for two cycles containing compounds and two base pairs for one cycle containing compounds and (ii) K_{app} values of ethidium bromide of 9.5×10^6 M^{-1} for poly(dA-dT) and 9.9×10^6 for poly(dG-dC).

Biological Methods. Cytostatic Activity. The cytostatic assays were performed according to previously established procedures.[71,72]

Antiviral Activity. The antiviral assays were performed as reported previously.[73,74]

Acknowledgment. The investigations were supported by grants from the French "Centre National de la Recherche Scientifique" (ARI Chimie-Biologie and PNR-SIDA), the Belgian "Fonds voor Geneeskundig Wetenschappelijk Onderzoek" (Project No. 3.0010.83), and the Belgian "Geconcerteerde Onderzoeksacties" (Project No. 85/90–79). We thank Professor J. W. Lown (Edmonton, Canada) for his interest and helpful discussions and for kindly providing a sample of netropsin. We are indebted to Anita Van Lierde, Frieda De Meyer, and Lizette Van Berckelaer for excellent technical assistance with the biological assays. The assistance of C. Duguet in typing this manuscript is also greatly appreciated.

Registry No. **1**, 6973-60-0; **2**, 13138-76-6; **3**, 2948-69-8; **4**, 13138-75-5; **5**, 1873-12-7; **7**, 2877-46-5; **8**, 2853-18-1; **10**, 13138-78-8; **11**, 28494-51-1; **12**, 3185-95-3; **13**, 24064-13-9; **14**, 69910-20-9; **15**, 97950-76-0; **16**, 3185-94-2; **17**, 14555-80-7; **18**, 119436-86-1; **19**, 119436-88-3; **19** (free base), 119436-87-2; **20**, 119454-85-2; **21**, 119436-89-4; **22**, 119436-91-8; **22** (free base), 119436-90-7; **23**, 119454-86-3; **24**, 119436-93-0; **24** (free base), 119436-92-9; **25**, 119436-94-1; **26**, 119436-96-3; **27**, 119436-98-5; **27** (free base), 119436-97-4; **28**, 119436-99-6; **29**, 119437-00-2; **30**, 119437-01-3; **31**, 119437-02-4; **32**, 119454-87-4; **33**, 119437-03-5; **34**, 119437-05-7; **34** (free base), 119437-04-6; **35**, 119437-07-9; **35** (free base), 119437-06-8; **36**, 119437-09-1; **36** (free base), 119437-08-0; **37**, 119437-11-5; **37** (free base), 119437-10-4; **38**, 119437-13-7; **38** (free base), 119437-12-6; **39**, 119437-15-9; **39** (free base), 119437-14-8; $H_2NCH_2CH_2CN$, 151-18-8; $(BOC)NH(CH_2)_4N(BOC)(CH_2)_3NH_2$, 68076-39-1; $H_2NC(NH)NHCH_2COOH \cdot HCl$, 14901-20-3; H-Ala-OBu-*t*·HCl, 13404-22-3; H-Ala-Ala-OBu-*t*·HCl, 60417-38-1; H-Ala-Ala-Ala-OBu-*t*·HCl, 119437-17-1; poly(dA-dT), 26966-61-0; poly(dG-dC), 36786-90-0; N^1-[[1-methyl-4-(*N'*,*N'*-dicyclohexylguanilyl)pyrrol-2-yl]carbonyl]-N^4,N^8-bis(*tert*-butoxycarbonyl)-spermidine, 119437-16-0.

(71) De Clercq, E.; Balzarini, J.; Torrence, P. F.; Mertes, M. P.; Schmidt, C. L.; Shugar, D.: Barr, P. J.; Jones, A. S.; Verhelst, G.; Walker, R. T. *Mol. Pharmacol.* **1981**, *19*, 321.
(72) Balzarini, J.; De Clercq, E.; Torrence, P. F.; Mertes, M. P.; Park, J. S.; Schmidt, C. L.; Shugar, D.; Barr, P. J.; Jones, A. S.; Verhelst, G.; Walker, R. T. *Biochem. Pharmacol.* **1982**, *31*, 1089.
(73) De Clercq, E.; Descamps, J.; Verhelst, G.; Walker, R. T.; Jones, A. S.; Torrence, P. F.; Shugar, D. *J. Infect. Dis. Rev.* **1980**, *5*, 563.
(74) De Clercq, E.; Montgomery, J. A. *Antiviral Res.* **1983**, *3*, 17.

New Hydrogen-Bond Potentials for Use in Determining Energetically Favorable Binding Sites on Molecules of Known Structure

David N. A. Boobbyer, Peter J. Goodford,* Peter M. McWhinnie, and Rebecca C. Wade

The Laboratory of Molecular Biophysics, The Rex Richards Building, University of Oxford, Oxford OX1 3QU, England.
Received May 31, 1988

An empirical energy function designed to calculate the interaction energy of a chemical probe group, such as a carbonyl oxygen or an amine nitrogen atom, with a target molecule has been developed. This function is used to determine the sites where ligands, such as drugs, may bind to a chosen target molecule which may be a protein, a nucleic acid, a polysaccharide, or a small organic molecule. The energy function is composed of a Lennard-Jones, an electrostatic and a hydrogen-bonding term. The latter is dependent on the length and orientation of the hydrogen bond and also on the chemical nature of the hydrogen-bonding atoms. These terms have been formulated by fitting to experimental observations of hydrogen bonds in crystal structures. In the calculations, thermal motion of the hydrogen-bonding hydrogen atoms and lone-pair electrons may be taken into account. For example, in a alcoholic hydroxyl group, the hydrogen may rotate around the C–O bond at the observed tetrahedral angle. In a histidine residue, a hydrogen atom may be bonded to either of the two imidazole nitrogens and movement of this hydrogen will cause a redistribution of charge which is dependent on the nature of the probe group and the surrounding environment. The shape of some of the energy functions is demonstrated on molecules of pharmacological interest.

Predictions of how ligands, such as drugs, might bind to biological molecules can be made by examining the energetic and steric properties of the system. In the method described here, the energies are calculated by a program called GRID,[1] which can be made available for use. They are displayed as energy contours around the chosen molecular target, using three-dimensional computer graphics with program FRODO[2] on an Evans and Sutherland PS300 display system. The energy is calculated between a probe group, which could be, for example, a water molecule or an amine nitrogen atom or an hydroxyl group, and a molecule of known structure, which is called the target molecule, at regular intervals throughout a target region. Large target molecules such as proteins or nucleic acids in an aqueous environment may be studied, and ligand molecules are treated by studying a number of probe groups individually. Other force fields have been developed in order to determine the interaction between whole molecules, but in order to calculate the interaction between a probe and a molecule, it was necessary to develop an independent force field.

Previous approaches to the calculation of binding energies have used either quantum mechanics[3–5] or molecular mechanics methods.[6–9] The latter approach is adopted

(1) Goodford, P. J. *J. Med. Chem.* **1985**, *28*, 849.
(2) Jones, T. A. *J. Appl. Crystallogr.* **1978**, *11*, 268. Jones, T. A. *Computational Crystallography*; Sayre, D., Ed.; Clarendon Press: Oxford, 1982; p 303.
(3) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. *J. Am. Chem. Soc.* **1985**, *107*, 3902.
(4) Clementi, E. *Lect. Notes Chem.* **1980**, 19.
(5) Pullman, B.; Pullman, A. *Prog. Nucl. Acid. Res. (Mol. Biol.)* **1969**, *9*, 327.
(6) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187.

0022-2623/89/1832-1083$01.50/0

here principally because molecular orbital calculations for large systems are very time-consuming and are usually done in vacuo, while molecular mechanics calculations are quicker, and can therefore take some account of the solvent present in biological systems either explicitly or by means of a dielectric constant.

Hydrogen bonds are important in determining the conformation of proteins, polysaccharides, nucleic acids, and organic molecules.[10] The strength of such bonds varies in a complex manner, and a wide variety of empirical energy functions[11] have been devised in order to compute their strength. In general terms, they may be considered as intermediate range interactions between an electron-deficient hydrogen atom and a nearby region of high electron density. Both the donor atom, which is covalently bonded to the hydrogen atom, and the acceptor atom are highly electronegative. When strong hydrogen bonds are formed, the donor and acceptor atoms approach closer than the sum of their van der Waals radii.[10]

The physical origins of the hydrogen bond are subject to debate. In 1928, Pauling[12] described it as being electrostatic and due to ionic forces between a partially positively charged hydrogen atom and a negative lone pair on the acceptor atom. More recently, quantum mechanical explanations, involving valence bond,[13,14] charge transfer,[15] or molecular orbital theories,[16-19] have been proposed. These suggest that, at small separations, the hydrogen bond is governed by a balance between attractive electrostatic, charge transfer, polarization, and dispersion components and repulsive electron exchange terms, while at large distances the hydrogen bond is almost completely electrostatic. Charge transfer is the second largest attractive component and it enables stabilization of hydrogen bonds such as O–H···O, where part of the electron charge occupies the antibonding orbital, causing weakening and lengthening of the O–H bond.[15] Most of the charge redistribution, though, is due to the movement of the σ electrons from the proton acceptor to the donor atom.[19] The large exchange repulsion prevents bond bending and favors the occurrence of linear hydrogen bonds.

Molecular mechanics potentials may be divided into two categories: those which do contain a hydrogen-bonding term in addition to van der Waals and electrostatic terms[6,7,9,20-24] and those which do not.[8,25,26] The former often employ an exponential[27-29] or 12–10[6,7,9] function to describe the interaction between the hydrogen-bonding atoms. The latter generally rely on the explicit representation of hydrogen atoms and sometimes lone-pair electrons to account for the hydrogen-bond directionality. However, the lone-pair electrons and hydrogen atoms are not normally observed in macromolecules studied by X-ray crystallography and their coordinates must therefore be estimated with standard known geometries.

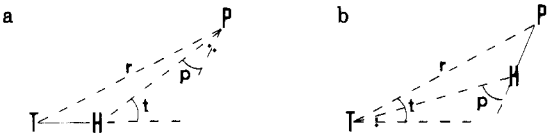


Figure 1. The definition of angles t and p when a hydrogen bond is (a) donated and (b) accepted by the target atom T. P represents the probe group. The direction of the lone pair electrons is represented by dots (:).

