ACS PHYSICAL CHEMISTRY Au
AN OPEN ACCESS JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

pubs.acs.org/physchemau
Article

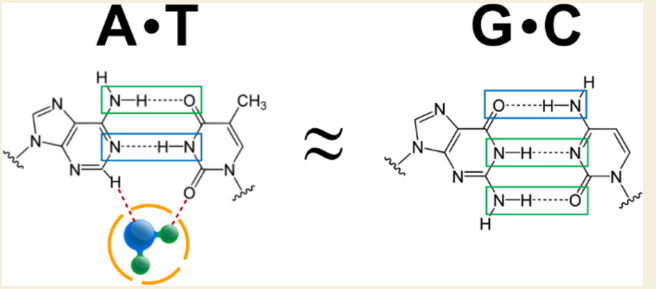
Hydration Waters Make Up for the Missing Third Hydrogen Bond in the A·T Base Pair

Chi H. Mak*

Cite This: ACS Phys. Chem Au 2024, 4, 180−190
Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Base pairing complementarity is central to DNA function. G·C and A·T pair specificity is thought to originate from the different number of hydrogen bonds the pairs make. Quantifying how many hydrogen bonds exist can be difficult because water molecules in the surrounding can make up for or disrupt direct hydrogen bonds, and the hydration structures around A·T and G·C pairs on duplex DNA are distinct. Large-scale computer simulations have been used here to create a detailed map for the hydration structure on A·T and G·C base pairs in water. The contributions of specific hydration waters to the free energy of each of the hydrogen bonds in the A·T and G·C pairs were computed. Using the equilibrium fractions of hydrated versus unhydrated states from the hydration profiles, the impact of specific bound waters on each hydrogen bond can be uniquely quantified using a thermodynamic construction. The findings suggest that hydration water in the minor groove of an A·T pair can provide up to about 2 kcal/mol of free energy advantage, effectively making up for the missing third hydrogen bond in the A·T pair compared to G·C, rendering the intrinsic thermodynamic stability of the A·T pair almost synonymous with G·C.



KEYWORDS: *DNA base pairs, hydration spine, hydrogen bonds, base pair complementarity, free energy*

INTRODUCTION

DNA base pair complementarity is the cornerstone of the molecular basis of life. The specificity with which adenine (A) pairs with thymine (T) and guanine (G) with cytosine (C) allows genetic information to be reliably encoded into the genome. At the same time, base pair complementarity enables DNA to be expressed and replicated by the cell virtually error-free. On a DNA double helix, a G·C base pair can be substituted for an A·T pair and vice versa without significantly altering the overall structure of the duplex. While there are three hydrogen bonds in the G·C pair but only two in A·T, the geometry of a Watson−Crick base pair step inside a duplex B-DNA is nearly geometrically isomorphic between the two canonical base pairs.[1−4]

The different number of hydrogen bonds that can be made by a G·C pair compared to A·T offers a simple rationale for base complementarity. The hydrogen bonding donor N6 and the acceptor N1 atoms on A complement the acceptor O4 and donor N3 atoms on T, whereas the donor O6 and acceptors N1 and N2 atoms on G match the acceptor N4 and donors N3 and O2 atoms on C. However, the precise number of hydrogen bonds that are present between two bases in aqueous solution is difficult to quantify because water molecules can accept and donate hydrogen bonds also. The free energy of formation of a base pair can offer a quantitative measure. In vacuum, the energy of an isolated G·C pair is indeed more stable than A·T by an approximately 3:2 ratio.[5,6] However, solution DNA melting studies have shown that the free energy of a G·C pair in water is no more than ∼1 kcal/mol stronger than an A·T pair,[5,7−10] suggesting that solvent effects are substantial; however, these measured duplex melting free energies also contain contributions from base stacking and backbone conformational entropy. Direct measurement of association free energy of the bases in water is also difficult due to their preference toward forming stacks rather than pairs.[11] Existing evidence suggests that base pair formation in aqueous solution appears to be governed by base−base hydrogen bonds that are only slightly more favorable than base−water hydrogen bonds.[3,12] Because of this, alternatives to the assumption that free energy difference of the base pairs is the molecular basis for complementarity have been proposed. For example, shape, geometric fit, and hydration have been suggested as possible alternative reasons for base pair complementarity.[12,13] However, regardless of whether the decisive driving force for base pair complementarity comes from one or more of these factors, the role of water molecules in

Received: November 13, 2023
Revised: January 5, 2024
Accepted: January 8, 2024
Published: January 24, 2024

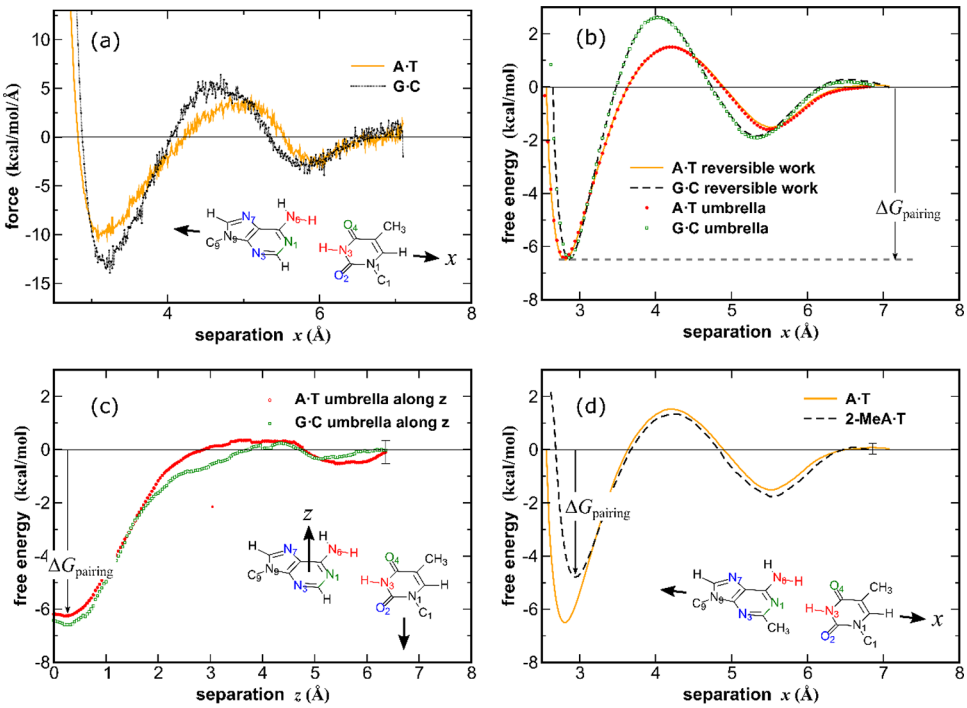


Figure 1. (a) Equilibrium forces as a function of base pair separation along the base pair plane for A·T (orange) and G·C (black) in water. (b) Base pairing free energy as a function of base separation computed by integrating the equilibrium forces for A·T (orange line) and G·C (black dashed line) from panel (a), compared to results from direct umbrella sampling for A·T (red circles) and G·C (open green squares). The free energy of each pair, $\Delta G_{pairing}$, is indicated by the downward arrow and is similar between A·T and G·C. Error bars are smaller than the plotting symbols. (c) Base pairing free energies of A·T (red circles) and G·C (green squares) computed by umbrella sampling by separating the bases in the vertical direction. Error bar shows cumulative error in the estimated pairing free energy. (d) Base pairing free energy for 2-methyl-adenine opposite to thymine (dashed line) compared to A·T (solid line). The 2-MeA·T pair is approximately 1.5 kcal/mol less stable. Error bar shows cumulative error in the estimated pairing free energy for the 2-MeA·T pair.

the aqueous solution is the key since an A·T pair can be hydrated by a different number of water molecules than G·C and this will not only affect their free energies but also how the bases can fit with each other.

