Energy Functions for Peptides and Proteins. I. Derivation of a Consistent Force Field Including the Hydrogen Bond from Amide Crystals

A. T. Hagler,* E. Huler, and S. Lifson

Contribution from the Chemical Physics Department, Weizmann Institute of Science, Rehovot, Israel. *Received December 27, 1973*

Abstract: An intermolecular force field for amides was derived by a least-squares fitting of the parameters of various trial functions to experimental data. These included six heats of sublimation, 72 unit cell vector components, and three dipole moments taken from the following set of molecules: oxamide, malonamide, succinamide, adipamide, urea, formamide, diketopiperazine, LL-3,6-dimethyl-2,5-piperazinedione, and *N*-methylacetamide. It was found that the essential feature of the amide hydrogen bond is the combination of the relatively large partial charges on the amide group and the short O···H contact distance. The latter is a result of the small repulsion of the amide hydrogen (small van der Waals radius), which the least squares indicated to be negligible. The final force field that emerged included Lennard-Jones type nonbonded interactions along with partial charges placed on the atoms. No explicit hydrogen bond function was found necessary to account for the data. Both 6-9 and 6-12 nonbonded functions were tried, and no significant distinction could be made between them. The resultant parameters of the force field are compared with those commonly used in conformational analysis. In particular the larger "van der Waals" radii found here are discussed in terms of the crystal compressive effect which was previously either ignored or only roughly estimated.

The analytic dependence of the hydrogen bond energy on the geometric relations between the participating atoms has been the subject of numerous studies.[1-4] Various energy sources have been suggested as contributing to the hydrogen bond, such as electrostatic interactions, nonbonded interactions including the attractive dispersion force, the short range repulsive force, and a covalent charge transfer[5] contribution. Unfortunately, there is not yet a consensus in the literature on the relative importance of these contributions and on the analytic representation of the hydrogen bond.[6-13]

The purpose of the present work is to obtain a consistent force field to describe systems containing the N−H···O=C hydrogen bond. Analyzing the available experimental information on amide crystals, one observes that the hydrogen bond geometry varies significantly. This variation is the result of the interplay of all crystal forces, including the hydrogen bond, which must cancel each other at the equilibrium crystal structure. The heat of sublimation of the crystal is similarly related to the work necessary to separate the molecules against all crystal forces, including the hydrogen bond.

In the present study we consider a set of amide crystals, composed only of C=O, NH, and CH_n groups. We have examined various hydrogen bond functions, as well as nonbonded and electrostatic energy parameters, for the atoms in the amide group (*i.e.*, N, H, C′, O). We transferred (and partly reexamined) the alkane force field from a previous study.[14,15] The parameters of the amide potential functions were subjected to a least-squares fitting analysis while the transferred alkane field was kept fixed. In this analysis a best fit was sought to the structures and the lattice energies of the amide crystals, as well as to the dipole moments of some amide containing molecules. Vibrational spectra, not included in the present work, are the subject of further research.

In optimizing the crystal structures a new method was used, which avoided repeated minimization and thus made a convergent optimization computationally feasible.[16]

After testing various possible formulations of the hydrogen bond energy function, a rather simple representation emerged. The least-squares method indicated that the Morse function, which was included to represent specific interactions involved in the hydrogen bond other than nonbonded and electrostatic energies, may be neglected. The data could be equally well fit by just electrostatic and van der Waals interactions of the Lennard-Jones type, while the Morse function added only a negligible and indeterminate contribution. This result is in agreement with conclusions on the nature of weak hydrogen bonds, arrived at by quantum mechanical methods.[12,13,17,18]

(1) E. R. Lippincott and R. Schroeder, *J. Chem. Phys.*, 23, 1099 (1955).
(2) H. A. Scheraga, *Advan. Phys. Org. Chem.*, 6, 103 (1968).
(3) R. F. McGuire, F. A. Momany, and H. A. Scheraga, *J. Phys. Chem.*, 76, 375 (1972).
(4) (a) D. A. Brant, *Macromolecules*, 1, 291 (1968); (b) A. M. Liquori, *Quart. Rev. Biophys.*, 2, 65 (1969).
(5) By covalent charge transfer we refer to an "interchange" of electrons either through sharing (covalent) or partial charge transfer, *i.e.*, bonded interaction, as opposed to usual nonbonded interactions. The bonded interaction is presumed to be much steeper in the attractive region than the nonbonded interaction.
(6) R. Chidambaram, R. Balasubramanian, and G. N. Ramachandran, *Biochem. Biophys. Acta*, 221, 196 (1970).
(7) The nature of the hydrogen bond has been a subject of controversy. Several reviews and books have treated the various aspects of the hydrogen bond.[8-13]
(8) G. C. Pimentel and A. L. McClellan, "The Hydrogen Bond," W. H. Freeman, San Francisco, Calif., 1960.
(9) D. Hadzi, Ed., "Hydrogen Bond," Pergamon Press, New York, N. Y., 1969.
(10) S. Bratoz, *Advan. Quantum Chem.*, 3, 209 (1967).
(11) G. C. Pimentel and A. L. McClellan, *Annu. Rev. Phys. Chem.*, 22, 347 (1971).
(12) P. A. Kollman and L. C. Allen, *Chem. Rev.*, 72, 283 (1972).
(13) A. T. Hagler and S. Lifson in "The Proteins," submitted for publication.
(14) S. Lifson, "Recent Developments in the Consistent Force Field Calculations," 23e Reunion Annuelle, Orleans, France, 1972, 19-22.9.
(15) A. Warshel and S. Lifson, *J. Chem. Phys.*, 53, 582 (1970).
(16) A. T. Hagler and S. Lifson, *Acta Crystallogr., Sect. B*, 30, 619 (1974).
(17) (a) G. N. G. Port and A. Pullman, *FEBS (Fed. Eur. Biochem. Soc.) Lett.*, 31, 70 (1973); (b) J. E. Del Bene, *J. Chem. Phys.*, 58, 3139 (1973); (c) P. Kollman and L. C. Allen, *J. Amer. Chem. Soc.*, 93, 4991 (1971); (d) M. Dreyfus and A. Pullman, *Theor. Chim. Acta*, 19, 20 (1970).
(18) C. A. Coulson and U. Danielson, *Ark. Fys.*, 8, 245 (1954).

Another major result which came out of the least squares is the requirement that the short range repulsion of the amide hydrogen be small.

The calculations were carried out for both 6–9 and 6–12 Lennard-Jones type potentials. Both potentials led to the above mentioned results, and their agreement with experiment was similar although the 6–12 potential was slightly better.

In the following paper[19] the resultant force field is tested by minimization of the energy of ten crystals with respect to all degrees of freedom of the unit cell. The angle and distance dependence of the hydrogen bond energy will be discussed in detail elsewhere.

