

# CALCULATION OF THE CRYSTAL STRUCTURES OF HYDROCARBONS BY MOLECULAR PACKING ANALYSIS\*

DONALD E. WILLIAMS and THOMAS L. STARR

Department of Chemistry, University of Louisville, Louisville, KY 40208, U.S.A.

(Received 14 November 1976)

**Abstract**—Several model potentials for H...H, H...C, and C...C nonbonded interactions were tested by calculation of the crystal structures of 18 assorted hydrocarbons. The (exp-6-1) potential was significantly superior to the (exp-6) potential, supporting the existence of net coulombic charges of about 0.15 electron unit on hydrogen in aromatics. The (exp-6-1) potential derived by the predicted structural fit derivation method was found to give about the same goodness of fit as the usual force fit derivation method. However, the force fit derived potential was more transferable to the calculation of lattice vibrational frequencies. The EPEN potential was found inferior in this application because of its poorer fit to the n-pentane and n-hexane crystal structures and its excessive requirements for computational time.

## INTRODUCTION

The structure of a molecular crystal is usually separable into two parts. The first part is the molecular structure, which is often identical to or at least very close to the structure of the molecule in the gas phase. The second part is the packing structure, which is unique to the crystal structure. The molecular structure is determined by strong interatomic forces (bonds), with their associated stretching and bending force constants, plus the weaker forces of intramolecular nonbonded interactions. In contrast, the packing structure is determined entirely by the weak nonbonded interactions. If the different parts of the molecule are sufficiently separated by intervening bonds, it is believed that nonbonded intramolecular forces should be expressible by the same potential functions as nonbonded intermolecular forces.

In the practice of determination of the nonbonded potentials, it is found that much shorter nonbonded interatomic contacts are found in the intramolecular case. There may also be interactions between the nonbonded parameters and bending and stretching force constants, as for example in a 1,3 nonbonded interaction. A similar problem can occur with intermolecular nonbonded interactions if the molecules are not sufficiently rigid. In either case a careful selection of reference structures can be made to minimize the ambiguity of definition of the nonbonded potentials. In the case of crystal structures, it is comparatively easy to choose a set of reference basis structures in which the molecules are effectively rigid, eliminating any possibility of significant interaction with intramolecular stretching, bending or twisting modes.

## THE PARAMETERIZATION OF THE NONBONDED POTENTIALS

A simple pairwise additive model is desired. Simplicity is a practical necessity if the potentials are to be used for large molecules. For neutral atoms, the London dispersion energy has an inverse sixth power dependence on the interatomic distance. This is the first term of a

series in  $r^{-n}$ , with  $n$  even. The next term,  $r^{-8}$ , might contribute perhaps 30% of the nonbonded energy at the potential minimum. But the shape of the curve is little changed, and an empirically adjusted increase in the  $r^{-6}$  coefficient can fairly well accommodate the higher terms.

If the molecule has net atomic charges, coulombic terms must be added to the potential. Even simple hydrocarbons are known to have significant net atomic charges. These charges, can easily exceed 0.1 electron unit. The coulombic contribution to the lattice energy of a hydrocarbon crystal can easily exceed 10% of the total lattice energy, if the potential field is parameterized for optimum net atomic charges.

At very close distances, as filled electron shells overlap, interatomic repulsion sets in. Various functions have been used to describe this repulsion: Morse potentials,  $r^{-12}$  potentials, or more generally  $r^{-n}$  ( $n > 6$ ) potentials, or exponential potentials have been used. The most accurate studies of repulsion between small molecules (e.g. He-He, H<sub>2</sub>-H<sub>2</sub>) indicate that the exponential form gives the most accurate description of the repulsion.