In the GRID method, an "extended" atom representation is used when large target molecules are studied. An "extended" atom is a sphere, centered on a point charge, consisting of one atom heavier than hydrogen and an appropriate number of hydrogen atoms which are considered implicitly by modification of the Lennard-Jones and electrostatic parameters of the heavy atom. The energy function includes a hydrogen-bond term which can account for the short-range dipole–dipole interactions due to the lone pairs and hydrogen atoms since these are not formulated explicitly in the "extended" atom representation. The hydrogen-bond term enables asymmetries in the intermolecular interaction to be modeled which cannot be evaluated by means of the atom-centered centrosymmetric van der Waals and electrostatic components of the energy function. In large molecules, hydrogen coordinates are only used in order to determine the geometry of hydrogen bonds. For small target molecules, an "all atom" representation may be used but a hydrogen-bonding function is still necessary as lone pairs are not treated explicitly. An "extended" atom representation is always used for probes.

In this paper, the formulation of an improved set of hydrogen-bond potentials for use in the GRID method is described. In order to match the experimental observations as closely as possible, the mobility of the target molecule's hydrogen atoms and lone pairs as well as the length and orientational dependence of the hydrogen bond are considered. The use of the new hydrogen-bond potentials, together with the Lennard-Jones and electrostatic potentials, to display probe–ligand interactions is demonstrated in certain cases of pharmacological interest where hydrogen bonding is important.

Method

The GRID energy function consists of Lennard-Jones, electrostatic, and hydrogen-bonding terms given by the following equations:

$$E = \sum E_{lj} + \sum E_{el} + \sum E_{hb}$$

(7) Weiner, S. J.; Kollman, P. A.; Case, D. A.; Singh, U. C.; Ghio, C.; Alagona, G.; Profeta, S.; Weiner, P. *J. Am. Chem. Soc.* **1984**, *106*, 765. Weiner, S. J.; Kollman, P. A.; Nguygen, D. T.; Case, D. A. *J. Comput. Chem.* **1986**, *7*, 230.
(8) Hagler, A. T.; Huler, E.; Lifson, S. *J. Am. Chem. Soc.* **1974**, *96*, 5319.
(9) Momany, F. A.; McGuire, R.; Burgess, A.; Scheraga, H. A. *J. Phys. Chem.* **1975**, *79*, 2361.
(10) Pauling, L. *The Nature of the Chemical Bond*, 3rd ed.; Cornell University Press: New York, 1960; Chapter 12.
(11) Hall, D.; Pavitt, N. *J. Comput. Chem.* **1984**, *5*, 441.
(12) Pauling, L. *Proc. Natl. Acad. Sci. U.S.A.* **1928**, *14*, 359.
(13) Coulson, C. A.; Danielsson, U. *Ark. Fys.* **1954**, *8*, 239, 245.
(14) Tsubomura, H. *Bull. Chem. Soc. Jpn.* **1954**, *27*, 445.
(15) Puranik, P. G.; Kumar, V. *Proc. Indian Acad. Sci.* **1963**, *58*, 29.
(16) Morokuma, K.; Pedersen, L. *J. Chem. Phys.* **1968**, *48*, 3275.
(17) Morokuma, K.; Winick, J. R. *J. Chem. Phys.* **1970**, *52*, 1301.
(18) Dreyfus, M.; Pullman, A. *Theor. Chim. Acta* **1970**, *19*, 20.
(19) Umeyama, H.; Morokuma, K. *J. Am. Chem. Soc.* **1977**, *99*, 330, 1316.
(20) Momany, F. A.; Carruthers, L. M.; McGuire, R. T.; Scheraga, H. A. *J. Phys. Chem.* **1974**, *78*, 1595.
(21) Momany, F. A. *Environmental Effects on Molecular Structures and Properties*; Pullman, B.; Ed.; D. Reidel: Dordrecht, Holland, 1976; p 437.
(22) Vedani, A.; Dunitz, J. D. *J. Am. Chem. Soc.* **1985**, *107*, 7653.
(23) Nilsson, L.; Karplus, M. *J. Comput. Chem.* **1986**, *7*, 591.
(24) Levitt, M. *J. Mol. Biol.* **1974**, *82*, 93.
(25) Snir, J.; Nemenoff, R. A.; Scheraga, H. A. *J. Phys. Chem.* **1978**, *82*, 2497.
(26) Hermans, J.; Berendsen, H. J. C.; van Gunsteren, W. F.; Postma, J. A. *Biopolymers* **1984**, *23*, 1512.
(27) Kroon-batenburg, L. M. J.; Kanters, J. A. *THEOCHEM* **1983**, *105*, 417.
(28) Taylor, R. *J. Mol. Struct.* **1981**, *71*, 311.
(29) Lippincott, E. R.; Schroeder, R. *J. Chem. Phys.* **1955**, *23*, 1099.

where

$$E_{lj} = A/r^{12} - B/r^6$$

$$E_{el} = q_p q_t / K\zeta\{1/r + [(\zeta - \epsilon)/(\zeta + \epsilon)]/(r^2 + 4s_p s_t)^{1/2}\}$$

$$E_{hb} = \mathrm{Er} \times \mathrm{Et} \times \mathrm{Ep}$$

The total energy, E, is computed as the sum of the pairwise interactions of the probe group with each atom in the target molecule. r is the distance between the probe group and an atom in the target, q_p is the electrostatic charge of the probe, q_t is the electrostatic charge of the target atom, K is a combination of geometrical factors and natural constants, ϵ is the dielectric constant of the solvent,[1] ζ is the dielectric constant of the target phase,[1] s_p is the depth of the probe in the target phase in Å,[1] s_t is the depth of the target atom in the target phase in Å,[1] A and B are parameters describing the types of interacting atoms, and Er, Et, and Ep are functions of r, t, and p, respectively, where the angle made by the hydrogen bond at the target is t and at the probe is p, as defined in Figure 1.

The new hydrogen-bond potential was formulated to give predicted structures which agreed with experimental data. It was necessary to make the energy the product of three terms: Er dependent on the separation of the hydrogen-bonding atoms, and Et and Ep dependent on the angle made by the hydrogen bond at the target or probe atom respectively. Et and Ep describe the orientational dependence of the hydrogen bonds. In this paper, however, situations are considered in which only one hydrogen-bond interaction occurs between the target and the probe at any one position. Therefore, Ep = 1.0 as the probe is assumed to orientate itself in order to form the strongest hydrogen bond possible.

It should be noted that the hydrogen-bond potential was developed for use as part of the GRID energy function[1] and was not designed to be transferable to other energy potentials.

Distance Dependence

For the distance dependence Er of the hydrogen bond, a 12–10 function has been used in a number of other potential functions[6,7,9,22,23] and this was tested first. However, when used as a term in the GRID energy function, it proved to give too narrow a range of hydrogen-bond lengths when compared with the range of lengths observed experimentally.[30] A 6–4 function, which had also been used previously,[6,23] was then tried, but this could produce hydrogen bonds of such a great length that it would be possible for another non-hydrogen atom to position itself between the hydrogen-bonding atoms. Eventually an 8–6 function, as suggested by Reiher,[31] was adopted and found to give the most satisfactory results and the closest agreement with experimental observations. This is given by the following equations:

$$\mathrm{Er} = C/r^8 - D/r^6$$

$$C = -3E_m r_m^8 \text{ kcal Å}^8/\text{mol}$$

$$D = -4E_m r_m^6 \text{ kcal Å}^6/\text{mol}$$

where r is the separation of the acceptor atom and the donor heavy atom in angstroms, E_m is the optimum hydrogen-bond energy in kilocalories/mole for the particular hydrogen-bonding atoms considered, r_m is the optimum hydrogen-bond length in angstroms for the particular hydrogen-bonding atoms considered (see Figure 2). E_m and r_m vary according to the chemical type of the hydrogen-bonding atoms. At this stage in the development of the potential function, it is assumed that for N···N hydrogen bonds E_m = −2.0 kcal/mol and r_m = 3.2 Å; for N···O hydrogen bonds E_m = −2.8 kcal/mol and r_m = 3.0 Å; and for O···O hydrogen bonds E_m = −4.0 kcal/mol and r_m = 2.8 Å. Further work will be required, but it is already clear that hydrogen bonds involving ether oxygens may be weaker than those to other oxygen atoms.

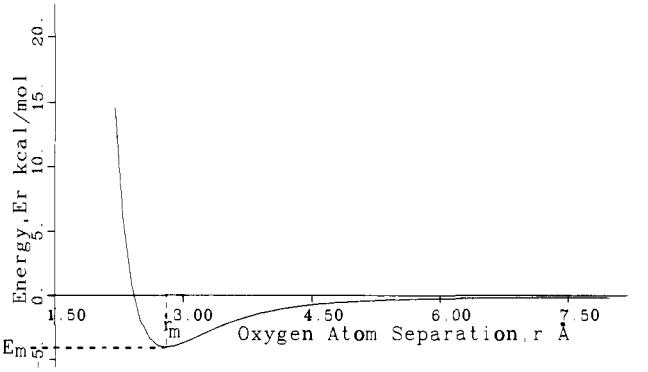


Figure 2. The variation of the hydrogen-bond energy Er, in kilocalories/mole, with the separation of the donor and the acceptor atoms r, in angstroms. See text for details.

A cutoff energy of −0.25 kcal/mol is imposed on long hydrogen bonds in order to exclude large numbers of weak hydrogen bonds from further consideration by program GRID.

Angular Dependence

The function Et is dependent on the chemical nature of the hydrogen-bonding atoms in the target molecule. It is chosen by fitting to experimental data on hydrogen-bond geometries in crystalline structures observed by X-ray or neutron diffraction. The observed structures are of organic small molecules and proteins as shown in Table I. Neutron studies tend to be slightly more accurate than those by X-ray because neutrons can easily detect hydrogen atoms. However, this difference is small when X-ray results have been "normalized" to account for the fact that the length of bonds to hydrogen atoms may be underestimated by 0.1–0.2 Å.[48,49] Spectroscopic studies also give

(30) Artymiuk, P. J.; Blake, C. C. F. *J. Mol. Biol.* **1981**, *152*, 737.
(31) Reiher, W. E. Ph.D Thesis, Harvard University, 1985.
(32) Ramakrishnan, C.; Prasad, N. *Int. J. Protein Res.* **1971**, *III*, 209.
(33) Ferraris, G.; Franchini-Angela, M. *Acta Crystallogr.* **1972**, *B28*, 3572.
(34) Pedersen, B. *Acta Crystallogr.* **1974**, *B30*, 289.
(35) Kroon, J.; Kanters, J. A.; van Duijnevelt-Van de Rijdt, J. G. C. M.; van Duijneveldt, F. B.; Vliegenthart, J. A. *J. Mol. Struct.* **1975**, *24*, 109.
(36) Mitra, J.; Ramakrishnan, C. *Int. J. Peptide Prot. Res.* **1977**, *9*, 27.
(37) Vinogradev, S. N. *Int. J. Peptide Prot. Res.* **1979**, *14*, 281.
(38) Vinogradev, S. N. *Int. J. Biopolymers* **1979**, *18*, 1559.
(39) Ceccarelli, C.; Jeffrey, C. A.; Taylor, R. *J. Mol. Struct.* **1981**, *70*, 255.
(40) Taylor, R.; Kennard, O. *J. Am. Chem. Soc.* **1982**, *104*, 5063.
(41) Jeffrey, G. A.; Maluszynska, H. *Int. J. Biol. Macromol.* **1982**, *4*, 17.
(42) Jeffrey, G. A.; Mitra, J. *Acta Crystallogr.* **1983**, *B39*, 469.
(43) Taylor, R.; Kennard, O.; Versichel, W. *J. Am. Chem. Soc.* **1983**, *105*, 5761.
(44) Taylor, R.; Kennard, O.; Versichel, W. *Acta Crystallogr.* **1984**, *B40*, 280.
(45) Murray-Rust, P.; Glusker, J. P. *J. Am. Chem. Soc.* **1984**, *106*, 1018.

