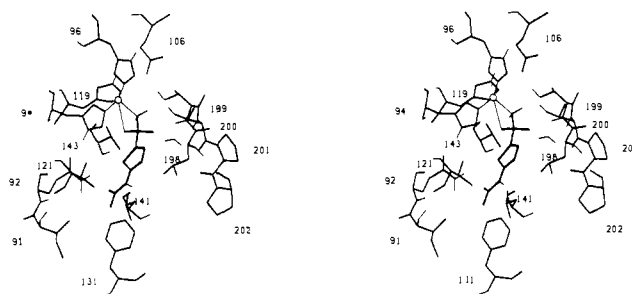

Figure 11. Stereoscopic view of the active site cleft of human carbonic anhydrase II with the inhibitor acetazolamide. The zinc atom is represented by a sphere; bonds from O and N atoms to H or Zn atoms are drawn light. The inhibitor molecule is drawn enhanced.
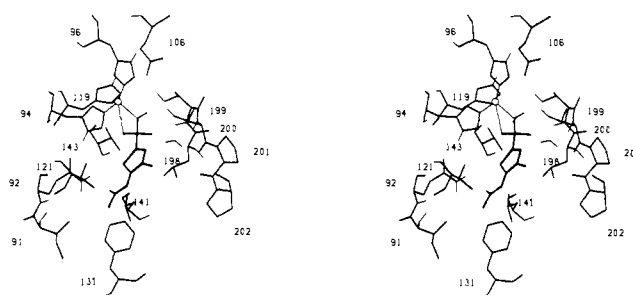

Figure 12. Stereoscopic view of the active site cleft of human carbonic anhydrase II with the inhibitor metazolamide. The zinc atom is represented by a sphere; bonds from O and N atoms to H or Zn atoms are drawn light. The inhibitor molecule is drawn enhanced.

ligands are 78°, 94°, 94°, and 98°; those between opposite equatorial ligands are 154° and 167°. The zinc is located almost exactly in the equatorial plane.[32] The hydrogen bond from the substrate hydroxyl O atom to the hydroxyl O atom of Thr 199 appears to be slightly weaker than the corresponding one found in the native enzyme (H$_2$O···OH Thr 199), although the O···O distance is shorter, 2.63 vs. 2.74 Å. The other geometrical parameters are O–H···O, 141° vs. 167°, angle at the acceptor, 85° vs. 81°, torsion angle, 124° vs. 116°. The orientation of the bicarbonate ion in the active site of human carbonic anhydrase II is shown in Figure 10.

3. Sulfonamide Inhibitors. The orientation of acetazolamide in the active site of HCA II has been investigated by Kannan et al.[16] by X-ray diffraction; a difference map revealed the positions of the two sulfur atoms of the inhibitor molecule. These positions were used to generate the starting orientation of the inhibitor for our molecular mechanics calculations.

(28) Pettitt, B. M.; Karplus, M. *J. Am. Chem. Soc.* **1985**, *107*, 1166–1173.
(29) 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*, 6424–6434.
(30) Korolkovas, A.; Senapeschi, A. N.; Stamato, F. M. L. G.; Da Silva, E. L. *Rev. Pharm. Bioquim.* **1981**, *17*, 13–27.
(31) Yeagle, P. L.; Lochmüller, C. H.; Henkens, R. W. *Proc. Natl. Acad. Sci. U.S.A.* **1975**, *72*, 454–458.