I. Method

(A) Choice of Force Field. At the beginning of our study, we tried the Lippincott and Schroeder potential as modified by Ooi, *et al.*,[20] and the Stockmayer potential as given by Liquori,[4b] in conjunction with the nonbonded potentials given in these studies. Neither potential fit the crystal data satisfactorily, and, in addition, the former does not merge smoothly with the van der Waals and electrostatic interactions at large distances.[2]

At this stage we decided to try the simplest possible potential for the hydrogen bond, which would include a covalent interaction and could account for possible angular and distance dependences[13] of the energy, and in which the covalent element would merge smoothly to a nonbonded potential as the hydrogen bond is broken. In addition we placed partial charges on all the atoms, as the simplest representation of electrostatic interactions.

All parameters related to the amide group were subjected to a least-squares optimization, in order to determine the relative importance of these terms in fitting the properties of the amide crystals. Thus our force field was chosen to be of the following form

$$2V = -\sum A/r^6 + \sum B/r^m + \sum qq'/r + 2V_{HB} \quad (1)$$

The molecules were considered as rigid bodies, and intramolecular interactions were therefore neglected. The sums ran over all interatomic distances between different molecules within a central unit cell (if it contained more than one molecule) and between the molecules in the central unit cell and all molecules in the other unit cells, up to a given cut-off distance (see subsection C).

The parameter m in eq 1 was given alternatively the values of 9 and 12 for comparison. The parameters A, B, and q for the atoms in the alkyl groups were transferred from a previous study[15] for $m = 9$ and from a fit to the crystal properties of hexane and octane for $m = 12$.

The covalent part of the hydrogen bond term was chosen as a Morse potential, attenuated by an exponential angular dependence, $f(\theta)$

$$V_{HB} = D\{e^{-2\alpha(r-r_0)} - 2e^{-\alpha(r-r_0)}\}f(\theta) \quad (2)$$

where D, α, and r_0 are the conventional Morse potential parameters and r is the O···H distance. Two alternatives of the angular dependence were considered. In the first $f(\theta)$ was given by $e^{-(\theta/\theta_0)^2}$ where θ is the supplement of the N–H···O angle, while in the second we tried $f(\theta) = e^{-(\theta_1/\theta_0)^2} + e^{-(\theta_2/\theta_0)^2}$ where θ_1 and θ_2 are the angles between the H···O vector and the assumed directions of the two lone pair orbitals on the oxygen atom.

The Lennard-Jones interaction between the N–H and O was attenuated by the given angular dependence in exactly the inverse manner, such that for a "perfect" hydrogen bond the nonbonded interaction was zero, and it grew correspondingly larger, the smaller the hydrogen bond contribution. Explicitly, the corresponding terms in eq 1 were replaced by $(-A/r^6 + B/r^m)(1 - f(\theta))$.

(B) Experimental Data. The crystallographic data for the least-squares optimization were taken from the experimental studies on the following eight molecules: oxamide,[21] malonamide,[22] succinamide,[23] adipamide,[24] urea,[25] formamide,[26] diketopiperazine (DKP),[27] LL-3,6-dimethyl-2,5-piperazinedione (LL-DMDKP).[28] This set of molecules represents a wide variety of hydrogen bond geometries, as can be seen from Table I. (Glutaramide and cyclopropanecarboxamide are also included in Table I and Figure 1, although they are not included

Table I. Hydrogen Bond Geometry[a]

	H···O, Å	180° − NH···O, deg	H···O=C, deg
Oxamide	2.02	29	155
	1.94	4	119
Malonamide	1.90	2	121
	1.97	13	122
	1.95	23	121
	1.96	5	116
	1.97	20	120
	2.13	29	144
	1.99	21	123
	2.31	41	149
Succinamide	1.94	3	138
	1.95	4	117
Glutaramide	2.02	28	151
	1.98	11	121
Adipamide	1.98	18	127
	2.06	35	146
Urea	2.07	12	106
	2.09	28	148
Formamide	1.91	19	125
	1.95	13	121
DKP	1.85	4	123
LL-DMDKP	1.91	11	120
	1.93	13	123
Cyclopropane-carboxamide	1.99	19	136
	1.92	10	138
	1.98	10	115
	2.03	18	122

[a] Based on hydrogen positions refined by minimization (see text).

(19) A. T. Hagler and S. Lifson, *J. Amer. Chem. Soc.*, **96**, 5327 (1974).

(20) T. Ooi, R. A. Scott, G. Vanderkooi, and H. A. Scheraga, *J. Chem. Phys.*, **46**, 4410 (1967).

(21) E. M. Ayerst and J. R. C. Duke, *Acta Crystallogr.*, **7**, 588 (1954).

(22) P. C. Chieh, E. Subramanian, and J. Trotter, *J. Chem. Soc. A*, 179 (1970).

(23) R. Davies and R. A. Pasternak, *Acta Crystallogr.*, **9**, 334 (1956).

(24) M. Hospital and J. Housty, *Acta Crystallogr.*, 20, 626 (1966).

(25) J. E. Worsham, Jr., H. A. Levy, and S. W. Peterson, *Acta Crystallogr.*, 10, 319 (1957).

(26) J. Ladell and B. Post, *Acta Crystallogr.*, **7**, 559 (1954).

(27) R. Degeilh and R. E. Marsh, *Acta Crystallogr.*, 12, 1007 (1959).

(28) E. Benedetti, P. Corradini, and C. Pedone, *Biopolymers*, **7**, 751 (1969).

