10 Fulton, D. *et al.* (1999) Regulation of endothelium-derived nitric oxide production by the protein kinase Akt. *Nature* 399, 597–601
11 Dimmeler, S. *et al.* (1999) Activation of nitric oxide synthase in endothelial cells by Akt-dependent phosphorylation. *Nature* 399, 601–605
12 Chen, Z.P. *et al.* (1999) AMP-activated protein kinase phosphorylation of endothelial NO synthase. *FEBS Lett.* 443, 285–289
13 Kim, F. *et al.* (2001) TNF-α inhibits flow and insulin signaling leading to NO production in aortic endothelial cells. *Am. J. Physiol. Cell Physiol.* 280, C1057–C1065
14 Shayman, J.A. (2000) Sphingolipids. *Kidney Int.* 58, 11–26
15 Lee, M.J. *et al.* (1998) Sphingosine-1-phosphate as a ligand for the G protein-coupled receptor EDG-1. *Science* 279, 1552–1555
16 Lee, M.J. *et al.* (1996) The inducible G protein-coupled receptor edg-1 signals via the G(i)/mitogen-activated protein kinase pathway. *J. Biol. Chem.* 271, 11272–11279
17 Liu, Y. *et al.* (2000) Edg-1, the G protein-coupled receptor for sphingosine-1-phosphate, is essential for vascular maturation. *J. Clin. Invest.* 106, 951–961
18 Igarashi, J. *et al.* (2001) Sphingosine 1-phosphate and activation of endothelial nitric-oxide synthase. Differential regulation of Akt and MAP kinase pathways by EDG and bradykinin receptors in vascular endothelial cells. *J. Biol. Chem.* 276, 12420–12426
19 Igarashi, J. and Michel, T. (2000) Agonist-modulated targeting of the EDG-1 receptor to plasmalemmal caveolae. eNOS activation by sphingosine 1-phosphate and the role of caveolin-1 in sphingolipid signal transduction. *J. Biol. Chem.* 275, 32363–32370
20 Brenman, J.E. and Bredt, D.S. (1997) Synaptic signaling by nitric oxide. *Curr. Opin. Neurobiol.* 7, 374–378
21 Wang, Y. *et al.* (1999) RNA diversity has profound effects on the translation of neuronal nitric oxide synthase. *Proc. Natl. Acad. Sci. U. S. A.* 96, 12150–12155

Michael A. Marletta
Dept of Chemistry, 211 Lewis Hall, University of California, Berkeley, CA 94720 1460, USA.
e-mail: marletta@cchem.berkeley.edu

More hydrogen bonds for the (structural) biologist

Manfred S. Weiss, Maria Brandl, Jürgen Sühnel, Debnath Pal and Rolf Hilgenfeld

Why does a given protein structure form and why is this structure stable? These fundamental biochemical questions remain fascinating and challenging problems because the physical bases of the forces that govern protein structure, stability and folding are still not well understood. Now, a general concept of hydrogen bonding in proteins is emerging. This concept involves not only N–H and O–H donor groups, but also C–H, and not only N and O as acceptor groups, but also π-systems. We postulate that the incorporation of the entirety of these interactions leads to a more complete description of the problem, and that this could provide new perspectives and possibly new answers.