The complexity of the microenvironment of water molecules around a DNA has been demonstrated by the classic crystallographic studies of Drew and Dickerson[14,15] in which X-ray structures of duplex B-DNA revealed a number of bound water molecules on the minor groove. These waters form a so-called "hydration spine" that is prominent along AT tracks on B-DNAs but absent or largely muted along GC tracks. Many subsequent experiments, including recent X-ray data,[16−19] NMR,[20−23] chiral nonlinear spectroscopy,[24] ultrafast dynamics,[25] and also simulations and theory[26−32] have confirmed the presence of these hydration spine waters. However, whereas these studies have demonstrated that the hydration structure around A·T pairs is distinct from G·C, the focus has been primarily on the structure of the spine instead of on the thermodynamics. What remains unaddressed is whether these hydration waters can impact base pair strengths and how. While it is known that hydration can alter the relative stability of B-, A-, and Z-form DNA,[33] the contribution of hydration to the free energy of individual DNA base pairs is unclear. Several hypotheses have been advanced to explain the origin of the difference between the hydration structure around A·T pairs versus G·C. These include: possible differences between A·T and G·C in their minor groove geometries[14] and their widths,[32] difference in the twist propensities of A·T versus G·C pairs,[32] the presence or absence of amino and carbonyl groups on the minor groove side of the bases,[34−36] and the possible role of counterions.[19,37] Despite all of these possibilities, the characteristics of the hydration structure around A·T pairs appear to be robust and nonsequence-specific.[23,38]

The goal of this paper is to quantify the role of hydration water molecules on the thermodynamic stability of canonical base pairs using computer simulations. Simulations provide reliable free energy estimates and can elucidate microscopic details of the hydration structure that are needed to address the questions that experiments are not yet able to answer. By a detailed analysis of the hydration structure around the A·T and G·C base pairs, the impact of bound waters on each hydrogen bond can be quantified via a simple thermodynamic formula. The paper is divided into three parts. The first part provides rigorous free energy estimates for an isolated A·T base pair and an isolated G·C pair in water. The second part examines the hydration structure around each base pair and compares it to the structure of waters around a doublet base pair and to that around a DNA duplex, to verify that the key water molecules that are involved in the complementarity in single base pairs are also the ones that make up the hydration spine in the X-ray structures of B-DNA helices. Finally, a simple thermodynamic construction is employed to quantify the contribution of specific water molecules in the hydration shell around a canonical base pair on its free energy. The results suggest that hydration water in the minor groove of an A·T pair can provide up to ∼2 kcal/mol of free energy advantage, effectively making up for the missing third hydrogen bond in the A·T pair compared to G·C.

https://doi.org/10.1021/acsphyschemau.3c00058
ACS Phys. Chem Au 2024, 4, 180−190

METHODS

All-atom Monte Carlo (MC) simulations in the constant-temperature constant-pressure ensemble were carried out using Amber ff99[39] potentials. Three sets of simulations were performed. The first involved a single base pair solvated in water to compute the unpairing free energy using umbrella sampling[40] by sliding one base out from the other with both remaining coplanar throughout. The second involved a doublet base pair in water, in which two base pairs were stacked on top of each other (according to their standard geometry inside a B-DNA), to calculate the free energy of the unpairing of one base pair by sliding one base out from the doublet. The third was a simulation of the Drew–Dickerson dodecamer d(CGCGAATTCGCG) B-DNA in water, for the analysis of the hydration structure around the helix. Relative to recent Amber force fields developed specifically for nucleic acid simulations such as BSC,[41] which have updated bond parameters for the nucleic acid backbone, all the nonbond parameters that define the base–base and base–water interactions have remained the same since the original parameter set in ff94,[42,43] and ff99 was used in these simulations so the new data could be cross-referenced to earlier studies.[43−47] All simulations were performed in a periodic box of either TIP3P or SPC/E waters. The majority of the simulations were carried out in TIP3P waters, with additional simulations in SPC/E waters to verify that the general validity of the results was solvent-model independent. In the Amber force field, base pair interactions and hydrogen bond are described by nonbonded electrostatic and dispersion terms in the potential. Electrostatic interactions were calculated in the MC simulations using an eighth order fast multipole method (FMM) with Ewald summation.[48] To maintain electrical neutrality in the base pair simulations, excess charges on each base was redistributed to the N9 atom on purines or the N1 atom on pyrimidines. For duplex DNA simulations, the excess −1 charge on each nucleotide was redistributed evenly throughout all the atoms in that monomer to avoid counterions,[26] resulting in a charge renormalization of no more than 0.03 on each atom. Using MC permitted straightforward umbrella sampling using hard walls, eliminating the restriction inherent in molecular dynamics simulations requiring strictly differentiable potentials.

RESULTS AND DISCUSSION

Free Energies of A·T and G·C Pairs in Water

A number of simulations were performed to compute the strength of the complementarity interactions in the two canonical singlet pairs, A·T and G·C and how the surrounding hydration structure affects them. Figure 1a shows the equilibrium force needed to pull apart a single A·T pair (orange line) and a G·C pair (black dotted line) in the direction x along the plane of the base pair. Starting with two infinitely separated bases ($x = \infty$) and integrating the force in the $-x$ direction then yielded the reversible work, which is the free energy of the base pair $G(x)$ as a function of their separation. Free energies derived from this reversible work are plotted in Figure 1b in the orange solid line for A·T and the black dashed line for G·C. These show that for both A·T and G·C, the equilibrium base pair distance is $x^* \sim 2.8$ Å (measuring between heavy atoms on the hydrogen bond donor and acceptor pair), corresponding to the distance where the force in Figure 1a is 0. The interaction free energy of the base pair $\Delta G_{pairing}$, which is the difference between $G(\infty)$ and $G(x^*)$, is shown in Figure 1b by the downward arrow. Somewhat surprisingly, $\Delta G_{pairing}$ for both A·T and G·C are almost identical at −6.4 kcal/mol even though G·C should have three hydrogen bonds and A·T only two. To check for internal consistency, these free energy curves derived from the equilibrium forces were compared to those from direct umbrella sampling, shown as red circles and green squares in Figure 1b. The resulting $G(x)$ from these two different methods were identical. In vacuum, the direct A·T and G·C pair energies would have been −14 and −23 kcal/mol; therefore, water molecules in the surroundings play a decisive role in modulating the free energies of the base pairs. These results are consistent with previous simulation studies.[43−47]

To verify the key results in Figure 1b demonstrating that A·T and G·C have almost identical base pairing free energies, additional simulations were performed where the same base pairs were unpaired in the z direction perpendicular to the base pair plane instead of the x direction. Because $\Delta G_{pairing}$ is path-independent, the computed free energy should be the same in both directions. Figure 1c shows that within the precision of the calculations, $\Delta G_{pairing}$ in the z direction are indeed identical to those in Figure 1b.

To further verify that the results were not an artifact of the solvent model, the base pair free energies were recomputed using a SPC/E model for water. The results are shown in Figure S1 in the Supporting Information compared to TIP3P. In SPC/E water, the base pair free energy was −6.0 kcal/mol for A·T and −6.4 kcal/mol for G·C, also approximately equal.

The geometries of the A·T and G·C pairs are shown in Figure 2. G·C has three hydrogen bonds between Gua:O6·Cyt:N4, Gua:N1·Cyt:N3, and Gua:N2·Cyt:O2. In the A·T pair, there are two hydrogen bonds between Ade:N6·Thy:O4 and Ade:N1·Thy:N3, but the "third" hydrogen bond is missing. This, however, does not mean that the C2 atom on adenine and O2 on thymine have a void between them. This space where the missing third hydrogen bond would have been can be filled by one or more water molecules from the solvent.

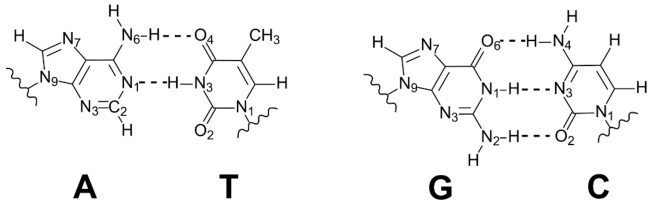


Figure 2. Expected hydrogen bonds in A·T and G·C base pairs.

To estimate the impact of this missing hydrogen bond and the influence of the surrounding water molecules, the Ade:H2 atom on A was mutated to a CH_3- group and the free energy of a 2-methylated-A·T pair was computed using umbrella sampling again. This extra methyl group would have excluded any water molecules from between Ade:H2 and Thy:O2. The resulting free energy is shown in Figure 1d. In vacuum, the direct A·T pair interaction of −14 kcal/mol is destabilized to −10 kcal/mol in the 2-methyl-A·T pair, but in water, the solvent compensates for this destabilization by making the 2-methyl-A·T pair only ∼1.5 kcal/mol weaker than A·T. This suggests that water molecules in the surroundings indeed play an indispensable role in modulating the free energy of base pairs.

Key results from the simulations on the free energies of the A·T and G·C pairs in water are summarized in Table 1. The top half of Table 1 refers to the singlet base pairs. The bottom half refers to the doublet pairs (to be discussed in the next section). Each of the values in Table 1 labeled "solvent contrib to pairing" corresponds to the *difference* between $\Delta G_{pairing}$ in vacuum versus in solution.

