C–H···O and other weak hydrogen bonds. From crystal engineering to virtual screening

Gautam R. Desiraju

DOI: 10.1039/b504372g

Hydrogen bonds, X–H···A, formed by weak donors (X = C) and acceptors (A = π system) were generally dismissed as being of little consequence before and even during the 1970s. This situation changed in the early 1980s, and during the two following decades they were implicated as being significant in many small molecule crystal structures, and also in solution. Today, knowledge gained about these interactions is being used to understand the structure of biomolecules with implications for structure based drug design.

My first encounter with the C–H···O hydrogen bond came in 1985 while trying to understand the crystal structures of some alkoxycinnamic acids.[1] These structures have a characteristic short axis of *ca.* 4.0 Å, which follows from the stacking of layers of flat aromatic molecules that are stitched laterally with directional interactions. That one of these lateral interactions was a C–H···O contact did not appear surprising. After all, a C–H group does have some acidity, and even if it is not highly activated like N–H and O–H groups, the idea that a continuous property like acidity would lead to hydrogen bonds X–H···O (X = C, N, O) with graded strengths did not seem particularly counter-intuitive or anti-scientific. After reading Leiserowitz's marvellous, and early, review of hydrogen bonding in carboxylic acids,[2] we tried to look at IR bathochromic shifts for ν_{C-H} in C–H···O bonds formed by terminal acetylenes and correlate them with *D*, the respective hydrogen bond C···O distances (Fig. 1). The correlation was good and suggested that the C–H···O is a genuine interaction with predictable consequences.[3] However, my real awakening to the possibilities for this interaction in structural chemistry came in 1989 when I attempted to correlate *D* values for contact geometries of the type $Cl_{3-n}(R_n)C–H\cdots O$ as one proceeds from the distinctly acidic $CHCl_3$ to the very feebly acidic $R_3C–H\cdots O$, *via* compounds of the type $Cl_2(R)C–H$ and $Cl(R_2)C–H$.[4] The correlation was very good and extended to the least acidic compounds in this series ($D \approx 4.0$ Å, Fig. 1), indicating that C–H···O interactions could be formed by many compounds, not just activated ones like acetylene and chloroform.[5] The strategy for these analyses was surely influenced by the inspiring 1982 Taylor–Kennard paper on the relevance and use of the Cambridge Structural Database in understanding the properties of the C–H···O hydrogen bond.[6] On the matter of C–H···O bonds and crystal packing, these authors wrote, "the frequency with which they occur suggests that they play a significant role". This sentence stands out after all these years: at that time (the early 1990s), it challenged us to embark on a major effort in which we tried to assess and establish the viability of the C–H···O bond in crystal engineering, the design of crystal structures with particular properties.

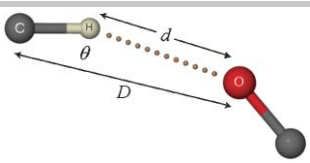

Fig. 1 Definitions of the geometrical parameters *d*, *D* and *θ* for a C–H···O hydrogen bond. The H-atom position should be neutron normalized for systematic analysis.

While attempting to evaluate the role of C–H···O bonds in crystal engineering, we came up against two hurdles almost immediately. The first was concerned

Gautam R. Desiraju

Gautam R. Desiraju (b. 1952 in Madras) studied chemistry at the Universities of Bombay and Illinois. He has been at the University of Hyderabad since 1979 and has made seminal contributions to the growth and development of the subject of crystal engineering, with particular reference to the properties of the weak hydrogen bond. He has written more than 250 scientific papers and is the author of 'Crystal Engineering. The Design of Organic Solids' (Elsevier, 1989) and 'The Weak Hydrogen Bond in Structural Chemistry and Biology' (with T. Steiner, OUP, 1999). He has lectured widely and is a recipient of the Alexander von Humboldt Forschungspreis (2000). He was on the Editorial Advisory Board of Chemical Communications between 2000 and 2004, and had exactly 40 papers in the journal at the beginning of its 40th year.

with the nature of the interaction: is the C–H···O interaction a hydrogen bond, in the usually understood sense, or is it merely a geometrical construct that is widespread because statistics favours a C–H group approaching an O-atom in organic crystals? The second pertains to the idea of engineering a crystal structure: even if the C–H···O is a hydrogen bond, it is clearly weaker than the typical N–H···O or O–H···O bond. So, does it really guide packing preferences, or is it content to just remain as a bystander watching the structural landscape? These questions are not completely independent: they are connected by the fact that the C–H···O interaction is weak. So, the discussion gradually moved to questions pertaining to strength *vs.* weakness. If the C–H···O interaction is weak, how is it that the interaction metrics (lengths, angles) are well conserved in whole groups of crystal structures? Are many weak interactions more effective than a few strong ones? What is the role of cooperativity in sustaining patterns of interactions? Is the C–H···O interaction indeed as weak as was presumed? Much of this discussion is elaborated in a book entitled *The Weak Hydrogen Bond in Structural Chemistry and Biology*[7] that I wrote with Thomas Steiner in 1999 and I will not repeat myself here. However, questions of the type I have posed above lead easily to biological issues, and studies of the weak hydrogen bond progressed naturally from chemistry to structural biology. Crystal engineering deals with supramolecular assembly of a certain type. Ligand–receptor recognition is supramolecular assembly of another type, and we and others have found that concepts about C–H···O bonds and other weak interactions that arose from crystal engineering and structural chemistry may be conveniently applied to rational drug design, in particular to understanding protein structure, ligand docking and virtual screening.