Thus our model potential is an (exp-6-1) potential. If  $r_{jk}$  is a nonbonded interatomic distance,

$$V_{jk} = B_{jk} \exp(-C_{jk}r_{jk}) - A_{jk}r_{jk}^{-6} + q_jq_k e^{-1}$$

where  $B_{jk}$ ,  $C_{jk}$ ,  $A_{jk}$ , and  $q_i$  may be empirically adjusted to give the best fit to structural data. As a check on the desirability of including net atomic charges in the model, the potential parameters were optimized by least squares both with zero net charges and with variable net charges.

Obviously the geometric-mean combining law applies to the net atomic charges  $q_i$ . It is a great simplification to also use the geometric-mean combining law for the other parts of the potential. If we do this we have

$$V = b_\alpha b_\beta \exp[-(c_\alpha + c_\beta)r] - a_\alpha a_\beta r^{-6} + q_\alpha q_\beta r^{-1}.$$

In this equation we use the subscripts  $\alpha$  and  $\beta$  to indicate classification of the empirical coefficients into atomic types. In the benzene structure, for example, 47 different coefficients would be required if no classification were possible. If we classify all C (and all H)

\*This work supported by NIH Research Grant GM16260 from the Institute of General Medical Sciences.

together, there are only 7 coefficients;  $b_H$ ,  $b_C$ ,  $c_H$ ,  $c_C$ ,  $a_H$ ,  $a_C$ , and  $q_H$ . The requirement of molecular electrical neutrality eliminates  $q_C$  as an adjustable parameter ( $q_C = -q_H$  in this case). Such a classification is highly desired to enable transferability of the potentials between structures and to new structures.

The repulsive part of the potential is doubly parameterized, and there exists a high correlation between numerical values of  $b$  and  $c$ . We use a procedure of setting  $c$  through use of the structure of the element. In the case of carbon, this structure is graphite (Crowell, 1958) and the value of  $C \approx 2c = 3.60 \text{ \AA}^{-1}$ . For hydrogen, we use quantum mechanical calculations (Williams, 1965) of  $H_2-H_2$  repulsion, where  $C = 3.74 \text{ \AA}^{-1}$ . Because of the high correlation between  $b$  and  $c$ , small errors in the choice of  $c$  can be compensated by an empirical shift in the value of  $b$ .

For hydrogen, all of the electron density is involved in bonding. This situation results in a significant shift of the apparent center of electron density into the bonding region. This shift is calculated and measured to be about  $0.07 \text{ \AA}$ . For heavier atoms (carbon, in this instance) this bonding shift is considered negligible.

The above treatment neglects thermal effects. The nonbonded potentials may be used to calculate the normal lattice frequencies. Conversely, the lattice frequencies may be used to define the potential. Many workers have calculated such lattice frequencies for various structures, but rarely have researchers fitted the potentials to the frequencies. We have found that the addition of lattice-frequency observational equations does not change the structurally-determined empirical potential very much (Starr & Williams, 1977). As far as the struc-

tural data is concerned the model incorporates the effect of an averaged thermal motion into the potential parameters.

#### THE REFERENCE STRUCTURES

A set of 18 well-determined hydrocarbon crystal structures was selected for the test of the model potential. Figure 1 shows the molecular structure of each compound. The molecules are all rather rigid, and the crystal packing exhibits a wide variety of symmetry (or lack of symmetry). Half of the structures are aromatic and half are saturated. Again, for transferability reasons we assign the same parameters to all types of carbon, and to all types of hydrogen, except for coulombic effects.

For the net atomic charges we use the conditions of C, CH,  $CH_2$ , or  $CH_3$  group neutrality and constant C-H bond moment. If we define the charge separation parameter,  $\Delta e$ , we have

$$\text{for CH } q_H + q_C = 0 \quad q_H - q_C = \Delta e$$

$$\text{for } CH_2 \quad 2q_H + q_C = 0 \quad q_H - q_C = \Delta e$$

$$\text{for } CH_3 \quad 3q_H + q_C = 0 \quad q_H - q_C = \Delta e$$

and  $q_H = \Delta e/2$ ,  $q_H' = \Delta e/3$ , and  $q_H'' = \Delta e/4$ . Note that this scheme does not place equal charges on all hydrogens, and that it avoids an excessively large negative charge on C in the methyl group.