The world of proteins used to be simple. Hydrogen atoms bound to nitrogen and oxygen atoms formed hydrogen bonds with lone electron pairs on other oxygen and nitrogen atoms. These 'classical' hydrogen bonds[1,2] (Fig. 1a) have been held responsible for the formation of secondary structural elements such as α helices and β sheets and, along with van der Waals and hydrophobic forces, they constitute one of the main pillars of overall protein stability and a principal determinant of protein conformation. However, this seemingly simple picture is not able to provide more than a qualitative explanation of protein structure, folding and stability. A quantitative description that would allow the calculation and prediction of the energetics of these phenomena needs to be more elaborate and, indeed, the 'classical' view on hydrogen bonds has evolved considerably over the years. In 1982, Taylor and Kennard presented unequivocal evidence for the existence of hydrogen bonds between C–H donor groups and oxygen acceptors[3], and in the 1990s these hydrogen bonds were discovered for proteins at first[4] and then for other biological macromolecules[5] (Fig. 1b). Early this year, two papers dealing with even more exotic hydrogen bonds were published: Steiner and Koellner described hydrogen bonds in proteins involving aromatic acceptors[6] (Fig. 1c), and Brandl *et al.* exhaustively surveyed the occurrence of interactions involving all possible C–H groups (C^{α}–H, $C_{aliphatic}$–H and $C_{aromatic}$–H) as donors and all possible side chain π-systems as acceptors[7] (Fig. 1d). Burley and Petsko have termed these interactions 'weakly polar interactions'[8] but they can also be classified as hydrogen bonds, although they are considerably weaker than 'classical' hydrogen bonds. Because of their frequent occurrence in proteins, these interactions can be expected to contribute significantly to the overall stabilization energy of a protein, which is often not more than a few kcal mol^{-1}. Of course there is no reason to restrict their potential importance to just the intramolecular protein context itself. As the respective donor and acceptor groups also occur on protein surfaces, these interactions can be used to recognize and selectively bind other proteins, ligands, substrates, inhibitors and so forth.

The nature of these various types of hydrogen bonds can be described qualitatively by Pearson's hard/soft–acid/base (HSAB) concept[9]. If N–H and O–H groups are considered hard acids, S–H groups intermediate acids and C–H groups soft acids, and oxygen and nitrogen atoms the hard bases, sulfur atoms intermediate bases and π-systems soft bases, then soft-acid···soft-base hydrogen bonds should form as do hard-acid···hard-base hydrogen bonds. A more quantitative description can be provided by quantum-chemical *ab initio* calculations, which suggest that at least four different attractive energy terms ought to be considered: (1) electrostatic energy arising from interactions between charges, partial charges and dipoles; (2) charge-transfer or delocalization energy; (3) polarization energy from interactions between permanent dipoles and induced dipoles; and (4) dispersion energy originating from interacting temporary dipoles and induced dipoles. These stabilizing energies are counterbalanced with a destabilizing, repulsive energy, termed exchange repulsion energy. The various types of hydrogen bonds can now be distinguished by how much the different energy terms contribute to the overall energy of a hydrogen bond. The hard-acid···hard-base hydrogen bonds are dominated by electrostatic and charge-transfer energy, and the soft-acid···soft-base hydrogen bonds by