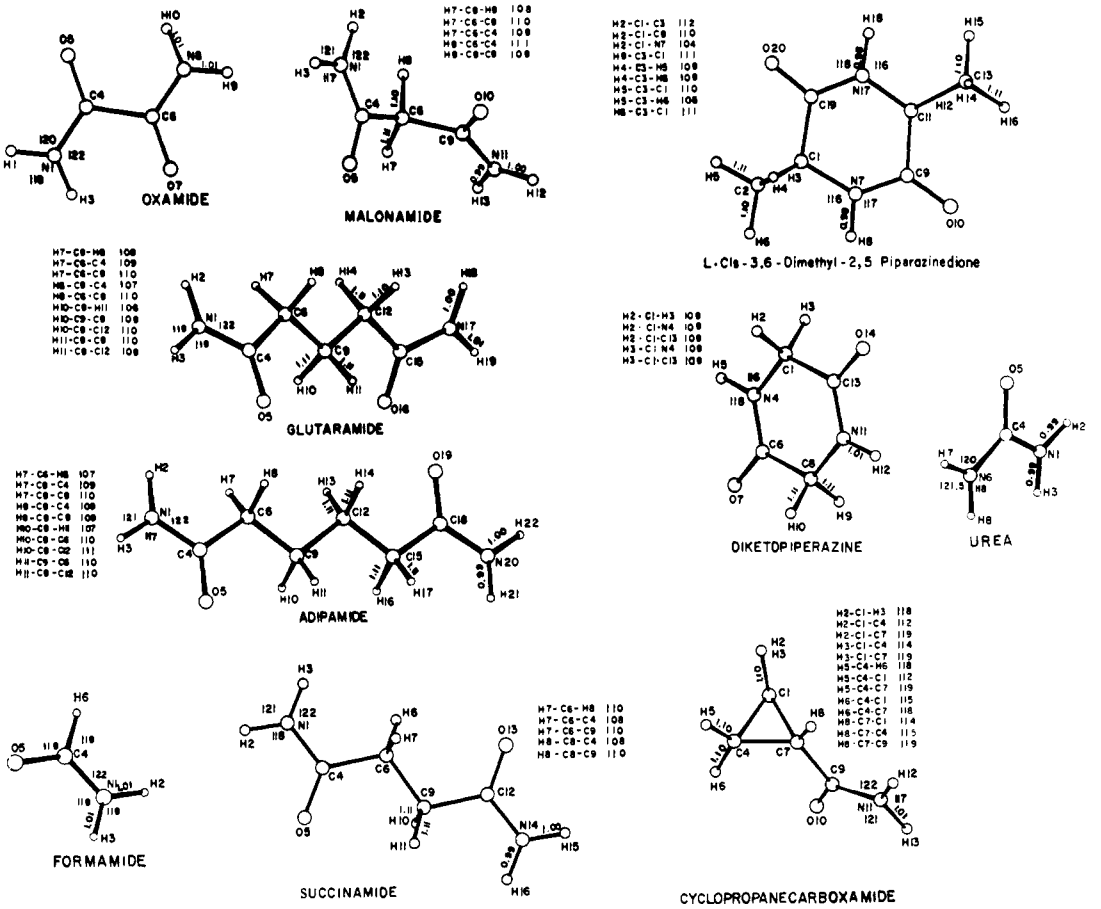
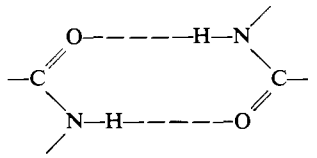


Figure 1. Hydrogen geometry of amide molecules as obtained by minimization of the intramolecular energy.

in the least-squares optimization. They are discussed in paper II.[19])

Furthermore, the overall crystal structure varies from crystal to crystal.[21-29] For example, most of the amides form hydrogen bonded centrosymmetric rings.



which pack along a translation axis (*e.g.*, oxamide[21]), a glide plane (*e.g.*, succinamide[23]), or a screw axis (*e.g.*, formamide[26]), whereas adipamide forms interconnected extended chains.[24,29] Adipamide is the only amide known to pack in a structure where the hydrogen bonds lie along a translation as well as a two-fold screw axis, except possibly the unstable form of chloroacetamide.[29,30] These packing arrangements result in an interconnected network of hydrogen bonds, which often forms a layer structure (oxamide, succinamide, formamide, and adipamide). Diketopiperazine[27] and LL-DMDKP[28] can participate in only four hydrogen bonds per molecule. They both form one-dimensional extended chains consisting of molecules hydrogen bonded by centrosymmetric rings. Urea is exceptional since it has only one oxygen but four hydrogens per molecule. This one oxygen is hydrogen bonded to four hydrogens of adjacent molecules.[25] Two of these hydrogen bonds involve N–H bonds in a molecule whose plane is perpendicular to the plane of the molecule containing the oxygen, a rather unusual situation.

The thermodynamic data[31] were taken from studies of the heat of sublimation of the following molecules: oxamide,[32] urea,[33] formamide,[34] and diketopiperazine.[35] The heats of sublimation of malonamide and succinamide were estimated from

$$\Delta H_{\text{subl}} = \Delta H_f^\circ(\text{gas}) - \Delta H_f^\circ(\text{solid}) \qquad (3)$$

by taking the heat of formation in the solid (ΔH_f°(solid)) from available data[31] and evaluating the heat of formation of the molecules in the gas phase (ΔH_f°(gas)) from bond and group contributions.[36]

The lattice energy per mole, V_L (which is given by V

(29) L. Leiserowitz and G. M. J. Schmidt, *J. Chem. Soc. A*, 2372 (1969).

(30) M. Katayama, *Acta Crystallogr.*, **9**, 986 (1956).

(31) J. D. Cox and G. Pilcher, "Thermochemistry of Organic and Organometallic Compounds," Academic Press, New York, N. Y., 1970.

(32) R. S. Bradley and T. G. Cleasby, *J. Chem. Soc.*, 1681 (1953); E. P. Egan, Jr., Z. T. Wakefield, and T. D. Farr, *J. Chem. Eng. Data*, 10, 138 (1965).

(33) K. Suzuki, S. Oniski, T. Koide, and S. Seki, *Bull. Chem. Soc. Jap.*, **29**, 127 (1956).

(34) G. Somsen and J. Coops, *Recl. Trav. Chim. Pays-Bas*, **84**, 985 (1965); A. Bander and H. Günthord, *Helv. Chim. Acta*, 41, 670 (1958).

(35) S. Seki, K. Suzuki, and T. Koide, *Kogyo Kagaku Zasshi*, **77**, 346 (1956).

(36) The value of the heat of formation of the group ($CCONH_2$) in the gas phase was obtained from the known heat of formation of hexanamide[31] (using the bond parameters in Table 50 p 542 of ref 31). The gaseous heats of formation of malonamide and succinamide were then calculated using the same parameters for the C–H and C–C bonds and this group contribution. These values were also used to calculate ΔH_f°(gas) of octanamide and acetamide and gave agreement within 2 kcal of the experimental values.

in eq 1 divided by the number of molecules per unit cell), was related to the heat of sublimation

$$\begin{aligned}\Delta H_{\text{subl}} &= H_{\text{gas}} - H_{\text{solid}} \\ &\cong pv + 3RT - (V_L + 6RT) \\ &\cong -V_L - 2RT\end{aligned} \tag{4}$$

where we have assumed H_{gas} to be given by the ideal gas law, and the vibrational energy of the crystal was estimated by $6RT$. A high estimate of the error of the lattice energy is of the order of 1–2 kcal/mol[37] except for malonamide and succinamide where it is higher due to the use of bond and group contributions.[36]

The dipole moments of urea,[38a] formamide,[38b] and *N*-methylacetamide[39] were also included in the least squares. The dipole moments of formamide and *N*-methylacetamide were measured in the gas phase while that of urea was taken from measurements in solution. The comparison of these with the calculated values implies the assumption that the charge distribution is the same as in the crystal state. Because of this, small relative weights were assigned to these quantities.

The locations of the hydrogens in the crystal structure are either omitted or inaccurately estimated in the X-ray diffraction studies used here. Since their location may significantly affect intermolecular energy calculations, we calculated the position of the hydrogens in all crystals. This was done by minimizing the intramolecular energy with respect to the hydrogen coordinates while all other atoms in the molecule were kept fixed at their crystallographic positions. The calculations were performed using an approximate force field previously derived for amides.[40,41] The results are given in Figure 1. This method is better than the method of attributing a "standard" geometry to the hydrogens because the angles and distances involving the hydrogens do actually vary, as do other angles and distances, both in the amide group and the rest of the molecule (see, *e.g.*, Table 6 of ref 23 and Table 7 of ref 22).

