

Modeling Crystal Structures by Numerical Methods: A Tool for Practical Structure Analysis

Wolfgang Linert* and Franz Renz

Institute of Inorganic Chemistry, Technical University Vienna, Getreidemarkt 9, A-1060 Vienna, Austria

Received May 24, 1993*

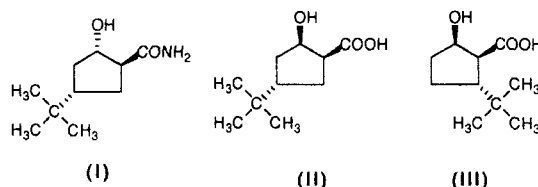
Molecular mechanics calculations on crystals have been performed by using a simplex optimization and a simulated annealing procedure. Empirical Coulomb-Lennard-Jones potentials with an additional set of 10-12 potentials to describe hydrogen bond interactions were used to model crystals (i) with periodic boundary conditions including the interactions of 27 unit cells and (ii) including 729 unit cells (called extended neighbor model). The restrictions imposed by the respective space group were taken into account. As test examples formamide, [1*S*-(1 α ,2 β ,4 β)]-4-(1,1-dimethylethyl)-2-hydroxycyclopentanecarboxamide, [1*S*-(1 α ,2 α ,4 β)]-4-(1,1-dimethylethyl)-2-hydroxy-cyclopentanecarboxylic acid, and [1*S*-(1 α ,2 α ,5 β)]-5-(1,1-dimethylethyl)-2-hydroxy-cyclopentanecarboxylic acid were used. Theoretically predicted structures are found to be in excellent agreement with the experimental structures, if extended neighbor interactions are taken into account.

1. INTRODUCTION

The modeling of inter- and intramolecular interactions by means of molecular mechanics calculations has been strongly advanced in the past decade. Especially the description of molecular interactions in the "gas phase" yields results which are in very good agreement with experimental findings. Similar methods can be used to investigate molecules packed into crystals. In this case it is felt that interactions should be taken into account which reach far out of the unit cell, while common investigations are based on next neighbor cell interactions, i.e. on periodic boundary conditions.¹ This in turn increases the number of local minima of the obtained potential surface so that least squares fitting procedures (applied to hydrocarbons),²⁻⁴ steepest descent, or gradient optimization techniques,⁵⁻¹⁰ or even more refined methods like the simplex technique¹² fail to obtain the absolute minimum. They require *a priori* information of the orientation and position of the molecules in the unit cell, because gradient methods will yield the next available minimum. Even in the latter methods, several trials have to be performed to verify that the absolute minimum has been reached.¹² Many investigations therefore start from already known X-ray structures (used, as the crystal state provides a rich source of information for deriving energy parameters,²⁻¹⁰ lattice dynamics,¹³ and for testing existing pair potentials,¹⁴ etc.). Statistical analysis of known crystal structures may be a solution to the problem.¹⁵ However, it appears as if the problem of finding absolute minima in complicated potential surfaces is often "quietly forgotten".

According to this we compared a simplex method and a statistical annealing approach for optimization of finite crystals, trying to include extended neighbor interactions (up to 729 neighboring unit cells) to optimize the orientation and position of the molecules within the unit cell and, if possible, the length of the three unit cell vectors *a*, *b*, *c*. As examples we have chosen the formamide crystal for which experimental data are available in literature¹⁶ and the compounds¹⁷ [1*S*-(1 α ,2 β ,4 β)]-4-(1,1-dimethylethyl)-2-hydroxy-cyclopentanecarboxamide (I), [1*S*-(1 α ,2 α ,4 β)]-4-(1,1-dimethylethyl)-2-hydroxycyclopentanecarboxylic acid (II), and [1*S*-(1 α ,2 α ,5 β)]-5-(1,1-dimethylethyl)-2-hydroxycyclopentanecarboxylic acid

(III), for which we have kindly been supplied with prepublished crystallographic data for this investigation.



2. METHODS

Formamide crystallizes in the space group $P2_1/c$ ($a = 3.61$ Å, $b = 9.05$ Å, $c = 8.42$ Å, $\beta = 125.39^\circ$).¹⁶ Compound I ($a = 13.25$ Å, $b = 6.99$ Å, $c = 13.30$ Å, $\beta = 113.78^\circ$)¹⁷ and compound II ($a = 16.86$ Å, $b = 6.10$ Å, $c = 10.52$ Å, $\beta = 107.03^\circ$)¹⁷ crystals belong also to this space group. Species III belongs to the space group $P2_1/n$ ($a = 6.17$ Å, $b = 21.75$ Å, $c = 7.89$ Å, $\beta = 105.57^\circ$).¹⁷ Both $P2_1/c$ and $P2_1/n$, being different settings of the same space group, contain 4 molecules/unit cell and have α and γ fixed to 90.0° . Three coordinates are needed to fix the center of mass of the first molecule, and three parameters (the Euler angles) were necessary to define its orientation. According to this a *six-parameter optimization* has been performed. Furthermore a *nine-parameter variation* was performed where the unit cell lengths have also been optimized.

