Hydration Heat Capacity of Nucleic Acid Constituents Determined from the Random Network Model

Bhupinder Madan and Kim A. Sharp
The Johnson Research Foundation, Department of Biochemistry and Biophysics, University of Pennsylvania, Philadelphia, Pennsylvania 19104

ABSTRACT The heat capacities of hydration (*dCp*) of the five nucleic acid bases A, G, C, T, and U, the sugars ribose and deoxyribose, and the phosphate backbone were determined using Monte Carlo simulations and the random network model. Solute-induced changes in the mean length and root mean square angle of hydrogen bonds between hydration shell waters were used to compute *dCp* for these solutes. For all solutes the *dCp* is significantly more positive than predicted from accessible surface area (ASA) models of heat capacity. In ASA models, nitrogen, oxygen, and phosphorus atoms are considered as uniformly polar, therefore making a negative contribution to *dCp*. However, the simulations show that many of these polar atoms are hydrated by water whose hydrogen bonds are less distorted than in bulk, leading to a positive *dCp*. This is in contrast to the effect of polar groups seen previously in small molecules and amino acids, which increase the water H-bond distortion, giving negative *dCp* contributions. Our results imply that *dCp* accompanying DNA dehydration in DNA-ligand and DNA-protein binding reactions may be significantly more negative than previously believed and that dehydration is a significant contributor to the large decrease in heat capacity seen in experiments.

INTRODUCTION

An understanding of the thermodynamics of RNA/DNA-protein and DNA-ligand binding reactions is crucial in understanding the origins of base sequence specificity. Base sequence specificity plays a role in gene expression and repression, replication, recombination, and drug specificity. The changes in the free energy, ΔG, enthalpy, ΔH, and entropy, ΔS, of binding reactions are of great value but are difficult to relate to specific structural interactions because they include contributions from various types of noncovalent interactions such as electrostatic forces, dispersion forces, and solvation effects. Sequence-specific binding of proteins to DNA and RNA is often accompanied by very large negative heat capacity changes (ΔCp), on the order of several hundred to a thousand cal/mol/K. Nonspecific binding usually has much smaller changes, so that a large ΔCp is often taken as a signature of specific binding. When both binding partners are fully folded the heat capacity of binding is believed to arise primarily from solvent effects. There is good evidence for this for protein/protein binding reactions (Privalov and Gill, 1988; Sturtevant, 1977), however, it is more of a working assumption in the analysis of DNA-protein binding. Nevertheless, heat capacity changes associated with DNA binding reactions have been used in analyses of the thermodynamics to de-convolute solvent and nonsolvent contributions to the energetics.

Heat capacity changes arising from solvation effects in proteins and small solutes are observed to scale with changes in solvent-accessible surface area (ASA). Area-based models have been shown to provide good estimates of heat capacity changes for protein folding and protein binding. In these models, the change in hydration heat capacity is estimated using:

$$\Delta C_p = C_{apolar}\Delta A_{apolar} + C_{polar}\Delta A_{polar}, \quad (1)$$

where ΔA_{apolar} and ΔA_{polar} are the changes in apolar and polar ASA, respectively, and C_{apolar} and C_{polar} are the corresponding heat capacity coefficients, parameterized using solute-transfer, model compound, and protein-folding thermodynamic data (Freire, 1995; Makhatadze and Privalov, 1990a,b; Murphy and Freire, 1992; Myers et al., 1995; Spolar et al., 1992). Values for C_{apolar} range from 0.32 to 42 cal/mol/K/Å^2, and for C_{polar} range from -0.14 to -0.265 cal/mol/K/Å^2, depending on the parameterization set. It should be noted that most polar groups of small molecules measured previously make a negative contribution to hydration heat capacity (Cabani et al., 1981), explaining the negative sign of C_{polar}.

The parameterization of Eq. 1 for proteins has been greatly aided by direct measurement of hydration heat capacities of peptides and other model compounds corresponding to amino acid fragments (Makhatadze and Privalov, 1990a,b; Murphy and Gill, 1991; Murphy et al., 1990). In contrast, there are no data for the corresponding nucleic acid model compounds, i.e., the base, sugar, and phosphate backbone moieties. Thus, the only approach used to date for calculating ΔCp for nucleic acid-protein binding has been to apply the ASA model assuming that the same protein-derived parameters can be used for the nucleic acid ASA. This assumption appears to work well for the binding of intercalating drugs to DNA (Ren et al., 2000) where observed decreases in *Cp* upon binding are consistent with the dominant effect of apolar surface burial. Single-strand

Received for publication 31 January 2001 and in final form 8 June 2001.
Address reprint requests to Dr. Kim A. Sharp, The Johnson Research Foundation, Department of Biochemistry and Biophysics, University of Pennsylvania, Philadelphia, PA 19104-6059. Tel.: 215-573-3506; Fax: 215-898-4217; E-mail: sharpk@mail.upenn.edu.

