

Free Energy Calculations of Watson–Crick Base Pairing in Aqueous Solution

E. Stofer, C. Chipot,[†] and R. Lavery*

Contribution from the Laboratoire de Biochimie Théorique, CNRS UPR 9080, Institut de Biologie Physico-Chimique, 13 rue Pierre et Marie Curie, Paris 75005, France, and Laboratoire de Chimie Théorique, CNRS UMR 7565, Université Henri Poincaré-Nancy I, BP 239, Vandoeuvre-les-Nancy, France

Received April 6, 1999

Abstract: The free energy of association of adenine-thymine and guanine-cytosine base pairs is estimated from potential of mean force computations in water using the AMBER potential energy function. The results, which lead to free energies per hydrogen bond of roughly 2 kcal/mol, highlight the existence of secondary water-bridged minima for both base pairs and also show that adenine-thymine pairs undergo frequent exchanges between canonical and reversed Watson–Crick states by contra-rotation of the bases around their common axis.

Introduction

Base pairing is of fundamental biological importance as the mechanism for homologous recognition between the strands of duplex DNA.¹ These highly specific interactions have become an archetype for molecular auto-assembly and, as such, they have inspired many attempts at engineering specifically binding, polydentate receptors.^{2,3} Even, within the biological world, base pairing interactions are very versatile, serving not only within the double helix, but also for the formation of more complex architectures for both DNA and RNA.⁴

The thermodynamics of base pairing has been studied for many years in various environments. Vacuum experiments lead to pairing enthalpies of 21 kcal/mol for guanine-cytosine (GC) pairs and 13 kcal/mol for adenine-thymine (AT) pairs.⁵ In nonpolar, or moderately polar, solvents these values are understandably reduced and, as an example, GC pairing yields enthalpies of roughly -11 kcal/mol in chloroform^{6,7} and -6 kcal/mol in dimethyl sulfoxide.⁸ Such data have been used to develop additive schemes for predicting pairing in solvents of this nature.^{9,10} However, in water, the biologically relevant solvent, pairing energies are weaker than stacking interactions and no direct measurements of pairing can be made. To overcome this difficulty it is possible to study the strand interactions of oligonucleotide DNA or RNA duplexes.¹¹ Such interactions are stable in water with moderate salt concentrations, although it

is clear that strand interactions involve not only base pairing, but also base stacking and factors linked to conformational changes in the backbone.¹² Stacking and pairing interactions are in fact coupled, stronger stacking competing with optimal pairing.¹³ On the basis of experiments concerning the terminal base pairs of oligonucleotides, the free energy increment associated with the a single hydrogen bond is estimated as 0.8 and 1.6 kcal/mol.¹⁴ However, in experiments on guanine to inosine substitutions, extrapolated to zero stacking, this increment increases to roughly 2.0 kcal/mol.¹⁵ A similar estimate (1.7–2.2 kcal/mol) was obtained in experiments on guanosine binding to a catalytic RNA.¹⁶

Concerning stacking, experimental values cover a relatively large range, although it is generally admitted that the strengths of the interactions are in the following order: $R-R > R-Y > Y-R$ ($R =$ purine, $Y =$ pyrimidine). Studies involving a short double-stranded DNA with a dangling 5'-nucleotide led to values of -2.0 kcal/mol for ApC and -1.1 kcal/mol for TpC.¹⁷ For free bases, spectroscopic studies gave larger values for AA (-5.7 kcal/mol) and UU (-1.7 kcal/mol) stacks.^{18–20}

Recently, single molecule force experiments on oligomeric^{21,22} and polymeric DNA,^{23,24} between complementary, adsorbed

* Address correspondence to this author at Institut de Biologie Physico-Chimique.

[†] Université Henri Poincaré-Nancy I.

(1) Saenger, W. *Principles of Nucleic Acid Structure*; Springer-Verlag: New York, 1984.

(2) Goswami, S.; Hamilton, A. D. *J. Am. Chem. Soc.* **1989**, *111*, 3425.

(3) Rotello, V. M.; Viani, E. A.; Deslongchamps, G.; Murray, B. A.; Rebek, J., Jr. *J. Am. Chem. Soc.* **1993**, *115*, 797.

(4) Lebrun, A.; Lavery, R. *Curr. Opin. Struct. Biol.* **1997**, *7*, 348.

(5) Yanson, I. K.; Teplitsky, A. R.; Sukhodub, L. F. *Biopolymers* **1979**, *18*, 1149.

(6) Kyogoku, Y.; Lord, R. C.; Rich, A. *J. Am. Chem. Soc.* **1967**, *89*, 496.

(7) Kyogoku, Y.; Lord, R. C.; Rich, A. *Biophys. Biochem. Acta* **1969**, *179*, 10.

(8) Newmark, R. A.; Cantor, C. R. *J. Am. Chem. Soc.* **1968**, *90*, 5010.

(9) Jorgensen, W.; Pranata, J. *J. Am. Chem. Soc.* **1990**, *112*, 2008.

(10) Schneider, H. *J. Chem. Soc. Rev.* **1994**, *22*, 227.

(11) Breslauer, K. J.; Frank, R.; Blöcker, H.; Marky, L. A. *Proc. Natl. Acad. Sci. U.S.A.* **1986**, *83*, 3746.