Table I. Summary of the Experimental Data on Hydrogen Bonding Which Was Used for Modeling the Angular Dependence of Hydrogen Bonds

ref no.	publication date	method of study[d]	no. of structures studied	total no. of hydrogen bonds studied	hydrogen bond type studied	compsition of crystals
32	1971	X	40	241	N–H···O	amino acid peptide polypeptide
33	1972	N	41	190	water H–O–H···A[a]	organic inorganic
34	1974	N	41	190	water H–O–H···A[a]	organic inorganic
35	1975	X	45	196	O–H···O	saccharide polyalcohol and related compounds
36	1977	X	90	356	O–H···O	amino acid peptide oligosaccharide
37, 38	1979	N, X	95	439	N–H···O O–H···O N–H···X$^-$[b] O–H···X$^-$	amino acid, peptide, and related molecules
39	1981	N	24	100	O–H···O	organic carbohydrate
40	1982	N	113	661	C–H···O C–H···N C–H···Cl	organic
41	1982		32	168	N–H···O	amino acid
42	1983	N, X	58	277	O–H···O	carbohydrate
43	1983	N, X	889	1509	N–H···O	organic
44	1984	N, X	889	1509	N–H···O	organic
45	1984	N, X	whole Cambridge Crystallographic Database[c]		N–H···O O–H···O	organic
46	1984	X	15	unknown	N–H···O O–H···O	protein
47	1985	X	86	529	N–H···O N–H···O	nucleoside nucleotide

[a] A is any atom accepting an hydrogen bond. [b] X$^-$ is a halide ion. [c] Reference 71. [d] N = neutron, X = X-ray diffraction.

some information about hydrogen-bond geometries.[50]

Data were collected from "good" crystal structures. For small molecules, these can be defined as structures in which the crystallographic R factor is less than 10%, hydrogen coordinates are available, and there is no disorder.[40,45] Therefore, only molecules which crystallize well are studied. As a result, the sample of hydrogen bonds may not be representative of those formed between molecules in solution. The hydrogen bonds in a crystal are likely to be shorter than in solution due to long-range attractive forces, but this is unlikely to exert a detectable influence on bond angles.[34]

The hydrogen-bond function has been designed to be applicable to general intermolecular interactions but not to intramolecular interactions which may have a different spatial distribution. However, experimental observations of intra- and intermolecular hydrogen bonds were both studied. In addition, some molecular systems show hydrogen-bonding patterns which are peculiar only to the particular molecule studied. For example, amino acids have a large number of three-centered hydrogen bonds due to a proton deficiency and the presence of charged ammonium and carboxylate groups,[51] while proteins have secondary structures, such as the α helix, with characteristic hydrogen-bond orientations which are not typical of intermolecular hydrogen bonds in general. Therefore, due consideration of the source of the experimental data must be taken when using the data as the basis for determining the terms occurring in the energy function.

The Derivation of Hydrogen-Bonding Functions

The observations in Table I have been analyzed to give the number of hydrogen bonds formed per 10° angular segment. This number is then converted into a probability by multiplying by a geometric factor,[34,49,52] since 10° segments of angle t do not occupy the same solid angle. For example, for a target amide nitrogen, a correction factor of $1/(\cos t' - \cos t'')$ is used where the 10° segment is between angles t' and t''. This alters the apparent distribution of hydrogen bonds as a function of angle t. For an N–H···O hydrogen bond, the most frequently observed value of angle t is about 15°. However, when the experimental data are converted into probabilities, it is seen that the most probable conformation of the N–H···O hydrogen bond is linear with angle $t = 0°$. The probabilities are then converted into relative energies assuming a Boltzmann distribution:

$$E_i = -RT(\ln \{P_i\} - \ln \{P_0\}) + E_0$$

where P_i is the probability of formation of a hydrogen bond of energy E_i and P_0 is the probability of formation of an optimally orientated hydrogen bond of energy E_0.

These energies are plotted against angle t and a suitable analytical function chosen to fit to them. The function must be continuous and of a simple form so that it is quick to compute.

In general, the orientational dependence Et of the hydrogen-bond energy varies according to the chemical na-

(46) Baker, E. N.; Hubbard, R. E. *Prog. Biophys. Mol. Biol.* **1984**, *44*, 97.
(47) Jeffrey, G. A.; Maluszynska, H.; Mitra, J. *Int. J. Biol. Macromol.* **1985**, *7*, 336.
(48) Taylor, R.; Kennard, O. *Acta Crystallogr.* **1983**, *B39*, 133.
(49) Olavsson, I.; Jonssen, P. G. *The Hydrogen-Bond*; Schuster, P., Zundel, G., Sandorfy, C., Eds.; North-Holland: Amsterdam, 1976; Vol. II, Chapter 8.
(50) Taylor, R.; Kennard, O. *Acc. Chem. Res.* **1984**, *17*, 320.
(51) Jeffrey, G. A.; Mitra, J. *J. Am. Chem. Soc.* **1984**, *106*, 5546.
(52) Kroon, J.; Kanters, J. A. *Nature* **1974**, *248*, 667.

Table II. Derived Functions Which Define Hydrogen-Bonding Characteristics

description of atom	no. of hydrogen bonds donated	no. of hydrogen bonds accepted	hybridization	variation of hydrogen/lone pair position	angular function, Et, for an atom which donates[a]	angular function, Et, for an atom which accepts[a]	angular range of Et
oxygen types							
carbonyl and carboxy O	0	2	sp^2	two alternative		$(0.9 + 0.1 \sin 2t_i) \cos t_0$	$0 < t_i < 90°$
						$K_1(K_2 - \cos^2 t_i)^3 \cos t_0$[b]	$90° < t_i < 110°$
						0	$t_i > 110°$
phenolic hydroxyl O	1	1	sp^2	two alternative[c]	$\cos^4 t$	$(0.9 + 0.1 \sin 2t_i) \cos t_0$	$0 < t_i < 90°$
						$K_1(K_2 - \cos^2 t_i)^3 \cos t_0$[b]	$90° < rt_i < 110°$
						0	$t_i > 110°$
				fixed	$\cos^4 t$	$\cos^2 t$	
phenolate O	0	2	sp^2	two alternative		$(0.9 + 0.1 \sin 2t_i) \cos t_0$	$0 < t_i < 90°$
						$K_1(K_2 - \cos^2 t_i)^3 \cos t_0$[b]	$90° < t_i < 110°$
						0	$t_i > 110°$
serine and	1	2	sp^3	rotating[d]	$\cos^6 t$	$\cos^4 t$	
threonine hydroxyl OH	1	2	sp^3	fixed	$\cos^6 t$	$\cos^2 t_0$	$t_i < 55°$[e]
						$\cos^2 t$	$t_i > 55°$
general case	1	2	sp^3	rotating	$\cos^4 t$	$\cos^2 t$	
hydroxyl OH	1	2	sp^3	fixed	$\cos^4 t$	$\cos^2 t_0$	$t_i < 55°$[e]
						$\cos^2 t$	$t_i > 55°$
ether O	0	2	sp^3	two alternative		$\cos^2 t_0$	$t_i < 55°$[e]
						$\cos^2 t$	$t_i > 55°$
water H_2O	2	2	sp^3	infinite	1	1	
				fixed	$\cos^4 t$	$\cos^2 t_0$	$t_i < 55°$[e]
						$\cos^2 t$	$t_i > 55°$
nitrogen types							
nitrogen N	0	1	sp, sp^2, sp^3	fixed		$\cos^2 t$	
nitrogen N	0	2	sp^2	two alternative		$(0.9 + 0.1 \sin 2t_i) \cos t_0$	$0 < t_i < 90°$
						$K_1(K_2 - \cos^2 t_i)^3 \cos t_0$[b]	$90° < t_i < 110°$
						0	$t_i > 110°$
NH	1	0	sp, sp^2, sp^3	fixed	$\cos^2 t$		
NH	1	1	sp^3	two alternative	$\cos^2 t$	$\cos^2 t$	
histidine N	1[f]	1[f]	sp^2	fixed	$\cos^2 t$	$\cos^2 t$	
phenolic NH	1	1	sp^2	two alternative[c]	$\cos^2 t$	$(0.9 + 0.1 \sin 2t_i) \cos t_0$	$0 < t_i < 90°$
						$K_1(K_2 - \cos^2 t_i)^3 \cos t_0$[b]	$90° < t_i < 110°$
						0	$t_i > 110°$
				fixed	$\cos^2 t$	$\cos^2 t$	
NH_2	2	0	sp^2, sp^3	two alternative	$\cos^2 t$		
amine NH_2	2	1	sp^3	rotating or fixed	$\cos^2 t$	$\cos^2 t$	
ammonium NH_3^+	3	0	sp^3	rotating or fixed	$\cos^2 t$		
fluorine and chlorine	0	3	sp^3	infinite		1	

[a] t is the angle made by the hydrogen bond at the target atom as defined in Figure 1. t_i is the deviation of the probe hydrogen atom from the bisector of the lone-pair orbitals within the plane of the lone pair orbitals. See Figure 3. t_0 is the deviation of the probe hydrogen atom from the plane of the lone pair orbitals. See Figure 3. [b] This is a smoothing function with $K_1 = 0.9/\cos^6 110°$ and $K_2 = \cos^2 110°$. [c] The hydrogen atom is at 64° to the C–O or C–N bond in the plane of the ring rather than the ideal trigonal angle of 60° adopted by hydrogens on other sp^2 target groups and by all lone pairs as experimental evidence suggests the suitability of this angle.[46] [d] The hydrogen atom is at 67° and the lone pair is at 58° to the C–O bond rather than the ideal tetrahedral angle of 70.5° at which hydrogens and lone pairs are assumed to rotate for other target groups. This difference is due the particular angular dependence observed experimentally for this target atom.[46] [e] For sp^3 target atoms, the lone pairs are assumed to lie at $t_i = 55°$ and different hydrogen-bonding functions are adopted depending on whether the probe is between or outside the lone-pair orbitals. [f] Both imidazole nitrogens are assumed able to donate or accept a hydrogen bond although they can only make one hydrogen bond at a time. The charge distribution of the histidine residue must be adjusted according to which nitrogen the hydrogen is assumed to be attached to.

ture of the target atom. This can be seen in Table II, which lists the functions used to describe the hydrogen bonds made by different types of target atom. These functions will now be described in detail for specific cases.

