Empirical Energy Functions for Energy Minimization and Dynamics of Nucleic Acids*

Lennart Nilsson[†] **and Martin Karplus**

Department of Chemistry, Harvard University, 12 Oxford Street, Cambridge, Massachusetts 02138

Received 16 April 1985; accepted 14 January 1986

An improved empirical energy function for energy minimization and dynamics calculations of nucleic acids is developed and evaluated by an examination of its representation of both static and dynamic properties of model systems. Among the properties studied and used for parameter optimization are base pairing interactions, sugar and phosphate energy surfaces, small crystal heats of sublimation, base, phosphate and sugar analogue vibration spectra, and the overall behavior of a DNA hexamer duplex in vacuum molecular dynamics simulations. The results obtained are compared with those from two other energy functions that have been used recently for nucleic acids. Parameters for two energy functions are given; one includes heavy atoms and only polar hydrogens and the other includes all atoms.

I. INTRODUCTION

One of the central problems in a molecular mechanics study, be it an energy optimization or a molecular dynamics simulation, is the calculation of the energy of the system as a function of the coordinates of all its constituents. During the last 15 years, much work has been done to obtain analytical forms for empirical energy functions[1–5] that are easy to use, even in a complex macromolecular system and at the same time are sufficiently good approximations to the true potential that the results are meaningful. Most of the effort has been applied to proteins, particularly as far as the use of energy functions for molecular dynamics is concerned;[6] it is more recently that the potential functions developed for nucleic acids[7–9] have been employed in dynamical simulations.[10–12]

Parameter sets now in use for nucleic acid dynamics have evolved from protein parameters. We have developed an improved empirical energy function (EF2) for nucleic acids in the absence of explicit solvent, i.e., shielding of charged groups that approximates the effect of solvent is included. We began with an earlier set (EF1) that has been employed for dynamical studies of DNA[11]; it was based on a parameterization by Kollman, Weiner, and Dearing,[8] which in turn utilized the protein results of Gelin and Karplus.[4] In addition, we present some comparisons with a newer energy function (EF3) by Kollman and coworkers.[5] All three energy functions include polar hydrogens explicitly, but use an extended atom representation for CH, CH_2, and CH_3 groups. In the present study both static and dynamic aspects of the energy functions were tested. Comparisons were made with known structural, thermodynamic, and spectroscopic properties of model compounds representative of the elementary building blocks of DNA and RNA. The five common bases (A, C, G, T, and U), ribose, deoxyribose, tetrahydrofuran, diethyl- and dimethylphosphate, and two small molecule crystals (*n*-octane and diketopiperazine) were used in the analysis. A method to use vibrational spectra to fit parameters, based on the approach of Warshel and Lifson,[13,14] is presented and applied to diethylphosphate, uracil, and tetrahydrofuran. Finally, to evaluate the overall efficacy of the energy functions, they were employed in molecular dynamics simulations of the double stranded DNA hexamer $d(CGTACG)_2$ with different electrostatic models that account for some of the effects of

*Supported in part by a grant from the National Institutes of Health.

[†]Present address: Department of Medical Biophysics, Karolinska Institute, S-104 01 Stockholm, Sweden

CCC 0192-8651/86/050591-26$04.00

solvent shielding. Most macromolecular molecular dynamics simulations to date have been made in vacuum, and the effects of surrounding solvent have largely been neglected, or only partially taken into account, e.g., through the use of reduced charges on phosphate groups and/or a distance dependent dielectric parameter[11,15,16]; both of these decrease long-range electrostatic interactions.

Section II describes the methods. In particular, it outlines the procedure used for refinement of parameters with vibrational data.

Section III presents the results for the nucleic acid components used for the duplex d(CGTACG)$_2$ oligomer. A concluding discussion is given in Section IV. Appendix 1 lists the parameter set based on heavy atoms and polar hydrogens; Appendix 2 lists an all-atom parameter set, including polar and aliphatic hydrogens.

II. METHODS

This section briefly describes the empirical energy functions and the approaches used to fit its parameters to the available data.

Empirical Energy Functions

The empirical energy function used to represent the systems studied has the general form

$$
\begin{aligned}
E(R;x) = & \sum_{\text{bonds}} K_b(b - b_0)^2 \\
& + \sum_{\substack{\text{bond}\\ \text{angles}}} K_\theta(\theta - \theta_0)^2 \\
& + \sum_{\substack{\text{dihedral}\\ \text{angles}}} K_\phi[1 + \cos(n\phi - \delta)] \\
& + \sum_{\substack{\text{improper}\\ \text{torsions}}} K_\omega(\omega - \omega_0)^2 \\
& + \sum_{\substack{\text{H-bond}\\ k,l\,\text{pairs}}} \left(\frac{A_{kl}}{s_{kl}^p} - \frac{B_{kl}}{s_{kl}^m}\right) \cos^\alpha(\phi_{D\text{-H-}A}) \cos^\beta(\phi_{AA\text{-}A\text{-H}}) \\
& + \sum_{\substack{\text{nonbond}\\ i,j\,\text{pairs}}} \left[4\epsilon_{ij}\left\{\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right\} + \frac{332 q_i q_j}{D r_{ij}}\right] \qquad (1)
\end{aligned}
$$

where b, K_b, and b_0 are the bond length, bond stretch force constant, and equilibrium distance, respectively; θ, K_θ, and θ_0 are a bond angle, its force constant, and equilibrium angle, respectively; ϕ, K_ϕ, n, and δ are a dihedral angle, its force constant, multiplicity (the sum may include terms with different multiplicities for any given dihedral angle), and phase, respectively; ω, K_ω, and ω_0 are an improper torsion (i.e., a dihedral angle between four nonconnected atoms, as used in some special cases[16]) its force constant, and equilibrium value, respectively; s_{kl} is the hydrogen bond length, A_{kl}, B_{kl}, and p, m define the shape and depth of the hydrogen bond well; $\phi_{D\text{-H-}A}$ and $\phi_{AA\text{-}A\text{-H}}$ are donor-hydrogen-acceptor and acceptor-antecedent-acceptor-hydrogen angles, with exponents $p = 12$, $m = 10$, and $\alpha = \beta = 0$ in EF1 and EF3, and $p = 6$, $m = 4$, $\alpha = 4$, and $\beta = 2$ in EF2. The quantities r_{ij}, ϵ_{ij}, σ_{ij}, and q_i, q_j are the nonbonded distance, the Lennard-Jones well depth, the Lennard-Jones diameter and the charges, respectively; for atom pairs i, j, respectively with D, an effective dielectric parameter.[11,15,16]

R is the Cartesian coordinate vector describing the geometry of the system, and x is the collection of parameters (force constants, equilibrium values, etc.) used in a given energy function.

Equilibrium values for bond lengths and bond angles are the same in EF1, EF2, and EF3; they all come from standard reference structures. In EF1 and EF2 only one dihedral term is assigned to a given bond, with the exception of the sugar-phosphate backbone where both two-fold and three-fold terms are used; whereas, in EF3 all possible dihedrals through a bond have energy terms associated with them. The choice of a harmonic, rather than cosine, improper dihedral term in EF1 and EF2 was made because the harmonic form prevents chiral centers from inverting; EF3 uses a cosine term. The same nonbonded parameters are employed in EF1 and EF2, and the hydrogen bond parameters correspond to the protein parameters in the program CHARMM (version 17).[16,18]

An r-dependent dielectric (i.e., $D = r$) with a nonbond cutoff switching function[16] applied between 9.5 and 10.5 Å, and a cutoff for hydrogen bonds with a switch between 3.5 and 4.0 Å for EF1 and EF3, and 4.0 and 7.0 Å for EF2 were used in all calculations, except

when otherwise noted in the text. EF3 introduces a scaling of 1–4 nonbonded interactions by 0.5. All calculations were done with the CHARMM program[16]; that EF3 calculations correspond to those of Kollman and coworkers[5] was verified by energy evaluations for identical systems with CHARMM[16] and AMBER[5] (results not shown).

Energy Minimizations and Dynamics

The standard CHARMM minimizers ABNR and NRAPH,[16] both of the Newton-Raphson type, were used in the minimizations, which were carried through until the rms average force per degree of freedom in the system was less than a specified value; the values chosen for different model systems are given in the text. The Verlet algorithm with a timestep of 0.001 ps and with covalent bonds to hydrogens fixed by means of the SHAKE[19] algorithm was used for the dynamics.

Parameter Fitting to Normal Mode Spectra

An important element in obtaining EF2 from EF1 involved improved fitting to the vibrational spectra of model compounds. Normal mode spectra are calculated assuming that the potential energy of the molecule is harmonic, in the neighborhood of an energy minimum. The system is then fully characterized by the mass weighted second derivative matrix,

$$F_{ij}(R;x) = \mathbf{M}^{-1/2}\frac{\partial^2 E(R;x)}{\partial R_i \partial R_j}\mathbf{M}^{-1/2} \qquad (2)$$

where $\mathbf{M}$ is a diagonal matrix with the atom masses as its elements. The normal modes of vibration are obtained by solving

$$F|\lambda_i\rangle = \lambda_i|\lambda_i\rangle \qquad (3)$$

where the eigenvectors $|\lambda_i\rangle$ are the normal modes, and the eigenvalues λ_i are related to the vibration frequencies ω_i (in cm^{-1}) by

$$\lambda_i = (2\pi c\omega_i)^2 \qquad (4)$$

where c is the speed of light.

The normal mode facility in the program CHARMM was augmented with a parameter fitting subprogram in which any set of parameters can be varied to optimize the fit between the calculated and an observed spectrum. This was used to improve the values of some of the force constants, though a corresponding approach could be used to optimize any parameters. Since the optimization is highly nonlinear an efficient iterative method has to be used, and the Marquardt algorithm[18] was chosen because it combines the large convergence radius of steepest descent with the rapid convergence of Newton-Raphson methods without requiring the computation of second derivatives.

The block diagram (Fig. 1) outlines the overall scheme used in the fitting process. There are a number of options that may be used; the most important of these is the ability to keep track of the modes during the iterations. If a procedure of this kind is to be successful it is necessary to have a well-assigned experimental spectrum. The initially computed spectrum will, in general, be different from the observed, not only in the peak positions, but also in the order and character of the modes. It is, therefore, important that there is a mechanism by which the correspondence between calculated and observed modes can be introduced and maintained. We have chosen the following simple approach. The quantity that the

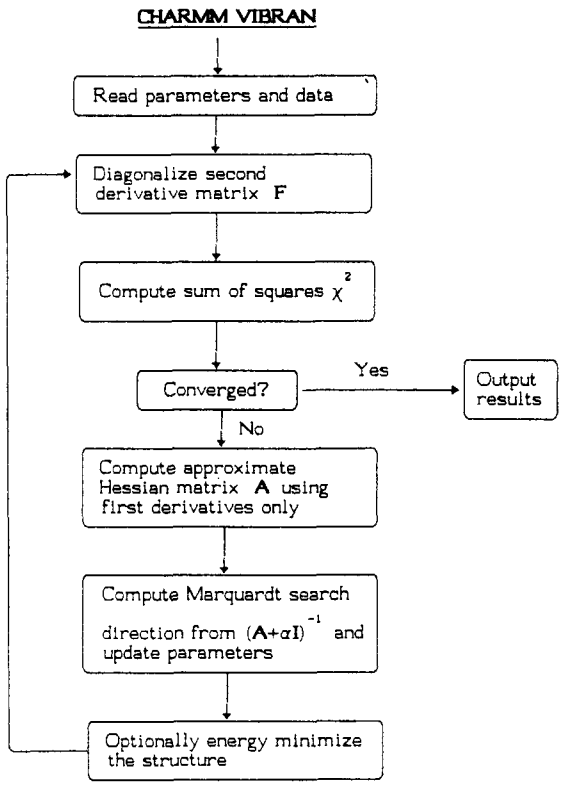


Figure 1. Block diagram of the parameter fitting module in CHARMM.

Marquardt algorithm is required to minimize is taken to be

$$\chi^2 = \sum_{i,j} w_{ij}(\lambda_i - \mu_j)^2 \tag{5}$$

where λ_i are the computed eigenvalues, μ_j the measured eigenvalues, and w_{ij} are weights. The weights are selected by the criterion

$$w_{ij} = \begin{cases} |\langle \lambda_i | \mu_j \rangle|^2, & \text{for the } j \text{ which maximizes } w_{ij} \\ 0, & \text{for all other } j \end{cases} \tag{6}$$

Here $|\lambda\rangle$ are the computed and $|\mu\rangle$ the observed normal mode eigenvectors. In case the calculated and observed modes are in the same order, this reduces to $w_{ij} = \delta_{ij}$.

The iterations are stopped when either the relative change in χ^2 ($\Delta\chi^2/\chi^2$) or the relative change in each of the parameters ($\Delta x_k/x_k$) from one iteration to the next is less than 10^{-3}. In the first case the fit is not improving anymore, and in the second case the parameters have stopped changing; in neither case is there anything to be gained by continuing the iterations.

The algorithm involves computing derivatives of the eigenvalues with respect to all free parameters. This can be done numerically without the need for a rediagonalization of the second-derivative matrix **F**, by using the Hellman-Feynmann like relation[13,21]:

$$\frac{\partial \lambda_i}{\partial x_j} = \left\langle \lambda_i \left| \frac{\partial \mathbf{F}}{\partial x_j} \right| \lambda_i \right\rangle \tag{7}$$

This relation is exact when the exact eigenvectors $|\lambda_i\rangle$ are used. The fact that the same $|\lambda_i\rangle$ can be used and only the derivatives $\partial \mathbf{F}/\partial x_j$ have to be evaluated numerically provides for substantial savings in computation time, especially for a large number of parameters.

