Acta Cryst. (1986). B42, 388-401

Model Anisotropic Intermolecular Potentials for Saturated Hydrocarbons

BY S. L. PRICE

University Chemical Laboratory, Lensfield Road, Cambridge CB2 1EW, England

(*Received* 4 *October* 1985; *accepted* 24 *January* 1986)

Abstract

The use of explicit orientation-dependent functions, to represent the effects of non-spherical features in the molecular charge distribution, in model intermolecular pair potentials is developed in order to produce more accurate transferable anisotropic site-site potentials for organic molecules. An isotropic atom-atom intermolecular potential-energy surface for methane is analysed using the orientation-dependent expansion functions appropriate to a pair of tetrahedral molecules, to determine the ability of a wide variety of one-centred anisotropic model potentials to reproduce the surface. The analysis is used to develop a simple anisotropic carbon potential scheme for CH_3 and CH_2 groups, which can predict the crystal structures of pentane, hexane and octane as accurately as the widely used isotropic atom-atom potential, and has the advantage of being more computationally efficient and more flexible for future development. The derivation of the appropriate orientation-dependent functions for tetrahedral molecules, using angular momentum theory, in a form which is very convenient for crystal structure analysis is given. These functions are also used to examine the use of van der Waals radii to predict molecular packing.

1. Introduction

Intermolecular forces determine most of the physical properties of matter. We need to develop a fuller quantitative understanding of intermolecular forces, as they control many vital biological processes, such as DNA replication and drug-receptor interactions, as well as determining the behaviour of technologically important phases, such as polymers and organic conductors.

At the simplest level, space-filling molecular models, representing the 'excluded volume' or van der Waals surface of the molecules, are very useful in determining sterically forbidden molecular conformations and packings. This approach is now being applied routinely, even for proteins, by using molecular graphics programs to build and manipulate the molecular models. Although this analysis pretends that molecules have hard edges, and gives no information on the energy associated with the packing, it is a valuable tool for the non-specialist, and so the capabilities of this simple approach are analysed and discussed in Appendix 2.

The next level of sophistication is to use models for the intermolecular forces between the atoms in the molecules, plus appropriate models for energy changes associated with the intramolecular degrees of freedom, to predict molecular conformations and crystal structures. Such model potentials can also be used in molecular dynamics calculations, or other simulation programs, to study the behaviour of molecules in any phase. These computer modelling techniques are limited by the computing resources available, which usually dictates the size of molecule which can be studied, and by the accuracy of the models used for the intermolecular forces. The first limitation is rapidly being overcome by the computer revolution, but the second limitation, which determines the reliability of the results, is a fundamental scientific problem. Although there are many rival sets of potential models and parameters in use, they are all based on the assumption of the isotropic atom-atom potential. In this model, the potential between two interaction sites, usually placed at all the atomic nuclei, depends only on the separation of the sites, and so models the molecule as a superposition of spherical charge distributions. This is a reasonable first approximation, but has been found to give unacceptably poor results for certain heteroatoms, unless the model is supplemented by extra sites to represent the lone-pair interactions (*e.g.* Profeta & Allinger, 1985; Williams & Cox, 1984). However, the addition of extra sites is a limited and inflexible method of improving the potential, and so the use of isotropic interaction sites is a major limitation on the accuracy of intermolecular potentials for organic molecules.

At the other extreme, chemical-physics studies of the properties of very small polyatomic molecules, for example scattering calculations, have described the intermolecular potential in terms of the centre-of-mass separation and the appropriate expansion functions for the orientation dependence of the potential, which are defined in terms of Wigner rotation matrices. This approach uses the power of angular-momentum theory to exploit the symmetry of the molecule, and is only useful for small highly sym-

0108-7681/86/040388-14$01.50

metric molecules such as diatomics and AX_n molecules.

These two approaches to modelling intermolecular forces have recently been combined, thereby discarding the assumption that the intermolecular potentials of organic molecules are composed of spherical interactions, to develop a powerful new approach to model potentials with the flexibility to be capable of great accuracy. In two highly successful pilot studies on Cl_2 and azabenzene molecules (Price & Stone, 1982, 1984), the appropriate orientation-dependent functions have been used to develop anisotropic site-site potentials for Cl, C and N (in aromatic rings) which model the significant effects of the Cl and N lone pairs and the bonded H atoms on the crystal packing. Thus the method has the advantage that it can model the potentials from non-spherical heteroatoms effectively, a pre-requisite for reliable studies on biological molecules. The azabenzene project also suggests that it may be possible to dispense with interaction sites on most hydrogen atoms, which drastically reduces the number of intersite vectors, and so increases the computational efficiency as well as the accuracy, which is also important for modelling large biological molecules.

In order to develop a transferable scheme for modelling the intermolecular potentials of organic molecules using anistropic site-site potentials, and fewer interaction sites than molecules, it is necessary to investigate which of the symmetry-allowed orientation-dependent terms will be most important, and how they should be incorporated into the functional form of the site-site potential. This cannot be done unambiguously, because there is, as yet, no definitively accurate intermolecular potential for a typical organic molecule. It is difficult to develop and test potentials by comparison with experimental data, as such data are either related to some orientational average of the potential, or only sample a limited number of relative orientations. However, we can test the types of anisotropy and functional forms of a one-centred model potential which are required to reproduce the isotropic atom-atom potential for methane, and use this to develop an anisotropic carbon site-site potential for saturated hydrocarbons. Although this procedure is obviously incapable of giving potentials which are more accurate than the current isotropic atom-atom model, it provides valuable insights into the use of anisotropic potentials in modelling intermolecular forces. These general guidelines will be important, not only for developing the next generation of transferable intermolecular potentials for organic molecules, but also for the development of highly accurate potentials for small molecules.

Methane is an important molecule for our understanding of intermolecular forces. It is amongst the most spherical and symmetric of molecules, and hence its intermolecular potential should be relatively conveniently described in terms of a one-centred expansion. Alternatively, it can be viewed as the extreme case of an atom with a non-spherical charge distribution, and so as a test case for developing modelling techniques for atoms with non-spherical features, such as lone-pair electrons. Methane is also a particularly suitable molecule for developing model intermolecular potentials, because we can construct a reasonable model for its intermolecular potential from a transferable atom-atom potential scheme for hydrocarbons, since the isotropic atom-atom potential has been found particularly successful for saturated hydrocarbons.

The first section of this paper uses the expansion functions for any scalar property of two identical tetrahedral molecules to investigate the anisotropy of an isotropic atom-atom potential-energy surface for methane. The expansion functions, which are based on Wigner rotation matrices, are derived in Appendix 1 in a novel simple form, which is particularly suitable for use in intermolecular potentials. These expansion functions are used to examine various characteristic measures of the anisotropy of the potential surface, which provides clues as to the most appropriate forms for model one-centre anisotropic potentials for methane. (The expansion functions are also used to examine the capabilities of van der Waals surface models in Appendix 2.)

Various one-centred anisotropic model potentials for methane are tested for their ability to fit the isotropic atom-atom potential-energy surface, including both commonly used and novel forms, and the use of the expansions for the minimum energy $[-\varepsilon(\Omega)]$ and corresponding separation $[\rho(\Omega)]$ as a function of orientation Ω in various extensions of the isotropic 'exp−6' potential to anisotropic systems. The comparison shows that some model potentials have significantly greater flexibility, and use parameters more effectively than others. These results are very pertinent to the problem of choosing the form of potential to input into scattering and simulation studies.

In the second section, one of these simple one-centre intermolecular potentials for methane, with anisotropic repulsion and isotropic dispersion contributions, is adapted to model CH_3 and CH_2 groups in saturated hydrocarbons, by simply adjusting the dispersion coefficient according to the number of attached hydrogen atoms. This anisotropic carbon site-site potential predicts the crystal structures of pentane, hexane and octane with the same high accuracy as the original isotropic atom-atom potential, but is computationally faster by a factor of three to four, because of the drastic reduction in the number of intersite vectors, which makes the model particularly attractive for calculations on large molecules.

Table 1. *Definition of the grid of points representing the potential surface*

(*a*) The relative orientations

A set of ten orientations of a regular tetrahedron is defined relative to space-fixed axes:

I: Standard orientation (Fig. 1)

$H_1(r/\sqrt{3})(1,1,1)$ $H_2(r/\sqrt{3})(-1,-1,1)$
$H_3(r/\sqrt{3})(1,-1,-1)$ $H_4(r/\sqrt{3})(-1,1,-1)$

II: Rotate I by 45° about x axis

$H_1(r/\sqrt{3})(1,0,\sqrt{2})$ $H_2(r/\sqrt{3})(-1,-\sqrt{2},0)$
$H_3(r/\sqrt{3})(1,0,-\sqrt{2})$ $H_4(r/\sqrt{3})(-1,\sqrt{2},0)$

III: Rotate II by $\frac{1}{2}\cos^{-1}(-1/3)$ about y axis

$H_1(r/3)(3,0,0)$ $H_2(r/3)(-1,-\sqrt{6},\sqrt{2})$
$H_3(r/3)(-1,0,-2\sqrt{2})$ $H_4(r/3)(-1,\sqrt{6},\sqrt{2})$

IV: Rotate I by 30° about z axis

$H_1(r/2\sqrt{3})(\sqrt{3}-1,\sqrt{3}+1,2)$
$H_2(r/2\sqrt{3})(1-\sqrt{3},-\sqrt{3}-1,2)$
$H_3(r/2\sqrt{3})(\sqrt{3}+1,1-\sqrt{3},-2)$
$H_4(r/2\sqrt{3})(-\sqrt{3}-1,\sqrt{3}-1,-2)$

V: Rotate IV by 30° about y axis

$H_1(r/4\sqrt{3})(5-\sqrt{3},2+2\sqrt{3},1+\sqrt{3})$
$H_2(r/4\sqrt{3})(\sqrt{3}-1,-2-2\sqrt{3},3\sqrt{3}-1)$
$H_3(r/4\sqrt{3})(1+\sqrt{3},2-2\sqrt{3},-1-3\sqrt{3})$
$H_4(r/4\sqrt{3})(-5-\sqrt{3},-2+2\sqrt{3},1-\sqrt{3})$

VI to X: Invert I to V.