In Figure 3, the free energy profiles in Figure 1b are compared to the in-vacuum free energy of the two single base pairs. The in-vacuum, or direct, hydrogen-bonding interactions between A·T

https://doi.org/10.1021/acsphyschemau.3c00058
ACS Phys. Chem Au 2024, 4, 180−190

Table 1. Summary of Calculated Base Pair Free Energies in Water[a]

summary of calculated free energy changes from simulations (kcal/mol)			
A·T		G·C	
pairing (along x, in TIP3P)	−6.38 ± 0.08	pairing (along x, in TIP3P)	−6.42 ± 0.08
pairing (along z, in TIP3P)	−6.26 ± 0.43	pairing (along z, in TIP3P)	−6.58 ± 0.51
solvent contrib to pairing	+7.03 ± 0.08	solvent contrib to pairing	+18.83 ± 0.08
pairing (in SPC/E)	−5.97 ± 0.65	pairing (in SPC/E)	−6.38 ± 0.53
2-MeA·T pairing (in TIP3P)	−4.79 ± 0.22		
A·T stacked upon A|T		G·C stacked upon G|C	
pairing (along x, in TIP3P)	−6.4 ± 0.8	pairing (along *x*, in TIP3P)	−6.4 ± 0.8
stacking (along z, in TIP3P)	−4.13 ± 0.82	stacking (along *z*, in TIP3P)	−4.59 ± 0.88

[a]The top half correspond to singlet pairs. The lower half corresponds to doublet pairs.

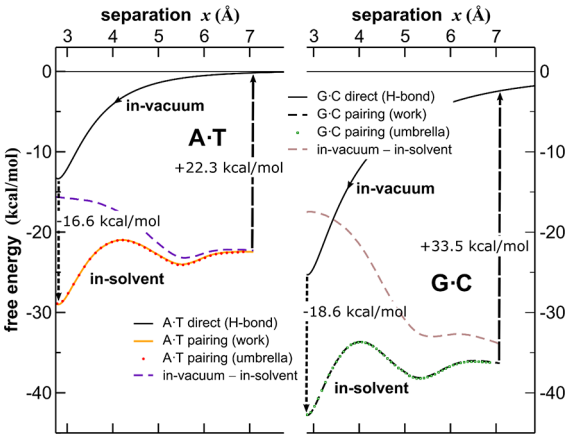


Figure 3. Free energy profiles for the A·T and G·C pair in water from Figure 1b, relative to their direct interactions in vacuum as a function of base separation. The solvent contribution, which is the difference between the in-solvent and in-vacuum free energies, is shown as the purple dashed line for A·T and the brown dashed line for G·C. The downward and upward dashed arrows indicate the contributions of hydrogen bonding from the solvent on the free energy of each pair, computed by thermodynamic integration.

and G·C are strictly monotonic as expected, but the base pair free energy in water is not, illustrating that the hydration structure surrounding the base pairs continuously adjusts itself depending on distance and how water molecules can fit between the bases. The differences between the in-vacuum and in-solution free energies are shown as the purple dashed line in Figure 3 for A·T and as the brown dashed line for G·C. The difference between in-vacuum and in-solvent free energies is much more pronounced for the G·C pair (brown dashed line) than A·T (purple dashed line), suggesting that hydration acts very differently toward the A·T pair compared to G·C.

To ascertain how much the hydrogen bonds offered by the solvent contribute to the pairing free energy, thermodynamic integration[49,50] was used to evaluate the free energy needed to turn off the electrostatics involving all the base atoms in the system reversibly in vacuum and in solution. While this eliminates all hydrogen-bonding interactions (base−water and base−base) as well as all long-range charge−charge interactions involving the bases, hydrogen bonds represent a large fraction of these interactions, and since the direct base−base hydrogen bond energies are known from the in-vacuum calculations, the base−water interactions can be determined easily. (The Supporting Information provides more details and the uncharging free energies are shown in Figure S2.) The magnitudes of the base−water interactions are illustrated by the vertical black arrows in Figure 3. On the left side of Figure 3, the downward black dashed arrow indicates that the solvent adds an additional 16.6 kcal/mol of free energy in terms of hydrogen bonds and electrostatic solvation to the A·T pair compare to vacuum, versus 22.3 kcal/mol for separate A and T indicated by the upward black dashed arrow. On the right, the downward black dashed arrow indicates that the solvent adds an additional 18.6 kcal/mol of free energy in terms of hydrogen bonds and electrostatic solvation to the G·C pair compare to vacuum, versus 33.5 kcal/mol for separate G and C. These results indicate that when fully hydrated, the solvent makes more hydrogen bonds to the separate bases compared to when the bases are paired, which is expected because there are more unsatisfied hydrogen-bonding donors and acceptors in the unpaired bases. However, the difference is significantly larger for the G·C pair than A·T. In fact, the net number of additional hydrogen bonds that are used to solvate the G·C pair and the A·T pair differ by only ∼2 kcal/mol (18.6 vs 16.6 kcal/mol), but the solvent provides many more hydrogen bonds to G and C separately compared to A and T.

Hydration Structure of Singlet and Doublet Pairs in Water

On a B-DNA duplex, the base pairs are not isolated. To understand whether the hydration structure observed around a singlet base pair indeed translates to the microenvironment of waters in the hydration structure on the helix, it is necessary to examine base pair stacks. Figure 4 shows umbrella sampling

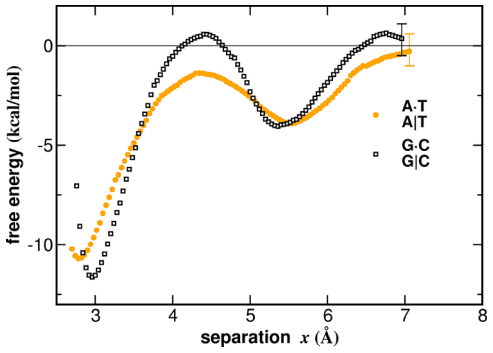


Figure 4. Free energy profiles of sliding a T out of an AA|TT doublet pair where an A·T is stacked on top of another A·T (orange circles) and sliding a C out of GG|CC doublet pair (black open squares) from umbrella sampling. Error bars show cumulative errors in the estimated pairing free energy.

results of the free energy of reversibly sliding one base out from a doublet base pair stack along the direction parallel to the base pair plane. The geometry of the bases in the initial doublet pair was chosen to be the same as in an ideal B-DNA[51] but with the sugar−phosphate backbone removed so the free energy of the base pair alone could be measured. The free energy measured this way contains contributions from both pairing and stacking, because when a base is removed from the doublet, stacking free energy is also sacrificed. Figure 4 shows the free energy profile of sliding a T out of an A·T stacked upon another A·T as the orange