 0968-0004/01/$ – see front matter PII: S0968-0004(01)01935-1

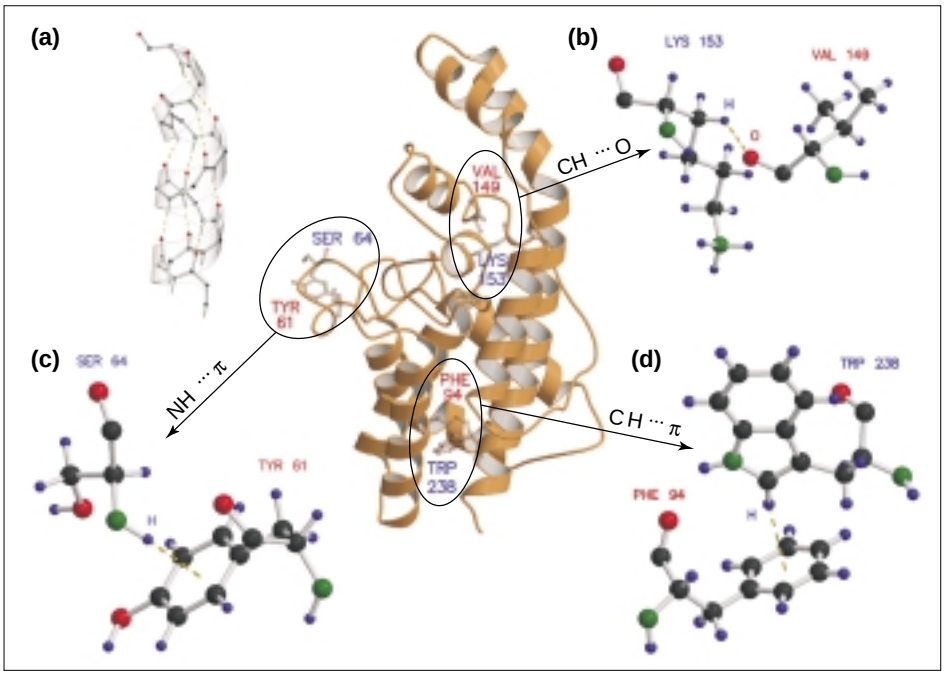

Fig. 1. Central panel: Molscript[20] diagram of phospholipase C (PLC, PDB entry 1AH7). The different hydrogen bond classes are highlighted as four enlarged views of the structure. (a) 'Classical' N–H···O hydrogen bonds as found in α helices (shown here are the main chain atoms of the helix 193–204 of PLC). (b) C^{β}–H···O hydrogen bond between Lys153 and Val149. (c) N–H···π hydrogen bond between Ser64 and the aromatic ring of Tyr61. (d) $C^{\delta1}$–H···π hydrogen bond between Trp238 and the aromatic ring of Phe94. The figure was rendered using the program Raster3D (Ref. 21).

dispersion and polarization energy[10]. As a consequence, soft-acid···soft-base hydrogen bonds are persistent in both highly polar solvents such as H_2O and apolar environments such as those found in the interior of proteins. The different types of hydrogen bonds also appear to exhibit a different behaviour when studied by vibrational spectroscopy. Observed frequency shifts of D–H (where D represents any atom) bands are usually a good indication for the formation of a hydrogen bond and they can also be used to distinguish between different hydrogen bond types. The 'classical' hydrogen bonds always show a red shift of the D–H (where D represents N or O) stretching frequency upon hydrogen bond formation. This implies that the D–H bond is weakened and concomitantly the D–H bond length increased. By contrast, the C–H stretching frequency can be either red- or blue-shifted and accordingly the C–H bond lengthened or shortened, respectively[10,11].

To be fair, it should be mentioned that the concept of hydrogen bonds involving C–H donor groups and/or π-acceptor groups is not new. For many years, it has been well known and widely accepted in the field of small organic and inorganic molecules, where inclusion compounds or host-guest complexes often form solely on the basis of these weak hydrogen bonds. The recently published books by Nishio *et al.*[12] and Desiraju and Steiner[13] provide excellent reviews on this topic. However, the protein community has been slow in recognizing the importance of these weak hydrogen bonds for proteins, although some papers of the past decade have pointed in this direction (see, for example, Ref. 14). It can only be hoped that the recent findings will initiate a re-evaluation of this scepticism.

To shed some light on the relative importance of the various types of hydrogen bonds in protein structures, a complete hydrogen-bond analysis of two high-resolution protein structures taken from the Protein Data Bank was carried out. Phospholipase C (PLC; PDB entry 1AH7, 1.5 Å resolution[15]) is a protein of mainly α-helical structure consisting of 245 amino acid residues (Fig. 1, central panel), and human heparin-binding protein (HBP; PDB entry 1A7S, 1.12 Å resolution[16]) is a β-sheet protein comprising 221 residues. The total numbers of identified hydrogen bonds in each class are given in Table 1; the criteria for their identification are outlined in the table legend. The four classes of hydrogen bonds described are also illustrated in Fig. 1 using examples identified in the PLC structure.

From the frequency of occurrence of the various interactions presented in Table 1, it can be inferred that for both all-α and all-β proteins, the 'classical' hydrogen bonds dominate the picture by their sheer number. The absolute numbers presented should certainly be taken with a grain of salt as they represent the numbers for individual proteins rather than those for whole groups of proteins. Nevertheless, they are able to reveal some general trends. Forming ~20–25% of the total number of hydrogen bonds, C–H···O interactions constitute the second most important group. It has already been described for proteins containing a large fraction of β-structure, that C^{α}–H···O hydrogen bonds occur almost ubiquitously in β sheets[4], but they also appear to occur frequently in α-helical proteins. Hydrogen bonds with π-acceptors constitute yet another considerable fraction. Even though π-systems are relatively infrequently observed to interact with O–H, N–H and S–H donors, they seem to exhibit a distinct preference to team up with C–H donors, thus forming C–H···π hydrogen bonds. The formation of these bonds is consistent with the HSAB concept. Again, we would like to emphasize that the observation that more C–H···π hydrogen bonds occur in PLC than in HBP, does not mean that these hydrogen bonds are more abundant in all-α proteins than in all-β proteins; it merely reflects the possible spread of the numbers. Brandl *et al.* had shown previously that proteins can contain anywhere between 0.0 and 22.6 C–H···π hydrogen bonds per 100 amino acids[7]. It has also been noted in some cases that these weaker hydrogen bonds occur at molecular interfaces and that they are important for molecular recognition[14,17,18].