The axes are defined as $\mathbf{x}=(\sqrt{3}/2)(\mathbf{H}_1+\mathbf{H}_3)$, $\mathbf{y}=(\sqrt{3}/2)(\mathbf{H}_1+\mathbf{H}_4)$, $\mathbf{z}=(\sqrt{3}/2)(\mathbf{H}_1+\mathbf{H}_2)$, for tetrahedra I to V. The **x** and **y** axes are interchanged for the inverted tetrahedra VI to X, to give a right-handed set of axes.

The intermolecular vector is defined by the product of the intermolecular distance and one of the following twelve vectors:

$\hat{\mathbf{R}}_1=(1,0,0)$ $\hat{\mathbf{R}}_7=(1/\sqrt{3})(-1,1,-1)$
$\hat{\mathbf{R}}_2=(0,1,0)$ $\hat{\mathbf{R}}_8=(1/\sqrt{2})(1,1,0)$
$\hat{\mathbf{R}}_3=(0,0,1)$ $\hat{\mathbf{R}}_9=(1/\sqrt{2})(1,0,1)$
$\hat{\mathbf{R}}_4=(1/\sqrt{3})(1,1,1)$ $\hat{\mathbf{R}}_{10}=(1/\sqrt{2})(0,1,1)$
$\hat{\mathbf{R}}_5=(1/\sqrt{3})(1,1,-1)$ $\hat{\mathbf{R}}_{11}=(\sqrt{3}/4,3/4,1/2)$
$\hat{\mathbf{R}}_6=(1/\sqrt{3})(-1,1,1)$ $\hat{\mathbf{R}}_{12}=(3/4,-1/2,-\sqrt{3}/4)$

A relative orientation is generated by taking molecule 1 at (0, 0, 0) and translating molecule 2 to $\hat{\mathbf{R}}$. This set of ten oriented tetrahedra and 12 intermolecular unit vectors gives 1200 relative orientations, of which 456 are unique.

(*b*) The relative separations

The potential surface is defined by evaluating the potential for each of the 456 relative orientations at centre-of-mass separations of 3·5, 3·75, 4·0, 4·25, 4·5, 4·75, 5·0 and 6·0 Å, and eliminating points with $U>0{\cdot}5$ kJ mol^{-1}. This gives a potential-energy surface of 2966 points with a r.m.s. energy of 0·91 kJ mol^{-1}.

(*c*) The potential

The potential is described by equation (1) and the following potential parameters, taken from Williams & Cox (1984):

$A_{CC}=2439{\cdot}8$ kJ mol^{-1} Å^6 $A_{HH}=136{\cdot}4$ kJ mol^{-1} Å^6
$B_{CC}=369743$ kJ mol^{-1} $B_{HH}=11971$ kJ mol^{-1}
$C_{CC}=3{\cdot}60$ Å^{-1} $C_{HH}=3{\cdot}74$ Å^{-1}

$A_{CH}=(A_{CC}A_{HH})^{1/2}$
$B_{CH}=(B_{CC}B_{HH})^{1/2}$
$C_{CH}=\frac{1}{2}(C_{CC}+C_{HH})$

and the effective C—H bond length $r=1{\cdot}02$ Å.

Thus, this theoretical study of model intermolecular potentials for methane leads to the development of a potential for CH_3 and CH_2 groups as a single anisotropic site, which is as accurate as current models, and confirms the practical utility of anisotropic site-site potentials. The possibilities for extending these models to give more accurate potentials for methane, and a transferable scheme of anisotropic site-site potentials for modelling biological molecules, are discussed in the final section.

2. Comparison of model potentials for methane

The potential surface used to compare various one-centred anisotropic intermolecular potentials for the repulsion and dispersion energy of two methane molecules was the isotropic atom-atom potential of Williams & Cox (1984), omitting the electrostatic interaction (which is discussed in § 4). The total potential is assumed to be a sum of isotropic exp − 6 potentials between all intermolecular pairs of atomic sites (the H sites are displaced 0·07 Å from the nucleus to model the shift of charge density in this bond), *i.e.*

$$U(R,\Omega)=\sum_{i,k=1,5} B_{\iota\kappa}\exp(-C_{\iota\kappa}R_{ik})-A_{\iota\kappa}/R_{ik}^6, \qquad (1)$$

where sites i and k are of type ι or κ (C or H) respectively. The parameters in this model (Table 1) were derived by fitting to a wide variety of hydrocarbon crystal structures. This and closely related atom-atom potentials have been widely used for modelling saturated hydrocarbons, and so this potential should be a reasonable approximation to the repulsion and dispersion contributions to the methane potential in the region of interest for packing problems, *i.e.* in the region of the minimum-energy well. In this potential scheme, the C C potential has a minimum of −0·40, C H of −0·14 and H H of −0·05 kJ mol^{-1}, which is a reasonable ratio, but nevertheless results in the C C interaction being smaller than the sum of either the C H or the H H interactions in the region of the minimum energy, for any orientation of two methane molecules. Hence, in this sense, the H-atom contributions dominate the potential.

2.1. *The anisotropy of the potential minimum*

The anisotropy of the potential-energy surface can be characterized by finding the minimum energy $-\varepsilon(\Omega)$, and corresponding separation $\rho(\Omega)$, as a function of orientation. This was done for a set of 456 relative orientations of two CH_4 molecules, defined in Table 1, and the set of values for $\varepsilon(\Omega)$ and $\rho(\Omega)$ were fitted to a linear expansion in the appropriate orientation-dependent functions, by a least-squares minimization. This expansion has the general form

$$\begin{aligned}x(\Omega)=x_{000}&+x_{303}Z_{303}(\Omega)+x_{330}Z_{330}(\Omega)\\&+x_{404}Z_{404}(\Omega)+x_{440}Z_{440}(\Omega)\\&+x_{332}Z_{332}(\Omega)+x_{334}Z_{334}(\Omega)\\&+x_{336}Z_{336}(\Omega)\end{aligned} \qquad (2)$$

for any scalar property x of the relative orientation of two like tetrahedral molecules, where the $Z_{l_1l_2j}(\Omega)$ are the members of the general set of functions $\bar{S}^{k_1k_2}_{l_1l_2j}$ with the appropriate symmetry (Appendix 1).

The results in Table 2 show that the minimum energy and corresponding separation vary with

Table 2. *Comparison of the expansions of the orientation dependence of various characteristics of the potential surface*

	Coefficients in the linear expansion of					
	(*a*) R for the energy contours (Å)			(*b*) U for the separations (kJ mol^{-1})		
Coefficients of	Min. $\rho(\Omega)$	$\sigma(\Omega)$ $U=0$	$U=1$ kJ mol^{-1}	$-\varepsilon(\Omega)$	$R=4$ Å	$R=5$ Å
Z_{000}	4·2324	3·7984	3·7007	−1·2615	−0·7497	−0·6743
Z_{303}	0·2500	0·2497	0·2306	0·2988	1·1643	−0·0185
Z_{330}	−0·0144	−0·0148	−0·0140	−0·0286	0·0515	0·0014
Z_{404}	−0·0760	−0·0763	−0·0696	−0·0866	−0·5155	−0·0015
Z_{440}	−0·0020	−0·0019	−0·0016	−0·0061	0·0311	0·0006
Z_{332}	0·0170	0·0177	0·0169	0·0348	−0·0894	−0·0021
Z_{334}	−0·0103	−0·0115	−0·0117	−0·0286	0·1747	0·0035
Z_{336}	−0·0522	−0·0557	−0·0512	−0·0346	−0·6836	−0·0083
R.m.s. error	0·0203	0·0205	0·0189	0·0412	0·2650	0·0042
R.m.s. value	4·2407	3·8073	3·7085	1·2830	1·2756	0·6740
Max value	4·8110	4·3810	4·2360	−0.7115	4·9568	−0·6261
(For Ω	I VI $\mathbf{R}_4$	I VI $\mathbf{R}_4$	I VI $\mathbf{R}_4$	I VI $\mathbf{R}_4$	I VI $\mathbf{R}_4$	I VI $\mathbf{R}_5$)
Min. Value	3·8080	3·3760	3·3070	−1·8726	−1·7506	−0.7002
(For Ω	I VI $\mathbf{R}_5$	I VI $\mathbf{R}_5$	I VI $\mathbf{R}_5$	I VI $\mathbf{R}_5$	I VI $\mathbf{R}_5$	I VIII $\mathbf{R}_1$)

orientation by 1·2 kJ mol^{-1} and 1 Å respectively, but can be described well by the above eight-term expansion, $\rho(\Omega)$ being described better than $\varepsilon(\Omega)$. The similarity of the anisotropy of $\rho(\Omega)$ and $\varepsilon(\Omega)$ implies that the orientations where the molecules can get close to each other (small ρ) have deeper wells (large ε); the global minimum energy occurs for the orientation with the smallest value of ρ, which occurs when two tetrahedral faces are parallel so that there are six equal shortest H H contacts. The coefficients in the expansion do not uniformly decrease in magnitude with the total order (l_1+l_2+j) of the term, contrary to the assumption that is usually made when using drastically truncated expansions. The two dominant anisotropic terms are Z_{303} and Z_{404} and, indeed, if the other five $Z_{l_1l_2j}$ terms are omitted, the r.m.s. error increases by less than a factor of two, to give a r.m.s. error of 0·0363 Å (0·9% of the r.m.s. value) for $\rho(\Omega)$ and 0·0607 kJ mol^{-1} (5%) for $\varepsilon(\Omega)$ for the three-term expansion. This is an acceptable error in comparison with the r.m.s. errors of 0·2024 Å and 0·2444 kJ mol^{-1} which occur if only the isotropic terms ρ_{000} and ε_{000} are used to approximate $\rho(\Omega)$ and $\varepsilon(\Omega)$. The success of the Z_{l0l} expansion is not surprising, since only the H H terms in the atom–atom potential can contribute to the other terms ($Z_{l_1l_2j}$ with $l_1, l_2 \neq 0$) in the expansion. Both the C H and H H interactions contribute to the Z_{l0l} and isotropic terms in the expansion. Thus this qualitative result can only be extended to weakly anisotropic systems, such as AH_n or non-spherical atoms. For other polyatomic molecules, the equivalent terms ($\bar{S}^{k_1k_2}_{l_1l_2j}$ with l_1 and l_2 non zero) will be far more important, and so it would be inadvisable to use a one-centred potential. [This point was investigated empirically for Cl_2 by Price & Stone (1982).]