(12) Vesnaver, G.; Breslauer, K. J. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 3569.

(13) Levitt, M. *Proc. Natl. Acad. Sci. U.S.A.* **1978**, *75*, 640.

(14) Freier, S. M.; Sugimoto, N.; Sinclair, A.; Alkema, D.; Neilson, T.; Kierzek, R.; Caruthers, M. H.; Turner, D. H. *Biochemistry* **1986**, *25*, 3214.

(15) Turner, D. H.; Sugimoto, N.; Kierzek, R.; Dreiker, S. D. *J. Am. Chem. Soc.* **1987**, *109*, 3783.

(16) Tanner, N. K.; Cech, T. R. *Biochemistry* **1987**, *26*, 3330.

(17) Guckian, K. M.; Schweitzer, B. A.; Ren R. X.-F.; Scheils, C. J.; Paris, P. L.; Tahmassebi, D. C.; Kool E. T. *J. Am. Chem. Soc.* **1996**, *118*, 8182.

(18) Iza, N.; Gill, M.; Montero, J. L.; Morcillo, J. *J. Mol. Struct.* **1988**, *175*, 19.

(19) Morcillo, J.; Gallego, E.; Peral, F. *J. Mol. Struct.* **1987**, *157*, 353.

(20) Friedman, R. A.; Honig, B. *Biophys. J.* **1995**, *69*, 1528.

(21) Lee G. U.; Chrisey, L. A.; Colton, R. J. *Science* **1994**, *266*, 771.

(22) Noy, A.; Vezenov, D. V.; Kayyem, J. F.; Meade, T. J.; Lieber, C. M. *Chem. Biol.* **1997**, *4*, 519.

(23) Essevaz-Roulet, B.; Bockelmann, U.; Heslot, F. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 11935.

bases²⁵ and also on chimeric molecules carrying nucleic bases as headgroups,²⁶ have added novel paths for investigating pairing interactions. However, these approaches also have their own difficulties (resolving single unpairing events, relationship between loading rate and measured forces, etc.).

Because of these difficulties, present estimates for base pairing in water come mainly from theoretical calculations. Although simplified continuum solvent models²⁷ could be used to get estimates of pairing, it is necessary to use explicit solvent models and molecular simulations or integral equation approaches to obtain relatively precise free energies. Early attempts^{28,29} using free energy perturbation theory confirmed a preference for base stacking in water, but led to weak, or even positive, pairing free energies in disagreement with the experimental results cited above.^{14–16} The latter of these two studies²⁹ involved physically separating a base pair in water, but due to limitations in computer resources cannot be considered to have converged. Since these early calculations, improvements in computer performance, simulation algorithms, and force fields have enabled more refined simulations to be carried out and, in a number of cases, the resulting free energy estimates are in good agreement with experiment (see, for example, refs 30–32). We have, therefore, decided to revisit the problem of pairing energies in aqueous solution in the hope of providing numbers which will serve for the parametrization of simpler theoretical approaches and will also help in understanding the role of water structure on base pairing interactions. We present results on both conventional (Watson–Crick) AT and GC pairs and, in the case of AT, we also consider reverse Watson–Crick interactions (Figure 1). The calculations are carried out by thermodynamic integration, which is a robust technique for determining the energy change directly from an ensemble average.³³ This approach enables the free energy change between two limiting states to be obtained by numerical integration with respect to an appropriate parameter that, in our case, controls either base separation or base rotation. We have been particularly careful to verify that the results obtained are stable with respect to both an increase in the integration time and the nonbonded cutoffs used.

Methodology

We have carried out potential of mean force calculations using the molecular simulation package AMBER 4.1.³⁴ The geometries of the bases were taken from the AMBER library. United atom methyl groups (with Lennard-Jones parameters, $\epsilon = 0.181$ kcal/mol, $R^* = 2.165$ Å) were used at the glycosidic sites (N9 of the purines, guanine and adenine, and N1 of the pyrimidines, cytosine and thymine) and a similar group was also used for the C5 methyl of thymine. The Parm 94 parameter set,³⁵ which has now been well validated for nucleic acids

(24) Rief, M.; Clausen-Schaumann, H.; Gaub, H. E. *Nature Struct. Biol.* **1999**, *6*, 346.

(25) Boland, T.; Ratner, B. D. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 5297.

(26) Pincet, F.; Rawicz, W.; Perez, E.; Lebeau, L.; Mioskowski, C.; Evans, E. *Phys. Rev. Lett.* **1997**, *79*, 1949.

(27) Friedman, R. A.; Honig, B. *Biopolymers* **1992**, *32*, 145.

(28) Cieplak, P.; Kollman, P. A. *J. Am. Chem. Soc.* **1988**, *110*, 3734.

(29) Dang, L. X.; Kollman, P. A. *J. Am. Chem. Soc.* **1990**, *112*, 503.

(30) Sun, Y.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 3599.

(31) Carlson, H. A.; Nguyen, T. B.; Orozco, M.; Jorgensen, W. L. *J. Comput. Chem.* **1993**, *14*, 1240.

(32) Chipot, C.; Jaffe, R.; Maignret, B.; Pearlman, D. A.; Kollman, P. A. *J. Am. Chem. Soc.* **1996**, *118*, 11217.

