Stacking Free Energy Profiles for All 16 Natural Ribodinucleoside Monophosphates in Aqueous Solution†

Jan Norberg and Lennart Nilsson*

Contribution from the Center for Structural Biochemistry, Department of Biosciences at Novum, Karolinska Institute, S-141 57 Huddinge, Sweden

Received June 30, 1995®

Abstract: The phenomenon of base stacking, which we treat here in full atomic detail for RNA fragments including solvent and the sugar phosphate backbone, is not yet completely understood. We present free energy profiles of stacking for all 16 natural ribodinucleoside monophosphates based on potential of mean force calculations for this unimolecular process. Stacking preferences follow the general sequence purine−purine > purine−pyrimidine ≥ pyrimidine−purine > pyrimidine−pyrimidine with the stacked state having 2−6 kcal/mol lower free energy than the unstacked states for purine−purine combinations. The ribodinucleoside monophosphates containing both a purine and a pyrimidine base showed in most cases higher stacking propensities if the purine base was in the 5′-position. No or a very small free energy barrier was observed for the unstacking of the pyrimidine−pyrimidine ribodinucleoside monophosphates. A large number of different structures, both stacked and unstacked, were obtained and major transitions of the backbone torsion angles were observed. The pseudorotation phase angle showed a preference for the ^{2}E mode in the unstacked states of ribodinucleoside monophosphates containing pyrimidine−pyrimidine combinations. The strongest interactions with solvent were observed for the phosphate oxygens, and the direct interactions of the 2′OH group of the 5′-ribose with the 3′-base and ribose destabilize the stacked state by about 2 kcal/mol. Both the solvent and the sugar phosphate backbone influence the stacking preferences of the dimers.

Introduction

The three-dimensional structures of DNA and RNA are to a large extent determined by base stacking. In contrast to the hydrogen bonding responsible for base pairing, stacking interactions are critically dependent on the aromatic base sequence. The sequence thus has a means, in addition to the functional groups on the base edges, of influencing interactions with other molecules. Base stacking has been studied in a variety of systems, ranging from isolated bases to long DNA duplexes. Even in the simplest nucleic acid system in which sequence dependent effects can be studied, a ribodinucleoside monophosphate, where a large number of experimental studies have been performed, a diversity of results is obtained.[1] Semiempirical calculations, molecular mechanics, molecular dynamics simulations, and free energy perturbation methods have been used to investigate base stacking;[2−6] most of these theoretical studies[2−4,6] of base stacking have focused on the base−base interactions but have left out the dynamical backbone and/or the influence of solvent interactions. In an earlier potential of mean force calculation of the stacking of a 9-methyladenine base and a 1-methylthymine base, the backbone was not included and the relative orientation of the bases was not allowed to vary.[7] The mechanism of base stacking is not yet completely understood, and different forces, such as hydrophobic effects or direct electrostatic interactions between the bases, have been suggested as making the dominant contribution to the stabilization of base stacks.[8] Proton nuclear magnetic resonance and spin−lattice relaxation time measurements have provided details of the conformational properties of ribodinucleoside monophosphates.[9−16] Optical spectroscopy experiments such as circular dichroism, optical rotatory dispersion, and UV absorption have been applied to determine thermodynamic parameters, such as the fraction that are in a stacked state, for the stacking interactions of ribodinucleoside monophosphates.[1,17−21] For the ribodinucleoside monophosphates, the experimental enthalpy ΔH appears to range from about −2 to −12 kcal/mol. Solvent interactions are known to influence the nucleic acid conforma-

* Author to whom correspondence should be addressed.

† Abbreviations: ApA, adenylyl-3′,5′-adenosine; ApC, adenylyl-3′,5′-cytidine; ApG, adenylyl-3′,5′-guanosine; ApU, adenylyl-3′,5′-uridine; CpA, cytidylyl-3′,5′-adenosine; CpC, cytidylyl-3′,5′-cytidine; CpG, cytidylyl-3′,5′-guanosine; CpU, cytidylyl-3′,5′-uridine; GpA, guanylyl-3′,5′-adenosine; GpC, guanylyl-3′,5′-cytidine; GpG, guanylyl-3′,5′-guanosine; GpU, guanylyl-3′,5′-uridine; UpA, uridylyl-3′,5′-adenosine; UpC, uridylyl-3′,5′-cytidine; UpG, uridylyl-3′,5′-guanosine; UpU, uridylyl-3′,5′-uridine; MD, molecular dynamics; PMF, potential of mean force.

® Abstract published in *Advance ACS Abstracts,* October 15, 1995.

(1) Frechet, D.; Ehrlich, R.; Remy, P.; Gabarro-Arpa, J. *Nucleic Acids Res.* **1979**, *7*, 1981−2001.
(2) Aida, M. *J. Theor. Biol.* **1988**, *130*, 327−335.
(3) Hunter, C. A. *J. Mol. Biol.* **1993**, *230*, 1025−1054.
(4) Langlet, J.; Claverie, P.; Caron, F.; Boeuve, J. C. *Int. J. Quantum Chem.* **1981**, *19*, 299−338.
(5) Norberg, J.; Nilsson, L. *Biophys. J.* **1994**, *67*, 812−824.
(6) Cieplak, P.; Kollman, P. A. *J. Am. Chem. Soc.* **1988**, *110*, 3734−3739.
(7) Dang, L. X.; Kollman, P. A. *J. Am. Chem. Soc.* **1990**, *112*, 503−507.
(8) Newcomb, L. F.; Gellman, S. H. *J. Am. Chem. Soc.* **1994**, *116*, 4993−4994.
(9) Ezra, F. S.; Kondo, N. S.; Ainsworth, C. F.; Danyluk, S. S. *Nucleic Acids Res.* **1976**, *3*, 2549−2562.
(10) Kondo, N. S.; Danyluk, S. S. *Biochemistry* **1976**, *15*, 756−768.
(11) Chachaty, C.; Perly, B.; Forchioni, A.; Langlet, G. *Biopolymers* **1980**, *19*, 1211−1239.
(12) Tran-Dinh, S.; Neumann, J. M.; Borrel, J. *Biochim. Biophys. Acta* **1981**, *655*, 167−180.
(13) Neumann, J.-M.; Guschlbauer, W.; Tran-Dinh, S. *Eur. J. Biochem.* **1979**, *100*, 141−148.
(14) Chachaty, C.; Yokono, T.; Tran-Dinh, S.; Guschlbauer, W. *Biophys. Chem.* **1977**, *6*, 151−159.
(15) Ezra, F. S.; Lee, C.-H.; Kondo, N. S.; Danyluk, S. S.; Sarma, R. H. *Biochemistry* **1977**, *16*, 1977−1987.
(16) Lee, C.-H.; Ezra, F. S.; Kondo, N. S.; Sarma, R. H.; Danyluk, S. S. *Biochemistry* **1976**, *15*, 3627−3639.
(17) Brahms, J.; Maurizot, J. C.; Michelson, A. M. *J. Mol. Biol.* **1967**, *25*, 481−495.
(18) Davis, R. C.; Tinoco, I., Jr. *Biopolymers* **1968**, *6*, 223−242.
(19) Guschlbauer, W.; Frič, I.; Holý, A. *Eur. J. Biochem.* **1972**, *31*, 1−13.
(20) Daniels, M.; Shaar, C. S.; Morgan, J. P. *Biophys. Chem.* **1988**, *32*, 229−237.
(21) Kang, H.; Chou, P.-J.; Johnson, W. C., Jr.; Weller, D.; Huang, S.-B.; Summerton, J. E. *Biopolymers* **1992**, *32*, 1351−1363.

