Role of Electrostatics in the Structure, Energy, and Dynamics of Biomolecules: A Model Study of *N*-Methylalanylacetamide†

B. Montgomery Pettitt‡ and Martin Karplus*

Contribution from the Department of Chemistry, Harvard University, Cambridge, Massachusetts 02138. *Received July 3, 1984*

Abstract: The contribution of electrostatic interactions to a range of structural, energetic, and dynamic properties of the alanine dipeptide is examined. An empirical energy function is employed to represent the dipeptide, and the partial atomic charges are varied from zero to values used in standard models for amino acids and proteins. It is demonstrated that, although there are large differences in the absolute energy of models with different charges, the relative energies of the various dipeptide conformers are less sensitive to the charges. The minimum energy structures of the conformers are only weakly dependent on the charges. A normal mode analysis shows that only a few modes are sensitive to the charges and that thermodynamic quantities, which represent sums over the modes, are essentially the same for all the models. Comparison of the harmonic dynamics results with those from an ensemble of molecular dynamics trajectories reveals that the electrostatic contributions to the potential surface introduce anharmonic effects.

I. Introduction

Empirical energy functions are now in widespread use for the study of the properties and interactions of molecules of biological interest.[1-5] Since most such molecules contain both polar and charged groups, the electrostatic contribution to the energy is expected to play an important role. Its proper treatment has been a basic concern of those attempting to develop empirical energy functions. Most of the existing models for electrostatic interactions employ the monopole approximation in which point charges are located on the atoms with their values determined by a mixture of theoretical and empirical criteria.[1-5] In a few cases higher moments[6] as well as polarization effects[7-9] have been included in the potential function. Within the monopole approximation, a wide range of values have been used for the charges,[1-5] including the limit of neglecting charges for calculations on systems as polar as proteins and nuclei acids.[10] In the other extreme, there are models in which all nonbonded interactions are represented by point charges.[11,12] A variety of approaches to dielectric shielding have been also introduced. They include use of a dielectric constant (e.g., values of 1-10 for the protein interior), a distant-dependent dielectric function,[13,14] as well as more detailed representations of polarization and solvent effects.[9,15-17]

In view of the wide range of empirical approaches to the electrostatic contributions to the potential energy, it is important to delineate their role in determining molecular properties of interest. They include the energy of a molecule and the interaction energy between molecules, the structures and energies of different conformational minima, and the vibrations and dynamics in the neighborhood of the minima as well as the magnitudes of barriers between them. Since the various properties depend in different ways on the energy and its derivatives, it is expected that their sensitivity to the electrostatics model will vary; e.g., structural and motional properties might be expected to be less sensitive than the energy itself. Once the properties of isolated molecules are understood, it will be desirable to study the system in solution. Clearly the effect of solvation, particularly in polar media, will vary with respect to the different electrostatic models. We reserve a discussion of the influence of partial charge distributions on the condensed-phase behavior for a future study.[17]

In this paper we examine the effects of the electrostatic interactions by applying a series of models to *N*-methylalanylacetamide, the alanine "dipeptide". This system is chosen for study because it is a small molecule that includes the backbone degrees of freedom (ϕ, ψ, ω) and polar groups (C=O and N—H) of proteins in a normal environment. Also, there are three methyl groups which give some information on "side-chain" behavior. The monopole approximation is used and the molecule is treated in the all-atom representation with a well-studied empirical potential[18,19] that is kept invariant except for the choice of partial atomic charges. These cover a range from zero (no charges) to the values commonly used in protein simulations.[1-5]

Section II outlines the methodology. The results are presented and discussed in Section III. Section IV summarizes the implications of this study for empirical energy function calculations on macromolecules.

II. Methodology

In this section we introduce the potential energy function and then outline the methods used for energy minimization, normal mode analysis, and molecular dynamics.

(a) Model. The empirical energy function chosen for the present study has the form used in a previous theoretical treatment of the dipeptide.[18,19] The total potential energy V consists of terms related to bond lengths,

† Supported in part by grants from the National Science Foundation and the National Institutes of Health.

‡ NIH Research Service Award Postdoctoral Fellow, 1983.

(1) Gibson, K. D.; Scheraga, H. A. *Proc. Natl. Acad. Sci. U.S.A.* **1967**, *58*, 420. Dunfield, L.; Burgess, A.; Scheraga, H. *J. Phys. Chem.* **1978**, *82*, 2609.
(2) Hagler, A.; Huber E.; Lifson, S. *J. Am. Chem. Soc.* **1974**, *96*, 5319.
(3) Levitt, N. *J. Mol. Biol.* **1983**, *168*, 595.
(4) Brooks, B.; Bruccoleri, R.; Olafson, B.; States, D.; Swaminathan, S.; Karplus, N. *J. Comput. Chem.* **1983**, *1*, 187.
(5) Weiner, S. P.; Kollman, P.; Case, D.; Singh, U.; Ghio, C.; Alagoma, G.; Profeta, S.; Weiner, P. *J. Am. Chem. Soc.* **1984**, *106*, 765.
(6) Perahia, D.; Pullman, A. *Theor. Chim. Acta* **1978**, *48*, 263. Pullman, A.; Zakrewska, C.; Perahia, D. *Int. J. Quantum Chem.* **1979**, *16*, 393.
(7) Claverie, P.; Rein, R. *Int. J. Quantum Chem.* **1969**, *3*, 537. Claverie, P. *J. Mol. Biol.* **1971**, *56*, 75.
(8) Barnes, P.; Finney, J. L.;f Nicholas, J. D.; Quinn, J. E. *Nature (London)* **1979**, *282*, 459.
(9) Warshel, A.; Levitt, M. *J. Mol. Biol.* **1976**, *103*, 227.
(10) Levitt, M. *Cold Spring Harbor Symp. Quant. Biol.* **1982**, *67*, 251.
(11) Nemenoff, R.; Snir, J.; Scheraga, H. *J. Phys. Chem.* **1978**, *82*, 2504.
(12) Marchese, F.; Mehrotra, P.; Beveridge, D. *J. Phys. Chem.* **1982**, *86*, 2592.
(13) Gelin, B. R.; Karplus, N. *Proc. Natl. Acad. Sic. U.S.A.* **1975**, *72*, 2002.
(14) Gelin, B. R. Ph.D. Thesis, Harvard University, Cambridge, MA, 1976.
(15) Matthew, J. B.; Friend, S. H.; Gurd, F. R. W. *Biochemistry* **1980**, *20*, 571.
(16) States, D.; Karplus, M. unpublished results.
(17) Pettitt, B. M.; Karplus, M.; Rossky, P. J., unpublished results.
(18) Rossky, P. J. Karplus, M.f Rahman, A. *Biopolymers* **1979**, *18*, 825.
(19) Rossky, P. J.; Karplus, M. *J. Am. Chem. Soc.* **1979**, *101*, 1931.