(33) Chipot, C.; Kollman, P. A.; Pearlman, D. A. *J. Comput. Chem.* **1996**, *17*, 1012.

(34) Pearlman, D. A.; Case, D. A.; Caldwell, J. C.; Ross, W. S.; Cheatham, T. C.; Seibel, G.; Singh, U. C.; Weiner, P.; Kollman, P. A. *AMBER 4.1*, University of California, San Francisco: San Francisco, CA, 1994.

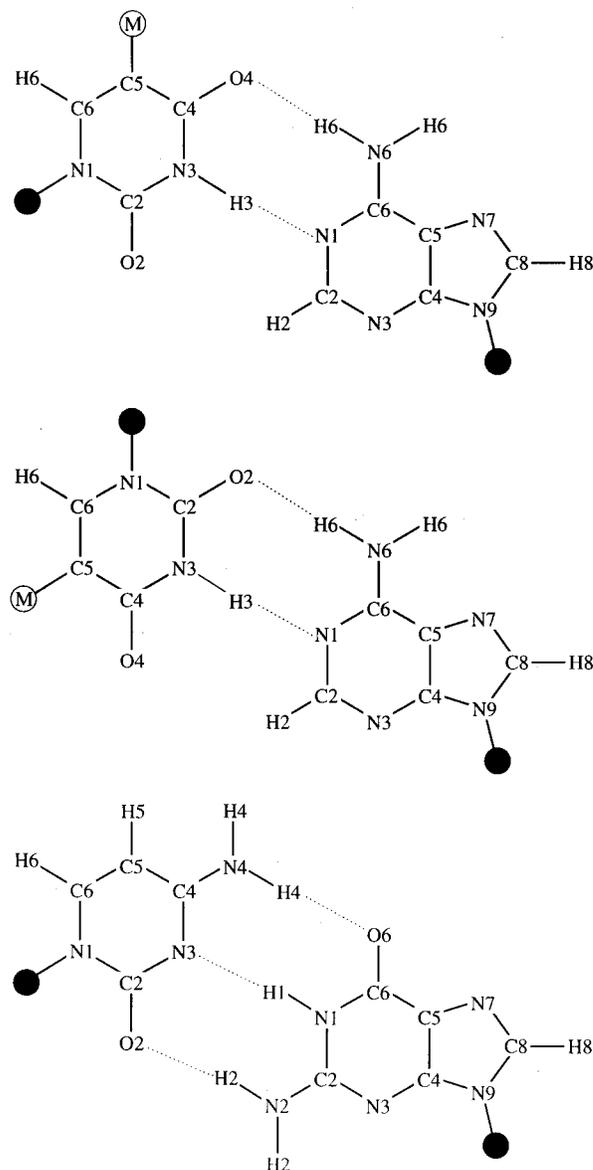


Figure 1. Schematic view of the base pairs studied: the Watson–Crick AT pair (top), the reversed Watson–Crick AT pair (center), and the Watson–Crick GC pair (bottom). The Black circles represent the united atom methyl groups bound to the glycosidic atoms of the bases. M represents the united atom C5 methyl group of thymine.

(see, for example, refs 36–39) was used for all calculations. The atomic charges of the bases were those of the AMBER library,⁴⁰ with a charge on the united atom glycosidic methyl group adjusted for overall neutrality.

To avoid the bases moving away from Watson–Crick pairing during the simulation, we maintained four atoms aligned R(C4)–R(N1)–Y(N3)–Y(C6), the first two belonging to the purine (R) and the latter two to the pyrimidine (Y). This involves using SHAKE⁴¹ to constrain

(35) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M., Jr.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. C.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179.

(36) Hobza, P.; Kabelac, M.; Šponer, J.; Mejzlik, P.; Vondrasek, J. *J. Comput. Chem.* **1997**, *18*, 1136.

(37) Flatters, D.; Zakrzewska, K.; Lavery, R. *J. Comput. Chem.* **1997**, *18*, 1043.

(38) Duan, Y.; Wilkosz, P.; Crowley, M.; Rosenberg, J. M. *J. Mol. Biol.* **1997**, *272*, 553.

(39) Young, M. A.; Beveridge, D. A. *Biophys. J.* **1997**, *73*, 2313.

(40) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Kollman, P. A. *J. Am. Chem. Soc.* **1993**, *115*, 9620.

(41) Ryckaert, J.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327.

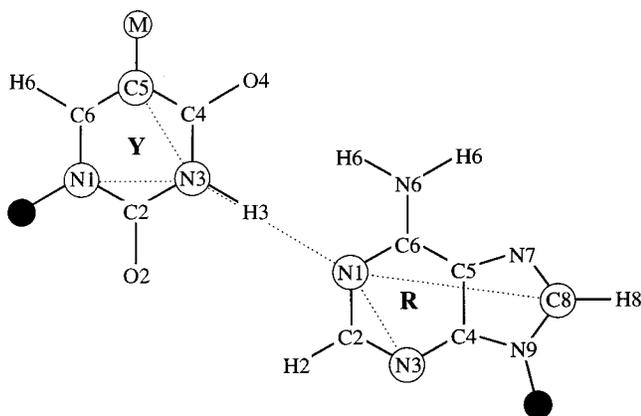


Figure 2. Angle restraints used to maintain the alignment of the base pairs: $R(N1)-Y(N3)-Y(N1) = 149^\circ/147^\circ$ (AT/GC), $R(N1)-Y(N3)-Y(C5) = 150^\circ/153^\circ$ (AT/GC), $R(N3)-R(N1)-Y(N3) = 150^\circ/149^\circ$ (AT/GC), $R(C8)-R(N1)-Y(N3) = 158^\circ/158^\circ$ (AT/GC).

