An All Atom Force Field for Simulations of Proteins and Nucleic Acids

Scott J. Weiner, Peter A. Kollman,

Department of Pharmaceutical Chemistry, School of Pharmacy, University of California, San Francisco, California 94143

Dzung T. Nguyen, David A. Case,

Department of Chemistry, University of California, Davis, California 95616

Received 7 March 1985; accepted 5 August 1985

We present an all atom potential energy function for the simulation of proteins and nucleic acids. This work is an extension of the CH united atom function recently presented by S. J. Weiner et al. (*J. Amer. Chem. Soc.*, **106**, 765 (1984). The parameters of our function are based on calculations on ethane, propane, *n*-butane, dimethyl ether, methyl ethyl ether, tetrahydrofuran, imidazole, indole, deoxyadenosine, base paired dinucleoside phosphates, adenine, guanine, uracil, cytosine, thymine, insulin, and myoglobin. We have also used these parameters to carry out the first general vibrational analysis of all five nucleic acid bases with a molecular mechanics potential approach.

INTRODUCTION

The development of a molecular mechanics force field appropriate for both proteins and nucleic acids is timely, given the recent advances in our understanding of protein-nucleic acid interactions. We have recently presented such a force field,[1] which was reasonably successful in reproducing structures, energies, and vibrational frequencies of model systems. For reasons of computational efficiency that force field used a united atom (spherical) representation of CH, CH_2, and CH_3 groups. Because of this approximation, compromises have to be made which lead in some cases to less than optimum fits with experiment.[1] Recent calculations on nucleic acid interactions[2] have also suggested that when one is examining small energy differences, a spherical representation of CH groups leads to poorer agreement with experiment than an all atom representation. Simulations of nmr relaxation or methyl group rotations should benefit from an explicit treatment of all hydrogen atoms, and such a representation also makes comparisons to observed vibrational spectra much more straightforward. Here, we extend our previous force field to allow a general all atom representation of proteins and nucleic acids. We also envision a "hybrid force field" in which one uses an all atom representation at the active site of an enzyme and a united atom formalism elsewhere, thus retaining the greater accuracy of the all atom force field in the important parts of the structure.

Some work towards an all-atom representation has been described earlier. In our previous article,[1] we presented results of all-atom calculations on dipeptide models and on a furanose sugar. In addition, in our quantum mechanical calculations[1,3] we had already determined a charge model appropriate for both all atom and united atom representations.[1,3] Thus, our task here mainly involves recalibration of torsional and angle parameters using experimental conformational energies. We also present a normal mode analysis of all five of the nucleic acid bases.

DEVELOPMENT OF PARAMETERS

Force Field Equation

The force field equation is the same as used previously:[1]

$$E_{\text{total}} = \sum_{\text{bonds}} K_R(R - R_o)^2 + \sum_{\text{angles}} K_\theta(\theta - \theta_o)^2 + \sum_{\text{dihedrals}} \frac{V_n}{2}[1 + \cos(n\phi - \gamma)] + \sum_{i<j}\left[\frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^{6}} + \frac{q_i q_j}{\varepsilon R_{ij}}\right] + \sum_{\text{H-bonds}}\left[\frac{C_{ij}}{R_{ij}^{12}} - \frac{D_{ij}}{R_{ij}^{10}}\right] \quad (1)$$

Journal of Computational Chemistry, Vol. 7, No. 2, 230–252 (1986)

CCC 0192-8651/86/020230-23$04.00

and all the terms have the same meaning as before.[1] The five new atom types required for an all-atom potential appear in Table I. The general approach used to determine force field parameters appears in ref 1: small molecules salient to nucleic acids and polypeptides were taken as test cases from which to best fit desired structural and energetic properties. As before[1] we try to describe our reasons for choosing particular test cases/parameters, so that the origins and biases in our parameters may be understood.

Aliphatic Small Molecule Test Cases

We began our analysis (Table II) of model systems with the simplest of all hydrocarbons containing at least one torsion: ethane.

Table I. List of atom types.

Atom	Type
CK	sp^2 aromatic carbon in five membered ring between two nitrogens and bonded to one hydrogen (in purines).
CQ	sp^2 carbon in six membered ring of purines between two "NC" nitrogens and bonded to one hydrogen.
CR	sp^2 aromatic carbon in five membered ring between two nitrogens and bonded to one hydrogen (in HIS).
CV	sp^2 aromatic carbon in five membered ring bonded to an N: and bonded to an explicit hydrogen (e.g. $C_\delta—N_\varepsilon=C_\varepsilon$ in HIS).
CW	sp^2 aromatic carbon in five membered ring bonded to an N—H and bonded to an explicit hydrogen (e.g. $C_\delta—N_\varepsilon—C_\varepsilon$ in HIS).

The bond stretching parameters CT—HC (R_o = 1.09 Å K_R = 310 kcal/mol Å^2) and bond bending parameters HC—CT—HC (θ_o = 109.5, K_θ=35 kcal/mol rad^2) and HC—CT—CT (θ_o = 109.5, K_θ=35 kcal/mol rad^2) were taken directly from MM2.[4] For atom type CT, which is the atom type used for all sp^3 C atoms in the all atom force field, we chose the same van der Waals radii and well depth (R^* = 1.80 Å, ε = 0.06 kcal/mol) as derived before.[1] Although we reported van der Waals parameters for HC in ref. 1, these were taken from Hagler, Euler, and Lifson[5] and never checked on any of our small molecule test cases. However, we found from base stacking calculations of N,N di-methyluracil (discussed below) that one must use a much smaller van der Waals well depth for HC to retain reasonable agreement with experiment. We, therefore, kept the repulsive A_{ij} parameter for HC the same, but reduced the well depth from 0.036 kcal/mol to 0.01 kcal/mol, thus leading to HC van der Waals parameters of R^* = 1.54 Å and ε = 0.01 kcal/mol. We have used these values for all of the simulations presented here.

With a $V_3/2$ torsional potential of 1.3 kcal/mol for CT—CT, we calculate an eclipsed-staggered energy difference of 2.8 kcal/mol for ethane, in good agreement with the experimental value of 2.9 kcal/mol.[6] At this point it is important to note that in our united atom force field we scaled all 1–4 nonbonded interactions (nonbonded atoms separated by

Table II. Results for hydrocarbons.

Parameter	1[a]	2[b]	Experiment[c]
	Ethane		
ΔE(eclipsed–staggered)	2.83	3.02	2.9[d]
	n-Butane		
ΔE(gauche–trans)	0.58	0.51	0.5–0.9[e]
ΔE(cis–trans)	4.57	4.76	4.5–6.0[e]
Structural Parameters			
Φ(gauche)	65.8	65.4	65.9[e]
ϑ(C—C—C)(cis)	115.7	115.9	117.0[e]
ϑ(C—C—C)(trans)	110.6	111.1	113.1[e]
ϑ(C—C—C)(gauche)	112.6	113.1	114.4[e]
	Propane		
ΔE(V1)[d]	2.99	3.24	3.3[d]
ΔE(V2)[d]	3.59	3.72	3.9[d]

[a]1-4 nonbonded interactions reduced by 50%; energy in kcal/mol.
[b]No scaling of the 1-4 nonbonded interactions; energy in kcal/mol.
[c]Experimental values.
[d]Ref. 6.
[e]Ref. 7.

three bonds) by 50%. We again address this scaling question (see ref. 1, p. 770) and present results on our small molecule test cases both with and without the 1–4 scale factor (compare the columns labeled '1' and '2' in Table II). As is clear from Table II, for ethane and for the other test cases discussed below, the results are rather insensitive to 1–4 scaling.

Next, we proceeded to *n*-butane. Using a general torsion (ref. 1, p. 769) $V_3/2$ = 1.3 kcal/mol for X—CT—CT—X, we calculated $\Delta E_{\text{gauche-trans}}$ = 0.6 kcal/mol; which is in reasonable agreement with the experimental value of 0.5-0.9 kcal/mol.[7] As a final check of our choice for torsional potentials about CT—CT bonds, we calculated the important rotational barriers in propane, which occur for the relative energy difference between the staggered vs. eclipsed conformations of the central CH_2 bond relative to a rotating CH_3 when (a) the other CH_3 is staggered with respect to the central CH_2 (V_1) and (b) the other CH_3 eclipses C2 (V_2). We found V_1 = 3.0 kcal/mol and V_2 = 3.6 kcal/mol; the experimental values are 3.3 kcal/mol and 3.9 kcal/mol, respectively.[6]

Next we focused on the aliphatic ethers dimethyl ether (DME), methyl ethyl ether (MEE), and tetrahydrofuran (THF). We again selected MM2 parameters for the equilibrium parameters for bonds and angles. To best reproduce the gauche-trans and cis-trans barriers in MEE and C_s—C_2 and C_{2v}—C_2 energy differences in THF, we used a $V_3/2$ = 1.15 kcal/mol for X—CT—OS—X and a $V_2/2$ = 0.2 kcal/mol for CT—CT—OS—CT. This gave a $\Delta E(\text{MEE})_{\text{gauche-trans}}$ = 1.36 kcal/mol and $\Delta E(\text{MEE})_{\text{cis-trans}}$ = 5.3 kcal/mol; the experimental values are 1.4 ± 0.2 kcal/mol and 5.9 kcal/mol, respectively.[8] We also calculated a $\Delta E(\text{DME})_{\text{eclipsed-staggered}}$ = 2.7 kcal/mol in agreement with experiment.[9] Finally, we applied these parameters to tetrahydrofuran. For THF (Table III) the calculated $\Delta E_{C_s-C_2}$ = 0.1 kcal/mol, $\Delta E_{C_{2v}-C_2}$ = 3.8 kcal/mol, $q(C_2)$ = 0.40 Å and $q(C_s)$ = 0.37 Å, in good agreement with the experimental values.[11] We do not report the vibrational frequencies for our models here, but they are similar to those previously reported.[1] In particular, the calculated frequencies for THF are similar to the values for FF1 (Table III in ref. 1), which are in reasonable agreement with experiment for all but the two ring bending modes near 500-600 cm^{-1}; which are calculated to be ~150 cm^{-1} too low in frequency.

In our earlier study, we examined the sugar pucker profile of deoxyadenosine and adenosine and we report the corresponding results with the all-atom representation in Table IV. The results reported there are analogous to those reported earlier[1] with the united atom model, although we present additional results here on the energetics and structures of O1′exo conformations of the nucleosides. In the case of deoxyadenosine, the two electrostatic models considered earlier ($\varepsilon = R_{ij}$ and $\varepsilon = 4R_{ij}$) give similar results, which are in reasonable agreement with available experiments: two local minima in the sugar pucker energy surface, C2′endo and C3′endo, with the former ~0.6 kcal/mol lower in energy; two maxima between them on the pseudorotation (W) surface, which are, respectively, 1.0-1.4 (O1′endo) and 3.7-4.2 (O1′exo) above the C2′endo energy. Although we did not report it in detail previously, the "O1′endo" maxima occur at a phase (W) = 64°–66° in both united atom and all atom models, considerably different from the 90° value labeled[12] as O1′endo. All of these minima correspond to the glycosidic angle (χ) in the anti region, H5′—O5′—C5′—C4′ in the trans region, O5′—C5′—C4′—C3′ in the "g^+" region, and C4′—C3′—O3′—H3′ in the trans region characteristic of A and B DNA. In our earlier study,[1] we used adiabatic mapping with constraints on C4′—C3′—C2′—C1′ (and for O1′exo also on O1′—C1′—C2′—C3′) to determine the energy profiles. Here, we did so as well, but we also used a transition state finding algorithm described elsewhere[18] to confirm that the O1′endo and O1′exo structures are true transition states. The detailed values reported in Table IV have been determined using the latter approach. We stress that all degrees of freedom have been energy optimized, but only the key features are reported there.

In the case of adenosine, it is clear that the orientation of the 2′OH and its strong electrostatic properties can strongly effect the sugar pucker profile. In the absence of the inclusion of water, which will compete for the intramolecular H bonding of the 2′OH and 3′OH we merely note that the only models with highly dampled intramolecular electrostatic

Table III. Results for aliphatic ethers.

Parameter	1[a]	2[b]	Experiment[c]
	Dimethyl Ether		
ΔE(eclipsed–staggered)	2.68	2.84	2.7[d]
Structural Parameters			
ϑ(C—O—C) (staggered)	112.0	112.9	111.8[d]
ϑ(C—O—C) (eclipsed)	112.9	114.0	—[d]
	Tetrahydrofuran		
$\Delta E(C_{2v}—C_2)$	0.13	0.28	≈0.1[g]
$\Delta E(C_s—C_2)$	3.83	3.89	3.5[g]
Structural Parameters			
C_2 conformation			
q^g	0.40	0.39	0.39[i]
ϑ(C—O—C)	108.9	109.1	110.5[i]
ϑ(C—C—O)	106.7	106.7	106.5[i]
ϑ(C—C—C)	100.5	100.7	101.8[i]
C_s conformation			
q^g	0.37	0.37	0.364[i], 0.38[j]
ϑ(C—O—C)	105.7	106.2	106.2[i]
ϑ(C—C—O)	105.0	105.0	105.0[j]
ϑ(C—C—C)	103.7	103.8	104.1[j]
	Methyl Ethyl Ether		
ΔE(gauche–trans)	1.36	1.24	1.4 ± 0.2[k]
ΔE(cis–trans)	5.27	5.21	(5.9)[l]
Structural Parameters			
gauche			
Φ	68.1	68.1	72, 85[m]
ϑ(C—O—C)	113.8	114.5	(113.2)[l]
ϑ(C—C—O)	112.6	112.8	(112.2)[l]
cis			
ϑ(C—O—C)	117.0	117.6	(116.3)[l]
ϑ(C—C—O)	117.1	117.4	(117.3)[l]

[a]1–4 nonbonded interactions reduced by 50%; energy in kcal/mol.
[b]No scaling of the 1–4 nonbonded interactions; energy in kcal/mol.
[c]Experimental values.
[d]Ref 8.
[e]Difference in energy between C_2 and C_s conformations of THF (kcal/mole).
[f]Difference in energy between C_2 and planar C_{2v} conformations of THF (kcal/mole).
[g]Ref. 11.
[h]Mean out of plane distance of ring, as defined in ref. 12.
[i]Ref. 13.
[j]Ref. 14.
[k]Ref. 9.
[l]MM2 calculations ref. 10.
[m]See ref. 15 for discussions on these parameters.

interactions, such as the $\varepsilon = 4R_{ij}$ model presented here lead to calculated properties in reasonable agreement with experiment. Here, interestingly, the "O1'*endo*" geometry is nearer the "classical" $W = 90°$ value than in the deoxyadenosine case.