#### THE CRYSTAL LATTICE SUMS

The repulsive part of the potential is essentially convergent by considering nonbonded interatomic contacts up to  $6 \text{ \AA}$  (about 500 contacts for the benzene structure). At the  $6 \text{ \AA}$  distance, the  $r^{-6}$  potential is only about 80%

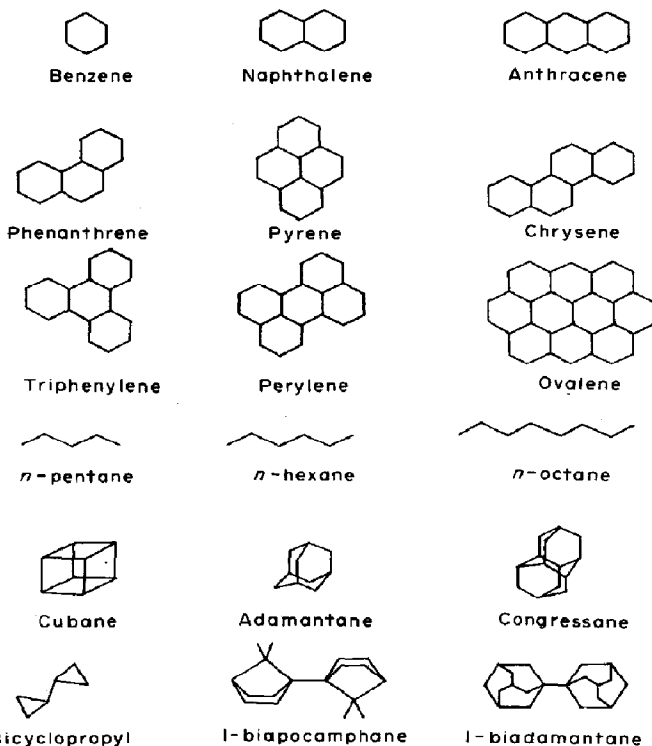


Fig. 1. The molecular structures of the 18 compounds whose crystal structures provided the experimental data for derivation of the hydrocarbon nonbonded potentials.

convergent, while the  $r^{-1}$  potential is even much less convergent. Increasing the summation limit to 15 Å (about 9000 contacts) reduces the truncation error of the  $r^{-6}$  terms to 1%. The  $r^{-1}$  sum is still poorly convergent at 15 Å. Molecular packing analysis can be expensive of computer time if such large summation limits are used.

Fortunately, a mathematical method (Williams, 1971) is available to drastically improve the convergence of the  $r^{-6}$  and  $r^{-1}$  lattice sums. This method multiplies the terms by an appropriate convergence function (in practice, an incomplete gamma function) and takes the Fourier transform of the remainder. The Fourier transform is also rapidly convergent. Using this convergence acceleration technique, a 6 Å sum limit can yield 99% accuracy in the lattice sum.

#### DERIVATION OF THE MODEL POTENTIALS

Three optimized model potentials were derived from the 18 observed hydrocarbon crystal structures. Many workers neglect net atomic charges in these compounds.

Potential Set I does not include any net atomic charges. It corresponds to Potential Set C of the Williams (1970) reference. The geometric-mean combining law was relaxed in this case to allow  $B_{CH}$  to vary independently. A better fit to the experimental data was obtained with this seven parameter potential, as compared to a six parameter potential which assumed the geometric-mean combining law. This procedure also equalized the number of adjustable parameters for Potentials I, II, and III, since the additional variation of  $B_{CH}$  in Potential I coincided with the nonvariation of  $\Delta e$  in that potential. As mentioned above the exponents ( $C_{\alpha\beta}$ ) were fixed externally, so that each calculation optimized five adjustable parameters.