The calculated normal mode eigenvectors were characterized by inspection of projections of the eigenvectors $|\lambda\rangle$ onto the internal coordinates u of the system

$$u = L|\lambda\rangle \tag{8a}$$

$$L_{ij} = \frac{\partial u_i}{\partial R_j} \tag{8b}$$

and of the potential energy distribution (PED) of each mode among the internal degrees of freedom. The energy in mode j along internal coordinate u_i is computed from[22]

$$E_{ij} = k_i \sum_{l,m} \Delta R^j_l \frac{\partial u_i}{\partial R_l} \Delta R^j_m \frac{\partial u_i}{\partial R_m} \tag{9}$$

where k_i is the force constant (K_b, K_θ, n^2K_ϕ, or K_ω) for internal coordinate i in eq. (1), ΔR^j_l is the component along Cartesian coordinate R_l of eigenvector j, and the summation is over all $3N$ Cartesian coordinates of the N-atom system.

One difficulty in the method is to obtain suitable initial estimates for the experimental $|\mu\rangle$. Ideally, the Cartesian coordinate representation of all the eigenvectors $|\mu\rangle$ would be used. However, in most cases even well-assigned spectra for larger molecules are presented only with verbal descriptions of the major features of each mode. It would be possible to try to construct the Cartesian eigenvectors from such descriptions. We approached the problem by calculating a spectrum with EF1 and characterizing the eigenvectors in terms of their principal contributions. This served to establish an approximate correspondence between the experimental and the calculated modes. The calculated set of eigenvectors was then rearranged in the same order with respect to frequency as the corresponding experimental modes. The thus rearranged set was used as a guide ($|\mu\rangle$) for the next cycle of refinement. As the optimization progresses the modes change character and the correspondence between calculated and experimental modes usually has to be altered one or more times. Well-assigned spectra for nucleic acid components are scarce, but we have found one spectrum representing each of the main groups: phosphate, sugar, and base.

Variation of Parameters

The available vibrational data from uracil, diethylphosphate, and tetrahydrofuran used in the parameter fitting are restricted to frequencies larger than 200 cm^{-1}; i.e., the known modes involve mainly bond stretchings and bond angle vibrations, with comparatively little information about torsions and out-of-plane motions. The parameters that were optimized in going from EF1 to EF2 were, therefore, limited to those that are relevant to these kinds of motion. For all three

model systems, all bond stretching and bond angle bending force constants were optimized; in addition the improper torsion force constants were optimized for uracil. No equilibrium values, torsional parameters (but see below), nonbonded parameters, or charges were changed in these fits. To increase the efficiency of the optimizations, especially in the beginning when the deviations from the observed data were largest, the variations were performed in several steps, optimizing different (loosely coupled) subgroups of the force constants, before a final cycle with all of them free to vary.

The parameters obtained with these model systems were carried over to their counterparts in the full nucleic acid parameter set; that is, uracil to the other bases, dimethylphosphate to the phosphate and tetrahydrofuran to the sugar groups.

To give a better representation of the sugar pucker potentials the dihedral force constants, and the number of dihedrals used, were changed by trial-and-error for both ribose and deoxyribose between EF1 and EF2.

Finally, the hydrogen bonding parameters in EF2, which uses a different functional form than EF1, were taken from the corresponding parameterization used for proteins in the CHARMM program (version 17).[16]

The full parameter set used in EF2 is given in Appendix 1, and Appendix 2 lists a corresponding set for a representation with all hydrogens explicitly treated; the details of the all-hydrogen parameterization will be given elsewhere. (L. Nilsson, B.M. Pettitt, and M. Karplus, to be published.)

III. RESULTS

The development of EF1 into EF2 is outlined and a comparative analysis of energy functions EF1, EF2, and EF3 with respect to several experimentally accessible quantities for small model systems (Fig. 2) of relevance for nucleic acids is presented.

Static Properties

Hydrogen Bonding Hydrogen bonding energies and geometries of Watson-Crick base-

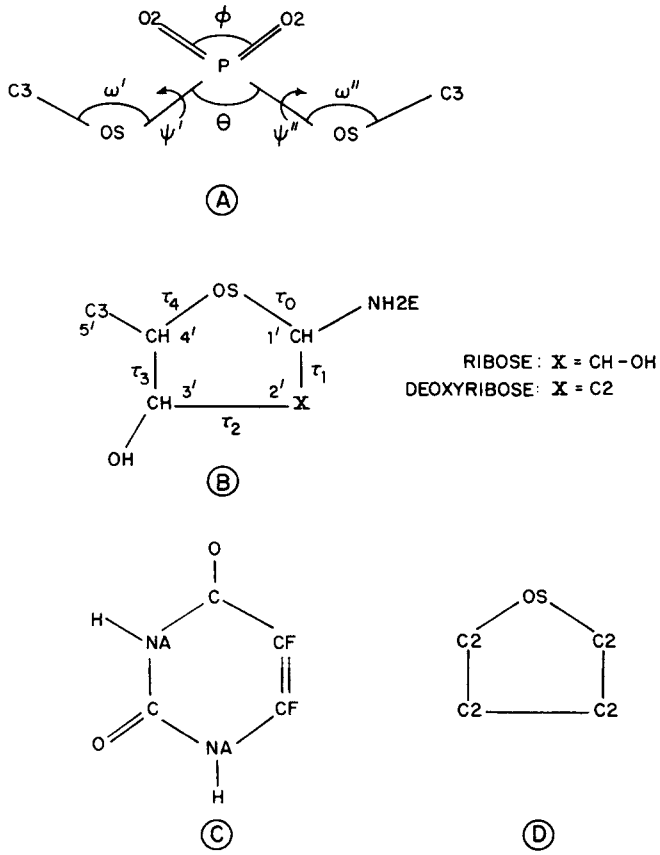


Figure 2. Model Compound nomenclature (internal coordinates, and atomtypes). A, Dimethylphosphate (DMP); B, Ribose and Deoxyribose, C, Uracil; D, Tetrahydrofuran (THF).

pairs (G-C and A-T) of the three energy functions were checked by a calculation in which the two bases (without any sugar or phosphate) were placed in the XY plane in the standard Watson-Crick position, and then one of them was moved away along the PurN1-PyrN3 axis. The total energy of the system was calculated at different $N1$-$N3$ separations in the range 2.5 to 12.0 Å (Fig. 3) without any minimizations. It can be seen that all energy functions give similar well depths and minimum positions, with the somewhat wider well in EF2 due to its less steep ($s^{-6} - s^{-4}$) hydrogen bond potential; such a distance dependence is in accord with *ab initio* calculations on related systems such as the formamide dimer.[18] The most important difference between the energy functions is EF3's use of the special hydrogen-bond term to provide the repulsive inner wall, and of other terms (mainly electrostatic) in the potential to give the attractive part. When used together with a switching function and nonbond cutoffs this leads to the large oscillations in energy at distances close to the cutoff. EF1 and EF2 on the other hand get 30%–50% of their well depth from the explicit hydrogen-bond energy term, so that they only have a small ripple in the potential in Fig. 3. A nonbond cutoff is used to reduce the number of interactions, and thereby the computational time; the ripple remaining in EF1 and EF2 is expected to have negligible effects on dynamics simulations.

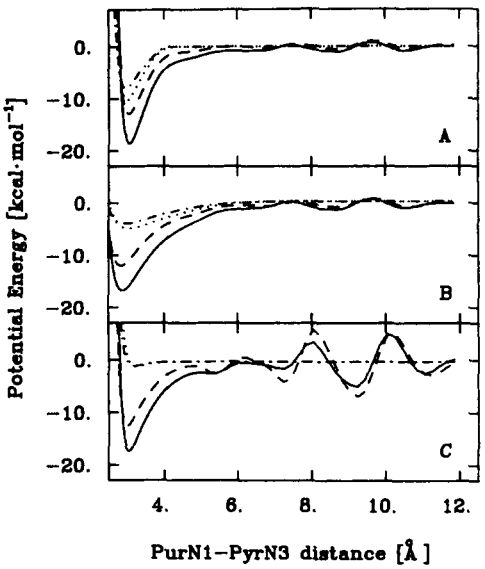


Figure 3. Watson-Crick basepairing potential energy as a function of the distance between the N1 (purines) and N3 (pyrimidine atoms. A, EF1: B, EF2: and C, EF3. G=C total energy (——), G=C hydrogen bond energy term (....). A=T total energy (– – –), A=T hydrogen bond energy term (–.–.).

Free minimizations of the energy of the basepairs were also performed, starting from the Watson-Crick position. Table I gives the optimal $N1$–$N3$ distances and the energy differences between this minimum energy conformation and one in which the bases are separated by 100 Å. The conformations[23,24] and energies[25] obtained are close to the experimental results in all cases. The exact distances are somewhat uncertain since the crystallographic distance data come from full nucleotides, with a sugar-phosphate backbone, while the model systems used here are the bases themselves.

Stacking The pure stacking energy of a basepair is a rather difficult quantity to assess experimentally. Further, the solvent contribution to the energy is not clear. A series of recent gas-phase mass-spectrometric,[25] calorimetric,[26,27] and spectrophotometric melting[28] experiments, as well as calculations[29,30] suggest ΔH values in the range 5–15 kcal/(mol of base pair). The base-base distance 3.4 Å found in standard B-DNA geometry, which is affected by factors other than stacking interaction energy minimum, is taken

Table I. Hydrogen bonding results.

Energy Function	GC pair ΔE[a] [kcal/mol]	r[b] [Å]	AT pair ΔE [kcal/mol]	r [Å]
EF1	22.6	2.93	14.0	2.93
EF2	19.8	2.82	12.7	2.79
EF3	21.9	2.87	12.9	2.83
Experiment[c]	21.0	2.95	13.0	2.82

[a] ΔE is the energy difference between the system at the optimum distance and at infinite separation.
[b] r is the minimum energy N1 (purine)-N3 (pyrimidine) distance.
[c] ΔE from gas phase data of Yanson *et al.*[25] distances from X-ray data of Seeman *et al.*[23] (AU) and Rosenberg *et al.*[24] (GC).

as an indicator of the appropriate distance for the pure base-base interaction that we have examined.

In simple model calculations, in which all ten different combinations of Watson-Crick base pairs were placed in a B-DNA conformation (i.e., with coplanar bases at helical twist angles of 36°) with the z axis as the helix axis, the energy of the system was calculated as a function of distance between the base pairs along the z axis. The stacking energy was estimated as the difference between the minimum energy, which in all cases occurred at Δz = 3.2–3.3 Å, and the energy at infinite separation (Table II). The spread in stacking energy of 6 kcal/(mol of base pair) between the most (5′-GC-3′)$_2$ and the least (5′-TA-3′)$_2$ favorable case is the same for all energy functions and similar to other calculations.[29] As to the absolute values, it is difficult to make a quantitative evaluation since the experimental data are in solution and the calculations are made *in vacuo*.; the lower energies (12–18 kcal/(mol of basepair)) given by EF3 are somewhat closer to the experimental solution estimates. The difference between EF2 and EF3 is due to the slightly larger van der Waals radii used in EF3.

Table II. Base stacking energy in B-DNA conformation.[a]

Bases	Energy Function EF1	EF2	EF3
(5′) G-C (3′) (3′) C-G (5′)	23	24	18
A-T C-G	23	24	16
A-T T-A	22	23	14
A-T G-C	23	23	15
C-G A-T	22	23	14
A-T A-T	22	22	14
C-G C-G	22	22	15
C-G G-C	22	22	14
G-C A-T	22	22	14
T-A A-T	18	18	12

[a]Energy per mol of base pair [kcal mol^{-1}]; only the bases are included so that there is no contribution from the backbone.

Small molecule crystals; heat of sublimation The heat of sublimation (ΔH_{subl}) of small molecules is a thermodynamic parameter that is sensitive to the electrostatic and van der Waals terms in the potential energy. We computed ΔH_{subl} for diketopiperazine (DKP), which forms hydrogen bonds in the crystal lattice,[31] and for n octane, where the crystal packing is solely determined by nonbonded contacts.[32] Although neither molecule is directly related to nucleic acid components, the two molecules were chosen as representatives of polar (DKP) and nonpolar (n octane) groups. Thus, they provide a test of the validity of the overall scale of the charges and van der Waals parameters. Corresponding structural and thermodynamic studies of nucleic acid crystals are in progress.

From the minimizations of DKP and n octane in vacuum and in the crystal lattice ΔH_{subl} was calculated from[14,33]:

$$\begin{aligned}\Delta H_{subl} &= (E_{vac} + pV) - E_{xtal} - 3RT \\ &= E_{vac} - E_{xtal} - 2RT \end{aligned} \quad (10)$$

where E_{vac} and E_{xtal} is the energy per molecule in the vacuum and crystal minimizations, respectively, and R is the gas constant; $3RT$ is a simple estimate of the loss of translation/rotation energy in going from the gas phase to the crystal. No correction is made for changes in the internal vibrational or overall molecular vibration or libration contribution to the enthalpy.

The crystal computations were constructed from eight DKP or four n-octane molecules in a complete unit cell, and the IMAGES facility[34] in the program CHARMM was used with interactions to neighboring cells included up to a cutoff distance of 12.0 Å to account for the rest of the lattice. Minimizations were carried out, with the crystallographic lattice parameters kept fixed, until the rms average force per degree of freedom was less than 0.3 kcal mol^{-1} Å^{-1}.