Base Stacking, Hydrogen Bonding, and Thermal Stability of Nucleic Acids

As a check for our van der Waals parameters, we have carried out base stacking calculations between two 1,3,N,N,-dimethyluracil molecules as was done in the united atom force field paper (ref. 1, p. 773). As mentioned above, we initially used Hagler, Euler, and Lifson[5] values for the radii and well depth of the aliphatic hydrogens ($R^* = 1.375$ Å, $\varepsilon = 0.038$ kcal/mol). We calculated a base stacking stabilization energy of −12.0 kcal/mol (dielectric constant $\varepsilon = 1$); too stable compared with the ≈ −9.5 kcal/mol found in the united atom simulation and the "experimental" value of about −9 kcal/mol (see Table V). To achieve a lower base stacking energy we decided to decrease ε and simultaneously increase R^* by retaining the repulsive (R^{-12}) coefficient while changing

Table IV. Calculations on adenosine and deoxyadenosine.

Deoxyadenosine	$\varepsilon = R_{ij}$	$\varepsilon = 4R_{ij}$	Exp.[a]
ΔE(C3'endo-C2'endo)[b]	0.69	0.58	0.66
ΔE(O1'endo-C2'endo)[c]	1.37	1.03	—
ΔE(O1'exo-C2'endo)[d]	3.68	4.19	—
W(C2'endo)[e]	151	148	165
q(C2'endo)[f]	0.38	0.39	(0.35–0.41)
W(C3'endo)[e]	14	20	(2–20)
q(C3'endo)[f]	0.39	0.40	(0.35–0.41)
W(O1'endo)[e]	64	66	—
q(O1'endo)[f]	0.36	0.37	—
W(O1'exo)[f]	288	276	—
q(O1'exo)[f]	0.21	0.27	—
Adenosine	$\varepsilon = R_{ij}$	$\varepsilon = 4R_{ij}$	Exp.[a]
ΔE(C3'endo-C2'endo)[b]	2.55	0.29	(0.19–0.42)
ΔE(O1'endo-C2'endo)[c]	2.56	1.31	—
ΔE(O1'exo-C2'endo)[c]	5.06	4.27	—
W(C2'endo)[e]	190	154	(150–170)
q(C2'endo)[f]	0.30	0.38	(0.35–0.41)
W(C3'endo)[e]	21	18	(2–20)
q(C3'endo)[f]	0.37	0.39	(0.35–0.41)
W(O1'endo)[e]	30	98	—
q(O1'endo)[f]	0.36	0.36	—
W(O1'exo)[f]	334	290	—
q(O1'exo)[f]	0.40	0.19	—

[a]Experimental data from Davies (ref. 16) and, Altona and Sundaralingham (ref. 17).
[b]Energy difference between energy minimized C3'endo and C2'endo conformations (kcal/mol).
[c]Energy difference between O1'endo and C2'endo conformations (kcal/mol).
[d]Energy difference between O1'exo and C2'endo conformations (kcal/mol).
[e]Energy refined pseudorotation angle (ref. 12) for given conformation (in degrees).
[f]Energy refined mean out of plane sugar distance in Angstroms (see ref. 12).

Table V. Hydrogen bonding and stacking for base pairs.

Complex	$\Delta E(\varepsilon = 1)$[a]	$\Delta E(\varepsilon = R_{ij})$[b]	ΔE(Langlet)[c]	ΔH(expt)[d]
GC Watson-Crick[e]	−21.3	−22.1	−23.7	−21.0
AT Watson-Crick[f]	−11.5	−13.7	−12.9	−13.0
AT Hoogsteen[g]	−12.0	−13.5	−13.6	−13.0
1,3-Dimethyluracil Stack[h]	−10.5	−9.9	−9.1	(−9.1)

[a]Energy of complex formation with $\varepsilon = 1$ in kcal/mole.
[b]Energy of complex formation with $\varepsilon = R_{ij}$ in kcal/mole.
[c]Energy calculated by Langlet et al (ref. 19).
[d]Experimental value for associations inferred from the experiments by Yanson et al (ref. 20). In the case of the 1,3-dimethyluracil stacking, the value in parentheses is the value calculated by Langlet *et al* (ref. 19), since these authors showed that there was an important electric field dependence in the experiments by Yanson.
[e]Watson and Crick H-bonded structure of 9-methylguanine and 1-methylcytosine. Model built using computer graphics and then energy refined. The (H---X) distances are 1.87, 1.90 and 1.87 Å ($\varepsilon = 1$) and 1.83, 1.84 and 1.83 ($\varepsilon = R_{ij}$) for $G2NH_2$---CO2, GN1H---CN3 and GO6---$C4NH_2$ hydrogen bonds respectively.
[f]Watson and Crick H-bonded structure of 9-methyladenine and 1-methylthymine. Model built using computer graphics and then energy refined. The (H---X) distances are 1.86 and 1.90 Å ($\varepsilon = 1$) and 1.82 and 1.86 ($\varepsilon = R_{ij}$) for AN1---TN3H and $A6NH_2$---TO4 hydrogen bonds respectively.
[g]Hoogsteen H-bonded structure of 9-methyladenine and 1-methylthymine. Model built using computer graphics and then energy refined. The (H---X) distances are 1.92 and 1.80 Å ($\varepsilon = 1$) and 1.87 and 1.78 ($\varepsilon = R_{ij}$) for $A6NH_2$---TO2 and AN7---TN3H hydrogen bonds respectively.
[h]Stacked complex of 1,3-dimethyluracil model built using figure 11 A1 in the paper by Langlet *et al* (ref. 17) and energy refined, base-base minimum energy distance = 3.55 Å for $\varepsilon = R_{ij}$ and 3.55 Å for $\varepsilon = 1$.

while changing the attractive (R^{-6}) one. As is seen in Table V, ΔE base stacking = −10.5 kcal/mol ($\varepsilon = 1$) and −9.9 kcal/mol ($\varepsilon = R_{ij}$) for the "harder" hydrogen.

Table V also summarizes the results of nucleic acid base hydrogen bonding calculations. As expected (since none of the hydrogen bond parameters nor charges of the proton

donor and acceptor atoms were changed), the results are nearly identical with those in the united atom simulation (ref. 1, p. 773).

As we have emphasized before, the two major bottlenecks in extracting useful information from molecular mechanics calculations on complex systems are the large number of local minima and the difficulties in accurately representing solvent/counterion effects. That was a motivation for developing a force field which gave similar H-bonding energies and structures when one used a unity dielectric constant ($\varepsilon = 1$), to be used when solvent and counterions were explicitly included in the calculation, and a distance dependent dielectric constant ($\varepsilon = R_{ij}$), to be used when solvent/counterions were not explicitly included. Since we have only added C-H groups here and the atomic partial changes for the H bonding groups have been changed little if at all, these features have been retained here.

We have also studied the relative thermal stabilities of short oligonucleotide strands of DNA (ref. 1, p. 773). The all-atom results are similar to those presented previously for the united atom representation so we do not present them in detail here. The relative thermal stabilities for the first three sets of polymers (i.e., 1 vs. 2, 3 vs. 4, and 5 vs. 6 in Table IX in ref. 1) correlate qualitatively with the calculated ΔE's in both the united and all atom simulations. However, for the polytrinucleotides (7–10), the qualitative order is not as consistent with experiment, as was also the case for the united atom model.

Charge Derivation for Proteins and Nucleic Acids

Consistent with the approach employed in our united atom force field, we have derived atom centered charges by using quantum mechanically derived electrostatic potentials fit to a point charge model.[1,3] When we derived charges for our united atom force field model systems, we also used the same quantum mechanical electrostatic potential to fit an all atom representation. The only difference in the method of derivation arises for the bridge atoms (e.g., those carbon atoms which are between the peptide backbone and the chromophore; C_β, C_γ, etc.) The excess charge [the amount needed to achieve charge neutrality (or unity) after the backbone and chromophore are calculated] was distributed with the same ratio as that found from Mulliken populations with an STO-3G basis set carried out on entire amino acids.

There was one exception to this procedure because of the study of Williams and Starr,[21] who showed that for aromatic hydrocarbons, a charge of -0.15 on the carbon and $+0.15$ on the corresponding hydrogen was necessary to optimally fit crystal properties of aromatic hydrocarbon crystals. We modified our charges for phenylalanine, accordingly but left the tyrosine and tryptophan charges as before. We also carried out molecular mechanics calculations on $(benzene)_2$, with this set of charges and found that our parameters ($\varepsilon = 1$) led to a "herringbone" benzene dimer, as is found in the crystal, to be stable relative to two benzene monomers by -2.7 kcal/mol, compared to a stabilization of -1.7 kcal/mol found for a stacked centrosymmetric dimer, qualitatively consistent with Janda *et al*'s observation[22] that benzene dimer is polar. However, with $\varepsilon = R_{ij}$, the calculated values are -2.8 and -2.6 kcal/mol, respectively, reflecting the fact that van der Waals forces favor the stacked dimer and electrostatic the herringbone dimer. One can still use $\varepsilon = R_{ij}$ and get qualitative agreement with experiment by changing our CA van der Waals parameters in a manner analogous to that described above for HC: with $R^* = 1.96$ Å and $\varepsilon = 0.06$ kcal/mol, both dielectric models find herringbone more stable than parallel. The complete charge models for all residues in the AMBER data base appear in Figures 1 and 2. The typical quality of fits to calculated electrostatic potentials are analyzed in ref. 3.

Test Cases on Insulin and Myoglobin

Since force fields for systems as complex as proteins are constantly evolving and being "fine tuned", we decided in our earlier work to select two structurally well resolved proteins for which to use as standards for comparing different parameter sets. In Table VI we present structural results on the effects of different force fields after energy refinement of insulin and myoglobin starting with the x-ray crystal structures.[23,24] An earlier discussion on the general question of protein compaction upon energy minimization and on the various

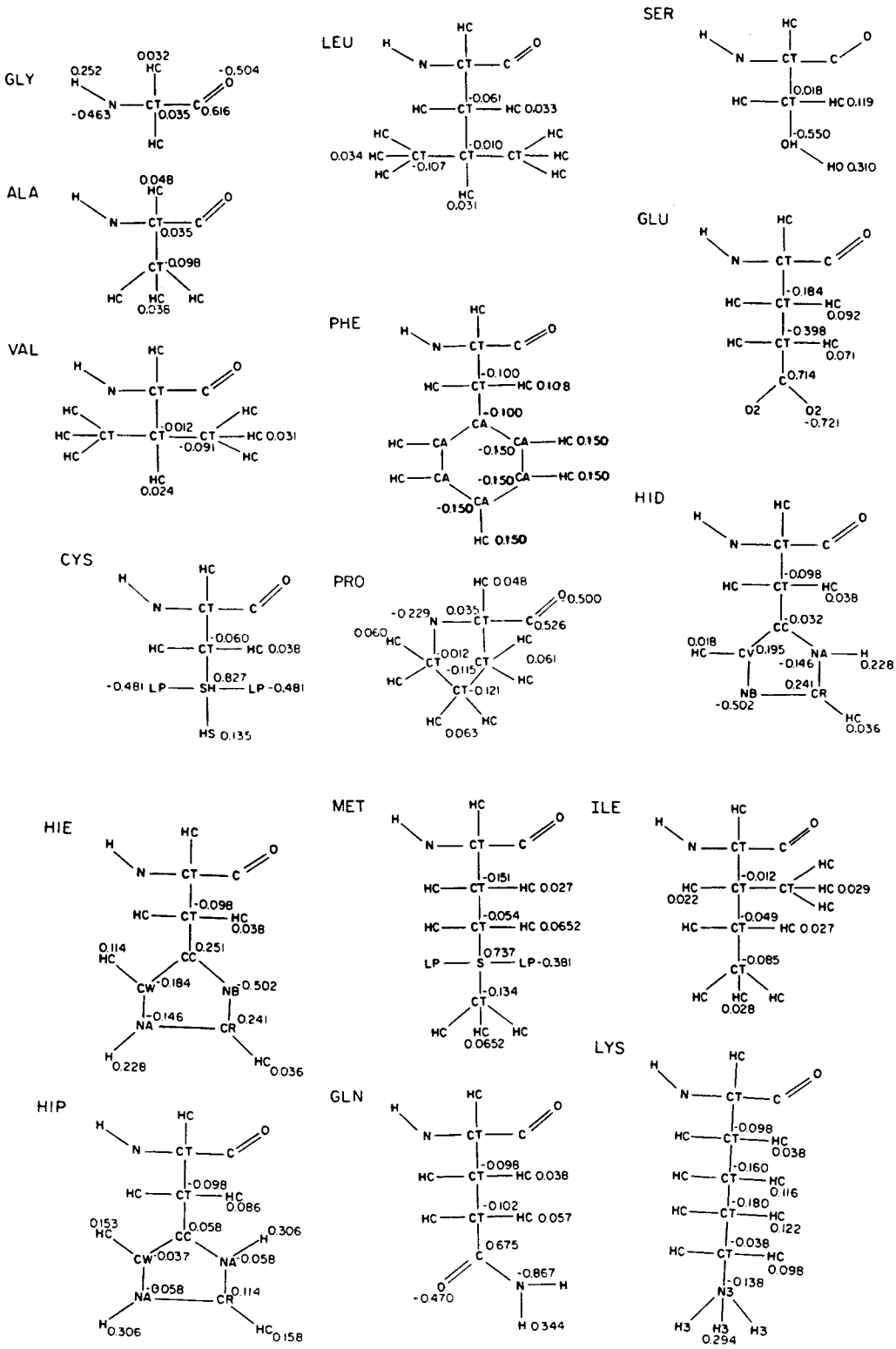

Figure 1. Charges and atom types for amino acid residues.