0006-3495/01/10/1881/07 $2.00

duplex DNA formation has a small heat capacity change, indicating compensation between polar and nonpolar area burial (Holbrook et al., 1999). The heat capacity of ion-induced DNA condensation, in contrast, is positive (Matulis et al., 2000), implying that burial of polar surface area dominates. For some protein-DNA binding, however, there is often a significant discrepancy between the experimental and ASA-calculated values of ΔCp (Connelly et al., 1993). In some cases of DNA-ligand binding, this can be explained by coupled protein folding and binding (Lundback et al., 2000; Spolar and Record, 1994) where the folding buries additional, primarily hydrophobic area. In other cases, there appears to be too large a heat capacity decrease upon binding the protein to be explained by this (Jin et al., 1993; Ladbury et al., 1994; Lundback et al., 1993; Merabet and Ackers, 1995; Oda et al., 1998). We previously investigated the contribution of long-range solvent and ion-screened electrostatic interactions to ΔCp (Gallagher and Sharp, 1998). These contributions do not scale with ASA and so constitute a possible additional source of the heat capacity decrease. However, our calculations showed that this can only partially explain the difference. A recent analysis of the heat capacity of DNA melting also shows that electrostatic effects alone cannot account for the large observed heat capacity changes (Rouzina and Bloomfield, 1999). A detailed review of protein unfolding thermodynamics (Robertson and Murphy, 1997) illustrates another difficulty: the heat capacity data for proteins may be fit equally well with a number of different area-based models. It has also been proposed that changes in the vibrational state of water molecules that are either released or trapped upon DNA-protein binding may contribute to the large decrease in *Cp* (Ladbury et al., 1994; Morton and Ladbury, 1996). Reduction in the vibrational or breathing motions of DNA upon ligand binding may also play a role. Unfortunately, it is not yet possible to calculate such vibrational contributions reliably enough to test this explanation. Thus, to date it is unclear whether large decreases in *Cp* upon ligand-DNA binding can be satisfactorily accounted for in terms of solvation contributions calculated using protein-derived ASA parameters. We consider here whether heat capacity parameters that have been derived from proteins are applicable to the nucleic acid's contribution to the hydration heat capacity. Clearly this requires a better understanding of the heat capacity changes associated with nucleic acid hydration.

In previous work we showed that a combination of explicit water simulations combined with the random network model may be used to obtain quantitative agreement between measured and calculated hydration heat capacities (Madan and Sharp, 1999, 1997; Sharp and Madan, 1997; Vanzi et al., 1998). In this work we apply the same combination of explicit water simulations and the random network model to calculate estimates of the hydration heat capacities of all the fragments comprising nucleic acids. We had two goals: first, to compute the heat capacity of hydration of the base, sugar, and phosphate backbone using this already established technique, and second, to compare the heat capacity of hydration of these molecules with those computed from ASA models and to explain discrepancies, if any, with the ASA model data with a goal to improving their accuracy and our understanding of the experimental data.

MATERIALS AND METHODS

The solvation of five nucleic acid bases A, G, C, T, and U, the sugars ribose and deoxyribose, and the phosphate backbone (which comprise all the constituents of DNA and RNA excepting the modified bases of tRNA) was simulated by inserting each of these molecules in turn into a box of 750 molecules of TIP4P water. The phosphate backbone was represented by the di-methyl phosphate ion ($(CH_3)_2PO_4^-$), the two capping methyl groups representing the adjoining ribose C3′ and C5′ groups. Solutes were represented by the OPLS solute potential function (Pranata et al., 1991; Jorgensen et al., 1983) and included solute flexibility. The simulation method and its analysis have been described in detail in our earlier work (Madan and Sharp, 1999, 1997; Sharp and Madan, 1997). Briefly, the initial solvent box dimensions were 37.5 Å × 25 Å × 25 Å. Periodic boundary conditions and a nonbond cutoff of 12.0 Å were used. A Metropolis Monte Carlo algorithm incorporated in the program BOSS (Jorgensen, 1992) was used to run the simulations at constant temperature and pressure. The solutions were first equilibrated at 1 atm pressure and 25°C for 5×10^7 Monte Carlo steps, and then data were collected over 10 consecutive runs of 1×10^7 steps each. The error estimates for the average quantities were computed from deviations of batch averages between the 10 runs.

An instantaneous snapshot of the system was analyzed every 1000 Monte Carlo steps during each of the 10 data collection runs. Solute atom-water oxygen radial distribution functions, $g(r)$, were computed from the simulations, and the extent of the first hydration shell of each solute atom was determined from the position of the first minimum of its $g(r)$. All waters lying within the solute's first hydration shell were identified in each snapshot, and each such water was assigned to the hydration shell of the closest solute atom. Two first-shell water molecules are deemed to be hydrogen bonded (H-bonded) if they lie within a distance of 3.4 Å, the first minimum in the O-O radial distribution function of pure water. For each pair of H-bonded first-shell waters, the H-bond length and angle were computed, and the angle and length probability distributions were accumulated throughout the simulation. Fig. 1 *A* illustrates the definition of H-bond length and angle used here. The three random network model parameters the average oxygen-oxygen distance between H-bonded water molecules, d, the standard deviation in the average oxygen-oxygen distance, s, and the root mean square (rms) hydrogen bond angle between two water molecules, θ, were then computed from the probability distribution functions.

