J|A|C|S
ARTICLES

The Elusive Atomic Rationale for DNA Base Pair Stability

Paul L. A. Popelier* and Laurent Joubert†

*Contribution from the Department of Chemistry, U.M.I.S.T., Manchester,
M60 1QD, United Kingdom*

Received November 8, 2001. Revised Manuscript Received February 19, 2002

Abstract: A systematic analysis of the electrostatic interaction between 27 natural DNA base pairs was carried out, based on ab initio correlated wave functions and the topology of the electron density. Using high rank multipole moments we show that the atomic partitioning of the interaction energy contains many substantial contributions between distant atoms. Profiles of cumulative energy versus internuclear distance show large fluctuations and provide an electrostatic fingerprint of the partitioning of interaction energy in a complex. A quantified comparison between each pair of energy profiles, one for each base pair, makes clear that there is no correlation between the total base pair interaction energy and the shape of the profile. In other words, base pairs with similar interaction energy are not stable for the same reasons in terms of atomic partitioning. In summary, simple rules to rationalize the pattern of energetic stability of naturally occurring base pairs in terms of subsets of atoms are elusive. Our work cautions against inappropriate use of Jorgensen's secondary interaction hypothesis.

Introduction

In his 1988 lecture[1] on supramolecular chemistry Lehn recognized intermolecular interactions as the basis of highly specific recognition, reaction, transport, and regulation processes in biology. He asserted that the design of artificial abiotic receptor molecules required the correct manipulation of the energetic features of noncovalent intermolecular interactions. Among those interactions[2] hydrogen bonding fulfils an important role in guiding the design of self-assembling nanostructures. It is often invoked[3] to rationalize molecular recognition or the association of biological molecules, areas of increasing importance to organic chemistry.[4] Hydrogen bonding also enables useful structural predictions in the growing field of abiotic oligomers[5] and foldamers.[6]

In this paper we seek an improved understanding of the electrostatic contribution to intermolecular interaction in base pairs formed by the natural nucleic acid bases guanine (G), cytosine (C), adenine (A), and thymine (T). An investigation of the natural base pairs is valuable in a wider context because the use of heteroatomic modules as recognition units for self-assembly is inspired by Nature's employment of nucleic acid bases as carriers of genetic information.[7]

For many decades it was thought that triply hydrogen bonded base pairs are invariably more stable than doubly hydrogen bonded ones. However, simply counting hydrogen bonds in a complex or folding pattern does not necessarily provide insight into stability.[8,9] This failure led to the secondary interaction hypothesis (SIH),[10–12] which involves short-range "cross-interactions" between the frontier atoms of each base pair as explained in more detail below. Further work[13] led to a general scheme based on empirical increments for the prediction of hydrogen-bonded associations of nucleobases.

In this paper we use advanced quantum chemical methods to prove that the SIH is not compatible with a complete, rigorous, and detailed analysis of the relative energetic stability of 27 natural base pair complexes in vacuo. After a section on the computational method used we systematically discuss the series of calculations we performed in our attempt to find support for the SIH. However, from our compelling evidence we were forced to conclude that the SIH has no general validity grounded in quantum chemistry and should hence be used with great caution, if at all.

Computational Details

The full details of the representation of base pairs used here are given in a previous paper[14] that focused on the development

---

* Corresponding author. E-mail: pla@umist.ac.uk.

† Current address: Laboratoire d'Electrochimie, ENSCP, 11 rue Pierre et Marie Curie, 75231 Paris Cedex 05, France.

(1) Lehn, J.-M. *Angew. Chem., Int. Ed. Engl.* **1988**, *27*, 89–112.
(2) Whitesides, G. M.; Mathias, J. P.; Seto, C. T. *Science* **1991**, *254*, 1312–1319.
(3) Prins, L. J.; Reinhoudt, D. N.; Timmerman, P. *Angew. Chem., Int. Ed. Engl.* **2001**, *40*, 2382–2426.
(4) Mammen, M.; Simanek, E. E.; Whitesides, G. M. *J. Am. Chem. Soc.* **1996**, *118*, 12614–12623.
(5) Archer, E. A.; Gong, H.; Krische, M. J. *Tetrahedron* **2001**, *57*, 1139–1159.
(6) Hill, D. J.; Mio, M. J.; Prince, R. B.; Hughes, T. S.; Moore, J. S. *Chem. Rev.* **2001**, *101*, 3893–4011.
(7) Zimmerman, S. C.; Corbin, P. S. In *Structure and Bonding*; Fujita, M., Ed.; Springer: New York, 2000; Vol. 96, p 63.
(8) Gardner, R. R.; Gellman, S. H. *J. Am. Chem. Soc.* **1995**, *117*, 10411–10412.
(9) Yang, J.; Gellman, S. H. *J. Am. Chem. Soc.* **1998**, *120*, 9090–9091.
(10) Jorgensen, W. L.; Pranata, J. *J. Am. Chem. Soc.* **1990**, *112*, 2008–2010.
(11) Jorgensen, W. L.; Severance, D. L. *J. Am. Chem. Soc.* **1991**, *113*, 209–216.
(12) Pranata, J.; Wierschke, S. G.; Jorgensen, W. L. *J. Am. Chem. Soc.* **1991**, *113*, 2810–2819.
(13) Sartorius, J.; Schneider, H.-J. *Chem. Eur. J.* **1996**, *2*, 1446–1452.
(14) Joubert, L.; Popelier, P. L. A. Submitted for publication, 2002.