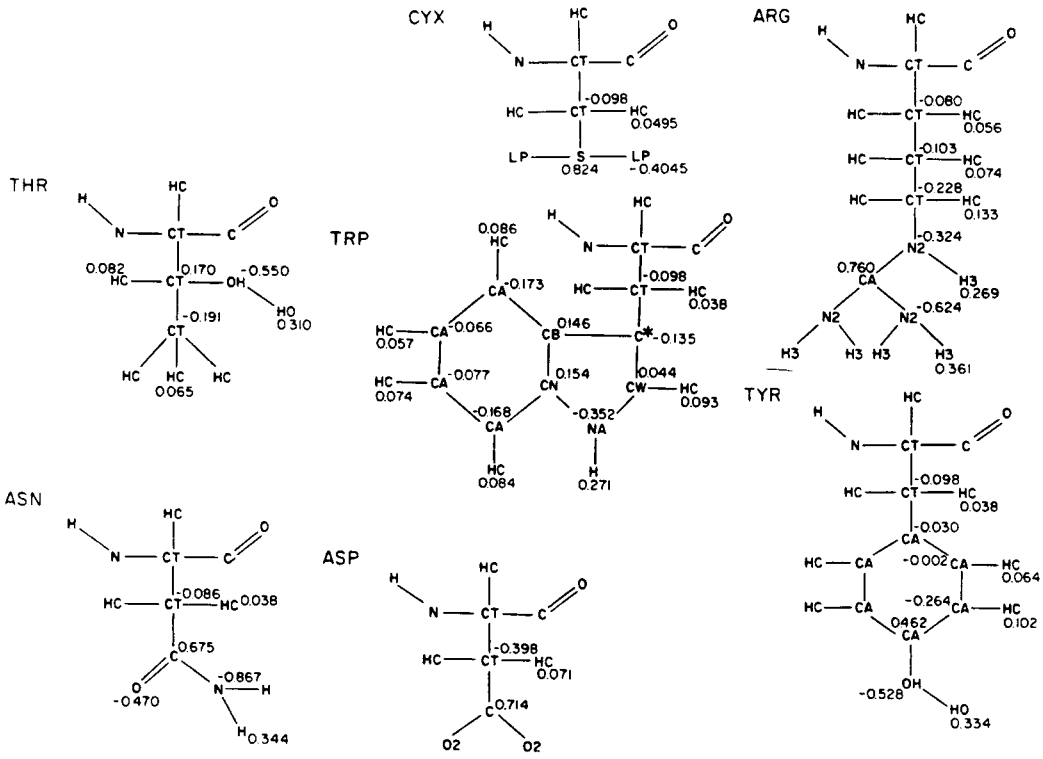

Figure 1 (Continued)

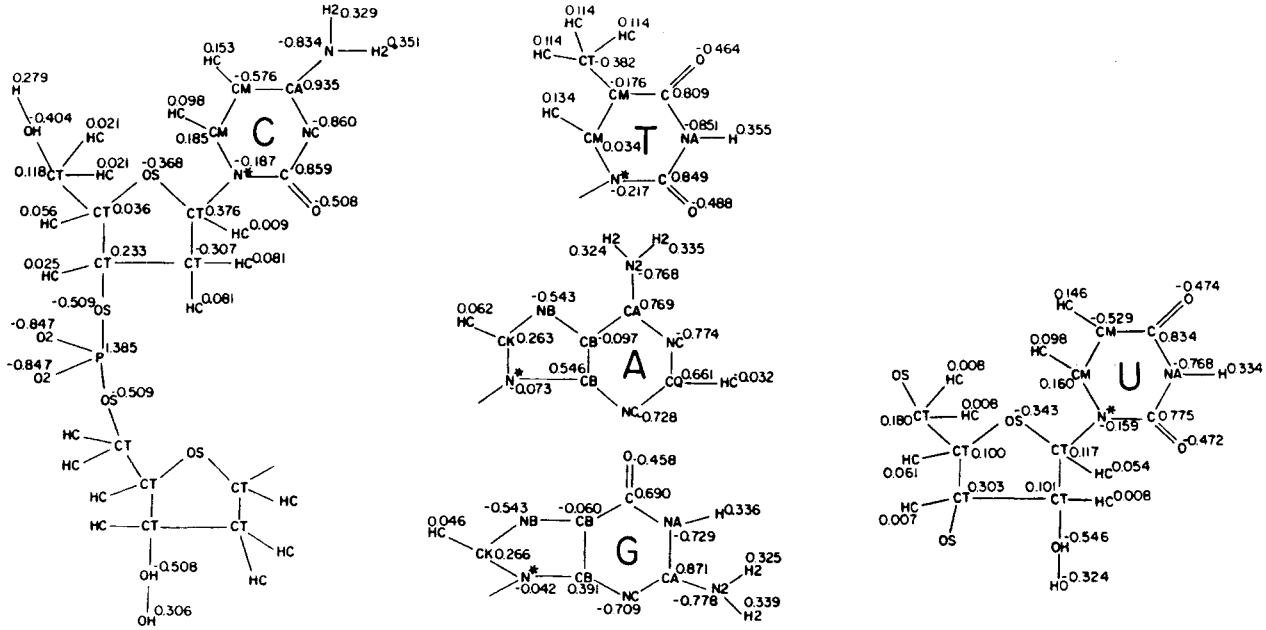

Figure 2. Charges and atom types for nucleic acid residues.

united atom results appears in ref. 1. For both insulin and myoglobin the all atom force field appears to compact the protein less than with the united atom parameter set with "normal" van der Waals radii (1.8% vs. 6.8% for myoglobin and 1.7% vs. 7.0% for insulin). Further, the root mean square (RMS) movement for backbone atoms appears to be slightly less in the all atom case. The use of the larger united atom van der Waals radii suggested by Jorgensen[25] leads to comparable compaction to the all-atom simulation. We emphasize that one could energy refine these structures to lower energy gradients to compare more

Table VI. Protein Refinement.

	Energy Evaluations	RMS[a] Gradient	RMS[b] Backbone	RMS[c] All Atoms	Compaction/ Expansion[d]
Refinement of Insulin					
		United Atom			
Cutoff[a] = 9.0 Å, $\varepsilon = R_{ij}$	1318	0.14	0.28	0.43	−1.6%
Cutoff = 12.0 Å, $\varepsilon = R_{ij}$	3937	0.09	0.56	0.72	−7.0%
		United Atom-Jorgensen Nonbonded[f]			
Cutoff = 9.0 Å, $\varepsilon = R_{ij}$	1080	0.20	0.26	0.41	+0.2%
		All Atom			
Cutoff = 8.0 Å, $\varepsilon = R_{ij}$	1780	0.097	0.27	0.44	−1.7%
Refinement of Myoglobin					
		United Atom			
Cutoff = 9.0 Å, $\varepsilon = R_{ij}$	2000	0.56	0.54	0.82	−6.8%
		United Atom-Jorgenson Nonbonded[f]			
Cutoff = 9.0 Å, $\varepsilon = R_{ij}$	2000	0.70	0.41	0.71	−1.8%
		All Atom			
Cutoff[g] = 5.0 Å, $\varepsilon = R_{ij}$	2000	0.25	0.35	0.65	−1.8%

[a]RMS gradient, in units of kcal/Å, calculated at end of the energy refinement.
[b]Root mean square fit in Å for the minimized insulin backbone atoms, compared with coordinates from the starting crystal structure.
[c]Root mean square fit in Å for all the minimized insulin atoms, compared with coordinates from the starting crystal structure.
[d]These values represent the ratio of the minimized volume to initial volume. The specific volumes were generated using the radius of gyration calculated from all insulin backbone atoms.
[e]Cutoff is the distance up to which all nonbonded interactions will be evaluated. This is analogous to the simulation reported in ref. 1. The number of non-bonded interactions in insulin is about 50,000.
[f]Ref. 25 non-bonded parameters used for CH, C2 and C3 united atoms; the remaining terms as in ref. 1.
[g]Used residue based rather than atom based cutoff; i.e. if any atom of two residues is within 5 Å of an atom in the other, all of the atom-atom non-bonded interactions between the two residues are included in the calculations. The use of a 5 Å residue based cutoff for myoglobin leads to a comparable number of non-bonded interactions as an atom based 8 Å cutoff. The number of non-bonded interactions included is about 100,000.

precisely with x-ray diffraction structures. Given that we have not included water or crystal neighbors in the calculation and the experiment does not correspond to 0° K structure, a more detailed comparison is not warranted at this time, but we do plan further study on this question.

Vibrational Frequencies of the Bases, Imidazole, and Indole

Our purpose in analyzing the normal mode spectra of the protein and nucleic acid fragments is different from that of many vibrational studies. We are not directly interested in assignments of observed lines, nor in obtaining the best possible fit to experimental data. The approximations inherent in our representation of the force field (e.g., we have no stretch-stretch or stretch-bend interaction terms), and our desire to maintain as much transferability as possible among molecules precludes the sort of precise fitting (to within a few cm^{-1}) that is possible for many small organic molecules. Instead, we wish to ensure that the general features of the vibrational spectra are reproduced in our calculations, particularly for the low frequency normal modes that are important in discussions of conformational flexibility.

Unfortunately, it is just for these low frequency modes that the experimental evidence is the least reliable. Even for uracil, the simplest of the nucleic acid bases, the experimental identification of the low frequency out-of-plane modes is only tentative; even greater uncertainties exist for the purines. The present calculations can thus be viewed as *predictions* that may spur increased experimental attention to this problem.

Vibrational analysis on the nucleic acid bases, imidazole and indole presented below generally support the simple interpolation model we employed earlier to derive force constants for bond stretch and bending and Fourier coefficients for the torsional energies. Such a model works surprisingly well for the nucleic acid bases, including both six-membered rings and fused 6/5 ring systems. The only substantive change from the united

atom force field has been the inclusion of improper torsions for out of plane bending for C—H bonds and for exocyclic NH_2 groups. The latter was not included in the united atom force field because at that time, we did not have definitive data on whether it was necessary, given the fact that in aniline, the equilibrium structure for the exocyclic N is nonplanar, and even in the nucleic acid bases, the energy for distortion from planarity would be expected to be small. Our normal mode calculations presented here have shown that a small ($V_2/2$ = 1.0 kcal/mole) improper torsion for these exocyclic NH_2 groups significantly improves the agreement with experiment.

Uracil

Calculated and experimental frequencies for uracil are collected in Table VII. For all but the lowest frequencies, we have listed the lines seen in a recent study of uracil trapped in a low-temperature nitrogen matrix.[26] These should have fewer intermolecular perturbations on the spectrum than do earlier studies in solution or in concentrated crystals. Since the low frequency modes were not identified in ref. 26, we have reported earlier tentative identifications by Beetz and Ascarelli,[27] based on Raman spectra. The in-plane frequencies have also been estimated by *ab initio* quantum mechanical calculations using an STO-3G basis set.[28] The force constants from this basis were "corrected" to those appropriate to the 431-G basis by comparison of fragment calculations in which both basis sets were used. The resulting frequencies are uniformly 10–14% higher than experimental ones and the *ab initio* values reported in Table VII have been corrected by a single scaling factor of 0.92, which bring

Table VII. Vibrational frequencies (cm^{-1}) for uracil.

# description[a]	exp.[b]	*ab initio*[c]	this work	
			united atom	all atom
In-plane vibrations:				
1 N1—H str	3470	3546(−9)	3310	3310(−8)
2 N3—H str	3423	3520(−9)	3307	3307(−8)
3 C5—H str	3130	3128(0)	—	2955(0)
4 C6—H str	2970	3090(0)	—	2953(0)
5 C2=O str	1779	1809(−7)	1788	1742(−10)
6 C4=O str (Ur I)	1727(−2)	1782(−1)	1736	1663(−12)
7 C=C str (Ur II)	1645(−3)	1696(−2)	1639	1820(−5)
8 ring str (Ur III)	1477(−13)	1512(−15)	1551	1552(−8)
9 ring str	1463	1424(−18)	1506	1533(−7)
10 N3—H bend	1405	1405(−8)	1454	1420(−7)
11 ring str (Ur IV)	1380(−7)	1398(−6)	1304	1320(−2)
12 ring str, N1—H bend (Ur V)	1247(−8)	1272(−3)	1155	1303(−5)
13 C6—H bend	1192	1194(−2)	—	1144(−1)
14 C5—H bend	1171	1090(−11)	—	1043(−5)
15 ring def. I	977	977(−15)	940	938(−14)
16 ring str	965	950(−15)	896	886(−16)
17 ring breathing (Ur VI)	724(−2)	754(−5)	715	714(−6)
18 ring def. III	557(−6)	561(−3)	578	567(−4)
19 ring def. II	518(−5)	542(−3)	545	544(−5)
20 C=O in phase bend	538	512(−5)	520	513(−3)
21 C=O out-of-phase bend	401(−2)	382(−2)	373	371(−2)
Out-of-plane vibrations:				
1 C6—H wag	846	—	—	908(−1)
2 C5—H wag	811	—	—	818(−1)
3 ring def.	762	—	945	804(−6)
4 N3—H wag	685	—	723	678(−3)
5 N1—H wag	677	—	645	653(−1)
6 ring def.	592	—	544	589(0)
7 ring def.	435[d]	—	516	453(−1)
8 C4=O wag	194[d]	—	182	183(−1)
9 C2=O wag	167[d]	—	168	166(−2)

[a]See ref. 32 for uracil nomenclature; values in parenthesis are ^{15}N isotope shifts.
[b]Ref. 26, except where noted. Isotope shifts are from Ref. 32.
[c]Determined from modified STO-3G force constants, scaled by 0.92. See Table V of Ref. 28.
[d]Ref. 27

them in line with experimental data.[29] The brief descriptive assignments for the experimental spectra are based on ^{2}H and ^{15}N isotopic substitution and on studies of thiouracils.[27] The final columns of this table report results from the present calculation and from our earlier, united atom force field.[1]

Overall, the agreement between theoretical and experimental results is satisfactory for the low frequency modes. Below 800 cm^{-1}, the largest deviation between the all atom results and experiment is 30 cm^{-1} (in plane mode No. 21). The average absolute relative error is 5.9% and the largest mismatch for any frequency is 175 cm^{-1} in an in-plane mode (No. 7) that has the C5=C6 stretch as a principal component. A similar discrepancy shows up in the other bases with differences ranging from 235 cm^{-1} for thymine to 257 cm^{-1} for adenine. This problem arises from inaccuracies in the linear interpolation model used to determine our force constants: a model based on benzene and ethylene data apparently cannot be transferred directly to these heterocyclic bases. While it would not be difficult to lower the stretching constants in an *ad hoc* fashion to obtain better agreement with experiment, we have not done so here, feeling that the advantages of a general procedure for assigning force constants outweigh the disadvantages of these errors in such high frequency vibrations. Similar comments could be made for the only other large frequency mismatch in the uracil calculation, which is 128 cm^{-1} in an in-plane mode (No. 14) involving the C5—H bend as a principal component.

Similar results, with slightly larger errors, are found with the united atom force field (Table VII). This is to be expected, since the only difference between the two potentials involves the parameters associated with the two hydrogens attached to C5 and C6. Except for these, the constants in both force fields were determined by the procedure outlined earlier,[1] which made use of the observed vibrational spectra of other molecules but not the nucleic acid bases.