The optimum values of the nonbonded potential parameters were derived by minimization of the equilibrium forces in the crystal. That is, the residual function

$$R(A_{HH}, A_{CC}, B_{HH}, B_{CH}, B_{CC}, \Delta e) = \sum_{\text{structures}} w_i \left[ \frac{\partial V}{\partial p_i} \right]^2,$$

where  $p_i$  is a structural parameter such as a lattice constant or a molecular rotation angle in the crystal, and  $w_i$  is an appropriate weighting factor. In Potential Set I,  $\Delta e$  was set to zero. Since there were 118 structural parameters and 5 adjustable parameters, there was a good ratio of observational equations to the number of variables.

Obviously  $R$  goes to zero as  $V$  goes to zero, so it is necessary to normalize the scale of  $V$ . This was done by requiring the calculated lattice energies of benzene (a representative aromatic structure) and n-hexane (a representative saturated structure) be equal to the observed heats of sublimation at absolute zero temperatures. Incidentally, a good correspondence was obtained to the other unfitted heats of sublimation.

Potential II is similar to I, except that  $\Delta e$  was allowed to vary, and  $B_{CH}$  was set equal to the geometric mean of  $B_{CC}$  and  $B_{HH}$ . The optimum value of  $q_H$  (CH group) was found to be about 0.15 electron unit. Potential II differs from an earlier potential (Williams, 1974) because a different weighting scheme was used. The details of the new weighting scheme are given by Starr & Williams (1977).

For Potential III the fitting procedure suggested by Hagler & Lifson (1974) was used (predicted parameter

shift method). In this procedure the function minimized was

$$R = \sum w_i (\Delta p_i)^2$$

where  $\Delta p_i$  is the predicted shift in the structural parameter given by

$$\Delta p = -H^{-1} \nabla V.$$

The gradient vector  $\nabla V$  is obtained as for Potentials I and II. The Hessian matrix,  $H$ , is the matrix of the second derivatives of  $V$  with respect to the structural parameters. The predicted shift vector  $\Delta p$  is thus found by a first-order Taylor series expansion. Normalization of the potential functions is less critical with this method. For comparison purposes, Potential III was also normalized to the benzene and n-hexane structures. Additional details of the potential derivations are reported elsewhere (Starr & Williams, 1977). Table 1 shows the values of the empirical potential parameters.

The potentials are calibrated only down to the shortest observed nonbonded contacts between molecules in the reference crystals. These shortest distances are approx 2.5 Å for H...H, 2.8 Å for C...H, and 3.5 Å for C...C. Thus, it would be dangerous to assume that the potentials are reliable at shorter distances, such as commonly encountered in molecular strain caused by intramolecular repulsions at short distances. The potentials are also not calibrated to dipole moment data.

#### COMPUTER TESTS OF THE MODEL POTENTIALS

The given three model potentials were tested by finding the calculated equilibrium structure of all 18 hydrocarbons. In this test, rather than hold the structural parameters at their observed values and vary the potentials, the optimum potentials were held fixed and the structural parameters varied until the equilibrium forces were zero. The same Taylor's series expansion formula shown above was used; this is an example of the Newton-Raphson minimization technique. In the few cases where the Hessian was not positive definite, steepest descents was used instead. Table 2 shows the calculated shifts away from the observed values for each of the 118 structural parameters.

Table 1. Optimized values (except C) of the potential parameters. Potentials I and II were derived by the force fit method\*

Parameter	I	II	III
$A_{HH}$	102	136	153
$B_{HH}$	9080	11,677	14,622
$C_{HH}$	3.74	3.74	3.74
$A_{HC}$	467	573	642
$B_{HC}$	35,600	65,485	84,788
$C_{HC}$	3.67	3.67	3.67
$A_{CC}$	2140	2414	2695
$B_{CC}$	300,000	367,250	490,290
$C_{CC}$	3.60	3.60	3.60
$q_H$ (CH group)	0.000	0.153	0.167

\*Potential I assumes zero net atomic charges while Potential II optimizes the model charges. Potential III also allows nonzero net atomic charges but was derived by the structural fit method. The geometric-mean combining law was assumed except that in Potential I,  $B_{CH}$  was independently varied. The units are kJ/mol, Å, and electrons.