10.1021/ja0125164 CCC: $22.00 © 2002 American Chemical Society

of a topological intermolecular potential tested on DNA base pairing. Here we mention a few salient features and refer for the mathematical formulas to our earlier work testing the topological potential on van der Waals molecules.[15] The geometries of the (monomeric) nucleic acid bases A, T, C, and G were optimized at the B3LYP/6-311+G(2d,p)[16,17] level by using GAUSSIAN98.[18] Topological atoms are defined according to the theory of "Atoms in Molecules" (AIM)[19,20] as regions in real space consisting of a bundle of electron density gradient paths attracted to a nucleus. Given its deep roots in quantum mechanics AIM can be regarded as a prime partitioning method that is currently applied by over a hundred laboratories.[21,22] A local version of the program MORPHY01[23] computed the AIM multipole moments[24] according to the spherical tensor formalism.[25] The topological atoms were capped by the $\rho = 10^{-7}$ au isodensity envelope. The integration errors measured via $L(\Omega)$[26] were all below 1×10^{-4} au for nonring atoms and about 1×10^{-3} au for ring atoms. The program ORIENT3.2j[27] calculated the intermolecular electrostatic energy including all possible multipole−multipole interactions up to R^{-6}, where R is the distance between two multipole origins or sites. Our anisotropic model clearly extends far beyond the point charge model that provided the context in which the SIH was formulated. The availability of analytical forces and Hessian[28,29] combined with a powerful search method[30] enabled ORIENT3.2j to generate the base pair complexes quickly and reliably. The repulsion was modeled via pseudohard spheres (i.e. slightly softened) by using the Pauling radii,[31] being 2.0 Å for C, 1.4 Å for O, and 1.5 Å for N, while H was not assigned a radius. The topological intermolecular potential yields small discrepancies in geometry compared to supermolecular calculations.[14] Hence we confirm the dominance of the electrostatic contribution to the total interaction energy.[32] It was noted before[33] that DFT interaction energies are trustworthy (in relation to MP2) except for stacked base pairs. Since the latter are not included in this study we are confident that the B3LYP method is adequate, which is further confirmed by our previous study.[14]

There is no official nomenclature for the naming of base pair complexes. Here we adopt the convention introduced by Hobza and Sandorfy,[34] which differentiates the Watson−Crick (WC), reverse WC (RWC), Hoogsteen (H), and reverse H (RH) configurations. An auxiliary number distinguishes some base pairs further according to the order of decreasing base pair stability obtained by the level of calculation of ref 34. For example, there are four guanine−guanine complexes denoted GG1, GG2, GG3, and GG4. However, GG2 is not a minimum at our level of calculation and hence this label does not appear in our work. For space reasons we cannot reproduce a diagram of all 27 base pairs we investigated but refer to diagrams published in earlier work.[33,34]

Finally it should be pointed out that the term topology strictly refers to quantum (chemical) topology and should not be confused with its use in the purely supramolecular context of catenanes and knots.[35]