0002-7863/85/1507-1166$01.50/0

Table I. Charge Parameters (in Units of the Electronic Charge)

model	C_L	H_{L1}	H_{L2}	H_{L3}	C'_L	O_L	N_L	H_L	C_α	H_α
1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
2	−0.092	0.031	0.028	0.025	0.433	−0.428	−0.274	0.183	0.070	0.008
3	−0.048	0.0246	0.0246	0.0246	0.433	−0.428	−0.274	0.183	0.070	0.0246
4	−0.300	0.100	0.100	0.100	0.480	−0.480	−0.360	0.260	0.100	0.100

model	C_β	$H_{\beta1}$	$H_{\beta2}$	$H_{\beta3}$	C'_R	O_R	N_R	H_R	C_R	H_{R1}	H_{R2}	H_{R3}
1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
2	−0.020	0.004	0.009	0.016	0.433	−0.428	−0.274	0.183	0.093	−0.007	0.008	−0.001
3	−0.048	0.0246	0.0246	0.0246	0.433	−0.428	−0.274	0.183	−0.048	0.0246	0.0246	0.0246
4	−0.300	0.100	0.100	0.100	0.480	−0.480	−0.360	0.260	−0.300	0.100	0.100	0.100

Table II. Electronic Structure Estimates of Charge for *N*-Methylacetamide[a,b]

model	C_L	H_1	H_2	H_3	N	H	C′	O	C_R	H_1	H_2	H_3
STO3g	−0.052	0.071	0.074	0.074	−0.380	0.192	0.297	−0.292	−0.197	0.079	0.067	0.067
CNDO/2	0.105	−0.005	−0.001	−0.001	−0.222	0.113	0.349	−0.337	−0.088	0.030	0.029	0.029

[a] Charges determined from the indicated calculation (see text) by a Mullikan popultion analysis. [b] Geometry used given in Kitano et al.; see ref 20.

bond angles, dihedral angles, hydrogen bonds, and nonbonded interactions including van der Waals and Coulombic contributions. It has the form

$$V = \sum_{\substack{1,2\\ \text{pairs}}} K_b(b - b_0)^2 + \sum_{\substack{\text{bond}\\ \text{angles}}} K_\theta(\theta - \theta_0)^2 + \sum_{\substack{1,3\\ \text{pairs}}} F(\rho - \rho_{\min})^2 + \sum_{\substack{\text{dihedral}\\ \text{angles}}} K_\phi[1 + \cos(n\phi - \delta)] + \sum_{\substack{\text{H bonded}\\ k,l \text{ pairs}}} \left[\frac{A_{kl}}{s^{12}} - \frac{B_{kl}}{s^{10}}\right] + \sum_{\substack{\text{nonbond}\\ i,j \text{ pairs}}} \left[4\epsilon_{ij}\left(\left[\frac{\sigma_{ij}}{r_{ij}}\right]^{12} - \left[\frac{\sigma_{ij}}{r_{ij}}\right]^{6}\right) + \frac{q_i q_j}{r_{ij}}\right] \quad (1)$$

where b, K_b, and b_0 are the bond length, bond-stretch force constant, and isolated equilibrium distance, θ, K_θ, and θ_0 are the bond angle, bond-angle force constant, and isolated equilibrium angle, ρ, F, and $\rho_{\min}$ are the 1,3 distance, Urey-Bradley force constant, and effective equilibrium value (see Appendix A), ϕ, K_ϕ, n, and δ are the dihedral angle, its force constant, multiplicity, and phase, s is the hydrogen bond length, A_{kl} and B_{kl} define the shape and depth of the hydrogen bond potential, and r_{ij}, ϵ_{ij}, σ_{ij}, and q_i and q_j are the nonbonded distance, the Lennard-Jones well depth and diameter, and the charges for atom pairs i and j. The list of parameter values has been given elsewhere.[18,19] All interactions are included without any distance truncation, as is appropriate for such a small molecule. In the present paper, we use these parameters as listed in ref 18 and 19, except for the charges (q_i), which are varied.

The atomic charges employed are listed in Table I; see Figure 1 for the atom definitions. To establish an extreme for comparison, model 1 has all charges set to zero. Model 2 is the charge set from the earlier dipeptide study, which was derived from a population analysis of the electronic structure given by a CNDO calculation. The third set of charges is closely related to model 2; it differs only in that identical charges are used for the three methyl groups. Model 4 uses a set of charges which have been proposed on the basis of dipole moment and other data for proteins.[4] The magnitudes of the charges in this model are larger than for the other cases considered; still larger charges appear in some empirical potentials.[5,11,12]

The investigation of model 3 and its comparison with model 2 was prompted by several considerations. Although the differences in the charges are small (because the individual C and H partial charges are themselves small), there is a qualitative change between the two in that for model 3, the methyl groups all have the same C–H polarity whereas they differ in model 2 (i.e., compare the rightmost and leftmost methyl groups in Table I). It seemed of interest to determine whether removal of this small asymmetry has significant effect on the behavior of the dipeptide. This is of some importance since quantum mechanical calculations for specific assumed geometries are often used in determining the atomic charges. In fact, the charges used in model 2 were evaluated from a population analysis of a CNDO calculation for one nuclear configuration of the dipeptide. That the methyl group polarity should be the same in different environments (as in models 3 and 4) appears reasonable on the grounds of the electronegativity arguments and is substantiated by ab initio calculations performed on amino acid analogues. Calculations for the experimental geometry[20] of *N*-methylacetamide (see Figure 1b) yielded ab initio charges at the STO-3G level given in Table II; they are compared to a CNDO/2 calculation at the same geometry. The calculations support the dipole direction used for the methyl groups in models 3 and 4. Apparently, the polarity of the N–H adjacent methyl in CNDO calculations of N-methylacetamide and N-methylalanylacetamide is anomalous, although there is some difference between the ab initio charge on the carbons of the two methyl groups. That the hydrogen charges of a given methyl group are not equal in these calculations (see Table II) is a consequence of using a particular geometry.

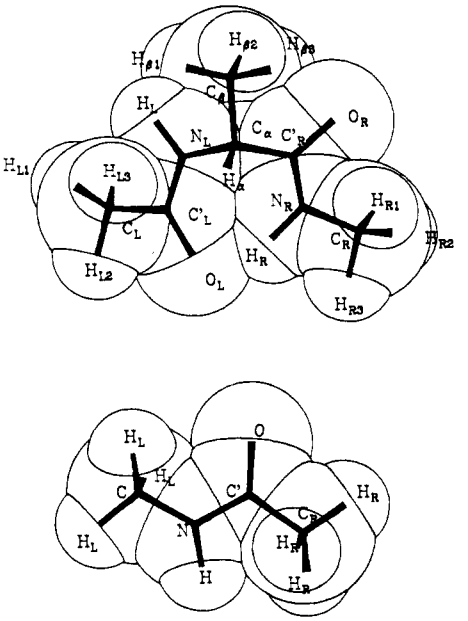

Figure 1. Structure of the C_{7eq} conformation of *N*-methylalanylacetamide. The axial and equitorial designations refer to the orientation of the β methyl with respect to the seven-membered hydrogen-bonded ring. Below is the structure of *N*-methylacetamide (see text). The dihedral angles of interest for this work are defined as ω for O_L–C'_L–N_L–H_L, ϕ for C'_L–N_L–C_α–C'_R, and ψ for N_L–C_α–C'_R–N_R.