To describe the unit cell, the Cartesian coordinates of one molecule are transformed into crystallographic coordinates by using the formulas of Dunitz.¹⁸ After this step the coordinates of the remaining three molecules were generated by using the symmetry operators¹⁹ of the respective space groups (Table I). Finally the crystallographic coordinates were transformed back into Cartesian coordinates. Due to this techniques it was necessary to store only $3N$ coordinates (N = number of atoms in the molecule) when interaction energies were calculated.

In addition to the molecule-molecule interactions within a single unit cell, first and extended neighbor interactions between the cells were taken into account. Two finite crystal models have been considered, the so called "*periodic boundary condition model*", containing the central unit cell plus 26 neighboring unit cells (i.e. a $3 \times 3 \times 3$ unit cell crystal), and the "*extended crystal model*", including 729 unit cells (i.e. a $9 \times 9 \times 9$ unit cell crystal). The former model refers to a unit

* Author to whom correspondence should be addressed.
* Abstract published in *Advance ACS Abstracts*, September 1, 1993.

Table I. Symmetry Operators of Space Groups $P2_1/c$ and $P2_1/n$

	$P2_1/c$			$P2_1/n$		
molecule 1	X	Y	Z	X	Y	Z
molecule 2	$1-X$	$1-Y$	$1-Z$	$1-X$	$1-Y$	$1-Z$
molecule 3	X	$0.5-Y$	$0.5+Z$	$0.5+X$	$0.5-Y$	$0.5+Z$
molecule 4	$1-X$	$0.5+Y$	$0.5-Z$	$0.5-X$	$0.5+Y$	$0.5-Z$

Table II. Types of Cell-Cell Interactions and the Number of Occurrences in the "Periodic Boundary Condition Model"

no.	translation	no. of cases	no.	translation	no. of cases
1	intra	27	33	$a+b-c$	8
2	a	18	34	$a-b+c$	8
3	b	18	35	$a-b-c$	8
4	c	18	36	$a+b+2c$	4
5	$a+b$	12	37	$a-b+2c$	4
6	$a+c$	12	38	$a+b-2c$	4
7	$b+c$	12	39	$a-b-2c$	4
8	$a-b$	12	40	$a+2b+c$	4
9	$a-c$	12	41	$a+2b-c$	4
10	$b-c$	12	42	$a-2b+c$	4
11	$2a$	9	43	$a-2b-c$	4
12	$2b$	9	44	$2a+b+c$	4
13	$2c$	9	45	$2a+b-c$	4
14	$2a+b$	6	46	$2a-b+c$	4
15	$2a-b$	6	47	$2a-b-c$	4
16	$2a+c$	6	48	$2a+2b+c$	2
17	$2a-c$	6	49	$2a+2b-c$	2
18	$2b+c$	6	50	$2a-2b+c$	2
19	$2b-c$	6	51	$2a-2b-c$	2
20	$a+2b$	6	52	$2a+b+2c$	2
21	$a-2b$	6	53	$2a+b-2c$	2
22	$b+2c$	6	54	$2a-b+2c$	2
23	$b-2c$	6	55	$2a-b-2c$	2
24	$a+2c$	6	56	$a+2b+2c$	2
25	$a-2c$	6	57	$a+2b-2c$	2
26	$2(a+b)$	3	58	$a-2b+2c$	2
27	$2(a-b)$	3	59	$a-2b-2c$	2
28	$2(a+c)$	3	60	$2a+2b+2c$	1
29	$2(a-c)$	3	61	$2a+2b-2c$	1
30	$2(b+c)$	3	62	$2a-2b+2c$	1
31	$2(b-c)$	3	63	$2a-2b-2c$	1
32	$a+b+c$	8			

cell where any interaction or molecule "leaving" the unit cell at one side enters at the opposite side. This is equivalent to the interaction of 26 next neighbor cells with a central unit cell.¹ Cell-cell and intracell energy contributions of this system were built up by translations rather than by explicit periodic boundary conditions to allow for construction of larger systems. First we calculated the molecule-molecule interactions within the reference unit cell. Following this step the cell-cell interactions were obtained by shifting the original coordinates into the neighboring cells. The shift vectors are listed for all 62 different cases in Table II (translations in positive and negative directions are equivalent), which lists the factors by which of the energy contributions had to be multiplied in order to obtain the sum of the 351 cell-cell contributions. Exactly the same translations are used to build up the extended neighbor model (i.e. the $9 \times 9 \times 9$ unit cell crystal) by shifting the above constructed $3 \times 3 \times 3$ model. In principle this recursive method could be used to construct the next larger model (for example a $27 \times 27 \times 27$ system).

In order to obtain absolute values of energy, an Ewald summation or a cutoff radius of at least 250 Å should be performed, avoiding divergence of the Coulombic terms in infinite crystals. If the atomic charges are lower about 0.3 elementary units the convergence radius reduces, up to our experience to some 20–30 Å, provided neutral molecules are considered and the dipole moments of the molecules do not exceed 5 D (i.e. yields convergent Coulomb contributions). This is supported by investigations with α - and β -glycine.^{20,21}

All of the investigated molecules fulfill this condition. As we are only interested in the conformation of the molecules, only intracell, first, and extended neighbor Coulombic interactions were taken into account, neglecting the contributions of more remote neighbors. The latter would contribute only a small constant term to the energy and are independent of the orientation of the molecules.