Results and Discussion

On the basis of accumulating evidence Jorgensen and Pranata[10] pointed out in 1990 that the energetic stability of a complex does not correlate with the number of hydrogen bonds. To solve the quandary they proposed the so-called secondary interaction hypothesis. This ad hoc explanation involves hydrogen bond donors (D) and acceptors (A) held near one another in the complex, but *not* hydrogens bonded to each other.[36] Figure 1 illustrates this rule showing the two original complexes Jorgensen and Pranata used to propose their hypothesis. The GC complex has four secondary interactions between G and C, two of which are attractive (D•A) and two repulsive (A•A and D•D). On the other hand the uracil-2,6-diaminopyridine (U-DAP) complex has four repulsive interactions and is hence expected to be less stable than GC although both complexes exhibit three hydrogen bonds (referred to as primary interactions). Although Jorgensen and Pranata used their hypothesis to explain the difference in K_a values of GC and U-DAP in chloroform, the SIH was triggered by an analysis of the isolated complexes. In their paper no precise definition of a secondary interaction is given other than stating that the atomic separations are still short (2.3−3.7 Å or 4.3−7.0 au) and explicitly illustrating the hypothesis on only two cases: GC versus U-DAP and the interaction between two peptides versus their retro-inverso isomer. Two years later the SIH was exploited to explain the great stability of DDD•AAA complexes[37] and was invoked in the context of nucleic acid base related aromatic heterocycles[37−48] and peptides and peptidomimetics.[8,9,49] However, it is also known to fail or to need further modification.[9,36−38,50]

---

(15) Popelier, P. L. A.; Joubert, L.; Kosov, D. S. *J. Phys. Chem. A* **2001**, *105*, 8254−8261.

(16) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648−5652.

(17) Foresman, J. B.; Frisch, A. *Exploring Chemistry with Electronic Structure Methods*, 2nd ed.; Gaussian Inc.: Pittsburgh, PA, 1996.

(18) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A., Jr.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Gonzalez, C.; Challacombe, M.; Gill, P. M. W.; Johnson, B. G.; Chen, W.; Wong, M. W.; Andres, J. L.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. *Gaussian 98*, revision A.7; Gaussian, Inc.: Pittsburgh, PA, 1998.

(19) Bader, R. F. W. *Atoms in Molecules. A Quantum Theory*; Oxford University Press: Oxford, GB, 1990.

(20) Popelier, P. L. A. *Atoms in Molecules. An Introduction*; Pearson Education: London, GB, 2000.

(21) Popelier, P. L. A.; Aicken, F. M.; O'Brien, S. E. In *Chemical Modelling: Applications and Theory*; Vol. 1, Royal Society of Chemistry Specialist, Periodical Report; Hinchliffe, A., Ed.; Royal Society of Chemistry: London, UK, 2000; Chapter 3, pp 143−198.

(22) Popelier, P. L. A.; Smith, P. J. In *Chemical Modelling: Applications and Theory*; Vol. 2, Royal Society of Chemistry Specialist, Periodical Report, Hinchliffe, A., Ed.; Royal Society of Chemistry: London, UK, 2002.

(23) *MORPHY01*, a program written by P. L. A. Popelier with a contribution from R. G. A. Bone, D. S. Kosov, and M. in het Panhuis, UMIST: Manchester, England, EU, 2000.

(24) Popelier, P. L. A. *Mol. Phys.* **1996**, *87*, 1169−1187.

(25) Stone, A. J. *The Theory of Intermolecular Forces*; Clarendon: Oxford, GB, 1996.

(26) Aicken, F. M.; Popelier, P. L. A. *Can. J. Chem.* **2000**, *78*, 415−426.

(27) Stone, A. J.; Dullweber, A.; Hodges, M. P.; Popelier, P. L. A.; Wales, D. J. *ORIENT3.2j*; Cambridge, 1996.

(28) Popelier, P. L. A.; Stone, A. J. *Mol. Phys.* **1994**, *82*, 411−425.

(29) Popelier, P. L. A.; Stone, A. J.; Wales, D. J. *Faraday Discuss.* **1994**, *97*, 243−264.

(30) Baker, J. *J. Comput.Chem.* **1986**, *7*, 385−395.

(31) Pauling, L. *The Nature of the Chemical Bond*, 3rd ed.; Cornell University Press: Ithaca, NY, 1960.

(32) Uchimaru, T.; Korchowiec, J.; Tsuzuki, S.; Matsumura, K.; Kawahara, S.-I. *Chem. Phys. Lett.* **2000**, *318*, 203−209.

(33) Sponer, J.; Leszczynski, J.; Hobza, P. *J. Phys. Chem.* **1996**, *100*, 1965−1974.

(34) Hobza, P.; Sandorfy, C. *J. Am. Chem. Soc.* **1987**, *109*, 1302−1307.

(35) Seeman, N. C. *Acc. Chem. Res.* **1997**, *30*, 357−363.

(36) Gardner, R. R.; Gellman, S. H. *Tetrahedron* **1997**, *53*, 9881−9890.