https://doi.org/10.1021/acsphyschemau.3c00058
ACS Phys. Chem Au 2024, 4, 180−190

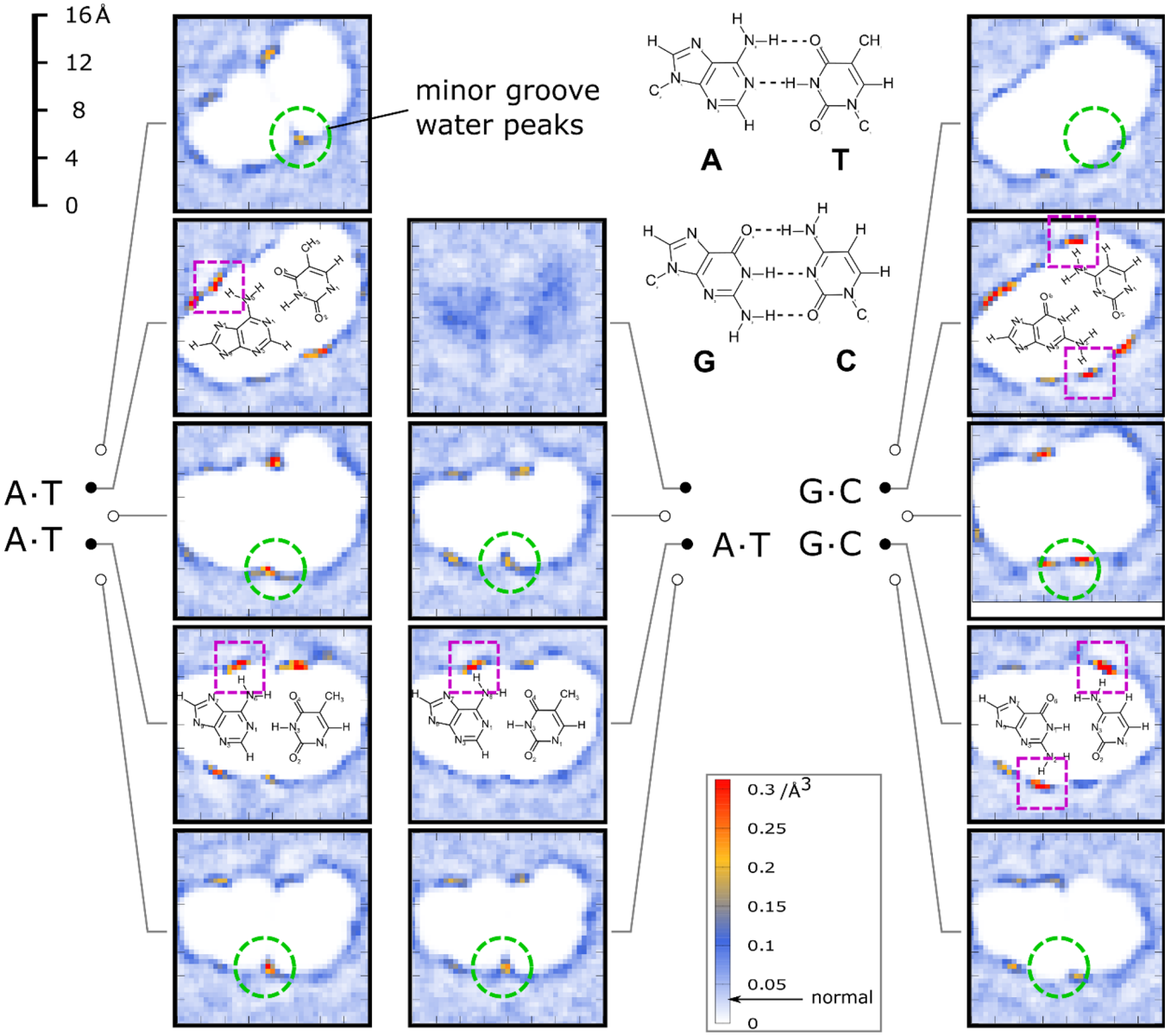


Figure 5. Water density maps illustrating hydration structure around an AA|TT doublet pair (left column), an A·T singlet pair (middle column) and a GG|CC doublet pair (right column). Slices show cross sections parallel to the base pair planes along each rung, at the midpoint between two rungs (1.7 Å from each base pair), as well as 1.7 Å above and below the stack. Water number densities in units of $Å^{-3}$ are represented using the false color scale shown, with the normal liquid water density indicated by the arrow. Each cell on the density maps is (0.4 Å).[3] Green broken circles indicate positions on the minor groove side half a base-pair step above and below and slightly offset from where the hydrogen bond would be. Pink squares highlight hydration peaks along the plane of the base pair on the non-hydrogen-bonding N−H in those amines that form direct hydrogen bonds with the carbonyls.

circles and of sliding a C out of a G·C stacked upon another G·C as the black squares. The results show that similar to the singlet base pairs, the total unstacking + unpairing energies of A·T and G·C pairs are similar to each other, to within ∼1 kcal/mol.

The lower half of Table 1 summarizes the results of these doublet free energy calculations, assuming the stacking and the pairing contribution to the free energy are additive. Using the pairing free energies for the singlet pairs, the stacking free energy of a C in GG|CC doublet is estimated to be ∼4.6 kcal/mol, whereas for a T in the AA|TT doublet, it is ∼4.1 kcal/mol. These are consistent with previous estimates.[46]

Figure 5 compares the hydration structure of the aqueous solvent around the AA|TT and GG|CC doublet pairs to the singlet A·T pair, showing slices in the water density map parallel to the base planes. The distance between the two rungs is 3.4 Å. On the far left are density slices for the AA|TT doublet, at a height of ∼1.7 Å above the top rung to 1.7 Å below the bottom rung, in increments of 1.7 Å, which is half the distance between the two rungs (i.e., half base-pair step). The slices which have a line drawing of the bases superimposed on them are along the base pair plane. The broken circles show where the minor groove waters would have been if this AA|TT doublet was inside a B-DNA. The densities are reported using false colors, with red being a factor of ∼8 above the normal density of liquid water indicated by an arrow on the color scale.

Ensemble-averaged water density maps for the A·T pair are shown in the middle column of Figure 5. The same minor groove water peaks in the AA|TT doublet are also observed in the same

https://doi.org/10.1021/acsphyschemau.3c00058
ACS Phys. Chem Au 2024, 4, 180−190

positions on the A·T pair highlighted by the broken circles. Notice that the minor groove hydration waters are found approximately at the middle between two rungs, but less prominently along the base pair plane on an A·T pair. A comparison between the A·T singlet and the AA|TT doublet suggests that the minor groove water peaks between two stacked A·T pairs at the midpoint between the two rungs seem to persist down to the single A·T base pair level.

Water density slices are shown on the right side of Figure 5 for the GG|CC doublet pair. The cells in which waters on what would be on the minor groove side if this was part of a B-DNA would be found are shown inside the broken circles. Like the AA|TT doublet, there is a high-density peak on the minor groove between the two rungs of the GG|CC doublet pair, but unlike the AA|TT doublet, the GG|CC doublet does not appear to have this peak above and below it. Furthermore, there are also high-density water peaks in the minor groove on the two planes along the two G·C base pairs. While intermediate slices are not shown, there is in fact a diffused hydration band on the minor groove side of the GG|CC doublet that stretches between the bottom rung and the top rung, but this high-density region does not extend much beyond the top or the bottom. In contrast, high-density water peaks on the minor groove of the AA|TT doublet are found halfway between the two rungs as well as half a base-pair step (1.7 Å) above the top rung and below the bottom rung, but there is no hydration water on the plane of either of the two rungs in the minor groove. Pink squares highlight hydration peaks along the plane of the base pair on the non-hydrogen-bonding N−H groups in those amines that form direct hydrogen bonds with the carbonyls on the other base. These hydration waters are observed for both G·C and A·T pairs.

The stereographs in Figure 6 illustrate the hydration structures in greater detail. Cells in the density map that have water density greater than 3 times the ambient are enclosed by the wire meshes. Cells that have water density greater than 6 times the ambient are further shown enclosed in yellow. The green broken circles highlight positions at half a base-pair step above and below the base pair on the minor groove side, the same positions highlighted by the green circles in Figure 5. Notice in Figure 6a that for the A·T pair, the hydration peaks inside these green circles are present above and below the base-pair plane, but they are absent along the base-pair plane. In contrast, Figure 6b shows that for the G·C pair, this hydration band is more diffuse and it stretches from above the base-pair plane to below it, with significant density (>3 times ambient) along the base-pair plane also. These are consistent with the densities shown in Figure 5. Finally, the pink squares in Figure 6 highlight hydration peaks along the plane of the base pair on the non-hydrogen-bonding N−H groups in those amine groups that form direct hydrogen bond with the carbonyl group on the other base. These hydration waters are also observed for both G·C and A·T pairs in the density maps in Figure 5 as well.

Hydration Profiles on the Minor Groove of AT Pairs in Duplex DNA

Figure 7 shows ensemble-averaged hydration structure along the AT track on the Drew−Dickerson dodecamer obtained from simulations. These results largely agree with previous simulations and theory,[26−32] and they provide details of the equilibrium hydration structure as a comparison to the singlet and doublet base pair data in Figure 5.

Compared to the water density peaks in Figure 5 for singlet and doublet base pairs, the hydration peaks on duplex DNA

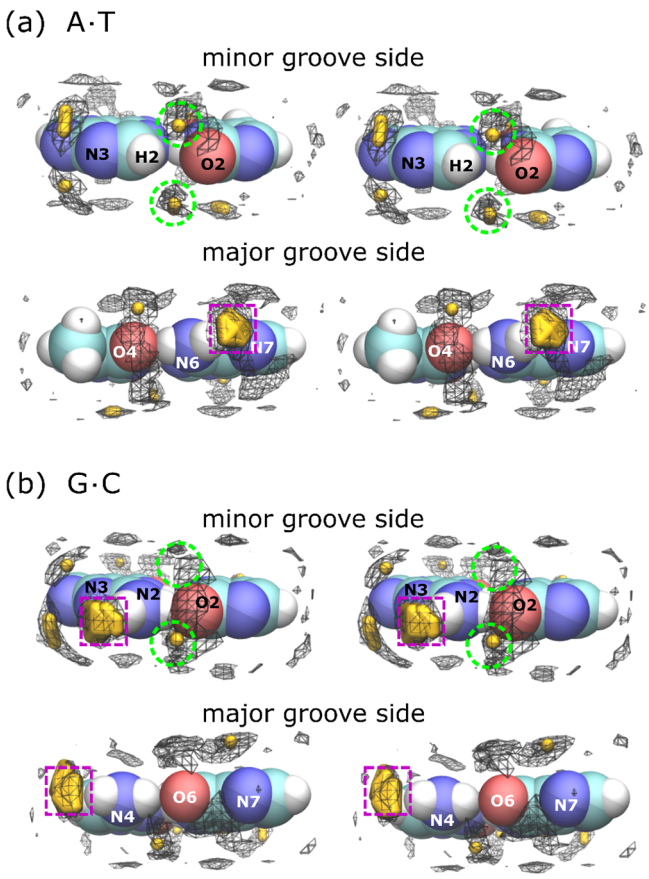


Figure 6. Stereographs showing details of the hydration structure around (a) the A·T singlet base pair and (b) the G·C pair on the minor groove side and the major groove side. Cells that have water density >3 times the ambient are shown by the wire meshes. Furthermore, regions that have water density >6 times the ambient are colored in yellow. Green broken lines show the positions of the hydration peak closest to the O2 atoms on T in A·T or on C in G·C at a distance ∼1.7 Å above and below the base pair plane. For the A·T pair, there are two distinct peaks, one above O2 and the other below. For the G·C pair, there is a more diffuse and continuous band of hydration that stretches from above the O2 atom to below it. The overall water density inside this hydration band on the G·C pair is also weaker than the two distinct peaks in the A·T pair. Pink squares highlight hydration peaks along the plane of the base pair on the non-hydrogen-bonding N−H groups in those amines that form direct hydrogen bonds with the carbonyls.

