



**HAL**  
open science

# A PM6 quantum chemical study of the H-bonded and stacked associates of the adenine and thymine DNA bases: The nature of base stacking

Victor I. Danilov, Tanja van Mourik

## ► To cite this version:

Victor I. Danilov, Tanja van Mourik. A PM6 quantum chemical study of the H-bonded and stacked associates of the adenine and thymine DNA bases: The nature of base stacking. *Molecular Physics*, 2008, 106 (12-13), pp.1487-1494. 10.1080/00268970802060682 . hal-00513195

**HAL Id: hal-00513195**

**<https://hal.science/hal-00513195>**

Submitted on 1 Sep 2010

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



**A PM6 quantum chemical study of the H-bonded and stacked associates of the adenine and thymine DNA bases:  
The nature of base stacking**

Journal:	<i>Molecular Physics</i>
Manuscript ID:	TMPH-2008-0048.R1
Manuscript Type:	Full Paper
Date Submitted by the Author:	11-Mar-2008
Complete List of Authors:	Danilov, Victor; Institute of Molecular Biology and Genetics, National Academy of Sciences of Ukraine, Department of Molecular and Quantum Biophysics van Mourik, Tanja; University of St Andrews, School of Chemistry
Keywords:	adenine-thymine, base pair, base stacking, PM6



1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

**A PM6 quantum chemical study of the H-bonded and stacked associates of the adenine and thymine DNA bases: The nature of base stacking**

For Peer Review Only

1  
2  
3  
4  
5  
6  
7  
8

# A PM6 quantum chemical study of the H-bonded and stacked associates of the adenine and thymine DNA bases: The nature of base stacking

9 VICTOR I. DANILOV<sup>a</sup> and TANJA VAN MOURIK<sup>b\*</sup>

10  
11  
12  
13  
14 Running Head: A PM6 study on H-bonded and stacked adenine-thymine.

15  
16  
17 TANJA VAN MOURIK (Corresponding author)

18 School of Chemistry  
19 University of St. Andrews  
20 North Haugh  
21 Fife KY16 9ST  
22 Scotland (UK)

23  
24  
25 Email: [tanja.vanmourik@st-andrews.ac.uk](mailto:tanja.vanmourik@st-andrews.ac.uk)  
26 Phone: +44 (0)1334 463822  
27 Fax: +44 (0)1334 463808  
28

29  
30  
31 VICTOR I. DANILOV

32  
33 Department of Molecular and Quantum Biophysics  
34 Institute of Molecular Biology and Genetics  
35 National Academy of Sciences of Ukraine  
36 150 Zabolotny Street, Kiev-143, 03143  
37 Ukraine

38  
39 E-mail: [vid@ipnet.kiev.ua](mailto:vid@ipnet.kiev.ua)  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57

58  
59  
60

---

\* Corresponding author. Email: [tanja.vanmourik@st-andrews.ac.uk](mailto:tanja.vanmourik@st-andrews.ac.uk)

# A PM6 quantum chemical study of the H-bonded and stacked associates of the adenine and thymine DNA bases: The nature of base stacking

VICTOR I. DANILOV<sup>a</sup> and TANJA VAN MOURIK<sup>b†</sup>

<sup>a</sup>Department of Molecular and Quantum Biophysics, Institute of Molecular Biology and Genetics, National Academy of Sciences of Ukraine, 150 Zabolotny Street, Kiev-143, 03143, Ukraine

<sup>b</sup>School of Chemistry, University of St Andrews, North Haugh, St. Andrews, Fife, KY16 9ST, Scotland, UK

The planar H-bonded adenine-thymine (A•T) base pair and four stacked hetero associates (A/T) of these bases in a cluster containing 200 water molecules were studied by the new semi-empirical PM6 method. The formation of the planar A•T base pair in the water cluster was found to be energetically unfavourable, primarily due to the destabilising contribution of the base-water interactions. Under complete geometry optimisation, the planar A•T base pair converted into a nonplanar propeller-twisted and buckled associate. In contrast, the formation of all stacked dimers studied was found to be favourable, with formation energies ranging from -11 to -20 kcal/mol. The preference for the formation of these stacks results from the favourable change in the water structure and distortion of the paired and isolated bases during the base association reaction.

**Keywords:** adenine-thymine, base pair, base stacking, PM6

---

<sup>†</sup> Corresponding author. Email: tanja.vanmourik@st-andrews.ac.uk

## 1. Introduction

DNA (deoxyribonucleic acid) is one of the most important biomolecules, as it encodes the genetic information in the nucleus of cells. The basic structural units of DNA are nucleotides, which consist of a deoxyribose sugar, a phosphate group, and a base. The most common form of DNA is the double helix, consisting of two intertwined polynucleotide strands. In the double helix, the bases on opposite strands form base pairs, which are stabilised by hydrogen bonds (H-bonds). In the canonical DNA structure, adenine (A) binds to thymine (T) and cytosine (C) binds to guanine (G). Uracil (U) takes the place of thymine in RNA (ribonucleic acid). It is generally assumed that the stability of the DNA and RNA double helix arises from base-pairing interactions between complementary bases on opposite strands, as well as from stacking interactions between adjacent bases on one strand. The factors providing the stability of nucleic acid secondary structure in water are already evident in the interaction between nucleotide base pairs. Thus, the investigation of the nature of base-associate stability in water is very important for understanding the stability of different conformational states in nucleic acids, as well as for a better understanding of the mechanisms of intramolecular structural transitions in DNA.

Theoretical investigations find that, in vacuum, H-bonded base pairs are generally more stable than their stacked analogues [1-4]. In contrast, thermodynamic and spectroscopic studies on nucleic acid bases, nucleosides and nucleotides in aqueous solution have shown the formation of exclusively stacked associates [5]. The underlying physical forces that lead to base stacking are, however, not well understood and, indeed, the phenomenon has been attributed alternatively to the hydrophobic effect [6], dispersion interactions [7] and electrostatic interactions [8].