Oxygen. Carbonyl and Carboxy Oxygen Atoms. A carbonyl or carboxy oxygen atom can accept one or two hydrogen bonds. A definite tendency for these to form in the direction of the lone pairs at ±60° to the C–O bond in the lone-pair plane has been reported.[49]

The orientation of a hydrogen bond accepted by a carbonyl or carboxy oxygen atom is described in terms of two angles: t_o, which is the angle subtended at the target oxygen atom by the hydrogen atom which is donated by the probe and the lone-pair plane of the target oxygen; and t_i, which is the angle of deviation within the oxygen lone pair plane of the probe hydrogen atom from the direction of the C–O bond. This is illustrated in Figure 3. In observed hydrogen bonds, angle t_0 tends to be less than 40° although larger angles are not unusual.[48,52] Within the lone-pair plane, hydrogen bonds in the lone pair orbital

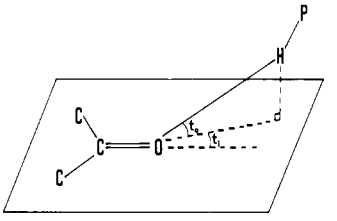


Figure 3. The geometry of the target carbonyl group. Angle t_0 is the angle subtended at the oxygen atom by the probe's hydrogen atom and the plane of the lone pair orbitals. Angle t_i is the angle subtended by the probe hydrogen atom and the C–O axis within the plane of the lone-pair orbitals.

directions are favored and more hydrogen bonds occur between the orbitals than outside them as can be seen from the histograms in Figure 4. The directionality is more marked if two hydrogen bonds are accepted than if only one hydrogen bond is accepted, when both lone pairs tend to attract the single donor atom to their bisector, sug-

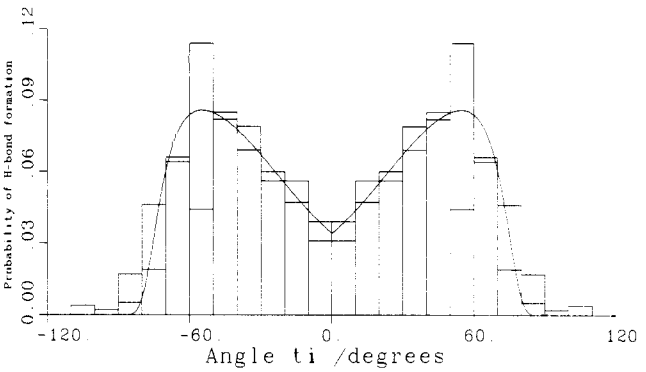


Figure 4. The two histograms show the observed probabilities of a hydrogen bond to a carbonyl oxygen atom derived from the experimental data[32,60] which was used for modeling the hydrogen-bonding function. The variation between the histograms is due to the different experimental samples studied. The graph shows the modeled variation of the probability of hydrogen-bond formation with angle t_i derived from the total interaction energy calculated for a target carbonyl oxygen atom interacting with a probe water molecule. The water molecule lies in the lone-pair plane, i.e. $t_0 = 0°$, at a distance of 2.8 Å from the oxygen.

gesting that both electrostatic and steric factors influence the hydrogen-bond geometry. There is no correlation between angle t_0 and angle t_i.

The chosen hydrogen-bonding term Et for a carbonyl or carboxy oxygen atom has maxima at $t_i = \pm 45°$ and this causes E_{hb} to be most attractive at this angle. However, when the Lennard-Jones interaction of the probe with the atoms covalently bonded to the target oxygen are also taken into account, the probability of hydrogen-bond formation (derived from the total energy calculated by program GRID) is found to be highest in the lone pair orbital directions at $t_i = \pm 60°$ as shown by the graph in Figure 4. The functional form of Et causes a variation in E_{hb} of about 10% over the range $t_i = 0$–$90°$ and this is not incompatible with molecular orbital calculations[18] for a formamide dimer where the variation in hydrogen-bond energy is about 15% over the range $t_i = 0$–$80°$. When angle $t_i > 80°$, the molecular orbital calculations show that the total interaction energy rises rapidly as exchange repulsion effects increase.

The Phenolic Hydroxyl Group. In proteins, the tyrosine residue is observed to donate hydrogen bonds more often than it accepts them.[46] It can be considered able to donate one hydrogen bond and to accept only one hydrogen bond, because there is a significant back-donation of the oxygen lone pairs which gives partial double-bond character to the C–O bond, restricting rotational motion around it. The hydroxyl oxygen resonates between sp^2 and sp^3 hybridization and can thus be considered to have its hydrogen atom and one lone pair in the plane of the aromatic ring at ±60° to the C–O bond direction, thus accounting for the observed hydrogen-bonding geometry. In calculating hydrogen-bond energies, the hydrogen atom is allowed to occupy the energetically most favorable of the two positions available to it, according to the position and chemical characteristics of the interacting probe. Similarly, the lone-pair orbital always takes up the most favorable of the two possible positions available to it. Hence, account is taken of the influence of the incoming probe on the conformation of the phenolic hydroxyl group.

The experimental data studied were limited to that on 150 hydrogen bonds to tyrosine residues from a survey of 15 protein crystal structures.[46] The geometries are described in terms of one angle only, the deviation from the C–O bond direction, and there are no data for when this angle is greater than 90°. Data are available as a plot of the variation of hydrogen-bond occurrence with angle t, the deviation from the O–H direction, for when tyrosine donates. $Et = \cos^4 t$ was found to be the most suitable function fitting this plot.

The experimental data on hydrogen bonds accepted by tyrosine are more limited because tyrosine is found to accept only about 20% of its experimentally observed hydrogen bonds.[46] However, where it is observed, tyrosine accepts hydrogen bonds preferentially at 59 ± 16° to the C–O bond. The same angular function is used for accepted hydrogen bonds as for a carbonyl oxygen atom because the available experimental data on tyrosine are inadequate for formulating a separate function specifically for this group. If better data on observations become available in the future, it may be possible to assign a more suitable function for this hydrogen bond.

An anionic phenolate oxygen is treated in the same manner as a phenolic hydroxyl group accepting a hydrogen bond. It may accept up to two hydrogen bonds but is unable to donate.

The sp^3-Hybridized Hydroxyl Groups. Crystallographic data show that hydrogen bonds accepted by the sp^3 hydroxyl group are strongly constrained to the group's lone pair plane which can be defined if the position of the hydroxyl hydrogen is known. This effect is more marked than the corresponding constraint for hydrogen bonds to form in the lone-pair plane of carbonyl oxygen atoms. The angle t_0 for a sp^3 hydroxyl group is generally less than 20°[52,53] compared to less than 40° for a carbonyl oxygen atom as mentioned above. Within the lone-pair plane, hydrogen bonds occur more often between the lone-pair orbitals than outside them. They have a fairly uniform distribution over the range $-60° < t_i < 60°$ if one hydrogen bond is formed, but are more strongly constrained to the lone-pair directions if two hydrogen bonds are accepted.[39,51]

However, the observed lone-pair directionality does vary according to the method of study. Microwave spectroscopy of hydrogen-bonded dimers[50,53] shows an energetic preference for the lone-pair directions for sp^3-coordinated atoms although this is smaller than for sp^2 coordination. In charge-density studies,[54] the sp^2 configuration shows some small deformation in electron density toward lone pairs whereas, for sp^3 hydroxyl groups and water, the lone pairs appear to exist as one broad peak in the electron density.

A large amount of crystallographic data is available on the geometry of hydrogen bonds to hydroxyl groups in proteins. The protein hydroxyl groups in serine and threonine are therefore described by angular functions which differ somewhat from those for general hydroxyl groups. The protein functions define hydrogen bonds which are more constrained to the lone pair or hydrogen directions than the functions for general hydroxyl groups, which make allowance for the incompleteness of the data from which the function for a general hydroxyl group is chosen.

In proteins, hydroxyl groups are more likely to donate a hydrogen bond than accept one. In a survey of 15 protein crystals, serine was observed to donate 67% of its hydrogen bonds and threonine 75% of its hydrogen bonds.[46] Hydrogen bonds in proteins are found to be accepted at 58 ± 35° to the C–O bond and donated at 67° ± 35°. It appears that the donated hydrogen bonds are closer to the tetrahedral angle because of the greater directional control

(53) Millen, D. J. *Croat. Chem. Acta* **1982**, *55*, 133.
(54) Olovsson, I. *Croat. Chem. Acta* **1982**, *55*, 171.