The Z_{l0l} terms only depend on the relative orientation of the intermolecular vector **R** and the molecular axes, whereas the $Z_{l_1l_2j}$ terms with $l_1, l_2 \neq 0$ also depend on the relative orientation of the two sets of molecular axes. Hence the Z_{l0l} terms can be thought of as expanding the shape of the isolated molecule. The contribution of either molecule to the dominant anisotropic term Z_{303} has a maximum when **R** is along a CH bond, and so increases $\rho(\Omega)$ along the four CH bonds, and reduces it in the opposite direction. The centrosymmetric term Z_{404} refines this picture by increasing ρ along a CH bond, and directly behind it, and decreasing ρ along the bond bisectors.

The terms in $Z_{l_1l_2j}$ with $l_1, l_2 \neq 0$ are required to describe the changes in energy of a pair of molecules as they are rotated about **R**. The most important of these secondary terms, Z_{336}, is the orientation dependence of the interaction of the first non-vanishing total multipole moment for tetrahedral molecules, the octupole component [Ω_{xyz} or Ω_{32s} as defined by Price, Stone & Alderton (1984)]. Hence a consideration of the shape, and not just the symmetry, of an AH_n molecule or non-spherical atom should give guidance as to the dominant anisotropic terms in the potential and, unless the molecular fragment is charged, these will not be the terms in the central multipole expansion of the electrostatic energy.

2.2. *Characterization of the potential surface*

The choice of $\rho(\Omega)$ and $\varepsilon(\Omega)$ to characterize the anisotropy of the potential surface is an obvious one, but nevertheless needs investigation. In Table 2, the expansion of $\rho(\Omega)$ is compared with the expansions for other energy contours of the surface, namely the zero-potential σ, [$U(\sigma)=0$], and the $U=1$ kJ mol^{-1} contour. The anisotropies of these characteristic separations are essentially the same, emphasizing the conclusion that the shape terms Z_{l0l} dominate the radial scale of the potential.

In contrast, the variation of the energy with orientation for fixed centre-of-mass separations $R=4{\cdot}0$ and

$R = 5{\cdot}0$ Å and the minimum energy $\varepsilon(\Omega)$ differ grossly, and are relatively poorly described by an eight-term expansion. Although Z_{303} dominates the anisotropy, the higher-order terms become increasingly important at smaller separations. This occurs because, although the anisotropy of the repulsion and dispersion are defined by the same assumption (*i.e.* the isotropic atom–atom model), the different radial forms lead to different anisotropies. The repulsive wall is extremely anisotropic, so that even at 4 Å, which is well within the range of minimum-energy separations, a potential of the form $U(R, \Omega) = \Phi(\Omega)f(R)$ is clearly inadequate. The required anisotropy could be produced with potentials of the form $U(R, \Omega) = f[R, \rho(\Omega)]$, such as $U_{\text{rep}} = A \exp\{-\alpha[R - \rho(\Omega)]\}$, because the exponential term produces powers of Z_{303} *etc.*, which correspond to higher terms in the anisotropy of the potential. Hence it is more useful, and compact, to discuss the anisotropy of a potential-energy surface in terms of its radial anisotropy, *i.e.* as an energy-contour hypersurface, than in terms of energy variations at fixed R.

2.3. *Anisotropic one-centred exp–6 potentials for methane*

The analysis of the anisotropy of the atom–atom potential for methane can be used to design model intermolecular potentials which are functions of the carbon–carbon separation R and the orientation-dependent functions $Z_{l_1 l_2 j}(\Omega)$, seeking effective representations of the potential-energy surface. One obvious approach is to use the expansions for the well depth $\varepsilon(\Omega)$ and corresponding separation $\rho(\Omega)$ in place of ε and ρ in the isotropic exp–6 potential;

$$U_{\text{scale}}(R, \Omega) = \varepsilon(\Omega)(1 - 6/\alpha)^{-1} \times \{(6/\alpha) \exp(\alpha[1 - R/\rho(\Omega)]) - [\rho(\Omega)/R]^6\}. \quad (3)$$

This potential has its minimum-energy contour $-\varepsilon(\Omega)$ at $\rho(\Omega)$, and includes one additional parameter α which has to be determined. The centre-of-mass separation R is scaled by the radial anisotropy $\rho(\Omega)$ and, since it is in the form $U = \Phi(\Omega)f[R/s(\Omega)]$, it is an example of a Corner (1948) potential. The Kihara (1978) and Gaussian Overlap (Berne & Pechukas, 1972) are also examples of Corner potentials, with specific models for $\Phi(\Omega)$ and $s(\Omega)$, which have been examined by Walmsley (1977). However, as Stone (1979) has pointed out, the Corner potential models are unsatisfactory in that where $s(\Omega)$ is large, the potential is softer, *i.e.* the force is inversely proportional to $s(\Omega)$. Stone (1979) proposed an alternative form $U = \Phi(\Omega)f[R - s(\Omega)]$, where the reduced shape of the potential curve (U/Φ) is the same for all orientations, but the radial scale is shifted by $s(\Omega)$. A shifted version of the exp–6 potential can be defined as

$$U_{\text{shift}}(R, \Omega) = \varepsilon(\Omega)(\alpha\rho_0 - 6)^{-1} \times (6 \exp\{-\alpha[R - \rho(\Omega)]\} - \alpha\rho_0^7/[R - \rho(\Omega) + \rho_0]^6), \quad (4)$$

which also has a minimum energy $-\varepsilon(\Omega)$ at $\rho(\Omega)$. Here ρ_0 is taken as the isotropic component of $\rho(\Omega)$, so the model has only one adjustable parameter α, but ρ_0 could be taken as an additional parameter.

These model potentials automatically fit the surface at the minimum-energy contour, within the error associated with fitting $\varepsilon(\Omega)$ and $\rho(\Omega)$, but will give different extrapolations away from this contour. The models can be compared with the potential-energy surface grid, defined in Table 1, which is designed to sample the potential in the well region. The value of the unknown potential parameter α was derived for each model by least-squares fitting to the 2966 points on the potential-energy surface. The r.m.s. errors in reproducing the surface are given in Table 3. This shows that both the shifted and the scaled form of $U(R, \Omega)$ give a reasonable fit to the surface, with the shifted model (r.m.s. error 0·06 kJ mol^{-1}) being somewhat better than the scaled model (r.m.s. error 0·08 kJ mol^{-1}). The simpler 'shape only' expansions for $\varepsilon(\Omega)$ and $\rho(\Omega)$, which contain only the three terms in Z_{l0l}, also give an acceptable fit to the surface, with r.m.s. errors of approximately 0·1 kJ mol^{-1}, compared with 0·34 kJ mol^{-1} for the isotropic potential using ε_{000} and ρ_{000}. This suggests that a simple model potential for small molecules or fragments can be produced by using the molecular shape to determine the appropriate S_{l0l}^{k0} to model $\rho(\Omega)$ and $\varepsilon(\Omega)$ within the shifted potential (4). A van der Waals surface could be used to estimate the anisotropic parameters, assuming $\varepsilon_{l0l}^{k0} \simeq \rho_{l0l}^{k0}$, leaving only ρ_0 and ε_0 to be determined by fitting to experimental data such as virial coefficients.

2.4. *Other potential models*

The exp–6 potentials provide a useful extrapolation of a potential surface away from the minimum-energy contour, but this extrapolation is very inflexible and, as the same parameters are used to describe the repulsion and dispersion contributions, it cannot be very accurate. Hence this type of model potential is of limited usefulness in deriving a potential by fitting to a set of experimental results or *ab initio* points, and other types of model potential may be able to fit the data better, with fewer parameters, and still give a reasonable extrapolation of the potential to the regions not sampled by the data being fitted. The choice of model potential is the major factor in determining the accuracy of the fitted potential, and hence the reliability of the properties calculated from

Table 3. *Comparison of model potential surfaces*

(*a*) Anisotropic extensions of the exp−6 potential

Expansion of $f(\Omega)$	Scaled potentials (equation 3) R.m.s. error (kJ mol^{-1})	Optimum α_{scale}	Shifted potentials (equation 4) R.m.s. error (kJ mol^{-1})	Optimum α_{shift} (Å^{-1})
Full expansion	0·081	14·9374	0·060	3·6832
Z_{l0l} expansion	0·104	14·5975	0·094	3·5771
Isotropic*	0·344	9·3367	0·346	1·011

(*b*) Other model potentials

	R.m.s. error (kJ mol^{-1})		Potential parameters (Å, kJ mol^{-1} *etc.*)	
Isotropic exp−6	0·307	$\alpha = 9·3468$	$\rho = 4·1694$	$\varepsilon = 1·0894$
Linear expansion (equation 5)				
All $U_{l_1l_2j}$ the same	0·261	$\alpha = 9·4033$	$\rho = 4·1660$	
		$\varepsilon_{000} = 1·0742$	$\varepsilon_{303} = -0·2690$	$\varepsilon_{330} = 0·0135$
		$\varepsilon_{404} = 0·0859$	$\varepsilon_{440} = 0·0117$	$\varepsilon_{332} = -0·0288$
		$\varepsilon_{334} = 0·0118$	$\varepsilon_{336} = 0·0532$	
Two term fixed $\alpha = 9·35$	0·217	$\rho_{000} = 4·3316$	$\varepsilon_{000} = 1·0461$	
		$\rho_{303} = 2·5878$	$\varepsilon_{303} = -2·3442$	
Three term fixed $\alpha = 9·35$	0·203	$\rho_{000} = 4·3502$	$\varepsilon_{000} = 1·0489$	
		$\rho_{303} = 2·2830$	$\varepsilon_{303} = -5·2794$	
		$\rho_{404} = 3·1272$	$\varepsilon_{404} = 0·3238$	
Type $U(R,\Omega) = \exp\{-\alpha(\Omega)[R-\rho(\Omega)]\} - A/R^6$ (equation 6)				
Two term $\rho(\Omega)$	0·161	$\alpha = 3·1431$	$A = 1·221 \times 10^4$	
		$\rho_{000} = 4·2243$	$\rho_{303} = 0·1026$	
Three term $\rho(\Omega)$	0·096	$\alpha = 3·4930$	$A = 1·1482 \times 10^4$	
		$\rho_{000} = 4·1747$	$\rho_{303} = 0·1292$	$\rho_{404} = -0·0408$
Three term $\alpha(\Omega)$	0·208	$\rho = 4·4161$	$A = 1·5849 \times 10^4$	
		$\alpha_{000} = 2·6693$	$\alpha_{303} = 0·3457$	$\alpha_{404} = -0·1397$