appear to be stronger. The stronger hydration peaks on the duplex may be due to coordination with the phosphate groups on the backbone,[14] but the hydration pattern around each A·T pair is fundamentally the same as the singlet and doublet base pairs. Along the planes of the A·T or T·A pairs, there are prominent water peaks inside the minor groove, in the area where the missing third hydrogen bond in an A·T pair would have been, and particularly close to the adenine H2 atom. The density profiles on the far right of Figure 7 show hydration patterns along the planes roughly midpoint between the A·T or T·A base pairs, whose positions are indicated roughly by the white circles on the helix. The green open circles highlight the strong water peaks in the minor groove between the base pair steps. These are even stronger than those along the planes of the base pair steps, with greater than 10-fold enhancement compared to the normal density of liquid water. These hydration peaks in between base pair steps are also held closer to the helix, and they are embedded deeper into the minor groove. On the

https://doi.org/10.1021/acsphyschemau.3c00058
ACS Phys. Chem Au 2024, 4, 180−190

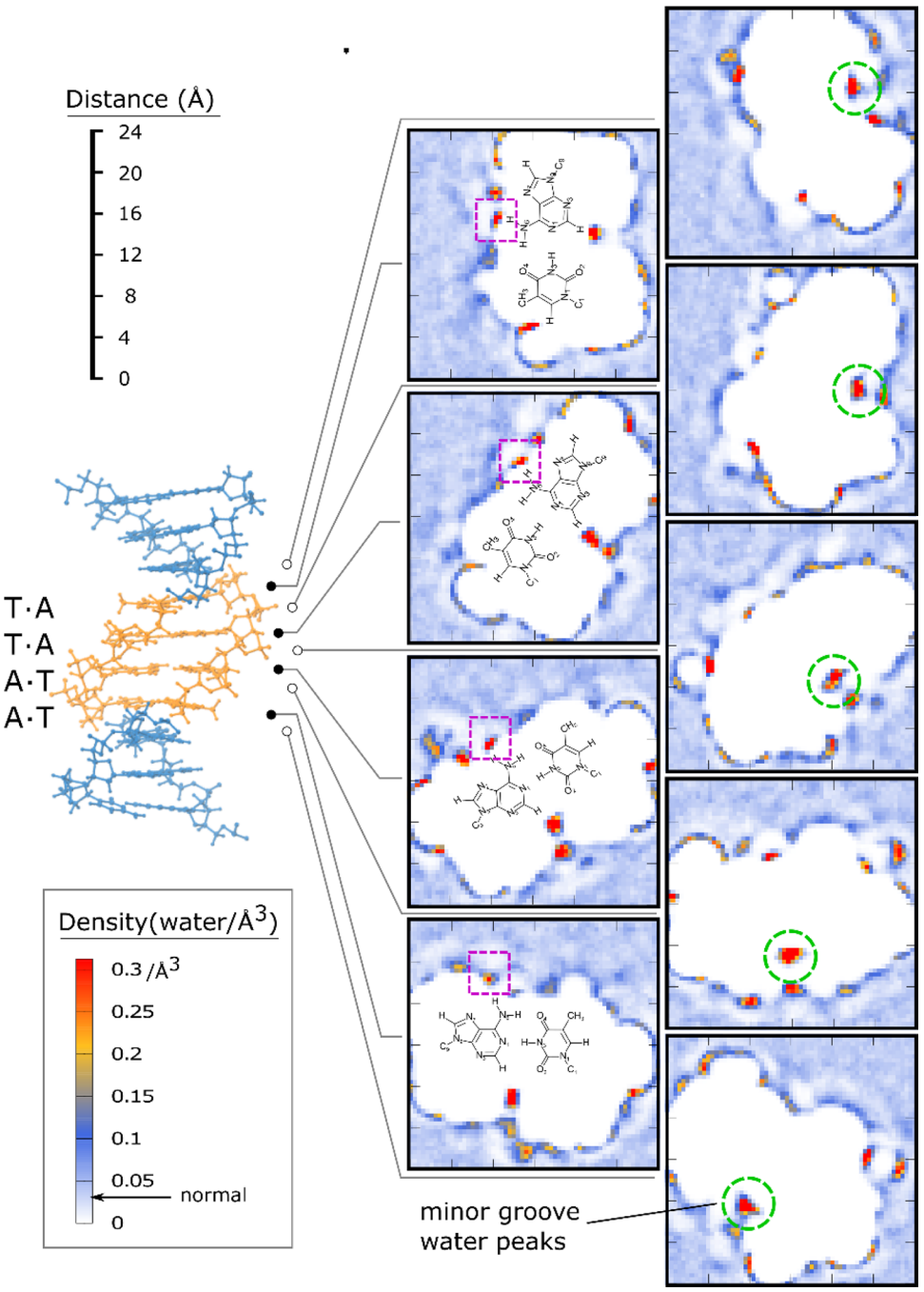


Figure 7. Equilibrium water density maps illustrating hydration structure around A·T and T·A pairs on the Drew−Dickerson dodecamer (PDB code: 1BNA). Slices show cross sections parallel to the base pair planes along each rung, at the midpoint between two rungs (1.7 Å from each base pair), as well as 1.7 Å above and below the stack. Green broken circles indicate positions on the minor groove side half a base -pair step above and below and slightly offset from where the third hydrogen bond would be. Pink squares highlight hydration peaks on the major groove along the plane of the base pair on the non-hydrogen-bonding N−H in those amines that form direct hydrogen bonds with the carbonyls.

major groove side, the strongest water peaks are the ones associated with the N−H on N6 of adenine, highlighted by the pink squares, similar to those in Figure 5.

Effects of Specific Hydration Waters on Base Paring Free Energies

Figure 8 shows effects of specific hydration waters on the A·T pair. The same are displayed for the G·C pair in Figure S3. The three columns in Figure 8 show the effects of *shared* hydration water(s) that are within 3.8 Å from *both* Ade:H2 and Thy:O2 on the left, Ade:N1 and Thy:N3 in the middle and Ade:N6 and Thy:O4 on the right. The two hydrogen bonds in the A·T pair are between Ade:N1·Thy:N3 and Ade:N6 ·Thy:O4, whereas Ade:H2 and Thy:O2 is where the missing third hydrogen bond would have been.

The circles in Figure 8a−c show the average number of hydration waters $\langle N \rangle$ between each atom pair as a function of separation between A and T. When the separation is large (>∼ 6.5 Å), there are no longer any shared hydration waters between the two bases. When A and T are paired, there are ∼2 shared hydration waters around Ade:H2 and Thy:O2 (Figure 8a), ∼0.4