In a minimization energy is consistently being removed from the system until no further changes in energy or structure are seen. In effect, a particular conformational substate has been frozen in, and the minimized structure corresponds to a T = 0° K state. As a further check on the results, E_{vac} and E_{xtal}

Table III. DKP Minimizations and Dynamics Averages.

Energy-Function	Vacuum R_G[a] [Å]	Vacuum rms[b] [Å]	Vacuum E[c] [kcal/mol]	Crystal R_G[a] [Å]	Crystal rms[b] [Å]	Crystal E[c] [kcal/mol]	ΔH_{subl}[d] [kcal/mol]	%diff[e]
Experimental								
Minimized				1.876			24.8	
EF1	1.86	0.345	−11.34	1.89	0.25	−32.93	20.4	−18
EF2	1.81	0.328	−11.06	1.83	0.25	−33.34	21.1	−14
EF3	1.90	0.198	1.34	1.89	0.12	−30.19	30.3	+22
Dynamics								
EF1	1.86	0.326	−5.24	1.86	0.25	−24.48	18.0	−27
EF2	1.81	0.323	−4.08	1.82	0.26	−26.22	20.0	−16
EF3	1.83	0.174	12.93	1.83	0.08	−14.14	25.9	+4

[a] R_G is the radius of gyration.
[b] rms is the root mean square difference between the atomic positions of the minimized structure and the crystal structure.
[c] E is the total minimized potential energy per molecule.
[d] ΔH_{subl} is the heat of sublimation at T = 300 K, as defined in the text. The experimental ΔH_{subl} is taken from ref. 33.
[e] %diff is the difference in percent between the calculated and the experimental value.

were also computed as the average potential energy of molecular dynamics simulations of DKP in vacuum and in the crystal lattice, at T = 300 K.

The results for DKP in Table III show that very similar ΔH_{subl} values, within 30% of the experimental result are obtained. The radius of gyration R_G is in all cases within 3% of the x-ray value. The rms deviations between the minimized structures (or the average structure in the MD simulations) and the x-ray structure vary from 0.25 to 0.35 Å. The DKP structure, which in the crystal, is planar ends up slightly puckered in a boat form as a consequence of minimization in vacuum and in the crystal; the atoms are in all cases within 0.5 Å (vacuum) or 0.4 Å (crystal) of the best-fit least-squares plane. The results for EF1, EF2, and EF3 are 0.38, 0.36, and 0.40 Å, respectively. The MD average structures are very similar to those from the minimizations. As expected the inclusion of the surrounding lattice decreases the rms shifts.

For octane (Table IV) EF1 and EF2 are 8% below, and EF3 is 38% below, the estimated value of ΔH_{subl} at T = 0° K. The initial all-trans conformation is conserved in both vacuum and crystal minimizations; the dihedral and bond angles change less than 2°.

Dimethylphosphate (DMP) Energy Contour Map The nucleic acid backbone conformation is to a large extent determined by the phosphodiester bridge between neighboring sugars. DMP is a model system for which a number of semiempirical and *ab initio* calculations have been made (see ref. 35, and references therein). Further, x-ray structural data on nucleic acids provide information on the observed range of conformations.

Table IV. Octane minimizations.

Energy-Function	Vacuum E_{min}[a]	Crystal	ΔH_{subl}[b]
EF1	−1.31	−16.3	15.0
EF2	−1.34	−16.4	15.1
EF3	−0.55	−10.6	10.0
Experimental[c]			16.3

[a] Total minimized potential energy per molecule [kcal mol^{-1}].
[b] Heat of sublimation at T = 0 K [kcal mol^{-1}].
[c] Reference 19.

Fully relaxed energy contour maps of DMP were computed by minimizing the structure with the phosphodiester dihedral angles ψ' and ψ'' (see Fig. 2 for definitions of angles) constrained at 30° intervals. The symmetry of the system was used to restrict the region investigated to $0° \leq \psi' \leq 360°, 0° \leq \psi'' \leq \min(\psi', 360°-\psi')$. At each (ψ', ψ'') point all other degrees of freedom were minimized using the NRAPH minimizer in the CHARMM program until convergence was achieved (rms force less than 0.001 kcal mol^{-1}Å^{-1}). This is important since it allows strain induced by unfavorable van der Waals contacts to be relieved by small adjustments in angles.

Figure 4 shows the maps obtained for EF2 and EF3. The torsion force constants dominate in this problem and so EF1 (data not shown) and EF2 give virtually identical maps. The *gg* conformation is in all cases the lowest; it is 1–2 kcal mol^{-1} below the metastable *tt* conformation with the *gt* and *gg*$^-$ in between (Table V). This is consistent with the relative numbers of the different conformations that are found in crystals.[5] It also agrees with the best quantum mechanical calculations[35] (CNDO and *ab initio* STO-3G with partial geometry optimization), except for the *gg*$^-$ conformation which in Gorenstein's calculation[35] is a local maximum about 2 kcal mol^{-1} above the global *gg* minimum. The internal coordinates at the three minima are close to the QM and x-ray data (Table V), with the QM angle values somewhat smaller than what is found in the present calculations and the x-ray analysis. The relaxation in θ from *gg* to *tt* reduces the $\Delta E = E_{tt} - E_{gg}$ in the same way here as in the QM calculations,[35] i.e., θ_{tt} is 3° (EF2) to 8° (QM) smaller than θ_{gg}; this reduces E_{tt} by relieving the van der Waals contacts of the methyl groups with each other and with the double-bonded oxygens.

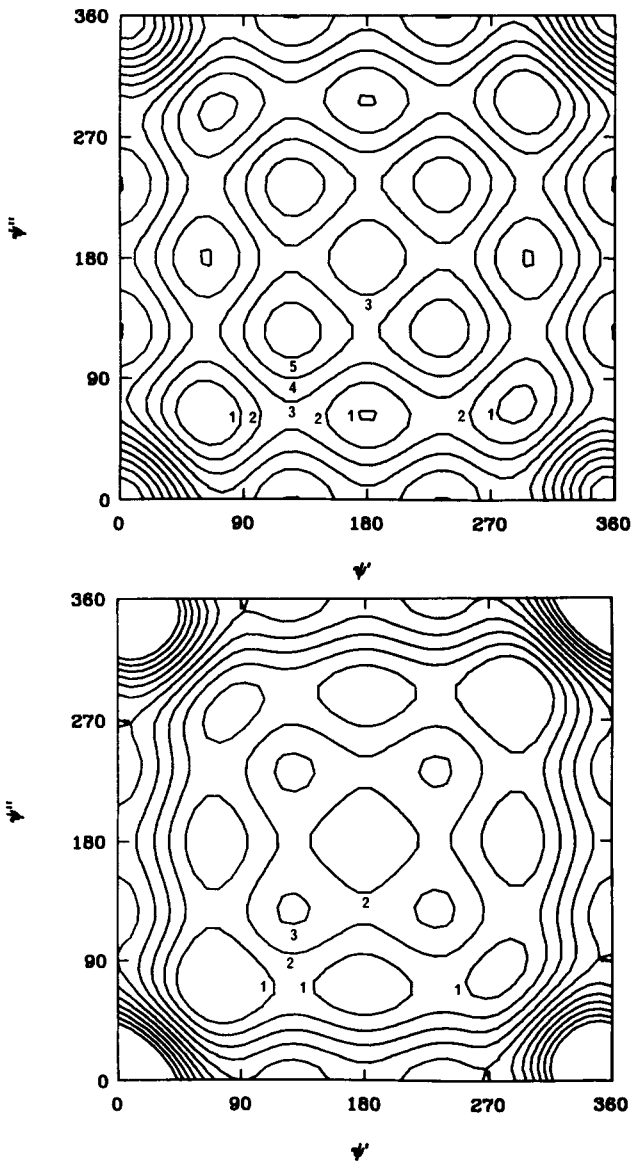


Figure 4. Dimethylphosphate energy contour maps; dependence on the phosphodiester dihedral angles ψ' and ψ'' (defined in Fig. 2). A, EF2 and B, EF3.

Sugar pucker energies The pseudorotation energy surface of a sugar model system for ribose and deoxyribose (Fig. 2) was studied by computing the total energy as a function of geometry. The ring dihedral angles τ_j (j = 0, 1, 2, 3, 4) were constrained to give puckered conformations at 15° intervals from 0° to 360° of the pucker phase P as defined by Altona and Sundaralingam[36]:

$$\tau_j = \tau_m \cos[P + 0.8\pi(j - 2)] \qquad (11)$$

where P is the pucker phase angle and the maximum torsion τ_m was set at 38°. Although this description of the pucker state is approximate, it was used here instead of the exact description of Cremer and Pople,[37] to facilitate comparisons with previous studies,[38,39] which used the same formulation and the same value for τ_m. At each of the conformations the energy was minimized with respect to the Cartesian coordinates of all atoms; harmonic constraints, with force constants equal to 400 kcal mol^{-1} Å^{-2}, were applied to the τ_j. Minimizations were carried through to convergence with NRAPH (rms energy gradient < 0.0001 kcal mol^{-1} Å^{-1}).

The minimization was started at P = 0° by applying the appropriate constraints on a set of initial coordinates obtained from a minimization of a C3′-endo structure, one of the known energy minima. At each point along the P circle, the final conformation from the minimization at the previous point was used as the new starting conformation; the convergence of the NRAPH minimizer is such that this procedure, which is faster than a grid search, should not introduce any artifactual behavior. This is verified by the fact that energies computed at the first (P = 0°) and last (P = 360°) points, which represent the same conformation, agree to within 0.1 %.

Table V. DMP energy and geometry at local minima.

(ψ', ψ'') Conformation	Energy-Function	ΔE	θ	ϕ	ω'	ω''	ψ'	ψ''
	EF1	0.00	104	118	124	124	63	63
	EF2	0.00	104	119	124	124	63	63
gg	EF3	0.00	103	120	121	121	70	70
	QM[a]	0.00	99	126	112	112	68	68
	x-ray[b]	45	105	120	122	122	73	73
	EF1	0.50	105	118	125	125	72	−72
	EF2	0.65	105	119	125	125	71	−71
gg^-	EF3	0.90	104	120	121	121	80	−80
	QM[c]	>2	—	—	—	—	—	—
	x-ray	—	—	—	—	—	—	—
	EF1	0.96	103	119	123	123	63	180
	EF2	0.94	103	119	124	123	62	180
gt	EF3	0.81	103	120	121	120	69	179
	QM	0.14	95	124	112	112	75	179
	x-ray	6	99	120	122	122	74	169
	EF1	1.99	101	120	123	123	180	180
	EF2	1.93	101	120	120	123	180	180
tt	EF3	1.69	103	120	120	120	180	180
	QM	0.89	91	123	111	111	180	180
	x-ray	—	—	—	—	—	—	—

[a]CNDO and STO-3G calculations with optimization of all angles in this table.[35]
[b]The x-ray data are the number of structures in a given conformation and average parameters from a survey of the Cambridge Crystal Data Bank (P. Murray-Rust unpubl.; see ref. 5).
[c]In the CNDO calculations only θ was relaxed at all (ψ', ψ'') points[35] and gg^- did not come out as a local minimum; therefore, no optimization of the other degrees of freedom was made for this case.

As has been pointed out by Olson and coworkers[38] the sugar energy surface as a function of P and τ_m is very sensitive to the exact geometry at each point. Some preliminary runs (not shown) using a larger system with the methyl at C5′ replaced by —C—O—H were very dependent on initial conditions, and had large fluctuations in the curve of the energy as a function of P. The difficulties were due to local minima, with the substituents locked into conformations that did not represent the global energy minimum at a given pucker phase. The problems disappeared when the size of the system was reduced to the system shown in Fig. 2. This model retains the main characteristics of a (deoxy)ribose ring, and has the hydroxyl groups represented by extended atoms; a similar model is used by Olson.[38,39] In the actual potential (see Appendix 1), all polar hydrogens are explicitly treated, also in the sugar hydroxyl groups. The partial atomic charges are the same as shown in the Figure in Appendix 1, with the following exceptions:

$$O3' = -0.15, C3' = 0.15, O2' = -0.15,$$
$$C1' = 0.25, \text{ and } NH2E = -0.05.$$

The potential curves (Fig. 5) show that EF1 has minima in the C3′-endo (P = 0°–36°) and C2′-endo (P = 144°–180°) regions with C3′-endo the more stable, for both ribose and deoxyribose, and a 1–2 kcal mol^{-1} barrier between the two states. The EF3 potentials are much flatter than the others, with both barriers of roughly equal height, 1–2 kcal mol^{-1}. The deoxyribose has both its minima considerably shifted from P = 0° toward P = 90°

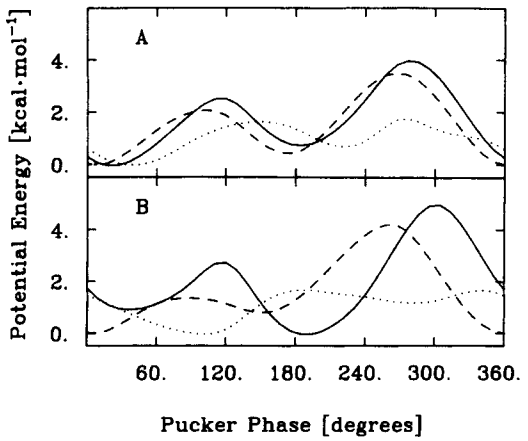


Figure 5. Sugar energy as a function of pucker phase angle. A, Ribose and B, deoxyribose, EF1 (– – –), EF2 (——), EF3 (....).

and from $P = 180°$ toward $P = 270°$, neither of which are normally found puckering states. The EF2 potential has been adjusted to give a behavior in accord with that recommended by Olson[38,39]; C2′-endo ($P = 144°–180°$) is the more stable minimum for deoxyribose and C3′-endo ($P = 0°–36°$) is slightly more stable than C2′-endo for ribose. The $P = 120°$ barrier (2 kcal mol^{-1}) is lower than the $P = 270°$ (4–5 kcal mol^{-1}) barrier.