0002-7863/95/1517-10832$09.00/0

tion, and stacked structures are most favored in aqueous solvents.[22-24]

We have used molecular dynamics (MD) simulations, which have the potential of providing very detailed information about interactions in macromolecular systems, to make as realistic a theoretical study as possible of the stacking problem. MD simulations were performed on all atom models of all 16 ribodinucleoside monophosphates containing combinations of the adenine (A), cytosine (C), guanine (G), and uracil (U) bases in water. This approach has previously been shown to be fruitful in very long molecular dynamics simulations used in investigations of the conformational flexibility and solvent interactions of stacked and unstacked conformations of the ribodinucleoside monophosphate GpU.[5,25] In the present study, the free energy of going from a stacked state to an unstacked was determined by potential of mean force (PMF) calculations along a reaction coordinate defined by the distance between the glycosidic nitrogen atoms of the bases. This reaction coordinate was chosen because it is a reasonable and simple criterion for stacking,[5] and at the same time allows for flexibility of both the bases and the backbone. Conformational changes, in terms of backbone torsions, sugar puckering, and base orientation as well as solvent binding sites and energetic contributions of the ribodinucleoside monophosphates, were analyzed for stacked and unstacked states.

The potential of mean force approach provides an estimate of the free energy difference between stacked and unstacked states and also gives the free energy profile for this process along the chosen reaction coordinate. This is obtained with a standard molecular mechanics potential energy function, without any adjustable parameters for these particular calculations. Together with the structural information for solute and solvent also available in the simulations we thus get a very complete description of the unimolecular stacking/unstacking equilibrium and its dependence on base sequence in ribodinucleoside monophosphates.

Methods

The initial stacked conformations of the ribodinucleoside monophosphates were generated as standard A-RNA, which has a base tilt of 16° and an axial rise per residue of 2.81 Å,[26] from X-ray fiber diffraction data.[27] Combinations of the adenine (A), cytosine (C), guanine (G), and uracil (U) bases were used to build the 16 ribodinucleoside monophosphates (MpN, where M, N = A, C, G, U). The distance between the nitrogen atoms (N1 for pyrimidine and N9 for purine), which connect the base to the sugar moiety, was the same 4.736 Å in all the 16 stacked conformations thus generated.

Potential Energy Function. All energy minimizations and MD simulations were performed using the CHARMM[28] program version 22 with the new all atom nucleic acid parameters.[29] The potential energy function, $U(\bar{R})$, where $\bar{R}$ is the vector of the coordinates of the atoms has the form

$$U(\bar{R}) = \sum_{\text{bonds}} {}^1/_2K_b(b - b_0)^2 + \sum_{\text{angle}} {}^1/_2K_\theta(\theta - \theta_0)^2 + \sum_{\text{Urey-Bradley}} {}^1/_2K_{UB}(S - S_0)^2 + \sum_{\substack{\text{dihedral}\\ \text{angles}}} K_\phi(1 + \cos(n\phi - \delta)) + \sum_{\substack{\text{improper}\\ \text{torsions}}} {}^1/_2K_\omega(\omega - \omega_0)^2 + \sum_{\substack{\text{nonbond}\\ i,j \text{ pairs}}} \epsilon_{ij}\left[\left(\frac{R_{\min\,ij}}{r_{ij}}\right)^{12} - \left(\frac{R_{\min\,ij}}{r_{ij}}\right)^{6}\right] + \frac{q_iq_j}{4\pi\epsilon_0\epsilon_r r_{ij}}$$

where K_b, K_θ, K_{UB}, K_ϕ, and K_ω are the bond, bond angle, Urey–Bradley, dihedral angle, and improper dihedral angle force constants, respectively. The quantity b is the bond length, θ the bond angle, S the Urey–Bradley 1,3 distance, and ω the improper dihedral angle, and the subscript zero is used for the equilibrium values. ϕ is the dihedral angle, n the multiplicity, and δ the phase. Nonbonded interactions, with the nonbonded distance r_{ij} between atoms i and j, are represented with a Lennard-Jones term, where ϵ_{ij} is the well depth and $R_{\min\,ij}$ is the distance to the Lennard-Jones minimum, and a Coulomb term, where q_i is the partial atomic charge, ϵ_0 is the dielectric constant, and ϵ_r is the relative dielectric constant.

Computational Details. To make the system electrically neutral, a sodium counterion was placed on the bisector of the phosphate oxygens. Each ribodinucleoside monophosphate with a sodium counterion was first energy minimized 100 cycles of steepest descent and the counterion moved away from the phosphate oxygens nearly symmetrically. Thereafter each ribodinucleoside monophosphate was solvated in a 25.0 Å side cubic box of TIP3P[30] water molecules. To remove bad contacts, the water molecules were energy minimized 100 cycles of steepest descent and 3000 cycles of adopted-basis set Newton–Raphson,[28] while the ribodinucleoside monophosphate was constrained using a harmonic potential with a force constant of 1.0 kcal·mol^{-1}·Å^{-2}. The integration of the equations of motion was carried out with the Verlet algorithm,[31] and to allow for a time step of 0.002 ps, all hydrogen atom-heavy atom bond lengths were constrained using the SHAKE algorithm.[32] Periodic boundary conditions were used in the MD simulations. All the MD simulations were carried out at a temperature of 300 K. A relative dielectric constant $\epsilon_r = 1.0$ was used, and the nonbonded interactions were smoothly shifted to zero at a cutoff of 11.5 Å.[28] Coordinates were saved every 0.04 ps, and the nonbonded list was updated every 20 steps. All the analysis, energy minimizations, and MD simulations were performed on IBM RISC 6000 and DEC AXP 3000/400 workstations.

Stacking Model. We used $R_{N_xN_y}$, the distance between the base glycosidic nitrogen atoms (N_x, N_y = N1-pyrimidine, N9-purine), as a reaction coordinate (Figure 1). This was implemented in the umbrella sampling procedure and gives a one-dimensional definition of stacking, which in geometric terms depends on base–base parallelism, overlap, and distance.

Other, more complex coordinates can be devised, but the advantage of using this simple reaction coordinate is that the bases and the backbone still can be flexible. Base stacking was in a previous MD simulation[5] of a dinucleoside monophosphate found to be well characterized by $R_{N_xN_y}$, and in the present investigation, a large number of different conformations was obtained by sampling along this reaction coordinate.

Potential of Mean Force. The potential of mean force $w(R)$ is the free energy, in our case the Helmholtz free energy since the simulations were run at constant volume, along a reaction coordinate R and takes the form

$$w(R) = -k_BT \ln \rho^*(R) - E_{\text{restr}}(R)$$

where k_B is the Boltzmann constant, T is the temperature, and $\rho^*(R)$ is the probability distribution along the reaction coordinate in the presence of the biasing potential $E_{\text{restr}}(R)$. Sampling of a reaction coordinate is

(22) Cantor, R. C.; Schimmel, P. R. *Biophysical Chemistry*; Freeman: San Francisco, CA, 1980.
(23) Pohorille, A.; Burt, S. K.; MacElroy, R. D. *J. Am. Chem. Soc.* **1984**, *106*, 402–409.
(24) Lowe, M. J.; Schellman, J. A. *J. Mol. Biol.* **1972**, *65*, 91–109.
(25) Norberg, J.; Nilsson, L. *Chem. Phys. Lett.* **1994**, *224*, 219–224.
(26) Saenger, W. *Principles of Nucleic Acid Structure*; Springer-Verlag: New York, 1988.
(27) Arnott, S.; Smith, P. J. C.; Chandrasekaran, R. In *CRC Handbook of Biochemistry; Molecular Biology: Nucleic Acids*, 3rd ed.; Fasman, G. D., Ed., CRC Press: Cleveland, OH, 1976; Vol. 2, pp 411–469.
(28) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187–217.
(29) MacKerell, A. D., Jr.; Wiórkiewicz-Kuczera, J.; Karplus, M. *J. Am. Chem. Soc.*, in press.
(30) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926–935.
(31) Verlet, L. *Phys. Rev.* **1967**, *159*, 98–103.
(32) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327–341.