(b) Energy Minimization and Vibrational Analysis. The optimized geometry and its associated energy for each model were obtained for three local minima (C_{7eq}, C_{7ax}, C_5) on the dipeptide (ϕ and ψ) potential energy surface. All degrees of freedom were included in the minimization. In each case a combination of 200 steps of ABNR[4,22] and 2 steps of Newton–Raphson minimizations were performed; the resulting root mean square (rms) gradient per atom was less than 0.01 kcal/mol Å in the optimized configurations. The second derivative matrix of the potential at the C_{7eq} minimum was computed and utilized for determining the vibrational frequencies and normal modes by diagonalizing the quadratic projection of the potential energy matrix in the mass-weighted Cartesian space; all degrees of freedom were included. As an overall measure of the variation in the vibrations of the models (say n and m),

(20) Kitano, M.; Fukuyama, T.; Kutchitsu, K. *Bull. Chem. Soc. Jpn.* **1973**, *46*, 384.

(21) Reiher, W., private communication.

(22) States, D.; Karplus, M.; Unpublished results.

Table III. Major Minima on the ϕ and ψ Surface[a]

model	ω	ϕ	ψ	rms_ω[b]	rms_ϕ[b]	rms_ψ[b]	E	ΔE[c]	E_e	E_{hb}
1										
C_{7eq}	177.8	-69.2	65.1	10.4	9.1	13.4	15.6	0.0	0.0	-3.0
C_{7ax}	-179.0	64.9	-63.5	10.6	8.6	11.6	15.7	0.1	0.0	-3.0
C_5	-179.2	-174.5	174.3	10.5	9.8	14.2	17.4	1.8	0.0	-1.6
2										
C_{7eq}	-179.1	-66.7	65.0	10.2	8.1	10.4	-3.9	0.0	-19.7	-3.0
C_{7ax}	-179.8	63.4	-61.8	10.2	7.8	10.1	-3.9	0.0	-19.8	-3.0
C_5	-179.5	-174.5	175.1	10.4	9.6	12.9	0.6	4.5	-16.9	-1.5
3										
C_{7eq}	178.3	-67.6	64.2	10.1	7.9	10.7	7.8	0.0	-8.0	-3.0
C_{7ax}	-179.0	64.1	-61.1	10.2	7.6	10.3	7.8	0.0	-8.1	-3.0
C_5	-179.6	-174.5	174.9	10.3	9.6	12.7	11.5	3.7	-6.0	-1.6
4										
C_{7eq}	177.2	-67.9	61.3	10.1	7.5	10.7	4.9	0.0	-11.6	-3.0
C_{7ax}	-179.0	64.0	-59.6	10.2	7.2	10.0	5.2	0.3	-11.4	-2.9
C_5	-179.4	-172.8	176.6	10.3	9.4	12.3	9.7	4.8	-8.4	-1.8

[a] All angles in degrees and energies in kcal/mol. [b] Root mean square harmonic fluctuations at 300 K. [c] Energy relative to absolute minimum for a given model.

the frequency difference of the ith normal mode, $\Delta_{n,m}{}^i$ was computed and the average magnitude of the differences, $\delta_{n,m}$, was obtained from

$$\delta_{n,m} = \frac{1}{N}\sum_{i=1}^{N}|\Delta_{n,m}{}^i| \tag{2}$$

From the frequencies, the harmonic vibrational contribution to the thermodynamic variables of state were calculated by use of the appropriate partition function and its derivatives; the expressions used for the Helmholz free energy, A, the enthalpy, H, the heat capacity at constant volume, C_v, and the entropy, S, are[23] (without zero-point contributions)

$$-\frac{A}{NkT} = \sum_i \ln\,[1 - \exp(-\theta_i/T)]^{-1}$$

$$\frac{H}{NkT} = \sum_i \frac{\theta_i/T}{[\exp(\theta_i/T) - 1]}$$

$$\frac{C_v}{Nk} = \sum_i \frac{(\theta_i/T)^2 \exp(\theta_i/T)}{[\exp(\theta_i/T) - 1]^2}$$

$$S = -(A - H)/T \tag{3}$$

where N is Avogadro's number, $\theta_i = h\nu_i/k$, and ν_i is the frequency of the ith mode. In the limit of low-frequency (or equivalently high temperature), the reduced enthalpy and heat capacity approach a constant whereas the entropy and free energy diverge logarithmically.

(c) Molecular Dynamics Simulation. Molecular dynamics simulations[24,25] were performed for the C_{7eq} conformation of each of the models. The Verlet algorithm was used to integrate the equations of motion.[26] Since the primary concern here is with the behavior of the lowest modes dominated by the dihedral angles, the equations of motion were simplified by performing the simulations with rigid bonds; the SHAKE algorithm[27] was used to apply the constraints. Because the system is relatively harmonic and energy relaxation is slow, a series of trajectories with different initial energy distributions in the internal degrees of freedom was used. For each model, 14 trajectories were run for 16 ps with a time step of 1 fs. Each trajectory was started from configuration near the C_{7eq} minimum. A set of velocities appropriate for 300 K was selected from a Boltzmann distribution by use of random numbers. The first half of the trajectory (8 ps) served as an equilibration period, and the second half (8 ps) was used for analysis; since the lowest frequency motion in the harmonic model corresponds to a period of ~0.7 ps, the trajectory length should be adequate.

III. Results and Discussion

In this section we consider the various properties of the alanine dipeptide and their dependence on the magnitude of the electrostatic interactions.

(a) Static Properties, Energies, and Geometries. In Table III, we present the energies, geometries, and harmonic fluctuations obtained from each of the models for the optimized structures of the two C_7 configurations (C_{7eq} and C_{7ax}) and for the extended conformer (C_5). Only the peptide dihedral angles are listed, because the bond lengths are fixed and the bond angles are negligibly affected by the electrostatic interactions. It is evident that the conformations of the three minima are very similar in all the models, although the largest charges (model 4) do lead to significant changes in the values of ϕ and ψ; e.g., in comparing the zero charge case (model 1) with model 4, we see that ψ, in particular, has decreased by nearly 4° in both the C_{7eq} and C_{7ax} geometry.