Theory and experiment both show only two normal modes below 300 cm^{-1}, and these are depicted in Figure 3. The largest energy component to these modes involves the out-of-plane motion of the exocyclic oxygens, but substantial distortion of the rest of the ring,

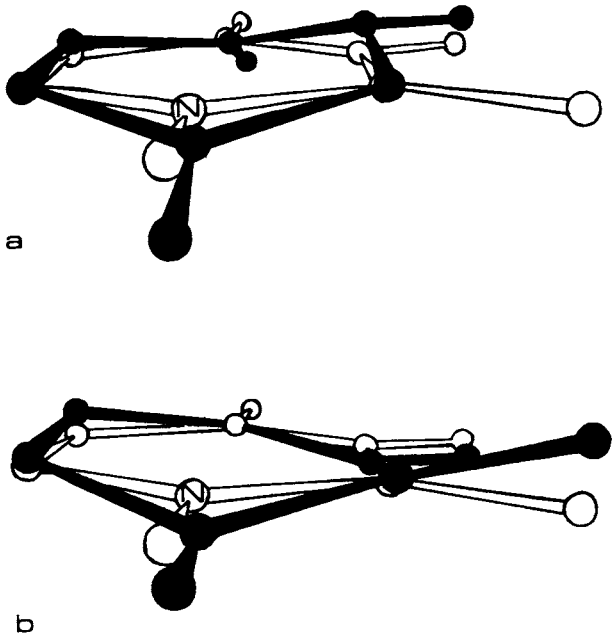

Figure 3. The two lowest normal modes of uracil.

including pyramidalization of the nitrogens, also occurs. Even these lowest frequencies are large compared with those less than 50 cm^{-1} that are responsible for most conformational fluctuations in large molecules.[30] Pictures of the remaining, higher frequency modes will be presented elsewhere.[31]

It should be noted that some of the detailed features of the frequencies in Table VII leave room for improvement. For example, the two C—H stretching frequencies seen experimentally are split by 60 cm^{-1}, whereas they are nearly degenerate in the calculation, a result of our using the same atom type for both C5 and C6. Similar comments apply to the two N—H stretches. We have also calculated the frequency shifts for isotopic substitution of both nitrogens with ^{15}N (Table VII). Our calculations predict fairly large (12 cm^{-1}) shifts for in-plane mode 6 (the so-called "UrI" band), whereas the observed shift[32] is much smaller (2 cm^{-1}). Both the experimental and *ab initio* results suggest that the large isotope shifts should occur instead for modes 8 and 9 (the "UrIII" band). This suggests that the ring deformations sensitive to the nitrogen mass are located at somewhat too high a frequency with the present potential. For the lower frequency modes, our results are in better agreement with the quantum calculations and experiment, giving large shifts (~15 cm^{-1}) only for the in-plane ring deformations 15 and 16.

Thymine

Experimental[33] and calculated frequencies for thymine are given in Table VIII. There is less extensive experimental data for this base, and our "assignments" for the 900–1500 cm^{-1} region are based only on alignment of observed and calculated frequencies. We now predict four out-of-plane frequencies below 400 cm^{-1}, all primarily involving exocyclic motion: wags of the two carbonyls and the methyl group, and the torsional rotation of the methyl group, whose exact frequency is not important, since this motion is expected to be very anharmonic. The ring deformation mode at 114 cm^{-1} is, in fact, the lowest frequency predicted for any of the five bases considered here. This motion is depicted in Figure 4, in which one can see that the deformation involves the methyl and carbonyl groups bending to the same side of the ring. This sort of deformation may be of low enough frequency to be accessible in DNA's, and could contribute to propeller twists in a base pair. In other respects the frequencies of thymine are similar to those of uracil.

Table VIII. Vibrational frequencies (cm^{-1}) for thymine.

#	description	exp.[a]	this work
In-plane vibrations:			
1	N1—H str	3210	3310
2	N3—H str	3175	3037
3	C6—H str	3063	2954
4	C—H (methyl) str	2993	2943
5	C—H (methyl) str	—	2940
6	C—H (methyl) str	2990	2827
7	C2=O str	1735	1741
8	C4=O str	1677	1679
9	C=C str	1600	1835
10	ring str	—	1611
11	ring str, N1—H bend	1495	1540
12	ring str	1483	1413
13	CH_3 def.	1461	1386
14	CH_3 def.	—	1385
15	N3—H bend, ring str.	1406	1302
16	CH_3 def.	1382	1369
17	C6—H bend	1366	1269
18	ring str	1245	1143
19	C—CH_3 str, ring str	1203	1099
20	ring str	1028	903
21	ring bend	984	856
22	ring breathing	815	719
23	ring bend	806	715
24	C=O bend (in-phase)	617	584
25	ring def.	560	542
26	ring def.	475	455
27	C=O bend (out-of-phase)	392	368
28	C—CH_3 bend	321	306
Out-of-plane vibrations:			
1	N1—H wag	885	705
2	N3—H wag	818	660
3	C6—H wag	764	903
4	ring def.	852	910
5	ring def.	818	801
6	ring def.	635[b]	588
7	ring def.	433[b]	456
8	C—CH_3 wag	285[b]	276
9	C=O wag	206[b]	172
10	C—CH_3, C=O wag	—	114
11	CH_3 rot.	—	78

[a]Ref. 33 except where noted.
[b]Ref. 27.

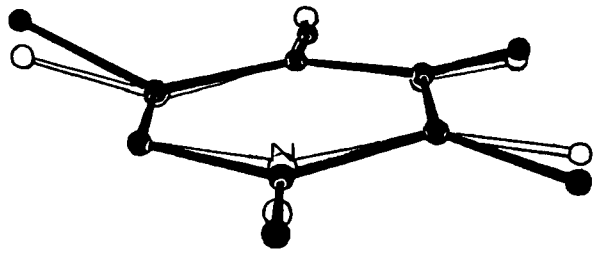

Figure 4. The lowest frequency ring deformation in thymine.

Cytosine

The experimental results given in Table IX are based on infrared and Raman spectra of polycrystalline cytosine.[34] The assignments are based on deuteration shifts, comparisons with uracil, and a normal mode calculation using a valence force field for bond angle and torsional terms.[35] Our calculated frequencies have an average absolute relative error of 5.6%. We again have two out-of-plane frequencies below 400 cm^{-1}. These involve the exocyclic C=O and C—NH_2 groups, and are at slightly higher frequencies than the corresponding modes in uracil. In analogy to the uracil-thymine difference, one expects 5-methyl cytosine to have a much lower frequency than cytosine itself, and perhaps be able to distort slightly from a planar structure in response to external forces. Calculated frequencies for 5-methyl-cytosine indeed show out-of-plane modes at 115 and 197 cm^{-1}, corresponding to the C—NH_2 and C=O wagging motions. These large decreases from the cytosine values (Table IX) support the notion that it is the heavy atom exocyclic substituents that dominate the lowest frequency motions in the DNA bases.

All of our calculations of the pyrimidines exhibit systematic errors for in-plane vibrations in the 700–1800 cm^{-1} region: calculated frequencies in the upper end of this region (from 1300–1800 cm^{-1}) tend to be too high

Table IX. Vibrational frequencies (cm^{-1}) for cytosine.

#	description	exp.[a]	prev. calc.[b]	this work
In plane vibrations:				
1	NH_2 str (out-of-phase)	3380	3343	3360
2	N1—H str	3169	3230	3309
3	NH_2 str (in-phase)	3230[c]	3167	3247
4	C6—H str	3117[c]	3120	2954
5	C5—H str	3117[c]	3106	2953
6	C=C str	1615	1710	1795
7	C=O str	1662	1664	1741
8	NH_2 bend	1703	1641	1694
9	ring str	1538	1541	1629
10	ring str	1505	1505	1592
11	C—NH_2, ring str	1465	1452	1499
12	N1—H bend	1364	1286	1413
13	C5—H bend	1277	1235	1302
14	C6—H bend	1236	1088	1148
15	ring def.	1100	1082	1048
16	ring def., NH_2 wag	1010	960	999
17	ring def.	994	882	963
18	NH_2 wag	966	798	904
19	ring breathing	793	656	712
20	ring def.	600	475	548
21	ring def.	549	440	531
22	C=O bend	533	525	496
23	C—NH_2 bend	400[c]	349	340
Out-of-plane vibrations:				
1	C6—H wag	894[c]	1365	911
2	C5—H wag	823	1039	821
3	ring def.	782	757	788
4	N1—H wag	760	823	652
5	ring def.	566	603	577
6	NH_2 rock	548[d]	543	641
7	NH_2 wag	485[d]	484	468
8	ring def.	421	408	392
9	C=O wag	232[d]	232	212
10	C—NH_2 wag	197[d]	168	177

[a]Ref. 34, except where noted.
[b]Ref. 35.
[c]Raman frequency from ref. 34.
[d]Ref. 27.

compared with experiment, while those in the lower range are too low. Fairly complex ring deformation and bending motions are involved in these modes, and the origin of this bias is not clear. It may be that the stretching force constants (which are relatively more important for the higher frequencies) are too high relative to the bending force constants (which dominate the lower frequencies).

Adenine

The experimental uncertainties in assignments for the pyrimidine spectra are magnified for the purines, which have a double ring system and a much larger number of fundamental vibrational excitations. Results for adenine are collected in Table X comparing infrared and Raman data for polycrystalline adenine[36,37] to the present calculations and to quantum chemical calculations analogous to those described above for uracil.[37] The agreement with experiment is not so close here as it was for the pyrimidines. For example, although both our results and the quantum mechanical calculations place the NH_2 scissoring mode close to its observed position at 1670 cm^{-1}, the experimental results indicate that this should be the highest frequency aside from the hydrogen stretches; the present calculations have three ring deformation modes above it, at 1718, 1809, and 1827 cm^{-1}. This is in line with the bias seen in the pyrimidines, which carries over to the purines: frequencies from 1300–1800 cm^{-1} tend to be too high in the present potential, while those below this range are too low. Hence, the average absolute relative error

Table X. Vibrational frequencies (cm^{-1}) for adenine.

#	description	exp.[a]	exp.[b]	*ab initio*[b]	this work
In-plane vibrations:					
1	NH_2 str (out-of-phase)	3294	3340	3684	3362
2	N9—H str	2800	3240	3465	3008
3	NH_2 str (in-phase)	3118	3280	3490	3248
4	C8—H str	—	3120	3143	2953
5	C2—H str	—	3050	3069	2950
6	C5=C6 str	1570	1580	1644	1827
7	ring str (py)	1604	1620	1694	1809
8	ring str (py)	1449	1490	1505	1718
9	NH_2 scissors	1673	1670	1713	1688
10	ring str (Im)	1507	1520	1549	1644
11	ring str (Im)	—	—	1500	1593
12	C—NH_2 str, NH_2 rock	1308	1312	1248	1514
13	N9—H bend	1252	1425	1408	1414
14	ring str (py)	1334	1332	1339	1358
15	ring str (Im)	1417	1255	1240	1196
16	C8—H bend	—	1225	1149	1154
17	C2—H bend	1235	1380	1349	1158
18	ring str (py)	1125	1124	1124	1073
19	ring str	1156	1180	1083	1062
20	NH_2 rock	1023	945	997	942
21	ring bend (py)	937	900	936	847
22	ring bend (Im)	847	820	911	780
23	ring bend (Im)	723	725	708	620
24	ring bend	621	620	626	562
25	ring bend (py)	530	535	542	538
26	ring bend (py)	—	—	516	471
27	C—NH_2 bend	250	334	274	288
Out-of-plane vibrations:					
1	C2—H wag	912	—	—	990
2	N9—H wag	—	—	—	797
3	ring (py) def.	871	—	—	788
4	C8—H wag	796	—	—	880
5	ring (Im) def.	639	—	—	610
6	NH_2 rock	645	—	—	644
7	ring (Im) def.	542	—	—	529
8	NH_2 wag	—	—	—	483
9	ring (py) def.	337	—	—	443
10	ring (butterfly)	248	—	—	274
11	ring (propeller)	237	—	—	229
12	ring def. + C—NH_2 wag	—	—	—	179

[a]Ref. 36.
[b]Ref. 37.

increases to 7.6%. As with the pyrimidines, the low frequency spectra of greatest interest to us appear to be fairly well represented with this force field.

The double ring system in adenine does not appear to give it enhanced conformational flexibility compared with the single ring pyrimidines. There are four frequencies below 300 cm^{-1}, with the lowest being estimated at 179 cm^{-1}. One of these is an in-plane bending of the NH_2 group, and the three lowest out-of-plane frequencies are shown in Figure 5. The lowest frequency consists primarily of a wag of the exocyclic NH_2 group, accompanied by some distortion of the five membered ring. The second lowest out-of-plane mode is a "butterfly" distortion about the ring junction, while the third-lowest mode is a "propellor twist", with the two rings being distorted in opposite directions. All three of these out-of-plane modes resemble vibrational modes of a rectangular membrane, and their relatively high frequencies indicate that the forces maintaining the planarity of the adenine ring system are fairly strong. As we saw above in the uracil-thymine comparison, low vibrational frequencies are closely related to the presence of heavy exocyclic substituents, of which adenine has only one, the NH_2 group on C4.

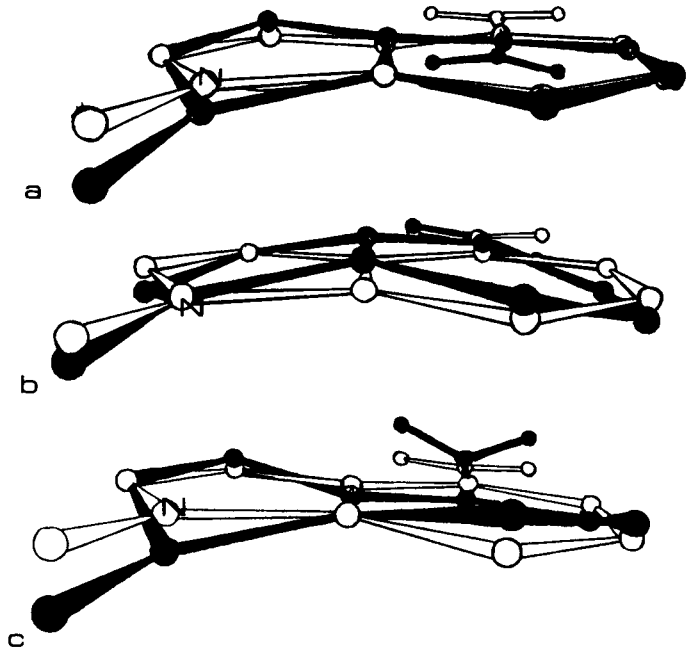

Figure 5. The three lowest frequency out-of-plane normal modes in adenine.