Each H-bond between two first-shell waters can be classified according to which solute atom(s) the two waters are hydrating. Each class has the potential to have different mean H-bond parameters, depending on the nature and position of the solute atoms and the effect of their neighbors. With solutes the size of nucleic acids, this results in many possible classes of H-bonds. To make the presentation and analysis of the results manageable, the atoms of each solute were grouped into a small number of classes based on similarity in their polarity and position in the solute. For example, with guanine, carbon atoms C4, C5, and C6 that lie along the center of the base formed one class (a), the other ring carbons formed a second class (b), the ring nitrogens formed a third class (c), etc. (Fig. 1 *B*). A complete list of the groupings is given in Table 1. The random network parameters d, s, and θ were computed for all possible classes of H-bonds (a-a, a-b, b-b, etc.) of each solute, where the a-b class indicates a hydrogen bond between a

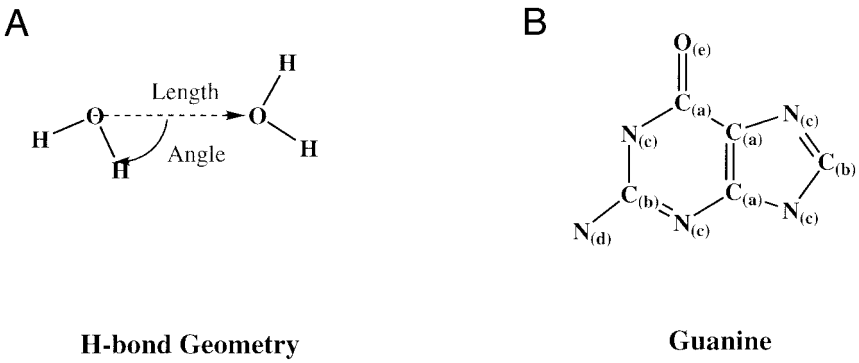

FIGURE 1 (*A*) Definition of H-bond angle and length; (*B*) Example of atom/group classification for the guanine base. The group for each atom is indicated in brackets.

water solvating one of the solute's a-class atoms and a water solvating one of the solute's b-class atoms.

The total heat capacity of hydration, ΔCp, for each of the solute molecules was determined from the changes in the random network model parameters, with respect to bulk water, of all the first shell H-bonds, by summing the contributions from all of the perturbed hydrogen bonds:

$$\Delta C_p^{hyd} = \sum_i N_i[C_p^{rn}(d_i,s_i\theta_i) - C_p^{rn}(d_o,s_o\theta_o)]$$

$$= \sum_i N_i \Delta C_p^{rn}(d,s,\theta), \quad (2)$$

where C_p^{rn} is the contribution to heat capacity arising from a group of N_i perturbed H bonds with average parameters d_i, s_i, and θ_i, where d_o, s_o, and θ_o are the corresponding values for the bulk water in the absence of solute. The expression for C_p^{rn} in terms of d_i, s_i, and θ_i was obtained from a modified version of the random network model of water developed by Henn and Kauzmann (1989) as described in our previous work (Madan and Sharp, 2000; Sharp and Madan, 1997). As we have repeatedly observed in our previous work, significant perturbations in d, s, or θ are seen only in the first hydration shell. Thus, we have not included the contributions from second- and higher-order hydration shells in the above equation.

The heat capacities of hydration for the various solutes were also computed from the ASA model using Eq. 1 with the parameters of Spolar et al. (1992) for C_{apolar} and C_{polar}, 0.32 and -0.14 cal/mol/K/Å^2, respectively. Following the usual practice, hydrogens were omitted from these calculations, and oxygen, nitrogen, and phosphorus atoms are taken to be polar and carbon atoms to be apolar. Solvent-accessible areas were calculated using the program SURFCV (Sridharan et al., 1992) using standard atomic radii (Bondi, 1968).

TABLE 1 Classification of solute atoms

Solute	Atom classes*
Adenine	(C4, C5, C6), (C2, C8), (N1, N3, N7, N9), (N6)
Guanine	(C4, C5, C6), (C2, C8), (N1, N3, N7, N9), (N2), (O6)
Cytosine	(C2, C4), (C5, C6), (N1, N3), (N4), (O2)
Thymine	(C2, C4), (C5, C6), (N1, N3), (CH_3), (O2, O4)
Uracil	(C2, C4), (C5, C6), (N1, N3), (O2, O4)
Ribose	(C1′, C4′), (C2′, C3′), (C5′), (O4′), (O1′, O2′, O3′, O5′)
Deoxyribose	(C1′, C4′), (C2′, C3′)(C5′), (O4′), (O1′, O3′, O5′)
Dimethyl phosphate	(P), (O1, O2), (O3′, O5′), (C3′, C5′)

*Classes of similar atoms are grouped within parentheses.