Focusing on the energies, we see that there is a large variation in the absolute values. Although these are not directly meaningful, the Coulomb contribution would enter as corrections in any attempt to estimate the heats of formation for such polar systems,[28] analogous to the empirical treatments available for hydrocarbons and related molecules.[29] More germane to the modeling of biomolecules, the electrostatic interactions play an essential role in the thermodynamics of substrate binding, in heats of sublimation of crystals and in heats of solutions. Thus, the variation of nearly 20 kcal for equivalent geometries between the uncharged system and the other models is of considerable importance.

Of concern also is the large difference in energies for models 2 and 3 whose charges vary only in a very minor way and whose geometries are almost identical. If we examine the seven-membered ring structures, C_{7ax} and C_{7eq}, formed by an intramolecular hydrogen bond, the opposite polarity of the methyl groups produces an attractive interaction between the two ends of the molecule. This attraction is not simply between the methyl moieties, although this is a part of the effect. It is largely due to the interactions of the right methyl with the highly charged hydrogen bonding atoms on the opposite side of the molecule. When these attractions are removed, as in model 3, by making the methyls electrostatically equivalent, the energy change is significant. This electrostatic background is why the C_5 energies also differ in models 2 and 3.

In contrast to the absolute energies, the energy differences among conformers (C_{7ax} vs. C_{7eq} vs. C_5) are relatively similar for the four models. In all cases, C_{7eq} and C_{7ax} are close in energy with the largest difference between the two occurring in model 4 that has the highest partial charges. The C_5 conformation is always significantly less stable. However, for model 1 (zero charge) the difference between the C_7 conformers and C_5 is considerably smaller than for the three other cases. This is due to a lack of H_L, C'_R and N_L, O_R repulsions within the C_5 ring structure.

Table III also lists separately the electrostatic and hydrogen bond contributions to the energy for each case. As expected from

(23) McQuarrie, D. "Statistical Mechanics"; Harper and Row: New York, 1976.
(24) Alder, B.; Wainwright, T. E. *J. Chem. Phys.* **1957**, *27*, 1208.
(25) Rahman, A. *Phys. Rev.* **1964**, *136*, A405.
(26) Verlet, L. *Phys. Rev.* **1967**, *159*, 98.
(27) Ryckaert, J. P.; Cicotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327.
(28) Tse, Y. C.; Newton, M. D.; Vishreshuara, S.; Pople, J. A. *J. Am. Chem. Soc.* **1978**, *100*, 4329.
(29) Allinger, N. L.; Tribble, M. T.; Miller, M. A.; Wertz, D. H. *J. Am. Chem. Soc.* **1971**, *93*, 1637. Allinger, N. L.; Tribble, M. T.; Miller, M. A. *Tetrahedron* **1972**, *28*, 1173.

Table IV. Normal Mode Frequencies for Different Models[a,b]

	model						
I	1	2	3	4	$\Delta_{1,2}$	$\Delta_{2,3}$	$\Delta_{2,4}$
7	41.0	57.3	56.4	59.2	−16.2	0.9	−1.9
8	71.3	72.2	72.0	71.6	−0.9	0.2	0.5
9	116.7	119.5	121.1	128.3	−2.8	−1.6	−8.9
10	127.6	130.2	131.0	138.0	−2.6	−0.8	−7.8
11	162.8	162.9	162.4	157.7	−0.1	0.5	5.2
12	179.0	179.4	178.9	178.2	−0.4	0.5	1.2
13	218.4	222.6	222.1	222.1	−4.1	0.4	0.5
14	233.9	237.0	237.6	239.1	−3.1	−0.6	−2.1
15	253.1	255.5	255.4	258.8	−2.4	0.1	−3.3
16	300.0	302.6	302.5	301.5	−2.6	0.1	1.1
17	324.0	331.3	331.5	337.2	−7.3	−0.1	−5.9
18	363.3	363.3	363.1	360.1	0.0	0.1	3.2
19	440.6	441.5	441.6	439.3	−1.0	−0.1	2.2
20	471.3	473.0	473.4	476.4	−1.7	−0.4	−3.4
21	554.9	546.3	546.3	542.9	8.6	0.0	3.4
22	607.7	606.4	605.7	600.2	1.3	0.7	6.2
23	646.1	649.8	648.5	644.3	−3.7	1.2	5.5
24	677.0	675.9	676.4	675.9	1.1	−0.5	0.0
25	732.1	732.3	731.9	729.2	−0.3	0.4	3.2
26	790.6	829.1	816.6	802.8	−38.5	12.5	26.3
27	894.4	897.7	896.8	897.5	−3.3	0.8	0.1
28	963.5	963.7	963.6	961.2	−0.3	0.2	2.5
29	974.2	976.2	976.0	977.4	−2.0	0.2	−1.2
30	1006.0	1004.7	1004.8	1002.5	1.3	0.0	2.2
31	1034.3	1032.9	1032.9	1026.4	1.3	0.0	6.6
32	1060.0	1059.9	1059.9	1061.5	0.1	0.0	−1.6
33	1064.8	1064.5	1064.6	1065.2	0.3	−0.1	−0.6
34	1074.1	1074.8	1074.1	1073.1	−0.6	0.6	1.7
35	1094.6	1093.1	1094.6	1098.6	1.5	−1.5	−5.4
36	1139.9	1142.1	1141.8	1144.2	−2.2	0.3	−2.1
37	1177.8	1184.4	1184.2	1187.4	−6.6	0.2	−3.0
38	1292.8	1292.6	1293.7	1291.4	0.2	−1.2	1.1
39	1303.3	1300.1	1300.1	1297.8	3.2	0.0	2.3
40	1327.8	1346.8	1345.5	1354.3	−19.0	1.3	−7.4
41	1368.0	1380.6	1379.8	1384.3	−12.6	0.8	−3.7
42	1410.7	1407.8	1407.9	1396.5	3.0	−0.1	11.3
43	1415.4	1414.7	1414.7	1414.3	0.7	0.0	0.3
44	1417.7	1418.7	1420.0	1428.7	−1.0	−1.3	−10.0
45	1458.5	1458.4	1458.5	1459.5	0.1	−0.1	−1.1
46	1461.1	1461.8	1461.7	1465.2	−0.8	0.1	−3.3
47	1467.5	1468.0	1468.2	1466.9	−0.5	−0.2	1.1
48	1469.6	1469.8	1469.5	1469.6	−0.3	0.3	0.2
49	1475.0	1474.9	1474.7	1472.2	0.2	0.2	2.7
50	1476.4	1476.1	1475.5	1473.5	0.2	0.6	2.6
51	1543.0	1540.8	1541.0	1536.3	2.2	−0.2	4.5
52	1588.7	1618.5	1618.1	1634.8	−29.8	0.4	−16.3
53	1644.2	1642.9	1643.6	1644.0	1.3	−0.6	−1.0
54	1650.8	1652.1	1650.9	1654.9	−1.3	1.2	−2.8
55	2878.7	2878.2	2878.3	2876.8	0.4	0.0	1.5
56	2887.2	2887.4	2887.6	2889.2	−0.2	−0.2	−1.8
57	2891.2	2891.1	2891.2	2891.4	0.1	−0.1	−0.3
58	2959.5	2959.5	2959.5	2959.6	0.0	0.0	−0.2
59	2960.7	2960.5	2960.6	2959.9	0.2	0.0	0.7
60	2968.8	2968.8	2968.8	2968.8	−0.1	0.1	0.0
61	2968.9	2968.9	2969.1	2970.0	−0.1	−0.2	−1.1
62	2971.1	2971.0	2971.1	2970.5	0.1	0.0	0.5
63	2972.3	2972.3	2972.4	2973.1	0.0	−0.1	−0.8
64	3033.7	3030.8	3031.1	3031.0	2.9	−0.3	−0.2
65	3323.4	3322.7	3322.7	3323.2	0.8	0.0	−0.5
66	3422.4	3457.4	3458.5	3498.8	−35.0	−1.1	−41.4