The geometrical data used to model formamide were identical with those used by Lifson et al.⁹ The geometries of molecules I–III have been obtained by using force field optimization.^{6,7} For intermolecular interactions van der Waals contributions were modeled by using the 1–6–12 (Coulomb–Lennard–Jones) potentials and the 10–12 potentials accounting for the hydrogen bonds.^{22,23} The interaction potential energy for Coulomb–van der Waals interaction, except for hydrogen interactions, is given by

$$U = \sum_i^N \sum_j^N \left(\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6} + \frac{q_i q_j}{r_{ij}} \right) \quad (1)$$

The hydrogen interactions U_H are modeled by

$$U_H = \sum_i^N \sum_j^N \left(\frac{C_{ij}}{r_{ij}^{12}} - \frac{D_{ij}}{r_{ij}^{10}} + \frac{q_i q_j}{r_{ij}} \right) \quad (2)$$

so that $U_{\text{tot.}} = U + U_H$. The radius r_{ij} denotes the distance between the atoms i and j ; A_{ij} , B_{ij} , C_{ij} , and D_{ij} are constants; q_i and q_j denote the charges on atoms i and j . It should be pointed out that according to our experience it is necessary to include the terms modeling hydrogen bonding for geometry optimization.^{22,23} The actual parameters were taken from the literature.²⁴ All atoms were taken into account, because a "united atom" approach may yield incorrect intermolecular distances.²⁵ The atomic charges were extracted from literature⁷ for which the constants A_{ij} , B_{ij} , C_{ij} , and D_{ij} have been set up and if necessary adopted CNDO/2 calculations²⁶ (Table III). AM1 charges are given for comparison. The potentials were already tested before for similar systems^{11,22,23} and by starting the optimization with X-ray diffraction data yielding only small deviations from the experimental data.

The *simplex minimization procedure*¹² can be visualized with the example of a surface in space. At the beginning of the search three points on the x - y plane are chosen that form the regular starting simplex (an equal-sided triangle). The functional value of each of the points is calculated. Now a new simplex is created by mirroring the corner with the highest function value at the remaining base line on the triangle. The procedure is restarted. If the algorithm oscillates (i.e. the mirrored point becomes the highest point of the resulting simplex), the second highest point is mirrored to prevent infinite loops. If it rotates about a minimum, the search is continued with a shortened simplex base line. For the present investigations the procedure was stopped when the change in the total energy was less than 10^{-6} kcal/mol per unit cell. Normally 2500–2800 optimization steps were necessary to reach the minimum. The calculations were repeated with different, stochastically chosen starting geometries until the absolute minimum could not be improved. Between 40 and 50 calculations were enough with six-parameter models. More than 150 trials were done for the nine-parameter model so that up to 420 000 configurations were calculated.

Parallel to the simplex procedure, a homogeneous *simulated annealing algorithm* was used for locating energy minima.^{22,27,28} This algorithm derives its name from a process in condensed matter physics where a solid is heated until its

Table III. Most Important Partial Charges (in Atomic Units) for Atoms and Resulting Dipole Moments of the Molecules (AM1-Charges Given for Comparison)

group	formamide		I		II		III	
	used	AM1	used	AM1	used	AM1	used	AM1
C in C=O	0.445	0.270	0.330	0.313	0.500	0.328	0.500	0.348
O in C=O	-0.369	-0.351	-0.500	-0.353	-0.500	-0.358	-0.500	-0.295
X in CO-X ^a	-0.415	-0.475	-0.570	-0.365	-0.685	-0.278	-0.685	-0.282
O in C-OH			-0.685	-0.626	-0.685	-0.318	-0.685	-0.333
C in C-OH			0.285	0.041	0.285	0.043	0.285	0.035
H in C-OH			0.400	2.204	0.400	0.203	0.400	0.214
H in CO-XH _n ^a	0.192 ^b	0.232 ^b	0.370	0.200	0.400	0.233	0.400	0.240
	0.176 ^c	0.227 ^b						
dipole moment ^d		3.47		1.89		2.35		3.1

^a X = O in the carboxylic acid; X = N in the carboxamide. ^b Cis hydrogen in amide. ^c Trans hydrogen in amide. ^d Dipole moments in debye. Experimental value for formamide 3.37 D.

particles can move freely. The temperature is gradually decreased, and the particles come to rest in their lowest energy ground states, provided the initial temperature is sufficiently high and the cooling speed is sufficiently low. This can be simulated by a Monte Carlo process that is controlled by the so called "metropolis criterion". Provided the energy (in our case $U_{\text{tot.}}$) in every state is known, this criterion selects random walk movements, stepping from a configuration Π_k to a adjacent configuration Π_{k+1} . The step is accepted with the probability

$$P(\Pi_k, \Pi_{k+1}) = \min(1; \exp(-\Delta U_{\text{tot.}}/RT)) \quad (3)$$

$U_{\text{tot.}}$ represents the intermolecular interactions and T the temperature of the system. In the limit $T \rightarrow 0$ this approach reduces to a simple Monte Carlo method. The simulated annealing runs were started at 241 K. The thermal motion was simulated by a random walk restricted to a maximum of 0.2 Å for the Cartesian coordinates of a molecule and 10° for its Eulerian angles. Markov chains at a fixed temperature with a length of 1200 trial moves were calculated; i.e. a "homogeneous simulated annealing algorithm" was used. The temperature was decreased after such an Markov chain (assuming that thermal equilibrium is approached) according to a predefined cooling program by about 2–5%. This yielded some 500–1500 temperature steps, depending on the criterion that an acceptance ratio reached zero for more than ten consecutive Markov chains (i.e. the system was "frozen"). Therefore several between 600 000 and 1 800 000 configurations have been calculated. The configurations generated by the this procedure were refined by a numeric-gradient search method, which usually yielded a further decrease in energy by up to 0.2 kcal/(mol·unit cell). Simplex and simulated annealing calculations were performed on IBM ES9021-720 and Siemens VP50EX machines.