(37) Murray, T. J.; Zimmerman, S. C. *J. Am. Chem. Soc.* **1992**, *114*, 4010−4011.

(38) Jeong, K. S.; Tjivikua, T.; Muehldorf, A.; Deslongchamps, G.; Famulok, M.; Rebek, J., Jr. *J. Am. Chem. Soc.* **1991**, *113*, 201−209.

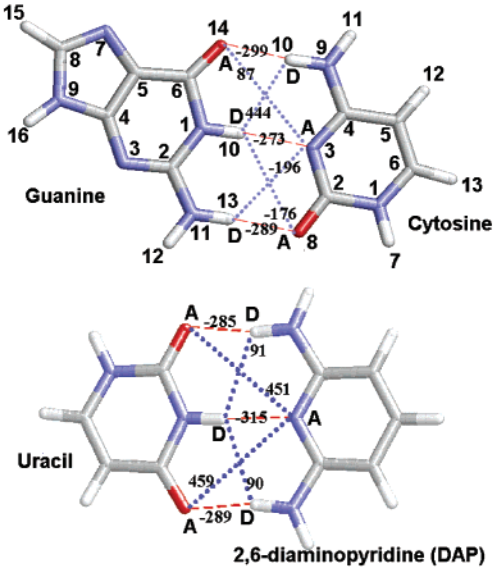

Figure 1. The GC-WC base pair complex (with labeling scheme referring to Figure 2) and the U-DAP complex. The three primary interactions (hydrogen bonds) are marked by a dashed line (red) and the four secondary interactions by a dotted line (blue). Donor (D) and acceptor (A) atoms are indicated. The electrostatic interaction energies for all primary and secondary interactions (kJ/mol) are marked near the relevant interaction line.

It can be verified from Figure 1 that according to our AIM multipole model (i.e. topological potential) the sum of the primary hydrogen bond interactions is −861 kJ/mol for GC-WC and −889 kJ/mol for U-DAP. As a result the interaction energy due to the three hydrogen bonds in GC-WC and U-DAP wrongly favors U-DAP by 28 kJ/mol. However, the difference in the secondary interaction energies is much more pronounced, namely +159 kJ/mol for GC-WC and +1091 kJ/mol for U-DAP, confirming that the observed difference in stability cannot be ascribed to the primary interactions (three hydrogen bonds). Our numerical results support the SIH in this case since GC-WC is experimentally found to be more stable than U-DAP. However, we are very concerned about drawing conclusions from an arbitrary subset of interactions involving just the frontier atoms. This is why we base our systematic and more complete investigation on *all* possible atom−atom interactions.

Figure 2 shows all atom−atom contributions to the total electrostatic interaction energy between G and C in the GC-WC configuration. The pattern of large and spread out contributions is typical for all 27 base pairs. The average of the absolute values of the atom−atom interactions is 141 kJ/mol, which is of the same order of magnitude as the total electrostatic interaction energy itself (−113 kJ/mol). The range of values is particularly wide as the standard deviation is 154 kJ/mol. This figure makes it clear that the four secondary interactions (G/H10-C/H10, G/H10-C/O8, G/O14-C/N3, and G/H13-C/N3) constitute an inconspicuous subset of interactions. Explaining the total energetic stability of a base pair by these four interactions only poses the inherent danger of over-interpretation. Our work is to scrutinize exactly the existence of remarkable subsets that would explain total stability.

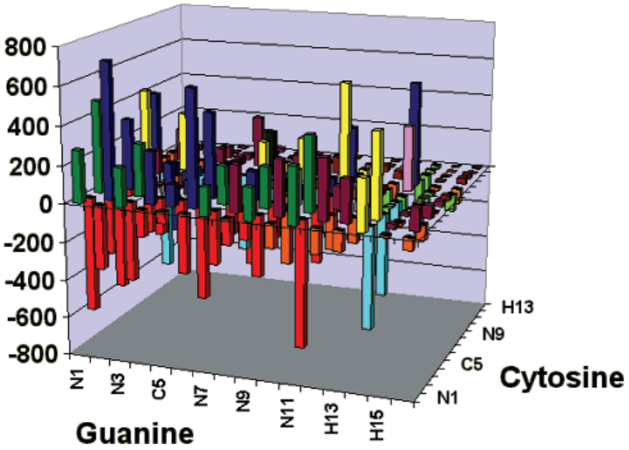

Figure 2. Histogram of the atom−atom contributions to the total electrostatic interaction energy (kJ/mol) marked according to atom type. Color code: CC (dark blue), NN (dark green), OO (pink), CN (red), ON (yellow), CO (light blue), OH (light green), NH (orange), CH (purple), and HH (brown). Many interactions contribute significantly to the total energy. The numerical labels correspond to those in Figure 1.