(C) Optimization of the Force Field. A linearized least-squares method[42] is used to obtain a best fit to the experimental data described above. The particular aspect which requires comment here is the way the crystal geometry is treated. All the molecules of the unit cell are kept fixed and the energy parameters are changed such as to give a best fit to the nine components of the three unit cell vectors. The method used for the optimization with respect to the unit cell vectors is discussed in detail elsewhere.[16] Here we shall give only a concise summary. Let **a** denote the vector composed of these nine components. Rather than calculating $\mathbf{a}_{\text{calcd}}$ by a tedious energy minimization, we estimate $\Delta\mathbf{a} = \mathbf{a}_{\text{calcd}} - \mathbf{a}_{\text{exptl}}$ by

$$\Delta\mathbf{a} = -\mathbf{F}(\mathbf{a}_{\text{exptl}})^{-1}\nabla_{\mathbf{a}}\mathbf{V}(\mathbf{a}_{\text{exptl}}) \tag{5}$$

based on the Newton–Raphson formula. Here **F** is the matrix of second derivatives, $F_{\alpha\beta} = \partial^2 V/\partial a_\alpha \partial a_\beta$, and $\nabla_{\mathbf{a}}\mathbf{V}$ is a vector whose components are $\partial V/\partial a_\alpha$. Both **F** and $\nabla\mathbf{V}$ are calculated at the experimental crystal structure ($\mathbf{a}_{\text{exptl}}$). The parameters, represented by a vector **p**, are changed iteratively according to the formula

$$\Delta\mathbf{p}^i = -\{(\mathbf{Z}'\mathbf{W}^2\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{W}^2\Delta\mathbf{y}\}^{i-1} \tag{6}$$

The vector $\Delta\mathbf{y}$ includes all the components of the vector $\Delta\mathbf{a}$ as given by eq 5, as well as the differences in the lattice energy (ΔV_L) and dipole moments ($\Delta\mu$) at the $(i-1)$ iteration. **Z** is the Jacobian matrix, $Z_{lm} = \partial y_l/\partial p_m$, and **W** is the diagonal matrix of the weights of the various observables.

Once the optimal force field has been derived, we further calculated the exact minimum geometry of the crystals using the Fletcher–Powell[43] method. The minimization was carried out over all independent variables of the unit cell, their number being $9 + 6\cdot(s-1)$, where s is the number of molecules per unit cell. In this way the symmetry properties of the unit cell were not imposed but rather derived, thus ensuring that the calculated structure be a true energy minimum (see the subsequent paper[19]).

The calculations involved lattice sums over the energy components and their first and second derivatives. Two unit cells on each side of the central unit cell were taken into account giving a total of $5^3 = 125$ unit cells. Since some of the unit cells dimensions are very large a cut-off criterion was employed. In order to ensure reliability of the results this cutoff was taken as 50 Å for the lattice sums in the least squares. If the distance between any two atoms of two molecules was less than the cutoff, the interactions of all atoms of the two molecules were included. This procedure was used to avoid the danger of creating artificial effects from parts of a molecule, which may exert spurious electrostatic forces due to their charges or dipole moments. It should be pointed out that such a large cut-off distance was feasible because the lattice sums were calculated only once at the experimental structure of each crystal.[16]

II. Results of the Least Squares

(A) Morse Potential's Failure. As stated above, our trial force field included, initially, a Morse potential, attenuated by an angular dependence, to represent an assumed covalent contribution to the O···H interaction. The purpose was to assess by the least-squares analysis of experimental data the relative importance of the electrostatic, nonbonded, and possible covalent contributions in fitting the crystal data. The results were unequivocally against the use of the Morse potential. All the standard deviations except that of r_0 were found to be larger than the parameters themselves, and therefore the potential as a whole appeared to be ill-determined. Furthermore, the potential well was small ($D = 0.7$ kcal/mol for the first $f(\theta)$ and 0.2 for the second). The only definite role it did play in the calculations was to supply a repulsive force, to offset the attractive electrostatic forces. The Morse potential appeared, therefore, to be superfluous, namely, that by omitting the angle-dependent Morse potential altogether and considering the Lennard-Jones potential as equally operative for all angles θ, we could get compar-

(37) The experimental error in the measured heat of sublimation is of the order of ±0.5 kcal/mol.[31-35] The correction to obtain the lattice energy amounts to ~1.2 kcal ($2RT$). Assuming a large uncertainty in this correction yields the quoted high estimate for the error in the lattice energy.

(38) (a) A. L. McClellan, "Tables of Experimental Dipole Moments," W. H. Freeman, San Francisco, Calif., 1963; (b) J. Kurland and E. B. Wilson, Jr., *J. Chem. Phys.*, 27, 585 (1957).

(39) R. M. Meighan and R. H. Cole, *J. Phys. Chem.*, **68**, 503 (1964).

(40) A. Warshel, M. Levitt, and S. Lifson, *J. Mol. Spectrosc.*, **33**, 84 (1970).

(41) S. Karplus and S. Lifson, *Biopolymers*, 10, 1973 (1971).

(42) W. C. Hamilton, "Statistics in Physical Science," Ronald Press, New York, N. Y., 1964, p 150.

(43) R. Fletcher and M. J. D. Powell, *Comput. J.*, **6**, 163 (1963).

able results with fewer parameters. This assumption was fully supported by subsequent results. Consequently our final choice of the force field consisted of just the nonbonded and electrostatic terms in the form given by eq 1. It should be noted that these results cannot be taken to imply that there is no function which might be used to represent charge transfer that would improve the fit to the data. They only indicate the inadequacy of the Morse potential and similar functions mentioned in section IA, on the one hand, and the possibility to obtain a reasonable agreement with structural and energetic properties of amides, when the hydrogen bond is represented by electrostatic and nonbonded interactions only, on the other hand. These results are in agreement with conclusions reached on the basis of quantum mechanical calculations as to the dominance of the electrostatic interaction in hydrogen bonding,[17] although the charge transfer contribution in these studies is not negligible.[12,17d] Many of the studies carried out to date have been on O–H···O hydrogen bonding systems[12,17] for which the O···H distance, ~1.7, may be compared to the O···H distance in amides of ~1.9, and thus charge transfer might not be expected to be as important for these longer bonds. (For further discussion on this subject, see ref 12, 17b, and 13).