In Olson's potential[38] the barriers are dominated by van der Waals interactions, which can be attributed to the fact that the ring substituents were held in fixed positions relative to the ring, whereas in the present minimizations the substituents can move to relax some of the bad contacts. In EF2 the main contribution to the barriers, therefore, comes from the torsion term in eq. (1), with the *vdW* term being significant only for the $P = 270°$ barrier. This is achieved by using two-fold dihedral potential terms for some of the C—C—C—O and O—C—C—O torsions in addition to the normal three-fold terms. The arrangement of the dihedrals is a modification of that used by Olson,[38] with a larger number of dihedral terms than are otherwise used in energy functions for proteins or in other parts of the nucleic acids; details are given in Appendix 1. All energy terms except the hydrogen bonding term in eq. (1) are present in the calculations, but only dihedral, van der Walls and electrostatics show any variation with pucker phase for the model system.

Dynamic Properties

In addition to the static properties analyzed in the preceding section, the energy functions were evaluated by employing them in test calculations of certain dynamic properties of nucleic acid components and a nucleic acid oligomer. First, normal mode analyses were made of some nucleic acid constituents, with diethylphosphate and tetrahydrofuran representing the backbone and uracil as a base. For these compounds well-assigned spectra are available and were used in deriving EF2 from EF1. Normal mode spectra calculated for 1-methyl uracil, adenine, and 9-methyl adenine are also presented, although experimental comparisons are less straightforward because there are no fully assigned results. Finally the influence of different electrostatic models on the behavior of MD simulations of a DNA hexamer duplex $d(CGTACG)_2$ was studied.

Diethylphosphate (DEP) Normal Modes DEP was the first model system to be subjected to the parameter fitting subsystem in CHARMM. With EF1 parameters and x-ray coordinates for DEP,[40] the bond length and bond angle force constants were adjusted to fit the calculated normal-mode spectrum to the well-assigned experimental Raman-spectrum measured by Brown and Peticolas.[41] Since in this spectrum only modes with wave numbers >200 cm^{-1} are present, the four lowest modes, which are torsional vibrations, were left out of the fitting process. The initial *tggt* conformation, from the crystal structure, is conserved to within 10° for dihedral angles and 5° for bond angles in the final energy minimized (EF2) structure.

From the eigenvalues and eigenvectors computed with EF1 a guess was made as to how these modes corresponded to the observed modes, and this guess was used to keep track of the modes during the initial phase of the optimization. The bond force constants were first allowed to vary, followed by a step in which the angle force constants were varied, which in turn was followed by a step during which all bond and angle force constants were optimized. This reduced the χ^2 value from 1869 to 268 [see eq. (5)].

The parameters resulting from this phase (*PAR-A*) were then used for a new mode assignment and the optimization was restarted with the C3-C2,*P*-O2 and *X-P-X* (*X* is any atom) force constants free, giving parameter set (*PAR-B*). As the PED and general characters of the modes now agreed quite well with the experimental analysis, a final pass was made through the optimizer without keeping track of the modes, and with all angle and bond force constants free. Attempts to vary the van der Waals and dihedral parameters did not improve on the spectrum obtained with this final parameter set (*PAR-C*), which was incorporated into EF2.

The course of the optimization is depicted in Figure 6, which shows the changes in the

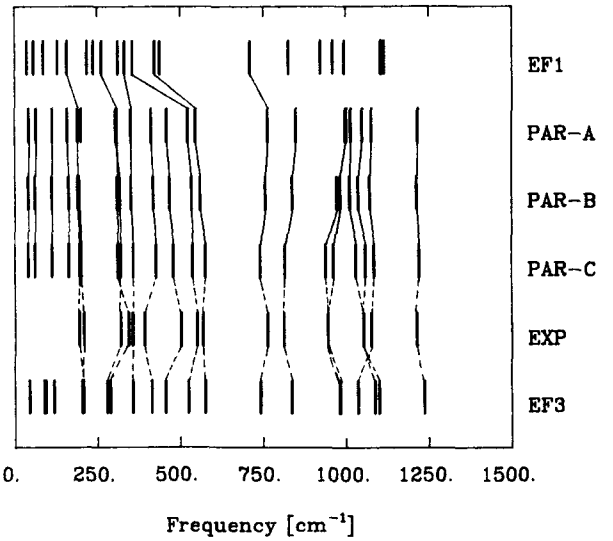


Figure 6. Diethylphosphate modes during fitting process, from EF1 to EF2. EF3 and experimental spectra also indicated. The thin solid line join modes whose eignevectors satisfy $|\langle\lambda_i|\lambda_i\rangle|^2 \geq 0.6$; the dashed lines joining the experimental spectrum to EF2 and EF3 are qualitative indicators of the correspondence between the spectra based on assignments in Table VI.

spectrum; Table VI gives the final parameter values. It is clear that the first phase, leading to *PAR-A*, and to a lesser extent the final phase is where the main improvement occurred. The four lowest (torsional) modes do not change appreciably, indicating that they are essentially decoupled from the remaining modes. These low modes correspond to motions in the shollow valleys of the DMP energy contour maps (Fig. 4). It would be very useful to have data for these low frequencies since they depend primarily on the dihedral angle potential, which is of paramount interest for conformational calculations.

Table VII lists the potential energy distributions (PED) of the EF2 DEP spectrum, as well as of the normal coordinate assignment made by Brown and Peticolas to their Raman spectrum.[42] The agreement in general is good, although there are some differences in the dominant contributions to the normal modes (e.g., the P—O bond stretching component of mode 13 in EF2 is absent in the assignment of Brown and Peticolas).

Uracil Normal Modes The recently published Raman spectra of uracil in Ar and N_2 matrices[40] provide an excellent opportunity to improve some of the force constants. In such an experiment perturbations from the environment are minimized, and the data are directly comparable to standard molecular mechanics vacuum calculations.

Table VI. DEP and DMP parameters.

Parameter	Equi.	Force Constants EF1	PAR-A	PAR-B	EF2	EF3
Bonds[a]						
P O2	1.480	450.	*533.[b]	*525.	*528.	525.
P OS	1.610	300.	*240.	240.	*237.	230.
OS C2	1.430	300.	*272.	272.	*292.	320.
OS C3	1.43	300.	272.	272.	292.	320
C3 C2	1.525	300.	*260.	*230.	*201.	260.
Angles[c]						
P OS C2	120.5	46.5	*42.6	42.6	*47.6	100.
P OS C3	120.5	46.5	42.6	42.6	47.6	100.
OS P OS	102.6	46.5	46.5	*45.6	*48.1	45.
OS P O2	108.2	46.5	*85.5	*92.9	*98.9	100.
O2 P O2	119.9	70.0	*121.5	*129.4	*140.5	140.
OS C2 C3	111.0	46.5	*154.3	154.2	*150.5	80.
Phi[d] (not varied)		Periodicity				
X C2 OS X		3			0.50	1.45
X OS P X		3			0.75	0.75
OS P OS C2		3			0.75	0.75
OS P OS C2		2			0.75	0.25
OS P OS C3		3			0.75	0.75
OS P OS C3		2			0.75	0.25

[a]Bond lengths in Å and force constants in kcal mol^{-1} Å^{-2}.
[b]The * indicates adjusted parameters.
[c]Bond angles in degrees and force constants in kcal mol^{-1} rad^{-2}.
[d]Torsion angle force constants in kcal mol^{-1}. All these dihedrals have a phase angle (δ) of 0 degrees. The double entries for the P—OS dihedrals indicate the use of two cosine terms, with periodicity 2 and 3.

Table VII. DEP normal mode spectra.

Mode	EF2 Frequency [cm^{-1}]	EF2 Character[a]	Experimental[b] Frequency [cm^{-1}]	Experimental[b] Character[a]
27	1224	P=O (0.90)a	1215	P=O (0.84)a
26	1081	P=O (0.50)s	1077	P=O (0.53)s
25	1051	C—O (0.69)a	1053	C—O (0.71)
24	1030	C—O (0.54)s	1053	C—O (0.65)
23	961	C—C (0.35)a	945	C—C (0.48)
		C—C—O (0.41)a		C—C—O (0.22)
22	937	C—C (0.58)s	945	C—C (0.59)
		C—C—O (0.28)s		C—C—O (0.28)
21	812	P—O (0.51)a	812	P—O (0.53)a
		C—C (0.35)a		C—C (0.32)
20	739	P—O (0.47)s	763	P—O (0.51)s
		C—C (0.25)s		C—C (0.26)
19	578	O=P=O (0.69)	569	O=P=O (0.55)
18	531	O—P=O (0.50)	551	O—P=O (0.64)
		C—C—O (0.27)a		C—C—O (0.17)
17	482	O—P=O (0.63)	503	O—P=O (0.53)
16	427	O—P=O (0.90)	393	O—P=O (0.39)
				C—C—O (0.36)
15	349	O—P—O (0.26)	357	C—P—O (0.17)
		O—P=O (0.20)		O—P=O (0.16)
				O=P=O (0.25)
14	320	P—O (0.16)s	345	P—O (0.33)
		O—P—O (0.26)		O—P=O (0.29)
13	304	P—O (0.30)a	321	O—P—O (0.41)
		O—P=O (0.29)		O—P=O (0.43)
12	193	C—O—P (0.66)a	210	C—O—P (0.45)
		O—P=O (0.17)		O—P=O (0.44)
11	176	C—O—P (0.54)s	195	C—O—P (0.23)
		O—P—O (0.22)s		O—P=O (0.63)
10	148	torsion		—
9	99	torsion		—
8	61	torsion		—
7	38	torsion		—

[a]The internal coordinates are given, followed by the PED contributions in parentheses; *a* and *s* indicate antisymmetric and symmetric combinations of the listed internal coordinates.
[b]Reference 41.

Starting with EF1 and a minimized uracil conformation generated from the equilibrium values of bond lengths and angles in EF1, a series of optimizations were done as outlined in Figure 7 and Table VIII; the approach is analogous to that used for DEP. The first spectrum computed from EF1 had a number of clearly displaced lines. Because of the limited radius of convergence of nonlinear optimization procedures, it is necessary to use as good a starting guess as possible; some parameters were, therefore, changed manually as suggested by the first EF1 spectrum before the optimization program was started. Since the uracil spectrum is more complex than the DEP spectrum, and the starting set was worse, the present case involved more manual interventions with inspection and reassignment of calculated to observed modes than did the DEP fitting.

The characters of the normal modes found with EF2 and in the assignment of the experimental spectrum[42] are given in Table IX. The final uracil parameters give a reasonable representation of the spectrum, although there are some differences in the middle of the spectrum (modes 18, 19, and 22 in the experimental assignment). These may be due, in part, to the extended atom representation used in the calculations; modes in this region are strongly coupled to hydrogen vibrations, e.g., this is evident from the changes in the calculated spectrum upon methylation of N1 (Fig. 8).

Because of correlations among the many parameters, this set is not unique; variations of 20%–30% in individual parameters can be compensated for by variations in certain of the others, and a similar spectrum would be produced. The uracil conformation stays vir-

Table VIII. Uracil parameters.

Parameter	Equi.	Force constants EF1	PAR-B	PAR-C	PAR-E	PAR-F	PAR-G	EF2	EF3
Bonds[a]									
H NA	1.010	300.	470.	*472.[b]	*472.	472.	*472.	*472.	434.
O C	1.229	600.	675.	*669.	*633.	633.	*498.	*612.	570.
CF C	1.444	450.	450.	*396.	*415.	415.	*450.	*386.	410.
CF CF	1.350	450.	600.	*671.	*671.	671.	*550.	*685.	549.
NA C	1.388	450.	400.	*243.	*284.	284.	*462.	*337.	418.
CF NA	1.365	450.	450.	*468.	*371.	371.	*321.	*301.	448.
Angles[c]									
C NA H	116.5	70.	35.	35.	*44.	*30.	30.	*35.	35.
CF NA H	115.5	70.	35.	35.	*33.	*45.	45.	*30.	70.
C NA C	126.4	70.	70.	70.	*100.	*99.	99.	*16.	70.
C NA CF	121.6	70.	70.	70.	*100.	*99.	99.	*116.	70.
NA C NA	115.4	70.	70.	70.	*70.	*69.	69.	*62.	70.
NA C O	120.6	70.	70.	70.	*70.	*94.	94.	*69.	80.
C C O	125.3	70.	70.	70.	*80.	*80.	80.	*86.	80.
CF C NA	114.1	70.	70.	70.	*70.	*69.	69.	*62.	70.
NA CF CF	121.2	70.	70.	70.	*100.	*99.	99.	*116.	70.
C CF CF	120.7	70.	70.	70.	*100.	*99.	99.	*116.	85.
Impropers[d]									
H X X NA	0.0	50.	15.	*40.	*44.	44.	44.	*43.	—[e]
C X X CF	0.0	50.	10.	*5.	*4.5	4.5	4.5	*4.5	—
C X X O	0.0	50.	125.	*118.	*110.	110.	110.	*113.	—
CF X X NA	0.0	50.	10.	*5.	*4.5	4.5	4.5	*4.5	—
C X X NA	0.0	50.	10.	*5.	*4.5	4.5	4.5	*4.5	—

[a]Bond lengths in Å and force constants in kcal mol^{-1} $Å^{-2}$.
[b]The * indicates adjusted parameters.
[c]Bond angles in degrees and force constants in kcal mol^{-1} rad^{-2}.
[d]Improper torsion angles force constants in kcal mol^{-1} rad^{-2}.
[e]No impropers are used for uracil in EF3.