3. RESULTS AND DISCUSSION

Numerical optimization of a crystal structure^{1–4} by using empirical potential functions is similar to the task of determining the structure of a single molecule,^{5,6} but the number of parameters to be optimized is much higher. For molecules with N atoms we have to optimize a maximum of $3N - 6$ parameters, whereas in crystals the number of parameters is $3N - 6 + 12 = 3N + 6$, provided that the crystallographic coordinates of just one molecule per unit cell have to be optimized, and the remaining molecules are generated by using the symmetry operations of the space group. The relationship $3N - 6$ equals the number of internal coordinates; six parameters are needed to fix the molecule in the unit cell, and six parameters are necessary to specify the length and angles

Table IV. Lowest Obtained Minimum Energies (kcal/(mol·unit cell)) for Formamide As Obtained from the Simplex Optimization Procedure As Obtained from Various Starting Geometries

	periodic boundary condition model six-parameter variation	extended neighbor model	
		six-parameter variation	nine-parameter variation
1	-31.000	-44.990	-50.080
2	-30.374	-44.737	-49.956
3	-30.222	-44.270	-49.704
4	-29.985	-43.989	-49.696
5	-29.159	-43.907	-49.626
6	-28.793	-43.552	-48.922
7	-28.604	-43.230	-48.641
8	-28.330	-43.174	-48.585
9	-28.100	-43.148	-48.367
10	-28.011	-42.948	-47.741
11	-27.952	-42.830	-47.719
12	-27.730	-42.452	-47.070
13	-27.341	-42.411	-47.048
14	-26.289	-42.300	-47.030
15	-26.196	-42.181	-46.644
16	-25.937	-42.170	-45.970
17	-25.570	-41.820	-45.840

between the unit cell axes. If the molecules are assumed to be rigid (internal coordinates have been optimized by a force field method or may be approximated by using standard bond lengths and bond angles) and if the length and the orientation of the unit cell vectors also do not vary (these can easily be determined by using simple X-ray techniques), there are six parameters left, which have to be optimized, namely, the three relative coordinates specifying the center of mass and the three Euler angles fixing the orientation of the molecule.^{22,23}

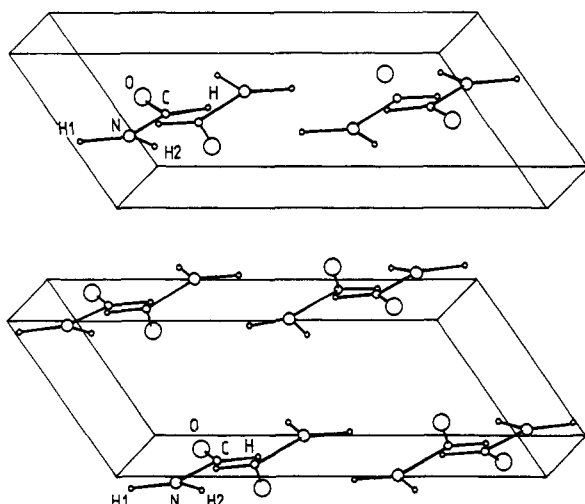
Therefore at least six parameters have to be optimized in the case where the unit cell parameters are given. Allowing for a variation of the unit cell size, we have to vary nine parameters for the present system. According to this both a *six-parameter* and a *nine-parameter optimization* have been performed.

To optimize the crystal structure of the test molecules, three schemes were followed, namely, (i) the *periodic boundary condition model* (i.e. taking into account interactions between 27 unit cells) with the variation of six parameters and fixed unit cell dimensions, (ii) the *extended neighbor approximation* (i.e. taking into account interactions between 729 unit cells) with the variation of six parameters using fixed unit cell parameters, and (iii) the *nine-parameter extended neighbor approximation* (interaction of 729 unit cells) in which the unit cell lengths were also optimized.

The calculated minimum energies for formamide resulting from the simplex method are listed in Table IV. The periodic boundary condition model yielded less favorable minimum energies than the extended neighbor model. As expected the

Table V. Experimental and Calculated Crystallographic Coordinates of Formamide

atom		periodic boundary condition model	extended neighbor model ^a	experiment
		six-parameter variation	six-parameter variation	
C	X	0.521	0.077	0.593
	Y	0.304	0.071	0.057
	Z	0.394	0.250	0.258
O	X	0.537	0.192	0.717
	Y	0.159	-0.058	-0.069
	Z	0.267	0.251	0.248
N	X	0.494	-0.080	0.421
	Y	0.159	0.165	0.157
	Z	0.408	0.104	0.119
H	X	0.528	0.094	0.611
	Y	0.353	0.110	0.087
	Z	0.486	0.350	0.362
H1	X	0.487	-0.144	0.341
	Y	0.137	0.253	0.293
	Z	0.511	0.132	0.154
H2	X	0.485	-0.114	0.398
	Y	0.095	0.139	0.140
	Z	0.322	-0.009	0.007