(B) The Optimized Results. The iteration history of the least-squares optimization of this force field is given in detail for the 6–12 potential in Table II. The

Table II. Iteration History of Weighted Deviations in the Least-Squares Calculations[a]

Iteration	$\Sigma(w\Delta a)^2$	$\Sigma(w\Delta V_L)^2$	$\Sigma(w\Delta\mu)^2$	$\Sigma(w\Delta y)^2$
		The 6–12 Potential		
1	2560	223	103	2886
2	988	157	68	1213
3	578	178	50	806
4	389	199	39	627
5	287	211	30	528
10	125	124	15	264
20	111	90	6	207
30	103	66	1	170
47	100.1	64.5	0.8	165.4
		The 6–9 Potential		
47	162	94	11	267

[a] w are the weights given in Table III; Δa, ΔV_L, and $\Delta\mu$ are the deviations in unit cell parameters, lattice energies, and dipole moments, respectively. The last column is the total weighted sum of squares ($\Delta y = \Delta a, \Delta V_L, \Delta\mu$).

iteration history for the 6–9 potential is similar, and only the final results are given. The results of the optimization with the 6–12 potential seem to be better than those of the 6–9 potential. However, the difference is reduced when the exact minimization of the crystal energy is carried out.[19] The remaining slight difference in the sum of squared deviations is not significant, and in any case is partly due to the fact that two different alkane force fields were used, as discussed above. Thus we conclude that the 6–9 and 6–12 are essentially equally successful in fitting the properties of these crystals.

The convergence of the least-squares method was estimated by both the diminishing change of the sum of squares per iteration and the vanishing of the gradient of the sum of squares with respect to the parameters. The least squares was terminated when the change in the sum of squares in three successive iterations was less than 0.1%. The derivatives of the sum of squares with respect to the parameters at this point were such that a change of 1% in the parameter corresponding to the largest derivative gave a change in the sum of squares of 0.1%, while the same change in most of the other parameters gave a change of less than 0.01%.

The results of the least squares are presented in Table III. The calculated $\Delta\mathbf{a}$, being derived by optimization,

Table III. Results of the Least-Squares Optimization[a]

	(1) Crystal Structure[b]			
	Oxamide	Malonamide	Succinamide	Adipamide
Δa_{1x}	0.17	−0.01	0.05	0.01
Δa_{1y}	−0.12	0.01	0.00	0.04
Δa_{1z}	0.05	−0.06	0.12	0.04
Δa_{2x}	−0.02	−0.03	0.00	−0.02
Δa_{2y}	−0.21	0.04	−0.01	−0.02
Δa_{2z}	0.17	−0.00	0.00	−0.03
Δa_{3x}	0.05	0.10	−0.19	0.04
Δa_{3y}	−0.17	−0.02	0.00	0.07
Δa_{3z}	0.02	0.01	−0.03	0.08
	Urea	Formamide	Diketo-piperazine	LL-DMDKP
Δa_{1x}	−0.10	−0.02	0.02	−0.73
Δa_{1y}	−0.03	0.01	−0.10	−0.16
Δa_{1z}	0.15	−0.01	0.23	0.48
Δa_{2x}	−0.05	−0.16	0.14	−0.06
Δa_{2y}	−0.08	0.13	0.13	0.09
Δa_{2z}	−0.15	0.12	−0.24	0.38
Δa_{3x}	−0.04	0.19	−0.15	0.07
Δa_{3y}	0.03	−0.11	0.08	−0.55
Δa_{3z}	0.04	−0.10	0.05	−0.20

	(2) Lattice Energy and Dipole Moments					
	Lattice energy, kcal/mol			Dipole moments, D		
	Exptl	Calcd	w	Exptl	Calcd	w
Oxamide[c]	−28.2	−25.6	2.0			
Malonamide	−28.8	−31.5	1.6			
Succinamide	−32.3	−33.3	1.6			
Urea	−22.2	−22.2	3.0	4.45	4.32	5
Formamide	−17.5	−16.5	2.0	3.71	3.81	5
DKP	−26.0	−26.4	3.0			
N-Methyl-acetamide				3.71	3.65	5

[a] For the 6–12 potential. [b] The Δa's (in Å) were derived by eq 4. The weights used in eq 5 were 10 for all molecules except LL-DMDKP for which it was 5. [c] Four unit cells in each direction were used to calculate the lattice energy, rather than the two normally used because of the small size of the unit cell. (The difference in V_L amounts to 0.4 kcal.)

without minimization of the energy, gives approximate values of the difference between $\mathbf{a}_{calcd}$ and $\mathbf{a}_{exptl}$. The calculated energy as given in this table relates to the experimental coordinates. Only oxamide exceeds the estimated experimental error of about 1–2 kcal/mol. In malonamide the difference is similar; however, the "observed" value involves an estimate of the heat of formation in the gas phase, from group contributions, as described above and therefore possesses a comparable error. The dipole moments are fit as well or better than should be expected since the model of partial charges is a simplification and since the dipole moments were obtained in the gas phase and in solution, where geometry and charge distribution may be different from those of the crystal phase. In fact we took the

Table IV. Potential Parameters[a]

Atom	6–12 potential			6–9 potential		
	A	$B \times 10^{-3}$	q[c]	A	$B \times 10^{-3}$	q[c]
H_C[b]	32.9	7.15	0.10	15.0	0.445	0.11
O	502 (269)	275 (70)	−0.38 (0.02)	1410 (463)	45.8 (7.1)	−0.46 (0.03)
N	1230 (266)	2271 (367)	−0.28 (0.03), −0.83 (0.06)	2020 (414)	86.9 (16.2)	−0.26 (0.04), −0.82 (0.07)
C	532	1811		1230	38.9	
C′	1340 (648)	3022	0.38	355 (401)	12.5 (11.5)	0.46
H_N[b]	0	0	0.28	0	0	0.26
			0.41			0.41

[a] Units: energy in kcal/mol, length in Å, charge in electrons. [b] H_C and H_N represent the hydrogen on carbon and on nitrogen, respectively. [c] The following groups were defined as neutral: CO, NH, NH_2, and CH_n. The charges for N and H_N refer to NH and NH_2, respectively.

Table V. Correlation Matrices for the 6–12 and 6–9 Potentials[a]

6–9 potential	6–12 potential								
	A_O	B_O	A_N	B_N	$A_{C'}$	$B_{C'}$	q_{NH}	$q_{C'}$	q_{NH_2}
A_O	1	0.85	−0.40	−0.23	−0.84	−0.80	−0.02	0.32	−0.03
B_O	0.86	1	−0.06	−0.19	−0.86	−0.91	−0.13	−0.06	−0.28
A_N	−0.23	−0.11	1	0.68	0.26	−0.03	0.32	0.00	0.30
B_N	−0.05	−0.04	0.81	1	0.03	−0.15	0.51	0.32	0.76
$A_{C'}$	−0.52	−0.60	−0.22	−0.26	1	0.94	0.17	0.07	0.14
$B_{C'}$	−0.71	−0.84	−0.11	−0.29	0.89	1	0.08	−0.02	0.05
q_{NH}	0.31	0.17	0.07	0.29	0.24	0.03	1	0.50	0.74
$q_{C'}$	0.51	0.11	−0.05	0.17	−0.03	−0.10	0.47	1	0.71
q_{NH_2}	0.42	0.14	−0.04	0.42	0.07	−0.14	0.73	0.72	1

[a] The correlation matrix is symmetric. The top half gives the correlation matrix for the 6–12 potential while the bottom half corresponds to the 6–9 potential.

dipole moments only as constraints on the charges and gave them relatively low statistical weights.