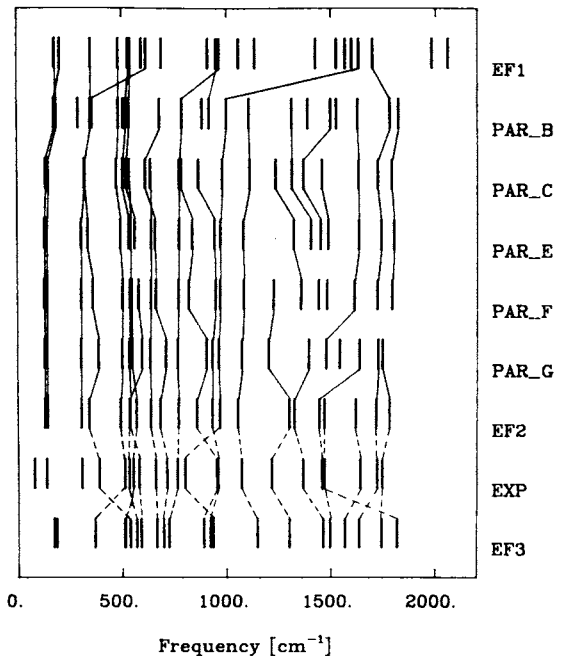


Figure 7. Uracil modes during fitting process from EF1 to EF2. Diagram details as in Figure 5. The N—H stretch modes with $\omega > 3000$ cm^{-1} are left out of the diagram.

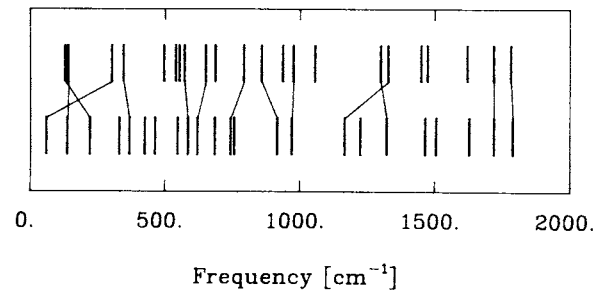


Figure 8. Couplings between normal mode spectra of uracil (upper spectrum) and 1-methyl uracil (lower spectrum) calculated with EF2. The thin solid lines join modes for which $w = |\langle \lambda_i | \lambda_i \rangle|^2 \geq 0.6$. The N—H stretch modes with $\omega > 3000$ cm^{-1} are left out of the diagram.

tually constant in going from EF1 to EF2; the rms displacement in atomic position between the minimized structures is only 0.037 Å.

In a calculation (data not shown) of uracil normal modes with a constant dielectric (instead of $D = r$ in eq. (1) as was used in the other computations) only very minor fre-

Table IX. Uracil normal mode spectra.

Mode	EF2 Frequency [cm^{-1}]	EF2 Character[a]	Experimental[b] Frequency [cm^{-1}]	Experimental[b] Character[a]
30	3451	N—H	3475	N—H
29	3449	N—H	3425	N—H
28	1783	C=O str	1750	C=O str
27	1719	C=O str	1725	C=O str
26	1622	C5—C6 (0.41)	1644	C5—C6 (0.57)
25	1474	ring str	1473	ring str, N—H vib
24	1451	ring str	1461	ring str, N—H vib
23	1329	ring str, N—H vib	1370	ring str, N—H vib
22	1301	N—H vib (0.70)	1219	ring str (0.63)
21	1058	ring str	1076	ring str
20	977	N—H oop, C=O oop	963	ring def (0.75)
19	938	ring str/def	958	ring str (0.61)
18	860	ring str	806	C=O oop
17	793	N—H oop, C=O oop	769	C=O oop
16	687	ring str	719	ring str
15	651	N—H oop, C=O oop	664	N—H oop
14	573	ring str	585	N—H oop
13	554	N—H oop, C=O oop	557	ring def
12	541	ring def, C=O vib	537	C=O vib
11	497	ring def	516	ring def
10	347	C=C vib	393	C=O vib
9	304	oop torsion	310[c]	oop torsion
8	142	oop torsion	140[c]	oop torsion
7	129	oop torsion	80[c]	oop torsion

[a]The internal coordinates with the main contributions to each mode are given. Numbers in parentheses represent the fraction of the mode's total energy along the given internal coordinate. Abbreviations: oop-out of plane motion; str-stretch; def-deformation; vib-in plane bending vibration.

[b]Reference 42.

[c]These frequencies are not seen experimentally. The values in the table are scaled frequencies from the CNDO/2 calculations in ref. 52 upon which Szczesniak *et al.*[42] based all their out of plane assignments.

quency shifts (<3 cm^{-1}) were observed, in agreement with recent calculations by Pettitt and Karplus.[43]

Normal Modes of Adenine and 9-Methyl Adenine The force constants obtained for uracil were transferred to adenine, and normal mode spectra were computed for adenine and 9-methyl adenine. Figure 9 shows these spectra and how they are related. It is evident that methylation changes the whole appearance of the spectrum. Adenine spectra have not been assigned as well as uracil, but we can at least compare some of the major peaks that have been identified in 9-methyl adenine in solution[44] with our calculation. The highest skeletal modes, strongly mixed with NH_2 scissor motions, are observed at 1580, 1623, and 1647 cm^{-1}; the calculated modes at 1525, 1537, 1574, and 1768 cm^{-1} have similar characters and occur in the same frequency range. The observed peak at 1340 cm^{-1} has been assigned to in-phase stretching motions of C^8-N9 and C2-N3; here the corresponding calculated mode is at 1409 cm^{-1}. The ring breathing and NH_2 out-of-plane bending that are assigned to the bands at 729 and 690 cm^{-1} have calculated counterparts at 820 and 625 cm^{-1}, respectively.

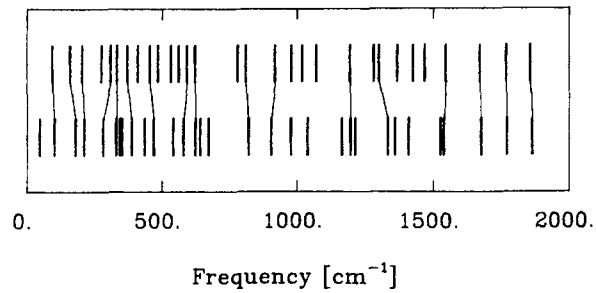


Figure 9. Couplings between normal mode spectra of adenine (upper spectrum) and 9-methyl adenine (lower spectrum) calculated with EF2. For details see caption of Figure 8. The N—H stretch modes with $\omega > 3000$ cm^{-1} are left out of the diagram.

Tetrahydrofuran (THF) normal modes With the THF Raman assignments of Eyster and Prohofsky[45,46] bond stretch and bond angle force constants were fitted from EF1. The optimization did not converge when all three (OS-C2-C2, C2-C2-C2, C2-OS-C2) angle force constants were allowed to vary; instead they were all set equal and were varied together. This resulted in the bond force constants changing from 300 kcal mol^{-1} $Å^{-2}$ in EF1 to 255 for (C2-C2) and 369 for (OS-C2) in EF2, while the angle force constant changed from 46.5 kcal mol^{-1} rad^{-2} in EF1 to 82.5 in EF2. The force constant $K_{OS\text{-}C2}$ had already been fitted to DEP with a somewhat different outcome (292); because of the better fit to the DEP spectrum this value was used in EF2 for THF as well. The THF fitting was then repeated with OS-C2 kept at 292, giving $K_{C2\text{-}C2}$ = 355. and the angle force constants once again came out as 82.5. The spectra in Figure 10 show that EF2 is an improvement over EF1 and EF3 (spectrum taken from ref. 5), especially in the low-frequency region. Further attempts to improve the spectrum by varying torsion force constants and van der Waals parameters either did not converge or gave no improvement. The final fit is not as good as that obtained by Eyster and Prohofsky,[45,46] which is not surprising since their energy function is more complicated (with Urey-Bradley terms) and also includes all hydrogens, in contrast to our extended atom representation.

Molecular Dynamics Simulations The long-range character $(1/r)$ of the electrostatic contribution to the energy, with the neglect of surrounding solvent and dissolved counter ions in a standard vacuum simulation pose more problems for elongated highly charged polyelectrolytes, such as nucleic acids, than for globular proteins that have an interior shielded from the outside. This is illustrated by one published MD simulation of DNA[10] in which the helical structure was unstable; it unwound when electrostatic terms were included in the energy function. This is, however, not always the case, and a number of simulations of *B* and *Z* DNA[11] and of *t*RNA[12] with electrostatic contributions have reported no such lack of stability. All of these stable molecular dynamics simulations used reduced charges for the phosphate group (−0.2e to −0.32e) and a distance dependent dielectric factor[11,15,16]; the electrostatic term in eq. (1) has the form

$$E_{ele}(r_{ij}) = \frac{C_{q_i q_j}}{r_{ij}^2} \qquad (12)$$

The use of a distance dependent dielectric, which is introduced to take into account the solvent and ion shielding in a simplified fashion, should not be extended to simulations where the solvent is included explicitly. Because of the interdependence of the many components of the energy function, some reparametrization may be necessary when the electrostatic model is changed; hydrogen bonding strength is one example where a change of dielectric can have an important effect. This leads to difficulties in comparing results, and may suggest that separate parameter sets are needed for accurate vacuum and solution simulations.

A series of MD simulations of the double stranded DNA hexamer $d(CGTACG)_2$, which has 270 atoms in the extended atom representation with explicit polar hydrogens, were made with different electrostatic options for both EF2 and EF3. Runs of between 5 and 20 ps length were performed starting from energy minimized idealized *B*-form structures.[47] The stability of the DNA helix was assessed from plots of the overall helix-rotation/base, $\Delta\phi$, measured by the angular displacement between successive N1/N9 atoms, with the global helix axis oriented along the *z* axis.

For the stable runs $\Delta\phi$ stays close 36° for all bases (Fig. 11A), whereas in the unstable

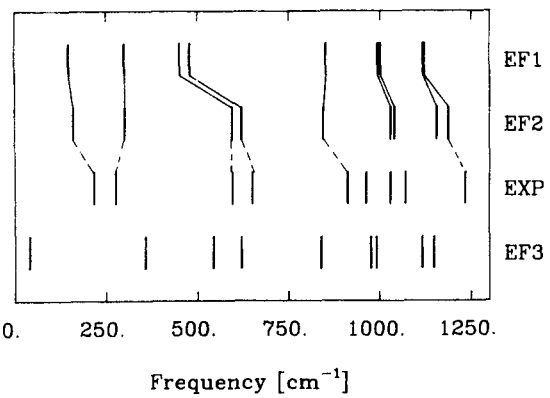


Figure 10. Tetrahydrofuran spectra with solid lines joining corresponding calculated modes in EF1 and EF2; the dashed lines indicate the connections between the calculated modes and the experimental assignments.[45]

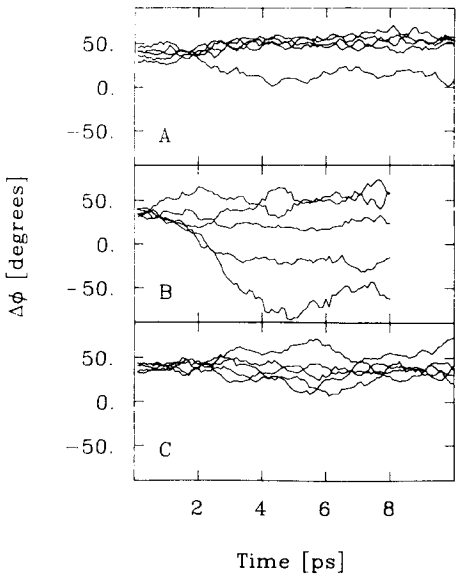


Figure 11. Helix rotation/basepair, $\Delta\phi$, for the middle four base pairs during MD simulation of $d(CGTACG)_2$ as a function of time. A, EF2 with $D = r$, B, EF2 with $D = 2$: C, EF3 with $D = 2$.

cases, $\Delta\phi$ for some of the bases had become negative within 5 ps (Fig. 11B), an indication of unwinding. The results, and details, of the simulations are summarized in Table X. EF2 seems to require a distance dependent dielectric, and EF3 with a constant dielectric leads to a helix, the ends of which are distorted; e.g., the plane of the terminal cytidine base forms a 90° angle to the rest of the base stack. It should be noted that the EF3 simulations use the full charges (−1.0e/nucl) as described by Weiner *et al.*[5]. The stick drawings[48] in Figure 12, showing average structures from three of the MD simulations, illustrate the kinds of stable and unstable helices that were obtained.

When the somewhat larger van der Waals radii used by Jorgensen[49] are introduced for the sugar carbons, unfavorable contacts between the C2′ and one of the doublebonded phosphate oxygens on the 3′-side give rise to a force promoting unwinding of the helix (Table X); this yields an unstable helix even with the otherwise stable EF2 set.