Is the C–H···O interaction a hydrogen bond? Of course, one could avoid answering this question by asking, "What is a hydrogen bond?" But little is gained by scoring debating points. What is undisputed is that there are certain types of hydrogen bonds, formed by highly activated C–H groups, that are practically indistinguishable from conventional or strong hydrogen bonds of the N–H···O or O–H···O type.[8] What happens as one descends in the scale of carbon acidity is more equivocal. C–H···O contact geometries are still found but the geometrical, spectroscopic and energetic criteria that are traditionally used to characterize hydrogen bonds become more fuzzy, till one is at the point where one asks if a C–H···O contact formed by an unactivated Me-group, with $D \cong 3.5$ Å, and which has a stabilization of around 0.5 kcal mol^{-1} is even worth considering. At this stage, there are no hard and fast rules, and circumstantial evidence becomes important. Consider for example, the case of 4-chlorocubanecarboxylic acid, **1** which forms the very rare *syn*,*anti* O–H···O catemer in its crystal structure (Fig. 2).[9] That the catemer is stabilized by an appropriate C–H···O bond donated by the (activated) cubyl C–H group is deduced not so much from the geometry of this latter interaction but rather from the fact that similar catemers are obtained in the seemingly unrelated indole-2-carboxylic acid, **2**, and phenylpyruvic acid, **3**, wherein equivalent support is provided by a corresponding N–H···O and O–H···O bond respectively. The rarity of these catemers (*vis-à-vis* the O–H···O hydrogen bonded dimer, which is the preferred motif for carboxylic acids) provides the necessary confidence in claiming that the C–H···O is a true hydrogen bond and not just a coincidental confluence of functionalities in the crystal. Other examples of interaction mimicry are equally convincing,[10] and hydrogen bonds formed by C–H donors are now accepted as genuine.

Much of the discussion on where the hydrogen bonding phenomenon stops, is provided by the weakest of these interactions, notably C–H···F–C and

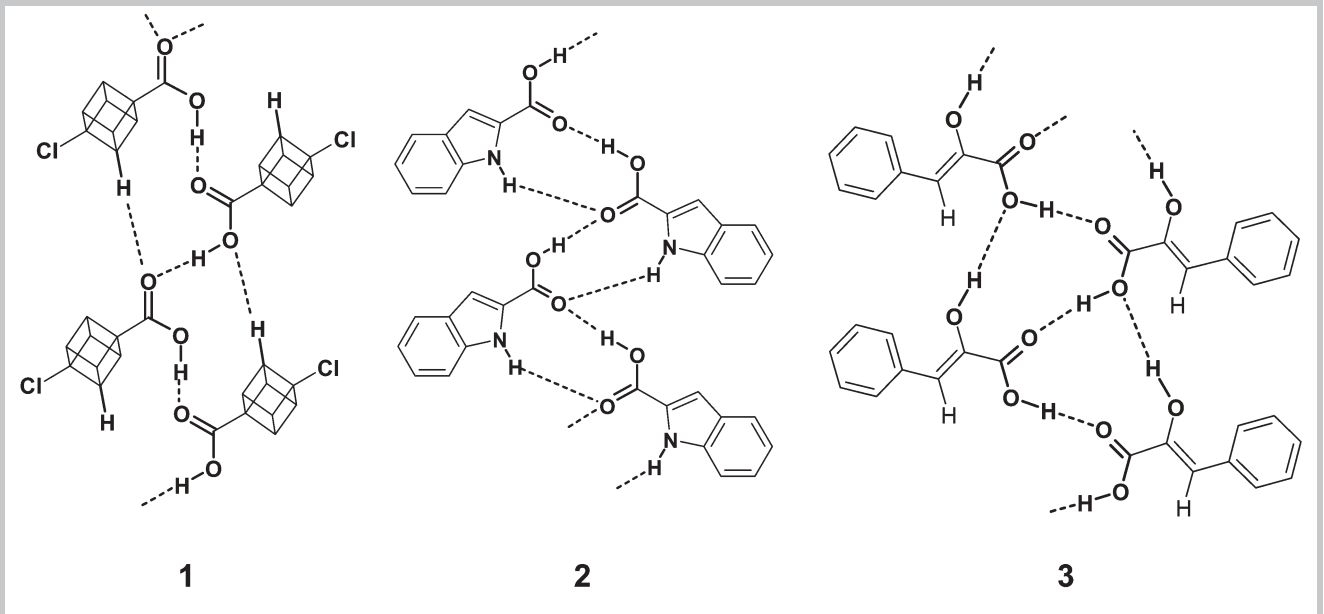

Fig. 2 Catemer formation with O–H···O hydrogen bonds in carboxylic acids. The catemers are stabilized by the supporting interactions, C–H···O in **1**, N–H···O in **2**, and O–H···O in **3**.