RESULTS

Examination of the random network parameters d, s, and especially θ provide a quantitative description of the distortions in water structure induced by a solute or solute group. The general picture of these distortions for the eight solutes studied here is presented in Fig. 2 in the form of a plot of d versus θ. Each point represents the mean values of these two parameters for one of the classes of H-bond in one of the solutes. The values for pure water (d = 2.95 Å, θ = 29.5°) are shown for comparison. For none of the H-bond classes does s vary significantly from its pure water value of 0.224 Å. The overall picture is that most of the first-hydration-shell H-bonds are less distorted than in bulk, having shorter mean lengths and smaller angles. This result is expected of H-bonds where either both or one of the waters is solvating an apolar group (referred to generically as the apolar and mixed H-bond classes, respectively) (Sharp and Madan, 1997). All the apolar and most mixed-class H-bonds indeed fall on the lower left region with respect to bulk water.

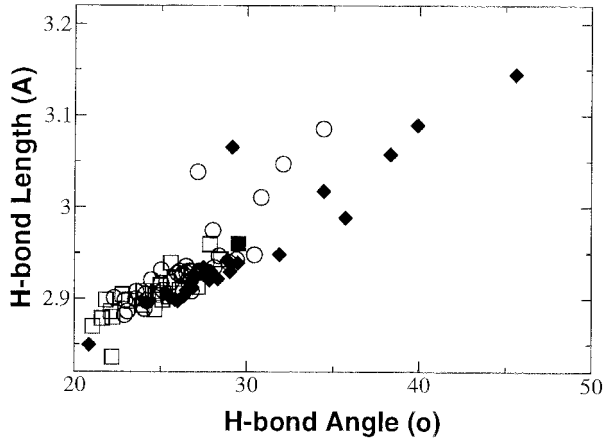

FIGURE 2 Plot of mean H-bond length versus rms H-bond angle for pure water (■) and for all eight solutes. H-bond class: mixed (○), apolar (□), and polar (◆).

TABLE 2 First-shell water-water H-bonds showing significantly increased angular distortion relative to bulk water

Solute	Solvated groups	θ	N
Adenine	N6-N6	32°	1.1
Guanine	O6-O6	38°	1.3
	O6-N2	34°	1.0
Cytosine	O2-O2	40°	1.2
	O2-C6	34°	1.0
Thymine	None		
Uracil	None		
Ribose	Hydroxyl oxygens	36°	4.7
Deoxyribose	Hydroxyl oxygens	36°	3.2
Dimethyl phosphate*	O1-O1, O2-O2, O1-O2	46°	2.1

H-bonds are classified according to the solute group(s) in whose first hydration shell the two waters lie. θ and N are the rms H-bond angle and mean number of H-bonds, respectively. These H-bonds make a negative contribution to the hydration *Cp*.

*No waters are found in the phosphorus atom's hydration shell because this atom is completely occluded by the four attached oxygen atoms.

There are relatively few classes of H-bond with increased distortion (higher θ values than in bulk), and these belong to the classes where both waters are solvating a polar group (the generic polar H-bond class). Most of these involve solvation of at least one, usually two, polar oxygens (Table 2); the others involve solvation of the NH_2 groups of A and G. Again, this result is expected of the polar H-bond class. A departure from this pattern is that many of the nominally polar class of H-bonds, predominantly those where both waters are solvating N, or NH groups of the bases, have decreased distortion. Although these decreases are not large on a per H-bond basis, they are significant, because there are many such H-bonds, particularly around the nitrogen-rich bases. Waters solvating the acetal oxygen of ribose and deoxyribose, which is polar, but less polar than a hydroxyl oxygen, also have decreased H-bond distortion.

The H-bond angle distribution for pure water is bimodal with a large peak at ~12° and a small peak at ~52° (Fig. 3 *A*). The first large peak is due to stronger, relatively straight hydrogen bonds formed by a water molecule with four quasi-tetrahedrally arranged neighbors. The second small peak is due to weaker, more bent hydrogen bonds formed with extra-tetrahedral mismatched water molecules. The increases and decreases in θ seen for various H-bond classes of water solvating the eight solutes arises from shifts in numbers of H-bonds in these two populations, rather than changes in the H-bond angle of either population. This is illustrated in Fig. 3 *A* with representative H-bond angle probability distributions for the apolar and polar class of H-bonds (water around the thymine methyl and ribose hydroxyl groups, respectively). For the apolar class where waters solvate methyls, methylene groups, and carbon atoms in aromatic rings, the low-angle peak is increased and the high-angle peak decreased. We attribute this to pushing away of the mismatched water by these nonpolar groups,