A variety of theoretical approaches have been used to account for solvent effects on base stacking. Sinanoğlu and Abdunur [9,10] deduced the important role of water on the formation of stacked dimers from estimates of the contribution of the solvent to the stability of nucleotide base associates. It was concluded that the main driving force for the

1  
2  
3 formation of the double helix is the hydrophobic effect, which was described in terms of a  
4 surface tension model. The important role of water on the formation of stacked dimers was  
5 also elucidated by one of us [1], on the basis of a careful analysis of the experimental and  
6 theoretical data available in the early 1970s. Pohorille and co-workers [11] used Monte  
7 Carlo (MC) simulations to study the Watson-Crick configuration of the H-bonded A•U  
8 base pair and stacked associates of the adenine and uracil bases. The geometries of the  
9 bases were taken from the minimum-energy structures obtained from quantum-mechanical  
10 perturbation calculations *i.e.*, fixed geometries were used in the simulations. The stacked  
11 homo-associates A/A and U/U were found to be energetically preferred, whereas the  
12 formation of the planar A•U base pair was unfavourable in water, in agreement with the  
13 experimental data. It was shown that the change in the water-water interaction associated  
14 with the hydrophobic effect is the main driving force for base stacking in solution.  
15 However, contrary to what is usually assumed [12], Pohorille's work found a destabilising  
16 water contribution to the enthalpy of association. Negative enthalpies were only obtained  
17 when the gas-phase solute-solute energy was added to the total energy. This surprising  
18 result may be an artefact of the use of rigid geometries for the stacked dimers. A MC study  
19 on the hydration of isolated uracil and thymine molecules, their stacked homo-associates  
20 (U/U and T/T) and H-bonded base pairs (T•T and U•U), conducted by us [13-16], also  
21 showed stacked base associates to be preferred over H-bonded base pairs in water. The  
22 stacking interactions of uracil and thymine in water are mainly caused by the increased  
23 water-water interaction from monomer to dimer, arising from the energetically more  
24 favourable water structure around the stacks. Thus, for the U/U and U•U associates, water  
25 causes an enthalpic stabilisation of the stacked dimers. Kollman and co-workers used the  
26 free energy perturbation/molecular dynamics method to study the solvation effect on the  
27 association of the N1- and N9-methylated DNA bases [17]. It was found that, in aqueous  
28 solution, the stacked associates are slightly more stable than the H-bonded complexes,  
29 whereas in vacuo, H-bonded complexes are favoured. The stacked associates were much  
30 better solvated than the H-bonded pairs. It was found that the base-base and base-water

1  
2  
3 interactions are primarily responsible for the formation of the A/A, A/T, and G/C stacked  
4  
5 complexes and that the Van der Waals energy is key in favouring the stacked complexes in  
6  
7 aqueous solution. However, only the standard B-DNA geometries were used for the  
8  
9 stacked associates. Dang and Kollman [18] presented the potential of mean force (PMF) for  
10  
11 the different associates of N9-methylated adenine and N1-methylated thymine in aqueous  
12  
13 solution. In agreement with the studies discussed above, the PMFs indicate that the stacked  
14  
15 configuration is more favourable than the H-bonded configuration in solution. A force field  
16  
17 and continuum solvation study of the free energy of association of nucleic acid bases in  
18  
19 aqueous solution found that electrostatic interactions oppose base stacking whereas  
20  
21 nonpolar interactions favour stacking [19]. Thus, nonpolar interactions, including the  
22  
23 hydrophobic effect and the enhancement of Van der Waals interactions by close-packing,  
24  
25 are responsible for base stacking, although their relative contribution is parameter  
26  
27 dependent and therefore difficult to isolate.

28  
29  
30 Direct experimental data on the nature of stacking interactions are difficult to obtain.  
31  
32 Nevertheless, substantial progress was attained in this direction over recent years. Guckian  
33  
34 *et al.* [20] used “dangling end” studies to evaluate the stacking behaviour of the four  
35  
36 natural nucleosides and nine non-natural analogues in the context of a hexamer DNA  
37  
38 duplex. Analysis of the obtained data showed that hydrophobic effects are larger than other  
39  
40 effects (electrostatic and dispersion) stabilising stacking. More recently, Frank-  
41  
42 Kamenetskii *et al.* obtained a complete set of free-energy parameters governing the  
43  
44 stacked–unstacked equilibrium at the site of the DNA nick [21,22], allowing the calculation  
45  
46 of the base-stacking and base-pairing contributions to the stability of A•T- and G•C-  
47  
48 containing DNA polymers from stacking parameters of individual contacts. As a result, it  
49  
50 was shown that base stacking is the main stabilising factor in the DNA double helix; A•T  
51  
52 pairing is always destabilising and G•C pairing barely contributes to its stabilisation. Thus,  
53  
54 as has been long anticipated, the double helix is stabilised by stacking interactions rather  
55  
56 than base pairing. Base-stacking interactions do not only dominate the overall stability of  
57  
58 the duplex, but also significantly contribute to the dependence of the stability on its base-  
59  
60



1  
2  
3 pair sequence. Therefore, base stacking is the main driving force responsible for the  
4  
5 stabilisation of the three-dimensional structure of DNA and RNA, a conclusion that is  
6  
7 qualitatively in accord with earlier theoretical predictions [17,18].  
8

9  
10 However, despite numerous experimental and theoretical investigations, the energetic  
11  
12 preferences of base association in water, the physical factors stabilising base stacking and,  
13  
14 in particular, the role of water as a solvent have not yet been adequately explored. It  
15  
16 remains as yet unclear why stacked dimers are preferred over H-bonded base pairs. The  
17  
18 role played by hydrophobic groups in base stacking is even less well understood.  
19  
20 Therefore, at present it is nearly impossible to predict how a specific structural change  
21  
22 made to an aromatic structure (such as a DNA base) would affect stacking. Generally, there  
23  
24 is no common picture of the relative significance of hydrophobic, electrostatic, or  
25  
26 dispersive effects on base stacking in water.  
27

