J. Am. Chem. Soc. 1993, 115, 7930–7931

Relative Hydration Free Energies of Nucleic Acid Bases

A. H. Elcock and W. G. Richards*

Physical Chemistry Laboratory, Oxford University South Parks Road, Oxford OX1 3QZ, United Kingdom

Received April 14, 1993

The relative hydration free energies of nucleic acid bases are crucial to any rationalization of nucleic acid stability.[1] Direct measurement encounters the problems of volatility, so reliance has to be placed on values derived from computer simulation.[2] Here the free energy perturbation technique (FEP)[3] is used with the carefully developed OPLS[4] (optimized potentials for liquid simulations) parameters to derive relative hydration free energies for the five naturally occurring bases. One check on reliability comes from computing the free energy change over a complete cyclical path when perturbing one molecule into another. The overall free energy change, which should be 0, here is within 1 kcal mol^{-1}, but there are large deviations from earlier simulations.[5]

Five perturbations were carried out according to the thermodynamic cycle shown in Figure 1.

Mutating a purine to a pyrimidine clearly involves a very large change in geometry, and so the use of FEP may not seem at all obvious. However, in this work we developed a novel mutation method which allows any perturbations of this type to proceed in a very smooth manner by effectively shrinking the two rings into one. This is illustrated for the 9-methylguanine to 1-methylcytosine perturbation in Figure 2. The perturbation is achieved by mapping out the pyrimidinic ring onto the purine template and mutating the intervening ring atoms to dummies so as to maintain overall topology. This use of endocyclic dummy atoms is expected to reduce significantly problems with end-point sampling. In addition, since, in contrast to previous approaches,[6,7] there is no disturbance of the glycosidic bond, extension of this scheme to DNA simulations is straightforward.

Molecular dynamics (MD) simulations were carried out within the AMBER4.0 suite of programs.[8] The OPLS parameter set was used to represent all nonbonded interactions. All intramolecular bond, angle, and dihedral terms for the 1-methylpyrimidines and 9-methylpurines were assigned standard values.[9] In each case, the solute was immersed in a box of 506 TIP3P[10] water molecules and simulations were conducted in the NPT ensemble at 298 K and 1 atm pressure. Bond lengths were constrained using the SHAKE algorithm,[11] and dihedrals involving only ring atoms were fixed by the application of holonomic constraints.[12] The intermolecular contribution of these constraints was calculated as implemented in AMBER4.0. A time step of 2 fs was used, and a 9-Å cutoff was applied to nonbonded interactions.

(1) Saenger, W. *Principles of Nucleic Acid Structure*; Springer-Verlag: New York, 1984.
(2) Cullis, P. M.; Wolfenden, R. *Biochemistry* **1981**, *20*, 3024–3028.
(3) Beveridge, D. L.; Di Capua, F. M. *Annu. Rev. Biophys. Biophys. Chem.* **1989**, *18*, 431–492.
(4) Pranata, J.; Wierschke, S. G.; Jorgensen, W. L. *J. Am. Chem. Soc.* **1991**, *113*, 2810–2819.
(5) Bash, P. A.; Singh, U. C.; Langridge, R.; Kollman, P. A. *Science* **1987**, *236*, 564–568.
(6) Ferguson, D. M.; Pearlman, D. A.; Swope, W. C.; Kollman, P. A. *J. Comput. Chem.* **1992**, *13*, 362–370.
(7) Wunz, T. P. *J. Comput. Chem.* **1992**, *13*, 667–673.
(8) Pearlman, D. A.; Case, D. A.; Caldwell, J. C.; Seibel, G. L.; Singh, U. C.; Weiner, P. A.; Kollman, P. A. *AMBER4.0* (UCSF); Department of Pharmaceutical Chemistry: University of California, San Francisco, CA, 1991.
(9) Weiner, S. J.; Kollman, P. A.; Nguyen, D. T.; Case, D. A. *J. Comput. Chem.* **1986**, *7*, 230–252.
(10) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926–935.
(11) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327–341.
(12) Tobias, D. J.; Brooks, C. L., III *J. Chem. Phys.* **1988**, *89*, 5115–5127.

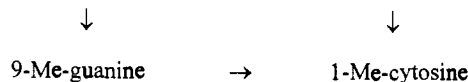

Figure 1. Thermodynamic cycle employed for calculation of relative hydration free energies.

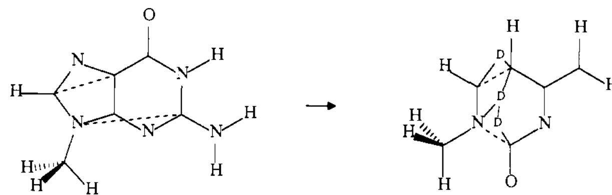

Figure 2. Schematic illustration of guanine-to-cytosine mutation. Additional bonds used to give correct pyrimidine geometry are indicated by dotted lines. Dummy atoms used to maintain topology between the two structures are represented by D. Ring dihedrals were constrained to maintain a planar geometry throughout the simulation.

The perturbations were performed over a minimum of 21 windows, each consisting of 6-ps equilibration, 9-ps data collection for the A→G, U→T, and T→C mutations and 10-ps equilibration, 10-ps data collection for the A→U and G→C mutations. In some cases, the larger ΔG values associated with changes near the end points required windows to be further subdivided to improve convergence. Table I presents the computed free energy changes in both directions along with the total simulation length for each perturbation.