[a] All values in wavenumbers (cm^{-1}). The six lowest translational and rotational modes are not included. [b] For definition of $\Delta_{n,m}$, see text.

the total energy results, the electrostatic contributions are large in all cases except in the zero-charge model but are relatively insensitive to the differences in conformation due to their long-range character and the resulting cancellations for relatively small changes in geometry. From the variation of the remainder of the energy for each conformer, it is clear that the small structural differences at the minima for the different models do have a small but significant effect on the energy difference between the two C_7 and the C_5 conformers.

It is important to note the presence of a nonelectrostatic hydrogen bond term in the energy function (eq 1); for models 2–4, the hydrogen bond has approximately equal contributions from the explicit term and from the Coulomb interactions. This reduces the effect of the electrostatic contribution on the structure and energy, relative to models in which hydrogen bonds are treated as purely electrostatic.[2]

(b) Harmonic Dynamics. To determine the effect of the electrostatic contribution on the harmonic dynamics, we consider the normal mode results for the lowest energy (C_{7eq}) conformation. Table IV gives the complete set of normal mode frequencies (excluding translation and rotation) for the four models. Although the energies were found to vary by as much as 15 kcal, the normal mode frequency differences are rather small. The effects of the electrostatic terms can be seen most easily by looking at the Δ

Table V. Normal Mode Assignments for Modes Sensitive to Electrostatics

mode	assignment
7	ϕ, ψ dihedral combination
26	N_R–H_R wag
40	C'_R–N_R, C'_R–C_α asymmetric stretch combination
41	H_α–C_α–N_L and χ_β combination
42	H–C_L–H symmetric combination with C'_L–C_L–H and H_L–N_L–C'_L
44	H–C_R–H symmetric combination with N_R–C_R–H
52	C_R–N_R–H_R angle
66	N_R–H_R stretch

values listed in the last three columns of the table. For most of the modes, the differences between any pair of models is less than 5 cm^{-1}; in fact, for a large number of modes, the variation is less than 1 cm^{-1}. An analysis of the eigenvectors projected on the internal coordinates indicates that the ordering of the modes is the same for all the models, with only a few exceptions, and that the relative contributions of the significant internal coordinates are generally very similar. Models 1–3 have identical state ordering, and model 4 is different in only two instances. Normal modes 47 and 48 (composed of χ_β motion coupled to either C_α–C_β–H_β or H_β–C_β–H_β) are mixed and 49 and 50 (composed of improper χ_β torsion (rocking) coupled to N–C_R–H_R or H_R–C_R–H_R) are reversed. In both cases, the pair of modes involved have similar frequencies.

As example of a mode that is only slightly perturbed is the amide I band which corresponds to modes 53 and 54 in the models; it consists mainly of the C=O stretch. The calculated values cover the range 1643 ± 2 cm^{-1} for mode 53 and 1652 ± 2 cm^{-1} for mode 54. The amide I band has been identified for the dipeptide in solution under basic and acidic conditions. The range of observed frequencies in these environments is from 1632 to 1668 cm^{-1}, respectively.[30]

A small number of modes show a larger dependence on the electrostatic contribution to the potential. Those that vary by more than 10 cm^{-1} are listed in Table V in terms of the frequencies and the dominant internal coordinate contributions calculated with model 4. The essential factor in the frequency change of most of these modes is the variation in the electrostatic component of the hydrogen bond. In model 1, the effective hydrogen bond energy is approximately half or less that of the other models; thus, model 1 differs more from the other models than they do among each other. Modes involving the hydrogen bond that appear in Table V are mode 1 (the lowest frequency mode dominated by the ϕ and ψ dihedral angles that is constrained by the internal hydrogen bond), mode 26 (the out-of-plane N_R–H_R wag), mode 40 (the C'_R–N_R, C'_R–C_α asymmetric stretch combination), mode 52 (the H_R–N_R–C_R angle), and mode 66 (the N_R–H_R stretch). It is of interest that mode 26 is the only one that shows a large difference between models 2 and 3. Mode 41 is composed of H_α–C_α–N_L and χ_β motion and shifts relative to model 1 due to an accumulation of angle strain in the charged models. Interestingly Δ_{24} also goes down for mode 66 due to larger charges in the H-bonded ring.

The changes in frequency between models 2 and 4 arise more from the change in magnitude of the charges (Table I) than the symmetry of the methyl groups. If symmetry were responsible there would have been a more drastic change in the Δ_{23} column of Table IV. The largest difference in frequencies occurs for modes 52 and 66 both of which involve the N_R–H_R group, which has significantly larger charges in model 4 than 2. The observed differences in modes 42 and 44 for models 2 and 4 are due in part to the sizable change in the dipole moment of the methyl groups. Both of these modes are a mixture of H–C–H bend with neighboring angles.

To characterize the overall behavior of the normal mode frequencies by means of eq 2, we list the various $\delta_{n,m}$ values in Table VI in matrix form. It is clear that models 2 and 3 are very similar

(30) Avignou, A.; von Huang, P. *Biopolymers* **1970**, *9*, 427.