* The scaled and shifted potentials are equivalent when ε and ρ are isotropic, with $\alpha_{scale} = \rho\alpha_{shift}$. However, the fit is so poor that the two numerical minimizations give different potentials with virtually the same r.m.s. error.

it, and so it is important to consider a variety of models in order to obtain the best one-centred representation of the methane potential. The choice of model potential has often been determined by theoretical or computational considerations, and so different types of model potentials have been used for different types of studies. We can test whether the different approaches generate realistic model potentials, by fitting the trial model potential to the assumed surface, in order to assess the flexibility of the model and its efficiency in the use of adjustable parameters. A model which cannot fit the assumed potential surface significantly better than an isotropic potential (r.m.s. error 0·3 kJ mol^{-1}) is very unlikely to be capable of giving a good representation of the actual potential. On the other hand, there is no point in constructing an elaborate model to reproduce the surface exactly, as a model which is reasonably similar to the assumed potential could be a better approximation to the true potential.

Linear expansion potentials. The form of model potential required for scattering calculations which study rotational-energy transfer is a linear expansion in the eigenfunctions of the total angular momentum (Pack, 1974) of the general form

$$U(R,\Omega) = \sum U^{k_1k_2}_{l_1l_2j}(R)\bar{S}^{k_1k_2}_{l_1l_2j}(\Omega). \qquad (5)$$

If the radial part of the potential is assumed to be independent of l_1l_2j, for example the exp−6 potential, then the corresponding model potential for methane is

$$U(R,\Omega) = \sum \varepsilon_{l_1l_2j} Z_{l_1l_2j}(\Omega)(1-6/\alpha)^{-1} \times\{(6/\alpha)\exp[\alpha(1-R/\rho)] - (\rho/R)^6\}.$$

Table 3 shows that even a ten-parameter model of this obsolete form is incapable of giving a reasonable fit to the potential. This follows from the large variation in the anisotropy with R, which was discussed in § 2.2. The changes in the anisotropy with R can, in principle, be modelled by allowing the parameters in $U_{l_1l_2j}$ (*e.g.* α and ρ) to differ for different l_1l_2j. Since many terms in the expansion are required in order to represent the repulsive wall accurately (see the expansion of the potential at $R = 4$ Å in Table 2), such a model would have an unmanageable number of parameters. However, a drastically truncated expansion is often used for scattering calculations (*e.g.* Buck, Schleusener, Malik & Secrest, 1981; Buck, Huisken, Kohlhase, Otten & Schaefer, 1983), with the justification that the important features in the scattering result from the lowest-order terms in the potential. This type of potential was examined by fitting both a two- and a three-term-expansion model potential to the isotropic atom–atom potential surface, optimizing the values of ε_{l0l} and ρ_{l0l} (the α exponent was held constant because of the high correlation of the potential parameters). Both potentials reproduced the surface relatively badly (Table 3), showing that this form of potential is unsuitable for

studies of properties which depend on the full potential surface rather than just specific terms in the linear expansion. The coefficients in U_{303} changed significantly when the term in Z_{404} was added to the model potential. Thus, when a truncated expansion is fitted to a potential-energy surface, the fitted $U^{k_1k_2}_{l_1l_2j}(R)$ will partially compensate for the omitted higher-order terms, resulting in significant errors in the determination of $U^{k_1k_2}_{l_1l_2j}(R)$. Thus, rather than fitting a linear potential directly it could be more accurate to derive it by fitting another model to the experimental or *ab initio* data, and then extracting the required terms by numerical integration.

Anisotropic potentials for crystal-structure analysis. In many simulations, such as crystal-packing analyses, the property being studied is insensitive to the anisotropy of the long-range dispersion forces. It is particularly convenient if the dispersion forces are isotropic, as this enables their long-range energy contribution to be evaluated very efficiently, using, for example, the Ewald-Bertaut-Williams method (Williams, 1971), which is included in programs such as *WMIN* (Busing, 1981). It is a reasonable approximation to assume that the dispersion is isotropic, as the atom-atom potential surface is only weakly anisotropic by $R = 5$ Å (Table 2), and the anisotropic repulsion model can partially absorb the anisotropy of the dispersion forces in the well region. Hence, models of the general form

$$U(R, \Omega) = \exp\{-\alpha(\Omega)[R - \rho(\Omega)]\} - A/R^6, \quad (6)$$

where $\alpha(\Omega)$ and $\rho(\Omega)$ are various short linear expansions in Z_{l0l}, were fitted to the potential surface. [If α is independent of orientation, this form is exactly equivalent to $U(R, \Omega) = \Phi(\Omega)\exp\{-\alpha[R - \rho'(\Omega)]\} - A/R^6$ (Price & Stone, 1980).] This general model was used to derive anisotropic site-site potentials for Cl_2 and the azabenzene molecules by crystal structure analysis (Price & Stone, 1982, 1984).

A potential of this form, with $\rho(\Omega) = \rho_{000} + \rho_{303}Z_{303}(\Omega) + \rho_{404}Z_{404}(\Omega)$ and $\alpha(\Omega) = \alpha$, gave one of the best fits to the potential surface, comparable to that obtained with the shape-only (Z_{l0l} expansion) shifted exp−6 potential, using fewer parameters. Thus, we have obtained a good one-centred potential for methane which is fortunately in a particularly convenient form for crystal structure analysis. This potential is used in the next section to develop anisotropic carbon site potentials for saturated hydrocarbons.

A similar potential, which uses the anisotropy to scale the effective separation, instead of shifting the repulsive wall, by making α orientation dependent instead of ρ, gives a much poorer fit to the surface. This, along with the results from the anisotropic exp−6 potentials (§ 2.3), confirms the conclusion from a study of an *ab initio* $H_2\,H_2$ potential-energy surface (Price & Stone, 1980) that 'shifted' potential models are superior to the 'scaled' Corner potentials.

The use of orientation-dependent parameters within the potential, such as in (6), is clearly a much more effective approach to modelling the surface than using a linear expansion of the orientation dependence (5), even for this highly spherical molecule. The general form of repulsive potential in (6), with both the α and ρ parameters being orientation dependent, enables the potential to be both shifted and scaled with changes in orientation. Such model potentials will be sufficiently flexible to represent the anisotropy of the repulsive wall very accurately when we are able to determine the intermolecular potentials of polyatomic molecules with anything approaching the accuracy which has been achieved for monatomics such as argon (Maitland, Rigby, Smith & Wakeham, 1981). However, the next section shows that the simple five-parameter shifted potential is adequate for modelling the crystal packing of saturated hydrocarbons.

3. Anisotropic carbon site–site potentials for saturated hydrocarbons

A typical saturated hydrocarbon such as hexane, or a fragment such as an ethyl group, can be approximated as a set of rigidly bonded methane molecules, with the hydrogens along the C–C bonds removed. The repulsion potential for methane should be a reasonable approximation to the *intermolecular* repulsion of a CH_2 or CH_3 group within a molecule, because the repulsion potential in the region of the C–C bond, which is the region where the methane molecule differs substantially from the fragments, cannot be sampled in any simulation because of the presence of the bonded molecular fragment. The isotropic dispersion coefficient for CH_4 is too large for a CH_3 or CH_2 group, as the number of electrons contributing to this long-range effect differs. An estimate of the appropriate ratio may be obtained from the dispersion coefficients in the atom-atom model at separations which are large compared with the CH bond length. The potential of Williams & Cox (1984; and Table 1) gives these long-range $-R^{-6}$ coefficients as

$$\begin{aligned} C_6^{\rm iso}(CH_4,CH_4) &= A_{CC} + 8A_{CH} + 16A_{HH} \\ &= 9237 \text{ kJ mol}^{-1} \text{ Å}^6, \\ C_6^{\rm iso}(CH_3,CH_3) &= A_{CC} + 6A_{CH} + 9A_{HH} \\ &= 7129 \text{ kJ mol}^{-1} \text{ Å}^6, \\ C_6^{\rm iso}(CH_2,CH_2) &= A_{CC} + 4A_{CH} + 4A_{HH} \\ &= 5293 \text{ kJ mol}^{-1} \text{ Å}^6. \end{aligned}$$

The effective isotropic $-R^{-6}$ coefficient, A, which was

fitted to the surface is approximately 25% larger than $C_6^{iso}(CH_4,CH_4)$, because it is also partially modelling the extra dispersion in R^{-n}, $n>6$, produced by the anisotropy of the atom–atom surface. Hence the fitted A coefficient should be scaled by the ratio of the values of the C_6^{iso} to obtain estimates for the effective dispersion coefficients for CH_3 and CH_2.

Thus, we may define a carbon site potential for saturated hydrocarbon fragments

$$U(R,\Omega)=\sum_{\text{C sites}}\exp\{-\alpha[R_{ik}-\rho_{000}-\rho_{303}Z_{303}(\Omega_{ik})-\rho_{404}Z_{404}(\Omega_{ik})]\}-a_\iota a_\kappa/R_{ik}^6, \quad (7)$$

where the repulsion parameters are taken directly from Table 3, and the dispersion-energy parameters, $a_{CH_3}=94{\cdot}13$ $(\text{kJ mol}^{-1}\text{ Å}^6)^{1/2}$ and $a_{CH_2}=81{\cdot}11$ $(\text{kJ mol}^{-1}\text{ Å}^6)^{1/2}$, are obtained by the appropriate scaling of the fitted coefficient A.