28  
29 In the current work we study the association of adenine and thymine in aqueous  
30  
31 solution at the quantum-chemical level, enabled by the new semi-empirical PM6 method  
32  
33 [23]. We used the PM6 method recently to investigate hydrated complexes of thymine and  
34  
35 adenine [24]. In contrast to the previous MC simulations, the base geometries can be fully  
36  
37 optimised in the PM6 calculations, thereby allowing more realistic estimates of the base-  
38  
39 associate geometry. We chose the hydration of the H-bonded A•T base pair and stacked  
40  
41 A/T associates as the object of the current study, as it was experimentally shown that in  
42  
43 water adenine and thymine form stacked associates in an aqueous environment [25]. The  
44  
45 breakdown of the total internal energy of the isolated bases and their associates into  
46  
47 different energy components (the water energy, the water-base interaction and the base-  
48  
49 base interaction [26]) may contribute towards the elucidation of the mechanism of base  
50  
51 stacking.  
52  
53  
54  
55  
56  
57  
58  
59  
60

## 2. Methodology

### 2.1. *The Gibbs energy of hydration*

Prior to the 1970s there were several distinct quantities that were referred to as the “standard Gibbs energy of hydration” (or Gibbs energy of solvation, when the solvent is not water). All of these were defined in terms of the transfer of a solute from a specified standard state in the gas phase to another standard state in the liquid phase. However, there was some debate which of these quantities is the correct measure of the Gibbs energy of interaction between the solute and the solvent. In 1978 Ben-Naim introduced new definitions of hydration and of the corresponding thermodynamic quantities [27-29]. The hydration process was defined as the process of transferring a solute from a fixed position in an ideal gas phase to a fixed position in a liquid phase (a process that cannot be carried out in the laboratory). With this new concept the hydration energy has a very simple interpretation. It consists of two contributions: an average binding of solute to the system and the change in the average total interaction energy of the  $N$  water molecules, induced by the hydration process. This change may be re-interpreted as a structural change induced by the solute on the solvent. These hydration quantities are powerful tools for probing the effect of a solute on the water molecules in the immediate surroundings of the solute.

### 2.2. *Electronic structure calculations on adenine-thymine associates*

The internal energy and its different components (water-water, water-base and base-base interaction) of the isolated adenine and thymine bases and of their H-bonded (A•T) and stacked (A/T) hetero-associates, in a water cluster consisting of 200 solvent molecules, were computed with the semi-empirical PM6 method [23]. PM6 is the latest parameterisation of the NDDO method. The starting structures for the PM6 geometry optimisations were taken from MC simulations of the single base or dimer in the presence of 400 water molecules by selecting the 200 solvent molecules that were closest to the centre of mass of the solute. The initial geometry of the A•T base pair was taken to be the planar MP2/6-311G(2df,pd)-optimised Watson-Crick structure [25]. The internal

1  
2  
3 coordinates of the base pair were not optimised in the MC simulations. Two different PM6  
4  
5 optimisations were performed: one in which the internal coordinates of the base pair were  
6  
7 frozen at the MP2/6-311G(2df,pd)-optimised geometry, and one in which all coordinates  
8  
9 were fully optimised. As the exact geometry of the A/T stacks in water is not known [25],  
10  
11 we followed the procedure of Ref. [30] to generate initial configurations for the stacked  
12  
13 associates. We considered several configurations, starting from undisplaced stacked dimers  
14  
15 (with the axis connecting the centres of mass of the two bases perpendicular to the base  
16  
17 planes). From the undisplaced stack additional configurations were created by rotation  
18  
19 around two twist angles. The first twist angle defines the rotation of the upper base  
20  
21 (thymine) around the axis passing through its centre of mass and perpendicular to its plane.  
22  
23 We considered stacked associates with twist angles of  $0^\circ$  (where the thymine N1-H1 bond  
24  
25 and adenine N9-H9 bond are collinear) and  $180^\circ$  (with the twist applied in the right-handed  
26  
27 sense). The  $0^\circ$  associate corresponds to a parallel ( $A/T_{\text{para}}$ ) stack, whereas the  $180^\circ$   
28  
29 associate corresponds to an antiparallel ( $A/T_{\text{anti}}$ ) arrangement of the bases. From these, two  
30  
31 additional configurations were created by a (right-handed)  $180^\circ$  rotation of the upper base  
32  
33 of the dimer around the axis through the N1-H1 bond of thymine (see Figure 1 for the  
34  
35 labelling of the atoms). These are labelled  $A/T_{\text{para/rot}}$  and  $A/T_{\text{anti/rot}}$ , respectively. Since the  
36  
37 vertical base separations observed in all high-resolved oligonucleotides crystals constitute  
38  
39  $3.3 - 3.5 \text{ \AA}$ , the initial vertical separation between the two stacked bases was taken to be  
40  
41  $3.4 \text{ \AA}$  for all starting configurations. The geometries of the stacks were fully optimised in  
42  
43 the PM6 calculations.  
44  
45

46  
47 The PM6 calculations were carried out with MOPAC2007 [31]. The convergence  
48  
49 criterion for the geometry optimisation was a gradient norm of 1.0, which is sufficiently  
50  
51 small for the studied aqueous solution. The entropy is not considered in our simulations  
52  
53 since it has been shown experimentally [5] and theoretically [17] that the association of the  
54  
55 bases and their derivatives in water is primarily determined by enthalpy.  
56  
57  
58  
59  
60

### 2.3. Energy decomposition

The total internal energy of the AT (H-bonded A•T or stacked A/T) associate solvated in N water molecules is given by [27,28]:

$$E_{AT+N}(\text{solution}) = E_N(\text{AT} + \text{N}) + E_{AT}(\text{AT} + \text{N}) + E_{N/AT}(\text{AT} + \text{N}) \quad (1)$$

where  $E_N$  is the internal energy of the N water molecules;  $E_{AT}$  is the internal energy of the AT associate; and  $E_{N/AT}$  is the interaction energy between the N water molecules and the A and T bases. The part in round brackets indicates that these quantities are computed at their geometry in the AT+N system.