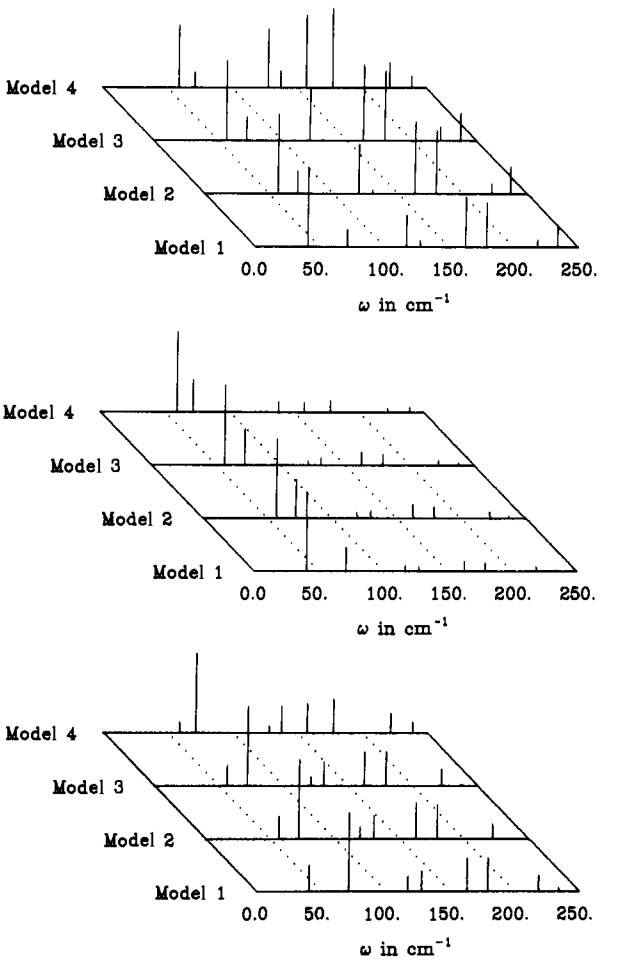

Figure 2. Harmonic power spectra of internal coordinates for the models: values are normalized to the maximum peak height: (top) dihedral angle ϕ, (middle) dihedral angle ψ, (bottom) left dihedral angle ω (see Figure 1).

Table VI. Mean Absolute Harmonic Deviations, δ, between Models[a]

	model			
model	1	2	3	4
1	0.0	3.9	3.6	6.6
2		0.0	0.6	4.0
3			0.0	3.7
4				0.0

[a] See eq 2.

on the average while models 1 and 4 differ more from each other and from both 2 and 3. For the 2nd and 4th models, the large average absolute difference arises in part from the change in the amide hydrogen stretching frequency (30 cm^{-1}, see Table IV).

The reduction of the charges and concomitant reduction in the hydrogen bond energy of model 1 leads to increased flexibility when compared with models 2 and 3. All of the ten lowest modes of model 1 are below those of model 2. For models 2 and 3, the results are very similar, although seven of the ten lowest modes of model 3 are below model 2. Model 4 has five of its lowest modes below those of 2 and five above.

As expected from the similarity of the normal modes, the contributions of the normal modes to the frequency spectrum of a given internal coordinate shows relatively little variation from model to model. Figure 2 presents the results for the dihedral angles ϕ, ψ, and ω. For all the dihedral angles, the intensity patterns are very similar for the four models. Both ϕ and ψ have large contributions from the lowest frequency mode (mode 7). Since mode 7 varies significantly in frequency in the different models, as is evident in the figures, some effect on the internal coordinate dynamics results. This is particularly true for ψ; for ϕ some of the higher frequency modes (e.g., ~162 and 179 cm^{-1})

Table VII. Harmonic Thermodynamic Properties[a]

case	T	A	S	H	C	zero point
			model 1			
C_{7eq}	100.0	−0.3871	0.0099	0.5993	0.0134	114.06
	200.0	−2.0274	0.0228	2.5233	0.0247	114.06
	300.0	−4.9072	0.0348	5.5184	0.0353	114.06
C_{7ax}	100.0	−0.3737	0.0097	0.5974	0.0133	114.19
	200.0	−1.9956	0.0225	2.5091	0.0245	114.19
	300.0	−4.8480	0.0344	5.4847	0.0352	114.19
C_5	100.0	−0.4430	0.0113	0.6919	0.0147	113.25
	200.0	−2.2753	0.0250	2.7150	0.0254	113.25
	300.0	−5.3887	0.0372	5.7708	0.0359	113.25
			model 2			
C_{7eq}	100.0	−0.3355	0.0091	0.5734	0.0132	114.30
	200.0	−1.8922	0.0219	2.4832	0.0245	114.30
	300.0	−4.6817	0.0338	5.4662	0.0352	114.30
C_{7ax}	100.0	−0.3309	0.0090	0.5741	0.0132	114.41
	200.0	−1.8813	0.0218	2.4740	0.0244	114.41
	300.0	−4.6571	0.0337	5.4407	0.0351	114.41
C_5	100.0	−0.4241	0.0111	0.6819	0.0147	113.23
	200.0	−2.2264	0.0247	2.7038	0.0254	113.23
	300.0	−5.3101	0.0369	5.7610	0.0359	113.23
			model 3			
C_{7eq}	100.0	−0.3367	0.0091	0.5734	0.0132	114.28
	200.0	−1.8946	0.0219	2.4836	0.0246	114.28
	300.0	−4.6859	0.0338	5.4685	0.0352	114.28
C_{7ax}	100.0	−0.3311	0.0090	0.5723	0.0131	114.40
	200.0	−1.8792	0.0217	2.4708	0.0244	114.40
	300.0	−4.6525	0.0336	5.4386	0.0351	114.40
C_5	100.0	−0.4200	0.0110	0.6801	0.0147	113.27
	200.0	−2.2163	0.0246	2.7010	0.0254	113.27
	300.0	−5.2930	0.0368	5.7558	0.0359	113.27
			model 4			
C_{7eq}	100.0	−0.3247	0.0089	0.5632	0.0131	114.36
	200.0	−1.8582	0.0216	2.4695	0.0246	114.36
	300.0	−4.6247	0.0336	5.4568	0.0353	114.36
C_{7ax}	100.0	−0.3129	0.0087	0.5548	0.0130	114.52
	200.0	−1.8203	0.0213	2.4433	0.0243	114.52
	300.0	−4.5498	0.0332	5.4086	0.0351	114.52
C_5	100.0	−0.4059	0.0108	0.6692	0.0146	113.31
	200.0	−2.1757	0.0243	2.6873	0.0254	113.31
	300.0	−5.2250	0.0366	5.7407	0.0358	113.31

[a] Calculated from eq 3; values in units of kcal/mol and kcal/mol/K. Zero point is the zero-point correction = $h\omega/2$ summed over all modes.

that are less effected by the model also make large contributions. The angle ω differs from both ϕ and ψ in that it is dominated by the second mode at 72 cm^{-1}; this mode is insensitive to the electrostatic potential.

Thermodynamic Properties. From the normal mode spectrum, the quantum-mechanical vibrational contribution to the thermodynamic properties can be evaluated within the harmonic approximation by use of eq 3. The results obtained for the four models are shown as a function of temperature in Figure 3, and values for selected temperatures are listed in Table VII. On the scale of the figure, the results are nearly indistinguishable, as expected from the similarity of the normal mode spectrum. The insert in the figure shows an expand version of the vibrational entropy calculation in the range 295–300 K from which the slightly higher entropy of model 1 is evident. This is due primarily to the considerably lower frequency of the first mode (see Table IV). More generally, since the low modes make the major contribution to the room temperature entropy, differences near 300 K are a measure of the flexibility in the harmonic approximation. The slight change in flexibility between models 2–4 is evident in the absolute entropy ordering. The heat capacity and enthalpy are not sensitive to these frequency changes, and there is correspondingly very little difference among the models for these quantities. This analysis is valid for the other minima as is evident from Table VII. However, additionally it seems that the rather shallow high-energy C_5 conformer exhibits the characteristics of a structure with considerably more flexibility. This is most easily seen in the entropy contributions. Also, it should be noted that the zero-point contribution to the enthalpy and free energy difference between C_5 and the other conformers is not negligible, though it is rather insensitive to the model.