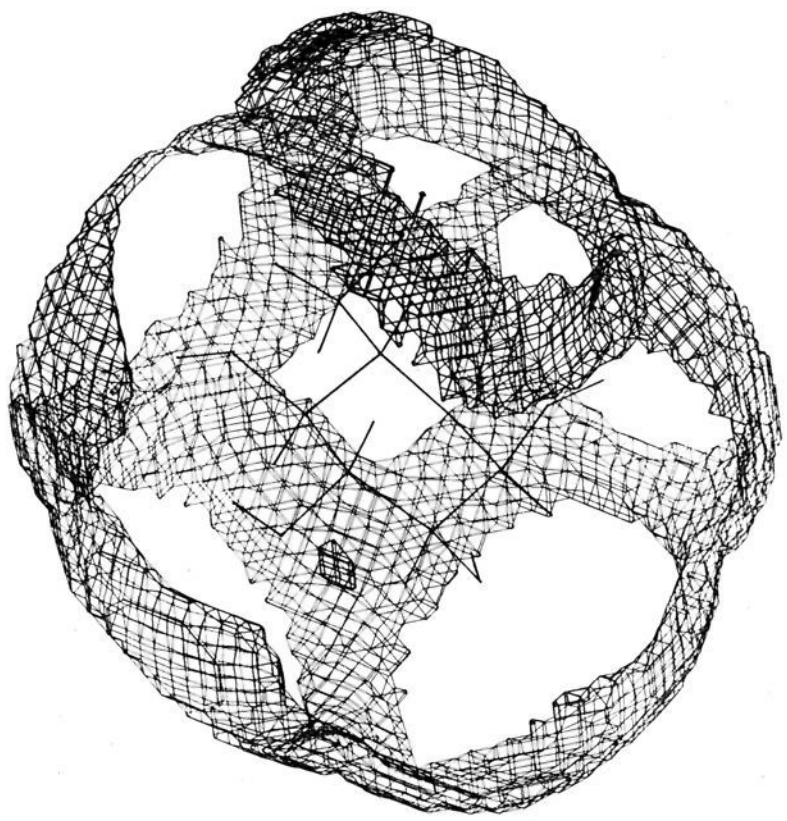


Figure 5. Energy contours at −3.2 kcal/mol for the interaction between an amide nitrogen probe able to donate one hydrogen bond and a glucose molecule. The hydroxyl hydrogens and lone pairs are allowed to rotate around the C–O axis, giving rise to circular regions of favorable energy where a strong linear hydrogen bond can be made.

of the hydrogen atom when compared to the lone-pair orbital.

In the calculation of hydrogen-bond energy, the hydrogen atom and lone pairs of a hydroxyl group are normally allowed to rotate at the tetrahedral angle around the projected C–O axis. The hydrogen atom and lone pairs are assumed to position themselves in the most energetically favorable positions according to the nature and position of the probe group. Angle t is then the deviation from this direction when the hydroxyl group donates. In Figure 5, the attractive energy contours obtained for the interaction of an amide nitrogen probe with glucose can be seen. The toroidal energy contours show the energetically favorable positions of the nitrogen following the circular locus of the hydroxyl lone pairs as it donates a linear hydrogen bond to the glucose molecule.

In some circumstances, such as if the hydroxyl group is already making an intramolecular hydrogen bond to another target atom, the explicit position of the hydrogen atom may be known and can therefore be fixed. The hydrogen bonds of the hydroxyl group are then calculated differently with donated hydrogen bonds favored in the predefined O–H direction while accepted hydrogen bonds are favored in the lone-pair plane where they are treated similarly to the hydrogen bonds of an ether oxygen atom (see below).

The Ether Oxygen Atom. This has sp^3 hybridization and is able to accept two hydrogen bonds. Its two lone pairs are in the plane perpendicular to the C–O–C plane and symmetrically placed with respect to it. The hydrogen bonds that are formed show a definite tendency to be in this lone-pair plane but have an approximately constant probability of occurrence between the lone-pair orbitals.[45] Unfortunately, there is very little experimental data on the hydrogen-bond geometry of the ether oxygen atom and a relatively weak angular dependence is chosen to fit this data.

Water. This can donate two and accept two hydrogen bonds. Generally, the hydrogen atom coordinates are unknown and in these circumstances the target waters are

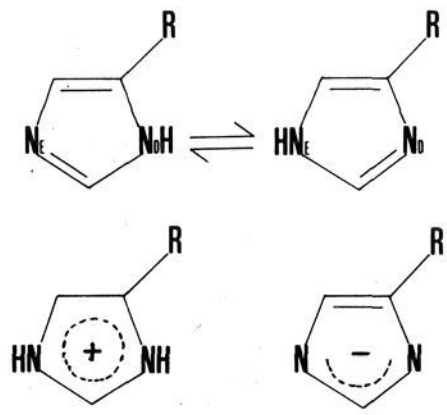


Figure 6. The two tautomers and the protonated and anionic forms of the histidine side chain.

treated as if they can rotate freely and no angular factor is applied to their hydrogen bonds. However, hydrogen positions can be assigned in program GRID when the hydrogen atom coordinates on target water molecules are unequivocally established, and the same angular constraints can then be applied as for a general target sp^3 hydroxyl group with its hydrogen atom coordinates fixed.

Nitrogen. Target nitrogen atoms may be able to donate up to three hydrogen bonds and may be able to accept one or two hydrogen bonds. Linear hydrogen bonds are favored[46,48,50] especially for shorter bonds because as the hydrogen bond becomes shorter, the nonbonded interaction becomes less favorable and so the hydrogen bond straightens out and thereby increases the distance between the heavy atoms. An $Et = \cos^2 t$ function centered on the N–H direction is adopted for all hydrogen bonds donated by these target nitrogen atoms.

For accepted hydrogen bonds, the same angular functions are applied as for oxygen atoms of the same geometry. However, E_m and r_m will be different for the two types of atom so the resultant hydrogen-bond energy for a target nitrogen atom will differ from that for a target oxygen atom. Movement of the nitrogen's hydrogen atoms and lone pairs is also allowed for in the same way as for oxygen. For example, the three hydrogens in an ammonium group may rotate about the C–N bond, or may be fixed at specified coordinates.

Histidine Nitrogen. The side chain of histidine can exist in four forms: two neutral, one protonated, and one anionic, as shown in Figure 6. The protonated form normally exists in significant concentrations only at low pH while the anionic form is generally found only at high pH. According to experiment[55–58] and theory,[59,60] the tautomer with NE protonated should predominate over that with ND protonated. This is because the R group is electron withdrawing and the effect of the R group is more marked for ND than NE.[61,62]

(55) Madden, J. J.; McGandy, E. L.; Seeman, N. C. *Acta Crystallogr.* **1972**, *B28*, 2377.

(56) Lehman, M. S.; Koetzle, T. F.; Hamilton, W. C. *Int. J. Peptide Prot. Res.* **1972**, *4*, 229.

(57) Reynolds, W. F.; Peat, I. R.; Freedman, M. H.; Lyerla, J. R. *J. Am. Chem. Soc.* **1973**, *95*, 328.

(58) Boschcov, P.; Seidel, W.; Muradian, J.; Tominaga, M.; Paiva, A. C. M.; Juliano, L. *Bioorg. Chem.* **1983**, *12*, 34.

(59) Ramani, R.; Boyd, R. J. *Can. J. Chem.* **1981**, *59*, 3232.

(60) Mezey, P. G.; Ladik, J. J.; Suhai, S. *Theor. Chim. Acta* **1979**, *51*, 323.

(61) Black, J. W.; Durant, G. J.; Emmett, J. C.; Ganellin, C. R. *Nature* **1974**, *248*, 65.

(62) Ganellin, C. R. *J. Appl. Chem. Biotechnol.* **1978**, *28*, 183.

The Protein Data Bank[63] files of 30 proteins were searched to assess the position of tautomeric equilibrium by examining the atoms neighboring the histidine residues. However, it was not possible to detect a significant preference for one or other of the tautomeric forms. The hydrogen-bonding function for histidine must therefore be devised to allow for the existence of both tautomers at physiological pH and to select the tautomer which gives the most favorable energetic interaction with the probe. It does this by allowing both nitrogens of the histidine residue to donate or accept a hydrogen bond, and making the way in which the hydrogen bond is actually formed depend on the nature and position of the probe group. If the probe group cannot make a hydrogen bond to either of these nitrogens, the hydrogen is assumed to be attached to NE. If a hydrogen bond is donated by NE or accepted by ND, the electrostatic charge of the histidine residue is unchanged. However, if a hydrogen bond is donated by ND or accepted by NE, there is an alteration in the electronic charge distribution of the histidine residue and, therefore, a compensating term must be applied to the electrostatic energy calculated. This allows not only the hydrogen position, as for a phenolic hydroxyl group, but also the electronic distribution of the histidine residue to be influenced appropriately by the nature of the probe group.

For calculations at high pH, both histidine nitrogens are assumed to be unprotonated while at low pH they are both assumed to be protonated. The histidine charge is adjusted accordingly, but its distribution at these pHs is assumed to be independent of the nature of the probe group.

There are insufficient observations of histidine unequivocally accepting a hydrogen bond from which to construct a function Et. However, as the nitrogen is sp^2 hybridized, we assume that it has the same characteristics as a carbonyl oxygen which is accepting two hydrogen bonds simultaneously, although only one hydrogen bond can be made by the histidine nitrogen. For carbonyl oxygen atoms with their full hydrogen-bonding capacity satisfied, most hydrogen bonds are observed within 20° of the lone pair orbital directions and their hydrogen-bond distribution can be modeled with use of the angular function $Et = \cos^2 t$. Nitrogen is less electronegative than oxygen so the lone pair cloud is likely to be more diffuse. However, the same angular function is used for the histidine nitrogen when it accepts an hydrogen bond, although E_m and r_m in the distance-dependent function Er are different.

The geometry of hydrogen bonds donated by histidine residues in proteins has been observed.[46] These data and general data on N–H···O bonds are fitted by the function $Et = \cos^2 t$ centered on the N–H direction.

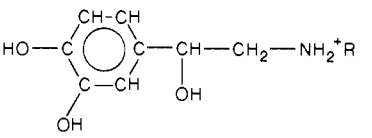
Halogens. Fluorine is able to accept up to three hydrogen bonds, but because of the wide range of angles over which it can accept, no angular function is adopted. Chlorine may also be able to accept weak hydrogen bonds with a similar angular dependence to fluorine. The hydrogen bonds to chlorine are observed to be generally longer than those to other acceptor atoms.[37]

Other Atoms. Hydrogen bonds to carbon atoms have been observed to be significant in small organic molecules containing nitrogen and in proteins where a carbon atom is adjacent to a protonated nitrogen.[40] Hydrogen bonds to sulfur atoms have also been described. However, hydrogen bonds to these atoms are not calculated by the present version of program GRID.

Examples of the Application of the Hydrogen-Bond Functions. The first two examples show applications of the GRID method to the catecholamines and the cardiac glycosides in order to demonstrate features of the hydrogen-bond potential. The GRID energy maps are calculated for uncharged probes so that the hydrogen-bond geometries predicted are dependent only on the centrosymmetric Lennard-Jones term and the directional hydrogen-bond term in the energy function. The third example, showing the application to cytochrome P450-cam, demonstrates that the GRID method is able to produce predictions in agreement with experimental observations for macromolecules.

The Catecholamines. These are derivatives of catechol with an aminoethyl side chain. As physiological effectors, they act on a number of different receptors, producing a variety of effects which are of great pharmacological interest. They are described here both because of their pharmacological importance and because they demonstrate the properties of the hydrogen-bond functions for hydroxyl groups and nitrogen atoms in program GRID.