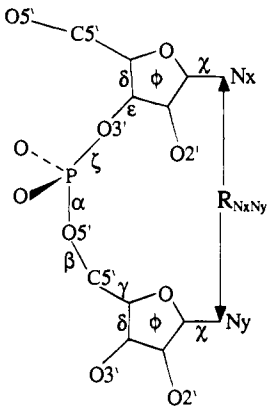

Figure 1. Description of the reaction coordinate, the restrained distance $R_{N_xN_y}$ between the nitrogen atoms (N_x, N_y = N1-pyrimidine, N9-purine), which connect the base to the sugar moiety of the backbone. The definitions of the backbone torsion angles (δ, ϵ, ζ, α, β, γ), the glycosidic torsion angles (χ), and the pseudorotation phase angles (ϕ) are given.

not usually adequate in a single unrestrained MD simulation; therefore, it is common to obtain a potential of mean force from a series of simulations, which use different restraining potentials to sample the reaction coordinate.[33] Potential of mean force profiles were generated using this umbrella sampling procedure with a harmonic restraining potential

$$E_{restr} = k(R_{N_xN_y} - R_{ref})^2$$

where k is the force constant and R_{ref} is the reference value used in a particular simulation window to restrain the distance $R_{N_xN_y}$ between the base nitrogen atoms (N_x, N_y = N1-pyrimidine, N9-purine), which connect the base to the sugar moiety of the backbone.

The force constant was set to 4.0–16.0 kcal·mol^{-1}·Å^{-2} such that neighboring sampling windows overlap. An equilibration period of 40 ps was performed for each of the initial 16 systems containing an aqueous solvated ribodinucleoside monophosphate with a sodium counterion. During this equilibration period, R_{ref} was set to 4.5 Å and a force constant of 4.0 kcal·mol^{-1}·Å^{-2} was used. Conformational space was sampled by performing MD simulations in 18 windows with R_{ref} varying from 3.5 to 12.0 Å in 0.5 Å intervals. In each window, the simulation was started from a snapshot of the preceding window, and after 5 ps of equilibration, conformations were sampled during 40–120 ps. The longer simulation times were required in the $R_{N_xN_y}$ = 5–10 Å range where the ribodinucleoside monophosphates are in transition between stacked and unstacked conformations. From a free MD simulation[5] of an unstacked conformation of a dinucleoside monophosphate, the majority of the conformations was obtained in the 6–9 Å range. The total MD simulation time for all 16 ribodinucleoside monophosphates was 18.4 ns (about 8000 CPU hours). The PMF profiles from the different simulation windows were spliced together with the weighted histogram analysis method[34,35] (WHAM). From prolonged simulations of GpU we estimate the error in the PMF profiles to be ±0.5 kcal/mol.

Fraction Stacked. To determine the fraction stacked, a distinction has to be made between stacked and unstacked conformations, and we choose to designate all states with $R_{N_xN_y} \leq R_{stack}$ as stacked, the rest being unstacked. The fraction stacked was calculated by integrating the PMF profile $w(R)$ to R_{stack}:

$$f_s(R_{stack}) = \frac{\int_0^{R_{stack}} \exp\left(-\frac{w(R)}{k_BT}\right) dR}{\int_0^{\infty} \exp\left(-\frac{w(R)}{k_BT}\right) dR}$$

where k_B is the Boltzmann constant and T the temperature. The selection of R_{stack} will be discussed below.

Energetic Calculations. A simple analysis of the energetics of the aqueous solvated ribodinucleoside monophosphates was performed for each MpN on a simulation of a stacked conformation, represented by the simulation using the restrained distance $R_{N_xN_y}$ = 4.5 Å, and on a simulation of an unstacked conformation, represented by the simulation using $R_{N_xN_y}$ = 9.0 Å. The energy difference between a stacked and an unstacked conformation was calculated from

$$\Delta E = \langle E \rangle_{4.5} - \langle E \rangle_{9.0}$$

where $\langle \cdots \rangle_{R_{N_xN_y}}$ is the average energy obtained in the simulations with $R_{N_xN_y}$ restrained to 4.5 and 9.0 Å as indicated. This difference was analyzed specifically for the base–base interaction energy (ΔE_{bb}), including all interaction terms between the atoms of one base with the atoms of the other base, and the internal backbone/backbone–bases interaction energy (ΔE_{bb-b}), which is the remaining internal energy in the solute.

Results and Discussion

For clarity the detailed analysis of conformational and energetic properties in the subsequent sections is restricted to the simulations performed with the restraining potential set up to keep the dimers close to $R_{N_xN_y}$ = 4.5 and 9.0 Å, representing stacked and unstacked conformations.

Free Energy Profiles. The free energy profiles (Figure 2a–f) of the ribodinucleoside monophosphates in aqueous solution are very sequence dependent, the most prominent characteristics being either a well-defined minimum around $R_{N_xN_y}$ = 4.5–5 Å or a rather flat profile with or without a small barrier. The deepest minima and the highest barriers and therefore the highest stacking abilities were obtained for the purine–purine dimers, especially for ApA (Figure 2a), which stacked substantially better than the pyrimidine–pyrimidine dimers (Figure 2b). This is in agreement with what has been observed experimentally.[1] We estimated the free energy of stabilizing the stacked state to be about 2–6 kcal/mol higher for purine–purine dimers than for pyrimidine–pyrimidine dimers as indicated by the plateau level in the PMF profiles (Figure 2a–f). The shallowest minimum for the purine–purine dimers was observed for the GpG dimer, and the highest stacking ability was observed for the adenine base containing purine–purine dimers.

In the PMF profiles of the pyrimidine–pyrimidine dimers no minimum was obtained indicating very low degree of stacking (Figure 2b). Most of the dimers containing combinations of purine and pyrimidine bases showed intermediate stacking propensities. Very good stacking ability was observed for the purine–pyrimidine dimers with a guanine base in the 5′ position, GpC and GpU (Figure 2d,f). An adenine base in the 5′ position was less proficient in promoting stacking (Figure 2c,e). Overall the purine–pyrimidine dimers showed higher stacking ability than the pyrimidine–purine dimers except for the A and C combinations, where CpA stacked better than ApC (Figure 2c). For the PMF profiles with a well-pronounced minimum, the PMF minimum was found close to the $R_{N_xN_y}$ distance in the starting structures of the ribodinucleoside monophosphates obtained from X-ray fiber diffraction data.[27] From this we could conclude that the reaction coordinate worked well and that the PMF minimum well represented the stable stacked and most favored structure. As we earlier have observed from unrestrained MD simulations of GpU, the stacked structure kept its conformation during the simulation.[5,25]

(33) Beveridge, D. L.; DiCapua, F. M. *Annu. Rev. Biophys. Biophys. Chem.* **1989**, *18*, 431–492.
(34) Kumar, S.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A.; Rosenberg, J. M. *J. Comput. Chem.* **1992**, *13*, 1011–1021.
(35) Boczko, E. M.; Brooks, C. L., III. *J. Phys. Chem.* **1993**, *97*, 4509–4513.
(36) Kraulis, P. J. *J. Appl. Crystallogr.* **1991**, *24*, 946–950.