Table 2. The results of complete lattice energy minimizations for the 18 basis hydrocarbon crystal structures. The molecular angle shift is given by  $\theta$ , and the translation shift by  $r$  (Å, deg)

(Structure) Parameter	Observed	Shift obtained with potential		
		I	II	III
(benzene)				
<i>a</i>	7.39	-0.65	-0.05	0.14
<i>b</i>	9.42	0.05	-0.07	0.00
<i>c</i>	6.81	0.24	0.19	0.30
$\theta$	0.0	19.2	2.6	2.4
(naphthalene)				
<i>a</i>	8.24	-0.68	-0.09	0.09
<i>b</i>	6.00	-0.20	-0.08	-0.03
<i>c</i>	8.66	-0.51	-0.06	-0.01
$\beta$	122.9	-9.1	-0.4	-0.8
$\theta$	0.0	12.2	2.0	2.0
(anthracene)				
<i>a</i>	8.44	-0.41	0.10	0.36
<i>b</i>	6.00	-0.03	-0.11	-0.08
<i>c</i>	11.12	-0.20	0.03	0.17
$\beta$	125.6	-2.5	-2.2	-1.8
$\theta$	0.0	3.9	3.8	4.8
(phenanthrene)				
<i>a</i>	8.46	-0.61	-0.19	-0.07
<i>b</i>	6.16	0.03	0.01	0.13
<i>c</i>	9.47	-0.20	0.10	-0.01
$\beta$	97.7	-3.0	-2.6	-2.2
<i>r</i>	0.00	0.12	0.03	0.01
$\theta$	0.0	5.1	3.0	3.0
(pyrene)				
<i>a</i>	13.65	-0.06	-0.01	-0.05
<i>b</i>	9.26	-0.11	-0.06	-0.04
<i>c</i>	8.47	-0.08	-0.13	0.00
$\beta$	100.3	1.3	-1.4	0.8
<i>r</i>	0.00	0.09	0.17	0.13
$\theta$	0.0	2.9	4.5	3.2
(chrysene)				
<i>a</i>	8.39	-0.48	0.02	0.18
<i>b</i>	6.20	-0.00	-0.07	-0.01
<i>c</i>	25.20	-0.40	-0.09	0.02
$\beta$	116.2	-1.3	-0.7	-1.2
$\theta$	0.0	4.0	1.6	1.7
(triphenylene)				
<i>a</i>	13.17	1.00	0.03	0.20
<i>b</i>	16.73	0.74	0.04	0.19
<i>c</i>	5.26	-0.81	-0.01	0.10
<i>r</i>	0.00	0.52	0.06	0.17
$\theta$	0.0	9.6	0.7	0.4
(perylene)				
<i>a</i>	11.28	0.09	0.01	0.28
<i>b</i>	10.83	-0.09	0.05	0.14
<i>c</i>	10.26	-0.04	-0.18	-0.12
$\beta$	100.6	0.6	-4.1	-3.8
<i>r</i>	0.00	0.15	0.07	0.10
$\theta$	0.0	0.8	4.2	4.2
(ovalene)				
<i>a</i>	19.47	2.05	0.65	0.58
<i>b</i>	4.70	-0.62	-0.14	0.03
<i>c</i>	10.12	0.04	-0.05	-0.01
$\beta$	105.0	2.7	-0.0	-0.3
$\theta$	0.0	9.7	2.4	1.9
(n-pentane)				
<i>a</i>	4.10	-0.04	0.08	0.16
<i>b</i>	9.04	-0.33	-0.08	0.08
<i>c</i>	14.70	0.02	0.13	0.17
<i>r</i>	0.00	0.03	0.01	0.03
$\theta$	0.0	0.8	0.4	0.0
(n-hexane)				
<i>a</i>	4.17	-0.09	0.02	0.09
<i>b</i>	4.70	-0.46	-0.17	-0.05
<i>c</i>	8.57	0.03	0.10	0.12
$\alpha$	96.6	0.4	-0.1	-0.6