C–H···π. Not coincidentally, both these interactions are important in ligand–receptor interactions with implications for drug design. The C–H···F–C interaction is indeed weak, and we could barely notice it in the crystal structures of the polyfluorinated benzenes—compounds wherein the dice had been loaded to observe this interaction.[11] But despite (or because of) its weakness, it is important in widely different areas of chemistry. Diederich and co-workers have obtained evidence for the interaction in the binding of thromboxane receptors and have implicated it in drug design strategies,[12] while Chan and co-workers have shown that *only an interaction as weak as the C–H···F–C* may successfully guide the course of stereocontrolled polymerization with Zr based agostic type catalysts.[13] The C–H···π interaction is much more ubiquitous; a comprehensive review by Nishio outlines its possibilities and scope.[14] What is important here is the idea that an interaction with much hydrophobic character, still retains some electrostatic nature—enough anyway so that vestiges of hydrogen bonding persist. Other manifestations of this hydrophobic character in a hydrogen bond are seen in the so-called blue shifted hydrogen bonds popularized by Hobza,[15] and metal atom based hydrogen bonds, where a transition metal acts as an acceptor (X–H···M) or as a donor (M–H···O); these have been studied and reviewed independently by Braga[16] and Brammer.[17] Computational studies of all these weak interactions are still in a state of infancy.[18] Experimental charge density studies followed by analysis with the Koch–Popelier criteria using Bader's atoms-in-molecules approach have provided some insights. But in the end, one is still not much closer to establishing where a hydrogen bond ends and a van der Waals interaction begins.[19] In part, the problem is linguistic and I advocated a return to the pre-Pauling term *hydrogen bridge* to describe (especially) the weaker variants of hydrogen bond.[20] After all, the word *bridge* carries with it no special chemical connotations; in effect, each one of us can interpret the term *hydrogen bridge* in whatever way we want, and this may not be so bad during this interim period when complete consensus on the nature of the interaction is absent.

Can one engineer crystal structures with C–H···O and other weak hydrogen bridges? We have suggested that the roles of these interactions in crystal packing may be classified as innocuous, supportive or intrusive.[7] This is a useful categorization but, arguably, subjective because it depends on what one feels the crystal structure ought to look like. Innocuous interactions are passive bystanders; they are very weak and merely exist in a structure that is almost wholly determined by other interactions. Supportive C–H···O interactions are not so weak but their directional preferences are satisfied within the geometrical constraints of the stronger interactions; the C–H···O bonds in 1,4-benzoquinone appear to be very satisfactory but the major interactions in terms of energetics arise from the stacking of planar layers. Intrusive interactions perturb the patterns and topologies of the stronger interactions. We have described a few examples: 3,5-dinitrocinnamic acid forms an O–H···O acid dimer but, on account of the numerous and strong C–H···O bonds, this synthon lies not on an inversion centre in the crystal but on a 2-axis;[21a] some geminal alkynols wherein the O–H and C≡C–H are sterically constrained do not form O–H···O–H hydrogen bonds at all but rather C–H···O and O–H···π interactions;[21b] some phenols have finite O–H···O–H patterns instead of the presumably preferred infinite patterns, because weak hydrogen bonds which act as chain stoppers are favoured by cooperative effects.[21c] In the end, when such crystal structures are better understood, it will be possible to treat all hydrogen bonds in a crystal together without unnecessarily discriminating between the (strong) O–H···O and (weak) C–H···O varieties. The implications of being able to do this would be especially beneficial in the study of the crystal structures of biological macromolecules.