This potential, derived purely from an analysis of the isotropic atom–atom potential for methane, can be tested by using it to predict the crystal structures of typical saturated hydrocarbons. The calculations were performed using the crystal structure analysis program *WMIN* (Busing, 1981), with the addition of two short subroutines to calculate the orientation-dependent functions. One subroutine, which is only called when the crystal structure is changed, calculates and stores the **x**, **y** and **z** vectors defining the orientation of each tetrahedral carbon atom, using a carefully ordered list of the numbers of the bonded carbon and hydrogen atoms (the connection table). The **z** axis is defined as the sum of unit vectors along the backbone bonds to atoms 1 and 2 (see Fig. 1), the **x** axis by the sum of unit bond vectors to atoms 1 and 3, followed by Schmidt orthogonalization to give a vector strictly perpendicular to **z**, and the **y** axis is given by the vector product $\mathbf{y}=\mathbf{z}\times\mathbf{x}$. This procedure gives a good positioning of the axes for CH_2 and CH_3 groups, which are rarely exactly tetrahedral. If the molecular unit deviates significantly from tetrahedral, for example in cubane, then this potential is not appropriate. However, more frequently, the positions of the hydrogen atoms are not determined, and this approach has the advantage of only requiring the input of the positions of one or two dummy hydrogen atoms in the expected directions of the C–H bonds. It is crucial that the atoms 1, 2, 3 are chosen so that they define a right-handed set of axes with an atom in the positive quadrant, as the interchange of two atoms would model the inverted molecular fragment. The potential can then be evaluated for every carbon–carbon intermolecular interaction, using an extra subroutine which evaluates the cosine of the angle between two vectors (**R** and **x**, **y** or **z** for either site), to obtain $\mathbf{x}_1.\hat{\mathbf{R}}$ *etc.* for substitution into the simple formulae for $Z_{303}(\Omega)$ and $Z_{404}(\Omega)$ which are given in Appendix 1.

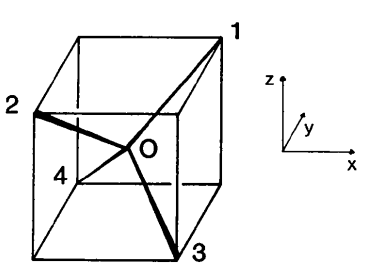

Fig. 1. Definition of axes for a tetrahedral molecule.

Table 4. *Crystal structure predictions from an anisotropic carbon potential*

	n-Pentane *Pbcn*		*n*-Hexane *P*$\bar{1}$		*n*-Octane *P*$\bar{1}$	
	Exp.	Error	Exp.	Error	Exp.	Error
a (Å)	4·100	0·186	4·17	0·12	4·22	0·08
b (Å)	9·076	−0·026	4·70	−0·07	4·79	−0·15
c (Å)	14·859	0·090	8·57	0·09	11·02	0·01
α (°)	—	—	96·6	−0·9	94·7	−0·6
β (°)	—	—	87·2	1·6	84·3	0·1
γ (°)	—	—	105·0	−1·6	105·8	−1·8
$\Delta\theta_x$ (°)		—		−0·07		0·00
$\Delta\theta_y$ (°)		−1·15		0·76		−0·53
$\Delta\theta_z$ (°)		—		−1·28		−0·44
Δt_y (Å)		−0·0007		—		—
U_l (kJ mol^{-1})	−41·5	2·8	−52·6	5·3	−66·4	3·8

The energy minimizations were carried out within the observed space group starting from the experimental results, using the program *WMIN* (Busing, 1981). The structural parameters $\Delta\theta_i$ and Δt_i give the symmetry-allowed rotations and translations of the rigid molecules from their experimental positions, defined relative to the crystal axes (as in *WMIN*; Busing, 1981). A positive error in the lattice parameters indicates that the predicted value is larger than the experimental value. The experimental structures for pentane and octane were taken from the work of Mathisen, Norman & Pedersen (1967), and the structure of hexane from Norman & Mathisen (1961). The hydrogen-atom coordinates are not fully determined for hexane, so the tetrahedral axes were defined using dummy sites. The lattice energies, U_l, were taken from Williams & Cox (1984).

These calculations, using orientation-dependent potentials, are very efficient, because the cost of calculating the $Z_{l_1l_2j}(\Omega)$ is a small price to pay for the omission of the hydrogen sites. The reduction in the number of intersite vectors is large; evaluating the intermolecular energy of a pair of hexane molecules requires summing over 400 interactions for the atom–atom model, compared with 36 for the anisotropic carbon model. Thus, the structure-prediction calculations for this potential required only a quarter to a third of the computer time needed for the isotropic atom–atom model.

The structures predicted by this anisotropic site–site potential for pentane, hexane and octane (Table 4) are in good agreement with the experimental crystal structures, with a r.m.s. error of only 2·3% in the lattice parameters. This is as accurate as the original isotropic atom–atom model, and is about as good as can be reasonably expected for these temperature-independent calculations. The predicted lattice energies are also reasonable, particularly considering the *ad hoc* method used to estimate the dispersion coefficients.

The hydrogen atoms play a major role in determining the packing of saturated hydrocarbons, allowing the close approach (<3·65 Å) of the terminal CH_3 groups in octane and hexane, where the molecules are parallel, and giving larger separations (≥3·9 Å) in the very different and unique structure adopted by pentane (Norman & Mathisen, 1972). Hence, it is not surprising that isotropic CH_3 and CH_2 potentials, whether fitted to the potential-energy surface or the crystal structures, cannot give even qualitatively correct predictions for the crystal structures. However, the packing effects of the hydrogen atoms can be adequately and efficiently modelled by this simple anisotropic carbon potential.

4. Discussion

This paper develops the use of orientation-dependent model intermolecular pair potentials for saturated hydrocarbons. The results have major implications for both the development of highly accurate potentials for methane, for use in chemical-physics studies, and also for the development of transferable anisotropic carbon site–site potentials to model saturated hydrocarbon fragments in biological molecules.

Although reliable intermolecular potentials for small polyatomics are urgently needed for a wide range of theoretical studies of their microscopic and macroscopic properties, there is, as yet, no recipe for obtaining them, although there has been considerable progress in developing both *ab initio* methods and the analysis of experimental data to obtain information on the potential. This paper mainly addresses the problem of what type of model potentials have the flexibility to enable them to describe the actual potential, an important preliminary, as there is no point in expending considerable effort to obtain accurate *ab initio* or experimental results, and then analysing them in terms of a model potential which is incapable of describing the actual surface. In particular, § 2.1 indicates that one-centred anisotropic potentials will only be appropriate for mildly anisotropic polyatomic molecules, such as AH_n, and § 2.3 showed that the choice of model is important; the shifted models, $U(R, \Omega) = \Phi(\Omega) f[R - s(\Omega)]$, are superior to the scaled Corner models, $U(R, \Omega) = \Phi(\Omega) f[R/s(\Omega)]$, and a linear expansion of the orientation dependence, $U(R, \Omega) = \Phi(\Omega) f(R)$, is a very poor model for the potential.

This study analyses a potential surface for CH_4 which assumes that the molecule is a superposition of spherical charge distributions. Since the atomic electron density rearranges on bond formation, the actual potential probably has a less 'bumpy' repulsive wall than the assumed model. Hence the one-centred anisotropic potentials may model the actual potential better than they model the assumed surface. The electrostatic contribution to the potential is neglected in this study, since saturated hydrocarbons are generally considered to be non-polar, and the point-charge model used by Williams & Starr (1977) gives a negligible (<0·5 kJ mol^{-1}) contribution to the lattice energy of the three saturated hydrocarbons. However, the electrostatic contribution to the methane potential requires careful consideration, since the octupole–octupole interaction has a maximum (destabilizing) energy for the dimer orientation predicted by the isotropic atom–atom potential. An electrostatic contribution can be added to the isotropic atom–atom potential, by using either a central octupole moment $\Omega_{xyz} = \sqrt{15/6} Q_{32s} = 6{\cdot}17 \times 10^{-50}$ C m^3 [from a SCF-CI calculation by Amos (1979)], or the corresponding point-charge model, which has positive charges of 0·189 e on the hydrogen-interaction sites. These two models give different predictions for the dimer structure, since the point-charge model generates higher multipole interactions, and so destabilizes the repulsion–dispersion minimum-energy orientation considerably less than the central octupole model. Thus, the electrostatic contribution, although small, modifies the anisotropy of the potential well, and so needs to be modelled accurately. Since the point-charge model is a gross approximation to the actual charge distribution in methane, and so will be a very poor model for the higher multipole moments, a more detailed knowledge of the electron distribution in methane is required. *Ab initio* methods can be used to calculate the long-range isotropic and anisotropic dispersion coefficients, which would improve the model of this contribution to the potential. The subtleties of the anisotropy of the potential for methane are manifest in the complexity of the orientational ordering phase transitions in CH_nD_{4-n} (Press & Hüller, 1974; Prager, Press & Heidemann, 1981), and a more accurate potential is needed for the analysis of the experimental data on these phases, but the situation is complicated by the possibility of metastable structures and by quantum-mechanical effects.