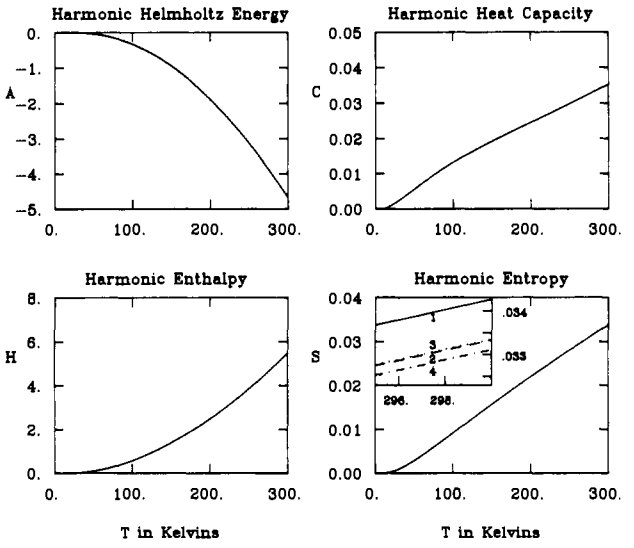

Figure 3. Vibrational contribution to thermodynamic properties in the harmonic approximation (see text).

(c) Molecular Dynamics. Molecular dynamics provides a more complete evaluation of the electrostatic effects on the internal motions of the dipeptide than is obtained from the harmonic

Table VIII. Molecular Dynamics Statistics for Charged Models in C_{7eq} Conformation Averaged over 14 Trajectories of 8 ps Each[a]

model	$\langle\omega\rangle$	rms_ω	$\langle\phi\rangle$	rms_ϕ	$\langle\psi\rangle$	rms_ψ	$\langle T\rangle$
2	179.4	10.6	−67.4	8.0	63.3	12.1	301.2
3	178.6	11.4	−68.2	8.4	62.9	12.3	301.1
4	177.7	11.4	−68.4	7.9	59.1	12.1	303.0

[a] All angles in degrees and energies in kcal/mol. Temperature and energy averaged over the entire 16 ps.

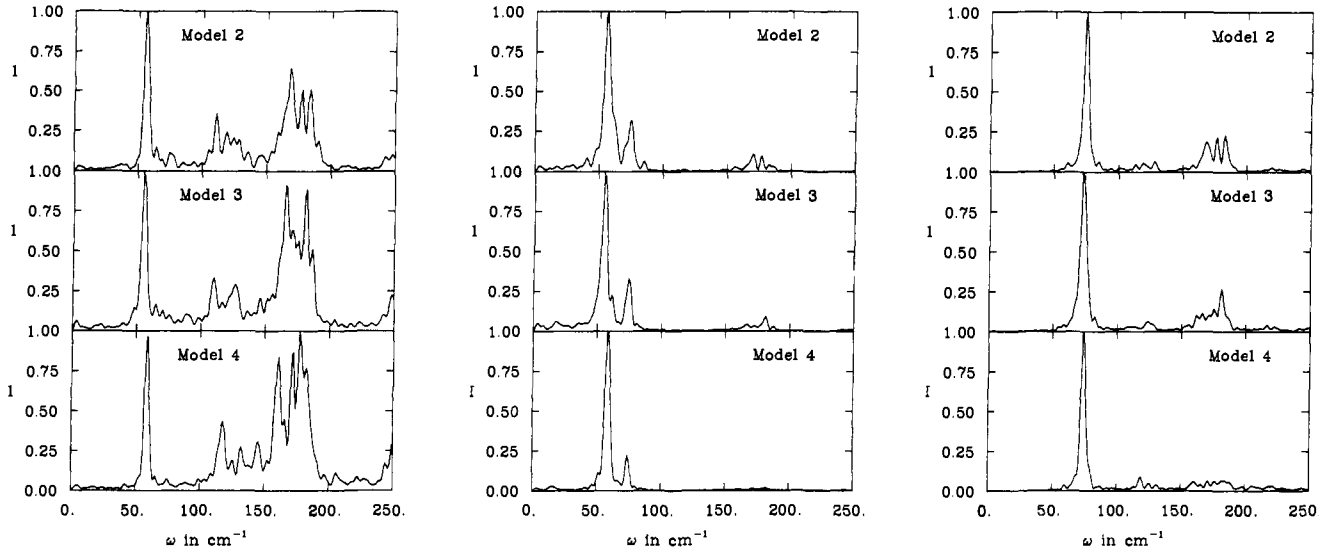

Figure 4. Molecular dynamics power spectrum of inernal coordinates for models 2–4; values are normalized to the maximum peak height: (left) dihedral angle ϕ, (middle) dihedral angle ψ, (right) left dihedral angle ω (see Figure 1).

approximation. The differences between the harmonic and full molecular dynamics results reflect the anharmonic attributes of the potential surface. Since the electrostatic contribution to the energy introduces anharmonic coupling terms, a comparison of the molecular dynamics results for the various models is of particular interest.

Table VIII shows some of the averages from the dynamics for the C_{7eq} configuration; they can be compared with the harmonic model results in Table III. We tabulate the average values and root mean square fluctuations from the trajectories for the dihedral angles ϕ, ψ, and ω. The average structures are very similar to those obtained by minimization (see Table III); the differences are less than 1°, except for the angle ψ in model 4. The pattern of relative $\langle\psi\rangle$ values (model $2 > 3 > 4$) found in the minimization is preserved in the dynamics. As to the rms fluctuations, the results from the dynamics and normal mode analysis are quite similar. The only significant difference is again in the dihedral angle ψ which shows slightly larger flucutations in the dynamics.

To provide a more sensitive test of anharmonic effects, the power spectra of the dihedral angles ϕ, ψ, and ω were examined. The power spectrum $I(\omega)$ was computed from the expression

$$I(\omega) = |\int e^{i\omega t}(\chi(t) - \langle\chi\rangle)\, dt|^2 \qquad (4)$$

where $\chi(t)$ is the time series for the dihedral angle and ω is the frequency. Figure 4 shows the resulting power spectra. We consider first the left ω dihedral angle (Figure 1) because it has the simplest behavior due to its relatively high force constant. In all cases the spectrum is dominated by a single peak whose frequency is essentially independent of the model ($\sim$73 cm^{-1}). Further, the peak is very close to the second mode ($\sim$72 cm^{-1}), which makes the dominant contribution in the harmonic model (Figure 3c). In the high frequency region, there appears to be a difference between models 2 and 3, which show clear secondary peaks between 150 and 200 cm^{-1}, and model 4, which seems to have a more broadly distributed set of intensities. By contrast the harmonic results show all three charged models to be similar to each other in this region. Thus, the anharmonic effects are most important in model 4 with the largest charges.