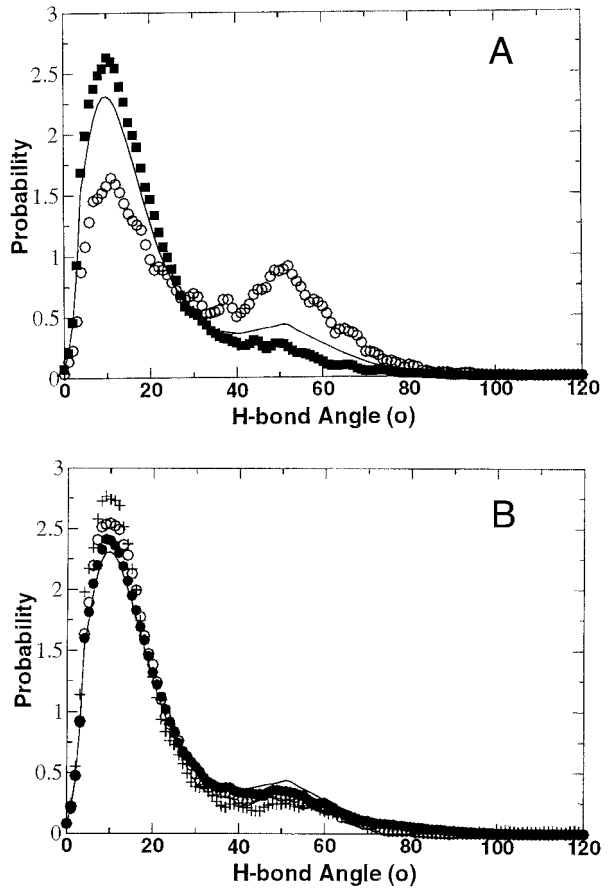

FIGURE 3 (*A*) Water-water H-bond angle distribution at 25°C for pure water (——); for the first hydration shell of ribose hydroxyl groups, d = 2.99 Å, θ = 36° (○), and for the first hydration shell of the thymine methyl group, d = 2.91 Å, θ = 25° (■). (*B*) Water-water H-bond angle distribution at 25°C for pure water (——); for the first hydration shell of cytosine ring nitrogens, d = 2.91 Å, θ = 27° (○), for the first hydration shell of adenine ring nitrogens, d = 2.92 Å, θ = 28° (●), and for the first hydration shell of ribose acetal oxygen, d = 2.90 Å, θ = 26°(+).

hence bringing more order to the hydration shell relative to that in the bulk water. The converse is true for hydrogen bonds formed among water molecules in the hydration shell of polar oxygen atoms. These show a diminished first peak and an enhanced second peak. This increase in more distorted H-bonds arises from the tendency of the polar atoms to align the water dipoles.

Representative examples of the H-bond angle probability distributions for waters solvating base nitrogens and sugar acetal oxygens are shown in Fig. 3 *B*. These show a small but significant increase in the low-angle H-bond peak at the expense of the high-angle peak, leading to a decrease in the rms H-bond angle.

In the random network model, solute-induced decreases in water-water H-bond angle result in an increase in the *Cp* of water, whereas solute-induced increases in angle produce a decrease in *Cp*. The reason for this is that the first situation corresponds to a mild perturbation in water structure, which

TABLE 3 Comparison of random network and ASA models for hydration heat capacity

Solute	Area (Å^2) Apolar	Area (Å^2) Polar	N*	Cp (hydration) (cal/mol/K) ASA†	Cp (hydration) (cal/mol/K) RNM‡
Adenine	89	189	66	2.02	22 ± 5
Guanine	60	227	69	−12.58	18 ± 4
Cytosine	79	165	45	2.18	21 ± 7
Thymine	115	141	52	17.06	31 ± 5
Uracil	87	148	53	7.12	28 ± 4
Ribose	100	167	37	8.62	21 ± 6
Deoxyribose	125	133	34	21	21 ± 4
Dimethyl phosphate§	0	80	7	−11	−1.4 ± 2

*Mean number of all classes of first-shell H-bonds.
†Using $C_{apolar} = 0.32$ and $C_{polar} = -0.14$ cal/mol/K/Å^2.
‡Predicted from the random network model (RNM), accounting for the previously observed systematic underestimate of the experimental values in this model by a factor of two (Sharp and Madan, 1997).
§Data for P, O1, and O2 atoms only, omitting the capping O3′/C3′, O5′/C5′ groups, which are included in the ribose and deoxyribose data.

results in net increase in the different energy states accessible to the water, with a corresponding increase in the energy fluctuation, (e.g., in Cp). The second situation corresponds to a large perturbation in the energy levels of different water configurations, such that the higher levels ($\gg kT$ energy gap) become less populated, producing a net decrease in the energy fluctuation and thus in Cp (Sharp and Madan, 1997).

Table 2 shows that there are relatively few H-bonds with increased angle compared with the total number of H-bonds between first-shell waters (Table 3). In each of the five bases, the intra-water H-bonds with the highest angle occur between waters that are hydrating the H-bonding groups that mediate Watson-Crick base pairing (Table 2). The large number of H-bonds with decreased H-bond angle is reflected in the net hydration Cp calculated for the bases using the random network Model (Table 3). The hydration Cp for the bases ranges from 18 to 31 cal/mol/K, considerably more positive, and even of opposite sign, to that expected using the area-based model with protein-derived parameters for heat capacity. This reflects the contribution from the many H-bonds between waters solvating base nitrogens. Although nitrogen atoms are usually considered as polar atoms in ASA models of solvation (Eisenberg and McLachlan, 1986; Spolar and Record, 1994), and based on their significant partial charge in potential functions such as OPLS (used here), our simulations show that overall they make a significant positive contribution to hydration heat capacity. This was observed in previous simulations of *N*-methyl acetamide and urea (Sharp and Madan, 1997; Vanzi et al., 1998) but is of more consequence for the nitrogen-rich bases. The exceptions are the NH_2 groups of A and G, involved in base pairing, which show a more polar character. The calculated hydration Cp for the two sugars is positive for both the random network and ASA models, and for the less polar deoxyribose there is quite close agreement between the models. Both models predict a negative hydration Cp for the ionic (and therefore highly polar) phosphate group, but again the random network model lies on the more positive side. The overall result is that the random network model predicts consistently less negative/more positive Cp of hydration for the entire base.