Guanine

Table XI gives calculated and experimental[38,39] frequencies for guanine. The difficulties in assignments mentioned above for cytosine and adenine are present here as well. Because it has one more heavy exocyclic atom than adenine, we now see four out-of-plane vibrations below 300 cm^{-1}, with frequencies somewhat lower than those in adenine. The three lowest of these are depicted in Figure 6. The lowest mode again involves primarily an out-of-plane wag of the NH_2 group, accompanied by some distortion of the pyrimidine and imidazole rings. In guanine the $C—NH_2$ is nearly parallel to the long axis of the double ring "rectangle", whereas in adenine it is parallel to the short axis of the rectangle. This may explain why the lowest frequency is lower in guanine than in adenine, since the accompanying distortion of the ring system is

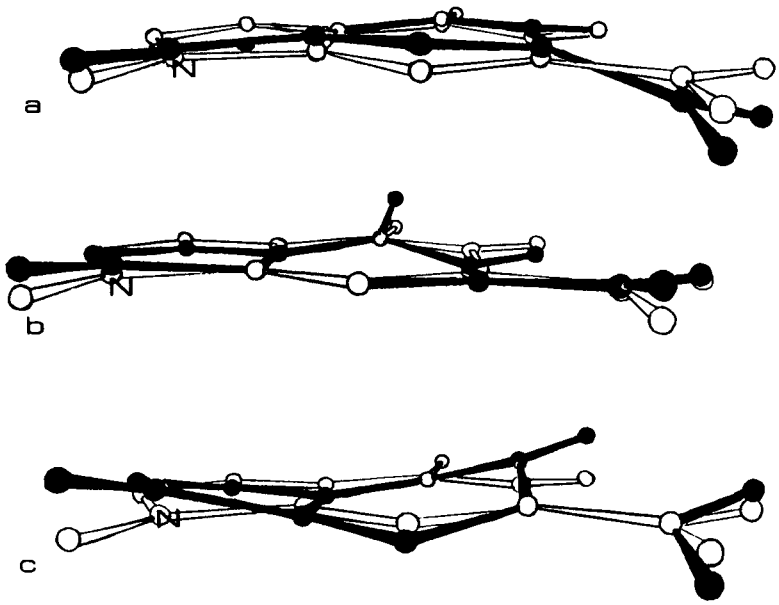

Figure 6. The three lowest frequency out-of-plane normal modes in guanine.

Table XI. Vibrational frequencies (cm^{-1}) for guanine.

#	description	exp.[a]	this work
In-plane vibrations:			
1	NH_2 str (out-of-phase)	3316	3358
2	N9—H str	2908	3308
3	N1—H str	2698	3306
4	NH_2 str (in-phase)	3114	3248
5	C8—H str	—	2953
6	C=C str	1638	1847
7	C=O str	1702	1793
8	NH_2 bend	1675	1719
9	ring str (py)	1565	1670
10	ring str (Im)	1550	1640
11	ring str (py)	1477	1620
12	ring str (Im)	1464	1576
13	ring str (py)	1420	1535
14	N9—H bend	1375	1394
15	C—NH_2 str	1390	1295
16	N1—H bend	1362	1290
17	ring str (Im)	1261	1173
18	C8—H bend	1174	1087
19	ring str (py)	1118	1074
20	ring bend (py, Im)	1042	984
21	ring bend, (NH_2 wag)	950	931
22	ring bend (Im)	850	817
23	ring breathing (Im, py)	778	724
24	ring def. (Im)	690	657
25	ring def. (py)	645	562
26	ring def. (py)	557	533
27	ring def. (py)	501	446
28	C=O bend	400	342
29	C—NH_2 bend	343	310
Out-of-plane vibrations:			
1	N9—H wag	884	795
2	C8—H wag	790	988
3	N1—H wag	726	780
4	ring (py) def.	690	672
5	NH_2 rock	654	654
6	ring (Im) def.	601	563
7	NH_2 wag	578	439
8	ring (py) def.	416	488
9	ring (Im) def.	490	450
10	ring (propeller) def.	243	282
11	ring (butterfly) def.	214	222
12	C=O wag	170	180
13	C—NH_2 wag	142	128

[a]Ref. 38 and Ref. 39.

easier to accomplish along the long direction. The second-lowest normal mode in guanine involves the C=O out-of-plane motion, and the butterfly and propellor distortions may be seen in the next two frequencies. As with adenine, these frequencies are large enough to suggest that distortions of these sorts will not contribute in a major way to conformational fluctuations in DNA.

In the uracil/thymine and cytosine/5-methyl-cytosine comparisons mentioned above, we observed the important effect of the masses of exocyclic groups on the low frequency out of plane modes. Since the bases are attached to the sugar-phosphate backbone of DNA by such a connection, these motions may move to lower frequency in a DNA-like environment. To simulate this, we have performed additional model vibrational calculations in which the mass of the hydrogen attached to N1 or N9 is changed to 99 amu. In every case, the frequency of the N—H out-of-plane wag is reduced to between 70 and 85 cm^{-1}. Other internal modes of the bases are shifted by only small amounts. This very crude model suggests that mass effects alone may move the frequency of base backbone motion into the 0–100 cm^{-1} range that dominates the conformational dynamics of large biomolecules. True internal motions of the bases, however, are expected to start at frequencies higher than this, as illustrated in the tables.

Imidazole

The calculated and experimental frequencies for imidazole are presented in Table XII. Normal mode calculations in which we proceed as previously[1] (column 2, Table XII) have some severe errors in the low frequency region: (a) we predict a frequency of 386 cm^{-1}

Table XII. Vibrational frequencies (cm^{-1}) of imidazole.

#	description	exp.[a]	calc.[b]	calc.[c]
In-plane vibrations:				
1	N1—H1 str	—	3309	3310
2	C5—H5 str	—	2956	2957
3	C2—H2 str	—	2953	2954
4	C4—H4 str	—	2951	2952
5	C=C str	1630	1658	1693
6	C—N str	1530	1639	1651
7	ring str, N1—H bend	1480	1549	1571
8	ring str	1405	1468	1490
9	ring str	1320	1247	1261
10	C5—H bend	1260	1133	1138
11	ring str	1120	1084	1096
12	C4—H bend	1084	1035	1060
13	C2—H bend	1055	1011	1021
14	ring bend	890	736	740
15	ring bend	855	720	725
Out-of plane vibrations:				
1	C5—H5 wag	965	1053	920
2	C2—H2 wag	809	934	839
3	C4—H4 wag	723	728	676
4	ring def.	660	634	654
5	ring def.	612	524	577
6	N1—H1 wag	517	386	509

[a]Ref. 40.
[b]Normal mode calculation w/o angle strain.
[c]Normal mode calculation w/angle strain.

for the N1—H out-of-plane wagging motion compared to the experimental value of 517 cm^{-1}; (b) The experimental frequency at 612 cm^{-1} has been assigned as an out-of-plane ring deformation mode yet our force field yields 524 cm^{-1}. Our analysis also gives frequencies more than 100 cm^{-1} too high for the out-of-plane C2—H2 and C5—H5 wagging motions. These discrepancies are of special concern to us since they occur at frequencies low enough to potentially influence conformational fluctuations in proteins.

To overcome these problems without deviating from our basic linear interpolation model, we decided to set the θ_0 for the C2—N3—C4 angle to 117.0 and the remaining θ_0 to 120°. By doing this we increase the sum of the internal angle θ_0's from 540° in our earlier force field[1] to 597°, introducing significant angle strain into the ring. The results are listed in the final column of Table XII. The difference between experimental and calculated frequency for the N1—H out-of-plane wagging mode is now only 8 cm^{-1}. At the same time, we increase the two out-of-plane ring deformation frequencies and decrease the out-of-plane C—H wagging motions. The average absolute error of this calculation compared to experiment is 5.9%. Here, as with our calculation with the previous five nucleic acid constituent bases, we observe the same systematic errors: the theoretical frequencies in the 1300–1800 cm^{-1} region tend to be high in comparison with experiment and those in the 700–1300 cm^{-1} range are low. The new θ_0 values have little effect on the calculated geometry. In comparing the theoretical structure with the neutron diffraction values,[41] average errors of less than 0.01 Å in bond lengths and 1.7 in bond angles are found.

Adding angle strain appears to be one way of improving the worst aspects of the vibrational frequencies and for imidazole (and presumably for other 5-membered planar rings as well) without making major alterations in the philosophy or parameters of our earlier force field.

Indole

Infrared frequencies and theoretical normal modes are gathered in Table XIII. The experimental vibrational frequencies[42] are assigned based on deuterium isotopic substitution studies of indole in the liquid phase at 70°C. A normal mode calculation of indole using the original[1] torsional parameters produces a very floppy ring system. For example, the calculated frequency of the ring propellor motion is lower than the experimental value by 80 cm^{-1}, and for the butterfly mode, our value is too low by 100 cm^{-1}. In general, as shown in Table XIII the calculated frequencies of the out-of-plane ring deformations are too low relative to the experimental data. The

Table XIII. Vibrational frequencies (cm^{-1}) of indole.

Assignment	exp.[a]	calc.[b]	calc.[c]
In-plane vibrations:			
N1—H str	3419	3308	3308
C2—H str	3123	2953	2953
C3—H str	3106	2958	2958
C—H str (benz.)	3050	2955	2955
C—H str (benz.)	3050	2954	2955
C—H str (benz.)	3050	2953	2953
C—H str (benz.)	3050	2952	2952
ring str (benz.)	1616	1838	1846
ring str (benz.)	1576	1746	1753
ring str (pyrr.)	1509	1672	1674
ring str (benz.)	1487	1736	1745
ring str (benz.)	1455	1607	1606
ring str (pyrr.)	1412	1510	1512
ring str (pyrr.)	1352	1393	1395
ring str (benz.)	1334	1542	1542
ring str (pyrr.)	1276	1300	1297
C—H bend (benz.)	1245	1242	1344
ring str (pyrr.)	1203	1227	1232
C—H bend (benz.)	1191	1145	1145
N—H bend	1147	1060	1057
C—H bend (benz.)	1119	1080	1079
C2—H bend	1092	1123	1124
C3—H bend	1064	895	895
C—H bend (benz.)	1010	991	989
ring bend (pyrr.)	895	804	808
ring bend (pyrr.)	767	586	585
ring breathing	758	761	759
ring bend (benz.)	607	533	534
ring bend (benz.)	542	465	466
Out-of-plane vibrations:			
C—H wag (benz.)	970	679	964
C—H wag (benz.)	930	629	796
C—H wag (benz.)	873	565	731
C2—H wag	848	948	977
C—H wag (benz.)	743	492	722
C3—H wag	725	788	867
ring def.	608	677	642
ring def.	575	419	534
N1—H wag	487	461	436
ring def.	423	321	372
ring def.	397	279	344
ring butterfly	254	154	221
ring propellor	224	144	205

[a]Ref. 42.
[b]This work, using torsional potentials from ref. 1.
[c]This work, using the modified torsional parameters described in the text.

predicted out-of-plane benzilic C—H wags are again too low compared to experiment while the out-of-plane pyrrolic C—H wags are too high.

We believe that the problem in predicting the vibrational modes of indole originates in the way we obtain our force constants. The bond lengths around the central C=C of indole are near the "break" of our linear interpolation model. A slight variation of a few hundredths of an Angstrom results in a difference of about 10 kcal/mole in the torsional barrier heights. To obtain reasonable results we have found it necessary to double all the dihedral force constants for the bonds connected to the central $C_5=C_6$ bond and to increase the $V_2/2$ value for the C5=C6 bond to 20 kcal/mole. With these changes in the torsional barrier, (final column of Table XIII) the calculated frequency for the ring propellor is different from the observed value by 19 cm^{-1}. For the ring butterfly motion, the theoretical value is 33 cm^{-1} lower than that of the experiment. The calculated values for the C2—H and C3—H motions are still high in comparison with the experimental data. Since we changed only torsional parameters, the calculated in-plane frequencies for the two calculations are similar. For indole, the average absolute error using the new torsional parameters is 8.8%.

DISCUSSION AND CONCLUSIONS

We have presented an extension of the C—H united force field developed by Weiner et al.[1] to include all atoms explicitly. Such an extension has involved analysis of the structure and energy of many small model test systems. The form of the potential has remained the same, with simple harmonic functions for bond stretching and bending Fourier series for torsions and a Lennard Jones + electrostatic representation of nonbonded interactions, and the reasonable agreement with experiment for the model systems is comparable to that found previously.

The inclusion of C—H groups appears to improve the agreement with experiment for calculated relative ethidium intercalation energies of different DNA sequences.[2] However, the structural differences between all-atom and united atom models are small, and, thus, the simpler model is likely to be satisfactory for many structural analyses. We envision the use of hybrid models in many simulations, with an all atom representation of, for example, the active site residues of enzymes, and a united atom representation of the remaining residues.

We have further evaluated the effect on the structure and energy of scaling 1–4 nonbonded interactions. Given that such a scaling has little effect on the structure and energy of the model systems and that such a scaling is used in the united atom force field, we have decided to employ such a scaling in our atom model as well. Physically, such a scaling makes sense for the van der Waals energies because the shorter distances of 1–4 interactions would be overestimated with a $1/R^{12}$ repulsive function. However, such an approach has a strictly empirical rationale for 1–4 electrostatic energies.

The problem of reproducing the frequencies of a 5-membered ring like imidazole led to a new problem having to do with the use of equilibrium data. In an isolated 5-membered ring, θ_0 values near 108° make the system "stainless" and it may be too easy to distort to smaller θ values in the ring. Such a problem can be ameliorated with the use of *standard* θ_0 values for acyclic systems (θ_0 = 120 for sp^2 XCX and 117° for XNX) and such an approach leads to reasonable agreement between calculated and observed frequencies.