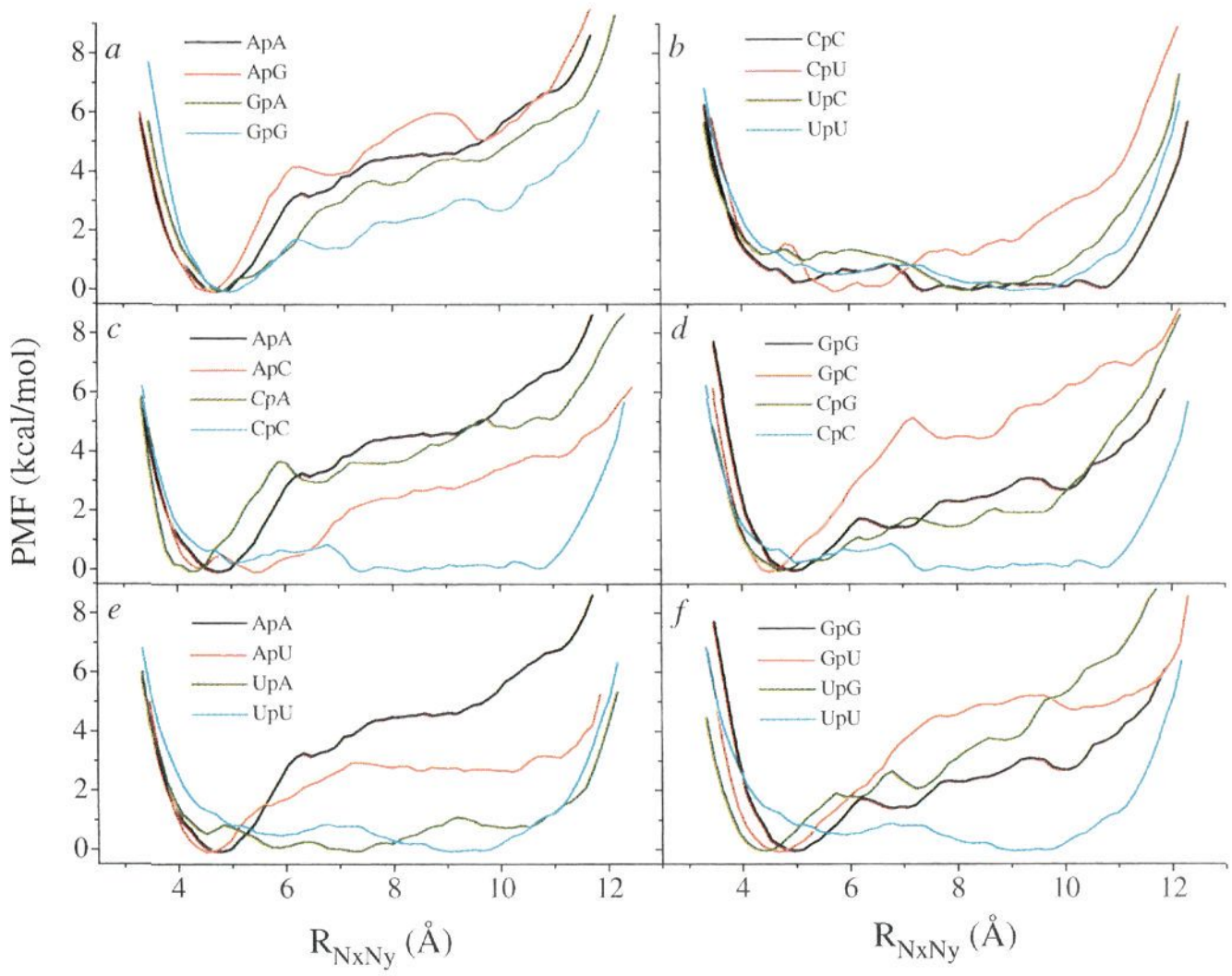

Figure 2. Free energy profiles of stacking for all the 16 combinations of bases in the ribodinucleoside monophosphates. PMF energy versus restrained distance $R_{N_xN_y}$ is shown for (a) the purine−purine combinations (A, G), (b) the pyrimidine−pyrimidine combinations (C, U), (c) the A, C combinations, (d) the G, C combinations, (e) the A, U combinations, and (f) the G, U combinations.

Stacking Equilibrium. From the PMF profiles we observed that many intermediate states were populated (Figure 2) and that the stacking/unstacking equilibrium cannot be well represented as a two-state system. It is also difficult to determine exactly which states along our reaction coordinate that are stacked or unstacked due to different distance, overlap, and parallelism between the bases. The equilibrium between stacked and unstacked states can be represented by an equilibrium constant $K_{eq} = f_s/f_u$, where f_s and f_u are the fraction stacked and unstacked, respectively. The particular value chosen for R_{stack} will obviously influence K_{eq} as can be seen in Figure 3, where we have plotted the fraction stacked f_s versus R_{stack}; it is possible to use different R_{stack} values for different sequences (see also the discussion of conformations below), but in order not to have this as a free parameter, we chose to use a sequence-independent R_{stack}. It should also be noted that even though the exact values of K_{eq} and f_s depend on R_{stack} the ranking of sequences according to K_{eq} or f_s is insensitive to R_{stack} since the curves in Figure 3 in general do not cross.

In the initial stacked structures of the ribodinucleoside monophosphates, the $R_{N_xN_y}$ distance was 4.736 Å and the criterion for stacking R_{stack} must be close to that. For the analysis R_{stack} distances of 4.5 and 5.0 Å were chosen. The fraction stacked at 4.5 Å, $f_s(4.5)$, was 0.4−2.1% for the pyrimidine−pyrimidine dimers and slightly higher for the UpA, GpG, and GpA dimers (Table 1). For most of the dimers $f_s(4.5)$ was between 7.0% and 15.6%, except for the UpG and CpA dimers (31.8% and 52.9%, respectively). At 5.0 Å the fraction stacked $f_s(5.0)$ was still low for the pyrimidine−pyrimidine dimers (1.9−6.0%), but an obvious increase was observed for all the other dimers, except UpA (6.1%). For instance the CpA dimer showed very high values of fraction stacked.

We calculated the equilibrium constants for the R_{stack} distances 4.5 Å, $K_{eq}(4.5)$, and 5.0 Å, $K_{eq}(5.0)$ (Table 1). Very small values of $K_{eq}(4.5)$ were observed for the pyrimidine−pyrimidine dimers, 0−0.02, but $K_{eq}(5.0)$ showed slightly higher values, 0.02−0.06. Highest values of $K_{eq}(4.5)$ and $K_{eq}(5.0)$ were obtained for the UpG and CpA dimers. The majority of the the dimers showed $K_{eq}(4.5)$ between 0.02 and 0.18 and $K_{eq}(5.0)$ from 0.29 to 2.21. Purine−purine dimers tended to be parallel at much longer $R_{N_xN_y}$ distances than pyrimidine−pyrimidine dimers (see Conformational Analysis below), indicating that the same stacking criterion should perhaps not be used for all ribodinucleoside monophosphates. The fraction unstacked could be divided into different parts, for instance it is possible to calculate the fraction of intermediate states and the fraction of unstacked states to have three separated states including the stacked state.