The information contained in Figure 2 can be compressed in a one-dimensional interaction energy profile via a successive addition of atom−atom interaction energies versus the internuclear distance. Such cumulative energy profiles are shown in Figure 3a for GC-WC and U-DAP and in Figure 3b for GG1, GG3, and GG4. Again it is clear that the atom−atom distance corresponding to the primary and secondary interactions is but a small part of the profiles and ignores substantial fluctuations beyond its range. The electrostatic "fingerprint" of each complex reveals alternating regions dominated by either attractive or repulsive interactions. Such global oscillations are observed for all base pairs. For example, the energy in GC-WC can reach values of more than 16 times (e.g. −1872 kJ/mol at 7.98 au) that of the total energy (−113 kJ/mol), which is reached at the utmost right of the profile. For all complexes primary and secondary interactions are located in a range of 3.6−7.0 au. However, from Figure 3a,b we learn that substantial energy fluctuations occur *beyond* 7 au. This observation is at variance with the widespread idea of the predominance of primary (hydrogen bonds) and secondary interactions in biological systems. Rather this picture is closer to the strong long-range electrostatic interactions observed in condensed matter. This conclusion is not unique to our current rigorous atomic partitioning scheme since alternative methods such as the anisotropic Distributed Multipole Analysis model[51] and the Merz−Kollman[52] and the Natural Population Analysis[53] point charge models reveal similar patterns, albeit less pronounced.

---

(39) Zimmerman, S. C.; Murray, T. J. *Tetrahedron Lett.* **1994**, *35*, 4077−4080.
(40) Beijer, F. H.; Sijbesma, R. P.; Vekemans, J. A. J. M.; Meijer, E. W.; Kooijman, H.; Spek, A. L. *J. Org. Chem.* **1996**, *61*, 6371−6380.
(41) Roberts, C.; Bandaru, R.; Switzer, C. *J. Am. Chem. Soc.* **1997**, *119*, 4640−4649.
(42) Beijer, F. H.; Sijbesma, R. P.; Kooijman, H.; Spek, A. L.; Meijer, E. W. *J. Am. Chem. Soc.* **1998**, *120*, 6761−6769.
(43) Kolotuchin, S. V.; Zimmerman, S. C. *J. Am. Chem. Soc.* **1998**, *120*, 9092−9093.
(44) Navarro, J. A. R.; Freisinger, E.; Lippert, B. *Inorg. Chem.* **2000**, *39*, 1059−1065.
(45) Söntjens, S. H. M.; Sijbesma, R. P.; van Genderen, M. H. P.; Meijer, E. W. *J. Am. Chem. Soc.* **2000**, *122*, 7487−7493.
(46) Lan, T.; McLaughlin, L. W. *J. Am. Chem. Soc.* **2001**, *123*, 2064−2065.
(47) Collinson, S. R.; Gelbrich, T.; Hursthouse, M. B.; Tucker, J. H. R. *Chem. Commun.* **2001**, 555−556.
(48) Hartel, C.; Goebel, M. W. *Helv. Chim. Acta* **2000**, *83*, 2541.
(49) Yang, J.; Christianson, L. A.; Gellman, S. H. *Org. Lett.* **1999**, *1*, 11−13.
(50) Zeng, H.; Miller, R. S.; Flowers, R. A., II; Gong, B. *J. Am. Chem. Soc.* **2000**, *122*, 2635−2644.

(51) Stone, A. J. *Chem. Phys. Lett.* **1981**, *83*, 233−239.
(52) Singh, U. C.; Kollman, P. A. *J. Comput.Chem.* **1984**, *5*, 129−145.
(53) Reed, A. E.; Weinstock, R. B.; Weinhold, F. *J. Chem. Phys.* **1985**, *83*, 735−746.