https://doi.org/10.1021/acsphyschemau.3c00058
ACS Phys. Chem Au 2024, 4, 180−190

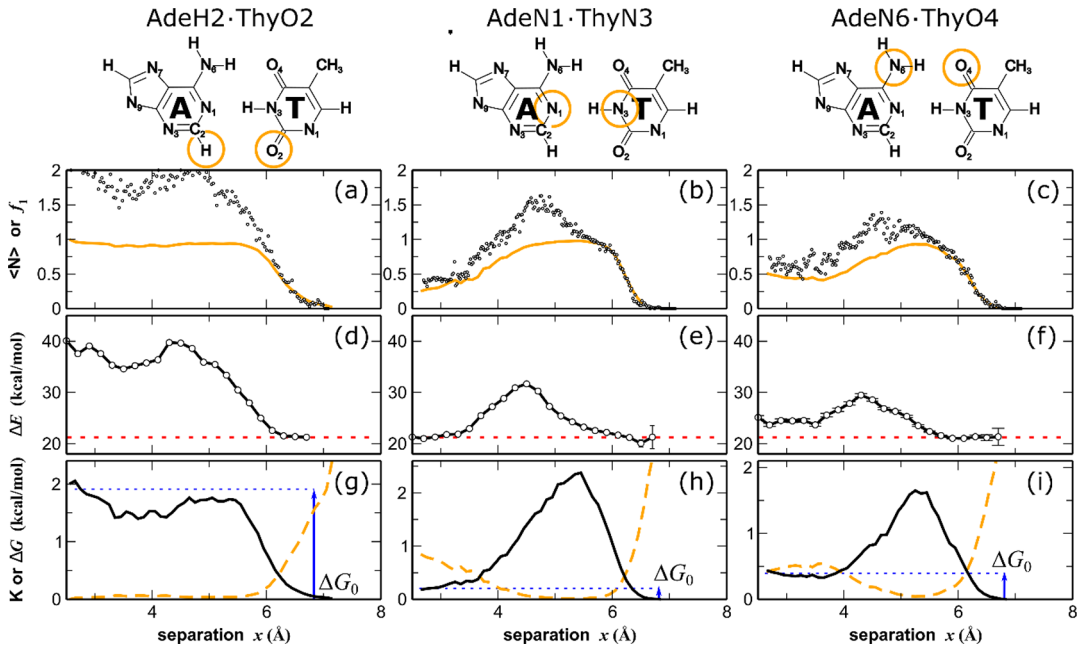


Figure 8. Effects of specific hydration waters on the A·T pair. Columns show the effects of shared hydration water(s) that are within 3.8 Å from both Ade:H2 and Thy:O2 on the left, Ade:N1 and Thy:N3 in the middle and Ade:N6 and Thy:O4 on the right. (a–c) Equilibrium number $\langle N \rangle$ of these hydration waters (open circles) and equilibrium fraction of hydrated states f_1 (orange line) as a function of base separation x. (d–f) Contribution to total energy due to these hydration waters (open circles). Red dashed line indicates the asymptotic value when the bases are separated. (g–i) Equilibrium constant $K = f_0/f_1$ between the unhydrated states f_0 and hydrated states f_1 (orange dashed line) and what the base pairing free energy would have been if the shared hydration waters between were absent (solid line). Blue arrows indicate the contribution of these hydration waters to the base pair's free energy.

shared hydration water around Ade:N1 and Thy:N3 and ∼0.6 shared hydration water around Ade:N6 and Thy:O4. The orange lines in Figure 8a–c show the fraction of microstates f_1 that are commonly hydrated by each atom pair, which was measured by the equilibrium average $f_1 = \langle \Theta \rangle$ where $\Theta = 1$ if the number of shared hydration waters is at least one, and 0 otherwise. Figure 8a shows that in the A·T pair, Ade:H2 and Thy:O2 are commonly hydrated almost all the time, with 2 hydration waters on the average, whereas Figure 8b shows that Ade:N1 and Thy:N3 are hydrated about 25% of the time, with about 1.5 waters shared by them when they are hydrated, and Figure 8c shows that Ade:N6 and Thy:O4 are commonly hydrated about 50% of the time, with about 1.4 shared hydration waters. Hydration structure data from in Figures 5 and 6 show that the hydration waters common to Ade:H2 and Thy:O2 are above and below the base pair plane on A·T, whereas those common to Ade:N1 and Thy:N3 are above or below the base pair plane, and those common to Ade:N6 and Thy:O4 are largely along the base pair plane. The fraction f_1 can also be used to establish an equilibrium constant between the unhydrated ($f_0 = 1 - f_1$) and hydrated (f_1) microstates, $K \equiv f_0/f_1$, which are plotted in Figure 8g–i as the orange dashed lines.

The contribution of the shared hydration water(s) to the total energy E is shown for each atom pair in Figure 8d–f. The free energy of the base pair consists of both enthalpic and entropic terms, but energy is almost the entirety of the enthalpic term. Figure 8d–f shows what the equilibrium energy $\langle E \rangle$ would have been if the hydrogen bonds offered by those shared hydration waters identified in Figure 8a–c were *absent*. The red dotted lines indicate $\langle E \rangle$ when A and T are not paired, showing that when the A·T pair forms, the hydration waters around Ade:H2 and Thy:O2, where the missing hydrogen bond would have been, is close to 20 kcal/mol, while the hydration water(s) around Ade:N1 and Thy:N3 have almost no effect and those around Ade:N6 and Thy:O4 have only minimal impact on the base pair's energy. The hydration waters common to Ade:N6 and Thy:O4 appear to be an integral part of the A·T pair, and the atom pair Ade:H2 and Thy:O2 are almost always hydrated by shared waters ($f_1 \sim 1$, $\langle N \rangle \sim 2$) as seen in Figure 8a. This suggests that hydration waters must be making up for the missing third hydrogen bond in the A·T base pair. Figure 8a,d suggests that the hydration waters that make up for the third hydrogen bond in the A·T base pair are incorporated into the pair to bridge Ade:H2 and Thy:O2 at a distance $x \sim 6$ Å between A and T. Close to two water molecules are held up to $x \sim 3$ Å, until the base pair is formed, and they appear to be integral to the base pair. On the contrary, those hydration water(s) picked up between Ade:N1 and Thy:N3 and Ade:N6 and Thy:O4 are released when the A·T pair forms, verifying that the hydration around Ade:H2·Thy:O2 is in fact unique. The same analysis for the G·C pair in Figure S3 in the Supporting Information shows that the hydration waters around all three atom pairs involved in hydrogen bonds in the G·C pair are also shed when the base pair is formed, and none of them appear to be integral to the base pair.

To derive a quantitative estimate for the impact of these shared hydration waters between Ade:H2 and Thy:O2 on the free energy of the A·T base pair, one can use the thermodynamic relationship

$$\overline{G}(x) = f_1(x)\cdot[\overline{G}_{\mathrm{hyd}}(x) + RT\ln f_1(x)] + f_0(x)\cdot [\overline{G}_{\mathrm{unhyd}}(x) + RT\ln f_0(x)] \quad (1)$$

for the molar free energy of the base pair $G(x)$ as a function of the separation x, in terms of the equilibrium fractions of hydrated states f_1 and its molar free energy $\overline{G}_{\mathrm{hyd}}$ and $f_0 = 1 - f_1$

https://doi.org/10.1021/acsphyschemau.3c00058
ACS Phys. Chem Au 2024, 4, 180–190

and $\bar{G}_{unhyd}$ of the unhydrated states. For every x, the hydrated and unhydrated states are in equilibrium; therefore, the free energy $G(x)$ in eq 1 has a minimum at the equilibrium value of $f_1(x)$. Minimizing $G(x)$ with respect to f_1 yields, not surprisingly, the thermodynamic relationship: $\bar{G}_{unhyd} - \bar{G}_{hyd} = -RT \ln K(x)$, where $K(x) = f_0/f_1$ is the equilibrium constant. This is plotted as the orange dashed line in Figure 8g. Using the free energy data in Figure 1b and $K(x)$ in Figure 8g, one can calculate what the base pairing free energy would have been if the shared hydration waters between Ade:H2 and Thy:O2 were absent by setting $f_1 = 0$ and $f_0 = 1$ in eq 1. The results are shown in Figure 8g as the solid black line. Notice how the apparent modulations in the free energy contribution from these hydration waters in the black line in Figure 8g mirror the modulations observed in the contribution of these hydration waters to the energy in Figure 8d, which are also mirrored in the equilibrium number of these hydration waters shown as circles in Figure 8a. The blue arrow indicates the contribution of these hydration waters to the A·T base pair free energy is approximately 2 kcal/mol. This accounts for approximately 1/3 of the A·T pairing free energy of ∼6 kcal/mol. Incorporating these hydration waters between Ade:H2 and Thy:O2, the total A·T pair free energy becomes similar to G·C. This suggests that the shared hydration waters between Ade:H2 and Thy:O2 in the A·T pair largely make up for the missing third hydrogen bond, producing an A·T pair free energy similar to G·C in water. The role of base pair hydration is also potentially relevant to the hydrogen bonding in isosteric DNA base pairs.[52]

Figure 8h,i shows free energy contributions from the hydration of the other atom pairs, Ade:N1 and Thy:N3 and Ade:N6 and Thy:O4, which are involved in direct hydrogen bonding in the A·T pair. The free energy impacts of these hydration waters are mild. The effects of hydration on the three atom pairs involved in direct hydrogen bonding in the G·C pair are also shown in Figure S3g−i. They are similarly mild. The common hydration waters between Ade:H2 and Thy:O2 in the A·T pair are therefore unique, and they appear to be intimately involved as an integral part of the base pair between adenine and thymine, producing a free energy advantage equivalent to approximately one hydrogen bond.