We can divide the 16 dimers in three groups with low ($f_s < 0.1$), medium ($0.1 < f_s < 0.5$), or high ($f_s > 0.5$) fraction stacked at room temperature. From the $f_s(5.0)$ in Table 1 we then get CpC, CpU, UpA, UpC, and UpU in the low group; ApA, ApC, CpG, GpA, GpG, and GpU in the medium group; and ApG, ApU, CpA, GpC, and UpG in the high group. A similar grouping based on experimental f_s estimates (see ref 1 and references therein) comes out with some differences from our analysis, which also depend on the experimental technique. The main differences, where our results differ from both UV and NMR data, are that we find that the UpG dimer should stack very well, whereas both UV and NMR in this case have medium-low stacking estimates, and we find the CpC dimer not to stack at all, when UV and NMR find medium stacking estimates for CpC. The conformational analysis for the UpG dimer below shows that it has an unusual conformation of the backbone α torsion in the extended state, which could contribute to a PMF favoring the stacked state. The CpC dimer on the other hand is very flexible, as are all the pyrimidine−pyrimidine dimers in our simulations, even at close distances.

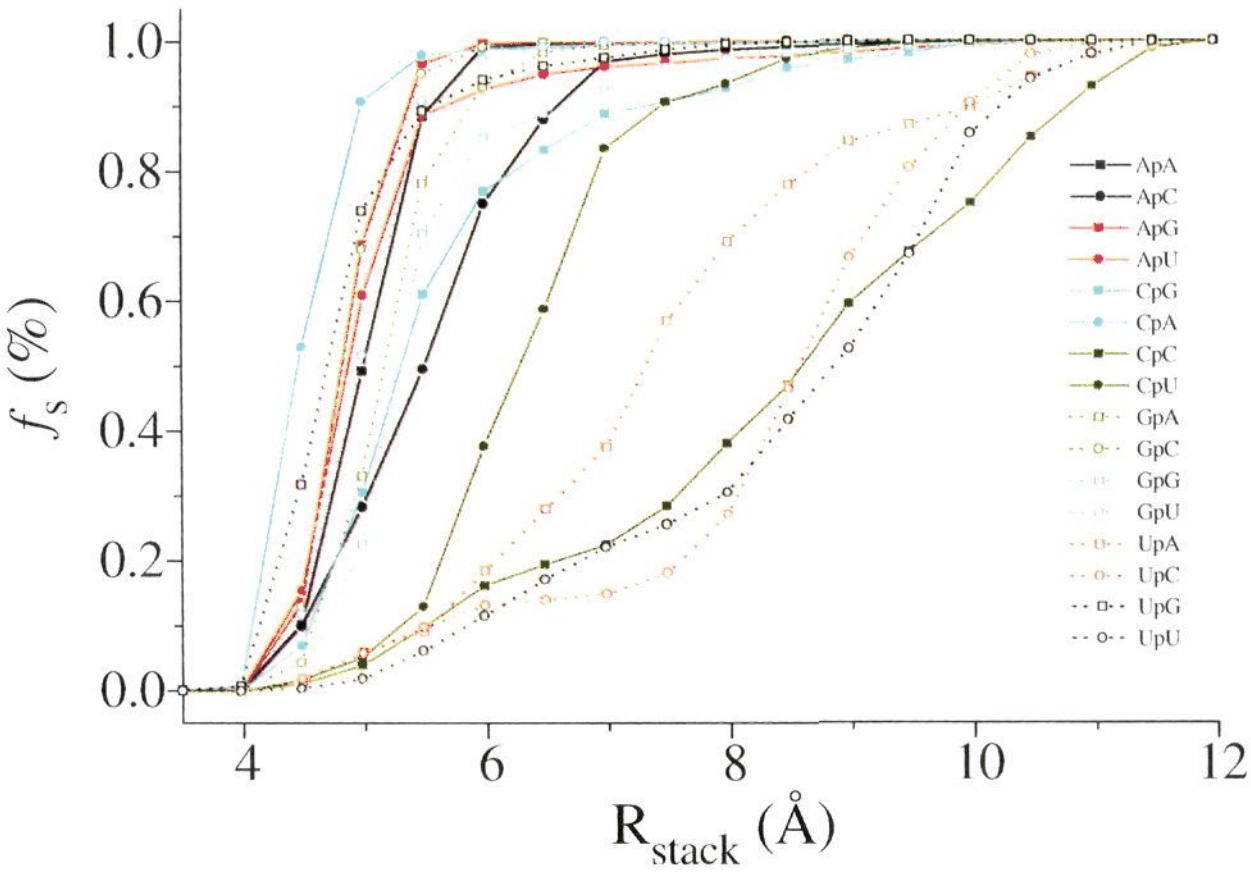

Figure 3. Fraction stacked, f_s, as a function of stacking distance, R_{stack}, which is a criterion to determine which states are stacked, for all 16 ribodinucleoside monophosphates.

Table 1. Equilibrium Constants, K_{eq}, and the Fraction Stacked, f_s, Calculated at the Stacking Distances R_{stack} = 4.5 and 5.0 Å

dimer	K_{eq}(4.5)	f_s(4.5) (%)	K_{eq}(5.0)	f_s(5.0 (%)
ApA	0.11	10.2	0.97	49.2
ApC	0.11	10.0	0.39	28.3
ApG	0.16	13.8	2.21	68.8
ApU	0.18	15.6	1.56	61.0
CpA	1.12	52.9	9.64	90.6
CpC	0.01	1.1	0.04	4.0
CpG	0.08	7.0	0.44	30.6
CpU	0.02	1.7	0.05	5.2
GpA	0.05	4.4	0.49	33.0
GpC	0.15	12.9	2.12	67.9
GpG	0.02	2.1	0.29	22.6
GpU	0.14	11.9	1.07	51.7
UpA	0.02	1.7	0.06	6.1
UpC	0.02	2.1	0.06	6.0
UpG	0.47	31.8	2.82	73.8
UpU	0	0.4	0.02	1.9

The stacking ability is not simply a function of base size as is evident from, e.g. the GpG dimer having a fairly small f_s, or from comparing how f_s depends on the base order in the six couples of heterodimers. Here we find that ApU stacks better than UpA, and that CpG stacks less well than GpC, both of which are in agreement with the ordering obtained from NMR data.[1] For ApC the ordering we observe is however the reverse of that seen in NMR.

One reason for discrepancies between the detailed sequence dependence in our PMF profiles and the sequence dependencies seen experimentally is that stacking as measured by our reaction coordinate $R_{N_xN_y}$ is not exactly the same as stacking measured by experimental spectroscopic techniques, which depend on electronic interactions. We can thus have situations where the bases are very parallel and overlap but with a distance between the bases of 6−7 Å, or bases that are close but not very parallel, or with little overlap, and which may be interpreted differently.

Conformational Analysis. Many different conformations, both stacked and unstacked, were obtained as can be seen for the ApC dimer in Figure 4. For $R_{N_xN_y}$ up to about 6 Å the bases of purine−purine dimers are parallel, but most of the pyrimidine-containing dimers leave the stacked state at shorter distances.