Table 1. Potential of Mean Force Calculations of the Separation of the Adenine–Thymine Pair^a

	run no.						
	1	2	3	4	5	6	7
N_{H_2O}	581	581	581	815	815	473	473
R_X	28.2	28.2	28.2	33.6	33.6	28.6	28.6
R_{YZ}	25.2	25.2	25.2	27.1	27.1	22.5	22.5
C_1	9.	9.	9.	11.	11.	9.	9.
C_2	10.	10.	10.	13.	13.	11.	11.
rotation	free	free	free	free	free	restrained	restrained
T_{win}	20	40	40	20	40	20	40
T_{tot}	1.3	2.6	2.6	1.0	1.6	1.2	2.4
R_0	9.1	9.1	2.9	7.9	6.8	2.9	2.9
R_f	2.7	2.7	9.3	2.9	2.8	8.7	8.9
M_1	4.5	4.4	4.3	4.4	4.1	4.4	4.4
B	1.3	1.5	1.5	1.6	1.7	0.9	1.7
M_2	0.8	0.8	0.7	1.1	1.0	1.2	0.9

^a M_1 , M_2 , and B (kcal/mol) are respectively the energies of the principle minimum, the secondary minimum, and the intervening barrier, with respect to an energy zero defined for the separated bases. N_{H_2O} is the number of solvating water molecules; R_X , R_{YZ} are the box dimensions (Å); C_1 , C_2 are the primary and secondary cutoff distances (Å); Rotation: treatment of overall rotation of the base pair. T_{win} is the sampling time per window (ps). T_{tot} is the total time for trajectory (ns). R_0 , R_f are the initial and final $R(N1)-Y(N3)$ distances.

four angles (Figure 2). This technique, an extension of that used earlier by Pearlman,⁴² nevertheless, allows the distance between the bases and their propeller angle to vary freely. Without such constraints the base pairs rapidly move toward stacked conformations by buckling and sliding.

The bases were solvated with a box of TIP3P⁴³ water molecules and periodic boundary conditions were used. For each simulation, a residue-based cutoff and a secondary cutoff (energies and forces updated at the same time as the pair list) were applied to truncate the solute–solvent and solvent–solvent interactions. No solute–solvent cutoffs were used, but interactions of the solute with its periodic images were not taken into account (see Tables 1 and 2 for details).

Two approaches were used concerning the rotational orientation of the base pair: (a) The orientation was left free, in which case it was necessary to guarantee that, as the bases were separated, a thick enough layer of solvent continued to separate them from their images. In the cases presented here, this thickness was always at least, 8 Å. (b) The principal axis of the base pair, characterized by the atoms $R(C4)-Y(C6)$, was restrained to lie along the principal axis of the rectangular solvent box by using $S_1-P_1-R(C4) = T_1-P_1-R(C4) = S_2-P_2-Y(C6) = T_2-P_2-Y(C6) = 90^\circ$, where P_i ($i = 1, 2$) are reference points on

Table 2. Potential of Mean Force Calculations of the Separation of the Guanine–Cytosine Pair^a

	run no.				
	1	2	3	4	5
N_{H_2O}	581	581	484	484	484
R_X	29.0	29.0	30.8	30.8	30.8
R_{YZ}	25.1	25.1	22.1	22.1	22.1
C_1	9.	9.	9.	9.	9.
C_2	10.	10.	10.	10.	10.
rotation	free	free	restrained	restrained	restrained
T_{win}	40	60	20	40	60
T_{tot}	2.2	3.9	1.0	2.0	3.9
R_0	2.8	2.8	2.9	2.9	2.9
R_f	8.3	9.3	7.8	7.8	9.3
M_1	5.8	6.1	6.4	5.1	5.7
B	1.5	1.6	0.7	1.6	1.6
M_2	1.2	1.2	1.6	1.0	1.1

^a M_1 , M_2 , and B (kcal/mol) are respectively the energies of the principle minimum, the secondary minimum, and the intervening barrier, with respect to an energy zero defined for the separated bases. Parameters as defined in Table 1.

the principal axis of the simulation box, 3 Å from the end planes of the box, and S_i and T_i define vectors orthogonal to the principal axis, lying in the planes defined by P_i . Points P_i , S_i , and T_i were given unit masses and restrained at their positions to a precision of 0.1 Å by using appropriate force constants. Although the second approach allows the use of an elongated, rectangular box with a smaller cross-section, and thus fewer solvent molecules, the time gained is partially offset by the cost of satisfying the supplementary constraints.

All the chemical bonds were constrained to their equilibrium values by means of the SHAKE⁴¹ procedure. Simulations were carried out in the (N, P, T) ensemble, at 300 K and 1 atm, using a 2 fs time step. The temperature was maintained using the Berendsen algorithm,⁴⁴ with a coupling constant to the external bath of 0.4 ps. When the rotation of the base pair was restrained this value was reduced to 0.1 ps.