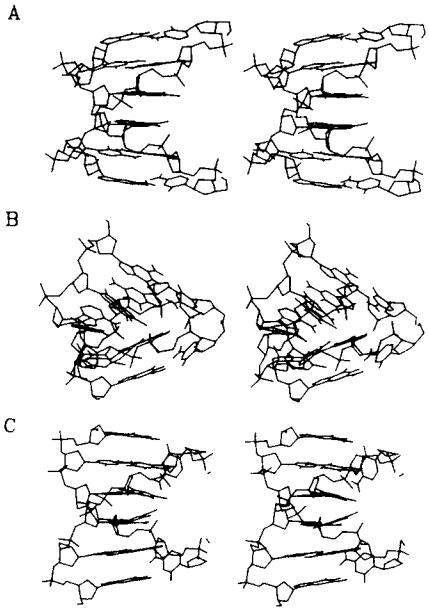


Figure 12. Stereoviews of 2 ps average coordinate sets from MD simulations of $d(CGTAG)_2$. A, EF2 with $D = r$; B, EF2 with $D = 2$; C, EF3 with $D = 2$.

Table X. Stability of $d(CGTACG)_2$ during MD simulations.

	Options: $D = r_{ij}$ Switching function[a]			$D = 2.0$ Shifted potential[b]		
Energy-function	bp/turn[c]	Rise[d]	Char.[e]	bp/turn[c]	Rise[d]	Char.[e]
EF2	9.1	3.20	Stable	16	1.64	Unstable
EF2[f]	—	—	—	14	2.54	Unstable
EF2[g]	38	3.95	Unwinds	18	—	Unstable
EF2[h]	—	—	—	12	3.17	Distorted
EF2[i]	—	—	—	8.8	3.36	Stable
EF3	9.8	3.33	Stable	10	3.26	Distorted

[a]The electrostatic energy term was switched off by a cubic smoothing function[16] between 9.5 and 10.5 Å.
[b]The electrostatic energy term was shifted[16] so that both the energy and the force go to zero at r_{ij} = 10.5 Å.
[c]Base pairs per turn, calculated from a 2 ps average from the MD simulation, for the four middle base pairs.
[d]Rise per base pair [Å], calculated from a 2 ps average from the MD simulation, for the four middle base pairs.
[e]The overall behavior of the oligomer during the simulation is characterized according to how well it retains its initial helical shape.
[f]The net charge on the phosphate group reduced to zero.
[g]The van der Waals radii of the extended carbons (CH, CH_2, CH_3) enlarged to the values used by Jorgensen.[49] The structure obtained with $D = 2.0$ is so deformed that a description in terms of helical parameters is difficult.
[h]D increased to $D = 10.0$.
[i]Switching function used instead of the shifted electrostatic potential.

IV. DISCUSSION

The development of a new empirical energy function (EF2) for nucleic acids in the present study, has led to an improved description of several experimentally accessible properties of importance for both the static (conformational), and dynamic behavior of DNA/RNA. An alternative recently developed energy function (EF3)[5] has made similar improvements over earlier functions. Although there are significant differences between the presently developed function EF2 and EF3, they are more similar to each other than to their common origins.[8,11] One area of difference is in the partial atomic charges. It is of interest, therefore, to compare the rms difference between the charges used in EF2 (EF3) and those recently estimated from x-ray electron densities. The results for cytosine[50] are 0.168e (0.379e) for the base, 0.162e (0.176e) for the dexyribose, and 0.348e (0.384) for the phosphate; thus, the EF2 charges are in somewhat better agreement, particularly for the base, than those from EF3.

All three energy functions that we have examined use the same basic assumptions and approximations. They all are valence type functions with the same general energy terms, include no Urey-Bradly corrections, and employ the extended atom representation for nonpolar hydrogens. Furthermore, the equilibrium values for bond lengths and angles are are the same. The differences are mainly to be found in charges and force constants, and to a lesser extent in the van der Waals parameters and in the choice of torsional angle potentials. Not surprisingly the overall behavior of nucleic acids and their components is rather similar for the three functions.

Nucleic acids have fewer nonpolar hydrogens than proteins, so for nucleic acids it is not quite as costly to include all hydrogens in the model; the number of atoms would typically increase by 30%–40% compared with approximately 100% for proteins. An all hydrogen nucleic acid energy function has (Appendix 2) been developed and is being employed in the analysis of proton NMR data by use of energy minimization and restrained dynamics techniques.[51]

The hard modes of motion, bond stretches, and angle vibrations, as shown in the normal mode analyses of DEP, uracil and THF, are in good agreement with experimental data for EF2 which is a significant improvement from EF1 in this respect. In comparison to other energy functions[41,44] EF2 and EF3 differ in some force constants by 30%–40%, which is also a variation seen between EF2 and EF3. Because of the couplings between the various internal degrees of freedom differences of this magnitude in individual force constants are not crucial, and the spectra are almost as well represented by EF2 as they are by the functions more specifically geared towards vibrational calculations. The STO-3G calculation of in-plane uracil modes,[52] upon which the experimental assignment was based, is more accurate than the EF2 spectrum.

For the softer modes associated with dihedral and nonbonded energy terms we have data only for static properties. Base pair hydrogen bonding and stacking energies and geometries are similar for all three functions and in good agreement with observations. Intermolecular van der Waals and electrostatic interactions in EF2 give the ΔH_{subl} for small molecule crystals within 20%. The dimethyl phosphate potential which is very important for the nucleic acid backbone is also very similar for all three functions and is in accord with observed structual data and QM calculations. The sugar pucker potential is best modelled by EF2.

The electrostatic interactions are the most problematic, as exemplified by the large oscillations in the EF3 Watson-Crick base pair potentials (Fig. 3) at distances where the electrostatic energy terms start getting switched off. The root of the problem is the impossibility to include all nonbonded pairs in big systems, making some sort of truncation necessary. These oscillations are at their worst in the case illustrated in Fig. 3, with the bases moving apart in the plane of the basepair, but there will be artefactual energy minima at long (6–10 Å) distances in other directions as well. This may make the structure artificially rigid, leading to difficulties in interpreting results of minimizations and MD simulations.

The molecular dynamics simulations on the DNA hexamer duplex $d(CGTACG)_2$ indicate that there is a delicate balance between the external (van der Waals, electrostatic. hydrogen bonds) and internal (mainly dihedrals) energy terms. With EF2 the helix

behaves very differently when an effective dielectric constant $D = 2$ is used, or when the van der Waals radii of extended carbons are increased by 0.3–0.5 Å. Either of these changes is sufficient to severely distort the helical structure of the oligomer. It is also clear that the treatment of electrostatic truncation effects can have a major influence on the overall behavior of the system. In cases where the individual atomic charges are fairly large, as in EF3, it seems unwise to use a simple truncation based on interatomic distances. Even if the truncation step is smoothed by a conventional cubic switching function or the whole electrostatic potential energy is shifted by a constant to go to zero at the cutoff distance, the force on the atoms may still fluctuate wildly at this cutoff distance. The choice of electrostatic potential form has less influence with EF3 than with EF2, and with EF3 a choice of $D = 2$, when used with a shifting function[16] applied over the whole range $0 < r_{ij} < r_{cut}$, only causes distortions of the ends of the helix, leaving the central four basepairs intact.

The analysis reported here suggests that the revised potential energy function EF2, an improvement over the earlier EF1, should be satisfactory for a range of energy minization and dynamical studies of nucleic acids. Further tests will have to be made in the presence of solvent to ascertain that the balance of interactions (DNA with water versus water with water) is correctly represented.

We thank Drs. B.R. Brooks, C.L. Brooks III, and B.M. Pettitt for many stimulating discussions. L.N. was the recipient of a Postdoctoral Fellowship from the Swedish Natural Science Research Council.

APPENDIX 1

The parameters in energy function EF2 are given below, followed by a figure showing charges and atom types for the common nucleic acid components: the five bases A, C, G, T, and U, the phosphate group and the sugars.

Bond		K_b kcal/mol/Å	b_0 Å
C	CB	386.0	1.419
C	CS	386.0	1.444
C2	CH	201.0	1.525
C2	C2	315.0	1.525
C3	C2	201.0	1.525
C3	CH	201.0	1.525
CA	CB	386.0	1.404
CB	CB	550.0	1.370
CF	CF	684.0	1.350
CF	C	386.0	1.444
CF	CA	386.0	1.433
CH	CH	201.0	1.525
CS	CF	386.0	1.343
CS	C3	201.0	1.525
H	NH3	472.0	1.040
H	NA	472.0	1.010
H	NS	472.0	1.010
H2	N2	472.0	1.010
HO	OH	450.0	0.960
HO	OS	450.0	0.960
N2	CA	400.0	1.333
NA	C	337.0	1.388
NA	CA	400.0	1.381
NB	CB	400.0	1.391
NB	CE	400.0	1.304
NC	CA	400.0	1.339
NC	CE	400.0	1.324
NC	CB	400.0	1.354
NC	C	400.0	1.358
NH2E	CH	422.0	1.490
NS	CE	300.0	1.371
NS	CF	337.0	1.365
NS	C	337.0	1.383
NS	CH	300.0	1.475
NS	CB	337.0	1.374
O	C	612.0	1.229
OH	C2	292.0	1.430
OH	CH	292.0	1.430
OS	CH	292.0	1.430
OS	C2	292.0	1.430
P	O2	528.0	1.480
P	OS	237.0	1.610
P	OH	237.0	1.610

Bond Angle			K_θ kcal/mol/rad^2	θ_0 degrees
CB	C	NA	62.0	111.3
CB	C	O	86.0	128.8
CF	C	NA	62.0	114.1
CF	C	O	86.0	125.3
CS	C	NA	62.0	114.1
CS	C	O	86.0	125.3
NA	C	O	70.0	120.6
NC	C	O	70.0	122.5
NS	C	NA	62.0	115.4
NS	C	O	70.0	120.9
NS	C	NC	70.0	118.6
C2	C2	C3	45.0	112.0
C2	C2	C2	82.0	110.0
C2	C2	NH3	65.0	109.5
CH	C2	CH	70.0	112.0
OH	C2	CH	70.0	112.0
OS	C2	C3	150.5	111.0
OS	C2	CH	70.0	112.0
OS	C2	C2	82.0	112.0
CB	CA	N2	70.0	123.5
CB	CA	NC	70.0	117.3
CF	CA	NC	70.0	121.5
N2	CA	CF	70.0	120.1
N2	CA	NA	70.0	116.0
N2	CA	NB	70.0	118.3
N2	CA	NC	70.0	119.8
NC	CA	NA	70.0	123.3
C	CB	CB	116.0	119.2
C	CB	NB	116.0	130.0
CA	CB	CB	116.0	117.3

Bond Angle			K_θ kcal/mol/rad^2	θ_0 degrees
CA	CB	NB	116.0	132.4
CB	CB	NA	70.0	106.2
CB	CB	NB	116.0	110.4
CB	CB	NS	70.0	106.2
NC	CB	NS	70.0	126.0
NC	CB	CB	116.0	127.7
NC	CB	NA	70.0	126.0
NB	CE	NA	70.0	113.9
NB	CE	NS	70.0	113.9
NC	CE	NC	70.0	129.1
C	CF	CF	116.0	120.7
CA	CF	CF	116.0	117.0
CS	CF	NS	116.0	121.2
NA	CF	CF	116.0	121.2
NS	CF	CF	116.0	121.2
C2	CH	NS	46.5	111.0
C2	CH	NH2E	46.5	111.0
C2	CH	OS	46.5	111.0
C2	CH	CH	46.5	111.0
C2	CH	OH	46.5	111.0
C3	CH	CH	46.5	111.0
C3	CH	OS	46.5	111.0
CH	CH	CH	46.5	111.0
CH	CH	OH	46.5	111.0
CH	CH	NH2E	46.5	111.0
CH	CH	OS	46.5	111.0
CH	CH	NS	46.5	111.0
OS	CH	NS	46.5	111.0
OS	CH	NH2E	46.5	111.0
C	CS	CF	70.0	120.7
C	CS	C3	70.0	119.7
C3	CS	CF	70.0	119.7
CA	N2	H2	35.0	120.0
H2	N2	H2	35.0	120.0
C	NA	C	116.0	126.4
C	NA	H	35.0	116.5
C3	NA	CF	35.0	119.3
C3	NA	C	70.0	119.3
CA	NA	C	116.0	125.2
CA	NA	H	35.0	119.3
CE	NA	CB	116.0	105.4
H	NA	CB	35.0	119.3
H	NA	CE	35.0	119.3
CB	NB	CE	70.0	103.8
C	NC	CA	70.0	120.5
CA	NC	CE	70.0	118.6
CA	NC	CB	70.0	112.2
CB	NC	CE	70.0	111.0
C2	NH3	C2	35.0	109.5
H	NH3	C2	35.0	109.5
H	NH3	H	35.0	109.5
C	NS	CF	116.0	121.6
C	NS	H	35.0	119.3
C	NS	CH	70.0	117.6
CB	NS	H	35.0	119.3
CB	NS	CE	116.0	105.4
CE	NS	H	35.0	119.3
CF	NS	H	35.0	119.3
CF	NS	CH	70.0	121.2
CH	NS	CE	70.0	128.8
CH	NS	CB	70.0	125.8
HO	OH	P	46.5	107.3
HO	OH	C2	46.5	107.3
HO	OH	CH	46.5	107.3
C2	OS	C2	82.0	111.5
CH	OS	CH	46.5	111.5
HO	OS	CH	46.5	107.3
HO	OS	C2	46.5	107.3
P	OS	C2	47.6	120.5
P	OS	CH	46.5	120.5
P	OS	C3	47.6	120.5
O2	P	O2	140.5	119.9
OH	P	OS	48.1	102.6
OH	P	O2	98.9	108.2
OS	P	O2	98.9	108.2
OS	P	OS	48.1	102.6