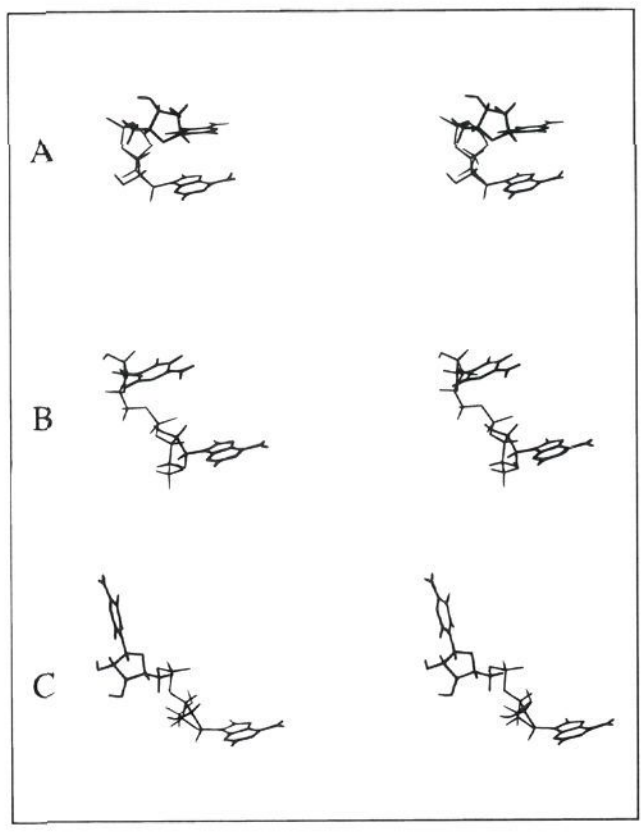

Figure 4. Stereoviews of stacked and unstacked conformations of the ribodinucleoside monophosphate ApC. The snapshots from the trajectories are taken from the molecular dynamics simulations using different restrained distances $R_{N_xN_y}$ (A) 4.0, (B) 7.5, and (C) 11.0 Å. The pictures in the figure were generated with the program Molscript.[36]

The pyrimidine−pyrimidine dimers were observed to be very flexible at short $R_{N_xN_y}$ distances and therefore showed very little stacking.

The dynamical behavior of the backbone and the glycosidic torsion angles and the pseudorotation phase angles was analyzed using the Dials and Windows programs.[37] Only for the CpC and UpU dimers we observed changes of the δ angle for the restrained distance $R_{N_xN_y}$ = 4.5 Å. The ApG and GpA dimers showed no transitions of the δ angle at $R_{N_xN_y}$ = 9.0 Å (Figure 5a), and the UpA and UpG dimers showed only changes of the δ angle in the 5′-ribose (Figure 5b). For a purine base in the 5′

(37) Ravishanker, G.; Swaminathan, S.; Beveridge, D. L.; Lavery, R.; Sklenar, H. *J. Biomol. Struct. Dyn.* **1989**, *6*, 669−699.

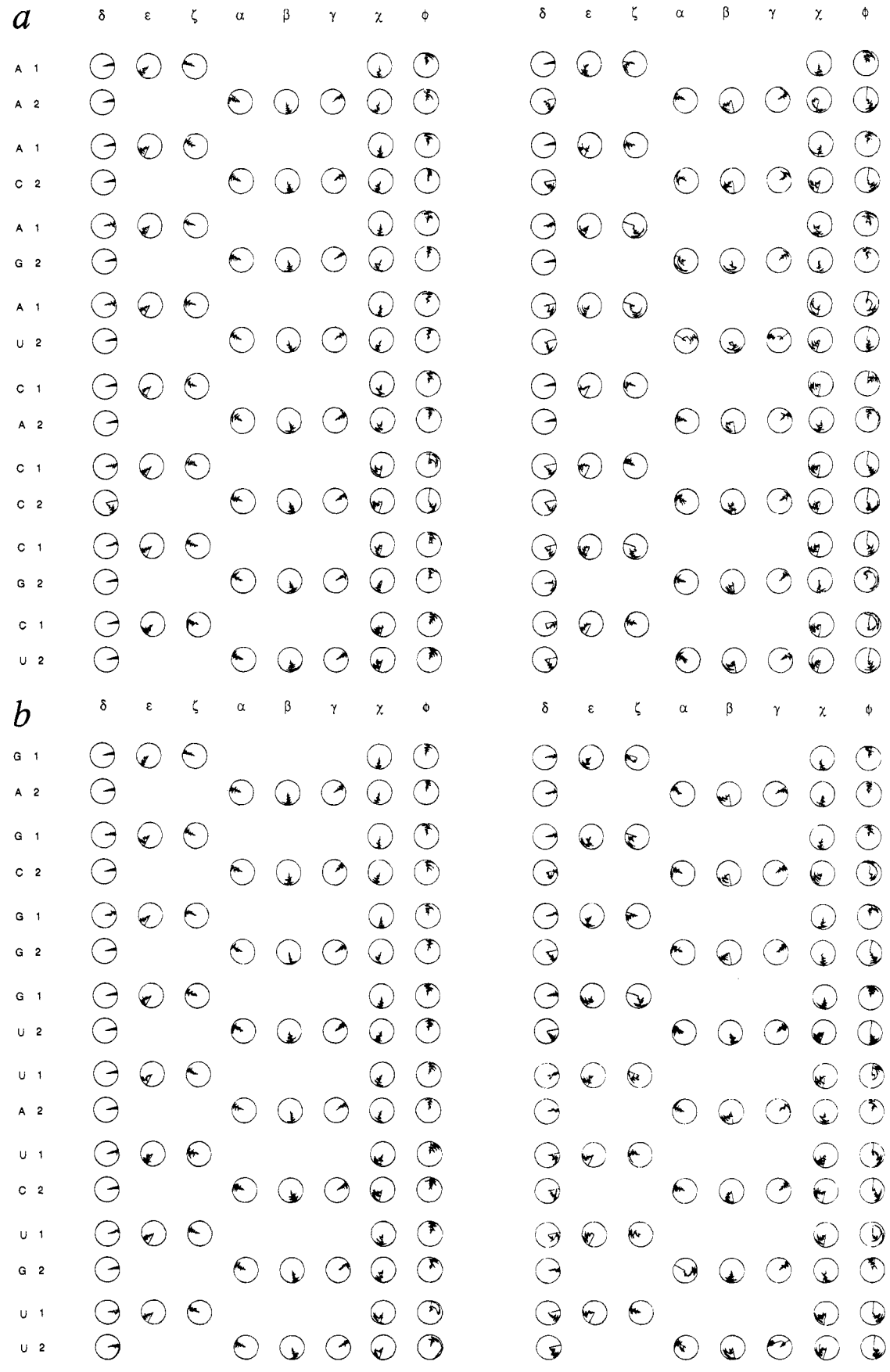

Figure 5. Time evolution of the backbone torsion angles (δ, ϵ, ζ, α, β, γ), the glycosidic torsion angles (χ), and the sugar pucker pseudorotation phase angles (ϕ) of the (a) ApN and CpN and (b) GpN and UpN (N = A, C, G, and U) ribodinucleoside monophosphates from the MD simulations using the restrained distances $R_{N_xN_y} = 4.5$ Å (stacked conformation, to the left) and 9.0 Å (unstacked conformation, to the right). The 5′-base is indicated by 1 and the 3′-base by 2. The angles are defined as in Figure 1.

position, no changes of the δ angle were observed, except for ApU, but for a pyrimidine base in the 5′ position, the δ angle changed for all ribodinucleoside monophosphates except CpA at $R_{N_xN_y} = 9.0$ Å. The δ angle changed for all dimers with a