TABLE 4 Correlation between accessible area, calculated hydration heat capacity, and number of first-shell H-bonds (R^2)

	Accessible area*	Cp(hydration) (RN)
Cp(hydration) (RN)	0.802 (0.803)†	1
Number of H-bonds	0.835	0.69

*Total accessible area of solute.
†Number in parentheses is for a fit to apolar/polar area model (Eq. 1).

DISCUSSION

From the simulations of the water structure around the constituent groups of nucleic acids the overall picture is that most of the first-hydration-shell H-bonds are less distorted than in bulk, having shorter mean lengths, and smaller angles. These provide a positive contribution to the hydration heat capacity. There are relatively few classes of H-bond with higher θ values than in bulk, notably those around the phosphate backbone, the sugar hydroxyls, and the base-pairing groups. These provide a smaller negative contribution to the hydration heat capacity. The large number of intra-water H-bonds with decreased H-bond angle results in calculated hydration Cp values for the bases considerably more positive than that expected using the area-based model with protein-derived parameters for heat capacity.

To more directly compare the random network and ASA models, the areas and random network values for Cp hydration in Table 3 were used to extract values for C_{apolar} and C_{polar} using a linear least-squares fit of the data to Eq. 1, yielding:

$$Cp(\text{hydration}) = 0.173A_{apolar} + 0.171A_{polar},$$

with $R^2 = 0.8$ (Table 4). Thus, within the resolution of the data Cp of hydration is simply given by:

$$Cp(\text{hydration}) \approx 0.17(A_{apolar} + A_{polar}) = 0.17A_{total}; \quad (4)$$

i.e., it is proportional to the total accessible surface area. If the one outlier, the phosphate backbone, is omitted, the coefficient increases slightly to 0.18 cal/mol/K/Å. Interestingly, this lack of distinction between the apolar and polar surface area contributions is very similar to the results of an analysis of Robertson and Murphy (1997), who found that Cp of unfolding for a set of 49 proteins could be fit as well by Eq. 4 with a coefficient of 0.15 cal/mol/K/Å as by Eq. 1.

The atoms of the nucleic acid constituents broadly fall into three classes, carbons, nitrogens/acetal oxygens, and

hydroxyl/phosphate oxygens, and their effect on H-bond structure of their solvating waters can be rationalized as follows. First, the carbon atoms have the largest radius and, on average, have a small magnitude of atomic or partial charge. Thus, their electrostatic interaction with water is the weakest, and their effect on the H-bond structure is mainly geometric: displacement of the mismatch water, with the concomitant decrease in mean angle, length, and decrease in water disorder (Madan and Sharp, 1996). Hydroxyl/phosphate oxygens have both the smallest radius and the largest magnitude of partial charge. They interact more strongly with water via electrostatic interactions and consequently increase the H-bond distortion. The nitrogens lie somewhere in between, with a radius and average partial charge magnitude between that of C and O. Although the direction of their effect on the H-bond structure cannot be predicted a priori, the simulations reveal that, qualitatively, many base nitrogens affect water structure in a similar fashion to carbons; i.e., there is a decrease in H-bond distortion.

Although the simulations we have described are well within the capabilities of available computing resources, they are far from trivial. It is also not likely, for technical reasons, that measured hydration *Cp* values for the nucleic acid constituents will be available in the near future. Thus, in an attempt to provide approximate parameters that may be of use with the ASA model (whose computational requirements are minimal), we have analyzed our results in terms of accessible surface areas (Eqs. 3 and 4). Of course it is known that there are limitations to area-based models, which we recognize in our own application of Eq. 1 to the analysis of the random network data. First, according to the random network model, the origin of heat capacity changes is distortion of H-bond structure; thus, it depends on the number of H-bonds perturbed by the solute, not the number of waters solvating the solute. Inherent in the random network model are contributions that do not scale with area. Consider two neighboring atoms A and B. There are three classes of solvating water-water H-bonds: AA, BB, and AB. Even if the number in the AA and BB classes scales with the ASA of the A and B atoms, there is still the AB class, which does not scale with the area of either A or B. In fact our results show that there are relatively few H-bonds around very polar oxygen atoms (Table 2), and on a per area basis these solute atoms have less water-water H-bonds in their first hydration shell than the apolar carbons. There is a moderate correlation between the total area of the solute and the total number of first-shell water-water H-bonds (Table 4), but because the heat capacity change depends on the extent of H-bond distortion as well as the number, and because this distortion is not the same for all the solutes' groups, the correlation between the number of H-bonds and the predicted hydration *Cp* is lower still, at $R^2 = 0.69$.