It is not clear why the interpolation model appears to work well for purines and not so well for indole, but perhaps the absence of exocyclic groups in the latter modifies the nature of the low frequency out-of-plane motions. In any case, the difficulties of interpolation between the appropriate torsional Fourier coefficients of ethylene (V_2 = 60 kcal/mole) and benzene (V_2 = 20 kcal/mole), involving a difference in C—C bond length of only ~0.08 Å are apparent, and one must resort to an empirical increase of the torsional barrier for the central C=C bond in indole to lead to a satisfactory agreement between the calculated and experimental frequencies of the low frequency out of plane modes. Thus, it is clear that the linear interpolation model is not as transferable as one had hoped, although in some cases the transferability works well. In any case, it does provide a good place to start in a valence force field vibrational analysis. In some cases empirical

adjustments are necessary, but in all the sp^2 ring systems in nucleic acids and proteins, we have now at our disposal a simple force field that gives a fair representation of the lowest and most important vibrational frequencies ($<400\ cm^{-1}$) undergone by these fragments.

In our previous article, we discussed the development of our force field and that of others, notably those described by Scheraga et al.[43] (ECEPP and UNICEPP). Karplus and co-workers[44] Levitt[45] and Hermans et al.[46] for proteins and by Sasisekharan,[47] Olson and Flory[48] and Levitt[49] for nucleic acids and have analyzed the similarities/differences between our force field and those. Pavitt and Hall[50] have compared a number of force fields (included our united atom one[1]) in their ability to reproduce the intra- and intermolecular structures of some cyclic peptides in crystals. They showed that our force field was the most effective one tested, but still led to larger deviations from experiment than typically found for similar studies of hydrocarbon crystals, where electrostatic effects are negligible. Based on the results presented here, we expect that the all atom model should give structural results comparable to the united atom model.

We have also used the united atom force field, with explicit inclusion of a TIPS3P H_2O model and $\varepsilon = 1$, in studies of formamide hydrolysis in solution,[51] in studies of trypsin catalysis[52] and in molecular dynamics studies of DNA including water and counterions[53] and the calculated results suggest the model is reasonable and has no "gross" flaws. Careful comparisons of our potentials with others that have been proposed will be of considerable value, and we have plans to carry out some of these. We also expect that improvements in our understanding of environmental effects such as solvent and the presence of counterions will enable us to make more secure and straightforward comparisons to experimental data. Nevertheless, the results presented here indicate that a variety of very interesting conformational problems can be studied using the potentials we have presented.

One of the authors (P. A. K.) would like to acknowledge the support of the NIH (GM-29072 and CA-25644) in this study. D. A. C. is an Alfred P. Sloan fellow and acknowledges support from the UC Cancer Research Coordinating Committee. We are also grateful to the UCSF Computer Graphics Lab, supported by RR-1081, for the use of their facilities.

Appendix

Bond Parameters

Bond	K_r	r_{eq}
C -C2	317	1.522
C -C3	317	1.522
C -CA	469	1.400
C -CB	447	1.419
C -CD	469	1.400
C -CH	317	1.522
C -CJ	410	1.444
C -CM	410	1.444
C -CT	335	1.522
C -N	490	1.335
C -N*	424	1.383
C -NA	418	1.388
C -NC	457	1.358
C -O	570	1.229
C -O2	656	1.250
C -OH	450	1.364
C*-C2	317	1.495
C*-CB	388	1.459
C*-CG	546	1.352
C*-CT	317	1.495
C*-CW	546	1.352
C*-HC	340	1.080
C2-C2	260	1.526
C2-C3	260	1.526
C2-CA	317	1.510
C2-CC	317	1.504
C2-CH	260	1.526
C2-N	337	1.449
C2-N2	337	1.463
C2-N3	367	1.471
C2-NT	367	1.471
C2-OH	386	1.425
C2-OS	320	1.425
C2-S	222	1.810
C2-SH	222	1.810
C3-CH	260	1.526
C3-CM	317	1.510
C3-N	337	1.449
C3-N*	337	1.475
C3-N2	337	1.463
C3-N3	367	1.471
C3-OH	386	1.425
C3-OS	320	1.425
C3-S	222	1.810
CA-CA	469	1.400
CA-CB	469	1.404
CA-CD	469	1.400
CA-CJ	427	1.433
CA-CM	427	1.433
CA-CN	469	1.400
CA-CT	317	1.510
CA-HC	340	1.080
CA-N2	481	1.340
CA-NA	427	1.381
CA-NC	483	1.339
CB-CB	520	1.370
CB-CD	469	1.400
CB-CN	447	1.419
CB-N*	436	1.374
CB-NB	414	1.391
CB-NC	461	1.354
CC-CF	512	1.375
CC-CG	518	1.371
CC-CT	317	1.504
CC-CV	512	1.375
CC-CW	518	1.371
CC-NA	422	1.385
CC-NB	410	1.394
CD-CD	469	1.400
CD-CN	469	1.400
CE-N*	440	1.371
CE-NB	529	1.304
CF-NB	410	1.394
CG-NA	427	1.381
CH-CH	260	1.526
CH-N	337	1.449
CH-N*	337	1.475
CH-NT	367	1.471
CH-OH	386	1.425
CH-OS	320	1.425
CI-NC	502	1.324
CJ-CJ	549	1.350
CJ-CM	549	1.350
CJ-N*	448	1.365
CK-HC	340	1.080
CK-N*	440	1.371
CK-NB	529	1.304
CM-CM	549	1.350
CM-CT	317	1.510
CM-HC	340	1.080
CM-N*	448	1.365
CN-NA	428	1.380
CP-NA	477	1.343
CP-NB	488	1.335
CQ-HC	340	1.080
CQ-HC	340	1.080
CQ-NC	502	1.324
CR-HC	340	1.080
CR-NA	477	1.343
CR-NB	488	1.335
CT-CT	310	1.526
CT-HC	331	1.090
CT-N	355	1.449
CT-N*	337	1.475
CT-N2	337	1.463
CT-N3	367	1.471
CT-OH	320	1.410
CT-OS	320	1.410
CT-S	222	1.810
CT-SH	222	1.810
CV-HC	340	1.080
CV-NB	410	1.394
CW-HC	340	1.080
CW-NA	427	1.381
H -N	434	1.010
H -N2	434	1.010
H -NA	434	1.010
H2-N	434	1.010
H2-N2	434	1.010
H2-NT	434	1.010
H3-N2	434	1.010
H3-N3	434	1.010
HO-OH	553	0.960
HO-OS	553	0.960
HS-SH	274	1.336
LP-S	600	0.679
LP-SH	600	0.679
O2-P	525	1.480
OH-P	230	1.610
OS-P	230	1.610
S -S	166	2.038

Angle Parameters

Angle	K_θ	θ_{eq}
C -C2-C2	63	112.4
C -C2-CH	63	112.4
C -C2-N	80	110.3
C -C2-NT	80	111.2
C -CA-CA	85	120.0
C -CA-HC	35	120.0
C -CB-CB	85	119.2
C -CB-NB	70	130.0
C -CD-CD	85	120.0
C -CH-C2	63	111.1
C -CH-C3	63	111.1
C -CH-CH	63	111.1
C -CH-N	63	110.1
C -CH-NT	80	109.7
C -CJ-CJ	85	120.7
C -CM-C3	85	119.7
C -CM-CJ	85	120.7
C -CM-CM	85	120.7
C -CM-CT	70	119.7
C -CM-HC	35	119.7
C -CT-CT	63	111.1
C -CT-HC	35	109.5
C -CT-N	63	110.1

Angle	K_θ	θ_{eq}	Angle	K_θ	θ_{eq}	Angle	K_θ	θ_{eq}	Angle	K_θ	θ_{eq}
C -N -C2	50	121.9	C3-SH-HS	44	96.0	CH-C2-OS	80	109.5	CB-C*-CG	85	106.4
C -N -C3	50	121.9	C3-SH-LP	600	96.7	CH-C2-S	50	114.7	CB-C*-CT	70	126.6
C -N -CH	50	121.9	CA-C -CA	85	120.0	CH-C2-SH	50	108.6	CB-C*-CW	85	106.4
C -N -CT	50	121.9	CA-C -OH	70	120.0	CH-CH-CH	63	111.5	CB-C*-HC	35	126.8
C -N -H	35	119.8	CA-C2-CH	63	114.0	CH-CH-N	80	109.7	CB-CA-HC	35	120.0
C -N -H2	35	120.0	CA-CA-CA	85	120.0	CH-CH-N*	80	109.5	CB-CA-N2	70	123.5
C -N*-CH	70	117.6	CA-CA-CB	85	120.0	CH-CH-NT	80	109.7	CB-CA-NC	70	117.3
C -N*-CJ	70	121.6	CA-CA-CN	85	120.0	CH-CH-OH	80	109.5	CB-CB-N*	70	106.2
C -N*-CM	70	121.6	CA-CA-CT	70	120.0	CH-CH-OS	80	109.5	CB-CB-NB	70	110.4
C -N*-CT	70	117.6	CA-CA-HC	35	120.0	CH-N -H	38	118.4	CB-CB-NC	70	127.7
C -N*-H	35	119.2	CA-CB-CB	85	117.3	CH-N*-CJ	70	121.2	CB-CD-CD	85	120.0
C -NA-C	70	126.4	CA-CB-CN	85	116.2	CH-N*-CK	70	128.8	CB-CN-CD	85	122.7
C -NA-CA	70	125.2	CA-CB-NB	70	132.4	CH-NT-H2	35	109.5	CB-CN-NA	70	104.4
C -NA-H	35	116.8	CA-CD-CD	85	120.0	CH-OH-HO	55	108.5	CB-N*-CE	70	105.4
C -NC-CA	70	120.5	CA-CJ-CJ	85	117.0	CH-OS-CH	100	111.8	CB-N*-CH	70	125.8
C -OH-HO	35	113.0	CA-CM-CM	85	117.0	CH-OS-HO	55	108.5	CB-N*-CK	70	105.4
C*-C2-CH	63	115.6	CA-CM-HC	35	123.3	CH-OS-P	100	120.5	CB-N*-CT	70	125.8
C*-CB-CA	85	134.9	CA-CN-CB	85	122.7	CJ-C -NA	70	114.1	CB-N*-H	35	127.3
C*-CB-CD	85	134.9	CA-CN-NA	70	132.8	CJ-C -O	80	125.3	CB-NB-CE	70	103.8
C*-CB-CN	85	108.8	CA-CT-CT	63	114.0	CJ-CA-N2	70	120.1	CB-NB-CK	70	103.8
C*-CG-NA	70	108.7	CA-CT-HC	35	109.5	CJ-CA-NC	70	121.5	CB-NC-CI	70	111.0
C*-CT-HC	35	109.5	CA-N2-CT	50	123.2	CJ-CJ-N*	70	121.2	CB-NC-CQ	70	111.0
C*-CW-HC	35	120.0	CA-N2-H	35	120.0	CJ-CM-CT	85	119.7	CC-C2-CH	63	113.1
C*-CW-NA	70	108.7	CA-N2-H2	35	120.0	CJ-N*-CT	70	121.2	CC-CF-NB	70	109.9
C2-C -N	70	116.6	CA-N2-H3	35	120.0	CJ-N*-H	35	119.2	CC-CG-NA	70	105.9
C2-C -O	80	120.4	CA-NA-H	35	118.0	CK-N*-CT	70	128.8	CC-CT-CT	63	113.1
C2-C -O2	70	117.0	CA-NC-CB	70	112.2	CM-C -NA	70	114.1	CC-CT-HC	35	109.5
C2-C*-CB	70	126.6	CA-NC-CI	70	118.6	CM-C -O	80	125.3	CC-CV-HC	35	120.0
C2-C*-CG	70	125.0	CA-NC-CQ	70	118.6	CM-CA-N2	70	120.1	CC-CV-NB	70	109.9
C2-C*-CW	70	125.0	CB-C -NA	70	111.3	CM-CA-NC	70	121.5	CC-CW-HC	35	120.0
C2-C2-C2	63	112.4	CB-C -O	80	128.8	CM-CJ-N*	70	121.2	CC-CW-NA	70	105.9
C2-C2-CH	63	112.4	CB-C*-CG	85	106.4	CM-CM-CT	70	119.7	CC-NA-CP	70	107.3
C2-C2-N	80	111.2	CB-C*-CT	70	126.6	CM-CM-HC	35	119.7	CC-NA-CR	70	107.3
C2-C2-N2	80	111.2	CB-C*-CW	85	106.4	CM-CM-N*	70	121.2	CC-NA-H	35	126.3
C2-C2-N3	80	111.2	CB-C*-HC	35	126.8	CM-CT-HC	35	109.5	CC-NB-CP	70	105.3
C2-C2-NT	80	111.2	CB-CA-HC	35	120.0	CM-N*-CT	70	121.2	CC-NB-CR	70	105.3
C2-C2-OS	80	109.5	CB-CA-N2	70	123.5	CM-N*-H	35	119.2	CD-C -CD	85	120.0
C2-C2-S	50	114.7	CB-CA-NC	70	117.3	CN-CA-HC	35	120.0	CD-C -OH	70	120.0
C2-CA-CA	70	120.0	CB-CB-N*	70	106.2	CN-NA-CW	70	111.6	CD-CA-CD	85	120.0
C2-CA-CD	70	120.0	CB-CB-NB	70	110.4	CN-NA-H	35	123.1	CD-CB-CN	85	116.2
C2-CC-CF	70	131.9	CB-CB-NC	70	127.7	CP-NA-H	35	126.3	CD-CD-CD	85	120.0
C2-CC-CG	70	129.0	CB-CD-CD	85	120.0	CR-NA-CW	70	107.3	CD-CD-CN	85	120.0
C2-CC-CV	70	131.9	CB-CN-CD	85	122.7	CR-NA-H	35	126.3	CD-CN-NA	70	132.8
C2-CC-CW	70	129.0	CB-CN-NA	70	104.4	CR-NB-CV	70	105.3	CE-N*-CH	70	128.8
C2-CC-NA	70	122.2	CB-N*-CE	70	105.4	CT-C -N	70	116.6	CE-N*-CT	70	128.8
C2-CC-NB	70	121.0	CB-N*-CH	70	125.8	CT-C -O	80	120.4	CE-N*-H	35	127.3
C2-CH-C3	63	111.5	CB-N*-CK	70	105.4	CT-C -O2	70	117.0	CF-CC-NA	70	105.9
C2-CH-CH	63	111.5	CB-N*-CT	70	125.8	CT-C*-CW	70	125.0	CF-NB-CP	70	105.3
C2-CH-N	80	109.7	CB-N*-H	35	127.3	CT-CC-CV	70	131.9	CF-NB-CR	70	105.3
C2-CH-N*	80	109.5	CB-NB-CE	70	103.8	CT-CC-CW	70	129.0	CG-CC-NA	70	108.7
C2-CH-NT	80	109.7	CB-NB-CK	70	103.8	CT-CC-NA	70	122.2	CG-CC-NB	70	109.9
C2-CH-OH	80	109.5	CB-NC-CI	70	111.0	CT-CC-NB	70	121.0	CG-NA-CN	70	111.6
C2-CH-OS	80	109.5	CB-NC-CQ	70	111.0	CT-CT-CT	40	109.5	CG-NA-CP	70	107.3
C2-N -CH	50	118.0	CC-C2-CH	63	113.1	CT-CT-HC	35	109.5	CG-NA-CR	70	107.3
C2-N -H	38	118.4	CC-CF-NB	70	109.9	CT-CT-N	80	109.7	CG-NA-H	35	126.3
C2-N2-CA	50	123.2	CC-CG-NA	70	105.9	CT-CT-N*	50	109.5	CH-C -N	70	116.6
C2-N2-H2	35	118.4	CC-CT-CT	63	113.1	CT-CT-N2	80	111.2	CH-C -O	80	120.4
C2-N2-H3	35	118.4	CC-CT-HC	35	109.5	CT-CT-N3	80	111.2	CH-C -O2	65	117.0
C2-N3-H3	35	109.5	CC-CV-HC	35	120.0	CT-CT-OH	50	109.5	CH-C -OH	70	115.0
C2-NT-H2	35	109.5	CC-CV-NB	70	109.9	CT-CT-OS	50	109.5	CH-C2-CH	63	112.4
C2-OH-HO	55	108.5	CC-CW-HC	35	120.0	CT-CT-S	50	114.7	CH-C2-OH	80	109.5
C2-OS-C2	100	111.8	CC-CW-NA	70	105.9	CT-CT-SH	50	108.6	CH-C2-OS	80	109.5
C2-OS-C3	100	111.8	CC-NA-CP	70	107.3	CT-N -CT	50	118.0	CH-C2-S	50	114.7
C2-OS-HO	55	108.5	CC-NA-CR	70	107.3	CT-N -H	38	118.4	CH-C2-SH	50	108.6
C2-OS-P	100	120.5	CC-NA-H	35	126.3	CT-N2-H3	35	118.4	CH-CH-CH	63	111.5
C2-S -C3	62	98.9	CC-NB-CP	70	105.3	CT-N3-H3	35	109.5	CH-CH-N	80	109.7
C2-S -LP	600	96.7	CC-NB-CR	70	105.3	CT-OH-HO	55	108.5	CH-CH-N*	80	109.5
C2-S -S	68	103.7	CD-C -CD	85	120.0	CT-OS-CT	60	109.5	CH-CH-NT	80	109.7
C2-SH-HS	44	96.0	CD-C -OH	70	120.0	CT-OS-P	100	120.5	CH-CH-OH	80	109.5
C2-SH-LP	600	96.7	CD-CA-CD	85	120.0	CT-S -CT	62	98.9	CH-CH-OS	80	109.5
C3-C -N	70	116.6	CD-CB-CN	85	116.2	CT-S -LP	600	96.7	CH-N -H	38	118.4
C3-C -O	80	120.4	CD-CD-CD	85	120.0	CT-S -S	68	103.7	CH-N*-CJ	70	121.2
C3-C -O2	70	117.0	CD-CD-CN	85	120.0	CT-SH-HS	44	96.0	CH-N*-CK	70	128.8
C3-C2-CH	63	112.4	CD-CN-NA	70	132.8	CT-SH-LP	600	96.7	CH-NT-H2	35	109.5
C3-C2-OS	80	109.5	CE-N*-CH	70	128.8	CV-CC-NA	70	105.9	CH-OH-HO	55	108.5
C3-CH-C3	63	111.5	CE-N*-CT	70	128.8	CW-C*-HC	35	126.8	CH-OS-CH	100	111.8
C3-CH-CH	63	111.5	CE-N*-H	35	127.3	CW-CC-NA	70	108.7	CH-OS-HO	55	108.5
C3-CH-N	80	109.5	CF-CC-NA	70	105.9	CW-CC-NB	70	109.9	CH-OS-P	100	120.5
C3-CH-NT	80	109.7	CF-NB-CP	70	105.3	CW-NA-H	35	125.3	CJ-C -NA	70	114.1
C3-CH-OH	80	109.5	CF-NB-CR	70	105.3	H -N -H	35	120.0	CJ-C -O	80	125.3
C3-CM-CJ	85	119.7	CG-CC-NA	70	108.7	H2-N2-H2	35	120.0	CJ-CA-N2	70	120.1
C3-N -H	38	118.4	CG-CC-NB	70	109.9	H2-NT-H2	35	109.5	CJ-CA-NC	70	121.5
C3-N*-CB	70	125.8	CG-NA-CN	70	111.6	H3-N -H3	35	120.0	CJ-CJ-N*	70	121.2
C3-N*-CE	70	128.8	CG-NA-CP	70	107.3	H3-N2-H3	35	120.0	CJ-CM-CT	85	119.7
C3-N*-CK	70	128.8	CG-NA-CR	70	107.3	H3-N3-H3	35	109.5	CJ-N*-CT	70	121.2
C3-N2-CA	50	123.2	CG-NA-H	35	126.3	HC-CK-N*	35	123.0	CJ-N*-H	35	119.2
C3-N2-H2	35	118.4	CH-C -N	70	116.6	HC-CK-NB	35	123.0	CK-N*-CT	70	128.8
C3-N3-H3	35	109.5	CH-C -O	80	120.4	HC-CM-N*	35	119.1	CM-C -NA	70	114.1
C3-OH-HO	55	108.5	CH-C -O2	65	117.0	HC-CQ-NC	35	115.4	CM-C -O	80	125.3
C3-OS-P	100	120.5	CH-C -OH	70	115.0	HC-CR-NA	35	120.0	CM-CA-N2	70	120.1
C3-S -LP	600	96.7	CH-C2-CH	63	112.4	HC-CR-NB	35	120.0	CM-CA-NC	70	121.5
C3-S -S	68	103.7	CH-C2-OH	80	109.5	HC-CT-HC	35	109.5	CM-CJ-N*	70	121.2