Table 2 (Contd)

(Structure) Parameter	Observed	Shift obtained with potential		
		I	II	III
$\beta$	87.2	0.5	0.8	1.3
$\gamma$	105.0	-2.6	-1.7	-1.1
$\theta$	0.0	2.7	3.3	2.9
(n-octane)				
<i>a</i>	4.16	-0.08	0.02	0.10
<i>b</i>	4.75	-0.30	-0.20	-0.07
<i>c</i>	11.00	-0.05	0.03	0.01
$\alpha$	94.8	0.4	0.1	-0.6
$\beta$	84.5	-0.3	0.3	0.3
$\gamma$	105.1	-2.7	-1.3	-0.8
$\theta$	0.0	5.2	5.8	4.3
(cubane)				
<i>a</i>	6.30	-0.14	-0.09	0.00
<i>c</i>	11.73	-0.19	0.58	0.87
$\theta$	0.0	1.1	2.3	2.3
(adamantane)				
<i>a</i>	6.60	0.12	-0.01	0.07
<i>c</i>	8.81	-0.08	0.11	0.25
$\theta$	0.0	0.3	0.7	0.8
(congressane)				
<i>a</i>	10.11	-0.22	-0.04	0.06
(bicyclopropyl)				
<i>a</i>	8.90	-0.37	0.08	0.07
<i>b</i>	5.14	-0.05	0.01	0.12
<i>c</i>	11.81	-0.46	-0.14	-0.01
$\theta$	0.0	2.8	2.8	2.8
(1-biapocamphane)				
<i>a</i>	6.89	-0.21	-0.09	-0.01
<i>b</i>	9.51	-0.21	-0.10	-0.03
<i>c</i>	6.58	-0.13	-0.01	0.07
$\theta$	0.0	0.5	0.5	0.5
(1-biadmantane)				
<i>a</i>	6.53	-0.14	-0.01	0.07
<i>b</i>	6.58	-0.18	-0.05	0.03
<i>c</i>	10.46	0.13	0.24	0.30
$\theta$	0.0	3.4	2.7	1.9
rms relative error in cell edges		0.05	0.02	0.02
rms error in cell angles		3.0	1.7	1.5
rms error in molecular position <i>r</i>		0.25	0.09	0.11
rms error in molecular angle $\theta$		7.0	2.9	2.7

To decrease calculation time, accelerated convergence was used. Contacts were considered up to 8 Å, corresponding to less than 1% error of convergence. The computer program, called PCK6, retains a table of contacts in memory for several cycles. At table makeup time the eigenvalues of the Hessian are checked for negative values. If all are positive, a Newton-Raphson sequence is initiated; otherwise, steepest descents is used. If overshifting occurs, the shifts are binary chopped as necessary. In a typical case (n-hexane) 2500 contacts were considered and six Newton-Raphson cycles were required for a complete lattice energy minimization, starting from the observed structure.

A perusal of Table 2 shows that there is a significant improvement of fit if net coulombic charges are allowed. For the cell edges, 5% rms relative error is obtained when no net charges are allowed. This figure is reduced to 2% when  $q_H$  (CH group) = 0.153e. This improvement of fit provides support for the existence of significant net coulombic charges in hydrocarbons. Quantum mechanical calculations as well as accurate X-ray electron density studies also support the existence of such net charges.