The interaction energy of an uncharged NH probe, which is able to donate one hydrogen bond, with L-noradrenaline (I) in a crystallographically observed conformation[64] was calculated. The energy contours obtained



I: R = H (noradrenaline)
II: R = CH_3 (adrenaline)

at –3 kcal/mol are shown in Figure 7a. The three small contours show where a hydrogen bond from the NH probe can be donated to the phenolic hydroxyl groups. This hydrogen bond is favored in the plane of the aromatic ring at the trigonal angle to the phenolic C–O bonds. The phenolic hydrogens are shown in their crystallographically observed positions, but program GRID allows them and the hydroxyl oxygen lone pairs to exchange positions in the plane of the aromatic ring so that they can make the most energetically favorable interaction with the nitrogen NH probe. These phenolic hydroxyl groups are important for β-agonist activity and both are required for full agonist activity.[65]

The dominating circular region shows where a linear hydrogen bond can be made to the alkyl–hydroxyl group as its lone pairs rotate at the tetrahedral angle around the C–O bond. The ring of contours is thickest near the benzene ring and the alkyl side chain, where the attractive Lennard-Jones and hydrogen-bond interactions of the probe summate. This hydroxyl group plays a very important role in determining the activity and the binding affinity of the β-adrenergic agonists and antagonists.[64] If it is absent or in a different stereochemical configuration, potency is generally reduced.

Figure 7b shows D-noradrenaline arranged in the same overall conformation as L-noradrenaline, so that the catechol and nitrogen moieties of both molecules are in the same spatial relationship to each other. The energy contours at –3 kcal/mol for the same probe group do not form

(63) Bernstein, F. C.; Koetzle, T. F.; Williams, G. J. B.; Meyer, E. F.; Bryce, M. D.; Rogers, J. R.; Kennard, O.; Shikanouchi, T.; Tasumi, M. *J. Mol. Biol.* **1977**, *112*, 535.

(64) Carlstrom, D.; Bergin, R. *Acta Crystallogr.* **1967**, *23*, 313.

(65) Caron, M. C.; Mukherjee, C.; Lefkowitz, R. J. *Receptors in Pharmacology*; Smythies, J. R., Bradley, R. J., Eds.; Marcel Dekker: New York, 1978; p 97.

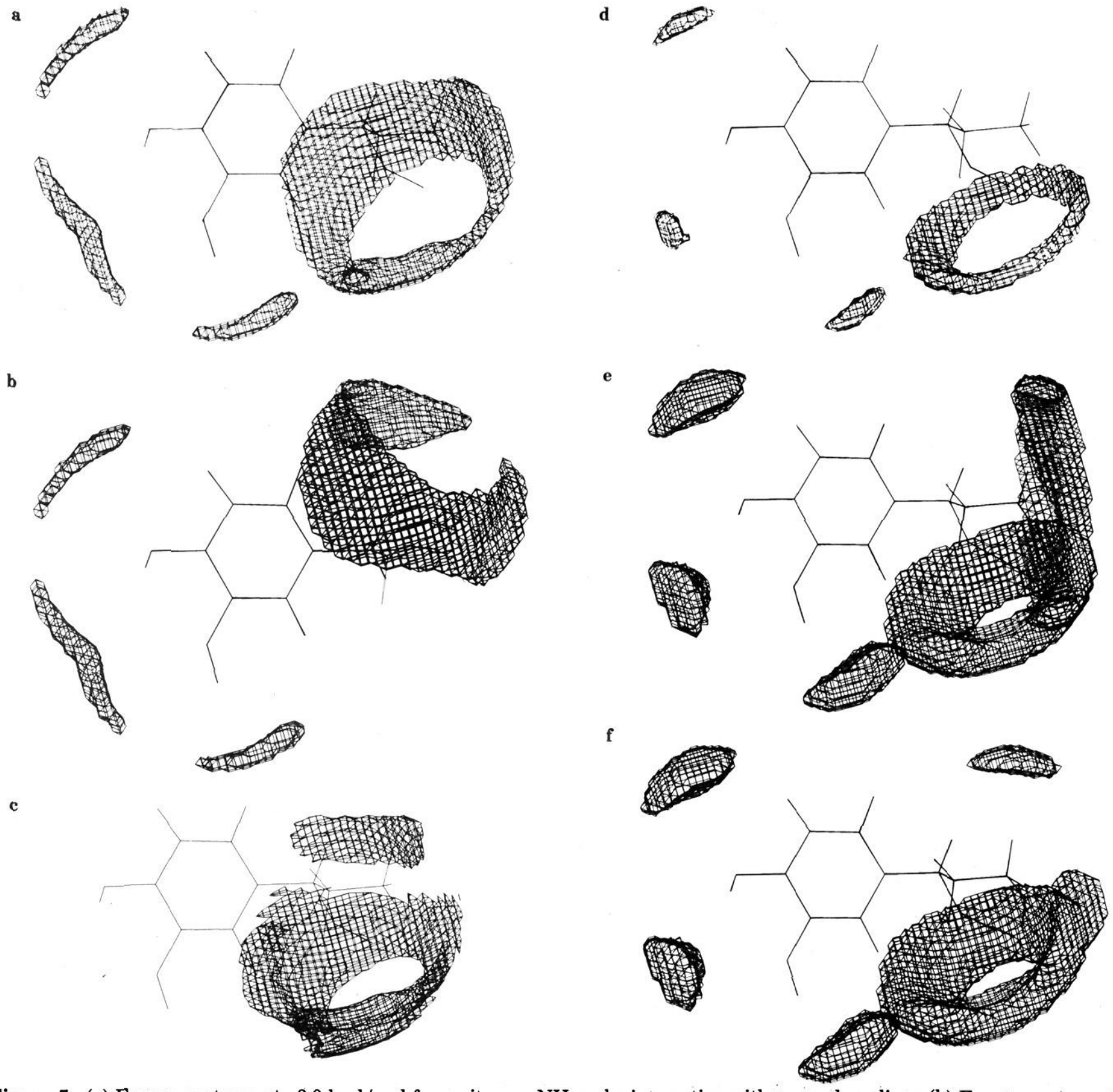


Figure 7. (a) Energy contours at −3.0 kcal/mol for a nitrogen NH probe interacting with L-noradrenaline. (b) Energy contours at −3.0 kcal/mol for a nitrogen NH probe interacting with D-noradrenaline. (c) Selectivity map with energy contours at −2.0 kcal/mol showing where the interaction of a nitrogen NH probe is more energetically favorable for L-noradrenaline than for D-noradrenaline. (d) Energy contours at −3.0 kcal/mol for a nitrogen probe which can accept one hydrogen bond, interacting with L-noradrenaline. (e) Energy contours at −2.0 kcal/mol for a nitrogen probe which can accept one hydrogen bond, interacting with L-noradrenaline. (f) Energy contours at −2.0 kcal/mol for a nitrogen probe which can accept one hydrogen bond, interacting with L-adrenaline. See text.

a full ring around the alkyl–hydroxyl group since there is now a repulsive steric interaction between the NH probe and the ammonium group of noradrenaline, which is closer to the alkyl–hydroxyl group in this conformation. The interaction of the two isomers can be compared in a selectivity map as shown in Figure 7c. This is constructed by calculating the difference in interaction energy of the probe with the L isomer and the D isomer at each point in the grid surrounding noradrenaline and contouring the difference energies at −2 kcal/mol. The two contoured regions in Figure 7c show where the probe interacts at least −2 kcal/mol more favorably with the L isomer than the D isomer. The uncontoured channel between the two regions shows where the two circular rings seen in parts a and b of Figure 7 overlap and mutually cancel each other out in the difference map. The volume excluded from the lower left-hand side of the main circular difference contour shows where a hydrogen bond can be made to the phenolic hydroxyl group in the D isomer but is made preferentially to the alkyl–hydroxyl group in the L isomer rather than the phenolic hydroxyl group. Such details are not immediately obvious when the conformations of the molecules are compared, but stand out clearly in appropriately contoured maps calculated by program GRID.

Figure 7d shows the interaction at −3 kcal/mol with L-noradrenaline of an uncharged nitrogen probe which is able to accept only one hydrogen bond. The small contoured regions show where the phenolic oxygens tend to donate an hydrogen bond in the plane of the aromatic ring at ±60° to the C–O bond. The phenolic hydrogens are allowed to adopt the position in the plane of the aromatic ring which gives the most energetically favorable interaction with the nitrogen probe. The contours at the phenolic oxygen are of a different shape from those in Figure 7a

because they are calculated with an angular dependence of a different functional form from that used when the phenolic oxygen accepts a hydrogen bond.

As before, the hydrogen of the alkyl–hydroxyl group on the target is able to rotate at the tetrahedral angle to the C–O bond, giving rise to the circular contour. The energy contours are at the same energy level as in Figure 7a, but the regions enclosed by the contours are seen to be smaller. This is because the hydroxyl hydrogen-bonding hydrogens of the target molecule have more influence on the direction of hydrogen bonds than the hydroxyl lone pairs do. A hydrogen bond is thus constrained to occupy a smaller range of angles when the target atom is donating.

Figure 7e again shows the interaction with L-nonadrenaline of an uncharged nitrogen N probe able to accept one hydrogen bond, but this time the map is contoured at −2 kcal/mol. This naturally makes all the original contours around the target oxygen atoms appear larger, but in addition to these contours there is an extra circular contour showing hydrogen bonding to the amine group which is assumed to exist in its ionized form as NH_3^+ at physiological pH.[66] These contours are circular because the three hydrogens of the ammonium group may rotate at the tetrahedral angle around the C–N bond. The contours are seen in this figure but not in Figure 7d, because the hydrogen bond to a nitrogen atom is weaker than that to an oxygen atom and does not show when a contour level of −3 kcal/mol is used. An idealized uncharged probe was used to generate Figure 7d,e, but in practice a probe nitrogen of this type would bear a partial negative charge, and this would result in a stronger interaction between the probe and the positively charged target nitrogen group.

Figure 7d,e demonstrates that it is important to choose a suitable contour level to display the interaction energies calculated by program GRID. When a new energy map is calculated, program GRID indicates where the deepest energy minimum occurs for the probe. The first GRID map should then be generated at an energy just above this minimum where it will show a small number of favorable probe binding sites. Maps contoured at higher energy levels can then be displayed in order to reveal further favorable binding sites. At higher but still negative contour energies the maps may become difficult to interpret, but maps contoured at slightly positive energies are useful because they show the shape of the target molecule as seen by the probe.