^a X-coordinate shifted by 0.5 unit cell length.**Figure 1.** Experimental crystal structure (upper part) and calculated crystal structure (lower part) of formamide resulting from the extended neighbor (six-parameter) model. The calculated relative X coordinates are shifted by $1/2$ unit with respect to those in the experimental crystal structure.

lowest energy minima were obtained by the extended neighbor model. Comparison of the calculated and the experimental¹⁶ crystallographic coordinates (Table V) indicate that the extended neighbor (six parameter) model yielded satisfactory coordinates, whereas the periodic boundary condition model did not. The periodic boundary condition model accounts inadequately for the cell-cell interactions. The nine-parameter extended neighbor model yielded incorrect values for the unit cell lengths ($a = 4.86$, $b = 6.93$, and $c = 6.87$ Å) associated with wrong positions of the molecules. There are obviously too many minima in the nine-parameter extended neighbor model. It appears to be hardly possible to find the absolute minimum by means of the simplex optimization technique in this case.

The calculated and the predicted crystal structures of formamide are shown in Figure 1. The numerical procedure shifted the molecule by $1/2$ unit along the X axes. This fact is not a shortcoming of the minimization procedure but simply indicates that there is no unique way to define a unit cell (Table I).

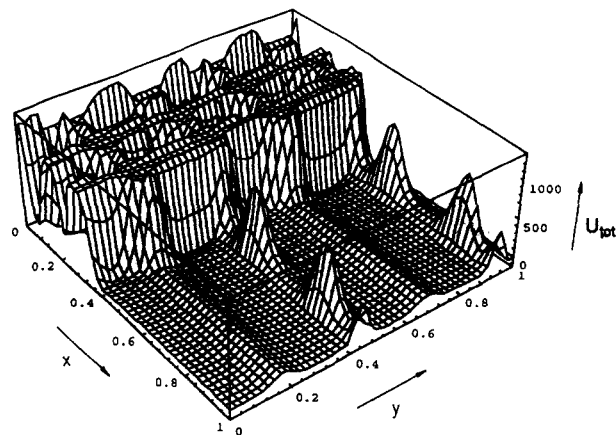
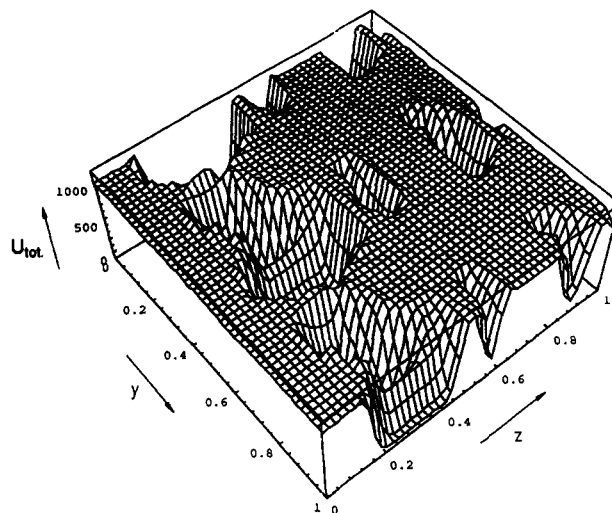
**Figure 2.** Projection of the energy surface for a variation of the point of gravity of formamide in the x versus y plane (in crystallographic units) with the remaining four parameters (i.e. z, ϕ, ψ, θ) taken at optimal positions to allow graphical illustration. The absolute minimum is found at $x = 0.402$ and $y = 0.044$ (respectively, $y = 0.544$ because of symmetry). ΔE is truncated at 1000 kcal/(mol-unit cell) when larger than 1000 kcal/(mol-unit cell).**Figure 3.** Projection of the energy surface for a variation of the point of gravity of formamide in the y vs z plane (in crystallographic units) with the remaining four (i.e. x, ϕ, ψ, θ) parameters taken at optimal positions to allow graphical illustration. The absolute minimum is found at $y = 0.044$ (respectively, $y = 0.544$ because of symmetry) and $z = 0.232$. ΔE is truncated at 1000 kcal/(mol-unit cell) when larger than 1000 kcal/(mol-unit cell).

Table III contains the 17 lowest energies, obtained from different starting geometries using the simplex method. The lowest refers to the absolute minimum, the others refer to local minima in the neighborhood of the absolute minimum. Generally we assumed that we have found the absolute minimum when at least six arbitrarily chosen starting geometries yielded this (lowest) value. The given values show, however, that the absolute minimum is to be found with increasing difficulties when the number of parameters (or even the atoms in the unit cell) is raised. To illustrate this, Figures 2-4 are given, showing the energy landscape along two variables with the remaining four variables being kept constant at the absolute minimum. Obviously, the real energy surfaces are much more complicated; all six variables had to be optimized simultaneously.

