Computational support for the suggested contribution of C—H···O=C interactions to the stability of nucleic acid base pairs

EVGENI B. STARIKOV† AND THOMAS STEINER* *at the Institut für Kristallographie, Freie Universität Berlin, Takustrasse 6, D-14195 Berlin, Germany. E-mail: steiner@chemie.fu-berlin.de*

Abstract

Ab inito quantum chemical calculations are performed to quantify the stabilizing role of long C—H···O=C contacts in nucleic acid base pairing, which was suggested by Leonard, McAuley-Hecht, Brown & Hunter [(1995). *Acta Cryst.* D51, 136–139]. For the Watson–Crick adenine–uracil pair, a contribution of about 6% to the total bond energy is obtained. This weakly bonding effect is primarily a result of electrostatic attraction between the total positive charge of adenine C(2)—H and the negative end of the dipole uracil O(2)=C.

1. Introduction

The hydrogen-bond patterns linking the Watson–Crick base pairs are among the most popularly known motifs of intermolecular interactions (see, *e.g.*, Jeffrey & Saenger, 1991). In conventional views, the guanine–cytosine (G–C) pair is linked by three hydrogen bonds, whereas the adenine–uracil pair (A–U) is linked by only two, Fig. 1. Recently, Hunter and coworkers provocatively suggested that the long contact C(2)—H···O(2)=C in the A–U pair might be of weakly hydrogen-bonding nature, thereby contributing to the stability of the base pair and completing a triple hydrogen-bond system similar to G–C (Leonard, McAuley-Hecht, Brown & Hunter, 1995). A related anticipated C—H···O hydrogen bond is the contact C(8)—H···O(2)=C in the Hoogsteen adenine–thymine (A–T) pair.

The idea itself that adenine forms C—H···O hydrogen bonds is not surprising: it is now well established that C—H···O interactions may be of hydrogen-bonding nature, and may play crucial roles in the architecture of organic and biological structures (see, *e.g.*, Desiraju, 1991, 1996; Derewenda, Lee & Derewenda, 1995; Steiner & Saenger, 1993; Steiner, 1995, 1996). Furthermore, purines and pyrimidines are well known to carry acidic C—H groups which are potent donors of C—H···O hydrogen bonds (recognised earlier by Sutor, 1963; Sundaralingam, 1966). In nucleic acid base pairing, a C—H···O interaction of close to ideal geometry (H···O = 2.3 Å, angle at H = 156°) was observed in a *trans* U–U pair in the crystal structure of the RNA fragment r(UUCGCG) by Wahl, Rao & Sundaralingam (1996). For non-Watson–Crick pairs involving modified bases, bonding contributions of C—H···O interactions were calculated using empirical potential functions (Ornstein & Fresco, 1983; Ornstein, 1988). The C(2)—H···O(2)=C contact in Watson–Crick A–U, however, has a geometry that is not immediately suggestive of hydrogen bonding: the H···O separation is long at ~2.7–2.8 Å, and the angle at H is

strongly bent, 131°; the more typical C—H···O bonds formed by purines and pyrimidines are close to linear and have H···O separations down to 2.1 Å. Hunter and coworkers stressed that the acceptor geometry is favourable with C—H pointing almost ideally in the C=O lone-pair direction. However, it has been shown that even for C—H···O=C bonds formed by the most acidic C—H donors, lone-pair directionality plays only a minor (though detectable) role (Steiner, Kanters & Kroon, 1996). This is a situation where it is difficult to judge
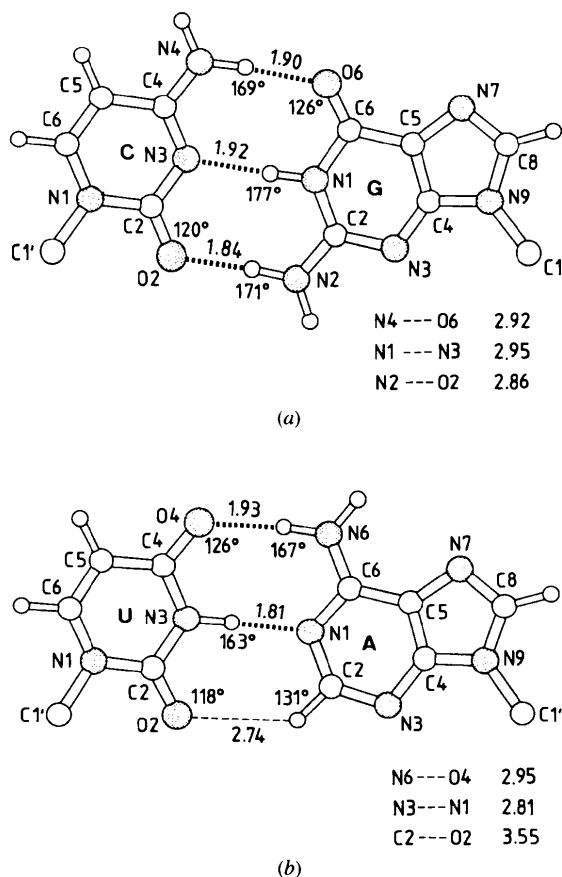

Fig. 1. Geometry of Watson–Crick nucleic acid base pairing as observed for (*a*) guanine–cytosine (G–C) in the X-ray crystal structure of sodium guanylyl-3′-5′-cytidine nonahydrate (Rosenberg, Seeman, Day & Rich, 1976), and (*b*) for adenine–uracil (A–U) in sodium adenylyl-3′-5′-uridine hexahydrate (Seeman *et al.*, 1976). Distances are given in Å, angles in °. H atoms are in theoretical positions with bond lengths of N—H = 1.03 and C—H = 1.09 Å.