The quantity of interest here is the internal energy change upon introducing the two bases A and T in a water cluster composed of N water molecules:

$$\Delta E^{\text{hydr}}(R_{AT}) = E_{AT+N}(\text{solution}) - E_N(\text{N}) \quad (2)$$

where  $E_N(\text{N})$  is the internal energy of the N water molecules in the water cluster N. For large distances between A and T, the energy of coupling the two bases to a system of N water molecules is given by:

$$\Delta E^{\text{hydr}}(R_{AT} = \infty) = [E_{A+N}(\text{solution}) - E_N(\text{N})] + [E_{T+N}(\text{solution}) - E_N(\text{N})] \quad (3)$$

where  $[E_{B+N}(\text{solution}) - E_N(\text{N})]$  is the internal energy change upon introducing base B (B = A or T) in a water cluster composed of N water molecules i.e., the internal energy of hydration of the base B.

The total internal energy change ( $\Delta E^{\text{form}}$ ) for the process of bringing the two bases A and T from infinity to a distance that within the liquid is of the order of magnitude of the molecular diameter of the solute (i.e., the formation energy of the base associate) can be written as:

$$\Delta E^{\text{form}} = E_{AT}(R_{AT} = \sigma) + \Delta E^{\text{hydr}}(R_{AT} = \sigma) - \Delta E^{\text{hydr}}(R_{AT} = \infty) \quad (4)$$

where the interaction energy of the bases,  $E_{AT}(R_{AT} = \sigma) \equiv \Delta E_{bb}(\text{gas})$ , is given by:

$$\Delta E_{bb}(\text{gas}) = E_{AT}(\text{gas}) - [E_A(\text{gas}) + E_T(\text{gas})] \quad (5)$$

Using equations (1) – (3) and (5), as well as expressions equivalent to (1) for the total internal energy of A and T, we can rewrite the total energy change  $\Delta E$  for process (4) as:

$$\Delta E^{\text{form}} = \Delta E_{\text{bb}}(\text{gas}) + \Delta E_{\text{dc}}(\text{wat}) + \Delta E_{\text{wb}} + \Delta E_{\text{ww}} \quad (6)$$

Here, the distortion cost,  $\Delta E_{\text{dc}}(\text{wat})$ , defined as the internal energy difference between the paired and isolated bases caused by the presence of the water cluster and by geometric changes during the base association reaction, is given by:

$$\Delta E_{\text{dc}}(\text{wat}) = E_{\text{AT}}(\text{AT+N}) - [E_{\text{A}}(\text{A+N}) + E_{\text{T}}(\text{T+N})] \quad (7)$$

The water-base interaction energy change,  $\Delta E_{\text{wb}}$ , is given by:

$$\Delta E_{\text{wb}} = E_{\text{N/AT}}(\text{AT+N}) - [E_{\text{N/A}}(\text{A+N}) + E_{\text{N/T}}(\text{T+N})] \quad (8)$$

and the water energy change,  $\Delta E_{\text{ww}}$ , is given by:

$$\Delta E_{\text{ww}} = E_{\text{N}}(\text{AT+N}) + E_{\text{N}}(\text{N}) - [E_{\text{N}}(\text{A+N}) + E_{\text{N}}(\text{T+N})] \quad (9)$$

### 3. Results and Discussion

Table 1 lists the total internal energy  $E_{\text{B+N}}(\text{solution})$  and its energy components (where B = A or T) for the A-(H<sub>2</sub>O)<sub>200</sub>, T-(H<sub>2</sub>O)<sub>200</sub> and (H<sub>2</sub>O)<sub>200</sub> systems. The decrease in the magnitude of the internal energy of the water cluster ( $E_{\text{N}}$ ) upon solvation of the adenine or thymine base shows that the process of transferring the base from vacuum into solution destroys the optimal arrangement of the water molecules in the water cluster. The smaller magnitude of the water energy for T-(H<sub>2</sub>O)<sub>200</sub> as compared to A-(H<sub>2</sub>O)<sub>200</sub> is evidence of the energetically less preferable structure of water around thymine. This is however compensated by a more favourable water-base interaction ( $E_{\text{N/B}}$ ).

Table 2 lists the total internal energy and its components for the A•T base pair solvated in a cluster consisting of 200 water molecules. The base pair was either fixed at its planar MP2/6-311G(2df,pd)-optimised geometry (labelled “Fixed” in Table 2) or fully optimised (labelled “Optimised”) in the PM6 calculations. The MP2/6-311G(2df,pd)-optimised geometry of the base pair is shown in Fig. 1. Complete geometry optimisation of the A•T

1  
2  
3 base pair in the presence of 200 water molecules led to an intrinsically nonplanar propeller-  
4  
5 twisted and buckled structure. The magnitude of the energy of the water cluster is larger for  
6  
7 the planar A•T base pair, indicating a more favourable arrangement of the water molecules  
8  
9 around this solute. However, the fully-optimised base pair is energetically preferred due to  
10  
11 the larger stabilising contribution of the water-base interaction.

12  
13 Table 3 lists the energy quantities related to the base pairing reaction of the planar and  
14  
15 fully-optimised A•T base pair in a cluster consisting of 200 water molecules, computed  
16  
17 using equations (5) – (9) and the data presented in tables 1 and 2. The results show that the  
18  
19 formation of the planar A•T base pair is not energetically favourable. Although  $\Delta E_{\text{ww}}$ ,  
20  
21  $\Delta E_{\text{dc}}(\text{wat})$  and  $\Delta E_{\text{bb}}(\text{gas})$  make stabilising contributions to the formation of the planar A•T  
22  
23 base pair, it does not form because of the large destabilising contribution of the water-base  
24  
25 interaction. Thus, two separated bases are energetically preferred over the planar base pair.  
26  
27 This is not surprising, since the loss of the stabilising contributions in the base pair can  
28  
29 easily be compensated by the stabilising interactions of the separated bases with the water  
30  
31 molecules. The formation energy of the fully-optimised nonplanar base pair constitutes  
32  
33 -10.0 kcal/mol, indicating that the formation of this base pair is energetically favourable.  
34  
35 This is mainly due to the much decreased destabilising contribution of the water-base  
36  
37 interaction. Thus, the nonplanar base pair is stable in a cluster of 200 water molecules,  
38  
39 whereas the planar base pair is not.  
40  
41