The geometry optimization for the variables x and y (see Figure 2) with well chosen starting positions ($x = 0.8$ and $y = 0.3$) finds the minimum energy in acceptable times with methods like simplex or gradient optimization. But other variables shown in Figures 3 and 4 have so many well separated

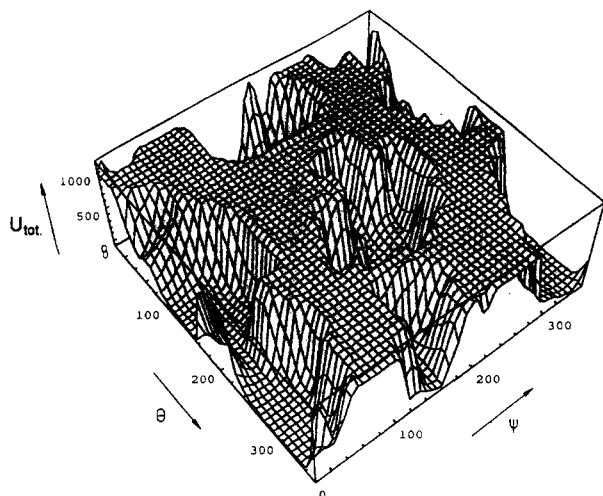


Figure 4. Projection of the energy surface for a variation of the point of gravity of formamide in the θ vs ψ plane (in degrees) with the remaining four (i.e. x, y, z, ϕ) parameters taken at optimal positions to allow graphical illustration. The absolute minimum is found at $\psi = 291.530$ and $\theta = 19.155$. ΔE is truncated at 1000 kcal/(mol-unit cell) when larger than 1000 kcal/(mol-unit cell).

local minima that special procedures like simulated annealing are necessary to find the minimum energy. Considering that formamide is a very simple molecule, one can imagine that simplex or gradient methods cannot possibly lead to the absolute minimum without further geometrical restrictions. In fact complicated force field calculations have been performed using the experimental crystal structure as starting parameters with the aim of testing force field potentials.⁹

On the basis of the results already explained for formamide, we have performed calculations on the cyclopentane derivatives I–III. We have used the extended neighbor model approximation with a six-parameter variation (i.e. fixed unit cell parameters) combined with the simulated annealing procedure. As in the case of formamide excellent agreement with the experimental crystal structures¹⁷ is found (Figures 5–7).

Minimum energies resulting from the simulated annealing procedure are given in Table VI. The energy minimum obtained for formamide is the same as that one resulting from the simplex method. The energies obtained for cyclopentane derivatives are lower by a factor of 3–4 (i.e. more stable) than the energies obtained for formamide. This is in agreement with the melting points of the compounds (formamide, 2.55 °C; cyclopentane derivatives, 90–120 °C¹⁷) and originates from the lower number of pair interactions due to different numbers of atoms in the case of formamide (6 atoms/molecule; 4 molecules/unit cell) compared to I (32 atoms/molecule; 4 molecules/unit cell), II (31 atoms/molecule; 4 molecules/unit cell), and III (31 atoms/molecule; 4 molecules/unit cell). The number of polar functional groups per unit cell is almost equal (see partial charges in Table VI) for all compounds. However, the cyclopentane derivatives contain many more nonpolar groups through the carbon ring and the *tert*-butyl group. The hydrogen bonding is strongest in formamide and least pronounced in II. The pronounced hydrogen bonding of the carboxylic group because of sterical screening of the *tert*-butyl group leads furthermore to higher energies being obtained for III than for I and II.

4. CONCLUSIONS

The present technique may be used to help in the interpretation of X-ray reflection data when the actual crystal structure is not known. This is of importance, especially when

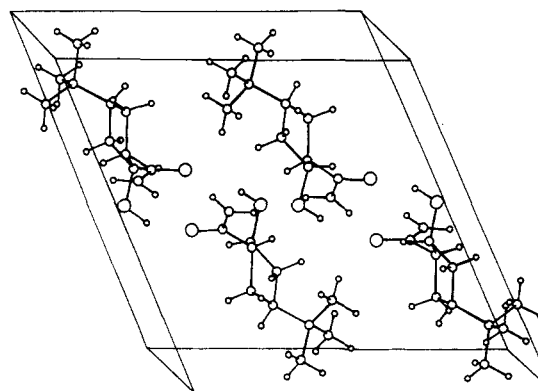
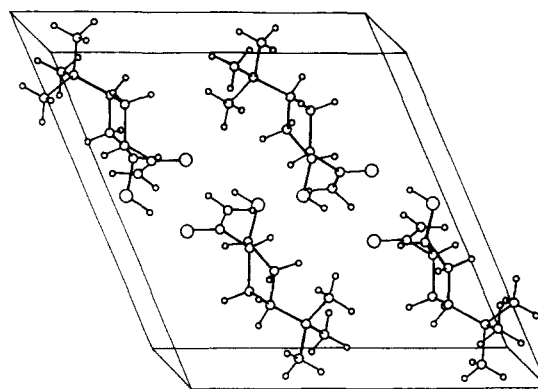


Figure 5. Experimental crystal structure (upper part) and calculated crystal structure (lower part) of $(1\alpha,2\beta,4\beta)$ -4-(1,1-dimethylethyl)-2-hydroxy-cyclopentanecarboxamide (I) resulting from the extended neighbor (six-parameter) model.