Table 2. Angle θ between the Normal Vector of the 5′-Base and the 3′-Base for the MD Simulations using the Restrained Distances $R_{N_xN_y}$ = 4.5 and 9.0 Å[a]

dimer	θ(4.5) (deg)	θ(9.0) (deg)
ApA	11.0 (6.4)	52.7 (33.7)
ApC	15.2 (7.2)	53.2 (27.0)
ApG	12.8 (6.8)	33.2 (16.6)
ApU	15.2 (7.1)	43.5 (20.8)
CpA	24.1 (10.1)	57.4 (22.2)
CpC	46.8 (20.0)	64.4 (17.4)
CpG	33.8 (14.6)	85.8 (32.6)
CpU	29.1 (13.3)	62.8 (21.3)
GpA	9.9 (5.4)	87.3 (18.8)
GpC	13.6 (6.8)	37.0 (16.1)
GpG	10.8 (5.5)	73.3 (27.0)
GpU	14.1 (7.2)	73.0 (20.7)
UpA	28.7 (13.6)	80.1 (27.7)
UpC	26.3 (10.4)	65.4 (22.7)
UpG	25.8 (10.6)	71.4 (21.6)
UpU	36.3 (18.6)	76.5 (20.5)

[a] In the parentheses are the standard deviations.

pyrimidine base in the 3′ position. Most of the dimers showed a change against a less negative value of the ϵ angle for both $R_{N_xN_y}$ = 4.5 and 9.0 Å. No changes of the ζ angle were observed at $R_{N_xN_y}$ = 4.5 Å, but at $R_{N_xN_y}$ = 9.0 Å, and ζ angle changed for all the dimers containing a purine base and the largest changes were obtained for the ApG, ApU, CpG, and GpU dimers. The ApU and UpG dimers showed major changes of the α angle at $R_{N_xN_y}$ = 9.0 Å, and the β angle was increased for a majority of the dimers. For $R_{N_xN_y}$ = 4.5 Å, no transitions were observed for the γ angle, but at $R_{N_xN_y}$ = 9.0 Å, the γ angle of the 3′-base changed dramatically for the ApU and UpU dimers. For the χ angle we observed slightly less negative values for the pyrimidine bases in both the 5′ and 3′ positions.

Two ranges of the pseudorotation phase angle ϕ are preferred for nucleotides: C3′-*endo* (^{3}E) at 0° ≤ ϕ ≤ 36° (in the north cycle, N) and C2′-*endo* (^{2}E) at 144° ≤ ϕ ≤ 190° (in the south cycle, S).[26] All the initial structures have the ϕ angle in the ^{3}E mode. For the CpC and UpU dimers we observed changes to the ^{2}E mode of the ϕ angle of the ribose group connected to the 3′ base for the MD simulation using the restrained distance $R_{N_xN_y}$ = 4.5 Å (Figure 5a,b). The ϕ angle of the 5′-ribose changed toward S and then back to ^{3}E for CpC and UpU, but all other dimers showed the ϕ angle in the ^{3}E mode for $R_{N_xN_y}$ = 4.5 Å. For the unstacked conformations obtained from the MD simulation using the restrained distance $R_{N_xN_y}$ = 9.0 Å, we observed no changes of the ϕ angle from the ^{3}E mode for the ApG and GpA dimers and for the ϕ angle close to the pyrimidine base of the GpC and UpA dimers showed changes to the ^{2}E mode and back to the ^{3}E mode. The ϕ angle showed a preference for the ^{2}E mode for both the sugar moieties of the ApU, CpC, CpG, UpC, and UpU dimers at $R_{N_xN_y}$ = 9.0 Å. The ϕ angle was observed in the ^{2}E mode for the sugar moiety of the pyrimidine base of the UpG dimer and for the sugar moiety of the 3′-base of the ApA, ApC, CpA, CpU, GpG, and GpU dimers at $R_{N_xN_y}$ = 9.0 Å.

To examine the flexibility of the bases relative each other and to understand the vibrational motions of the bases we calculated the angle θ between the normal vectors of the bases for the restrained distances $R_{N_xN_y}$ = 4.5 Å, θ(4.5), and 9.0 Å, θ(9.0). For the purine−purine dimers, we observed an angle θ(4.5) between 9.9° and 12.8° (Table 2), indicating the purine base planes to be strongly attracted to each other. For the purine and pyrimidine containing dimers with a purine base in the 5′ position, we observed an angle θ(4.5) of 13.6°−15.2°, but for a purine base in the 3′ position, and angle θ(4.5) was larger, 24.1°−33.8°. The pyrimidine−pyrimidine dimers were most

Figure 6. Base−base interaction energy difference between stacked ($R_{N_xN_y}$ = 4.5 Å) and unstacked ($R_{N_xN_y}$ = 9.0 Å) conformations. The error bars displayed are the largest standard deviation of the interaction energy for the stacked or unstacked form.

flexible at $R_{N_xN_y}$ = 4.5 Å, with θ(4.5) = 26.3°−46.8°. The purine−purine dimers showed the smallest root mean square fluctuations, 5.4°−6.8°, of θ(4.5). θ(9.0) was in the range 33.2°−87.3° for all 16 sequence combinations, which means that θ(9.0) > θ(4.5) for all the dimers. The root mean square fluctuations of the angle θ were larger for θ(9.0) than for θ(4.5) for all the dimers, except for the CpC dimer.

Energetic Analysis. Solvent interactions are known to influence stacking propensities, but the exact nature of this influence and its relation to the direct base−base interactions remains unclear.[24] The ΔE_{bb} energies from the direct base−base interactions (Figure 6) were in accord with other theoretical calculations,[2] which showed the highest stacking propensity for the GC pair, where we found the lowest base−base interaction energy for the GpC dimer. The lowest base−base interaction energies were overall obtained for dimers containing one guanine base, with the GpG dimer being the least favorable, as shown in other studies.[2] Energy differences of the internal backbone and backbone−bases interaction energy, ΔE_{bb-b}, for the dimers were mainly due to the backbone−base interaction energy. The ΔE_{bb-b} energy varied between −14 and 7 kcal/mol with standard deviations of about 7−10 kcal/mol for all the dimers. The lowest values of ΔE_{bb-b} were obtained for the ApU, UpA, CpG, and GpC dimers. Although the ΔE_{bb-b} values are less certain than those for ΔE_{bb}, it is clear that they are of the same magnitude and that their sequence dependencies are not the same.

Solvent Interaction Sites. The average number of hydrogen bonds were calculated for each hydrogen bond donor or acceptor atom in the solute. A hydrogen bond was assumed to exist if H···X ≤ 2.4 Å. The strongest interactions with water were formed by the phosphate oxygens, which had on average 2−3 hydrogen bonds to water molecules (Figure 7a−d, supporting information). As expected, due to that the phosphate oxygens are sticking into the water, these interactions are quite similar for the ribodinucleoside monophosphates, stacked or unstacked. The most exposed atom of the adenine bases in both the 5′ and 3′ positions was the N1 atom, which showed 1.1−1.4 hydrogen bonds to water and slightly more exposed for $R_{N_xN_y}$ = 9.0 Å than for $R_{N_xN_y}$ = 4.5 Å. The N7 atom of all the adenine bases showed 0.5−1.0 hydrogen bonds and was more exposed at $R_{N_xN_y}$ = 9.0 Å, but the N7 atom of the guanine bases formed more hydrogen bonds, on average 1.5−2.0. For the O6 atoms of the guanine bases we observed 1.1−1.7 hydrogen bonds to water. The O2 atom of the cytosine bases showed about 1.5−1.8