42  
43 The PM6-optimised configurations of the stacked associates in a water cluster  
44  
45 consisting of 200 water molecules are presented in Fig. 2. The bases in the stacks are not  
46  
47 exactly parallel. PM6 predicts the deviation from parallel bases to be up to about  $10^\circ$ . The  
48  
49 total internal energy and its components of the solvated stacks are listed in table 4. The  
50  
51  $E_{\text{AT+N}}(\text{solution})$  values show that the  $A/T_{\text{para/rot}}-(\text{H}_2\text{O})_{200}$  system is the energetically most  
52  
53 stable of the four different structures considered. This appears to be entirely due to the  
54  
55 larger base-water interaction; the internal energies of both the base associate and the water  
56  
57 cluster are in fact less favourable than those of the other three structures.  
58  
59  
60

1  
2  
3 The larger base-water interaction energies of the stacked dimers as compared to the H-  
4 bonded base pairs indicate that the stacked dimers are much better hydrated by the water  
5 cluster. The reason for this is that the atoms that participate in the formation of the H-bonds  
6 in the H-bonded base pairs are exposed to the water cluster in the stacked dimers. This  
7 creates an additional stabilisation, which is stronger than the energy of the H-bonds in the  
8 base pair. The same conclusion was also reached in a previous molecular dynamics study  
9 of the association of the nucleotide bases in water [17]. Application of equation (2) using  
10 the data given in tables 1, 2 and 4 allows the calculation of the internal energy of hydration  
11 of the planar and stacked dimers. The calculated values (not shown) support this  
12 conclusion.  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23

24 Although the H-bonded base pairs are not observed experimentally in water, the  
25 formation energy changes for the transition from the H-bonded to the different stacked  
26 dimers are of interest. These values (computed from the data given in tables 2 and 4) are  
27 presented in Table 5. It follows from Table 5 that all stacked associates are energetically  
28 preferred over the planar H-bonded base pair. The energy values give direct evidence that  
29 the water-base interaction energy is the determining factor in favouring the A/T stacked  
30 associates over the A•T base pair in an aqueous cluster. This may be due to the smaller  
31 hydrophobic surface available in the stacks for (unfavourable) base-water interaction. The  
32 change in the energy of the water cluster destabilises the stacks compared with the planar  
33 base pair, related to the energetically less favourable arrangement of the water molecules  
34 around the stacks. The stacks are also destabilised by the distortion factor. Thus, the greater  
35 stability of the stacked dimer is entirely due to the differing water-base interactions in the  
36 stacked and H-bonded dimers.  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50

51 Expressions (5) – (9), together with the data presented in tables 1 and 4, allow us to  
52 calculate the formation energy and its various contributions of the four different stacked  
53 associates in the water cluster. These results are listed in Table 6. In contrast to the planar  
54 A•T base pair, the formation of all stacked associates is highly favourable. The water  
55  
56  
57  
58  
59  
60

energy change  $\Delta E_{\text{ww}}$ , associated with the structural rearrangement of the water molecules around the bases during their association, contributes most to the stabilisation of the stacks, whereas the change of the internal energy of the bases,  $\Delta E_{\text{dc}}(\text{wat})$ , has the second-largest contribution. The only exception is  $A/T_{\text{para/rot}}-(\text{H}_2\text{O})_{200}$ , for which the water-base interaction is more stabilising than the water energy change. The stacked associates are noticeably less stabilised by the base-base interaction than the H-bonded base pair. It should be emphasised that the  $\Delta E_{\text{dc}}(\text{wat})$  contribution was absent in Monte Carlo studies of the hydration of the base associates [11,13-16], and considering the important contribution of  $\Delta E_{\text{dc}}(\text{wat})$  in the base stacking process, the PM6 calculations are expected to provide more realistic results. Thus, the water-water interaction is one of the main factors promoting stacked dimer formation, and the obtained data are a direct confirmation of the crucial role of the water-water interaction in base stacking earlier reported in Ref. [1].

#### 4. Conclusion

The current PM6 study on hydrated base associates shows that the base-pairing reaction in a water cluster is energetically unfavourable, in agreement with the results of Monte Carlo simulations [11,13-16]. The change in the structure of the water cluster during this reaction makes a stabilising contribution to the formation of the A•T base pair, whereas the water-base interaction is destabilising. In contrast, the base-stacking reaction in water is favourable. The water energy change, associated with the structural rearrangement of the water molecules around the bases during base association, is the main contribution stabilising the stacks. The water-base interaction is in this case destabilising (except for the  $A/T_{\text{para/rot}}$  stack). The transition of the planar H-bonded base pair to the stacked associate in an aqueous cluster of 200 water molecules is energetically favourable. The PM6 results show that this is entirely caused by the large stabilising contribution of the water-base interaction. The water energy change destabilises the stacks as compared with the planar base pair. The Monte Carlo simulations explain this by the energetically more favourable



1  
2  
3 structure of the water molecules around the stacks and to a lesser degree by the stabilising  
4  
5 action of the water-base interaction.  
6

7 The current PM6 study provides a first step in the quantum-chemical investigation of  
8  
9 base stacking in aqueous solution. However, further investigation into the physical basis of  
10  
11 nucleic acid base stacking may be warranted to validate the PM6 results presented in this  
12  
13 work. For example, we recently found that, although PM6 predicted the correct relative  
14  
15 stability of different hydrated clusters of adenine and thymine, there were notable  
16  
17 differences in the water network arrangement in the PM6- and B3LYP-optimised structures  
18  
19 [24]. Quantum chemical methods that explicitly take into account electron correlation are  
20  
21 expected to more accurately describe the dispersion energy, which plays a major role in the  
22  
23 base-base interactions in the stacked structures. Efficient correlated ab initio methods like  
24  
25 df-LMP2 (density-fitting local MP2 [32-35]) or density functional theory with functionals  
26  
27 that correctly describe Van der Waals interactions (like Truhlar's M05 [36] and M06 [37]  
28  
29 series of density functionals) may be able to provide benchmark results for the semi-  
30  
31 empirical values presented in the current paper.  
32  
33

### 34 Acknowledgements

35  
36  
37 One of us (V. I. D.) is grateful to Prof. J. J. P. Stewart for helpful discussions on the  
38  
39 PM6 method and for additional output data, and to Prof. A. Ben-Naim for helpful  
40  
41 discussions on the hydration of biomolecules and for copies of his publications. T.v.M.  
42  
43 gratefully acknowledges the Royal Society for their support under the University Research  
44  
45 Fellowship scheme.  
46  
47