Angle	K_θ	θ_{eq}
CM-CM-CT	70	119.7
CM-CM-HC	35	119.7
CM-CM-N*	70	121.2
CM-CT-HC	35	109.5
CM-N*-CT	70	121.2
CM-N*-H	35	119.2
CN-CA-HC	35	120.0
CN-NA-CW	70	111.6
CN-NA-H	35	123.1
CP-NA-H	35	126.3
CR-NA-CW	70	107.3
CR-NA-H	35	126.3
CR-NB-CV	70	105.3
CT-C -N	70	116.6
CT-C -O	87	120.4
CT-C -O2	70	117.0
CT-C*-CW	70	125.0
CT-CC-CV	70	131.9
CT-CC-CW	70	129.0
CT-CC-NA	70	122.2
CT-CC-NB	70	121.0
CT-CT-CT	40	109.5
CT-CT-HC	35	109.5
CT-CT-N	80	109.7
CT-CT-N*	50	109.5
CT-CT-N2	80	111.2
CT-CT-N3	80	111.2
CT-CT-OH	50	109.5
CT-CT-OS	50	109.5
CT-CT-S	50	114.7
CT-CT-SH	50	108.6
CT-N -CT	50	116.0
CT-N -H	38	116.4
CT-N2-H3	35	118.4
CT-N3-H3	35	109.5
CT-OH-HO	55	108.5
CT-OS-CT	60	109.5
CT-OS-P	100	120.5
CT-S -CT	62	98.9
CT-S -LP	600	96.7
CT-S -S	68	103.7
CT-SH-HS	44	96.0
CT-SH-LP	600	96.7
CV-CC-NA	70	105.9
CW-C*-HC	35	126.8
CW-CC-NA	70	106.7
CW-CC-NB	70	109.9
CW-NA-H	35	125.3
H -N -H	35	120.0
H2-N2-H2	35	120.0
H2-NT-H2	35	109.5
H3-N -H3	35	120.0
H3-N2-H3	35	120.0
H3-N3-H3	35	109.5
HC-CK-N*	35	123.0
HC-CK-NB	35	123.0
HC-CM-N*	35	119.1
HC-CQ-NC	35	115.4
HC-CR-NA	35	120.0
HC-CR-NB	35	120.0
HC-CT-HC	36	109.5
HC-CT-N	38	109.5
HC-CT-N*	35	109.5
HC-CT-N2	35	109.5
HC-CT-N3	35	109.5
HC-CT-OH	35	109.5
HC-CT-OS	35	109.5
HC-CT-S	35	109.5
HC-CT-SH	35	109.5
HC-CV-NB	35	120.0
HC-CW-NA	35	120.0
HO-OH-HO	47	104.5
HO-OH-P	45	108.5
HS-SH-HS	35	92.1
HS-SH-LP	600	96.7
LP-S -LP	600	160.0
LP-S -S	600	96.7
LP-SH-LP	600	160.0
N -C -O	80	122.9
N*-C -NA	70	115.4
N*-C -NC	70	118.6
N*-C -O	80	120.9
N*-CB-NC	70	126.2
N*-CE-NB	70	113.9
N*-CH-OS	80	109.5
N*-CK-NB	70	113.9
N*-CT-OS	50	109.5
N2-CA-N2	70	120.0
N2-CA-NA	70	116.0
N2-CA-NC	70	119.3
NA-C -O	80	120.6
NA-CA-NC	70	123.3
NA-CP-NA	70	110.7
NA-CP-NB	70	111.6
NA-CR-NA	70	110.7
NA-CR-NB	70	111.6
NC-C -O	80	122.5
NC-CI-NC	70	129.1
NC-CQ-NC	70	129.1
O2-C -O2	80	126.0
O2-P -O2	140	119.9
O2-P -OH	45	108.2
O2-P -OS	100	108.2
OH-P -OS	45	102.6
OS-P -OS	45	102.6

Torsional Parameters

Torsion	$V_n/2$	γ	n
X -C -C2-X	0.0	180	3
X -C -CA-X	5.3	180	2
X -C -CB-X	4.4	180	2
X -C -CD-X	5.3	180	2
X -C -CH-X	0.0	0	2
X -C -CJ-X	3.1	180	2
X -C -CM-X	3.1	180	2
X -C -CT-X	0.0	0	2
X -C -N -X	10.0	180	2
X -C -N*-X	5.8	180	2
X -C -NA-X	5.4	180	2
X -C -NC-X	8.0	180	2
X -C -OH-X	1.8	180	2
X -C*-C2-X	0.0	0	2
X -C*-CB-X	4.8	180	2
X -C*-CG-X	23.6	180	2
X -C*-CT-X	0.0	0	2
X -C*-CW-X	23.6	180	2
X -C2-C2-X	2.0	0	3
X -C2-CA-X	0.0	0	2
X -C2-CC-X	0.0	0	2
X -C2-CH-X	2.0	0	3
X -C2-N -X	0.0	0	3
X -C2-N2-X	0.0	0	3
X -C2-N3-X	1.4	0	3
X -C2-NT-X	1.0	0	3
X -C2-OH-X	0.5	0	3
X -C2-OS-X	1.45	0	3
X -C2-S -X	1.0	0	3
X -C2-SH-X	0.75	0	3
X -CA-CA-X	5.3	180	2
X -CA-CB-X	10.2	180	2
X -CA-CD-X	5.3	180	2

Torsion	$V_n/2$	γ	n
X -CA-CJ-X	3.7	180	2
X -CA-CM-X	3.7	180	2
X -CA-CN-X	10.6	180	2
X -CA-CT-X	0.0	0	2
X -CA-N2-X	6.8	180	2
X -CA-NA-X	6.0	180	2
X -CA-NC-X	9.6	180	2
X -CB-CB-X	16.3	180	2
X -CB-CN-X	20.0	180	2
X -CB-N*-X	6.6	180	2
X -CB-NB-X	5.1	180	2
X -CC-CF-X	14.3	180	2
X -CC-CG-X	15.9	180	2
X -CC-CT-X	0.0	0	2
X -CC-CV-X	14.3	180	2
X -CC-CW-X	15.9	180	2
X -CC-NA-X	5.6	180	2
X -CC-NB-X	4.8	180	2
X -CD-CD-X	5.3	180	2
X -CD-CN-X	5.3	180	2
X -CE-N*-X	6.7	180	2
X -CE-NB-X	20.0	180	2
X -CF-NB-X	4.8	180	2
X -CG-NA-X	6.0	180	2
X -CH-CH-X	2.0	0	3
X -CH-N -X	0.0	0	3
X -CH-N*-X	0.0	0	2
X -CH-NT-X	1.0	0	3
X -CH-OH-X	0.5	0	3
X -CH-OS-X	1.45	0	3
X -CI-NC-X	13.5	180	2
X -CJ-CJ-X	24.4	180	2
X -CJ-CM-X	24.4	180	2
X -CJ-N*-X	7.4	180	2
X -CK-N*-X	6.7	180	2
X -CK-NB-X	20.0	180	2
X -CM-CM-X	24.4	180	2
X -CM-CT-X	0.0	0	3
X -CM-N*-X	7.4	180	2
X -CN-NA-X	12.2	180	2
X -CP-NA-X	9.3	180	2
X -CP-NB-X	10.0	180	2
X -CQ-NC-X	13.5	180	2
X -CR-NA-X	9.3	180	2
X -CR-NB-X	10.0	180	2
X -CT-CT-X	1.3	0	3
X -CT-N -X	0.0	0	3
X -CT-N*-X	0.0	0	2
X -CT-N2-X	0.0	0	3
X -CT-N3-X	1.4	0	3
X -CT-OH-X	0.5	0	3
X -CT-OS-X	1.15	0	3
X -CT-S -X	1.0	0	3
X -CT-SH-X	0.75	0	3
X -CV-NB-X	4.8	180	2
X -CW-NA-X	6.0	180	2
X -OH-P -X	0.75	0	3
X -OS-P -X	0.75	0	3
O -C -C2-N	0.200	180	3
O -C -CH-C2	0.100	180	3
O -C -CH-N	0.100	180	3
O -C -CH-CH	0.100	180	3
OS-C2-C2-OH	2.000	0	3
OS-C2-C2-OH	0.500	0	2
OH-C2-C2-OH	2.000	0	3
OS-C2-C2-OS	2.000	0	3
OS-C2-C2-OS	0.500	0	2
OH-C2-C2-OH	0.500	0	2
OS-C2-CH-OS	0.500	0	2
OS-C2-CH-OH	0.500	0	2
OH-C2-CH-OH	0.500	0	2
OH-C2-CH-OH	1.000	0	3
OS-C2-CH-OS	1.000	0	3
OS-C2-CH-OH	1.000	0	3
C2-C2-S -LP	0.000	0	3
CH-C2-SH-LP	0.000	0	3
OS-CH-C2-OH	0.500	0	2
OS-CH-C2-OH	1.000	0	3
OH-CH-CH-OH	0.500	0	2
OS-CH-CH-OH	0.500	0	3
OH-CH-CH-OH	0.500	0	3
OS-CH-CH-OH	0.500	0	2
OS-CH-CH-OS	0.500	0	3
OS-CH-CH-OS	0.500	0	2
HC-CM-CM-CT	1.710	180	2
C -CM-CM-HC	6.590	180	2
N*-CM-CM-CT	6.590	180	2
CA-CM-CM-HC	6.590	180	2
N*-CM-CM-CA	9.510	180	2
HC-CM-CM-HC	1.710	180	2
N*-CM-CM-C	9.510	180	2
N*-CM-CM-HC	6.590	180	2