Some recent papers[53−55] have suggested that a direct nonclassical hydrogen bond between Ade:C2 and Thy:O2 may account for part of the stability of Watson−Crick and Hoogsteen A·T pairs. This suggestion was based on the experimental vibrational frequency measurement of the O2-containing carbonyl group in Thy and its solvatochromatic shift in duplex DNA[54] as well as theoretical calculations of vibrational signatures of A·T pairs in vacuum.[54−56] While vibrational shifts suggest that the O2 carbonyl may be involved directly in the A·T pair, they may also arise from interaction with weakly bound water molecules in the minor groove.[54] The results in the current study not only point to the central role of water molecules as the facilitator of a third hydrogen bond between Ade:H2 and Thy:O2 in an A·T pair but they also provide a quantitative measure of the impact of these hydration waters on the A·T base pair free energy.

CONCLUSIONS

The findings in this study clarify the role of hydration waters on the strength of DNA canonical base pairs. While DNA base pair complementarity is based on the different number of hydrogen bonds that are formed in the G·C pair compared to the A·T pair, the number of hydrogen bonds that exist in a base pair can be difficult to quantify because water molecules in the surrounding can interrupt as well as make up for direct hydrogen bonds between the bases. While base pairing specificity is determined by the complementarity between the hydrogen-bonding donor/acceptor motif on one base with the motif on its partner, the exact number of hydrogen bonds that exist in canonical DNA base pair is still largely unclear.

Large-scale computer simulations were used to create a detailed map for the hydration structure on A·T and G·C base pairs in water. The contributions of specific hydration waters to the free energy of each of the hydrogen bonds in the A·T and G·C pairs were also computed. The total free energy of one A·T pair in water is very similar to a G·C pair, both at ∼6.4 kcal/mol according to the free energy calculations. To understand how hydration waters influence the strength of the base pairs and how many hydrogen bonds exist in them, detailed equilibrium hydration profiles around single A·T and G·C pairs, stacked doublet pairs and on duplex DNA were sampled from the simulations. The hydration structure on the minor groove side of an A·T pair is consistent among all three scenarios. There is on the average one water molecule above and one water molecule below the A·T base pair plane occupying the space where the missing third hydrogen bond in the A·T pair would have been. And to derive a quantitative measure of how much free energy these hydration waters contribute to the A·T pair's stability, the equilibrium constant between hydrated and unhydrated states in the equilibrium ensemble was used to measure the free energy of hydration in the minor groove of the A·T pair. This corresponds to ∼2 kcal/mol, which is approximately 1/3 of the free energy of a G·C pair. The free energy contribution from these hydration waters therefore largely make up for the missing third hydrogen bond in the A·T pair, rendering it almost synonymous with the strength of an G·C pair, which has three direct hydrogen bonds.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acsphyschemau.3c00058.

Additional simulation details, including geometries and atomic charges of the bases, effect of solvent model on calculated free energies, solvent electrostatics calculation by thermodynamic integration, and effects of specific hydration waters on GC pair free energy (PDF)

AUTHOR INFORMATION

Corresponding Author

Chi H. Mak − *Departments of Chemistry and Quantitative and Computational Biology, and Center of Applied Mathematical Sciences, University of Southern California, Los Angeles, California 90089, United States;* orcid.org/0000-0002-5516-3304; Phone: 1-213-740-4101; Email: cmak@usc.edu; Fax: 1-213-740-3972

Complete contact information is available at: https://pubs.acs.org/10.1021/acsphyschemau.3c00058

Notes

The author declares no competing financial interest.

ACKNOWLEDGMENTS

This material is based in part upon work supported by the National Science Foundation under Grant Number CHE-

1664801. The author thanks Helen Berman and Bohdan Schneider for helpful discussions.

REFERENCES

(1) Watson, J. D.; Crick, F. H. C. Molecular Structure of Nucleic Acids - a Structure for Deoxyribose Nucleic Acid. *Nature* **1953**, *171* (4356), 737−738.

(2) Donohue, J.; Trueblood, K. N. Base Pairing in DNA. *J. Mol. Biol.* **1960**, *2* (6), 363−371.

(3) Olson, W. K.; Colasanti, A. V.; Lu, X.-J.; Zhurkin, V. B. Watson−Crick Base Pairs: Character and Recognition. In *Wiley Encyclopedia of Chemical Biology*; John Wiley & Sons, Ltd, 2008; pp 1−11.

(4) Saenger, W. *Principles of Nucleic Acid Structure*; Springer advanced texts in chemistry; Springer-Verlag: New York, 1984.

(5) Sukhodub, L. F. Interactions and Hydration of Nucleic Acid Bases in a Vacuum. *Experimental Study. Chem. Rev.* **1987**, *87* (3), 589−606.

(6) Yanson, I. K.; Teplitsky, A. B.; Sukhodub, L. F. Experimental Studies of Molecular Interactions between Nitrogen Bases of Nucleic Acids. *Biopolymers* **1979**, *18* (5), 1149−1170.

(7) Breslauer, K. J.; Frank, R.; Blöcker, H.; Marky, L. A. Predicting DNA Duplex Stability from the Base Sequence. *Proc. Natl. Acad. Sci. U. S. A.* **1986**, *83* (11), 3746−3750.

(8) Turner, D. H.; Sugimoto, N.; Kierzek, R.; Dreiker, S. D. Free Energy Increments for Hydrogen Bonds in Nucleic Acid Base Pairs. *J. Am. Chem. Soc.* **1987**, *109* (12), 3783−3785.

(9) Turner, D. H. Thermodynamics of Base Pairing. *Curr. Opin. Struct. Biol.* **1996**, *6*, 299−304.

(10) SantaLucia, J. A Unified View of Polymer, Dumbbell, and Oligonucleotide DNA Nearest-Neighbor Thermodynamics. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95* (4), 1460−1465.

(11) Paul, OP Ts'o. *Basic Principles in Nucleic Acid Chemistry V1*; Elsevier Science, 2012.

(12) Kool, E. T. Hydrogen Bonding, Base Stacking, and Steric Effects in DNA Replication. *Annu. Rev. Biophys. Biomol. Struct.* **2001**, *30* (1), 1−22.

(13) Kool, E. T.; Morales, J. C.; Guckian, K. M. Mimicking the Structure and Function of DNA: Insights into DNA Stability and Replication. *Angew. Chem.-Int. Ed.* **2000**, *39* (6), 990−1009.

(14) Drew, H. R.; Dickerson, R. E. Structure of a B-DNA Dodecamer: III. Geometry of Hydration. *J. Mol. Biol.* **1981**, *151* (3), 535−556.

(15) Kopka, M. L.; Fratini, A. V.; Drew, H. R.; Dickerson, R. E. Ordered Water Structure around a B-DNA Dodecamer: A Quantitative Study. *J. Mol. Biol.* **1983**, *163* (1), 129−146.

(16) Schneider, B.; Cohen, D.; Berman, H. M. Hydration of DNA Bases: Analysis of Crystallographic Data. *Biopolymers* **1992**, *32* (7), 725−750.

(17) Schneider, B.; Cohen, D. M.; Schleifer, L.; Srinivasan, A. R.; Olson, W. K.; Berman, H. M. A Systematic Method for Studying the Spatial Distribution of Water Molecules around Nucleic Acid Bases. *Biophys. J.* **1993**, *65* (6), 2291−2303.

(18) Schneider, B.; Berman, H. M. Hydration of the DNA Bases Is Local. *Biophys. J.* **1995**, *69* (6), 2661−2669.

(19) Tereshko, V.; Minasov, G.; Egli, M. The Dickerson-Drew B-DNA Dodecamer Revisited at Atomic Resolution. *J. Am. Chem. Soc.* **1999**, *121* (2), 470−471.

(20) Liepinsh, E.; Otting, G.; Wüthrich, K. NMR Observation of Individual Molecules of Hydration Water Bound to DNA Duplexes: Direct Evidence for a Spine of Hydration Water Present in Aqueous Solution. *Nucleic Acids Res.* **1992**, *20* (24), 6549−6553.

(21) Kubinec, M. G.; Wemmer, D. E. NMR Evidence for DNA Bound Water in Solution. *J. Am. Chem. Soc.* **1992**, *114* (22), 8739−8740.

(22) Maltseva, T. V.; Roselt, P.; Chattopadhyaya, J. Elucidation of the Origin of NOEs And ROEs That Show the Hydration in the Minor and Major Grooves of DNA Duplex with ATTAAT Tract by a Combination of NOESY and ROESY Experiments. *Nucleosides Nucleotides* **1998**, *17* (9−11), 1617−1634.

(23) Jóhannesson, H.; Halle, B. Minor Groove Hydration of DNA in Solution: Dependence on Base Composition and Sequence. *J. Am. Chem. Soc.* **1998**, *120* (28), 6859−6870.