As for engineering structures based on weak hydrogen bonds, I will confine myself to two examples from our work; length restrictions prevent me from mentioning the many papers that have been published by crystal engineering groups worldwide. Fig. 3 shows the crystal structure of 1,3-dibromo-2,4,6-trinitrobenzene, **4**. This compound is a two-dimensional charge transfer hyperpolarisable chromophore and crystallises in the non-centrosymmetric space group *C*2 in perfect polar order leading to an intense powder SHG signal at 1.06 μm.[22] The crystal structural analysis revealed the formation of layers parallel to the (302) plane, within which the chromophores adopt a hexagonal packing. Within each layer all the chromophores are arranged in a head-to-tail fashion oriented along the C(1)–C(4) molecular axis that is also the molecular dipole axis. The supramolecular hexagons are assembled with bifurcated C–H···O (2.60 Å, 155.2°) hydrogen bonded and Br···O_2N (2.93 Å, 169°) supramolecular synthons. The layers are interconnected

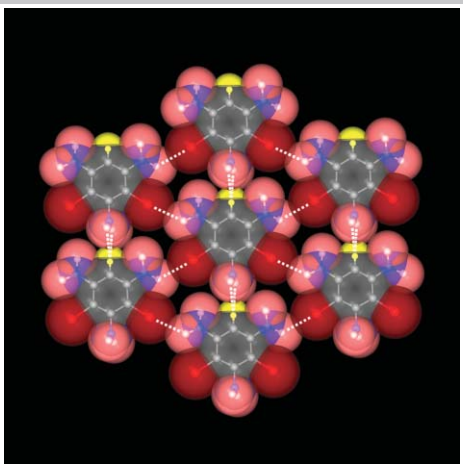

Fig. 3 Hexagonal arrangement of 1,3-dibromo-2,4,6-trinitrobenzene, **4**, in the crystal structure. Notice the bifurcated C–H···O interactions, and the perfect polar order of NLO chromophores. Note also that the C–H group is highly activated.

by very weak additional Br···O_2N interactions (3.48 Å), and stacked in a parallel fashion so that all the chromophores are oriented in the same direction. This crystal packing leads to a complete additivity of the molecular β tensor components for the macroscopic $\chi^{(2)}$ susceptibility. For a bifurcated interaction, the C–H···O in **4** is very short, and because the C–H group is highly activated, one may safely call this a proper hydrogen bond, and not a consequence or result of other packing features.

In the realm of host–guest chemistry, we have found that tetrakis(4-nitrophenyl)methane, **5**, is a versatile host that gives three types of host–guest complexes which we have termed guest-rich, host-rich and guest-excess.[23] The molecule contains many activated C–H groups and nitro groups (which are known to be good C–H···O acceptors[24]). In the guest-rich complexes (host : guest ratio between 1 : 1 and 2 : 1), the host molecules are linked with pairs of C–H···O bridges to give large diamondoid nets that are interpenetrated (Fig. 4). There is enough room in the channels that are thus created along the direction of interpenetration for the location of guest molecules like THF, dioxane, nitrobenzene and anisole. The host-rich structures (host : guest ratio 3 : 1) have rhombohedral symmetry. The host molecules are connected to each other and to the guest molecule with strong C–H···O bonds. These solids may be heated to melting without loss of solvent. However, when guest-rich complexes are heated or evacuated under mild conditions, they undergo partial solvent loss to yield complexes with the host-rich structure. This transformation is reversible provided the solvent loss does not exceed a threshold. The idea of a flexible host framework is a novelty. Usually, host–guest complexes are of two varieties: (1) the host forms a robust and invariant scaffold and guest molecules may be removed, added or exchanged at will or; (2) the assembly of the host is guest-induced and any attempt at removal of the guest results in collapse of the entire structure. Host–guest complexes formed by compound **5** belong to an interesting intermediate category. The C–H···O interactions that constitute the diamondoid host structure are strong enough with respect to exchange of guest molecules, but they are weak enough so that the host framework deforms reversibly from the diamondoid structure to the rhombohedral structure upon loss of solvent. Such behaviour is not characteristic of (typical) host–guest complexes wherein the host framework is built up with O–H···O and N–H···O bonds, which are too strong for host framework flexibility. It is the weakness of the C–H···O interaction that gives host–guest complexes of **5** their most characteristic property. Indeed one could say that crystal engineering based on weak interactions should target properties based on the weakness of these interactions and not on their strength (which is anyway marginal). Wuest has shown that in tetrahedral nitro aromatics based on pentaerythrityl tetraaryl ether cores, the open framework structure collapses because the molecules are significantly more flexible than **5**;[25] such a result is, to some extent, expected and illustrates what I have said above—one should not expect a C–H···O interaction to do exactly what an N–H···O or an O–H···O does. The C–H···O and similar hydrogen bridges are weak, flexible and hydrophobic. They come into their own when these attributes are desired—for example, in biomolecular processes wherein reversible but specific transformations are required.