We have calculated the free energy differences related to two different processes using the thermodynamic integration approach:^{33,45}

$$\Delta G = \int_0^1 \left\langle \frac{\partial H(r; \lambda)}{\partial \lambda} \right\rangle_\lambda d\lambda$$

where $H(r; \lambda)$ is the potential energy function describing the system.

(1) Separating the two bases along their principal axis: Separating the two bases was accomplished by either increasing the $R(N1)-Y(N3)$ distance in steps of 0.1 Å or decreasing the $P_1-R(C4)$ and $Y(C6)-P_2$ distances in 0.05 Å steps, as a function of the “coupling” parameter λ (between the limits indicated in Tables 1 and 2), leaving the propeller angle between the bases free to vary. It is remarked that propeller rotation was apparently not allowed in the earlier work of Dang and Kollman.²⁹

(2) Scanning the propeller angle between the two bases: Scanning the propeller angle was accomplished by increasing the dihedral angle $R(C5)-R(N1)-Y(N3)-Y(C5)$ in steps of either 2.5° or 5°. During this procedure, the distance separating the bases was free to vary.

It is remarked that using the points P_i to pull the bases apart does not significantly change the fluctuation of the distance between the bases at a given separation. As an example, the mean fluctuation of the distance $A(N1)-T(N3)$ close to the optimal separation of the base pair was 0.06 Å when using P_i restraints and 0.1 Å otherwise. It should also be recalled that using the potential force method^{33,42,45} to link λ to the constrained variable allows a fast and accurate determination of the holonomic constraint contribution to the free energy.^{33,45}

Prior to each of the simulation runs listed in Tables 1 and 2, the system was equilibrated for 100 ps. Various sampling times per window were tested for their effect on the final free energy changes for base

(42) Pearlman, D. A. *J. Chem. Phys.* **1993**, *98*, 8946.

(43) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926.

(44) Berendsen, H. J. C.; Postma, J. P.; van Gunsteren, W. F.; Di Nola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684.

(45) Mezei, M.; Beveridge, D. L. *Ann. N.Y. Acad. Sci.* **1986**, *482*, 1.

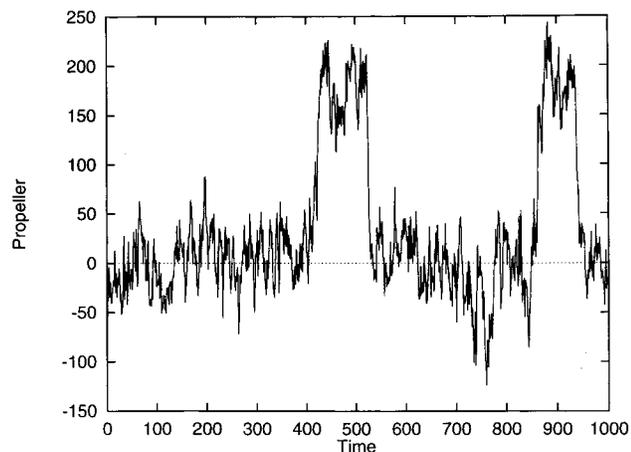


Figure 3. Propeller twist (ϕ) fluctuations (deg) of the AT pair as a function of time (ps).

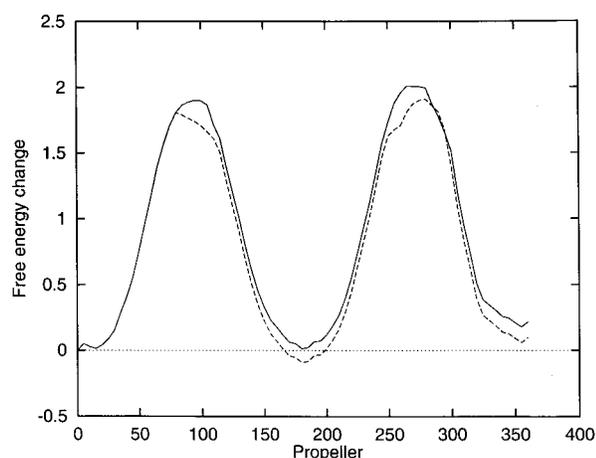


Figure 4. Potential of mean force (kcal/mol) as a function of the AT propeller twist (deg). Dotted line: 60 ps sampling (including 10 ps equilibration) per 5° window. Solid line: 80 ps sampling (including 10 ps equilibration) per 2.5° window in the regions around the free energy maxima ($80\text{--}110^\circ$ and $250\text{--}300^\circ$).

separation. For the propeller rotations, 60 or 80 ps sampling times, including 10 ps of equilibration, were used.