The geometry of the crystal has been included in the least-squares optimization by fitting the vector **a** of the nine components of the unit cell vectors. These determine fully the crystal structure of a crystal with one molecule per unit cell. However, six out of the eight optimized crystals have more than one molecule per unit cell. In these crystals the relative positions of the molecules in the unit cell were not subject to the optimization. Instead, only the packing of the content of the unit cells as rigid structures was optimized. Thus the total number of degrees of freedom of most of the crystals was significantly larger than the number of degrees of freedom included in the optimization. Subsequently, we carried out the total minimization of the crystal energy with respect to *all* degrees of freedom. This constitutes a more severe test of the force field than has been previously applied. The results are given and discussed in the following paper.[19]

(C) Energy Parameters. The optimized parameters of the attractive and repulsive terms of the Lennard-Jones potential, A and B, respectively, and the partial charges, q, of the atoms N, O, C′, and H_N are given in Table IV, with their standard deviations in parentheses. The parameters of C and H, transferred from alkanes, are also included. The correlation matrix for the optimized parameters is given in Table V. Its elements are measures of how well a change in one parameter can chmpensate for the effect of a change in the second on the sum of the squared deviations. (For a further discussion of the correlation coefficient see ref 42, p 31.)

(1) Partial Charges. Both the standard deviations and the correlation coefficients of the partial charges are relatively small. It is of interest to compare these charges both with those used by other authors in empirical conformational analysis and with those derived by molecular orbital methods.[44–48] The charge distributions resulting from a number of molecular orbital calculations on formamide are given in Table VI, in which we have again included our results for ease of comparison. We have taken the charges on the formyl hydrogen to be the same as that of an alkane hydrogen in order to avoid the introduction of an extra parameter for a single molecule.

Table VI. Charge Distributions in Formamide

	CFF (6–12)	*Ab initio* I[44]	*Ab initio* II[45]	*Ab initio* III[46]	CNDO/2[47]	Del Re method[48]
H_1[a]	0.101	0.152	0.194	0.17	0.039	0.063
C[b]	0.282	0.258	0.170	0.36	0.450	0.278
O	−0.383	−0.377	−0.383	−0.41	−0.365	−0.413
N	−0.829	−0.758	−0.639	−0.86	−0.409	−0.376
H_2[a]	0.415	0.368	0.337	0.37	0.188	0.224
H_3	0.415	0.357	0.321	0.37	0.174	0.224

[a] H_1 is the carbon hydrogen and H_2 is the amide hydrogen cis to the oxygen. [b] In formamide the HC=O group was taken as neutral.

It is surprising to see how well the calculated charges correspond to those of the *ab initio* calculations. It should be noted that the derivation of partial charges from molecular orbital calculations by population analysis is not meant to represent the contribution of

(44) M. B. Robin, F. A. Bovey, and H. Basch in "The Chemistry of the Amides," J. Zabicky, Ed., Interscience, New York, N. Y., 1970.
(45) M. Dreyfus, B. Maigret, and A. Pullman, *Theor. Chim. Acta*, 17, 109 (1970).
(46) M. A. Robb and I. G. Csizmadia, *Theor. Chim. Acta*, 10, 269 (1968).
(47) J. F. Yan, F. A. Momany, R. Hoffman, and H. A. Scheraga, *J. Phys. Chem.*, 74, 420 (1970).
(48) D. Poland and H. A. Scheraga, *Biochemistry*, 6, 3791 (1967).

the partial charges to the molecular force field in the same way as in the CFF. The one depends on the various inherent assumptions in the population analysis and in the molecular orbital theories, while the other depends on the fit between observed properties and those calculated from the intermolecular forces and on the assumption of the additivity and transferability of atom–atom interactions. Therefore, the similarity of the results is rather striking.

It is also interesting to note that the partial charges on the NH_2 group of formamide, as derived by various CNDO/2 calculations[47] and by the Del Re method[48] are significantly smaller than ours, as can be seen from Table VI. The Del Re method has also been used extensively for deriving charges for conformational calculations in molecules of biological interest.[13] As for the charges on the NH group we initially assumed the same charge on its hydrogen as on that of the NH_2. However, this led to a poor fit with experiment, involving not only dipole moments but also heats of sublimation. When we attributed independent parameters, the charges diverged significantly and the calculated properties of both the NH_2 containing molecules and those with NH agreed better with experiment. Our values for the charges on the C=O (± 0.38)[49] and the NH (± 0.28) groups agree very well with those estimated by Brant, *et al.*[50a] (C=O = ± 0.39, N–H = ± 0.28) from dipole moments and widely used in conformational calculations of peptides.[2,13,50b] This is interesting because the dipole moments were given relatively small weights in our calculation, while the optimized charges were very sensitive to the sublimation energy and also to the crystal structure.

The partial charges of the NH (± 0.26) in the force field with the 6–9 potential differ slightly from those in the 6–12 force field while the C=O charges (± 0.46) differ more. This difference may be due in part to the partial correlation between the various parameters, as seen in Table V.

(2) The Nonbonded Parameters. The presentation of the nonbonded parameters as in eq 1 is very convenient mathematically since the energy is linearly dependent on the parameters A and B if these are attributed independently for each pair of atoms. It is customary to reduce the number of independent parameters for two atoms say X and Y by the relations $A_{XY} = (A_X A_Y)^{1/2}$ and $B_{XY} = (B_X B_Y)^{1/2}$. Another equivalent presentation of the Lennard-Jones type potential is

$$V_{6-12} = \epsilon\{(r^*/r)^{12} - 2(r^*/r)^6\} \qquad (7)$$

or

$$V_{6-9} = \epsilon\{(2(r^*/r)^9 - 3(r^*/r)^6\}$$

where r^* is the distance at the minimum of the potential, commonly identified with twice the "van der Waals radius," and ϵ is the depth of the potential well. Here the combination rule commonly used is $\epsilon_{XY} = (\epsilon_X \epsilon_Y)^{1/2}$ and $r_{XY}^* = {}^1/_2 (r_X^* + r_Y^*)$. It is obvious that the two combination rules are not identical (except when $r_X^* = r_Y^*$).

(49) The $\pm$ notation represents the opposite signs on the two atoms in the electroneutral C=O and NH bonds (see Table IV).

(50) (a) D. A. Brant, W. G. Miller, and P. J. Flory, *J. Mol. Biol.*, 23, 47 (1967); (b) D. A. Brant, *Annu. Rev. Biophys. Bioeng.*, **1**, 369 (1972).