Torsional Parameters

Torsion	$V_n/2$	γ	n
N -CT-C -O	0.067	180	3
HC-CT-C -O	0.067	180	3
CT-CT-C -O	0.067	180	3
CT-OS-CT-CT	0.200	180	2
CT-OS-CT-CT	0.383	0	3
OS-CT-CT-OS	0.144	0	3
OS-CT-CT-OH	0.500	0	2
OH-CT-CT-OH	0.144	0	3
OS-CT-CT-OH	0.144	0	3
OS-CT-CT-OS	0.500	0	2
OH-CT-CT-OH	0.500	0	2
H -N -C -O	0.650	0	1
H -N -C -O	2.500	180	2
C2-OS-C2-C3	0.100	0	2
C2-OS-C2-C2	0.100	0	2
C3-OS-C2-C3	0.100	0	2
C3-OS-C2-C3	1.450	0	3
C2-OS-C2-C3	0.725	0	3
C2-OS-C2-C2	1.450	0	3
CH-OS-CH-C2	0.725	0	3
CH-OS-CH-CH	0.100	0	2
CH-OS-CH-CH	0.725	0	3
C2-OS-CH-C2	0.100	0	2
C3-OS-CH-C3	0.725	0	3
CH-OS-CH-N*	0.725	0	3
C3-OS-CH-C3	0.100	0	2
C2-OS-CH-C3	0.100	0	2
C2-OS-CH-C2	0.725	0	3
CH-OS-CH-C2	0.100	0	2
CH-OS-CH-N*	0.000	0	2
C2-OS-CH-C3	0.725	0	3
OH-P -OS-C3	0.750	0	2
OS-P -OS-C2	0.250	0	3
OS-P -OS-C2	0.750	0	2
OH-P -OS-C2	0.750	0	2
OS-P -OS-CT	0.250	0	3
OS-P -OS-CH	0.750	0	2
OS-P -OS-C3	0.750	0	2
OH-P -OS-C2	0.250	0	3
OS-P -OS-CH	0.250	0	3
OH-P -OS-CH	0.250	0	3
OH-P -OS-CH	0.750	0	2
OH-P -OS-CT	0.750	0	2
OH-P -OS-CT	0.250	0	3
OS-P -OS-CT	0.750	0	2
OH-P -OS-C3	0.250	0	3
OS-P -OS-C3	0.250	0	3
LP-S -S -LP	0.000	0	3
LP-S -S -C2	0.000	0	3
C2-S -S -C2	0.600	0	3
CT-S -S -CT	0.600	0	3
LP-S -S -CT	0.000	0	3
CT-S -S -CT	3.500	0	2
C2-S -S -C2	3.500	0	2

Improper Torsional Parameters

Torsion	$V_n/2$	γ	n
C2-CH-C -N3	7.0	180	3
C3-CH-CA-C3	7.0	180	3
C3-CH-NT-C	14.0	180	3
CH-CH-C -N3	7.0	180	3
H2-CH-N2-H2	0.0	180	3
X -C2-CH-X	14.0	180	3
X -CH-CH-X	14.0	180	3
X -CH-N -C	14.0	180	3
X -CH-N -C2	1.0	180	2
X -CT-N -CT	1.0	180	2
X -H2-N -H2	1.0	180	2
X -N2-CA-N2	10.5	180	2
X -O2-C -O2	10.5	180	2
X -X -C -O	10.5	180	2
X -X -CA-HC	2.0	180	2
X -X -N -H	1.0	180	2
X -X -N2-H3	1.0	180	2
X -X -NA-H	1.0	180	2

Non-Bonded Parameters

Atom	R*	ε
C	1.85	0.120
C*	1.85	0.120
C2	1.92	0.120
C3	2.00	0.150
CA	1.85	0.120
CB	1.85	0.120
CC	1.85	0.120
CD	1.85	0.120
CE	1.85	0.120
CF	1.85	0.120
CG	1.85	0.120
CH	1.85	0.090
CI	1.85	0.120
CJ	1.85	0.120
CK	1.85	0.120
CM	1.85	0.120
CN	1.85	0.120
CP	1.85	0.120
CQ	1.85	0.120
CR	1.85	0.120
CT	1.80	0.060
CV	1.85	0.120
CW	1.85	0.120
H	1.00	0.020
H2	1.00	0.020
H3	1.00	0.020
HC	1.54	0.010
HO	1.00	0.020
HS	1.00	0.020
LP	1.20	0.016
N	1.75	0.160
N*	1.75	0.160
N2	1.75	0.160
N3	1.85	0.080
NA	1.75	0.160
NB	1.75	0.160
NC	1.75	0.160
NP	1.75	0.160
NT	1.85	0.120
O	1.60	0.200
O2	1.60	0.200
OH	1.65	0.150
OS	1.65	0.150
P	2.10	0.200
S	2.00	0.200
SH	2.00	0.200

Hydrogen Bond Parameters

Acceptor	Donor	C	D
H	NB	7557	2385
H	NC	10238	3071
H	O2	4019	1409
H	O	7557	2385
H	OH	7557	2385
H	S	265720	35029
H	SH	265720	35029
HO	NB	7557	2385
HO	NC	7557	2385
HO	O2	4019	1409
HO	O	7557	2385
HO	OH	7557	2385
HO	S	265720	35029
HO	SH	265720	35029
H2	NB	4019	1409
H2	NC	4019	1409
H2	O2	4019	1409
H2	O	10238	3071
H2	OH	4019	1409
H2	S	265720	35029
H2	SH	265720	35029
H3	NB	4019	1409
H3	NC	4019	1409
H3	O2	4019	1409
H3	O	7557	2385
H3	OH	7557	2385
H3	S	265720	35029
H3	SH	265720	35029
HS	NB	14184	3082
HS	NC	14184	3082
HS	O2	14184	3082
HS	O	14184	3082
HS	OH	14184	3082
HS	S	265720	35029
HS	SH	265720	35029

References

1. S. J. Weiner, P. A. Kollman, D. A. Case, U. C. Singh, C. Ghio, G. Alagona, S. Profeta, and P. Weiner, *J. Amer. Chem. Soc.*, **106,** 765 (1984).
2. T. Lybrand and P. Kollman, "A molecular mechanical study of ethidium bromide interactions with base-paired dinucleosides and hexanucleosides," Biopolymers (in press).
3. U. C. Singh and P. Kollman, *J. Comp. Chem.*, **5,** 129 (1984).

4. N. Allinger, *J. Amer. Chem. Soc.*, **99,** 8127 (1977).
5. A. Hagler, E. Euler, and S. Lifson, *J. Amer. Chem. Soc.*, **96,** 5319 (1974).
6. See review by P. Payne and L. C. Allen in *Modern Theoretical Chemistry: Applications of Electronic Structure Theory,* H. F. Schaefer, Ed., Plenum, New York, 1977, Chap. 2.
7. A. Verma, W. Murphy, and H. Bernstein, *J. Chem. Phys.*, **69,** 1540 (1974); K. Kuchitsu, *Bull. Chem. Soc. Jpn.*, **32,** 7481 (1959); K. Raghavachari, *J. Chem. Phys.*, **81,** 1383 (1984).
8. P. Kasai and R. Myers, *J. Phys. Soc. Jpn.*, **30,** 1096 (1959).
9. T. Kitayama and T. Miyazawa, *Bull. Chem. Soc. Jpn.*, **41,** 1976 (1968).
10. S. Profeta, unpublished MM2 Calculations on methyl ethyl ether.
11. G. Engelsholm, A. Luntz, W. Gwinn, and D. Harris, *J. Chem. Phys.*, **50,** 2446 (1969).
12. D. Cremer and J. Pople, *J. Amer. Chem. Soc.*, **97,** 1354 (1975).
13. A. Almenningen, H. Seip, and A. Walladsen, *Acta Chim. Scand.*, **23,** 2748 (1969).
14. H. Geise, W. Adams, and L. Bartell, *Tetrahedron,* **25,** 3045 (1969).
15. W. Jorgensen and M. Ibrahim, *J. Amer. Chem. Soc.*, **103,** 3976 (1981).
16. D. Davies, *Prog. Nucl. Mag. Res. Spectros.*, **12,** 135 (1978).
17. C. Altona and M. Sundaralingham, *J. Amer. Chem. Soc.*, **94,** 8205 (1972).
18. C. J. Cerjan and W. H. Miller, *J. Chem. Phys.*, **75,** 2800 (1981); J. Simons, P. Jorgensen, H. Taylor, and J. Ozment, *J. Phys. Chem.*, **87,** 2745 (1983); D. T. Nguyen and D. A. Case, *J. Phys. Chem.*, **89,** 4020 (1985).
19. J. Langlet, P. Claverie, and F. Caron, *Intermolecular Forces,* B. Pullman, Ed., 14th Jerusalem Symposium, Reidel, Dordrecht (Holland), 1981,
20. I. Yanson, A. Teplisky, and L. Sukhodub, *Biopolymers,* **18,** 1149 (1979).
21. D. Williams and T. Starr, Comp. and Chem., **1,** 173 (1977).
22. K. C. Janda, J. Hemminger, J. Winn, S. Novick, S. Harrison and W. Kemperer, J. Chem. Phys., **63,** 1419 (1973).
23. H. Frauenfelder, G. Petsko, and D. Tsernoglu, *Nature,* **280,** 558 (1979).
24. G. Dodson, E. Dodson, D. Hodgkin, and C. Reynolds, *Can. J. Biochem.*, **57,** 469 (1979).
25. W. Jorgensen, *J. Amer. Chem. Soc.*, **103,** 335 (1981).
26. N. Szczesniak, M. J. Nowak, H. Rostkowska, K. Szczepaniak, W. B. Person, and D. Shugar, *J. Am. Chem. Soc.*, **105,** 5959 (1983).
27. C. P. Beetz, Jr., and G. Ascarelli, *Spect. Acta.*, **36A,** 299 (1980); see also H. Susi and J. S. Ard. *Spect. Acta.*, **27A,** (1971).
28. Y. Nishimura, M. Tsuboi, S. Kato, and K. Morokuma, *J. Am. Chem. Soc.*, **103,** 1354 (1981).
29. M. Tsuboi and Y. Nishimura in *Raman Spectroscopy, Linear and Nonlinear,* J. Lascombe and V. Huong, Eds., Wiley, Chichester, U.K., 1982, p. 683.
30. B. Brooks and M. Karplus, *Proc. Natl. Acad. Sci. USA* **80,** 6571 (1983).
31. D. T. Nguyen, Ph.D. Thesis, University of California, Davis, 1986.
32. Y. Nishimura, H. Harayama, K. Nomura, A. Y. Hirakawa, and M. Tsuboi, *Bull. Chem. Soc. Jpn.*, **52,** 1340 (1979).
33. N. K. Sanyal, S. L. Srivastava, and R. K. Goel, *Indian J. Phys.* **52B,** 108 (1977).
34. H. Susi, J. S. Ard, and J. M. Purcell, *Spectrochim. Acta.*, **29A,** 725 (1973).
35. B. F. Putnam and L. L. VanZandt, *J. Comp. Chem.*, **3,** 305 (1982).
36. A. Lautie and A. Novak, *J. Chim. Phys.*, **71,** 415 (1971).
37. Y. Nishimura, M. Tsuboi, S. Kato, and K. Morokuma in ref. 29, p. 703.
38. J.-M. Delabar and M. Majoube, *Spectrochim. Acta,* **34A,** 129 (1978).
39. J.-M. Delabar, *J. Raman Spect.*, **7,** 261 (1978).
40. C. Perchaud, A.-M. Bellocq, and A. Novak, *J. Chim. Phys.*, **62,** 1344 (1965).
41. B. M. Craven, R. K. McMullan, J. D. Bell, and H. C. Freeman, *Acta Cryst.*, **B33,** 2585 (1977).
42. A. Lautie, M. F. Lautie, A. Gruger, and S. A. Fakhri, *Spectrochim.*, **36A,** 85 (1980).
43. F. Momamy, R. McGuire, A. Burgess, and H. Scheraga, *J. Phys. Chem.*, **75,** 2361 (1975).
44. See e.g., B. Gelin and M. Karplus, *Biochem.*, **18,** 1256 (1979); for information about the Harvard group all-H force field see B. R. Brooks, R. E. Brucoleri, B. D. Olafson, D. J. States, S. Swaminathan, and M. Karplus, *J. Comput. Chem.*, **4,** 187 (1983).
45. M. Levitt, *J. Mol. Biol.*, **166,** 595, 617, 621 (1983).
46. J. Hermans, D. Ferro, J. McQueen and S. Wei, in *Environmental Effects on Molecular Structure and Properties,* B. Pullman, Ed., Reidel, Dordrecht, Holland, 1976.
47. V. Sasisekharan, in *Conformation of Biological Molecules and Polymers,* E. Bergmann and B. Pullman, Ed., Jerusalem Press, Jerusalem 1973.
48. W. Olson and P. Flory, *Biopolymers,* **11,** 25 (1972).
49. M. Levitt, *Cold Spring Harbor Symp. Quant. Biol.*, **47,** 271 (1983).
50. N. Pavitt and D. Hall, *J. Comp. Chem.*, **5,** 441 (1984).
51. S. J. Weiner, U. C. Singh and P. A. Kollman, Simulation of formamide hydrolysis by hydroxide ion in the gas phase and in aqueous solution, *J. Amer. Chem. Soc.*, **107,** 2219 (1985).
52. S. J. Weiner, Ph.D. thesis, U. C. San Francisco, Dec., 1984.
53. G. Seibel, U. C. Singh and P. Kollman, A molecular dynamics simulation of double helical B DNA including counterions and water, *Proc. Nat. Acad. Sci.*, **82,** 6537 (1985).