Results

(a) AT Base Pair Inversion. The fact that the normal hydrogen bonding face of thymine is symmetric with respect to the axis of the AT base pair means that a reversed Watson–Crick state containing an A(HN6)–T(O2) hydrogen bond can be generated by a 180° rotation of one of the bases about this axis (see Figure 1, center). During our molecular dynamics simulations, such inversions occur spontaneously, with a characteristic time for the transition of roughly 10 ps (Figure 3). This occurs even at the optimal separation between the two bases, A(N1)–T(N3) = 3 Å. One can attempt to calculate the free energy change for this process, and, thus, the difference between Watson–Crick and reversed Watson–Crick AT pairs, by performing a controlled rotation around the A(N1)–T(N3) axis of the AT base pair, as described in the methodology. The results are shown in Figure 4. The initial free energy curves obtained in this way showed some irregularity at the barriers between the normal and reversed states. This can be ascribed to statistical inaccuracies connected to insufficient sampling in the regions with higher conformational freedom. This was confirmed by longer sampling in the ranges $\phi = 80\text{--}110^\circ$ and

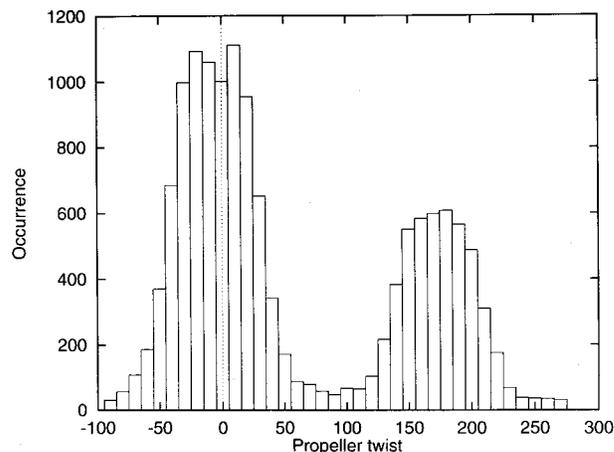


Figure 5. Histogram showing the propensity for various AT propeller twists (deg). Values around 0° correspond to Watson–Crick pairing and values around 180° to reverse Watson–Crick pairing. Note that the Watson–Crick form shows two maxima on either side of $\phi = 0^\circ$ corresponding to a preference for slight propeller twisting.

$250\text{--}300^\circ$. It should be noted that, after a complete rotation of the bases, these curves close to a precision of better than 0.2 kcal/mol, which is a good test of their quality. These results lead to three conclusions. First, the Watson–Crick and the reversed Watson–Crick forms of the AT pair have very similar stabilities. Second, the barrier for base rotation is roughly 1.9–2.0 kcal/mol. Third, the optimal conformation of the Watson–Crick form shows a preference for propeller twisting, with an angle of roughly 15° between the base planes, whereas this is not seen for the reversed state.

To obtain a more refined estimate of the relative stabilities of the two AT pairing configurations, we have used a single, long simulation of 14 ns, which leads to repeated sampling of both states. This simulation was carried out with 592 water molecules and with primary and secondary cutoffs of 9 and 10 Å, respectively. Neither the separation nor the overall rotation of the pair were restrained. A histogram of the resulting propeller twist angles is given in Figure 5. These results confirm that the Watson–Crick state is characterized by a small propeller twisting ($\pm 15^\circ$) not seen in the reversed state. On the basis of the mean residence times in each state (Watson–Crick: 400 ps, reversed Watson–Crick: 220 ps), it is possible to calculate that there is a free energy difference of roughly $0.6 \ln(220/400) = -0.36$ kcal/mol, in favor of the Watson–Crick state.

(b) Base Pairing Energies. The free energy curves for base separation are shown in Figures 6 and 7. They are all adjusted to zero for the maximal separation between the bases studied (see Tables 1 and 2). Both base pairs present the same general behavior with a primary minimum M_1 at the hydrogen bonding distance R_1 , corresponding to the free energy of base pairing, an energy barrier B , a secondary minimum M_2 at R_2 , where a water molecule forms a bridged hydrogen bond between A(N1) and T(HN3) or between G(HN1) and C(N3), and a final energy barrier. Note that all energies are given with respect to the zero point for separated bases.

For the specific case of the AT pair (Figure 6 and Table 1), the primary minimum occurs at $R_1 = 3$ Å and the water-bridged minimum at roughly $R_2 = 5.6$ Å. As noted above, the Watson–Crick and reversed Watson–Crick forms of this base pair exchange regularly along the separation pathway. The results given in the table do not depend significantly on either the sampling time per window or the cutoff values, suggesting that the simulations can be considered to have converged. We

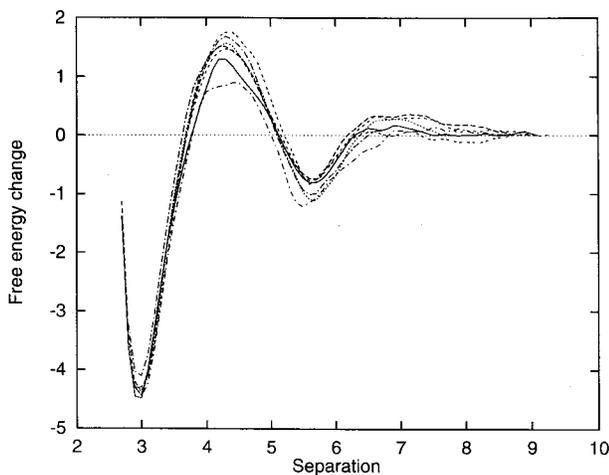


Figure 6. Free energy profiles (kcal/mol) for the AT pair as a function of the separation distance (Å) measured between A(N1) and T(N3). (See Table 1 for simulation protocols.)