### 48 References

- 49  
50 [1] V. I. Danilov, Mol. Biol. Repts **2**, 263 (1975).  
51 [2] A. Pullman, B. Pullman, Adv. Quantum Chem. **4**, 267 (1968).  
52 [3] J. Šponer, P. Jurečka, P. Hobza in *Computational studies of RNA and DNA*, edited  
53 by J. Šponer and F. Lankaš (Springer, Dordrecht, 2006), pp. 343.  
54 [4] J. Šponer, P. Jurečka, I. Marchan, F. J. Luque, M. Orozco, P. Hobza, Chem. Eur. J.  
55 **12**, 2854 (2006).  
56 [5] P. O. P. Ts'o, in *Basic Principles in Nucleic Acid Chemistry*, edited by Paul O.P.  
57 Ts'o (Academic Press, London and New York, 1974) p. 453.  
58  
59  
60

- 1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60
- [6] T. T. Herskovits, S. J. Singer, E. P. Geiduschek, *Arch. Biochem. Biophys.* **94**(1961).
- [7] S. Hanlon, *Biochem. Biophys. Res. Commun.* **23**, 861 (1966).
- [8] L. F. Newcomb, S. H. Gellman, *J. Am. Chem. Soc.* **116**, 4993 (1994).
- [9] O. Sinanoğlu, in *Molecular Associations in Biology*, edited by B. Pullman (Academic Press, New York, 1968), p. 427.
- [10] O. Sinanoğlu, S. Abdulnur, *Fed. Proc.* **24**, **Suppl. 15**, S12 (1965).
- [11] A. Pohorille, L. R. Pratt, S. K. Burt, R. D. MacElroy, *J. Biomol. Struct. Dyn.* **1**, 1257 (1984).
- [12] C. R. Cantor and P. R. Schimmel, *Biophysical chemistry, Part I. The conformation of biological macromolecules* (W.H. Friman and Co., San Francisco, 1980).
- [13] V. I. Danilov, I. S. Tolokh, *J. Biomol. Struct. Dyn.* **2**, 119 (1984).
- [14] V. I. Danilov, I. S. Tolokh, V. I. Poltev, G. G. Malenkov, *FEBS Lett.* **167**, 245 (1984).
- [15] V. I. Danilov, I. S. Tolokh, V. I. Poltev, *FEBS Lett.* **167**, 325 (1984).
- [16] V. I. Danilov, I. S. Tolokh, *FEBS Lett.* **173**, 347 (1984).
- [17] P. Cieplak, P. A. Kollman, *J. Am. Chem. Soc.* **110**, 3734 (1988).
- [18] L. X. Dang, P. A. Kollman, *J. Am. Chem. Soc.* **112**, 503 (1990).
- [19] R. A. Friedman, B. Honig, *Biophys. J.* **69**, 1528 (1995).
- [20] K. M. Guckian, B. A. Schweitzer, R. X. F. Ren, C. J. Sheils, D. C. Tahmassebi, E. T. Kool, *J. Am. Chem. Soc.* **122**, 2213 (2000).
- [21] E. Protozanova, P. Yakovchuk, M. D. Frank-Kamenetskii, *J. Mol. Biol.* **342**, 775 (2004).
- [22] P. Yakovchuk, E. Protozanova, M. D. Frank-Kamenetskii, *Nucleic Acids Res.* **34**, 564 (2006).
- [23] J. J. P. Stewart, *J. Mol. Model.* **13**, 1173 (2007).
- [24] T. van Mourik, V. I. Danilov, E. Gonzales, A. Deriabina, V. I. Poltev, *Chem. Phys. Lett.* **445**, 305 (2007).
- [25] N. I. Nakano, S. J. Igarashi, *Biochemistry* **9**, 577 (1970).
- [26] H. Sapper, W. Lohmann, *Biophys. Struct. Mech.* **4**, 327 (1978).
- [27] A. Ben-Naim, *Molecular theory of solutions* (Oxford University Press, New York, 2006).
- [28] A. Ben-Naim, *Solvation Thermodynamics*, (Plenum Press, New York, 1987).
- [29] A. Ben-Naim, *J. Phys. Chem.* **82**, 792 (1978).
- [30] J. Šponer, J. Leszczynski, P. Hobza, *J. Phys. Chem.* **100**, 5590 (1996).
- [31] J. P. P. Stewart, MOPAC2007, <http://OpenMOPAC.net> (Stewart Computational Chemistry, Colorado Springs, CO, USA, 2007).
- [32] G. Hetzer, P. Pulay, H.-J. Werner, *Chem. Phys. Lett.* **290**, 143 (1998).
- [33] G. Hetzer, M. Schütz, H. Stoll, H.-J. Werner, *J. Chem. Phys.* **113**, 9443 (2000).
- [34] M. Schütz, G. Hetzer, H.-J. Werner, *J. Chem. Phys.* **111**, 5691 (1999).
- [35] H.-J. Werner, F. R. Manby, P. J. Knowles, *J. Chem. Phys.* **118**, 8149 (2003).
- [36] Y. Zhao, N. E. Schultz, D. G. Truhlar, *J. Chem. Theor. Comput.* **2**, 364 (2006).
- [37] Y. Zhao, D. G. Truhlar, *Theor. Chem. Acc.*(in press).

Table 1. The total internal energy and energy components (in eV) of the A-(H<sub>2</sub>O)<sub>200</sub>, T-(H<sub>2</sub>O)<sub>200</sub> and (H<sub>2</sub>O)<sub>200</sub> systems.

Energy quantity <sup>a</sup>	A-(H <sub>2</sub> O) <sub>200</sub>	T-(H <sub>2</sub> O) <sub>200</sub>	(H <sub>2</sub> O) <sub>200</sub>
E <sub>B+N</sub> (solution)	-65473.82	-65533.61	-63896.56
E <sub>B</sub> (B+N)	-1575.53	-1633.83	
E <sub>N/B</sub> (B+N)	-4.61	-8.09	
E <sub>N</sub> (B+N)	-63893.68	-63891.64	-63896.56

<sup>a</sup>B = A (adenine) or T (thymine)

Table 2. The total internal energy and energy components (in eV) of the (A•T)-(H<sub>2</sub>O)<sub>200</sub> complex.