Closure of the thermodynamic cycle gives an overall error of 1.0 kcal mol^{-1}, which, given the size of the perturbations, is highly satisfactory and inspires considerable confidence in the results. Equal distribution of this error between all perturbations results in the relative hydration free energies given in Table II.

The FEP results reported here are in close correspondence with those obtained using a finite-difference Poisson–Boltzmann (FDPB) method for the same charge set[13] (column 2). Good agreement between FDPB and FEP results has been noted previously for simple polar solutes[14] and suggests that, at least for the OPLS parameter set, electrostatics play a dominant role in determining the relative ordering of hydration free energies. It is interesting to note that no such agreement is observed between FEP and FDPB results for the AMBER parameter set.[13] The origins of such large discrepancies are not obvious, as it seems unlikely that they would be due simply to nonelectrostatic factors.

Agreement with the predictions of the semiempirical AM1-SM2 solvation model[15] is poor (column 3), the reasons for which are not clear. The AM1-SM2 model's use of an empirical expression to account for all cavitation, dispersion, and first-hydration-shell effects may contribute to the differences obtained, as in the case of thymine such contributions account for 37% of the absolute hydration free energy.[16] The absence of energy breakdowns for the other bases prevents further examination of this point. Closer correspondence with the AM1-SM2 results is found for the AMBER FEP simulations, although this should be viewed in light of the following discussion.

The two sets of FEP results are strikingly dissimilar, especially with regard to the relative energies of adenine, cytosine, and thymine (columns 4 and 5). The magnitude of these differences is surprising, as the charge sets are similar and both give dipole moments in good agreement with the experimental gas-phase

(13) Mohan, V.; Davis, M. E.; McCammon, J. A.; Pettitt, B. M. *J. Phys. Chem.* **1992**, *96*, 6428–6431.
(14) Arald, J. C.; Nicholls, A.; Sharp, K.; Honig, B.; Tempczyk, A.; Hendrickson, T. F.; Still, W. C. *J. Am. Chem. Soc.* **1991**, *113*, 1454–1455.
(15) Cramer, C. J.; Truhlar, D. G. *Science* **1992**, *256*, 213–217.
(16) Cramer, C. J.; Truhlar, D. G. *Chem. Phys. Lett.* **1992**, *198*, 74–80.

0002-7863/93/1515-7930$04.00/0 © 1993 American Chemical Society

Table I. Computed Free Energy Changes (kcal mol⁻¹)

perturbation	simulation length (ps)	forward	reverse
A→U	460	−1.2	−0.9
U→T	300	−0.3	0.3
T→C	300	−7.0	−6.7
A→G	400	−10.2	−10.4
G→C	600	1.1	1.6

Table II. Relative Hydration Free Energies (kcal mol⁻¹)

solute	FDPB[a]	AM1-SM2[b]	AMBER FEP[c]	OPLS FEP[d]
9-Me-adenine	0.0	0.0	0.0	0.0
1-Me-cytosine	−6.0	2.2	−0.1	−8.5
9-Me-guanine	−8.9	−3.4	−7.0	−10.1
1-Me-thymine	0.4	7.6	5.1	−1.5
1-Me-uracil		6.1		−1.2

[a] Reference 13. [b] Reference 16. [c] Reference 5. [d] This work.

values,[4] although the slightly greater polarity of the AMBER parameter set is reflected in the more negative hydration free energies obtained with the FDPB method (Table IV of ref 13).

Given the similarity of the parameter sets, the most likely explanation of the differences in FEP results is as follows. The AMBER FEP simulations involved adenine–guanine and thymine–cytosine mutations. The results reported here for these perturbations are in reasonable agreement with those of the earlier simulations. The discrepancies arise, then, in relating the purine hydration free energies to those of the pyrimidines. In the present work, this was accomplished by direct perturbation of purine to pyrimidine with a simulation length of at least 460 ps. In the previous work, the relationship was established by annihilating adenine and thymine to methane over a period of 60 ps.[5] Recent FEP work has shown that the limited sampling of configuration space associated with short simulation lengths can be a significant source of error in the calculation of free energies.[17,18] The relatively short simulation times of the earlier work, coupled with the rather drastic nature of the mutations, suggests that the results may thus be subject to considerable statistical error. On the other hand, the long simulation times used in this work, together with the small overall error and the close correspondence with FDPB results, argue strongly for the precision of the present results. It should be noted that if sampling problems are the source of the discrepancies, then no real conclusion as to the relative merits of the two parameter sets can be drawn, as the differences obtained are of a statistical rather than a systematic nature. In fact, the good agreement obtained for the adenine–guanine and thymine–cytosine relative hydration free energies would seem to suggest that there should be no great difference between the two parameter sets.

Experimental determination of the relative hydration free energies has in the past proved impossible due to problems of volatility,[2] yet this would obviously be the best way to resolve the discrepancies between theoretical estimates. Until such experiments are performed, the results presented here probably represent the best available estimates of the hydration free energies of the nucleic acid bases.

Acknowledgment. This work was conducted in part pursuant to a contract with the National Foundation for Cancer Research.

(17) Pearlman, D. A.; Kollman, P. A. *J. Chem. Phys.* **1991**, *94*, 4532–4545.
(18) Mitchell, M. J.; McCammon, J. A. *J. Comput. Chem.* **1991**, *12*, 271–275.