Potential III, derived by the Hagler-Lifson procedure, gives a rather similar goodness of fit as Potential II. This potential also includes coulombic terms, with a somewhat larger  $q_H$  (CH group) = 0.167e. We have found, however, that Potential III gives poor results for the lattice vibrational frequencies, while Potential II gives good agreement. Thus, the potential derived by minimization of the forces resulted in better transferability to the vibrational problem. We are further studying this question of transferability of the potentials.

#### COMPARISON WITH THE EPEN POTENTIAL MODEL

This empirical model potential (Shipman *et al.*, 1975) is rather different from the (exp-6-1) potential we have been considering. It is a highly coulombic model, with point electronic charges being placed in bonding and lone pair positions. The number of attraction-repulsion centers is greater than in the (exp-6-1) model, requiring more terms in the lattice sum. An advantage of the EPEN model is that it requires only four adjustable parameters for saturated hydrocarbons. However, no provision is made for aromatic hydrocarbons in the model. A strong disadvantage of the model is that a very large number of contacts must be considered to get convergence of the

lattice sum, since the model is so highly coulombic. As with the (exp-6-1) potential, the number of nonbonded interatomic contacts considered was reduced by using convergence acceleration. In addition, because of the very large coulombic terms, the lattice sum was taken over an electrically neutral domain to reduce truncation error. Our experience indicates that crystal lattice energy minimizations with the EPEN potential require about ten times as much computer time as (exp-6-1) calculations for the same crystal structure.

Table 3 shows results with the EPEN potential compared to (exp-6-1) Potential II. We did not use the EPEN model for all 18 crystal structures since a lot of computer time would be needed and the EPEN potential is not defined for aromatic hydrocarbons. Potential II gives a better fit than the EPEN potential to both n-pentane and n-hexane (which have quite different crystal structures). Note that Shipman *et al.* (1975) did not vary the molecular rotation and translation in n-pentane. We have repeated their calculation and have obtained good agreement with their calculated values for the lattice constants of n-pentane from the EPEN potential.

#### CONCLUSION

The coulombic Potential Set II gives a significantly better fit to the 18 hydrocarbon crystal structures than the noncoulombic Potential Set I. Potential Set III, although giving as good a structural fit as II, was not as good in the calculation of lattice vibrational frequencies. The EPEN model in this application yielded a poorer fit than II, and required a much larger calculation time for lattice energy minimizations. We recommend Potential Set II for the calculation of the crystal structures of hydrocarbons, used with convergence acceleration.

#### REFERENCES

- Crowell, A. D. (1958) *J. Chem. Phys.* **29**, 446.
- Hagler, A. T. & Lifson S. (1974) *Acta Crystallogr.* **B30**, 1336.
- Shipman, L. L., Burgess, A. W. & Scheraga, H. A. (1975), *Proc. Nat. Acad. Sci. U.S.A.* **72**, 543.
- Starr, T. L. & Williams, D. E. (1977) *Acta Crystallogr. A*, in press.
- Williams, D. E. (1965) *J. Chem. Phys.* **43**, 4424.
- Williams, D. E. (1970) *Trans. Am. Cryst. Ass.* **6**, 21.
- Williams, D. E. (1974) *Acta Crystallogr.* **A30**, 71.

Table 3. Calculated crystal structures of n-pentane and n-hexane with the EPEN potential (A, deg)

(Structure) Parameter	Observed	Calculated shift EPEN
(n-pentane)		
<i>a</i>	4.10	-0.17
<i>b</i>	9.08	-0.30
<i>c</i>	14.86	-0.14
<i>r</i>	0.00	0.03
$\theta$	0.00	2.3
(n-hexane)		
<i>a</i>	4.17	-0.02
<i>b</i>	4.70	-0.25
<i>c</i>	8.57	0.15
$\alpha$	96.6	1.7
$\beta$	87.2	0.6
$\gamma$	105.0	-6.0
$\theta$	0.00	12.4