Energy quantity <sup>a</sup>	Fixed <sup>b</sup>	Optimised <sup>c</sup>
E <sub>AT+N</sub> (solution)	-67109.40	-67110.88
E <sub>AT</sub> (AT+N)	-3211.32	-3210.88
E <sub>N/AT</sub> (AT+N)	-5.90	-9.35
E <sub>N</sub> (AT+N)	-63892.18	-63890.65
E <sub>AT</sub> (gas)	-3214.05	-3214.05
E <sub>A</sub> (gas)	-1576.89	-1576.89
E <sub>T</sub> (gas)	-1636.77	-1636.77

<sup>a</sup>B = A (adenine) or T (thymine)

<sup>b</sup>The A•T base pair fixed at its planar MP2/6-311G(2df,pd) optimised structure.

<sup>c</sup>The A•T base pair fully optimised in the presence of the water molecules.

Table 3. Energetics (in kcal/mol) of the base pairing reaction in a water cluster consisting of 200 water molecules.

Energy quantity	Fixed <sup>a</sup>	Optimised <sup>b</sup>
$\Delta E^{\text{form}}$	24.0	-10.1
$\Delta E_{\text{bb}}(\text{gas})$	-8.8	-8.8
$\Delta E_{\text{dc}}(\text{wat})$	-45.2	-34.9
$\Delta E_{\text{wb}}$	156.9	77.3
$\Delta E_{\text{ww}}$	-78.9	-43.7

<sup>a</sup>The A•T base pair fixed at its planar MP2/6-311G(2df,pd) optimised structure.

<sup>b</sup>The A•T base pair fully optimised in the presence of the water molecules.

Table 4. The total internal energy and energy components (in eV) of different structures of the stacked A/T-(H<sub>2</sub>O)<sub>200</sub> complex.

Energy quantity <sup>a</sup>	A/T <sub>para</sub> -(H <sub>2</sub> O) <sub>200</sub>	A/T <sub>anti</sub> -(H <sub>2</sub> O) <sub>200</sub>	A/T <sub>para/rot</sub> -(H <sub>2</sub> O) <sub>200</sub>	A/T <sub>anti/rot</sub> -(H <sub>2</sub> O) <sub>200</sub>
E <sub>AT+N</sub> (solution)	-67111.37	-67111.08	-67111.48	-67111.17
E <sub>AT</sub> (B+N)	-3210.08	-3210.80	-3209.36	-3210.23
E <sub>N/AT</sub> (B+N)	-10.63	-9.26	-13.07	-10.88
E <sub>N</sub> (B+N)	-63890.66	-63891.02	-63889.06	-63890.05
E <sub>AT</sub> (gas)	-3213.89	-3213.89	-3213.89	-3213.85
E <sub>A</sub> (gas)	-1576.89	-1576.89	-1576.89	-1576.89
E <sub>T</sub> (gas)	-1636.77	-1636.77	-1636.77	-1636.77

<sup>a</sup>B = A (adenine) or T (thymine)

Table 5. Energy values (in kcal/mol) of the transition from the planar H-bonded base pair to the different stacked associates, in a cluster consisting of 200 water molecules.

Transition	$\Delta E$	$\Delta E_{dc}(\text{wat})$	$\Delta E_{wb}$	$\Delta E_{ww}$
$A \cdot T \rightarrow A/T_{\text{para}}$	-45.5	28.5	-109.1	35.1
$A \cdot T \rightarrow A/T_{\text{anti}}$	-38.8	11.9	-77.5	26.8
$A \cdot T \rightarrow A/T_{\text{para/rot}}$	-48.1	45.1	-165.3	72.1
$A \cdot T \rightarrow A/T_{\text{anti/rot}}$	-40.7	25.1	-114.9	49.1

Table 6. Energy values (in kcal/mol) of the base stacking reaction.

Energy quantity <sup>a</sup>	A/T <sub>para</sub> -(H <sub>2</sub> O) <sub>200</sub>	A/T <sub>anti</sub> -(H <sub>2</sub> O) <sub>200</sub>	A/T <sub>para/rot</sub> -(H <sub>2</sub> O) <sub>200</sub>	A/T <sub>anti/rot</sub> -(H <sub>2</sub> O) <sub>200</sub>
$\Delta E^{\text{form}}$	-17.8	-11.2	-20.2	-12.1
$\Delta E_{\text{bb}}(\text{gas})$	-5.2	-5.2	-5.1	-4.2
$\Delta E_{\text{dc}}(\text{wat})$	-16.7	-33.3	0.0	-20.1
$\Delta E_{\text{wb}}$	47.8	79.4	-8.3	42.0
$\Delta E_{\text{ww}}$	-43.7	-52.1	-6.8	-29.8



**Figure Captions**

Figure 1. The planar MP2/6-311G(2df,pd)-optimised Watson-Crick base pair.

Figure 2. The structure of the different adenine-thymine stacks optimised by the PM6 method in the presence of 200 water molecules. (a).  $A/T_{\text{para}}$ . (b).  $A/T_{\text{anti}}$ . (c).  $A/T_{\text{para/rot}}$ . (d).  $A/T_{\text{anti/rot}}$ . The stick diagrams show a top view of the stacks.