Table 3. Interaction Energies between the 2'OH Group of the 5'-Ribose and the 3'-Base Including the Sugar Moiety for the MD Simulations using the Restrained Distances $R_{N_xN_y}$ = 4.5 Å, $E_{2'OH}(4.5)$, and 9.0 Å, $E_{2'OH}(9.0)$[a]

dimer	$E_{2'OH}(4.5)$ (kcal/mol)	$E_{2'OH}(9.0)$ (kcal/mol)
ApA	3.81 (1.58)	2.69 (1.87)
ApC	3.77 (1.80)	2.93 (1.48)
ApG	4.29 (1.27)	3.45 (1.38)
ApU	3.18 (1.14)	1.76 (2.80)
CpA	3.91 (1.15)	2.11 (1.09)
CpC	4.12 (2.21)	2.27 (1.32)
CpG	4.45 (1.90)	1.78 (1.32)
CpU	3.90 (2.04)	1.89 (1.61)
GpA	4.05 (1.39)	1.88 (1.33)
GpC	4.07 (1.54)	2.24 (1.90)
GpG	4.15 (1.17)	2.94 (2.16)
GpU	3.14 (1.20)	0.15 (0.73)
UpA	3.96 (1.37)	1.65 (1.90)
UpC	4.34 (2.18)	3.13 (2.10)
UpG	4.49 (1.04)	1.90 (3.76)
UpU	3.15 (1.94)	1.15 (1.83)

[a] In the parentheses are the standard deviations.

hydrogen bonds to water for most of the dimers containing a cytosine base and for the uracil bases were the O2 and O4 atoms most exposed. Very few hydrogen bonds to water were obtained for the O3' and O5' atoms bonded to the phosphate atom. A majority of the O2' atoms in the sugar moiety of the dimers formed around 1.0 hydrogen bonds to water. The O4' atoms of the dimers were more exposed at $R_{N_xN_y}$ = 9.0 Å, especially for the O4' atoms of the 3'-ribose. The 5' and 3' terminals were observed to be very flexible, and therefore, the average number of hydrogen bonds to water varied.

Most of the observed differences for the average number of hydrogen bonds formed to water between the stacked states and the unstacked were about 0.5 hydrogen bonds, but some atoms could clearly be seen to be more exposed in the unstacked conformation, e.g. in ApA, GpG, UpA, GpC, UpG, and CpC. In a few cases, the opposite situation, with atoms more exposed in the stacked conformation, was found, most notably for the G in CpG and for the U in UpC.

We also calculated the interaction energy between the sodium counterion and the solute, including the bases and the backbone, for the MD simulations using the restrained distances $R_{N_xN_y}$ = 4.5 and 9.0 Å. The counterion was observed to be very mobile, and this particular interaction energy component varied throughout the simulations between 0 and −60 kcal/mol in a random fashion. We would thus conclude that in these simulations the ion in itself had no direct effect on the solute.

Differences in conformational stability and physical properties between DNA and RNA depend on the ribose 2'OH group and on the C5-methyl in the DNA base thymine.[26,38] In a MD simulation study[5] of a stacked conformation of GpU, the hydrogen bond between the 2'OH group of the 5'-ribose and the O4' atom of the 3'-ribose was present 83% of the simulation time. To investigate the influence of the 2'OH group on stabilizing stacking, we calculated the interaction energy between the 2'OH group and the adjacent base and sugar moiety (the phosphorus atom and the phosphate oxygens were not included) for the MD simulations using the restrained distances $R_{N_xN_y}$ = 4.5 Å, $E_{2'OH}(4.5)$, and 9.0 Å, $E_{2'OH}(9.0)$ (Table 3).

The interaction energy $E_{2'OH}(4.5)$ was observed to be 1−3 kcal/mol higher than $E_{2'OH}(9.0)$, indicating that the direct interactions of the 2'OH group of the 5'-ribose with the 3'-base and ribose do not stabilize stacking; a recent study[38] of 2'OH effects on the stability of RNA and DNA double and triple helices showed that the 2'OH group could have both stabilizing and destabilizing effects. The ApU, GpU, and UpU dimers showed the smallest $E_{2'OH}(4.5)$, 3.14−3.18 kcal/mol, but for the other dimers, $E_{2'OH}(4.5)$ varied between 3.77 and 4.49 kcal/mol. For $E_{2'OH}(9.0)$ we obtained larger variations depending on the populated conformations and the flexibility of the 2'OH group. Aside from these direct interactions the 2'OH group could also have an effect on stacking stability through its influence on the ribose conformation.

Conclusions

Potential of mean force calculations of the free energy of ribodinucleoside monophosphates have given a detailed picture of the stacking/unstacking equilibrium. The free energy of the stacking/unstacking process for all 16 combinations of dimers are presented (Figure 2). The PMF minimum represents well the most favored stacked structure in aqueous solution. Free energy differences of 2−6 kcal/mol were observed between the stacked state and the unstacked states for the purine−purine dimers. The general sequence dependence of stacking abilities, purine−purine > purine−pyrimidine ≥ pyrimidine−purine > pyrimidine−pyrimidine, was found to be in agreement with experimental data. We have in a previous PMF study[39] of ApA at different temperatures calculated that in this case stacking is enthalpy driven, with ΔH = −6 kcal/mol and ΔS = −13 cal/mol/K for the unstacked to stacked transition which is within the range of experimental values for ApA. In the present simulations, many different conformations, with different degrees of stacking, were observed, and together with the PMF profiles, this clearly shows the gradual nature of the stacking phenomenon. The variation in equilibrium constants and fraction stacked of ribodinucleoside monophosphates reported in the literature[1,15-21] is also likely to be a reflection of this.

A number of major transitions of the backbone torsion angles were observed when going from a stacked state to an unstacked. The glycosidic torsion angles showed a preference for less negative values, and for the stacked conformations, the pseudorotation phase angle showed small changes from the ^{3}E mode commonly found in riboses but had a preference for the ^{2}E mode in the unstacked states for the ribose groups connected to pyrimidine bases and for some ribose groups connected to purine bases.

The atoms making the strongest interactions with solvent are the phosphate oxygens, which form on average 2−3 hydrogen bonds. Only smaller differences of the number of hydrogen bonds to water were observed between the stacked and the unstacked states. The direct interactions of the 5'-ribose 2'OH group with the 3'-base and ribose of the molecule destabilize the stacked states by 1−3 kcal/mol.

These results have shown that it is possible to obtain quantitative information about the thermodynamics of stacking in small solvated model systems from potential of mean force calculations using a standard empirical energy function, with no specific stacking energy term. Stacking preferences are the result of a rather subtle balance between a number of interactions, and base size is by no means the only important factor. The sequence dependence of the free energies stabilizing stacking is not the same as for the base−base interaction energies, showing that both the solvent and the backbone influence the stacking energetics and thermodynamics in these systems and that their roles also have to be considered for a full understanding of stacking and its implications for the structure and stability of larger DNA and RNA molecules.

(38) Wang, S.; Kool, E. T. *Biochemistry* **1995**, *34*, 4125−4132.

(39) Norberg, J.; Nilsson, L. *J. Phys. Chem.* **1995**, *99*, 13056−13058.

Acknowledgment. This work was supported by the Swedish Natural Science Research Council and by the Magnus Bergvall Foundation.

Supporting Information Available: Figure 7a–d, illustrating the average number of hydrogen bonds formed between water and the ribodinucleoside monophosphates containing (a) purine–purine combinations, (b) pyrimidine–pyrimidine combinations, (c) adenine–pyrimdine and pyrimidine–adenine combinations, and (d) guanine–pyrimidine and pyrimidine–guanine combinations for the MD simulations using the restrained distances $R_{N_xN_y}$ = 4.5 and 9.0 Å (atoms of the 5′-base to the left and of the 3′-base to the right) (4 pages). This material is contained in many libraries on microfiche, immediately follows this article in the microfilm version of the journal, can be ordered from the ACS, and can be downloaded from the Internet; see any current masthead page for ordering information and Internet access instructions.

JA952156I