This work uses a simple anisotropic site–site carbon potential to predict the crystal structures of pentane, hexane and octane, with the same accuracy as and greater efficiency than the original isotropic atom–atom potential. Since these were the only crystal structures containing only unstrained CH_3 and CH_2 groups which were used to derive and test the established isotropic atom–atom potential (Williams & Starr, 1977; Williams & Cox, 1984), our potential is as well proven for these functional groups as the isotropic potential. This paper is not concerned with unsaturated hydrocarbon fragments, because they require a different type of anisotropy to represent the bonded hydrogen atoms. The isotropic atom–atom potential is simpler in that it is used for all hydrocarbons, but it is inflexible, and so the accuracy of this model is limited. In contrast, the anisotropic potentials have the great advantage that they are able to

model the non-sphericity of the bonding electron density, and so can be adapted for different chemical environments. Since the intermolecular potential contribution from any site is primarily determined by the local electron density, we can only confidently transfer site-site potentials between sites which are chemically equivalent, *i.e.* have essentially the same local electron distribution. The π-electron density of the aromatic hydrocarbons gives rise to strong electrostatic quadrupolar interactions which play a major role in determining the molecular packing, and this effect is only crudely modelled by charges on the C and H sites in the isotropic atom-atom potential (Price, 1985). Hence, we expect that anisotropic site-site potentials, which take account of the differences in local bonding environment between saturated, aromatic and strained hydrocarbons (*e.g.* cubane), should be more accurately transferable than any averaged isotropic potential.

In the original derivation of the isotropic atom-atom potential, the molecules were assumed to be rigid and this approximation was used in this work, but is not necessary. The anisotropic carbon potentials could be used to represent the van der Waals forces in molecular-mechanics predictions of molecular conformations and packing. The bending, stretching and torsion about the backbone bonds (C–C–C *etc.*) could be modelled in the usual manner (Burkert & Allinger, 1982), the only difference being that the anisotropic tetrahedral carbon potential would not allow the modelling of deformations within the CH_3 or CH_2 units. However, in most studies, particularly on larger molecules, such deformations are not important, and the saving of computer time in being able to model a methyl group by one site, and yet reproduce its anisotropy, makes this type of potential particularly advantageous.

It must be emphasized that this work only demonstrates the feasibility of modelling CH_3 and CH_2 units by a one-centred potential, at current levels of accuracy. The method of obtaining the potential is extremely derivative, and although the ability to predict the observed crystal structure is a useful test of a potential, it only samples a very limited number of relative orientations. Thus, crystal structure analysis leaves much to be desired as a means of establishing intermolecular potentials. The next stage is to develop a more reliable general method of determining the appropriate potential model and parameters for different molecular units, *via* an examination of the shape of the electron density distribution around each site, and the analysis of experimental sources of information about intermolecular forces.

The overall conclusion from this work, and the azabenzene and chlorine studies, is that it is possible to develop a new simple transferable scheme for the intermolecular potentials of organic molecules, using anisotropic functions, and fewer interaction sites than atoms. The new approach has the compelling attraction that it discards the gross approximation that the potential is composed of spherical atom-atom potentials, and it has the flexibility to be extended to reflect our increasing knowledge of intermolecular forces. Hence, the next generation of model potentials for organic molecules should be more accurate than those currently available, increasing the realism of molecular simulations.

I wish to thank Dr A. J. Stone for useful discussions on this work, the Science and Engineering Research Council for financial support, and the Royal Society for the award of a 1983 University Research Fellowship.

APPENDIX 1

Expansion functions for the orientation dependence of any scalar property of two identical tetrahedral molecules

A general definition of a complete set of expansion functions for the orientation dependence of any scalar property of two molecules of arbitrary shape has been given by Stone (1978):

$$S^{k_1k_2}_{l_1l_2j}(\Omega) = i^{l_1-l_2-j} \times D^{l_1}_{m_ak_1}(\Omega_1)^* D^{l_2}_{m_2k_2}(\Omega_2)^* D^{j}_{m0}(\Omega_{12})^* \begin{pmatrix} l_1 & l_2 & j \\ m_1 & m_2 & m \end{pmatrix},$$

where the $D^l_{mk}(\Omega)$ are Wigner rotation matrices, the term in brackets is a Wigner $3j$ coupling coefficient (Brink & Satchler, 1968) and Ω_1, Ω_2 and Ω_{12} define the orientations respectively of the two molecules and the intermolecular vector $\mathbf{R} = \mathbf{R}_2 - \mathbf{R}_1$, with respect to an arbitrary space-fixed axis system. [We have used Ω to denote the set of orientations $(\Omega_1, \Omega_2, \Omega_{12})$.] Summation over m_1, m_2 and m is implied. This orthogonal set of S functions has many convenient properties for describing bimolecular potentials, being easy to evaluate and to differentiate to give the associated forces and torques analytically (Stone, 1978). These functions arise naturally in the perturbation expansion of the energy of a pair of molecules at long range, which is expressed in terms of the permanent electrostatic multipole moments, static and dynamic polarizabilities of the individual molecules (Stone & Tough, 1984). This Appendix is only concerned with deriving some of the members of this set of functions which are appropriate for tetrahedral molecules, extending the analysis by Stone (1978), and expressing the functions in a form which is useful for implementation in crystal-structure-analysis programs, as described in § 3.

The only S functions which can appear in the description of a scalar property of a pair of molecules

are those combinations of the $D^l_{mk}(\Omega)$ which are symmetric under the symmetry groups of both molecules, and of the combined system. In order to take account of the symmetry of the isolated molecules, we note that, under molecular symmetry operations, $D^l_{mk}(\Omega)$ transforms in the same way as the spherical harmonic $Y_{lk}(\Omega)$ (Brink & Satchler, 1968), so we can use standard group-theoretical methods, or published tables (*e.g.* Stone, 1979; Watanabe, 1966) to determine the allowed values of l_1, k_1, l_2 and k_2. The fourfold screw axis of a tetrahedral molecule implies that $l+k+k/2$ is even (Watanabe, 1966), and so the allowed combinations of l_1 with k_1, and also l_2 with k_2, are $l=0$ $k=0$, $l=3$ $k=\pm 2$, $l=4$ $k=0, \pm 4$ *etc.* The twofold rotation axis $C_2(x)$ implies that the coefficients of $S^{-k_1k_2}_{l_1l_2j}$ are related to those of $S^{k_1k_2}_{l_1l_2j}$ by a factor of $(-)^{l_1}$. Secondly, we need to consider the properties of the entire system. Since we only will be considering properties which are invariant under inversion, l_1+l_2+j must be even (Stone, 1978), and we can define a general normalization

$$\bar{S}^{k_1k_2}_{l_1l_2j} = S^{k_1k_2}_{l_1l_2j}\begin{pmatrix} l_1 & l_2 & j \\ 0 & 0 & 0 \end{pmatrix}^{-1},$$

which gives the useful interpretive property that the magnitude of $\bar{S}^{k_1k_2}_{l_1l_2j}$ is less than or equal to unity (Stone & Tough, 1984). Since the molecules are identical, the coefficients of $\bar{S}^{k_1k_2}_{l_1l_2j}$ and $\bar{S}^{k_2k_1}_{l_2l_1j}$ must be equal. Thus the expansion functions generated by considering l_1, $l_2=0,3$ are $\bar{S}_{000}=1$, $(\bar{S}^{20}_{303}-\bar{S}^{-20}_{303}+\bar{S}^{02}_{033}-\bar{S}^{0-2}_{033})$, and $(\bar{S}^{22}_{33j}-\bar{S}^{-22}_{33j}-\bar{S}^{2-2}_{33j}+\bar{S}^{-2-2}_{33j})$ for $j=0,2,4,6$. The symmetry properties of a tetrahedral molecule transform Y_{40} and $Y_{4\pm4}$ between each other, and so there is only one independent term arising from $\bar{S}^{00}_{404}$ and $(\bar{S}^{40}_{404}+\bar{S}^{-40}_{404}+\bar{S}^{04}_{044}+\bar{S}^{0-4}_{044})$, which can be obtained by projecting out the totally symmetric component under the molecular symmetry operators of both molecules from either expression. Similarly, there is only one independent term arising from $\bar{S}^{k_1k_2}_{440}$. These eight terms give all the terms in the expansion with $l_1, l_2 \le 3$ and all terms of total rank up to eight, and were sufficient for the purposes of this work. Naturally there are further members, based on $\bar{S}_{44j}$, $j=2,4,6,8$, and $\bar{S}^{20}_{343}$ *etc.*, as well as terms involving higher values of l_1 and l_2.

The advantage of this general definition is that we may choose an axis system which is convenient for the application. For many calculations which use intermolecular potentials, including crystal structure analysis, it is convenient to describe the relative orientation of the molecules in terms of the unit intermolecular vector $\hat{\mathbf{R}}$, and the unit vectors $\mathbf{x}$, $\mathbf{y}$, $\mathbf{z}$, which define the molecule fixed axes. (The axis vectors can be calculated from the positions of the atoms bonded to the origin atom, as described in § 3.) Then $\bar{S}^{00}_{101}=\mathbf{z}_1 . \hat{\mathbf{R}}$, $\bar{S}^{00}_{011}=-\mathbf{z}_2 . \hat{\mathbf{R}}$ and $\bar{S}^{00}_{110}=\mathbf{z}_1 . \mathbf{z}_2$. The formulae for non-zero values of k_1, k_2 can be derived using shift operators, and the higher-rank S functions can be derived hierarchically, using the formulae for the products of two S functions. These methods of deriving the $\bar{S}^{k_1k_2}_{l_1l_2j}$ in terms of the scalar products of unit vectors are demonstrated in some detail by Price, Stone & Alderton (1984), where they are used to derive explicit formulae for the electrostatic energy, forces and torques for a pair of molecules of any symmetry.

The relationships between the direction cosines $[(\mathbf{x}_1 . \hat{\mathbf{R}})^2+(\mathbf{y}_1 . \hat{\mathbf{R}})^2+(\mathbf{z}_1 . \hat{\mathbf{R}})^2=1$, $\mathbf{x}_1 . \mathbf{y}_2=(\mathbf{z}_1 . \mathbf{x}_2)\times(\mathbf{y}_1 . \mathbf{z}_2)-(\mathbf{y}_1 . \mathbf{x}_2)(\mathbf{z}_1 . \mathbf{z}_2)$ *etc.*] lead to a plurality of equivalent expressions for the higher expansion functions. The formulae are given here in the form which shows explicitly the equivalence of the three axes for a tetrahedron, which can be obtained from any expression, derived from a set of product formulae, by projecting out the totally symmetric component, using the operator $P \equiv P^{A_1}_{\text{mol}\,1} P^{A_1}_{\text{mol}\,2}$.