For Peer Review Only

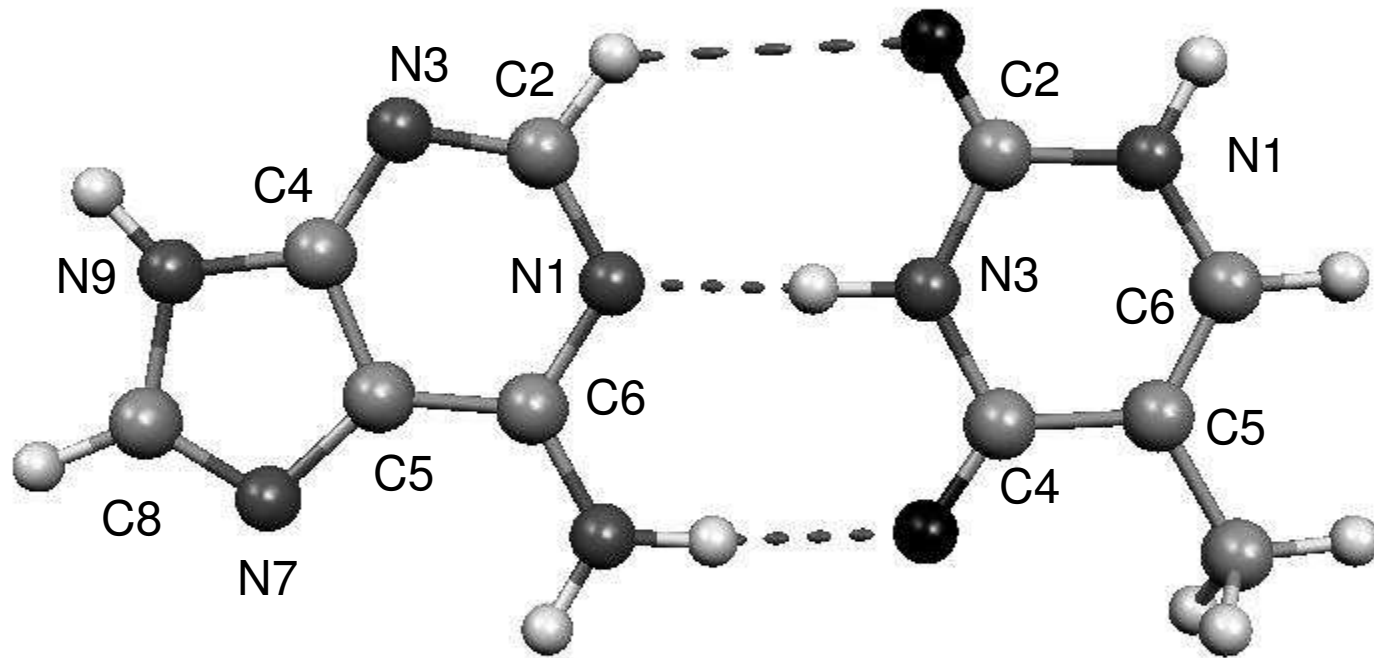


Figure 1

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49

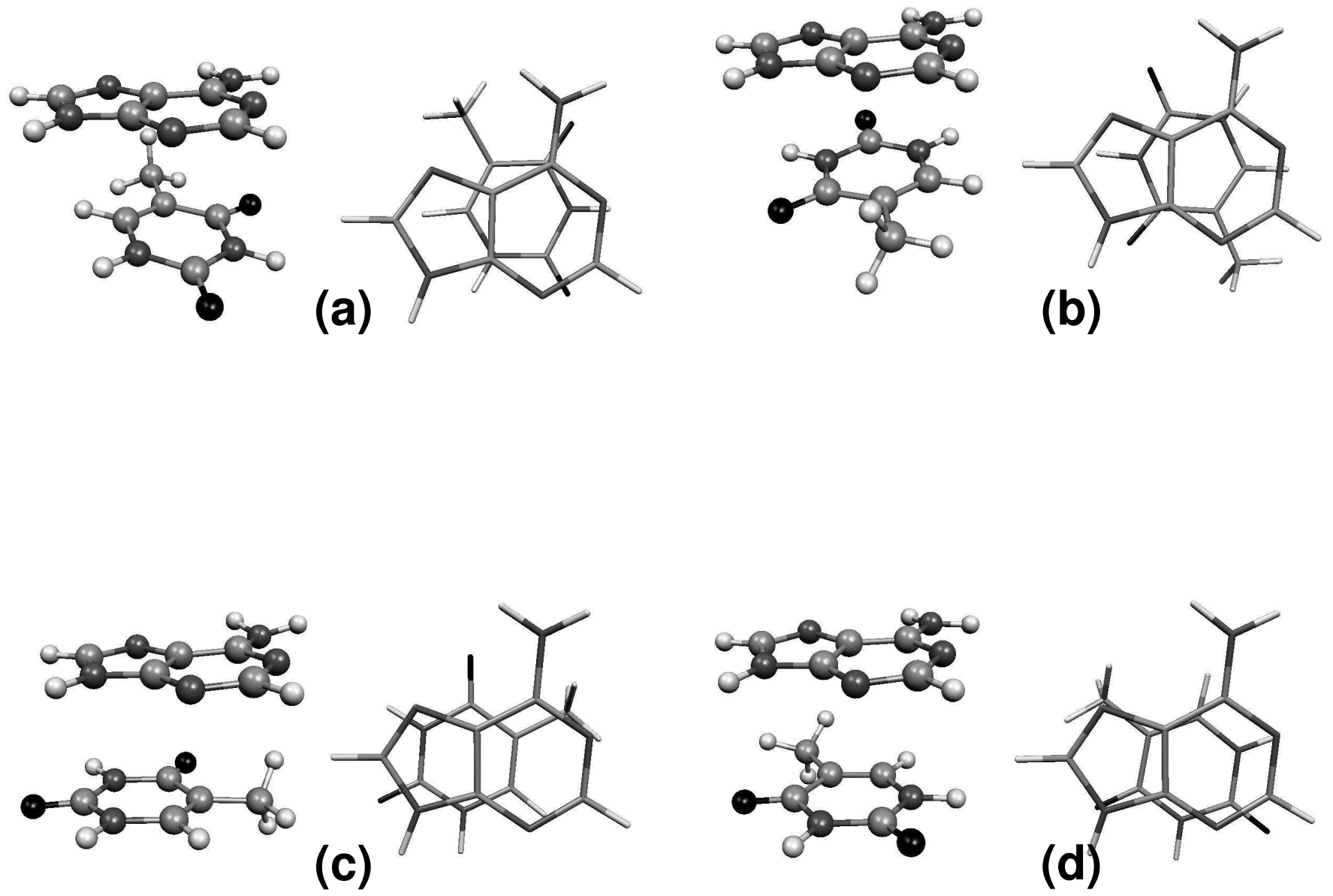


Figure 2

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49