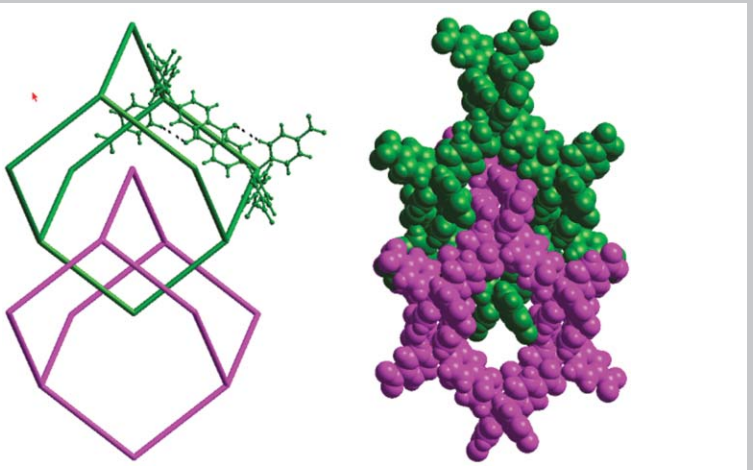

Fig. 4 Interpenetrated C–H···O diamondoid networks in the crystal structure of tetrakis(4-nitrophenyl)methane, **5**. Notice that each linkage in the tetrahedral superstructure is constituted with a pair of C–H···O hydrogen bonds.

Weak hydrogen bonds in biological molecules have been studied since the 1980s but it is only in recent years, with near atomic resolutions becoming a reality in macromolecular crystallography, that meaningful conclusions have been possible. With respect to the C–H···O bond, work by Derewenda on proteins,[26a] Sundaralingam on nucleic acids[26b] and Steiner on water[26c] is noteworthy. Every protein contains a very large number of C–H···O hydrogen bonds and for the larger proteins they occur in the thousands. There are three main configurations of weak C–H···O bonds in proteins: side chain to side chain, main chain to side chain and protein–ligand. Most of these interactions are weak to very weak and their functions are normally supportive at best. The most common of these interactions are C_α–H···O=C interactions in parallel and anti-parallel β-sheets. Other C–H···O=C contacts are found in α-helices, buried polar side chains and buried water molecules. An interesting residue is Pro which cannot donate N–H···O hydrogen bonds. If inserted in an α-helix, the regular pattern of N(*i*)–H···O=C(*i* − 4) hydrogen bonds is disrupted, leading to a kink in the helix. Chakrabarti has noted that in this situation, the activated proline $C_\delta H_2$ group is often involved in C–H···O interactions with carbonyl acceptors at positions (*i* − 3), (*i* − 4) or (*i* − 5) depending on the local conformation.[27] C–H···O hydrogen bonds from amino acid side chains are even weaker than hydrogen bonds formed by C_α–H groups. Other types of weak hydrogen bonds in proteins are formed to π-acceptors.

Examples are known with all strong donor types that are present in proteins: main chain and side chain N–H, side chain O–H, water molecules and O/N–H groups of substrate molecules. The acceptors are the side chains of Phe, Tyr, Trp and occasionally His residues. The energy range of O/N–H···π hydrogen bonds is about 2–4 kcal mol^{-1} for uncharged systems, in other words they are more significant than typical C–H···O interactions.[28]

The binding properties of proteins are the essence of functional genomics. It is necessary to know where a protein is localized and when it is expressed, but to find out what it does, one needs to find out to what it binds, and how. The specificity of biological processes suggests that the intermolecular interactions involved in the underlying recognition events are also specific, with conserved orientation. Hydrogen bonds, even the weakest ones, are electrostatic and therefore of long-range character; this is what makes them so important in the whole domain of biomolecular recognition. Hydrogen bonds, X–H···A, are instrumental not only in mediating drug–receptor binding, but they also affect physico-chemical properties, like solubility, partitioning, distribution, and permeability, that are crucial to drug development. The treatment of these interactions as hydrogen *bridges* recognizes their complex and composite nature.[20] The complexity arises because this is a many atom interaction: three (X, H and A) at the very least. The composite nature is manifested in its variable covalent, electrostatic and van der Waals character. Such variability follows from the nature of X and A, and the corresponding energy range is from $\sim$0.25 to 40 kcal mol^{-1}. The interaction is therefore chemically 'tunable' with the corresponding implications for function.

No less important in biological processes than specificity is reversibility. Weaker interactions can be made and broken more easily than stronger interactions. Accordingly, it is of interest to compare the significance of strong and weak interactions in the macromolecular recognition process. Is protein–ligand binding governed by conventional, that is electrostatic, N–H···O and O–H···O hydrogen bonds or do weaker interactions with a greater dispersive component like C–H···O also play a role? If so, to what extent are they significant? Noting that several recent studies have identified and validated the presence of C–H···O and other weak hydrogen bonds in macromolecular structures,[29] we undertook a database study of 28 selected high resolution protein–ligand crystal structures so that we could assess strong and weak hydrogen bonds simultaneously in a category of biological structures that is of importance in drug design.[30] We found that both strong (N–H···O, O–H···O) and weak (C–H···O) hydrogen bonds are involved in ligand binding and that multifurcation is common. Therefore, the restrictive geometrical criteria set up for hydrogen bonds in small molecule crystal structures may need to be relaxed in macromolecular structures. For example, there are definite deviations from linearity ($\theta \sim 180^{\circ}$) for both strong and weak hydrogen bonds. In contrast to small-molecule structures, anti-cooperative geometries are common in biomolecular structures. We found that C–H···O bonds formed by Gly, Phe, and Tyr are noteworthy and that the numbers of hydrogen bond donors and acceptors agree with the Lipinski rules that predict drug-like properties. Hydrogen bonds formed by water are also seen to be relevant in that ligand C–H···O_w interactions are abundant when compared to N–H···O_w and O–H···O_w. This suggests that ligands prefer to use their stronger hydrogen bond capabilities for use with the protein residues, leaving the weaker interactions to bind with water. In summary, the interplay between strong and weak interactions in ligand binding possibly leads to a satisfactory enthalpy/entropy balance.