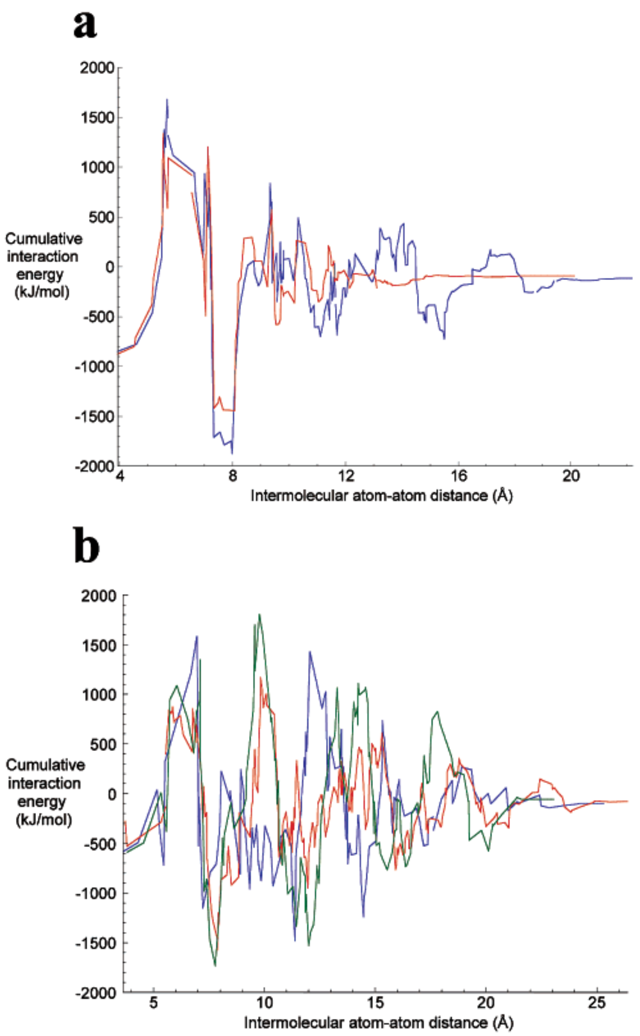

Figure 3. Cumulative electrostatic interaction energy (kJ/mol) versus the intermolecular atom−atom distance between two bases (au). Consider this distance as a parameter that is gradually increased, such that the energy of a new atom−atom pair is added to the cumulative energy whenever the corresponding interatomic distance of the pair is hit. (a) Energy profile of GC-WC (blue) and U-DA P(red). (b) Energy profiles for GG1 (blue), GG3 (red) and GG4 (green).

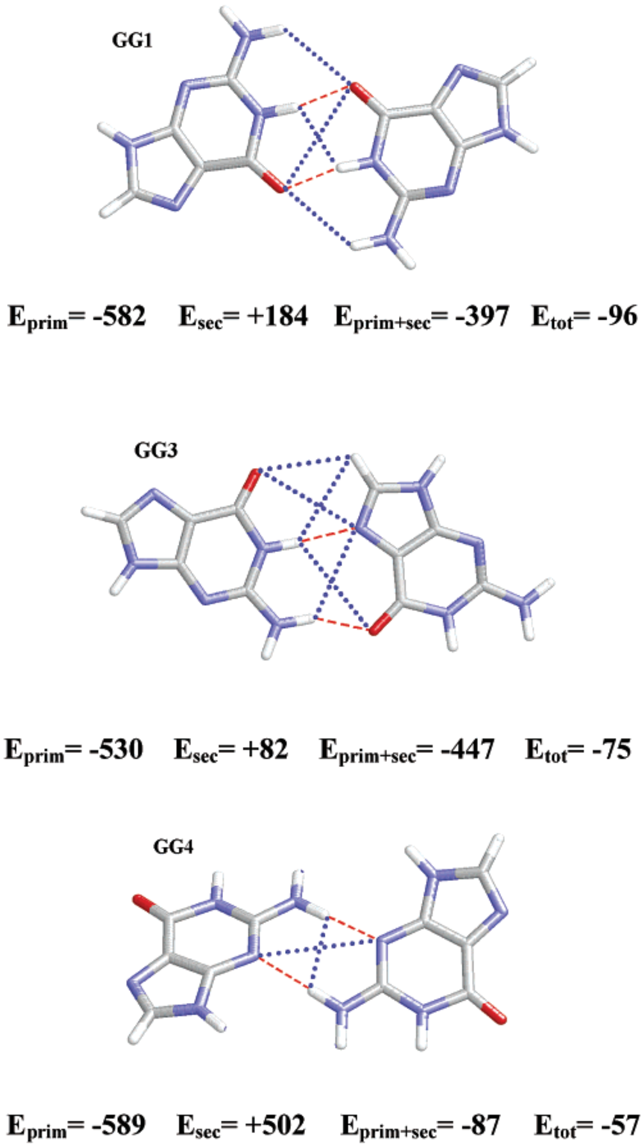

Figure 4. The GG1, GG3, and GG4 base pair complexes. Primary interactions (hydrogen bonds) are marked by a dashed line (red) and secondary interactions by a dotted line (blue). An energy (kJ/mol) breakdown in terms of primary and secondary interactions and total electrostratic energy is added.