Torsion Angle				K_ϕ kcal/mol	n	δ degrees
X[a]	C2	CH	X	1.4	3	0.0
X	C2	C2	X	1.6	3	0.0
X	CH	CH	X	1.4	3	0.0
X	CH	OS	X	0.9	3	0.0
X	CH	OH	X	0.5	3	0.0
X	C2	OH	X	0.5	3	0.0
X	C2	OS	X	0.5	3	0.0
X	OS	P	X	0.75	3	0.0
X	OH	P	X	0.75	3	0.0
X	CH	NS	X	0.3	2	0.0
X	CA	N2	X	6.0	2	180.0
CA	CB	CB	NA	2.5	2	180.0
OS	CH	CH	OS[b]	0.5	2	0.0
OS	CH	CH	OS	1.4	3	0.0
OH	CH	CH	CH	0.5	2	0.0
OS	CH	CH	CH	0.5	2	0.0
OS	CH	CH	CH	1.4	3	0.0

OH	CH	CH	OS	0.5	2	0.0
OH	CH	CH	OS	1.4	3	0.0
OS	CH	C2	CH	1.0	2	0.0
OS	CH	C2	CH	1.4	3	0.0
OS	CH	CH	C2	1.4	3	0.0
OS	CH	CH	C2	0.5	2	0.0
OS	C2	C2	C2	1.4	3	0.0
OS	C2	C2	C2	0.5	2	0.0
OS	P	OS	C2	0.75	3	0.0
OS	P	OS	C2	0.75	2	0.0
OS	P	OS	C3	0.75	3	0.0
OS	P	OS	C3	0.75	2	0.0
OS	P	OS	CH	0.75	3	0.0
OS	P	OS	CH	0.75	2	0.0
CE	CB	CB	NC	5.0	2	180.0
CE	CB	CB	CE	5.0	2	180.0
CE	CB	CB	CA	5.0	2	180.0
CE	CB	CB	NA	5.0	2	180.0
NC	CB	CB	NB	2.5	2	180.0
CA	CB	CB	NS	2.5	2	180.0
C	CB	CB	NS	2.5	2	180.0

[a]The atomtype '*X*' denotes any atom.
[b]When two entries with identical atom descriptors are given, two energy terms are included for that dihedral; one is threefold and the other is twofold.

Improper Torsion Angle				K_ω kcal/mol/rad^2	ω_0 degrees
C2	*X*	*X*	CH	31.5	35.26
C3	*X*	*X*	CH	31.5	35.26
OH	*X*	*X*	CH	31.5	35.26
OS	*X*	*X*	CH	31.5	35.26
CH	OS	C2	NS	31.5	35.26
CH	OS	CH	NS	31.5	35.26
CH	OS	C2	NH2E	31.5	35.26
CH	OS	CH	NH2E	31.5	35.26
H	*X*	*X*	NA	43.0	0.0
H	*X*	*X*	NS	43.0	0.0
H2	*X*	*X*	N2	43.0	0.0
C	*X*	*X*	CB	4.5	0.0
C	*X*	*X*	CS	4.5	0.0
C	*X*	*X*	CF	4.5	0.0
C	*X*	*X*	O	113.3	0.0
C	*X*	*X*	NC	4.5	0.0
C	*X*	*X*	NS	4.5	0.0
CA	*X*	*X*	CB	4.5	0.0
CA	*X*	*X*	NC	4.5	0.0
CA	*X*	*X*	N2	43.0	0.0
CA	*X*	*X*	NS	4.5	0.0
CB	*X*	*X*	CE	4.5	0.0
CB	*X*	*X*	NA	4.5	0.0
CB	*X*	*X*	NB	4.5	0.0
CB	*X*	*X*	NC	4.5	0.0
CB	*X*	*X*	NS	4.5	0.0
CF	*X*	*X*	NC	4.5	0.0
CF	*X*	*X*	NA	4.5	0.0
CH	*X*	*X*	NS	4.5	0.0
C3	*X*	*X*	CS	43.0	0.0
NB	*X*	*X*	NS	4.5	0.0

Nonbonded Parameters[a]

	α_i	N_i	R_i [Å]
C	1.65	5.0	1.8
CA	1.65	5.0	1.8
CB	1.65	5.0	1.8
CS	1.65	5.0	1.8
CH	1.35	6.0	1.85
C2	1.77	7.0	1.90
C3	2.17	8.0	1.95
CF	2.07	6.0	1.9
CE	2.07	6.0	1.9
H	0.044	1.0	0.8
H2	0.044	1.0	0.8
HO	0.044	1.0	0.8
N	1.15	6.0	1.65
NA	1.15	6.0	1.65
NB	1.15	6.0	1.65
NC	1.15	6.0	1.65
NS	1.15	6.0	1.65
N2	1.15	6.0	1.65
NH2E	1.7	8.0	1.7
O	0.84	7.0	1.6
O2	2.14	7.0	1.6
OS	0.64	7.0	1.6
OH	0.59	7.0	1.6
P	3.0	14.0	1.9

[a]The parameters in the table are used with the Slater-Kirkwood formula (refs. 17, 53, 54) to generate the coefficients $\epsilon_{ij} = \dfrac{181.173\alpha_i\alpha_j}{(R_i + R_j)^6 \cdot [(\alpha_i/N_i)^{1/2} + (\alpha_j/N_j)^{1/2}]}$ and $\sigma_{ij} = (R_i + R_j)/2^{1/6}$ for nonbonded interactions in eq. (1); α_i is the polarizability, N_i is the effective number of electrons, and R_i is the van der Waals radius.

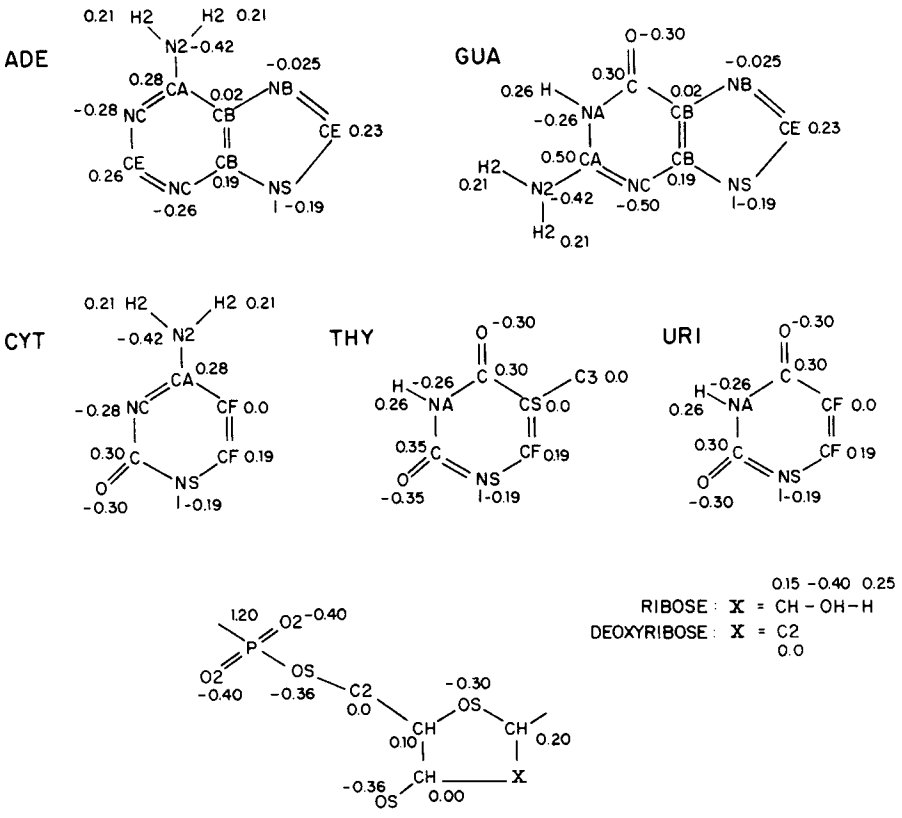
Hydrogen Bond		E^a [kcal mol^{-1}]	R^b_{min} [Å]
N[c]	N[d]	−3.0	3.0
N	O	−3.5	2.9
O	N	−4.0	2.85
O	O	−4.25	2.75

[a]Well depth.
[b]Minimum energy distance (donor − acceptor).
[c]Donor.
[d]Acceptor.

APPENDIX 2

The parameters in an extension of energy function EF2 including all hydrogen atoms are given below.

Bond		K_b kcal/mol/Å	b_0 Å
C	CB	302.0	1.419
C	CS	302.0	1.444
C2	C2	201.0	1.525
C2	CH	201.0	1.525
C3	C2	201.0	1.525
C3	CH	201.0	1.525
CA	CB	302.0	1.404
CB	CB	550.0	1.370
CF	C	302.0	1.444
CF	CA	302.0	1.433
CF	CF	517.0	1.350
CH	CH	201.0	1.525



CS	C3	201.0	1.525
CS	CF	302.0	1.343
H	C2	350.0	1.090
H	C3	350.0	1.090
H	CE	350.0	1.090
H	CF	350.0	1.090
H	CH	350.0	1.090
H	NA	472.0	1.010
H	NH3	472.0	1.040
H	NS	472.0	1.010
H2	N2	472.0	1.010
HO	OH	450.0	0.960
HO	OS	450.0	0.960
N2	CA	400.0	1.333
NA	C	302.0	1.388
NA	CA	400.0	1.381
NB	CB	400.0	1.391
NB	CE	400.0	1.304
NC	C	400.0	1.358
NC	CA	400.0	1.339
NC	CB	400.0	1.354
NC	CE	400.0	1.324
NH2E	CH	422.0	1.490
NS	C	302.0	1.383
NS	CB	302.0	1.374
NS	CE	300.0	1.371
NS	CF	302.0	1.365
NS	CH	300.0	1.475
O	C	612.0	1.229
OH	C2	292.0	1.430
OH	CH	292.0	1.430
OS	C2	292.0	1.430
OS	CH	292.0	1.430
P	O2	528.0	1.480
P	OH	237.0	1.610
P	OS	237.0	1.610

Bond Angle			K_θ kcal/mol/rad^2	θ_0 degrees
CB	C	NA	67.0	111.3
CB	C	O	80.0	128.8
CF	C	NA	67.0	114.1
CF	C	O	80.0	125.3
CS	C	NA	67.0	114.1
CS	C	O	80.0	125.3
NA	C	O	70.0	120.6
NC	C	O	70.0	122.5
NS	C	O	70.0	120.9
NS	C	NC	70.0	118.6
NS	C	NA	67.0	115.4
C2	C2	C3	45.0	112.0
C2	C2	NH3	65.0	109.5
C2	C2	C2	45.0	110.0
CH	C2	CH	70.0	112.0
H	C2	CH	30.0	109.5
H	C2	OS	30.0	109.5
H	C2	OH	30.0	109.5
H	C2	C3	30.0	109.5
H	C2	C2	30.0	109.5
H	C2	H	30.0	109.5
OH	C2	CH	70.0	112.0
OS	C2	CH	70.0	112.0
OS	C2	C3	150.5	111.0
H	C3	OS	30.0	109.5
H	C3	CH	30.0	109.5
H	C3	CS	30.0	109.5
H	C3	H	30.0	109.5
CB	CA	NC	70.0	117.3
CB	CA	N2	70.0	123.5
CF	CA	NC	70.0	121.5
N2	CA	NC	70.0	119.8
N2	CA	NB	70.0	118.3
N2	CA	NA	70.0	116.0
N2	CA	CF	70.0	120.1
NC	CA	NA	70.0	123.3
C	CB	CB	125.0	119.2
C	CB	NB	125.0	130.0
CA	CB	CB	125.0	117.3
CA	CB	NB	125.0	132.4
CB	CB	NA	70.0	106.2
CB	CB	NS	70.0	106.2
CB	CB	NB	125.0	110.4
NC	CB	CB	125.0	127.7
NC	CB	NS	70.0	126.0
H	CE	NC	30.0	120.0
H	CE	NS	30.0	108.0
H	CE	NB	30.0	108.0
NB	CE	NS	70.0	113.9
NC	CE	NC	70.0	129.1
C	CF	CF	125.0	120.7
CA	CF	CF	125.0	117.0
CS	CF	NS	125.0	121.2
H	CF	CA	30.0	120.0
H	CF	CF	30.0	120.0
H	CF	NS	30.0	120.0
H	CF	C	30.0	120.0
H	CF	CS	30.0	120.0
NA	CF	CF	125.0	121.2
NS	CF	CF	125.0	121.2
C2	CH	OS	46.5	111.0
C2	CH	CH	46.5	111.0
C2	CH	NS	46.5	111.0
C2	CH	OH	46.5	111.0
C2	CH	NH2E	46.5	111.0
C3	CH	CH	46.5	111.0
C3	CH	OS	46.5	111.0
CH	CH	CH	46.5	111.0
CH	CH	OS	46.5	111.0
CH	CH	NH2E	46.5	111.0
CH	CH	NS	46.5	111.0
CH	CH	OH	46.5	111.0
H	CH	OH	30.0	109.5
H	CH	OS	30.0	109.5
H	CH	CH	30.0	109.5
H	CH	NH2E	30.0	109.5
H	CH	NS	30.0	109.5
H	CH	C3	30.0	109.5
H	CH	H	30.0	109.5
H	CH	C2	30.0	109.5
OS	CH	NS	46.5	111.0
OS	CH	NH2E	46.5	111.0
C	CS	C3	70.0	119.7
C	CS	CF	70.0	120.7
C3	CS	CF	70.0	119.7
CA	N2	H2	40.0	120.0
H2	N2	H2	40.0	120.0
C	NA	H	40.0	116.5
C	NA	C	125.0	126.4
CA	NA	H	40.0	119.3
CA	NA	C	125.0	125.2
CB	NB	CE	70.0	103.8
C	NC	CA	70.0	120.5
CA	NC	CB	70.0	112.2
CA	NC	CE	70.0	118.6
CB	NC	CE	70.0	111.0
C2	NH3	C2	40.0	109.5
H	NH3	C2	40.0	109.5
H	NH3	H	40.0	109.5