The importance of C–H···O hydrogen bonds in protein–ligand binding has been demonstrated by Pierce *et al.* in a recent study of 200 liganded kinase structures.[31] The evidence is most convincing for activated C–H groups such as are found adjacent to heteroatoms in kinase ligands (heterocycles). While kinase ligands have been optimized for high-affinity binding using other criteria, the strong C–H···O hydrogen bonds that result are a serendipitous added value that is expected to be of considerable utility in protein modelling, ligand design, and structure–activity analysis. A question that arises immediately is, "What is the penalty in binding affinity for replacing a traditional protein–ligand hydrogen bond with an aromatic or heterocyclic C–H···O hydrogen bond?" This penalty appears to be surprisingly small and is rationalized on the basis that N–H and O–H groups must pay a larger desolvation price to leave the aqueous environment to form their hydrogen bonds with the protein. So, perhaps these two effects (hydrogen bond formation in the protein–ligand complex and desolvation) largely counterbalance one another, resulting in similar binding affinities for conventional hydrogen bonds and their C–H···O analogues.[32] Pierce *et al.* conclude that if N–H···O and C–H···O hydrogen bonds are interchangeable, the impact on ligand design would be tremendous, because N–H to C–H donor swaps would allow the design of novel inhibitors with similar binding affinity but potentially improved non-binding-related properties such as cell permeability or metabolic stability.

A final example of the interchangeable nature of these hydrogen bonds is provided in a recent study where we carried out virtual screening (VS) of 128 EGFR kinase inhibitors based on the 4-anilinoquinazoline fragment.[33] We chose this system because of the known importance of C–H···O hydrogen bonding in this case.[31] VS is a sequence of computational techniques that allows selection and ranking of possible leads from a library of compounds and is of significance in the current drug design scenario wherein high throughput screening is proving to be increasingly expensive and perhaps even unreliable.[34] VS is an evolving challenge and needs to be studied in depth and transformed into a tool of greater confidence and utility. Structure based VS consists of two parts, namely the accurate prediction of pose (docking) and the estimation of tightness of binding (scoring). Our goal was to arrive at the best combination of docking and scoring for the EGFR target and to develop a robust VS model. Since both docking and scoring operations would benefit from a better understanding of ligand–receptor recognition, we aimed to provide a chemical model that might be used to improve the overall efficiency of VS. This is why we chose this small group of 128 ligands; the aim of our

study was not to screen a very large library of ligands with a black box approach, but rather to evaluate current VS software and methodologies in the context of weak intermolecular interactions.

The docking of ligands for the VS was done in the active site as obtained in the experimental crystal structure of the erlotinib–EGFR complex.[35] Erlotinib is an anti-cancer drug from Genentech, belonging to the 4-anilinoquinazoline class. The 128 ligands were docked in the active site and the respective scores were obtained. The obtained poses, which represent positional and orientational information of the ligands, were classified into one of three categories: close, shifted and misoriented. We identified three key hydrogen bonds (N–H···N, O_w–H···N and C–H···O), of comparable stabilization energy, as responsible for anchoring the ligand in the active site (Fig. 5), and a ligand in the close category is docked with all three hydrogen bonds appearing correctly. A shifted ligand has one or more of the hydrogen bonds in place but the metrics are incorrect. A ligand with a misoriented pose is in a completely wrong orientation and/or position. While the N–H···N bond between Met769 and N(1) (*d*, 1.81 Å) and the O_w–H···N between water10 and N(3) (*d*, 2.01 Å) are of moderate strength, the C–H···O to Gln767 is very short (*d*, 2.19 Å) and involves a highly activated donor. Indeed, it is the best conserved interaction in the group. In the currently available docking software, the C–H···O bonds are not modelled explicitly; they fortuitously appear correctly