For the angle ψ, the power spectrum is again dominated by peaks in the low-frequency region. The largest contribution is from a peak between 50 and 60 cm^{-1} whose exact position is rather sensitive to the model. There is a secondary contribution from the peak that was dominant for the ω dihedral angle. These results correspond closely to the harmonic spectrum. The higher frequency contributions are in all cases rather small for both the molecular dynamics and harmonic calculations.

It is the angle ϕ with the more important higher frequency contribution, in addition to the lowest frequency peak, that shows the largest variations as a function of the electrostatic contribution. In models 2 and 3, there is some trading of intensities amongst the peaks between 150 and 200 cm^{-1}. Also, model 2 has less intensity between 150 and 200 cm^{-1} than models 3 and 4. More generally, model 4 differs considerably from models 2 and 3, which are relatively similar. Comparing with the normal mode results, we see that as for ω and ψ, the major relative intensity features are preserved. However, there are considerable differences in the 100–200-cm^{-1} region. Between 100 and 150 cm^{-1}, the harmonic treatment shows only one important peak in all models while there are several in the dynamics; further these vary considerably from model to model. For the 150–200-cm^{-1} region, there are only two peaks in the normal mode spectrum and a complicated and variable multipeak structure in the dynamics; i.e., the dynamics behavior changes significantly for the different models while the harmonic results are quite similar.

IV. Conclusions

An analysis has been made of the electrostatic contribution to a range of structural, energetic, and dynamic properties of *N*-methylalanylacetamide (alanine "dipeptide"), a model system for the interactions that occur in proteins. A monopole model was used for the electrostatic term, and the atomic partial charges were varied from zero to values used in standard models for empirical energy calculations in macromolecules.

Overall, it is found that, due to the slow variation with distance ($1/R$) of the electrostatic contribution, most structural and dynamic quantities are relatively insensitive to the choice of the model; even for the zero-charge case, the differences from the other models are not very great. However, when the absolute energy is being considered, energy differences as large as 15 kcal/mol are found. Since corresponding terms are important for binding energy calculations (e.g., substrate binding) and for evaluation

of the heats of solution and related quantities, it is clear that for the analysis of such problems, careful calibration of the electrostatic contribution to the energy is required.

On a more detailed level, comparison of three relatives minima (C_{7eq}, C_{7ax}, and C_5) of the potential surface showed that in no case was there a change in the relative ordering of stability of the conformers. Also, the geometry of these minima was nearly unaffected, even for the completely uncharged case; e.g., for the relatively high energy C_5 conformer, there were differences of about 2° in ϕ and ψ from model to model. The harmonic thermodynamics are correspondingly insensitive; only the entropy shows a slight variation which is due to the contributions of the lowest modes. Comparison of the normal modes indicates that most of them depend only weakly on the electrostatic model; the variations generally are less than 2 cm^{-1}. However, there are a number of frequencies, particularly those having a significant contribution from internal coordinates involved in hydrogen bonding, that are sensitive to the electrostatic model. This includes the lowest frequency mode, composed in a large part of ϕ and ψ, which varies by nearly 20 cm^{-1} out of 60 cm^{-1} and the highest frequency mode (the NH stretch), which varies by 40 cm^{-1}.

The dependence of the molecular dynamics results on the models generally follows the normal mode behavior. There are small frequency shifts in the power spectra which correspond to shifts in the normal mode frequencies. However, the changes in the intensity patterns in the power spectra appear not to arise from alterations in the normal modes themselves. Instead, they are a consequence of anharmonicity contributions that are significant at finite temperatures.[31] This is not surprising since charge–charge interactions represent a large part of the anharmonic character of the model. It is particularly for the relatively soft dihedral angle degrees of freedom that the anharmonic effects are important and that the electrostatic model dependence is most significant, though still relatively small.

Extrapolation to macromolecules of biological interest suggests that corresponding behavior will be found in the low-frequency motions that are expected to play the dominant role in protein function. However, because these macromolecules have many degrees of freedom that permit larger scale geometry changes (e.g., motions of side chains in proteins and alteration of the helical repeat distance in DNA), the electrostatic effect on the average geometry is likely to be more important than that for the alanine dipeptide.

We have not addressed the problem of finding suitable charges in this study. One approach is based on accurate quantum-mechanical calculations of the electron distribution for the system in the appropriate environment. The results would then be reduced to the classical point charge or a more refined representation that best characterizes the internal and external multipolar electrostatic potential and field of the system.[32,33] Such results should be supplemented by a more empirical approach making use of properties (e.g., lattice energies[2] and radial distribution functions)[34] that are sensitive to the values of charges.

Acknowledgment. We thank Dr. C. Brooks for helpful discussions and Dr. B. Brooks for computational assistance. We also thank Keith Davies for use of the CHEMGRAF program to produce Figure 1.

Appendix

In this appendix we show how the Urey–Bradley terms of the angle potential (1–3 interactions) can be simplified without loss of relevant information. The Urey–Bradley contribution was given by the functional form in ref 16 and 17,

$$V_{UB}^{\circ}(\rho) = F(\rho - \rho_0)^2 + F'(\rho - \rho_0) \quad \text{(A1)}$$

where F and F' are force field constants, ρ is the 1,3 distance, and ρ_0 is a parameter. Since this functional form represents a shifted parabola, we can substract the constant offset energy without loss of generality. Solving for the minimum, we find

$$\rho_{min} = \frac{-F'}{2F} + \rho_0, \; E_{min} = -\frac{(F'^2)}{4F} \quad \text{(A2)}$$

Consequently, eq A1 can be rewritten as

$$V_{UB}^{\circ}(\rho) = F(\rho - \rho_{min})^2 + E_{min} \quad \text{(A3)}$$

Since the absolute energy of the molecular models is generally not used, we may discard the E_{min} term from eq A3. Use of the quadratic

$$V_{UB}(\rho) = F(\rho - \rho_{min})^2 \quad \text{(A4)}$$

in eq 1 does not effect the equations of motion or the geometry of the potential energy minima. This form also is more convenient for parametrization due to the elimination of a physically uninteresting constant.

Registry No. *N*-Methylalanylacetamide, 19701-83-8.

(31) Dickey, J. M.; Paskin, A. *Phys. Rev.* **1969**, *188*, 1407.

(32) Cox, S. R.; Williams, D. E. *J. Comput. Chem.* **1981**, *2*, 304.
(33) Singh, U. C.; Kollman, P. *J. Comput. Chem.* **1984**, *5*, 129.
(34) Stillinger, F. H.; Rahman, A. *J. Chem. Phys.* **1974**, *60*, 1545.