In terms of hydrogen bonding energies, many battles have been fought over the years. The commonly accepted numbers now range from >10 kcal mol^{-1} for the strongest (e.g. O–H···O^-) bonds[2] to ~0.5–1.0 kcal mol^{-1} for C–H···π hydrogen bonds[12]. The 'classical' N–H···O hydrogen bonds commonly observed in biological macromolecules are considered to be of intermediate strength, ~5–6 kcal mol^{-1} (Ref. 2). Historically, this wide range of

Table 1. Complete hydrogen-bond analysis for phospholipase C (PLC) and human heparin-binding protein (HBP)

Hydrogen-bond type	PLC (all-α; 245 amino acids) (number of hydrogen bonds)[a]	HBP (all-β; 221 amino acids) (number of hydrogen bonds)[a]
D–H···A (D = N,O; A = N,O,S)[b]	333 (135.9)	192 (86.9)
S–H···A (A = N,O,S)[c,d]	–	–
C–H···A (A = N,O,S)[e]	82 (33.5)	81 (36.7)
D–H···π (D = N,O)[f]	2 (0.8)	–
S–H···π[d,g]	–	–
C–H···π[h]	50 (20.4)	20 (9.0)
Total	467 (190.6)	293 (132.6)

[a]Numbers represent the total number of hydrogen bonds found (numbers in parentheses represent the number of hydrogen bonds per 100 amino acid residues).
[b]Hydrogen bonds were selected when the distance between the donor atom D and the acceptor atom A (d_{D-A}) was ≤3.5 Å, the angle at the hydrogen atom ($\angle_{D-H-A}$) ≥90°, and the angle at the acceptor atom also ≥90°. These are the standard Baker-and-Hubbard criteria for hydrogen bonds[1].
[c]d_{S-A} ≤4.5 Å, $\angle_{S-H-A}$ ≥120°.
[d]Hydrogen bonds with S–H donor groups could not occur in the two cases studied because PLC does not contain a cysteine residue and all eight cysteine residues in PLC are involved in disulfide bridges.
[e]d_{C-A} ≤3.5 Å, $\angle_{C-H-A}$ ≥120°. To discriminate between a C–H···O interaction and a C–H···π interaction in cases when the acceptor O atom was part of a π-system, the selection criteria had to be expanded by an additional geometric constraint. For a C–H···O interaction to occur, the donor C atom had to be within a distance of 2.0 Å to the plane of the π-system containing the acceptor O atom.
[f]d_{D-X} ≤4.3 Å, where X is the center of the π-system[6], $\angle_{D-H-X}$ ≥120° and d_{Hp-X} ≤1.0 or 1.2 Å depending on the π-system in question[7].
[g]d_{S-X} ≤4.5 Å, $\angle_{S-H-X}$ ≥120° and d_{Hp-X} ≤1.0 or 1.2 Å.
[h]d_{C-X} ≤4.5 Å, $\angle_{C-H-X}$ ≥120° and d_{Hp-X} ≤1.0 or 1.2 Å.

interaction energies was one of the reasons why the weaker hydrogen bonds have been neglected for such a long time. Interestingly however, Scheiner *et al.*[19] recently published a quantum-mechanical study on C–H···O hydrogen bonds between C^{α}–H atoms of amino acids and H_2O molecules. They found that such a hydrogen bond is about half as strong as a 'classical' one between two H_2O molecules and, in the case of the positively charged lysine, the C^{α}–H···O bond is thought to be even stronger.