(24) McDermott, M. L.; Vanselous, H.; Corcelli, S. A.; Petersen, P. B. DNA's Chiral Spine of Hydration. *ACS Cent. Sci.* **2017**, *3* (7), 708−714.

(25) Pal, S. K.; Peon, J.; Bagchi, B.; Zewail, A. H. Biological Water: Femtosecond Dynamics of Macromolecular Hydration. *J. Phys. Chem. B* **2002**, *106* (48), 12376−12395.

(26) Subramanian, P. S.; Ravishanker, G.; Beveridge, D. L. Theoretical Considerations on the Spine of Hydration in the Minor Groove of D(Cgcgaattcgcg).D(Gcgcttaagcgc) - Monte-Carlo Computer-Simulation. *P Natl. Acad. Sci. USA* **1988**, *85* (6), 1836−1840.

(27) Subramanian, P. S.; Beveridge, D. L. A Theoretical Study of the Aqueous Hydration of Canonical B d(CGCGAATTCGCG): Monte Carlo Simulation and Comparison with Crystallographic Ordered Water Sites. *J. Biomol. Struct. Dyn.* **1989**, *6* (6), 1093−1122.

(28) Miaskiewicz, K.; Osman, R.; Weinstein, H. Molecular Dynamics Simulation of the Hydrated d(CGCGAATTCGCG)2 Dodecamer. *J. Am. Chem. Soc.* **1993**, *115* (4), 1526−1537.

(29) Feig, M.; Pettitt, B. M. Modeling High-Resolution Hydration Patterns in Correlation with DNA Sequence and Conformation. *J. Mol. Biol.* **1999**, *286* (4), 1075−1095.

(30) Orozco, M.; Luque, F. J. Theoretical Methods for the Description of the Solvent Effect in Biomolecular Systems. *Chem. Rev.* **2000**, *100*, 4187−4226.

(31) Furse, K. E.; Corcelli, S. A. Molecular Dynamics Simulations of DNA Solvation Dynamics. *J. Phys. Chem. Lett.* **2010**, *1* (12), 1813−1820.

(32) Zubatiuk, T.; Shishkin, O.; Gorb, L.; Hovorun, D.; Leszczynski, J. Structural Waters in the Minor and Major Grooves of DNA—A Major Factor Governing Structural Adjustments of the A−T Mini-Helix. *J. Phys. Chem. B* **2015**, *119* (2), 381−391.

(33) Dickerson, R. E.; Drew, H. R.; Conner, B. N.; Wing, R. M.; Fratini, A. V.; Kopka, M. L. The Anatomy of A-, B-, and Z-DNA. *Science* **1982**, *216* (4545), 475−485.

(34) Woods, K. K.; Lan, T.; McLaughlin, L. W.; Williams, L. D. The Role of Minor Groove Functional Groups in DNA Hydration. *Nucleic Acids Res.* **2003**, *31* (5), 1536−1540.

(35) Sun, Z.; McLaughlin, L. W. Probing the Nature of Three-Centered Hydrogen Bonds in Minor-Groove Ligand−DNA Interactions: The Contribution of Fluorine Hydrogen Bonds to Complex Stability. *J. Am. Chem. Soc.* **2007**, *129* (41), 12531−12536.

(36) Salandria, K. J.; Arico, J. W.; Calhoun, A. K.; McLaughlin, L. W. Stability of DNA Containing a Structural Water Mimic in an A-T Rich Sequence. *J. Am. Chem. Soc.* **2011**, *133* (6), 1766−1768.

(37) Tereshko, V.; Minasov, G.; Egli, M. A "Hydrat-Ion" Spine in a B-DNA Minor Groove. *J. Am. Chem. Soc.* **1999**, *121* (15), 3590−3595.

(38) Halle, B.; Denisov, V. P. Water and Monovalent Ions in the Minor Groove of B-DNA Oligonucleotides as Seen by NMR. *Biopolymers* **1998**, *48* (4), 210−233, DOI: 10.1002/(Sici)1097-0282(1998)48:4<210::Aid-Bip3>3.3.Co;2-P.

(39) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *J. Am. Chem. Soc.* **1995**, *117* (19), 5179−5197.

(40) Torrie, G. M.; Valleau, J. P. Nonphysical Sampling Distributions in Monte Carlo Free-Energy Estimation: Umbrella Sampling. *J. Comput. Phys.* **1977**, *23* (2), 187−199.

(41) Pérez, A.; Marchán, I.; Svozil, D.; Sponer, J.; Cheatham, T. E.; Laughton, C. A.; Orozco, M. Refinement of the AMBER Force Field for Nucleic Acids: Improving the Description of α/γ Conformers. *Biophys. J.* **2007**, *92* (11), 3817−3829.

(42) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *J. Am. Chem. Soc.* **1995**, *117* (19), 5179−5197.

(43) Dans, P. D.; Ivani, I.; Hospital, A.; Portella, G.; González, C.; Orozco, M. How Accurate Are Accurate Force-Fields for B-DNA? *Nucleic Acids Res.* **2017**, *45* (7), 4217−4230.

https://doi.org/10.1021/acsphyschemau.3c00058
ACS Phys. Chem Au 2024, 4, 180−190

(44) Stofer, E.; Chipot, C.; Lavery, R. Free Energy Calculations of Watson- Crick Base Pairing in Aqueous Solution. *J. Am. Chem. Soc.* **1999**, *121* (41), 9503−9508.
(45) Giudice, E.; Várnai, P.; Lavery, R. Base Pair Opening within B-DNA: Free Energy Pathways for GC and AT Pairs from Umbrella Sampling Simulations. *Nucleic Acids Res.* **2003**, *31* (5), 1434−1443.
(46) Mak, C. H. Unraveling Base Stacking Driving Forces in DNA. *J. Phys. Chem. B* **2016**, *120* (26), 6010−6020.
(47) Li, R.; Mak, C. H. A Deep Dive into DNA Base Pairing Interactions Under Water. *J. Phys. Chem. B* **2020**, *124* (27), 5559−5570.
(48) Challacombe, M.; White, C.; HeadGordon, M. Periodic Boundary Conditions and the Fast Multipole Method. *J. Chem. Phys.* **1997**, *107*, 10131−10140.
(49) Straatsma, T. P.; Berendsen, H. J. C.; Postma, J. P. M. Free Energy of Hydrophobic Hydration: A Molecular Dynamics Study of Noble Gases in Water. *J. Chem. Phys.* **1986**, *85* (11), 6720−6727.
(50) Straatsma, T. P.; Berendsen, H. J. C. Free Energy of Ionic Hydration: Analysis of a Thermodynamic Integration Technique to Evaluate Free Energy Differences by Molecular Dynamics Simulations. *J. Chem. Phys.* **1988**, *89* (9), 5876−5886.
(51) Olson, W. K.; Bansal, M.; Burley, S. K.; Dickerson, R. E.; Gerstein, M.; Harvey, S. C.; Heinemann, U.; Lu, X. J.; Neidle, S.; Shakked, Z.; Sklenar, H.; Suzuki, M.; Tung, C. S.; Westhof, E.; Wolberger, C.; Berman, H. M. A Standard Reference Frame for the Description of Nucleic Acid Base-Pair Geometry. *J. Mol. Biol.* **2001**, *313* (1), 229−237.
(52) Khakshoor, O.; Wheeler, S. E.; Houk, K. N.; Kool, E. T. Measurement and Theory of Hydrogen Bonding Contribution to Isosteric DNA Base Pairs. *J. Am. Chem. Soc.* **2012**, *134* (6), 3154−3163.
(53) Derewenda, Z. S. C-H Groups as Donors in Hydrogen Bonds: A Historical Overview and Occurrence in Proteins and Nucleic Acids. *Int. J. Mol. Sci.* **2023**, *24* (17), 13165.
(54) Fick, R. J.; Liu, A. Y.; Nussbaumer, F.; Kreutz, C.; Rangadurai, A.; Xu, Y.; Sommer, R. D.; Shi, H.; Scheiner, S.; Stelling, A. L. Probing the Hydrogen-Bonding Environment of Individual Bases in DNA Duplexes with Isotope-Edited Infrared Spectroscopy. *J. Phys. Chem. B* **2021**, *125* (28), 7613−7627.
(55) Beiranvand, N.; Freindorf, M.; Kraka, E. Hydrogen Bonding in Natural and Unnatural Base Pairs—A Local Vibrational Mode Study. *Molecules* **2021**, *26* (8), 2268.
(56) Krakauer, D. C.; RodriguezGirones, M. A. Searching and Learning in a Random Environment. *J. Theor. Biol.* **1995**, *177* (4), 417−429.

https://doi.org/10.1021/acsphyschemau.3c00058
ACS Phys. Chem Au 2024, 4, 180−190