We have therefore examined both combination rules. We optimized the parameters A and B as well as r^* and ϵ. The two results were basically the same, the differences both in the sum of squares and in the parameters themselves (when converted to the same form) being insignificant. The values of r^* and ϵ for both the 6–9 and the 6–12 potentials (corresponding to A and B in Table IV) are presented in Table VII along with values of these parameters commonly used in conformational calculations.

Table VII. Nonbonded Parameters of Atoms in Peptides

	6–9 (this work)		6–12 (this work)		—Ref 2—		—Ref 50a—	
Atom	r^*	ϵ	r^*	ϵ	r^*	ϵ	r^*	ϵ
H_C	3.54	0.0025	2.75	0.038	2.4	0.122	2.6	0.076
O	3.65	0.198	3.21	0.228	3.04	0.23	3.2	0.172
N	4.01	0.161	3.93	0.167	3.1	0.20	3.3	0.212
C	3.62	0.184	4.35	0.039	3.4	0.11	3.6	0.138
C′	3.75	0.042	4.06	0.148	3.4	0.11	3.6	0.138
H_N	0	0	0	0	2.4	0.122[a]	2.6	0.076
CH_2							3.9	0.160

[a] When special H bond potential is used, the nonbonded O···H interaction is omitted.

One of the most striking features of our results as can be seen in both Tables IV and VII is the fact that the nonbonded parameters for the amide hydrogen H_N are zero. Initially these parameters were included in the least-squares procedure. During the course of the optimization both the diameter r^* and the depth of the well ϵ of H_N decreased to values of the order of 1 Å and 0.01 kcal/mol, respectively. Furthermore the standard deviations of these parameters were larger than the values of the parameters themselves. These results indicated that the nonbonded parameters of the amide hydrogen could be ignored altogether, without affecting the results. Thus we set r^* and ϵ of H_N to zero and reoptimized, and indeed the sum of squares was essentially unaffected while the number of independent parameters was reduced by two. In this connection, it should be noted that the insensitivity to the nonbonded parameters of H_N hold only for r^* less than approximately 2 Å. For larger values than this (if ϵ of H_N is nonzero) there is no way of fitting the properties of these crystals satisfactorily. This arises from the fact that for larger values than 2.0 the equilibrium distance between O and H would necessarily be larger than the experimentally observed O···H distance $\sim$1.85–2.0 Å, in the N–H···O hydrogen bond. Force fields which do contain larger values for r^* of the amide hydrogen forbid the nonbonded interaction between O and H to occur.

This feature is in qualitative agreement both with previously proposed concepts of the hydrogen bond and with quantum mechanical calculations of the electronic distribution around the hydrogen when it is bonded to electronegative atoms. Thus Coulson[51] and others[10,52] have attributed the special nature of the hydrogen bond to the fact that the exchange repulsion of the hydrogen (equivalent to a small B or combination of small r^* and ϵ) is small. Furthermore, calculations of the electronic

(51) C. A. Coulson, *Res. Appl. Ind.*, 10, 149 (1957).

(52) J. N. Murrell, *Chem. Brit.*, **5**, 107 (1969).

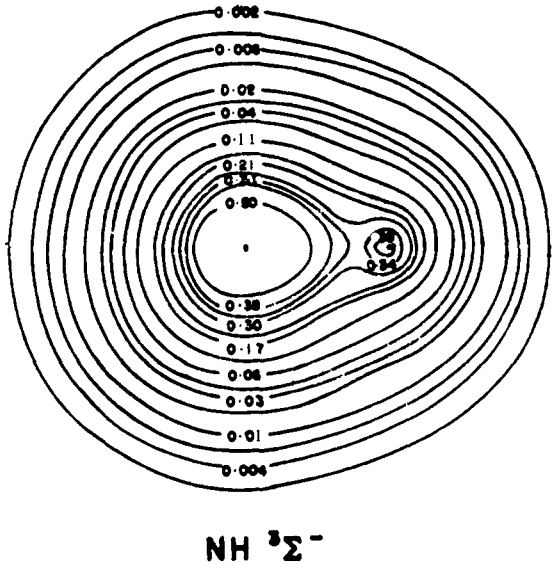


Figure 2. Electronic distribution of N–H, taken from ref 53.

distributions[53] of XH diatomic "molecules" show that the "size" of the hydrogen decreases significantly with the electronegativity of X. In fact for highly electronegative atoms X the molecule is approximately spherical around the center of X. The electronic distribution of N–H is given in Figure 2.[53] The small size of the hydrogen, and why it can be ignored relative to the nonbonded interaction of the nitrogen, can be further seen from this figure.

The above consideration may raise the question whether the nonbonded parameters of the atoms forming the NH group should not be independent of those of the NH_2 group, since the partial charges of the atoms are quite different. To answer this question by the methods proposed here many more molecules containing NH groups should be included in the analysis. This awaits further study.

One of the results of omitting the nonbonded parameters of H_N was that the nonbonded parameters of N have the lowest correlation and standard deviation among the nonbonded parameters as is seen in Tables IV and V. The highest correlations and standard deviations are those of the parameters of the amide carbon, C′. Indeed the sum of squares is relatively insensitive to whether the nonbonded parameters of C′ are varied independently or assigned the same values as those of C. The nonbonded parameters of the C′ were allowed to vary independently, since its nature should be different from the alkane carbon, which has a negative charge, while C′ is positive. The relative insensitivity of the crystal properties to the C′ nonbonded parameters may be due to the fact that the C′ is not on the perimeter of any of these molecules. It therefore makes fewer contacts with neighboring molecules and is less well determined by least-squares fitting to the molecular packing. (See, however, the discussion of oxamide in the following paper.[19])

Comparing our nonbonded parameters with those used by other authors (Table VII), one notes that our radii are generally larger. Scheraga, *et al.*,[2,20] took their values from Bondi's[54] compilation of experimental contact distances in crystals. Both Kitaygorodsky[55] and Brant, *et al.*,[50a] have noted that the contact distances as observed in crystals cannot be associated with the minimum of the nonbonded pair potential, because of the compressive effect of the rest of the lattice. Kitaygorodsky assumed the minimum distance to be 0.25, 0.225, and 0.2 Å larger than the observed contact distances for H···H, H···C, and C···C interactions, respectively. Brant, *et al.*, assumed an increment of 0.2 Å for all interactions (*e.g.*, H, O, N, C, C′, etc.). In the least-squares calculation performed here, the nonbonded parameters were derived taking all lattice forces and interactions into account and with no *a priori* restrictions on the location of the minimum of the nonbonded potential. It may therefore be concluded from the comparison of the "van der Waals radii" in Table VII that the effect of lattice compression is larger than has previously been assumed and certainly may not be neglected. It might be noted in this connection that the atomic radii, r^* of N and O in N_2 and O_2, are 3.80[56] and 3.32 Å,[57] respectively. These radii have been determined from the crystal structures of α nitrogen and α oxygen, taking full account of crystal forces. These numbers also substantiate the fact that the r^* are larger than the short contact distances previously used.