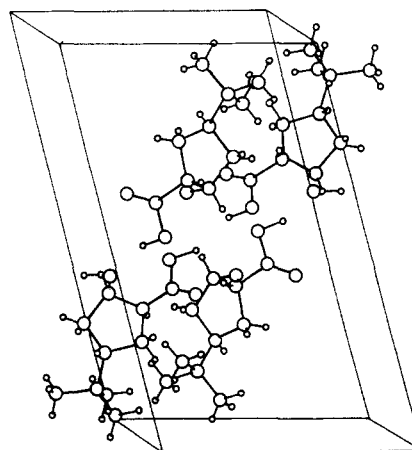


Figure 6. Calculated crystal structure of $(1\alpha,2\alpha,4\alpha)$ -4-(1,1-dimethylethyl)-2-hydroxycyclopentanecarboxylic acid (II) resulting from the extended neighbor (six-parameter) model.

only powder reflection data are available. In most cases it is relatively simple to measure the unit cell parameters which support the presented method, and it is much more difficult to place the molecule within the unit cell. The present approach helps to solve this problem. The method might also be used to explain isostructuralism and polymorphism. Cartesian coordinates of rigid molecules can be generated by using standard values¹⁹ of bond lengths and bond angles or by using standard quantum-mechanical or semi-empirical methods.²⁶ Then the minimization procedure may be performed. Taking the main pair interaction potentials into account and realizing that at least one, better sphere of neighbor-cell interactions in addition to second neighbor interactions has to be considered,

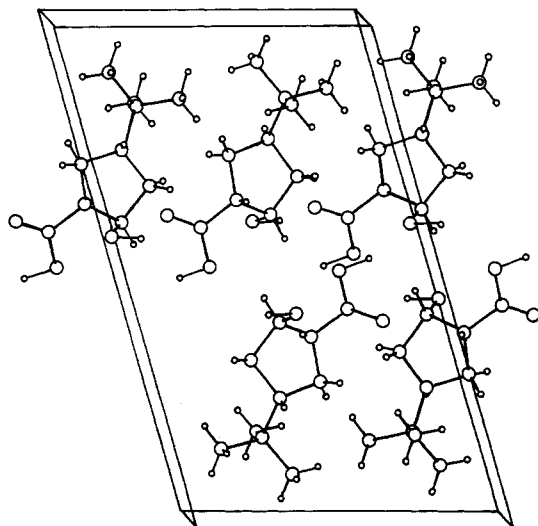


Figure 7. Calculated crystal structure of (1 α ,2 α ,4 β)-4-(1,1-dimethylethyl)-2-hydroxycyclopentanecarboxylic acid (III) resulting from the extended neighbor (six-parameter) model.

Table VI. Calculated Minimum Energies (U_{tot}) and Contributions to U_{tot} (kcal/(mol-unit cell)) and Mean Deviations between Calculated and Experimental Structures Obtained from the Extended Neighbor Approximation Using the Six-Parameter Model

compd	U_{tot}	r^{-1}	r^{-6}	r^{-12}	U_{H}	$\overline{\Delta x^a}/\text{\AA}$
formamide	-44.99	-34.1	-80.2	40.3	-6.41	0.10
I	-152.6	-93.1	-96.9	41.6	-4.14	0.23
II	-158.6	-90.6	-101.9	36.9	-2.90	0.17
III	-130.9	-65.7	-118.4	56.8	-3.55	0.13

$a \overline{\Delta x} = [\sum_i^n |x_{\text{calc}}^i - x_{\text{exp}}^i|]/n.$

complicated energy surfaces occur so that simulated annealing or genetic algorithms^{22,23} should be used to find lowest energy minima.

ACKNOWLEDGMENT

We thank Prof. A. Kalman (Budapest), Dr. I. Lukovits (Budapest), and Dr. Gy. Argay (Budapest) for suggesting the optimization of crystal structures, for supplying experimental results, and for many helpful discussions. Thanks are due to the "Fonds zur Förderung der Wissenschaftlichen Forschung" (Project 8795) for financial support.