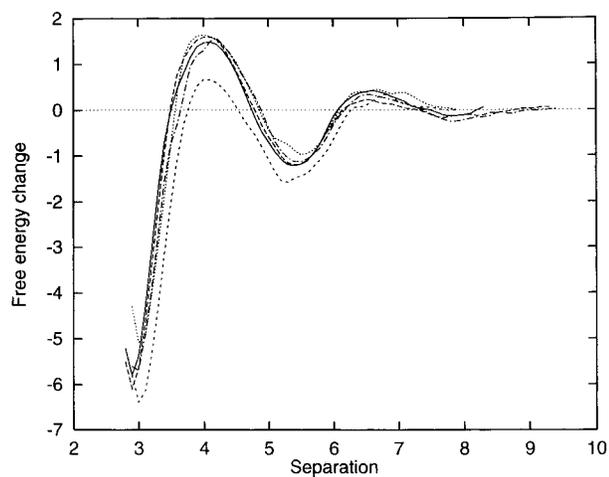


Figure 7. Free energy profiles (kcal/mol) for the GC pair as a function of the separation distance (Å) measured between G(N1) and C(N3). (See Table 2 for simulation protocols.)

consequently find a pairing free energy (M_1) of -4.3 ± 0.2 kcal/mol and a secondary minimum (M_2) of -0.9 ± 0.2 kcal/mol separated by a barrier (B) of 1.5 ± 0.3 kcal/mol. It is remarked that the error estimates are simply standard deviations calculated over the seven simulations presented in Table 1.

For the GC pair (Figure 7 and Table 2), the values are, as expected, considerably higher with a primary minimum, $M_1 = -5.8 \pm 0.5$ kcal/mol, a secondary minimum $M_2 = -1.2 \pm 0.3$ kcal/mol, and a barrier $B = 1.4 \pm 0.4$ kcal/mol. Error estimates are again standard deviations over the simulations carried out. Note that both the primary and the secondary minima occur at slightly shorter distances than for the AT pair, with $R_1 = 2.9$ Å and $R_2 = 5.4$ Å. Longer simulation times were necessary for the more strongly interacting GC pair. Although no transition to a reversed Watson–Crick form can occur for the GC pair, a strongly propeller twisted state with a dihedral angle between the base planes of roughly 100° was seen occasionally for short distances (3.0–3.8 Å) when a water molecule moved to form a bridged G(O6)–C(HN4) interaction.

To illustrate how a secondary minimum is formed between the paired bases in water, Figure 8 shows the evolving profile of water molecules around the bases of the Watson–Crick AT pair. This view, situated between the bases and perpendicular

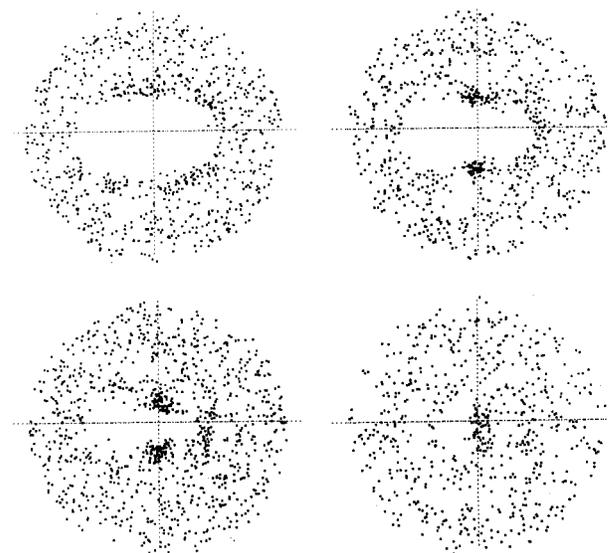


Figure 8. Water distribution in a plane midway between adenine and thymine, perpendicular to the principal axis of the base pair. Each point indicates the oxygen position of a water molecule. The images correspond to increasing distances of separation between the bases: 3 (top left), 4 (top right), 5 (bottom left), and 5.6 Å (bottom right). Only states with propeller twists less than $\pm 20^\circ$ were included in the sampling which was carried out every 0.4 ps.

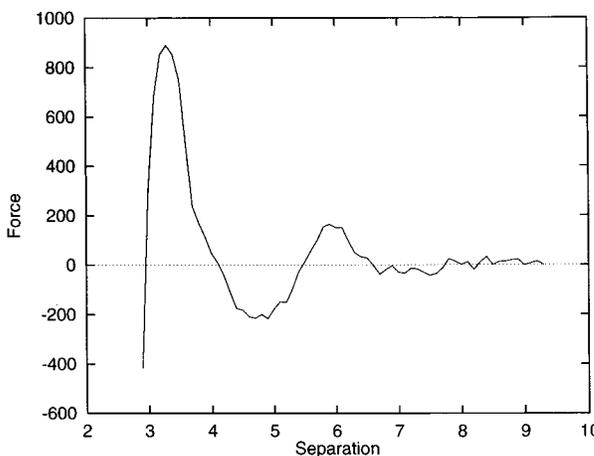
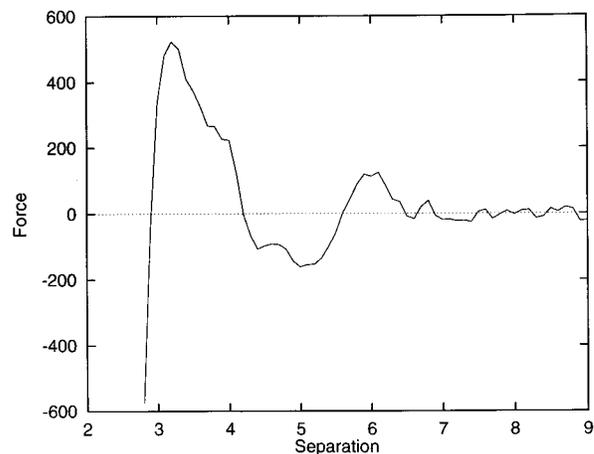