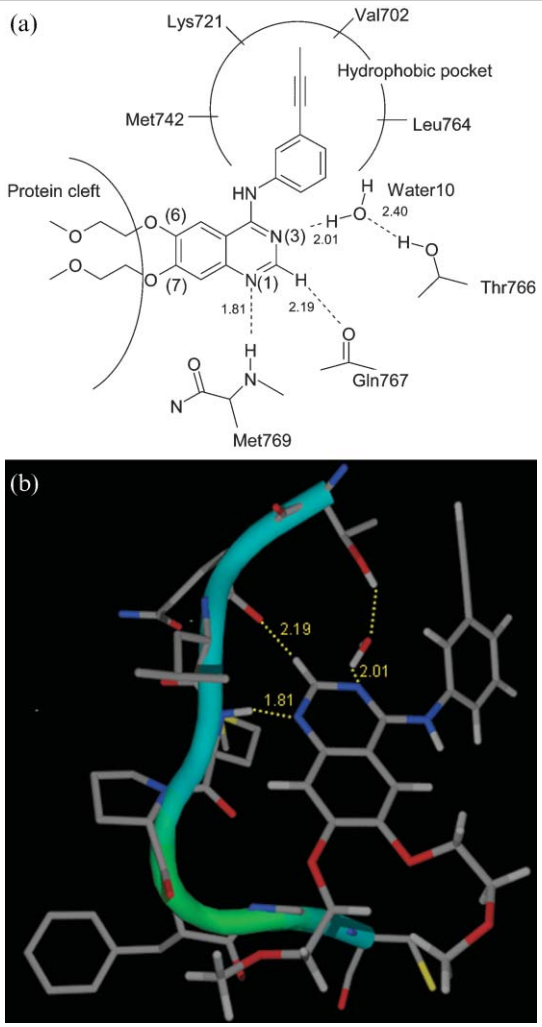

Fig. 5 Binding of erlotinib in the EGFR kinase active site. Notice the C–H···O bond formed by the activated donor. (a) Schematic diagram. (b) As in crystal structure.

for the close category ligands. The shifted and misoriented ligands could well be false negatives. We argue accordingly that if weak hydrogen bonds and other interactions *are* explicitly incorporated into the software, the efficiency of VS would increase greatly. VS is supposed to rapidly screen large chemical libraries and to 'cherry pick' and rank the few active ones, from the very large number of moderately active and inactive compounds—the so-called needle-in-the-haystack problem. Till today, VS approaches have concentrated on speed and automation. We suggest that future software should explicitly seek out hydrogen bond forming ability of a ligand, in other words address chemical issues directly so that structure based VS becomes increasingly accurate and reliable.

The C–H···O hydrogen bond was first invoked in the 1930s but it is only during the last 25 years or so that it and other weak interactions have been studied intensively and documented properly. Today, the question is not so much whether this interaction exists, or whether it is important in crystal packing as a structure determinant—these questions have long since been answered in the affirmative—but more about how it may be used and applied. In this regard, possibilities in the biological world appear to be very promising. Future work will show to what extent this promise is realised.

Acknowledgements

I thank Dr J. A. R. P. Sarma who, over the years, has been a co-author of several papers in this area.

Gautam R. Desiraju
School of Chemistry, University of Hyderabad, Hyderabad, 500046, India.
E-mail: gautam_desiraju@yahoo.com

Notes and references

1 J. A. R. P. Sarma and G. R. Desiraju, *J. Chem. Soc., Perkin Trans. 2*, 1987, 1195.
2 L. Leiserowitz, *Acta Crystallogr., Sect. B*, 1976, **32**, 775.
3 G. R. Desiraju and B. N. Murty, *Chem. Phys. Lett.*, 1987, **139**, 360.
4 G. R. Desiraju, *J. Chem. Soc., Chem. Commun.*, 1989, 179.
5 V. R. Pedireddi and G. R. Desiraju, *J. Chem. Soc., Chem. Commun.*, 1992, 988.
6 R. Taylor and O. Kennard, *J. Am. Chem. Soc.*, 1982, **104**, 5063.

7 G. R. Desiraju and T. Steiner, *The Weak Hydrogen Bond in Structural Chemistry and Biology*, Oxford University Press, Oxford, 1999.

8 (*a*) H. Bock, R. Dieneldt, H. Schödel and Z. Havlas, *J. Chem. Soc., Chem. Commun.*, 1993, 1792; (*b*) A. Cappelli, G. Giorgi, M. Anzini, S. Vomero, S. Ristori, C. Rossi and A. Donati, *Chem. Eur. J.*, 2004, **10**, 3177; (*c*) G. R. Desiraju, *J. Chem. Soc., Chem. Commun.*, 1990, 454.

9 (*a*) S. S. Kuduva, D. C. Craig, A. Nangia and G. R. Desiraju, *J. Am. Chem. Soc.*, 1999, **121**, 1936; (*b*) D. Das, R. K. R. Jetti, R. Boese and G. R. Desiraju, *Cryst. Growth Des.*, 2003, **3**, 675.

10 (*a*) P. G. Jones and F. Vancea, *CrystEngComm*, 2003, **5**, 303; (*b*) A. Anthony, M. Jaskólski, A. Nangia and G. R. Desiraju, *Chem. Commun.*, 1998, 2537.