(3) Comments on Sensitivity of Structure and Energy to the Potential Parameters. The sensitivity of various properties such as conformation or energy to a given term in a potential function is often of use in making or judging approximations, such as the early "hard sphere" calculations of dipeptides. The effect on the properties studied here of each of the potential parameters can be assessed by consideration of the elements of the **Z** matrix (eq 6). When these elements are multiplied by p, they give $\partial y/\partial \ln p$ which is the change in the property y with a given fractional change $(\Delta p/p)$ in the parameter. Consideration of this matrix for both the 6–9 and 6–12 potential leads to the following observations.

(i) For the most part the crystal structures (as reflected by the unit cell parameters) are more sensitive to the repulsive coefficients, B, of O and N than to the attractive coefficients, A, the ratio being in general $^1/_3$ to $^2/_3$ for both the 6–9 and 6–12 potential. The nonbonded parameters of the C′ have a smaller effect on the crystal structure in all cases for the 6–9 potential and in most cases (with oxamide as a noteworthy exception to be discussed[19]) for the 6–12 potential as well.

(ii) In general the charges are as, or more, important than the nonbonded parameters in determining the crystal structure and lattice energies.

(iii) The lattice energy is in every case more sensitive to the attractive coefficient, A, than to the repulsive coefficient, B, although the difference is not so pronounced in the 6–9 potential (~10–20%) as in the 6–12 potential (~33%). The magnitude of the numbers involved in the above observations may be sensed by considering some typical examples. Thus the changes in a_x (the x component of unit cell vector a) in adipamide with respect to the parameters for oxygen are $\partial a_x/\partial \ln A_0 = -0.82, \partial a_x/$

(53) R. W. F. Bader, I. Keavany, and P. E. Cade, *J. Chem. Phys.*, 47, 3383 (1967).

(54) A. Bondi, *J. Phys. Chem.*, **68**, 441 (1964).
(55) A. I. Kitaygorodsky, *Tetrahedron*, 14, 230 (1961).
(56) M. C. A. Donkersloot and S. H. Walmsley, *Mol. Phys.*, **19**, 183 (1970).
(57) J. C. Laufer and G. E. Leroi, *J. Chem. Phys.*, **55**, 993 (1971).

$\partial \ln B_0 = 1.34$, and $\partial a_x/\partial \ln q_0 = -0.99$. The corresponding elements of the **Z** matrix for the lattice energy of succinamide are $\partial V_L/\partial \ln A_0 = -13.57$, $\partial V_L/\partial \ln B_0 = 9.24$, and $\partial V_L/\partial \ln q_0 = -30.3$.

(4) The Hydrogen Bond. The energy functions derived through the CFF method to describe the amide hydrogen bond are seen to differ in no way from other interatomic interaction potentials. This raises a semantic question; is there such a thing as a hydrogen bond energy of the system C=O···H-N? If such energy could be defined exclusively as the extra energy, say due to charge transfer, which should exist over and above the various nonbonded interactions, then our results could not confirm its significance. It is more appropriate to consider the hydrogen bond as the sum of all interactions which make the system C=O···H—N particularly stable. In this connection it is worth noting that the most general and useful definitions of hydrogen bonds are geometric.

In fact, the energy and geometric dependence of the C=O···H—N system as given by these potential functions are, not surprisingly, very similar to those which have long been assigned to the hydrogen bond.[8] The minimum energy of this interaction is −2.4 kcal and occurs at an O···H distance of 2.1 Å. This feature along with the angular dependence of this interaction will be discussed in detail elsewhere.[58]

(58) A. T. Hagler, E. Huler, and S. Lifson, Proceedings of the Rehovot Symposium on Polyamino Acids, Polypeptides and Proteins, and their Biological Implications, 1974.

III. Summary

A force field for amides has been derived by a least-squares fit to the unit cell parameters, crystal energy, and dipole moments of amides. It was found that no explicit function need be included to represent the hydrogen bond, in order to fit the properties of these crystals. Instead the qualitative features of the hydrogen bond were seen to be natural outcomes of the usual nonbonded and electrostatic energy functions. What makes the hydrogen bond interaction particular is the negligible radius of the amide hydrogen, which allows a short contact distance between the NH and the CO groups, resulting in a strong electrostatic interaction. The nonbonded van der Waals radii, r^*, as derived by the least squares, with no *a priori* assumptions as to minimum energy distances, are larger than those currently used for conformational calculations. This is the result of taking full account of the effect of compression due to intermolecular forces on the observed contact distances. The charges as derived here are in general agreement with those commonly used in conformational calculations of peptides as derived from the amide dipole moment. The charges on the NH_2 group as derived here are in better agreement with those obtained by *ab initio* calculations of formamide than with those obtained from CNDO/2 and Del Re method calculations on the same molecule.

Acknowledgment. We are pleased to acknowledge the programming assistance of Mrs. Ruth Sharon and Miss Miriam Harel.

Energy Functions for Peptides and Proteins. II. The Amide Hydrogen Bond and Calculation of Amide Crystal Properties

A. T. Hagler* and S. Lifson

Contribution from the Chemical Physics Department, Weizmann Institute of Science, Rehovot, Israel. *Received December 27, 1973*

Abstract: The potential functions derived in the previous paper are tested by minimization of ten amide crystals (oxamide, malonamide, succinamide, glutaramide, adipamide, urea, formamide, diketopiperazine, LL-3,6-dimethyl-2,5-piperazinedione, and cyclopropanecarboxamide). The structures of these crystals are calculated by minimizing the energy with respect to all crystalline degrees of freedom with no constraints except for the experimentally determined number of molecules per unit cell. The potentials were found to account satisfactorily for the structure, including the symmetry which was derived and not imposed, and the energy of these crystals. An analysis was made of the deviations between the calculated and experimental crystal structures in some compounds in terms of their structural features and in terms of possible weaknesses in the potential functions. Deviations in formamide, along with other compounds, indicated that inclusion of the effect of the lone pair orbitals of oxygen may possibly give a better description of the hydrogen bond energy.

In the previous paper[1] (hereafter referred to as I) a force field for intermolecular interactions in amides was derived through a least-squares fit to experimental structure and heat of sublimation of amide crystals and to dipole moments of some amides. This force field is intended to be part of a general force field for use on biological macromolecules, as well as on other systems. The hydrogen bond was found to be adequately represented by partial charges placed on the atoms C′, O, H_N, and N and nonbonded parameters of the Lennard-Jones type between atom C′, O, and N. The nonbonded parameters for the amide hydrogen H_N were found to be negligible.

In the present work we shall examine how well the force field derived in I fits experimental properties of amide crystals, beyond those used for the optimization in I. To this end, the energies of ten amide crystals were minimized with respect to all degrees of freedom

(1) A. T. Hagler, E. Huler, and S. Lifson, *J. Amer. Chem. Soc.*, **96**, 5319 (1974).