REFERENCES AND NOTES

- (1) Allen, M. P.; Tildesley, D. J. *Computer Simulations of Liquids*; Clarendon Press: Oxford, U.K., 1987.
- (2) Williams, D. E. Crystal Packing of Molecules. *Science* **1965**, *147*, 605. Williams, D. E. Crystal Structure of Dibenzoylmethane. *Acta Crystallogr.* **1966**, *21*, 340. Williams, D. E. Computer Calculation of Molecular Crystal Structures. *Science* **1968**, *159*, 645.
- (3) Williams, D. E. Nonbonded potential parameters derived from crystalline hydrocarbons. *J. Chem. Phys.* **1967**, *47*, 4680. Williams, D. E. Calculating molecular crystal structures. *Acta Crystallogr. Sect. A* **1969**, *25*, 464. Williams, D. E.; Starr, L. T. Calculation of the crystal structures of hydrocarbons by molecular packing analysis. *Comput. Chem.* **1977**, *1*, 173.
- (4) Kitaigorodskii, A. I. *Molecular Crystal and Molecules*; Elsevier: Amsterdam, 1973.
- (5) Lifson, S.; Warshel, A. Consistent force field for calculations of conformations, vibrational spectra, and enthalpies of cycloalkane and *n*-alkane molecules. *J. Chem. Phys.* **1968**, *49*, 5116.
- (6) Lii, J. H.; Allinger, N. L. Molecular Mechanics. The MM3 Force Field for Hydrocarbons. 3. The van der Waals' Potentials and Crystal Data for Aliphatic and Aromatic Hydrocarbons. *J. Am. Chem. Soc.* **1989**, *111*, 8576.
- (7) Jorgensen, W. L.; Tirado-Rives, J. The OPLS Potential Functions for Proteins. Energy Minimization for Crystals of Cyclic Peptides and Crambin. *J. Am. Chem. Soc.* **1988**, *110*, 1657.
- (8) Hagler, A.; Huler, E.; Lifson, S. Energy Functions for Peptides and Proteins. I. Derivation of a Consistent Force Field Including the Hydrogen Bond from Amide Crystals. *J. Am. Chem. Soc.* **1974**, *96*, 5319.
- (9) Lifson, S.; Hagler, A. T.; Dauber, P. Consistent Force Field Studies of Intermolecular Forces in Hydrogen-Bonded Crystals. 1. Carboxylic Acids, Amides, and the C=O...H-Hydrogen Bond. *J. Am. Chem. Soc.* **1979**, *101*, 5111.
- (10) Hagler, A. T.; Lifson, S.; Dauber, P. J. Consistent Force Field Studies of Intermolecular Forces in Hydrogen-Bonded Crystals. 2. A Benchmark for the Objective Comparison of Alternative Force Fields. *J. Am. Chem. Soc.* **1979**, *101*, 5122.
- (11) Linert, W.; Lukovits, I. The Use of the Isokinetic Relationship and Molecular Mechanics to Investigate Molecular Interactions in Inclusion Complexes of Cyclodextrins. *Chem. Phys.* **1989**, *139*, 441.
- (12) Nelder, J. A.; Mead, R. A Simplex Method for Function Minimization. *Comput. J.* **1965**, *7*, 308.
- (13) Huler, E.; Warshel, A. On the Consistent Calculation of Lattice Dynamics Using Semiempirical Potential Functions. *Chem. Phys.* **1975**, *8*, 239.
- (14) Morse, M. D.; Rice, S. A. Tests of Effective Pair Potentials for Water: Predicted Ice Structures. *J. Chem. Phys.* **1982**, *76*, 650.
- (15) Gavezzotti, A. Generation of Possible Crystal Structures from the Molecular Structure of Low-Polarity Organic Compounds. *J. Am. Chem. Soc.* **1991**, *113*, 4622.
- (16) Ottersen, T. On the Structure of Peptide Linkage. The Structures of Formamide and Acetamide and -165 °C and an ab initio Study of Formamide, Acetamide and *N*-Methyl-Formamide. *Acta Chem. Scand.* **1975**, *A29*, 939.
- (17) Bernáth, G.; Svoboda, M. Stereochemical Studies XIII: Cyclic Aminoalcohols and Related Compounds-VI: Synthesis of Four 2-Amino-4-*t*-butylcyclopentanols. *Tetrahedron* **1972**, *28*, 3475. Bernáth, G. To be published.
- (18) Dunitz, J. D. *X-Ray Analysis and the Structure of Organic Molecules*; Cornell University Press: Ithaca, NY, 1979.
- (19) *International Tables for X-Ray Crystallography*; Hahn, T., Ed.; International Union of Crystallography, Reidel: Dordrecht, The Netherlands, 1983; Vol. A.
- (20) Momany, F. A.; Carruthers, L. M.; Scheraga, H. A. Intermolecular Potentials from Crystal Data. IV. Application of Empirical Potentials to the Packing Configuration and Lattice Energies in Crystals of Amino Acids. *J. Phys. Chem.* **1974**, *78*, 1621.
- (21) Derissen, J. L.; Smit, P. H. Voogd, J. On the Calculation of the Electrostatic Lattice Energies of α -, β -, and γ -Glycine. *J. Phys. Chem.* **1977**, *81*, 1474.
- (22) Linert, W.; Margl, P.; Lukovits, I. Numerical Minimization Procedures in Molecular Mechanics: Structural Modeling of the Solvation of β -Cyclodextrin. *Comput. Chem.* **1992**, *16*, 61.
- (23) Linert, W.; Margl, P.; Renz, F. Solute-solvent Interactions between Cyclodextrin and Water: A Molecular Mechanical Study. *Chem. Phys.* **1992**, *161*, 327.
- (24) Weiner, S. J.; Kollman, P. A.; Case, D. A.; Singh, U. C.; Ghio, C.; Allagona, G.; Profeta, S., Jr.; Weiner, P. A New Force Field for Molecular Mechanical Simulation of Nucleic Acids and Proteins. *J. Am. Chem. Soc.* **1984**, *106*, 765.
- (25) Lukovits, I. A Simplified Normal Coordinate Analysis of Vibration of Adsorbed Molecules. *Chem. Phys. Lett.* **1989**, *155*, 297.
- (26) Pople, J. A.; Beveridge, D. L. *Approximate Molecular Orbital Theory*; McGraw-Hill: New York, 1970.
- (27) Van Laarhoven, P. J. M.; Aarts, E. H. L. *Simulated Annealing: Theory and Applications*; Reidel: Dordrecht, Boston, Lancaster, Tokyo, 1987.
- (28) Wille, L. T. Minimum-energy configuration of atomic clusters: new results obtained by simulated annealing. *Chem. Phys. Lett.* **1987**, *133*, 405.