Figure 9. Force profiles (pN) derived from the free energy curves as a function of the separation distance $R(N1)$ – $Y(N3)$ (Å): (a) AT pair and (b) GC pair.

to the principal axis of the pair, contains one point for each water molecule position sampled every 0.4 ps. To obtain a clear

view, sampling was limited to states with small propeller twist angles ($\pm 20^\circ$). The results show that, as the base pair begins to separate, two favorable regions for water molecules appear above and below the space between the A(N1) and T(HN3) atoms, which constitute the axial hydrogen bond of the pair. This corresponds to water molecules beginning to bridge the increasing distance between these atoms and stabilize the weakening base–base bond. Once the distance between the bases is large enough to accommodate a water molecule (roughly, $R_1 + 2.5 \text{ \AA}$), the out-of-plane water clusters coalesce at a point on the base pair axis and a linear, bridged hydrogen bond is formed. Similar results were obtained in the case of the GC pair.

(c) Force Curves. Single molecule experiments of the type cited in the Introduction are now giving access to the forces associated with the rupture of base pairs. Converting the measured forces to free energies of binding is, however, complicated by the dependence of the measurements on loading rate.⁴⁶ To generate a set of reference values for the standard base pairs, we have numerically differentiated the free energy curves obtained from our simulations. The results for representative runs on the AT and GC pairs are shown in Figure 9. The maximal rupture force of roughly $550 \pm 35 \text{ pN}$ for AT and $860 \pm 40 \text{ pN}$ for GC occurs as the pair confronts the barrier between the primary and the secondary minima at a distance of around 3.3 \AA . It should be noted that these values are considerably higher than estimates based on a linear free energy change over an assumed rupture distance, namely, 125 pN (for an AT pairing energy of 3.5 kcal/mol and a rupture distance of 2 \AA).²⁵ It should, nevertheless, be recalled that our simulations apply to the longitudinal separation of Watson–Crick paired bases, maintaining a common base axis and prohibiting buckling or stacking of the base pairs. Recent micromanipulation experiments on polymeric DNA's also lead to much smaller forces of around $9\text{--}10 \text{ pN}$ for AT and $15\text{--}20 \text{ pN}$ for GC,^{23,24} but in these cases, the elasticity of the DNA strands surrounding the stretched base pair is expected to strongly diminish the maximal force measured.⁴⁷

Conclusions

Molecular dynamics simulations of adenine–thymine and guanine–cytosine base pairs in water have enabled us to calculate the Watson–Crick pairing free energies as -4.3 and -5.8 kcal/mol , respectively. These values are larger than those found in early free energy simulations,^{28,29} or in calculations based on a Langevin dipole model of the solvent.⁴⁸ They are, however, compatible with recent experimental studies which put the free energy increment at roughly 2 kcal/mol per hydrogen bond.^{14–16} A similar value has been proposed by Ahora and Jayaram using continuum solvent models and various molecular mechanics force fields.⁴⁹

As the base pairs are longitudinally separated, both exhibit a secondary minimum where a water molecule bridges the gap between the axial R(N3) and Y(N1) atoms. These minima are stabilized by -0.9 and -1.2 kcal/mol , respectively. In the case of the AT pair, regular exchanges occur between Watson–Crick and reversed Watson–Crick states, although the former is slightly more stable (-0.4 kcal/mol).

We have not investigated the free energy of base stacking in this study, but it was clear from unconstrained dynamic runs on base pairs that deformation toward a stacked conformation occurred rapidly if the two bases were not appropriately constrained. While it is not easy to judge which base stacking interactions would be likely to form from a mixture of A/T or G/C bases, the large experimental values seen for R–R stacks would be competitive with our pairing free energies.^{18,19}

A differentiation of the free energy curves leads to maximal forces for base pair rupture of 550 pN for AT and 860 pN for GC. These forces occur at a separation of roughly 3.3 \AA and correspond to the energy barrier between the directly hydrogen bonded and the water separated minima of the base pairs.

JA991092Z

(46) Evans, E.; Ritchie, K. *Biophys. J.* **1997**, *72*, 1541.

(47) Viovy, J.-L.; Heller, C.; Caron, F.; Cluzel, P.; Chatenay, D. C. R. *Acad. Sci. Paris, Life Sci.* **1994**, *317*, 795.

(48) Florian, J.; Sponer, J.; Warshel, A. *J. Phys. Chem. B* **1999**, *103*, 884.

(49) Ahora, N.; Jayaram, B. *J. Phys. Chem. B* **1998**, *102*, 6139.