Second, our simulations indicate that not all polar groups are equivalent, and extracting a single area coefficient for the polar area has the effect of averaging out these differences. Most notably, in our simulations, the base-pairing groups have more polar character regarding their effect on water structure and their hydration heat capacity than, say, other base nitrogen atoms. This has implications for analyzing the heat capacity changes in processes where different portions of the DNA surface are buried. Melting of duplex DNA is accompanied by a larger increase in *Cp* (Chalikian et al., 1999), as would be expected from the exposure of large amounts of apolar surface of the bases (as well as from expected increases in fluctuations of conformational enthalpy), but Holbrook et al. (1999) found that there was relatively little heat capacity change from association of two already base-stacked helical DNA strands to form a duplex, which they attributed to a cancellation of the positive contribution from apolar surface burial by negative contributions from polar surface burial. This is consistent with our finding that the base-pairing groups are the most polar, in terms of hydration heat capacity, because the area buried upon duplex association would contain a large contribution from these groups. Binding of proteins to fully duplex DNA, in contrast, would involve burial of polar groups that, in our simulations, have a smaller negative, or even positive, hydration heat capacity, implying a larger decrease in *Cp* per unit area upon binding.

Third, we have been able to study only the constituent fragments of nucleic acids to date, so neighbor effects in entire nucleotides between base, sugar, and phosphate groups and between nucleotides in polynucleotides remain to be determined. Equation 4 should thus be applied with caution, recognizing its highly empirical basis and its neglect of heterogeneous solute group effects.

In summary, the results of the simulations presented here strongly suggest that the hydration heat capacity of nucleic acids is considerably more positive than previously thought. This in turn would provide an additional source of negative heat capacity change when proteins and ligands bind specifically to DNA or RNA, providing another reason for the large negative ΔCp observed in many of these binding reactions. We also provide approximate parameters for nucleotides for use in the accessible-area model for heat capacity changes, with the caveat that not all polar atoms have the same effect. We are currently using these results to reanalyze the area changes for DNA-ligand binding reactions for which large heat capacity changes have been observed. Another future direction for this research is the simulation of the water structure around entire nucleotides and, ultimately, entire DNA and DNA/ligand complexes. This will require considerably more computation than the simulations described here, but it will provide a more detailed picture of the water structure and its effect on the hydration heat capacity, including a full account of neighbor effects.

Financial support is acknowledged from National Institutes of Health (GM54105).

REFERENCES

Bondi, A. 1968. Molecular Crystals, Liquids, and Glasses. John Wiley and Sons, New York.

Cabani, S., P. Gianni, V. Mollica, and L. Lepori. 1981. Group contributions to the thermodynamic properties of non-ionic organic solutes in dilute aqueous solution. *J. Solution Chem.* 10:563–595.

Chalikian, T., J. Volker, G. Plum, and K. Breslauer. 1999. A more unified picture for the thermodynamics of nucleic acid duplex melting: a characterization by calorimetric and volumetric techniques. *Proc. Natl. Acad. Sci. U.S.A.* 96:7853–7858.

Connelly, P. R., J. A. Thomson, M. J. Fitzgibbon, and F. J. Bruzzese. 1993. Probing hydration contributions to the thermodynamics of ligand binding by proteins. Enthalpy and heat capacity changes of tacrolimus and rapamycin binding to FK506 binding protein in D2O and H_2O. *Biochemistry.* 32:5583–5590.

Eisenberg, D., and A. D. McLachlan. 1986. Solvation energy in protein folding and binding. *Nature.* 319:199.

Freire, E. 1995. Forces and factors that contribute to the structural stability of membrane proteins. *Biochim. Biophys. Acta.* 1241:295–322.

Gallagher, K., and K. A. Sharp. 1998. Electrostatic contributions to heat capacity changes of DNA-ligand binding. *Biophys. J.* 75:769–776.

Henn, A. R., and W. Kauzmann. 1989. Equation of state of a random network, continuum model of liquid water. *J. Phys. Chem.* 93: 3770–3783.

Holbrook, J., M. Capp, R. Saecker, and M. Record. 1999. Enthalpy and heat capacity changes for formation of an oligomeric DNA duplex: interpretation in terms of coupled processes of formation and association of single-stranded helices. *Biochemistry.* 38:8409–8422.

Jin, L., J. Yang, and J. Cary. 1993. Thermodynamics of ligand binding to trp repressor. *Biochemistry.* 32:7302–7309.

Jorgensen, W. L. 1992. BOSS, Version 3.3, Copyright Yale University, New Haven, CT.

Jorgensen, W. L., J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein. 1983. Comparison of simple potential functions for simulating liquid water. *J Chem. Phys.* 79:926.

Ladbury, J. E., J. G. Wright, J. M. Sturtevant, and P. B. Sigler. 1994. A thermodynamic study of the *trp* repressor-operator interaction. *J. Mol. Biol.* 238:669–681.