The numbers presented unequivocally show that the weaker interactions cannot and must not be neglected. Albeit weak, they are numerous and therefore might help explain the well-known problem that protein stabilities, interaction energies and folding energies cannot be calculated very accurately. The consideration of these important interactions might enhance the usefulness of these calculations in general, and further our understanding of protein structures and their functions.

References

1 Baker, E.N. and Hubbard, R.E. (1984) Hydrogen bonding in globular proteins. *Prog. Biophys. Mol. Biol.* 44, 97–179
2 Jeffrey, G.A. and Saenger, W. (1994) *Hydrogen Bonding in Biological Systems*, Springer Verlag
3 Taylor, R. and Kennard, O. (1982) Crystallographic evidence for the existence of C–H···O, C–H···N, and C–H···Cl hydrogen bonds. *J. Am. Chem. Soc.* 104, 5063–5070
4 Derewenda, Z.S. *et al.* (1995) The occurrence of C–H···O hydrogen bonds in proteins. *J. Mol. Biol.* 252, 248–262
5 Wahl, M.C. and Sundaralingam, M. (1997) C–H···O hydrogen bonding in biology. *Trends Biochem. Sci.* 22, 97–102
6 Steiner, T. and Koellner, G. (2001) Hydrogen bonds with π-acceptors in proteins: frequencies and role in stabilizing local 3D structures. *J. Mol. Biol.* 305, 535–557
7 Brandl, M. *et al.* (2001) C–H···π-interactions in proteins. *J. Mol. Biol.* 307, 357–377
8 Burley, S.K. and Petsko, G.A. (1988) Weakly polar interactions in proteins. *Adv. Prot. Chem.* 39, 125–189
9 Pearson, R.G. (1963) Hard and soft acids and bases. *J. Am. Chem. Soc.* 85, 3533–3539
10 Hobza, P. and Havlas, Z. (2000) Blue-shifting hydrogen bonds. *Chem. Rev.* 100, 4253–4264
11 Brandl, M. *et al.* (2001) Quantum-chemical analysis of C–H···O and C–H···N interactions in RNA base pairs – H-bond versus anti-H-bond pattern. *J. Biomol. Struct. Dyn.* 18, 545–555
12 Nishio, M. *et al.* (1998) *The CH/π Interaction. Evidence, Nature and Consequences*, Wiley VCH, New York
13 Desiraju, G. and Steiner, T. (1999) *The Weak Hydrogen Bond in Structural Chemistry and Biology*, Oxford University Press
14 Perutz, M.F. (1993) The role of aromatic rings as hydrogen-bond acceptors in molecular recognition. *Proc. R. Soc. Ser. A* 345, 105–112
15 Hough, E. *et al.* (1989) High resolution (1.5 Å) crystal structure of phospholipase C from *Bacillus cereus*. *Nature* 338, 357–360
16 Karlsen, S. *et al.* (1998) Atomic resolution structure of human HBP/CAP37/azurocidin. *Acta Crystallogr. D* 54, 598–609
17 Umezawa, Y. and Nishio, M. (1998) CH/π interactions as demonstrated in the crystal structure of guanine-nucleotide binding proteins, Src homology-2 domains and human growth hormone in complex with their specific ligands. *Bioorg. Med. Chem.* 6, 493–504
18 Umezawa, Y. and Nishio, M. (1998) CH/π interactions in the crystal structure of class I MHC antigens and their complexes with peptides. *Bioorg. Med. Chem.* 6, 2507–2515
19 Scheiner, S. *et al.* (2001) Strength of the $C^{\alpha}H$···O hydrogen bond of amino acid residues. *J. Biol. Chem.* 276, 9832–9837
20 Kraulis, P.J. (1991) MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.* 24, 946–950
21 Merritt, E.A. and Murphy, M.E.P. (1994) Raster3D Version 2.0. a program for photorealistic molecular graphics. *Acta Crystallogr. D* 50, 869–873

Manfred S. Weiss*
Maria Brandl
Jürgen Sühnel
Debnath Pal
Rolf Hilgenfeld
Institute of Molecular Biotechnology, Beutenbergstr. 11, D-07745 Jena, Germany.
*e-mail: msweiss@imb-jena.de