The geometries and energy breakdown (in terms of primary and secondary interactions) of GG1, GG3, and GG4 are given in Figure 4. The total electrostatic energy ranks the base pairs (starting with the most negative energy and hence the most stable) as GG1 < GG3 < GG4, consistent with the nomenclature we use in this article. Due to the rather vague definition[10] of a secondary interaction it is difficult to make an unambiguous assignment. To the best of our knowledge we have assigned two hydrogen bond interactions in every GG complex, and three "outer" secondary interactions (one in GG3, two in GG1) reminiscent of Zimmerman's "over-hanging" secondary interactions.[39] According to the SIH the ranking is GG3 < GG1 < GG4, which is consistent with the sum of the energy of the primary and secondary interactions $E_{\mathrm{prim+sec}}$. However, the SIH only predicts correctly that GG4 is the least stable complex.

To obtain further insight from an energy profile we disentangle it according to atom types as shown in Figure 5. It is convenient to return to GC-WC since Figure 2 shows the different atom type subclasses such as ON, NN, CC, etc. The interaction energy grouped according to atom type in Figure 5 is characteristic for all base pairs. Each curve represents the cumulative energy of an interaction between two atoms with a given atomic number. As the internuclear distance increases each such curve either monotonically increases (for N−N, C−H, O−N, C−C, and H−H) or monotonically decreases (for C−N, N−H, C−O, and O−H). This feature could not be determined for O−O since there is only one such interaction in GC-WC. We see that many atom−atom interactions contribute substantially at high distance, again weakening support for the secondary interaction hypothesis. The three primary and four secondary interactions, marked in Figure 5, clearly constitute a rather arbitrary subset.

In another attempt to find a rationale for the pattern of energetic stability of *all* base pairs we investigate the correlation between energy profiles. Do certain features of the shape of the energy profiles determine the stability? The Pearson correlation coefficient was calculated for each of the 351 = 27 × (27 − 1)/2 possible pairs of profiles to be compared. The more similar the energy profiles of two base pairs, the closer this coefficient approaches unity. In Figure 6 the value of the

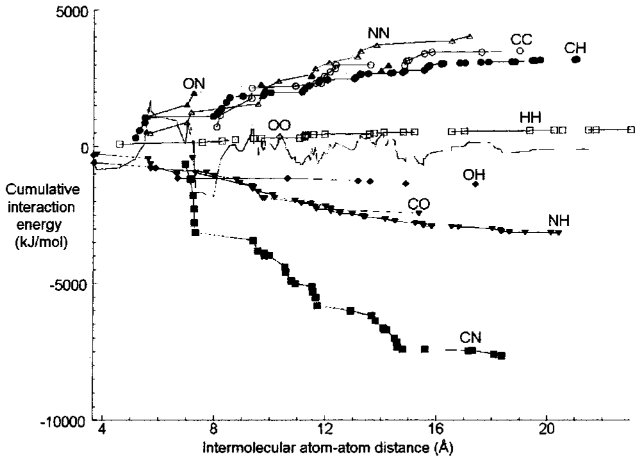

Figure 5. Energy profiles of GC-WC partitioned according to atom type. Cumulative electrostatic interaction energy (kJ/mol) versus the atom−atom distance between two bases (au). Open symbols: CC (circle), HH (square), OO (diamond), NN (upward triangle), CO (downward triangle). Solid symbols: CH (circle), CN (square), OH (diamond), ON (upward triangle), NH (downward triangle). The three primary interactions are highlighted in red and the four secondary interactions in blue. Total cumulated electrostatic energy (Figure 3a) is superimposed.

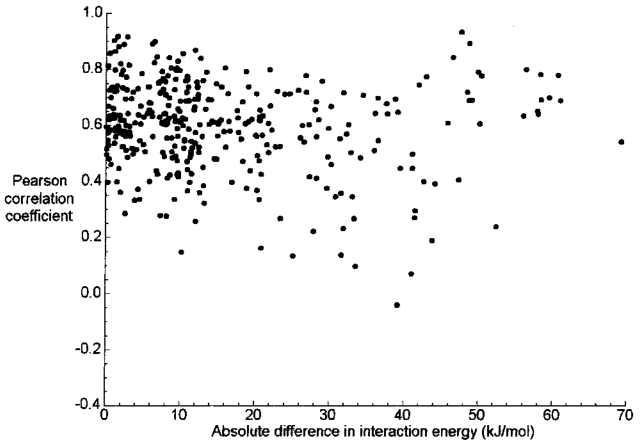

Figure 6. Scatter plot for all 351 base pair comparisons. The Pearson correlation coefficient between two different base pairs versus the absolute difference between the base pair interaction energies (kJ/mol). Each point represents one base pair. The overall correlation between the shapes of the energy profiles (as measured by the Pearson coefficient) and the difference in the interaction energy is poor ($r^2 = 0.02$).