In Figure 7f, the interaction with L-adrenaline (II) of an uncharged nitrogen able to accept one hydrogen bond is shown contoured at −2 kcal/mol. Adrenaline differs from noradrenaline in having a methyl group substituted at its nitrogen, and its conformation was altered from that observed in the adrenaline crystal[67] in order to superimpose it on the previous noradrenaline conformations. The catecholamine nitrogen is again treated as being positively charged and has two hydrogens whose locations are now assumed to be explicitly determined by the two bonds from the nitrogen to two carbon atoms. This causes the energy contours around the target nitrogen atom to form two discrete regions rather than a continuous ring: one of these regions can be distinctly seen; the other is partly hidden behind the ring of contours around the rotating alkyl–hydroxyl group.

These energy maps of noradrenaline and adrenaline demonstrate the way in which a small change in the structure or conformation of a molecule can have a large impact on the molecule's capacity to interact favorably with a ligand. This is due to the specificity of the hydrogen bonds formed, which have a geometry which is sensitively dependent on the chemical nature and orientation of the hydrogen-bonding atoms. The GRID contours show both the overall character of the interactions and their fine detail in a clear and meaningful way.

Cardiac Glycosides. The cardiac glycosides are used in the treatment of heart failure as they cause positive inotropic action. Their activity varies according to their structure and alternative hypotheses have been proposed regarding their mode of binding. These suggest that binding occurs either by interactions with specific sites[68] or via a general charge interaction with the molecular dipole of the steroid[69] or by means of hydrophobic interactions with the steroidal nucleus.[70] Two compounds, digitoxigenin and gitoxigenin, consisting of the portion of the cardiac glycoside without the sugar residues, were studied because of the possible importance of hydrogen bonding in their therapeutic action, and because they demonstrate the properties of the GRID hydrogen-bond functions for a carbonyl oxygen target atom. Their structures, which were obtained from the Cambridge Crystallographic Database,[71] have previously been shown to be in minimum-energy conformations[72] and were first superimposed by doing a least-squares fit of atoms C1–C17, O3, and O4.

Gitoxigenin is much less active than digitoxigenin.[68,69] The directionality of a hydrogen bond to the carbonyl oxygen O23 on the lactone ring of these two compounds can be seen clearly by using an amide nitrogen probe which is able to donate one hydrogen bond. The energy contours, at −3 kcal/mol, near the lactone carbonyl oxygen atom, are quite different for digitoxigenin (Figure 8a) and gitoxigenin (Figure 8b). This is highlighted by the selectivity map (Figure 8c), contoured at −3 kcal/mol, and shown with the structure of gitoxigenin, which indicates where a stronger interaction with the probe could occur for digitoxigenin than for gitoxigenin. In addition to the two contours which occur at the anticipated position near the lactone oxygen on this map, there is a third contour due to the extra hydroxyl group O16 in gitoxigenin. This hydroxyl group is able to form an intramolecular hydrogen bond to O14, thus preventing the probe from hydrogen bonding to O14 when it is positioned near the coordinates of O16 and giving rise to a region where digitoxigenin can interact more favorably with the probe than gitoxigenin. In fact, the intramolecular hydrogen bond between O14 and O16 in gitoxigenin is thought to cause the lactone ring to be tilted and rotated so that the carbonyl oxygen O23 is 2.4 Å from its position in digitoxigenin,[72] and the selectivity map explicitly draws attention to this interesting feature. The decreased binding energy of gitoxigenin indicated by program GRID may contribute to its reduced activity compared to that of digitoxigenin.

Cytochrome P450-cam. Cytochrome P450-cam catalyzes the hdyroxylation of camphor to 5-*exo*-hydroxy-

(66) Ganellin, C. R. *J. Med. Chem.* **1977**, *20*, 579.
(67) Carlstrom, D. *Acta Crystallogr.* **1973**, *B29*, 161.
(68) De Pover, A.; Godfraind, T. *Naunyn-Schmiedeberg's Arch. Pharmacol.* **1982**, *321*, 135.
(69) Dittrich, F.; Berlin, P.; Kopke, K.; Repke, K. R. H. *Curr. Top. Membr. Transp.* **1983**, *19*, 251.
(70) Dzimin, N.; Fricke, U. *Br. J. Pharmacol.* **1988**, *93*, 281.
(71) Allen, F. H.; Bellord, S.; Brice, M. D.; Cartwright, B. A.; Doubleday, A.; Higgs, H.; Hummelink, T.; Hummelink-Peters, B. G.; Kennard, O.; Motherwell, W. D. S.; Rogers, R.; Watson, D. G. *Acta Crystallogr.* **1979**, *B35*, 2331.
(72) Griffin, J. F.; Rohrer, D. C.; Ahmed, K.; From, A. H.; Hashimoto, T.; Rathmore, H.; Fullerton, D. S. *Mol. Pharmacol.* **1986**, *29*, 270.

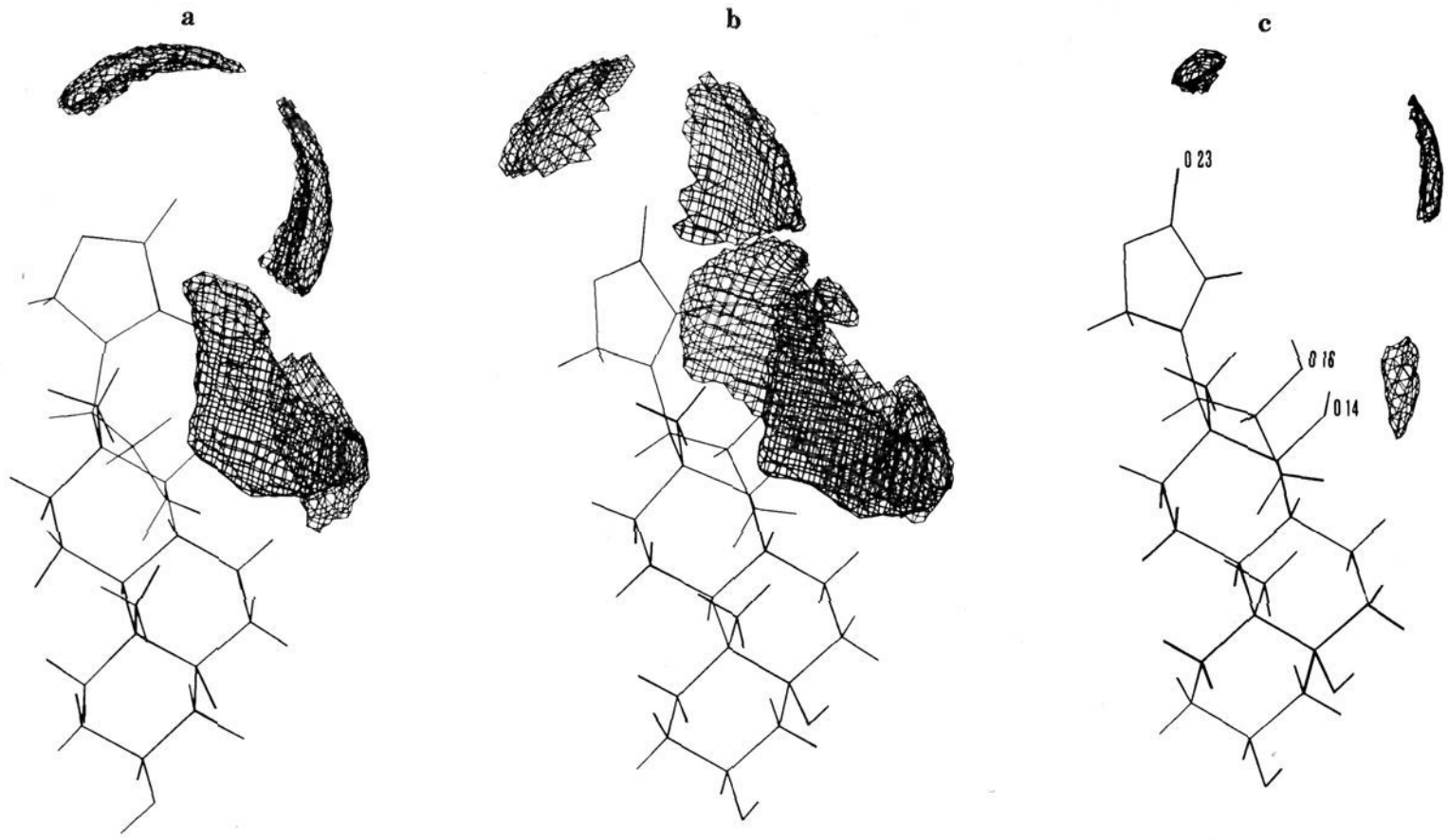


Figure 8. Energy contours at −3 kcal/mol for an amide nitrogen NH probe interacting with (a) digitoxigenin and (b) gitoxigenin. (c) The gitoxigenin molecule with a selectivity map with difference energy contours at −3 kcal/mol for an amide nitrogen NH probe showing regions where binding to digitoxigenin is favored by comparison to binding to gitoxigenin. For clarity, only energy contours near the lactone ring region of the molecules are shown. See text.

camphor using molecular oxygen. The protein structure has been observed by X-ray crystallography with[73,74] and without[75] the substrate bound. The use of program GRID on this enzyme is described in order to demonstrate how the overall energy potential of program GRID with the new hydrogen-bond terms can be applied to macromolecules and can produce results in accordance with known experimental observations.

It is thought[73] that the formation of a hydrogen bond between the camphor carbonyl oxygen and the hydroxyl oxygen of tyrosine 96 may hold the camphor molecule in the correct orientation in the binding site of cytochrome P450-cam with its 5-carbon exposed to the molecular oxygen and the heme group. The experimental structure shows that the camphor oxygen is in the plane of the tyrosine ring at 65.8° to the C–O axis and at a distance of 2.65 Å from the tyrosine oxygen atom. This arrangement would indicate favorable hydrogen-bond formation. However, it has also been suggested[74] that the interaction of the keto group with the tyrosine 96 hydroxyl group may only make a modest contribution to the free energy of binding because solvent will bind to tyrosine 96 in the absence of camphor.