(32) This type of coordination geometry combines features of two frequently observed types: a trigonal bipyramid with the zinc located exactly in the equatorial plane and a tetragonal pyramid with the metal about 0.3 Å above the basal plane. Out of 15 small-molecule structures retrieved from the CCDF, 8 belong to the first type and 6 to the second. We plan to introduce appropriate terms in the potential energy expression to allow for variations from these preferred arrangements.[35]
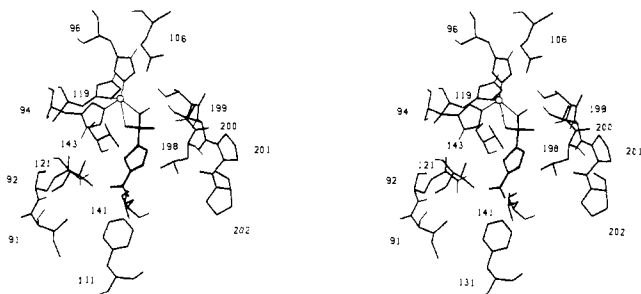
Figure 13. Stereoscopic view of the active site cleft of human carbonic anhydrase II with the inhibitor nitrothiophenesulfonamide. The zinc atom is represented by a sphere; bonds from O and N atoms to H or Zn atoms are drawn light. The inhibitor molecule is drawn enhanced.

We also analyzed the intermolecular environments of the sulfonamide group ($R-SO_2NH_2$) in accurate small-molecule crystal structures using data retrieved from the CCDF. Of the 19 sulfonamide structures with $R < 0.08$ in the database, all 19 act as hydrogen-bond donors through the N atom, 14 accept at least one additional hydrogen bond at an O atom, and 6 have all three heteroatoms hydrogen-bonded. Moreover, the amide N atom avoids hydrophobic contacts (shortest N⋯C = 3.48 Å) more than the O atoms do (shortest O⋯C = 3.19 Å). The orientation proposed by Kannan,[16,33] with the sulfonamide coordinated through N and one O atom to the zinc and the N hydrogen-bonded to the hydroxyl oxygen of Thr 199, is consistent with these results.

Our refinements suggest that the three sulfonamide inhibitors studied in detail (acetazolamide, AAA; metazolamide, MAA; and nitrothiophenesulfonamide, NTS) bind to the active site in a rather similar way (see Figure 11–13).

HETEROCYCLIC SULFONAMIDE INHIBITORS

R - -SO₂NH₂
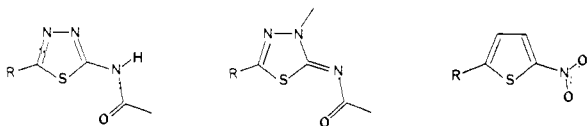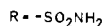
ACETAZOLAMIDE   METAZOLAMIDE   NITRO-THIOPHENE-SA

The sulfonamide N atom occupies the fourth coordination site at the zinc, replacing the zinc-bound water molecule in the native enzyme: Zn⋯N, 2.16 Å (AAA), 2.17 Å (MAA), and 2.19 Å (NTS). One oxygen atom of the sulfonamide group binds weakly to the metal to give a distorted (4 + 1) "tetrahedron": Zn⋯O, 3.30 Å (AAA), 3.24 Å (MAA), and 2.80 Å (NTS). Like the zinc-bound water molecule in the native enzyme, the amide NH_2 forms a hydrogen bond to the hydroxyl oxygen of Thr 199: N⋯O, 2.86 Å (AAA), 2.87 Å(MAA), and 2.84 Å (NTS); N–H⋯O, 172° (AAA), 176° (MAA), and 167° (NTS). For all three complexes, the H atoms are located close to the direction of the (sp³) lone-pair orbital at the hydroxyl O atom. The third O atom is hydrogen-bonded to the main chain –NH– of Thr 199: N⋯O, 3.09 Å (AAA), 3.20 Å (MAA), and 2.93 Å (NTS). The N–H⋯O angles are almost linear: 170° (AAA), 168° (MAA), and 174° (NTS). The H⋯O=S angles at the acceptor sulfonamide O atom are found to be 146° (AAA), 142° (MAA), and 154° (NTS), all three being within 20° of the assumed optimal value ($\chi_0 = 135°$).