coefficient is plotted against the absolute value of the energy difference of the base pairs it compares. The resulting scatter plot makes clear that similar electrostatic profiles do not automatically imply a small energy difference between the base pairs, since the upper part of the plot extends along the whole energy range. Conversely, the left part of the plot, which corresponds to base pairs with very similar energy, extends to rather low values of the correlation coefficient. As a consequence of these two observations we conclude that base pairs with similar interaction energies are not necessarily stable for the same reason.

Finally we scrutinize the support for the secondary hypothesis by quantum chemistry in a direct way. The energy profile corresponding to the sum of primary and secondary interaction energy throughout the series of base pairs does by no means reflect the *total* energy profile as can be seen from Figure 7. Hence the secondary interaction cannot be invoked to explain the pattern of natural base pair stability, not even in a semiquantitative manner.

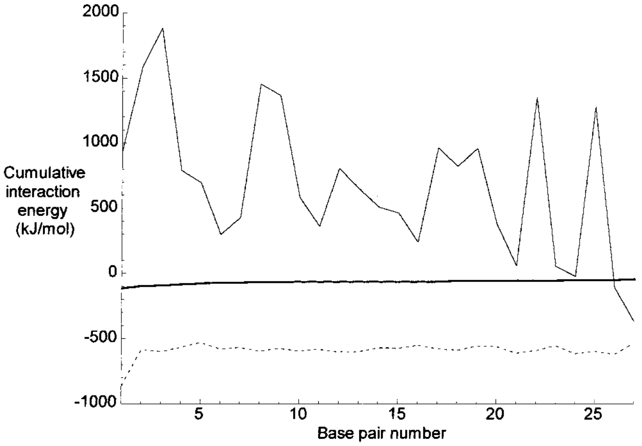

Figure 7. Cumulative interaction energy (kJ/mol) for primary interactions (up to 4 au) (dashed curve), the sum of primary and secondary interactions (up to 7 au) (plain curve), and the total interaction (bold). The base pairs are numbered from 1 to 27 corresponding to GCWC, GG1, GGNEW, CC, GG3, GC1, AC1, GT1, GA1, ATWC, GA4, GT2, ATRWC, AA1, ATH, AC2, ATRH, GG4, GA3, GA2, TT2, TC2, AA2, TT1, TC2, TT3, and AA3.

Possible concerns about the effect of polarization on the topological potential may be addressed by a simple calculation that places a unit test charge at a reasonable distance from one of the monomers. We carried out this test for GC and put the test charge roughly in the middle of the ring of C. The response of the electron density on G resulted in a change of 20% in the electrostatic potential energy of H10 in G (see Figure 1) at the site of the test charge. This atom is separated from the test charge by only 6.2 au. The next closest atom is subject to a change of 9%, followed by further removed atoms experiencing a change of around 1%. It is clear that the majority of long-range interactions are not substantially influenced by polarization. Although these calculations are not necessary to refute Jorgensen's purely electrostatic argument, they are useful in confirming the reliability of the current topological potential.

Conclusion

Our main question concerned the existence of subsets of atoms in two different bases forming a complex whose interaction energy parallels the *total* interaction energy. After a series of calculations designed to find support for the secondary interaction hypothesis we fail to find a physical basis for it in quantum chemistry when applied across molecular aggregates consisting of a very different number and type of atoms. Only in comparisons between highly similar chemical environments could the secondary interaction be invoked for the right reason.[32] However, in general simple rules to rationalize the pattern of energetic stability across naturally occurring base pairs in terms of subsets of atoms remains elusive. This work cautions against unjustified use of secondary interactions, which may lead to the same quandary that the hydrogen bond once introduced by its over-generalized use.

Acknowledgment. This work was supported by the Engineering and Physical Science Research Council (EPSRC), grant GR/M18119.

Supporting Information Available: Tables of energies and equilibrium structures, total interaction energies, and intermolecular geometric parameters (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

JA0125164