11 V. R. Thalladi, H.-C. Weiss, D. Bläser, R. Boese, A. Nangia and G. R. Desiraju, *J. Am. Chem. Soc.*, 1998, **120**, 8702.

12 F. Hof and F. Diederich, *Chem. Commun.*, 2004, 484 and the references cited therein.

13 S. C. F. Kui, N. Zhu and M. C. W. Chan, *Angew. Chem., Int. Ed.*, 2003, **42**, 1628.

14 M. Nishio, *CrystEngComm*, 2004, **6**, 130.

15 (*a*) P. Hobza and Z. Havlas, *Chem. Rev.*, 2000, **100**, 4253; (*b*) A. J. Barnes, *J. Mol. Struct.*, 2004, **704**, 3.

16 D. Braga, *J. Chem. Soc., Dalton Trans.*, 2000, 3705 and the references cited therein.

17 L. Brammer, in *Crystal Design. Structure and Function*, (Perspectives in Supramolecular Chemistry, vol. 7), ed. G. R. Desiraju, Wiley, Chichester, 2003, pp. 1–75.

18 Y. Gu, T. Kar and S. Scheiner, *J. Am. Chem. Soc.*, 1999, **121**, 9411.

19 T. Steiner and G. R. Desiraju, *Chem. Commun.*, 1998, 891.

20 (*a*) G. R. Desiraju, *Acc. Chem. Res.*, 2002, **35**, 565; (*b*) G. R. Desiraju, *Curr. Sci.*, 2005, **88**, 374.

21 (*a*) G. R. Desiraju and C. V. K. M. Sharma, *J. Chem. Soc., Chem. Commun.*, 1991, 1239; (*b*) M. A. Viswamitra, R. Radhakrishnan, J. Bandekar and G. R. Desiraju, *J. Am. Chem. Soc.*, 1993, **115**, 4868; (*c*) P. K. Thallapally, A. K. Katz, H. L. Carrell and G. R. Desiraju, *Chem. Commun.*, 2002, 344.

22 P. K. Thallapally, G. R. Desiraju, M. Bagieu-Beucher, R. Masse, C. Bourgogne and J.-F. Nicoud, *Chem. Commun.*, 2002, 1052.

23 R. Thaimattam, X. Feng, J. A. R. P. Sarma, T. C. W. Mak and G. R. Desiraju, *J. Am. Chem. Soc.*, 2001, **123**, 4432.

24 C. V. K. Sharma and G. R. Desiraju, *J. Chem. Soc., Perkin Trans. 2*, 1994, 2345.

25 D. Laliberté, T. Maris and J. D. Wuest, *CrystEngComm*, 2005, **7**, 158.

26 (*a*) Z. S. Derewenda, L. Lee and U. Derewenda, *J. Mol. Biol.*, 1995, **252**, 248; (*b*) M. Wahl and M. Sundaralingam, *Trends Biochem. Sci.*, 1997, **22**, 97; (*c*) T. Steiner, *Acta Crystallogr., Sect. D*, 1995, **51**, 93.

27 R. Bhattacharyya and P. Chakrabarti, *J. Mol. Biol.*, 2003, **331**, 925.

28 T. Steiner and G. Koellner, *J. Mol. Biol.*, 2001, **305**, 535.

29 (*a*) B. P. Klaholz and D. Moras, *Structure*, 2002, **10**, 1197; (*b*) L. Jiang and L. Lai, *J. Biol. Chem.*, 2002, **277**, 37732; (*c*) G. F. Fabiola, S. Krishnaswamy, V. Nagarajan and V. Pattabhi, *Acta Crystallogr., Sect. D*, 1997, **53**, 316.

30 S. Sarkhel and G. R. Desiraju, *Proteins*, 2004, **54**, 247.

31 A. C. Pierce, K. L. Sandretto and G. W. Bemis, *Proteins*, 2002, **49**, 567.

32 J. C. Adrian and C. S. Wilcox, *J. Am. Chem. Soc.*, 1992, **114**, 1398.

33 V. Aparna, G. Rambabu, S. K. Panigrahi, J. A. R. P. Sarma and G. R. Desiraju, *J. Chem. Inf. Model.*, 2005, DOI: 10.1021/ci049676u.

34 (*a*) S. Grüneberg, M. T. Stubbs and G. Klebe, *J. Med. Chem.*, 2002, **45**, 3588; (*b*) P. D. Lyne, *Drug Discovery Today*, 2002, **7**, 1047; (*c*) K. H. Bleicher, H.-J. Böhm, K. Müller and A. I. Alanine, *Nat. Rev. Drug Discovery*, 2003, **2**, 369.

35 J. Stamos, M. X. Sliwkowski and C. Eigenbrot, *J. Biol. Chem.*, 2002, **48**, 46265.