The heterocyclic ring is embedded in the hydrophobic lower part of the active site cleft. The interaction of the ring S atom with the side chain of Val 121 seems to be less pronounced than proposed by Kannan[16] for acetazolamide. The S⋯CH_3 distances

of 3.48 Å (AAA), 3.54 Å (MAA), and 3.61 Å (NTS) found in our calculations are in good agreement with structural data retrieved from the CCDF (shortest S⋯CH_3 contact = 3.46 Å). The main interaction of the heterocyclic S atom is clearly of *intra*-molecular S⋯O type. The crystal structure determination of molecular AAA[34] shows that the acetamido ($R-NHCOCH_3$) side chain is almost coplanar with the thiadiazole ring ($\omega = 4.9°$), giving a very short intramolecular contact to the ring S atom (S⋯O, 2.751 Å). For acetimido ($R=NCOCH_3$) side-chains, intramolecular S⋯O contacts as short as 2.519 Å are found in crystal structures. In our calculations, we allowed for such strong intramolecular interactions by assigning a torsional potential function with a barrier height of 1.0 kcal/mol (at $|\omega| = 90°$), replacing the van der Waals term.

We propose that in the complexes of AAA and MAA, a co-ordinated water molecule bridges the thiadiazole ring to the amide O atom of Thr 200. The presence of such a water molecule could explain the different binding strength of AAA and its thiazole analogue ($K_b = 7.1 \times 10^7$ vs. 3.2×10^6 mol^{-1}[20]) where the C atom in the ring 4-position cannot engage in hydrogen bonding. In the AAA and MAA complexes, this H_2O molecule could form two strong hydrogen bonds with calculated parameters:

$H_2O⋯O=C<$	Don⋯Acc: 2.70 Å (AAA), 2.71 Å (MAA)
	angle at H: 161° (AAA), 163° (MAA)
	angle at Acc: 123° (AAA), 122° (MAA)
	out-of-plane: 15° (AAA), 14° (MAA)
$H_2O⋯N_{ring}$	Don⋯Acc: 2.76 Å (AAA), 2.77 Å (MAA)
	angle at H: 161° (AAA), 162° (MAA)
	angle at Acc: 108° (AAA), 106° (MAA)
	out-of-plane: 36° (AAA), 34° (MAA)

As shown in Figure 11 for AAA, the postulated water molecule fits nicely into an interstice between the inhibitor and the main chain of the protein (residues Leu 198 through Pro 201). The hydrogen bond formed by the second sulfonamide O atom with the main chain –NH– of Thr 199 is weakened by the presence of this water molecule: O⋯N = 3.09 Å (AAA), 3.20 Å (MAA), but 2.93 Å for NTH where no such water molecule is present.

So far, the calculated internal binding energies of the enzyme–sulfonamide complexes are at best in qualitative agreement with the experimental determined binding constants. Our model can certainly be improved, however, by inclusion of a solvation shell around the inhibitor, sealing the active site cavity. In addition, we plan to improve our force field by introducing new terms to allow specifically for distortions from frequently occurring types of coordination geometry at metal centers.[35]

Conclusions

Lone-pair directionality of hydrogen bonds, clearly demonstrable in small-molecule crystal structures, needs to be incorporated in force-field calculations that model biological complexes. An appropriate potential function is modeled in (3) and parameters for various acceptor types are proposed in Table I. Results obtained by force-field calculations for enzyme–substrate complexes should always be checked against structural features retrieved from structural databases (such as the CCDF or the Brookhaven Protein Data Bank). Systematic comparisons should lead to recognition of energetic favorable and unfavorable atomic arrangements and hence provide a way of improving both the form and the parametrization of force-field models. The importance of electrostatic contributions to the force field has probably been overestimated, and no great harm seems to be introduced when these contributions are reduced or even set equal to zero.

(33) Kannan, K. K.; Vaara, I.; Notstrand, B.; Lovgren, S.; Borell, A.; Fridborg, K.; Petef, M. In "Drug Action at the Molecular Level"; Roberts, G. C. K., Ed.; Mc Millan: London, 1977; pp 73–91.

(34) Mathew, M.; Palenik, G. J. *J. Chem. Soc., Perkin Trans. 2* **1974**, 532–536.
(35) Vedani, A.; Dobler, M.; Dunitz, J. D. *J. Comput. Chem.*, submitted.