Lundback, T., C. Cairns, J. Gustafsson, J. Carlstedt-Duke, and T. Hard. 1993. Thermodynamics of gluticorticoid receptor-DNA interaction. *Biochemistry.* 32:5074–5082.

Lundback, T., J. F. Chang, K. Phillips, B. Luisi, and J. E. Ladbury. 2000. Characterization of sequence-specific DNA binding by the transcription factor Oct-1. *Biochemistry.* 39:7570–7579.

Madan, B., and K. Sharp. 1999. Changes in water structure induced by a hydrophobic solute probed by simulation of the water hydrogen bond angle and radial distribution functions. *Biophys. Chem.* 78:33–41.

Madan, B., and K. A. Sharp. 1996. Heat capacity changes accompanying hydrophobic and ionic solvation: a Monte-Carlo and random network model study. *J. Phys. Chem.* 100:7713–7721.

Madan, B., and K. A. Sharp. 1997. Molecular origin of hydration heat capacity changes of hydrophobic solutes: perturbation of water structure around alkanes. *J. Phys. Chem.* 101:11237–11242.

Madan, B., and K. A. Sharp. 2000. Heat capacity changes accompanying hydrophobic and ionic solvation: a Monte-Carlo and random network model study. *J. Phys. Chem.* 105:2256.

Makhatadze, G., and P. Privalov. 1990a. I. Partial molar heat capacity of individual amino acid residues in aqueous solution: hydration effect. *J. Mol. Biol.* 213:375–384.

Makhatadze, G., and P. Privalov. 1990b. II. Partial molar heat capacity of individual amino acid residues in aqueous solution: protein unfolding effects. *J. Mol. Biol.* 213:385–391.

Matulis, D., I. Rouzina, and V. A. Bloomfield. 2000. Thermodynamics of DNA binding and condensation: isothermal titration calorimetry and electrostatic mechanism. *J. Mol. Biol.* 296:1053–1063.

Merabet, E., and G. Ackers. 1995. Calorimetric analysis of the lambda cI repressor binding to DNA operator sites. *Biochemistry.* 34:8554–8563.

Morton, C. J., and J. E. Ladbury. 1996. Water mediated protein-DNA interactions: the relationship of thermodynamics to structural detail. *Protein Sci.* 5:2115–2118.

Murphy, K., and E. Freire. 1992. Thermodynamics of structural stability and cooperative folding behavior in proteins. *Adv. Protein Chem.* 43: 313–361.

Murphy, K. P., and S. J. Gill. 1991. Solid model compounds and the thermodynamics of protein unfolding. *J. Mol. Biol.* 222

Murphy, K. P., P. L. Privalov, and S. J. Gill. 1990. Common features of protein unfolding and dissolution of hydrophobic compounds. *Science.* 247:559–561.

Myers, J. K., C. N. Pace, and J. M. Scholtz. 1995. Denaturant m values and heat capacity changes: relation to changes in accessible surface areas of protein unfolding. *Protein Sci.* 4:2138–2148.

Oda, M., K. Furukawa, K. Ogata, A. Sarai, and H. Nakamura. 1998. Thermodynamics of specific and non-specific DNA binding by the c-myb DNA-binding domain. *J. Mol. Biol.* 276:571–590.

Pranata, J., S. Wierschke, and W. Jorgensen. 1991. The OPLS potential function for nucleotide bases. *JACS.* 113:2810–2819.

Privalov, P. L., and S. J. Gill. 1988. Stability of protein structure and hydrophobic interaction. *Adv. Protein. Chem.* 39:191–234.

Ren, J., T. C. Jenkins, and J. B. Chaires. 2000. Energetics of DNA intercalation reactions. *Biochemistry.* 39:8439–8447.

Robertson, A. D., and K. P. Murphy. 1997. Protein structure and the energetics of protein stability. *Chem. Rev.* 97:1251–1267.

Rouzina, I., and V. A. Bloomfield. 1999. Heat capacity effects on the melting of DNA. 1. General aspects. *Biophysical J.* 77:3242.

Sharp, K. A., and B. Madan. 1997. The hydrophobic effect, water structure and heat capacity changes. *J. Phys. Chem.* 101:4343–4348.

Spolar, R., J. Livingstone, and M. T. Record. 1992. Use of liquid hydrocarbon and amide transfer data to estimate contributions to protein folding. *Biochemistry.* 31:3947–3955.

Spolar, R. S., and M. T. Record. 1994. Coupling of local folding to site-specific binding of proteins to DNA. *Science.* 263:777–784.

Sridharan, S., A. Nicholls, and B. Honig. 1992. A new vertex algorithm to calculate solvent accessible surface areas. *Biophys. J.* 61:(Abstract 995).

Sturtevant, J. 1977. Heat capacity and entropy changes in processes involving proteins. *Proc. Natl. Acad. Sci. U. S. A.* 74:2236–2240.

Vanzi, F., B. Madan, and K. Sharp. 1998. Effect of the protein denaturants urea and guanidinium on water structure: a structural and thermodynamic study. *JACS.* 120:10748–10753.