---

† Present address: Fritz-Haber-Institut der Max-Planck-Gesellschaft, Abteilung Theorie, Faradayweg 4–6, D-14195 Berlin, Germany.

from only geometric parameters whether the C—H···O contact in A–U is actually bonding or not.

Before the suggestion of Hunter and coworkers can be accepted, the bonding nature of the contact under discussion must be verified by theoretical and, possibly, also by experimental work. This is a non-trivial task which is approached here by *ab initio* quantum chemical calculations on this system. The aim of the study is not to dwell on the question whether the contact deserves the classification as a 'hydrogen bond', but to clarify whether it is stabilizing or destabilizing the base pair, or whether it is rather null and nothing.

2. Computation

The study is based on the Watson–Crick A–U base-pairing geometry as experimentally observed in the crystal structure of sodium adenylyl-3'-5'-uridine (Seeman, Rosenberg, Suddath, Kim & Rich, 1976). Since no experimental H-atom positions are available, the H atoms have to be calculated to the theoretical positions. In this case, the H-atom positions are determined well by the non-H skeleton, so that this procedure is fairly reliable and accurate. The idealized bond lengths used are C—H = 1.09 and N—H = 1.03 Å.

In hydrogen-bonded systems like those shown in Fig. 1, the contribution of the individual constituents cannot be calculated directly. It is only possible to calculate total bond energies between molecules, *i.e.* the joint contribution of all intermolecular interactions. A common method used to study the contribution of an individual constituent is to calculate the total intermolecular bond energies for series of polyatomic fragments which do and do not contain the particular interaction under study. Comparison of the fragments yields an estimate for the contribution of the individual interaction being searched for. Computational fragmentation should stay as closely as possible to the realistic covalent configuration.
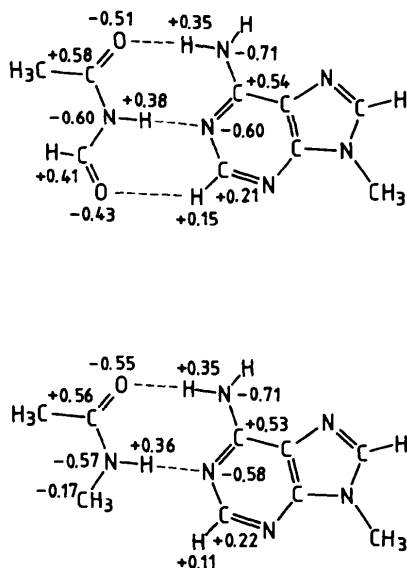

3. Results

The structural approximations used for the Watson–Crick A–U pair with and without the contacts C(2)—H···O(2)=C, and the calculated Mulliken point charges are shown in Fig. 2. The bond energy obtained for the top fragment (with C—H···O=C contact) is $12.4\,\mathrm{kcal\,mol^{-1}}$ ($51.9\,\mathrm{kJ\,mol^{-1}}$); this is close to the experimental gas phase value of $13.0\,\mathrm{kcal\,mol^{-1}}$ ($54.4\,\mathrm{kJ\,mol^{-1}}$) (Yanson, Teplisky & Sukhodub, 1979; Sukhodub, 1987), indicating that the approximation is realistic. After deletion of the O(2) atom (Fig. 2b), the bond energy is reduced to $11.7\,\mathrm{kcal\,mol^{-1}}$ ($49.0\,\mathrm{kJ\,mol^{-1}}$). The reduction of $0.7\,\mathrm{kcal\,mol^{-1}}$ ($2.9\,\mathrm{kJ\,mol^{-1}}$) amounts to about 6% of the total energy. This energy is small, not much larger than dispersion energies, but corresponds to what one must expect for weaker and distorted types of C—H···O hydrogen bonds (*e.g.* Steiner, 1996).†

An inspection of the distribution of partial charges in Fig. 2 shows that the C—H···O=C contact has essentially different electrostatic properties from the conventional N—H···O/N hydrogen bonds of the base pair. The latter have the typical charge distribution $N^{\delta-}$—$H^{\delta+}\cdots O/N^{\delta-}$, whereas in the discussed C—H···O contact, H and C carry positive charges, that on C may even be slightly larger. Therefore, the dipole moment of C—H is very small, whereas the total charge on the C—H group is relatively large. From the electrostatic point of view, this arrangement represents the situation where a partially charged moiety $(CH)^{\delta+}$ contacts the negative end of a dipole $O^{\delta-}=C^{\delta+}$. This clearly is a bonding situation with a positive effect on base-pair stability, but one can dispute whether it can be called a proper 'hydrogen bond'.