Bond Angle			K_θ kcal/mol/rad^2	θ_0 degrees
C	NS	CH	70.0	117.6
C	NS	H	40.0	119.3
C	NS	CF	125.0	121.6
CB	NS	H	40.0	119.3
CB	NS	CE	125.0	105.4
CE	NS	H	40.0	119.3
CF	NS	H	40.0	119.3
CF	NS	CH	70.0	121.2
CH	NS	CE	70.0	128.8
CH	NS	CB	70.0	125.8
HO	OH	P	46.5	107.3
HO	OH	C2	46.5	107.3
HO	OH	CH	46.5	107.3
CH	OS	CH	46.5	111.5
HO	OS	C2	46.5	107.3
HO	OS	CH	46.5	107.3
P	OS	CH	46.5	120.5
P	OS	C2	47.6	120.5
P	OS	C3	47.6	120.5
O2	P	O2	140.5	119.9
OH	P	O2	98.9	108.2
OH	P	OS	48.1	102.6
OS	P	O2	98.9	108.2
OS	P	OS	48.1	102.6

Torsion Angle				K_ϕ kcal/mol	n	δ degrees
X[a]	C2	CH	X	1.5	3	0.0
X	C2	C2	X	1.6	3	0.0
X	CH	CH	X	1.5	3	0.0
X	CH	OS	X	0.5	3	0.0
X	CH	OH	X	0.5	3	0.0
X	C2	OH	X	0.5	3	0.0
X	C2	OS	X	0.5	3	0.0
X	OS	P	X	0.75	3	0.0
X	OH	P	X	0.75	3	0.0
X	CH	NS	X	0.3	2	0.0
X	CA	N2	X	6.0	2	180.0
X	C2	NH3	X	0.6	3	0.0
X	NS	CE	X	7.0	2	180.0
X	NS	CF	X	7.0	2	180.0
X	NC	CE	X	7.0	2	180.0
X	CF	CF	X	8.0	2	180.0
X	CS	C3	X	0.5	3	0.0
X	OS	C3	X	0.5	3	0.0
OH	CH	CH	CH	0.5	2	0.0
C3	CH	CH	OS	0.2	2	0.0
OS	CH	CH	OS[b]	0.5	2	0.0
OS	CH	CH	OS	1.4	3	0.0
OS	CH	CH	CH	0.5	2	0.0
OS	CH	CH	CH	1.4	3	0.0
OH	CH	CH	OS	0.5	2	0.0
OH	CH	CH	OS	1.4	3	0.0
OS	CH	C2	OH	0.5	2	0.0
OS	CH	C2	CH	0.5	3	0.0
OS	CH	CH	C2	0.2	2	0.0
OS	CH	CH	C2	1.0	3	0.0
OS	P	OS	C2	0.75	3	0.0
OS	P	OS	C2	0.75	2	0.0
OS	P	OS	C3	0.75	3	0.0
OS	P	OS	C3	0.75	2	0.0
OS	P	OS	CH	0.75	3	0.0
OS	P	OS	CH	0.75	2	0.0
CE	CB	CB	NC	5.0	2	180.0
CE	CB	CB	CE	5.0	2	180.0
CE	CB	CB	CA	5.0	2	180.0
CE	CB	CB	NA	5.0	2	180.0
NC	CB	CB	NB	2.5	2	180.0
CA	CB	CB	NS	2.5	2	180.0
C	CB	CB	NS	2.5	2	180.0

[a]The atomtype 'X' denotes any atom.
[b]When two entries with identical atom descriptors are given, two energy terms are included for that dihedral; one is theefold and the other is twofold.

Improper Torsion Angle				K_ω kcal/mol/rad^2	ω_0 degrees
C2	X	X	CH	31.5	35.26
C3	X	X	CH	31.5	35.26
OH	X	X	CH	31.5	35.26
OS	X	X	CH	31.5	35.26
CH	OS	C2	NS	31.5	35.26
CH	OS	CH	NS	31.5	35.26
CH	OS	C2	NH2E	31.5	35.26
CH	OS	CH	NH2E	31.5	35.26
H	X	X	NA	40.0	0.0
H	X	X	NS	40.0	0.0
H2	X	X	N2	40.0	0.0
H	X	X	CE	40.0	0.0
H	X	X	CF	40.0	0.0
C	X	X	CB	4.5	0.0
C	X	X	CS	4.5	0.0
C	X	X	CF	4.5	0.0
C	X	X	O	90.3	0.0
C	X	X	NC	4.5	0.0
C	X	X	NS	4.5	0.0
CA	X	X	CB	4.5	0.0
CA	X	X	NC	4.5	0.0
CA	X	X	N2	40.0	0.0
CA	X	X	NS	4.5	0.0
CB	X	X	CE	4.5	0.0
CB	X	X	NA	4.5	0.0
CB	X	X	NB	4.5	0.0
CB	X	X	NC	4.5	0.0
CB	X	X	NS	4.5	0.0
CF	X	X	NC	4.5	0.0
CF	X	X	NA	4.5	0.0
CH	X	X	NS	40.0	0.0
C3	X	X	CS	40.0	0.0
NB	X	X	NS	4.5	0.0

Nonbonded Parameters[a]

	ϵ_i [kcal/mol]	R_i [Å]
C	0.09	1.8504
CA	0.09	1.8504
CB	0.09	1.8504
CH	0.09	1.8504
CE	0.09	1.8504
CF	0.09	1.8504
CS	0.09	1.8504
C2	0.09	1.8504
C3	0.09	1.8504
H	0.004489	1.4682
H2	0.004489	0.9002
HO	0.004489	0.9002
N2	0.16	1.6046
NA	0.16	1.6046
NB	0.16	1.6046
NC	0.16	1.6046
NH2E	0.16	1.7
NH3	0.16	1.6046
NS	0.16	1.6046
O	0.2304	1.5316
O2	0.2304	1.5316
OH	0.2304	1.4316
OS	0.2304	1.5316
P	0.59	1.9

[a]From Rossky, Rahman, and Karplus.[55]

Hydrogen Bond		E^a [kcal mol^{-1}]	R^b_{min} [Å]
N[c]	N[d]	−3.5	3.05
N	O	−4.0	2.95
O	N	−4.50	2.90
O	O	−4.75	2.80

[a]Well depth.
[b]Minimum energy distance (donor-acceptor).
[c]Donor.
[d]Acceptor.

References

1. A. Warshel, M. Levitt, and S. Lifson, *J. Mol. Spect.*, **33,** 84 (1970).
2. M. Levitt, *J. Mol. Biol.*, **82,** 93 (1974).
3. F. Momany, R. McGuire, A. Burgess, and H. Scheraga, *J. Phys. Chem.*, **79,** 2361 (1975).
4. B. Gelin and M. Karplus, *Biochemistry,* **18,** 1256 (1979).
5. S.J. Weiner, P.A. Kollman, D. Case, U.C. Singh, C. Ghio, G. Alagona, and P. Weiner, *J. Am. Chem. Soc.*, **106,** 765 (1984).
6. M. Karplus and J. A. McCammon, *Annu. Rev. Bioch.*, **53,** 263 (1983).
7. M. Levitt, *Proc. Natl. Acad. Sci. USA,* **75,** 640 (1978).

8. P. Kollman, P. Weiner, and A. Dearing, *Biopolymers,* **20,** 2583 (1981).
9. W.K. Olson, in *Topics in Nucleic Acid Structure, Part 2,* S. Neidle, Ed., Macmillan Press, London, 1982, p. 1.
10. M. Levitt, *Cold Spring Harbor Symp. Quant. Biol.,* **47,** 251 (1983).
11. B. Tidor, K.K. Irikura, B.R. Brooks, and M. Karplus, *J. Biomol. Struct. Dyn.,* **1,** 231 (1983).
12. S. Harvey, M. Prabhakaran, B. Mao, and J.A. McCammon, *Science,* **223,** 1189 (1984).
13. S. Lifson, and A. Warshel, *J. Chem. Phys.,* **49,** 5116 (1968).
14. A. Warshel and S. Lifson, *J. Chem. Phys.,* **53,** 582 (1970).
15. B. Gelin, Ph.D. Thesis, Harvard University, Cambridge, 1976.
16. B.R. Brooks, R.E. Bruccoleri, B.D. Olafson, D.J. States, S. Swaminathan, and M. Karplus, *J. Comp. Chem.,* **4,** 187 (1983).
17. P.K. Weiner and P.A. Kollman, *J. Comp. Chem.,* **2,** 287 (1981).
18. W.E. Reiher and M. Karplus, in preparation.
19. J.P. Ryckaert, G. Cicotti, and H.J.C. Berendsen, *J. Comp. Phys.,* **23,** 327 (1977).
20. D.W. Marquardt, *J. SIAM,* **11,** 431 (1963).
21. J. Overend and J.R. Scherer, *J. Chem. Phys.,* **32,** 1289 (1960).
22. Y. Morino and K. Kuchitsu, *J. Chem. Phys.,* **20,** 1809 (1952).
23. N.C. Seeman, J.M. Rosenberg, F.L. Suddath, J.J.P. Kim, and A. Rich, *J. Mol. Biol.,* **104,** 109 (1976).
24. J.M. Rosenberg, N.C. Seeman, R.O. Day, and A. Rich, *J. Mol. Biol.,* **104,** 145 (1976).
25. I.K. Yanson, A.N. Teplitsky, and L.F. Sukhdoub, *Biopolymers,* **18,** 1149 (1979).
26. J. Suurkuusk, J. Alvarez, E. Freire, and R. Biltonen, *Biopolymers,* **16,** 2641 (1977).
27. L.A. Marky and K.J. Breslauer, *Biopolymers,* **21,** 2185 (1982).
28. M. Petersheim and D.H. Turner, *Biochemistry,* **22,** 256 (1983).
29. R.L. Ornstein, R. Rein, D.L. Breen, and R.D. Macelroy, *Biopolymers,* **17,** 2341 (1978).
30. T.E. Haran, Z. Berkovich-Yellin, and Z. Shakked, *J. Biomol. Struct. Dyn.,* **2,** 397 (1984).
31. R. Degeilth and R.E. Marsh, *Acta Cryst.,* **12,** 1007 (1959).
32. H. Mathisen, N. Norman, and B. Pedersen, *Acta Chem. Scand.,* **21,** 127 (1967).
33. A.T. Hagler, E. Huler, and S. Lifson, *J. Am. Chem. Soc.,* **96,** 5319 (1974).
34. B.R. Brooks and M. Karplus, in preparation.
35. D.G. Gorenstein, B.A. Luxon, and J.B. Findlay, *Biochim. Biop. Acta,* **475,** 184 (1977).
36. C. Altona and M. Sundaralingam, *J. Am. Chem. Soc.,* **94,** 8205 (1972).
37. D. Cremer and J.A. Pople, *J. Am. Chem. Soc.,* **97,** 1354 (1975).
38. W.K. Olson, *J. Am. Chem. Soc.,* **104,** 278 (1982).
39. W.K. Olson and J.L. Sussman, *J. Am. Chem. Soc.,* **104,** 270 (1982).
40. Y. Kyogoku and I. Iitoka, *Acta Crystallogr.,* **21,** 49 (1966).
41. E. Brown and W.L. Peticolas, *Biopolymers,* **14,** 1259 (1975).
42. M. Szczesniak, M.J. Nowak, H. Rostkowska, K. Szczepaniak, W.P. Person, and D. Shugar, *J. Am. Chem. Soc.,* **105,** 5969 (1983).
43. B.M. Pettitt and M. Karplus, *J. Am. Chem. Soc.,* in press.
44. M. Tsuboi and S. Takahashi, in *Physico-Chemical Properties of Nucleic Acids,* J. Duchesne (Ed.), Academic, London, 1973, Vol. 2, p. 91.
45. J.M. Eyster and E.M Prohofsky, *Spectrochimica Acta,* **30A,** 2041 (1974).
46. A. Palm and E.R. Bissel, *Spectrochim. Acta,* **16,** 459 (1960).
47. S. Arnott, P.J. Campbell Smith, and R. Chandrasekaran, in *Handbook of Biochemistry and Molecular Biology,* 3rd ed. G.D. Fasman, Ed., CRC Press, Cleveland, 1975, p. 419.
48. Molecular graphics program HYDRA (Harvard York Drawing Program), developed by R. Hubbard at Harvard University.
49. W.L. Jorgensen and M. Ibrahim, *J. Am. Chem. Soc.,* **103,** 3976 (1981).
50. D.A. Pearlman and S.H. Kim, *Biopolymers,* **24,** 327 (1985).
51. L. Nilsson, M. Clore, A. Gronenborn, A.T. Brunger, and M. Karplus, *J. Mol. Biol.* **188,** 455 (1986).
52. Y. Nishimura, M. Tsuboi, S. Kato, and K. Morokuma, *J. Am. Chem. Soc.,* **103,** 1356 (1981).
53. J.C. Slater and J.G. Kirkwood, *Phys. Rev.,* **37,** 682 (1931).
54. K.S. Pitzer, in *Advances in Chemical Physics,* I., Prigogine, Ed., Interscience, New York, 1959, Vol. II, p. 59.
55. P.J. Rossky, M. Karplus, and A. Rahman, *Biopolymers,* **18,** 825 (1979).