Program GRID demonstrates the importance of the directionality of the hydrogen bonds made by the tyrosine 96 residue in this protein. The most energetically favorable position for a carbonyl oxygen probe in the active site is shown by the contours at −2.5 kcal/mol in Figure 9. These show where the carbonyl probe can hydrogen bond to the tyrosine 96 hydroxyl group and are at precisely the observed position of the carbonyl oxygen of the camphor substrate. An energy contour is not seen on the other side of the tyrosine C–O bond because the hydroxyl group of

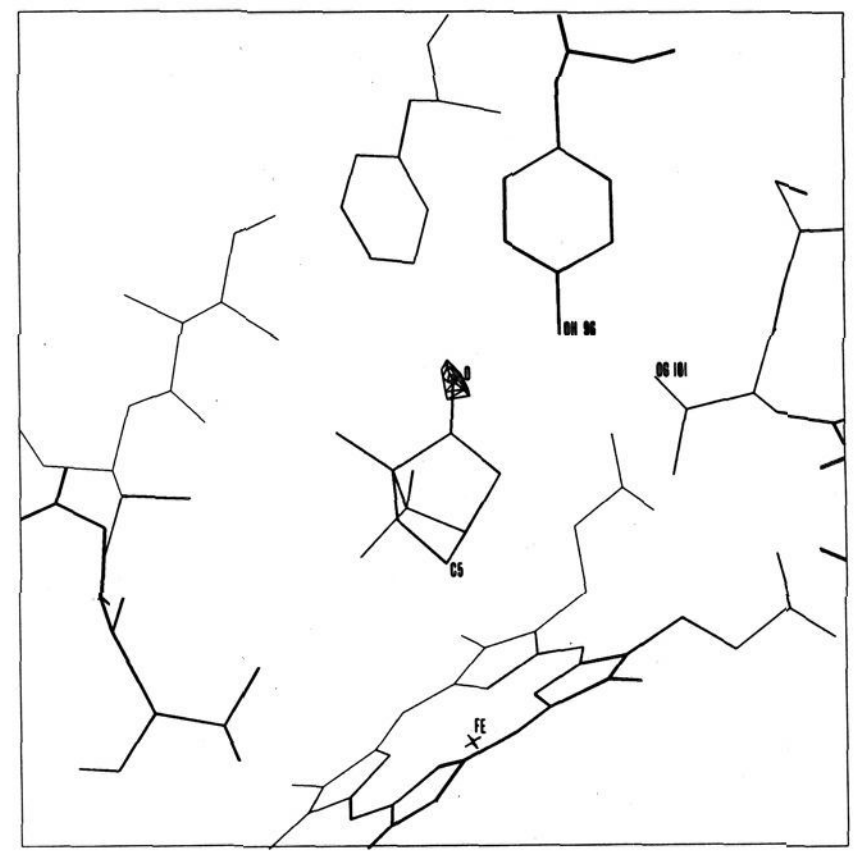


Figure 9. An energy contour map at −2.5 kcal/mol for a carbonyl oxygen probe at the substrate binding site of cytochrome P450-cam. See text.

threonine 101 is positioned to make an intramolecular hydrogen bond with tyrosine 96. Thus, program GRID distinguishes correctly between the hydroxyl groups of residues 96 and 101 and uniquely defines the tyrosine 96 binding site as the position where a favorable and specific interaction can be made.

The camphor molecule is hydrophobic apart from its carbonyl oxygen. In fact, the whole hydrocarbon skeleton of camphor is found within a region where a methyl probe can interact energetically favorably with the protein, and an energy contour at −2.0 kcal/mol for this methyl probe, as shown in Figure 10, defines the camphor position much more clearly than does the van der Waals surface of the

(73) Poulos, T. L.; Finzel, B. C.; Gunsalus, I. C.; Wagner, G. C.; Kraut, J. *J. Biol. Chem.* **1985**, *260*, 16122.
(74) Poulos, T. L.; Finzel, B. C.; Howard, A. J. *J. Mol. Biol.* **1987**, *195*, 687.
(75) Poulos, T. L.; Finzel, B. C.; Howard, A. J. *Biochemistry* **1986**, *25*, 5314.

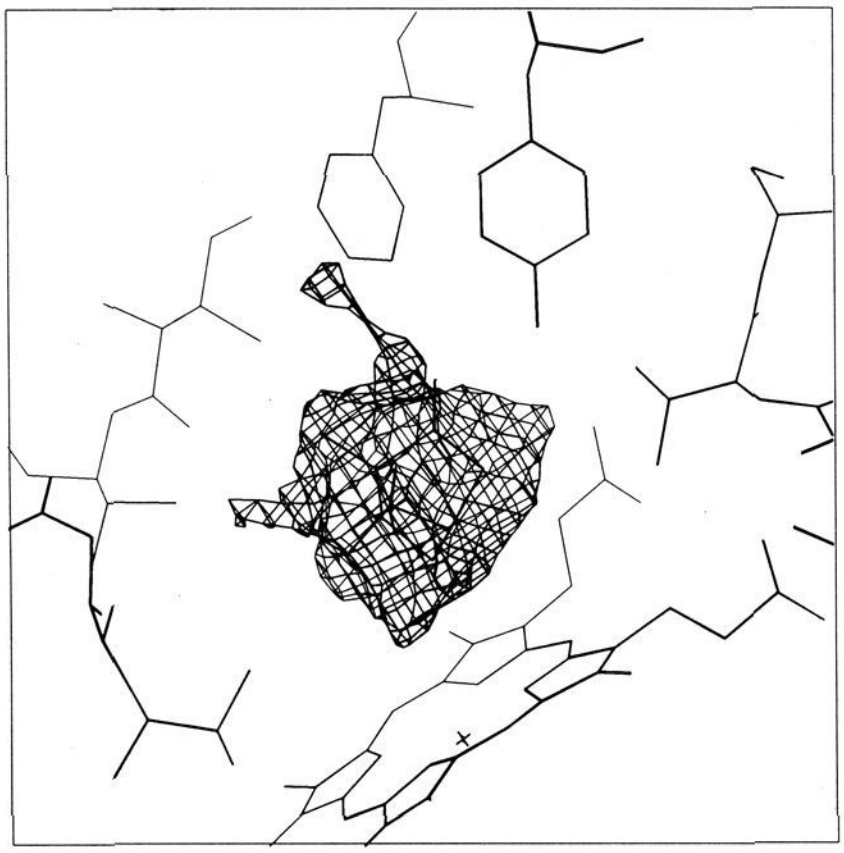


Figure 10. The energy contours at −2.0 kcal/mol for a methyl group probe interacting with cytochrome P450-cam. These energy contours surround the camphor substrate except for its polar carbonyl oxygen atom, and they define the position of the camphor molecule much more clearly than the van der Waals surface of the protein does.

protein itself. This shows the way in which program GRID can be used in sterically fitting a small molecule to a receptor site. It can be seen that the camphor oxygen atom alone is outside these methyl contours, showing that it is in a less hydrophobic region of the active site than the rest of the substrate molecule. The orientation of the camphor is apparently controlled by the precisely defined geometry of the hydrogen bond between the camphor oxygen and Tyr 96, while the position of the hydrophobic part of the substrate is determined by steric interactions with the surrounding protein.

Conclusion

An energy function has been developed for use in calculating the interaction between a probe group and a target molecule. The energy function consists of a Lennard-Jones, an electrostatic, and a hydrogen-bond term. There are sufficient experimental observations of hydrogen bonds in the literature to enable a set of hydrogen-bond functions to be modeled and fitted to experimental data. These hydrogen-bond functions are specific to different types of atom and model the variation in strength and geometry of the hydrogen bond according to the chemical nature of the donor and acceptor atoms as well as their position and orientation in the system under study. However, current experimental data for some types of atom are sparse and so functions to describe the hydrogen-bonding characteristics of certain atoms are still poorly defined. As more experimental data become available in the future, it should be possible to refine and improve such hydrogen-bond functions.

When appropriate, the hydrogen-bond term takes account of the mobility of the hydrogens and lone pairs analytically. Consideration of this property is vital in order that realistic predictions may be made. When a tautomeric change occurs on the approach of a ligand molecule, the functions allow for the appropriate charge redistribution in the target molecule. This treatment of molecular polarization is only applied to histidine residues in the present work, but it exemplfies an important principle which can be extended to other target systems.

A few examples of the energy maps have been shown in order to demonstrate the general shape of the predicted interaction between hydrogen-bonding groups. Predictions of substrate binding in cytochrome P450-cam demonstrate that GRID predictions can be accurate and specific even for a complicated system containing protein, a prosthetic group, a transition metal and water. The new energy function produces better results than the function originally used in program GRID, which only had a very simple and general hydrogen-bond term. It should be of value in the assessment of many different types of intermolecular interactions including those between drugs and their receptors.

Acknowledgment. Support from the Science and Engineering Research Council and the Medical Research Council is gratefully acknowledged.

Registry No. Tyrosine, 60-18-4; water, 7732-18-5; histidine, 71-00-1; L-noradrenaline, 51-41-2; D-noradrenaline, 149-95-1; L-adrenaline, 51-43-4; digitoxigenin, 143-62-4; gitoxigenin, 545-26-6; glucose, 50-99-7; cytochrome P450, 9035-51-2.

Analogues of Atriopeptin(103–125)amide Having High Binding Selectivity

Kerry L. Spear,*[,1] Monica S. Brown, Gillian M. Olins, and Dennis R. Patton

Cardiovascular Diseases Research Department, Searle Research and Development Division, G. D. Searle and Company, c/o Monsanto Company, 700 Chesterfield Village Parkway, Chesterfield, Missouri 63198. *Received May 20, 1988*

Analogues of atriopeptin(103–125)amide were prepared having a disulfide bridge at positions different from that found in the natural product. Most of these conformationally perturbed peptides were found to bind selectively to one subclass of binding sites. Binding affinity to a class of specific binding sites that is not associated with any known biological activity (nonvasorelaxant or NVR binding sites) is unaffected or even modestly improved. Affinity for the receptor subclass that is associated with vasorelaxation (VR subclass) decreases in most examples. In several cases, binding to the VR subclass is below the limits of detection for the assay used here. The data demonstrate that binding of atrial peptides to VR receptors requires rigidly defined receptor/ligand interactions. In contrast, the NVR subclass of binding sites appears to tolerate changes in peptide structure quite well.

The atrial peptides, a related group of cyclic hormones, are known to be potent vasorelaxants, natriuretics, and diuretics. These effects indicate a role in the regulation of fluid volume and blood pressure as well as electrolyte balance. As a result, this group of hormones, variously known as atriopeptin (AP), atrial natriuretic factor (ANF), and atrial natriuretic peptide (ANP), has been the subject

(1) Present address: The Marlstone Corporation, 24 Belvedere, Suite H, San Rafael, CA 94901.

0022-2623/89/1832-1094$01.50/0