4. Conclusions

The relevant result is that the C(2)—H···O(2)=C contact enhances the total bond energy of the Watson–Crick A–U base pair, *i.e.* it has a stabilizing function. Alternative views that the C(2)—H···O(2)=C contact is a 'forced' one and destabilizes the base pair, are disfavoured. Numerically, we

---

† A referee noted that the difference in total bond energies might be not due to a favourable C—H···O=C interaction in the top fragment, but due to a C—H···Me repulsion in the bottom fragment. This concern is not unjustified in principle, but we note that the distance of C(2)A—H to the Me group is even larger than that to C=O [C(2)A···O(2)U = 3.55 Å; C(2)A···C(2)U = 4.04 Å]. The interaction of a partially charged moiety $(CH)^{+}$ with a more or less apolar group at long distance will not exceed that with a strong C=O dipole at medium separation.

After selection of suitable structural models, quantum chemical calculations were performed in the *ab initio*, MO LCAO SCF (HF + MP2) approximation using the GAUSSIAN92/DFT package (Frisch *et al.*, 1993). Atomic partial charges and the total intermolecular bond energies were calculated using the 6-31G** basis set by taking into account electron correlation and the basis set superposition error (BSSE). In this algorithm, partial atomic charges are calculated according to the Mulliken scheme. No theoretical geometry optimization was performed (hardware: CRAY Y-MP4D/464 of the Konrad-Zuse-Zentrum, Berlin).

Fig. 2. Molecular fragments used for approximating the Watson–Crick A–U base pair with and without the C(2)—H···O(2)=C interaction. Relevant calculated atomic partial charges are given.

estimate an ~6% contribution to the total inter-pair bond energy. This is not a large effect, and much less than C—H···O interactions can do in principle, but it operates supporting the conventional hydrogen bonds of the base pair. Topologically, the interaction completes a cooperative triple hydrogen-bond system related to that in the Watson–Crick G-C pair. In essence, this backs up the suggestions of Hunter and coworkers.

By its nature, the discussed stabilizing effect is primarily a result of electrostatic attraction between the total positive charge on C(2)—H and the negative end of the O(2)=C dipole. This is not the classical hydrogen-bonding situation; we see no problem, however, to consider this interaction *cum grano salis* as a weak 'hydrogen bond'.

References

Derewenda, Z. S., Lee, L. & Derewenda, U. (1995). *J. Mol. Biol.* **252**, 248–262.

Desiraju, G. R. (1991). *Acc. Chem. Res.* **24**, 290–296.

Desiraju, G. R. (1996). *Acc. Chem. Res.* **29**, 441–448.

Frisch, M. J., Trucks, G. W., Schlegel, H. B., Gill, P. M. W., Johnson, B. G., Wong, M. W., Foresman, J. B., Robb, M. A., Head-Gordon, M., Replogle, E. S., Gompertz, R., Andres, J. L., Raghavachari, K., Birkley, J. S., Gonzalez, C., Martin, R. L., Fox, D. J., Defrees, D. J., Stewart, J. J. P. & Pople, J. A. (1993). *GAUSSIAN92/DFT*, Revision G.3, Gaussian Inc., Pittsburgh, PA, USA.

Jeffrey, G. A. & Saenger, W. (1991). *Hydrogen Bonding in Biological Structures*. Springer, Berlin.

Leonard, G. A., McAuley-Hecht, K., Brown, T. & Hunter, W. N. (1995). *Acta Cryst.* D**51**, 136–139.

Ornstein, R. L. (1988). *J. Mol. Struct. (Theochem)*, **179**, 311–317.

Ornstein, R. L. & Fresco, J. R. (1983). *Proc. Natl Acad. Sci. USA*, **80**, 5171–5175.

Rosenberg, J. M., Seeman, N. C., Day, R. O. & Rich, A. (1976). *J. Mol. Biol.* 145–167.

Seeman, N. C., Rosenberg, J. M., Suddath, F. L., Kim, J. J. P. & Rich, A. (1976). *J. Mol. Biol.* **104**, 109–144.

Steiner, Th. (1995). *Acta Cryst.* D**51**, 93–97.

Steiner, Th. (1996). *Cryst. Rev.* **6**, 1–57.

Steiner, Th., Kanters, J. A. & Kroon, J. (1996). *Chem. Commun.* pp. 1277–1278.

Steiner, Th. & Saenger, W. (1993). *J. Am. Chem. Soc.* **115**, 4540–4547.

Sundaralingam, M. (1966). *Acta Cryst.* **21**, 495–506.

Sukhodub, L. F. (1987). *Chem. Rev.* **87**, 589–606.

Sutor, D. J. (1963). *J. Chem. Soc.* pp. 1105–1110.

Wahl, M. C., Rao, S. T. & Sundaralingam, M. (1996). *Nature Struct. Biol.* **3**, 24–31.

Yanson, I. K., Teplisky, A. B. & Sukhodub, L. F. (1979). *Biopolymers*, **18**, 1149–1170.