Hence, we derive the following formulae for the independent symmetry-adapted expansion functions for the scalar properties of two tetrahedral molecules as:

$$\begin{aligned} Z_{000}(\Omega) &= \bar{S}^{00}_{000} \\ &= 1 \end{aligned}$$

$$\begin{aligned} Z_{303}(\Omega) &= (3\sqrt{10}/10)i(\bar{S}^{20}_{303}-\bar{S}^{-20}_{303}+\bar{S}^{02}_{033}-\bar{S}^{0-2}_{033}) \\ &= 3\sqrt{3}[(\mathbf{x}_1 . \hat{\mathbf{R}})(\mathbf{y}_1 . \hat{\mathbf{R}})(\mathbf{z}_1 . \hat{\mathbf{R}}) \\ &\quad -(\mathbf{x}_2 . \hat{\mathbf{R}})(\mathbf{y}_2 . \hat{\mathbf{R}})(\mathbf{z}_2 . \hat{\mathbf{R}})] \end{aligned}$$

$$\begin{aligned} Z_{330}(\Omega) &= (\bar{S}^{22}_{330}-\bar{S}^{2-2}_{330}-\bar{S}^{-22}_{330}+\bar{S}^{-2-2}_{330}) \\ &= -12P[(\mathbf{x}_1 . \mathbf{x}_2)(\mathbf{y}_1 . \mathbf{y}_2)(\mathbf{z}_1 . \mathbf{z}_2)] \end{aligned}$$

$$\begin{aligned} Z_{404}(\Omega) &= (12/7)P[\bar{S}^{00}_{404}+\bar{S}^{00}_{044}] \\ &\propto P[\bar{S}^{40}_{404}+\bar{S}^{-40}_{404}+\bar{S}^{04}_{044}+\bar{S}^{0-4}_{044}] \\ &= \tfrac{1}{2}\{5[(\mathbf{x}_1 . \hat{\mathbf{R}})^4+(\mathbf{y}_1 . \hat{\mathbf{R}})^4+(\mathbf{z}_1 . \hat{\mathbf{R}})^4+(\mathbf{x}_2 . \hat{\mathbf{R}})^4 \\ &\quad +(\mathbf{y}_2 . \hat{\mathbf{R}})^4+(\mathbf{z}_2 . \hat{\mathbf{R}})^4]-6\} \end{aligned}$$

$$\begin{aligned} Z_{440}(\Omega) &= (12/7)P[\bar{S}^{00}_{440}] \\ &\propto P[\bar{S}^{40}_{440}+\bar{S}^{-40}_{440}+\bar{S}^{04}_{440}+\bar{S}^{0-4}_{440}] \\ &\propto P[\bar{S}^{44}_{440}+\bar{S}^{4-4}_{440}+\bar{S}^{-44}_{440}+\bar{S}^{-4-4}_{440}] \\ &= \tfrac{1}{6}\{5[(\mathbf{x}_1 . \mathbf{x}_2)^4+(\mathbf{x}_1 . \mathbf{y}_2)^4+(\mathbf{x}_1 . \mathbf{z}_2)^4 \\ &\quad +(\mathbf{y}_1 . \mathbf{x}_2)^4+(\mathbf{y}_1 . \mathbf{y}_2)^4+(\mathbf{y}_1 . \mathbf{z}_2)^4+(\mathbf{z}_1 . \mathbf{x}_2)^4 \\ &\quad +(\mathbf{z}_1 . \mathbf{y}_2)^4+(\mathbf{z}_1 . \mathbf{z}_2)^4]-9\} \end{aligned}$$

$$\begin{aligned} Z_{332}(\Omega) &= (\bar{S}^{22}_{332}-\bar{S}^{2-2}_{332}-\bar{S}^{-22}_{332}+\bar{S}^{-2-2}_{332}) \\ &= 15\{3P[(\mathbf{z}_1 . \hat{\mathbf{R}})(\mathbf{z}_2 . \hat{\mathbf{R}})(\mathbf{x}_1 . \mathbf{x}_2)(\mathbf{y}_1 . \mathbf{y}_2)] \\ &\quad -P[(\mathbf{x}_1 . \mathbf{x}_2)(\mathbf{y}_1 . \mathbf{y}_2)(\mathbf{z}_1 . \mathbf{z}_2)]\} \end{aligned}$$

$$
\begin{aligned}
&Z_{334}(\Omega)\\
&\quad=(\bar{S}^{22}_{334}-\bar{S}^{2-2}_{334}-\bar{S}^{-22}_{334}+\bar{S}^{-2-2}_{334})\\
&\quad=-105P[(\mathbf{x}_1\cdot\hat{\mathbf{R}})(\mathbf{y}_1\cdot\hat{\mathbf{R}})(\mathbf{x}_2\cdot\hat{\mathbf{R}})(\mathbf{y}_2\cdot\hat{\mathbf{R}})(\mathbf{z}_1\cdot\mathbf{z}_2)]\\
&\qquad+60P[(\mathbf{z}_1\cdot\hat{\mathbf{R}})(\mathbf{z}_2\cdot\hat{\mathbf{R}})(\mathbf{x}_1\cdot\mathbf{x}_2)(\mathbf{y}_1\cdot\mathbf{y}_2)]\\
&\qquad-6P[(\mathbf{x}_1\cdot\mathbf{x}_2)(\mathbf{y}_1\cdot\mathbf{y}_2)(\mathbf{z}_1\cdot\mathbf{z}_2)]\\
&Z_{336}(\Omega)\\
&\quad=(\bar{S}^{22}_{336}-\bar{S}^{2-2}_{336}-\bar{S}^{-22}_{336}+\bar{S}^{-2-2}_{336})\\
&\quad=(\tfrac{3}{10})\{231(\mathbf{x}_1\cdot\hat{\mathbf{R}})(\mathbf{x}_2\cdot\hat{\mathbf{R}})(\mathbf{y}_1\cdot\hat{\mathbf{R}})\\
&\qquad\times(\mathbf{y}_2\cdot\hat{\mathbf{R}})(\mathbf{z}_1\cdot\hat{\mathbf{R}})(\mathbf{z}_2\cdot\hat{\mathbf{R}})\\
&\qquad-189P[(\mathbf{x}_1\cdot\hat{\mathbf{R}})(\mathbf{y}_1\cdot\hat{\mathbf{R}})(\mathbf{x}_2\cdot\hat{\mathbf{R}})(\mathbf{y}_2\cdot\hat{\mathbf{R}})(\mathbf{z}_1\cdot\mathbf{z}_2)]\\
&\qquad+42P[(\mathbf{z}_1\cdot\hat{\mathbf{R}})(\mathbf{z}_2\cdot\hat{\mathbf{R}})(\mathbf{x}_1\cdot\mathbf{x}_2)(\mathbf{y}_1\cdot\mathbf{y}_2)]\\
&\qquad-2P[(\mathbf{x}_1\cdot\mathbf{x}_2)(\mathbf{y}_1\cdot\mathbf{y}_2)(\mathbf{z}_1\cdot\mathbf{z}_2)]\},
\end{aligned}
$$

where the totally symmetric projections are given by:

$$
\begin{aligned}
&6P[(\mathbf{x}_1\cdot\mathbf{x}_2)(\mathbf{y}_1\cdot\mathbf{y}_2)(\mathbf{z}_1\cdot\mathbf{z}_2)]\\
&\quad=(\mathbf{x}_1\cdot\mathbf{x}_2)(\mathbf{y}_1\cdot\mathbf{y}_2)(\mathbf{z}_1\cdot\mathbf{z}_2)+(\mathbf{x}_1\cdot\mathbf{x}_2)(\mathbf{y}_1\cdot\mathbf{z}_2)(\mathbf{z}_1\cdot\mathbf{y}_2)\\
&\qquad+(\mathbf{x}_1\cdot\mathbf{y}_2)(\mathbf{y}_1\cdot\mathbf{z}_2)(\mathbf{z}_1\cdot\mathbf{x}_2)+(\mathbf{x}_1\cdot\mathbf{y}_2)(\mathbf{y}_1\cdot\mathbf{x}_2)(\mathbf{z}_1\cdot\mathbf{z}_2)\\
&\qquad+(\mathbf{x}_1\cdot\mathbf{z}_2)(\mathbf{y}_1\cdot\mathbf{x}_2)(\mathbf{z}_1\cdot\mathbf{y}_2)+(\mathbf{x}_1\cdot\mathbf{z}_2)(\mathbf{y}_1\cdot\mathbf{y}_2)(\mathbf{z}_1\cdot\mathbf{x}_2)\\
&18P[(\mathbf{z}_1\cdot\hat{\mathbf{R}})(\mathbf{z}_2\cdot\hat{\mathbf{R}})(\mathbf{x}_1\cdot\mathbf{x}_2)(\mathbf{y}_1\cdot\mathbf{y}_2)]\\
&=(\mathbf{z}_1\cdot\hat{\mathbf{R}})(\mathbf{z}_2\cdot\hat{\mathbf{R}})[(\mathbf{x}_1\cdot\mathbf{x}_2)(\mathbf{y}_1\cdot\mathbf{y}_2)+(\mathbf{x}_1\cdot\mathbf{y}_2)(\mathbf{y}_1\cdot\mathbf{x}_2)]\\
&\quad+(\mathbf{z}_1\cdot\hat{\mathbf{R}})(\mathbf{x}_2\cdot\hat{\mathbf{R}})[(\mathbf{x}_1\cdot\mathbf{y}_2)(\mathbf{y}_1\cdot\mathbf{z}_2)+(\mathbf{x}_1\cdot\mathbf{z}_2)(\mathbf{y}_1\cdot\mathbf{y}_2)]\\
&\quad+(\mathbf{z}_1\cdot\hat{\mathbf{R}})(\mathbf{y}_2\cdot\hat{\mathbf{R}})[(\mathbf{x}_1\cdot\mathbf{z}_2)(\mathbf{y}_1\cdot\mathbf{x}_2)+(\mathbf{x}_1\cdot\mathbf{x}_2)(\mathbf{y}_1\cdot\mathbf{z}_2)]\\
&\quad+(\mathbf{y}_1\cdot\hat{\mathbf{R}})(\mathbf{y}_2\cdot\hat{\mathbf{R}})[(\mathbf{x}_1\cdot\mathbf{x}_2)(\mathbf{z}_1\cdot\mathbf{z}_2)+(\mathbf{z}_1\cdot\mathbf{x}_2)(\mathbf{x}_1\cdot\mathbf{z}_2)]\\
&\quad+(\mathbf{y}_1\cdot\hat{\mathbf{R}})(\mathbf{x}_2\cdot\hat{\mathbf{R}})[(\mathbf{z}_1\cdot\mathbf{y}_2)(\mathbf{x}_1\cdot\mathbf{z}_2)+(\mathbf{z}_1\cdot\mathbf{z}_2)(\mathbf{x}_1\cdot\mathbf{y}_2)]\\
&\quad+(\mathbf{y}_1\cdot\hat{\mathbf{R}})(\mathbf{z}_2\cdot\hat{\mathbf{R}})[(\mathbf{z}_1\cdot\mathbf{x}_2)(\mathbf{x}_1\cdot\mathbf{y}_2)+(\mathbf{z}_1\cdot\mathbf{y}_2)(\mathbf{x}_1\cdot\mathbf{x}_2)]\\
&\quad+(\mathbf{x}_1\cdot\hat{\mathbf{R}})(\mathbf{x}_2\cdot\hat{\mathbf{R}})[(\mathbf{y}_1\cdot\mathbf{y}_2)(\mathbf{z}_1\cdot\mathbf{z}_2)+(\mathbf{y}_1\cdot\mathbf{z}_2)(\mathbf{z}_1\cdot\mathbf{y}_2)]\\
&\quad+(\mathbf{x}_1\cdot\hat{\mathbf{R}})(\mathbf{y}_2\cdot\hat{\mathbf{R}})[(\mathbf{y}_1\cdot\mathbf{z}_2)(\mathbf{z}_1\cdot\mathbf{x}_2)+(\mathbf{y}_1\cdot\mathbf{x}_2)(\mathbf{z}_1\cdot\mathbf{z}_2)]\\
&\quad+(\mathbf{x}_1\cdot\hat{\mathbf{R}})(\mathbf{z}_2\cdot\hat{\mathbf{R}})[(\mathbf{y}_1\cdot\mathbf{x}_2)(\mathbf{z}_1\cdot\mathbf{y}_2)+(\mathbf{y}_1\cdot\mathbf{y}_2)(\mathbf{z}_1\cdot\mathbf{x}_2)]\\
&9P[(\mathbf{x}_1\cdot\hat{\mathbf{R}})(\mathbf{y}_1\cdot\hat{\mathbf{R}})(\mathbf{x}_2\cdot\hat{\mathbf{R}})(\mathbf{y}_2\cdot\hat{\mathbf{R}})(\mathbf{z}_1\cdot\mathbf{z}_2)]\\
&=(\mathbf{x}_1\cdot\hat{\mathbf{R}})(\mathbf{y}_1\cdot\hat{\mathbf{R}})[(\mathbf{x}_2\cdot\hat{\mathbf{R}})(\mathbf{y}_2\cdot\hat{\mathbf{R}})(\mathbf{z}_1\cdot\mathbf{z}_2)\\
&\qquad+(\mathbf{y}_2\cdot\hat{\mathbf{R}})(\mathbf{z}_2\cdot\hat{\mathbf{R}})(\mathbf{z}_1\cdot\mathbf{x}_2)+(\mathbf{z}_2\cdot\hat{\mathbf{R}})(\mathbf{x}_2\cdot\hat{\mathbf{R}})(\mathbf{z}_1\cdot\mathbf{y}_2)]\\
&\quad+(\mathbf{x}_1\cdot\hat{\mathbf{R}})(\mathbf{z}_1\cdot\hat{\mathbf{R}})[(\mathbf{x}_2\cdot\hat{\mathbf{R}})(\mathbf{y}_2\cdot\hat{\mathbf{R}})(\mathbf{y}_1\cdot\mathbf{z}_2)\\
&\qquad+(\mathbf{y}_2\cdot\hat{\mathbf{R}})(\mathbf{z}_2\cdot\hat{\mathbf{R}})(\mathbf{y}_1\cdot\mathbf{x}_2)+(\mathbf{z}_2\cdot\hat{\mathbf{R}})(\mathbf{x}_2\cdot\hat{\mathbf{R}})(\mathbf{y}_1\cdot\mathbf{y}_2)]\\
&\quad+(\mathbf{y}_1\cdot\hat{\mathbf{R}})(\mathbf{z}_1\cdot\hat{\mathbf{R}})[(\mathbf{x}_2\cdot\hat{\mathbf{R}})(\mathbf{y}_2\cdot\hat{\mathbf{R}})(\mathbf{x}_1\cdot\mathbf{z}_2)\\
&\qquad+(\mathbf{y}_2\cdot\hat{\mathbf{R}})(\mathbf{z}_2\cdot\hat{\mathbf{R}})(\mathbf{x}_2\cdot\mathbf{x}_2)+(\mathbf{z}_2\cdot\hat{\mathbf{R}})(\mathbf{x}_2\cdot\hat{\mathbf{R}})(\mathbf{x}_1\cdot\mathbf{y}_2)].
\end{aligned}
$$

The normalization adjustments have been made to give Z_{303} a maximum value of 2 when the C—H bonds are pointing directly at each other, Z_{404} a maximum of 2 when two of the bond bisectors are collinear, and Z_{440} a maximum value of 1 when the axes are parallel.

Thus, we have derived the first eight terms in the expansion of the scalar properties of two like tetrahedral molecules, so that they implicitly include all the symmetry contained in the usual definitions which use tetrahedral rotor functions (James & Keenan, 1959; Maki, Kataoka & Yamamoto, 1979), but are in a simple form which can be computed very efficiently. This approach has already been used for $C_{\infty h}$ and approximate D_{3h}-symmetry molecular fragments, in the studies on Cl_2 and the azabenzenes (Price & Stone, 1982, 1984), but the power, as well as maximum complexity, of this type of analysis is shown particularly clearly for the higher symmetry of tetrahedral molecules.

APPENDIX 2

An investigation of the van der Waals philosophy for predicting molecular packing

The use of van der Waals surfaces, either as physical space-filling molecular models or as the concept that two non-bonded atoms cannot be closer than the sum of their van der Waals radii, has a long history as a quick but effective way of understanding, checking and predicting molecular packing. The value of this approach is exemplified by Kitaigorodsky's (1973) use of space-filling molecular models to predict the crystal structures of hydrocarbons and, more recently, by the use of molecular graphics to analyse possible drug–receptor dockings. The use of this 'excluded-volume' model is equivalent to representing the intermolecular potential as a step function, which is a very crude model for the forces which determine the packing. However, it is easily applied by the non-specialist, and so will continue to be widely used when only a qualitative understanding of the molecular packing is required, such as checking the plausibility of X-ray structural refinements, or when a full simulation with a better model intermolecular potential is too difficult. The orientation-dependent functions, which were used to analyse the potential-energy surface for methane in the main paper, can also be used to investigate the use of van der Waals surfaces as an approximation to the more realistic atom–atom potential.

The first test is to determine whether the minimum-energy separation for a fixed relative orientation of two methane molecules, on the isotropic atom–atom potential-energy surface, corresponds to a constant sum of van der Waals radii. This was investigated by finding the shortest intersite separation κ at the minimum-energy separation for each orientation, and expanding $\kappa(\Omega)$ in terms of the $Z_{l_1l_2j}(\Omega)$, as in (2). κ always corresponded to an H H intersite separation. Although κ varied between 2·726 and 3·273 Å, the r.m.s. error in the minimum-energy separation predicted by assuming that the shortest intersite distance was constant [i.e. $\kappa(\Omega)=\kappa_{000}=2{\cdot}8497$ Å] was only

0·0985 Å, showing that assuming that the shortest H H distance is constant at the mininum-energy contour is a far better approximation than assuming that the centre-of-mass separation is constant (r.m.s. error 0·20 Å). Since the isotropic atom–atom potential assumes that the potential from each H site is spherically symmetric, $\kappa(\Omega)$ would be constant if the only contribution to the potential came from the closest intermolecular pair of sites. Hence, the error in treating $\kappa(\Omega)$ as constant is a measure of the importance of the other site–site interactions in determining the minimum-energy separation. The convergence of the expansion of the variations in $\kappa(\Omega)$ with orientation is very poor, with r.m.s. error of 0·0679 Å for an eight-term expansion. Hence, there is no simple model for an anisotropic van der Waals radius for hydrogen which would model the effects of all the interactions, and give a significantly better prediction of the minimum-energy separation corresponding to an isotropic atom–atom potential.

The usefulness of the van der Waals radii approach to approximating the anisotropy of the potential surface, and predicting packing, can be investigated more directly by comparing the centre-of-mass separations corresponding to the first contact between the van der Waals surfaces of the molecules $R_v(\Omega)$ with the minimum-energy separation $\rho(\Omega)$. As discussed above, the most appropriate van der Waals surface model for an isotropic atom–atom potential surface is the usual superposition of spheres of radii r_C and r_H centred at the C and H interaction sites. However, the choice of values for r_C and r_H is not well defined. As well as using published radii, such as those of Bondi (1964), we can also derive radii specifically appropriate to the surface. We have already derived $r_H = \frac{1}{2}\kappa_{000}$ as an appropriate choice of radii to reproduce $\rho(\Omega)$ from the investigation of $\kappa(\Omega)$, the shortest intersite separation at the minimum-energy separation $\rho(\Omega)$. The choice of corresponding carbon radius is immaterial, since all the closest intersite contacts were H H. Other appropriate radii can be defined by noting that the assumption that the intermolecular potential is dominated by the shortest distance between shapes representing the molecules is implicit in the use of van der Waals radii to analyse the packing of the molecules. This assumption is reasonable, because the closest parts of the charge clouds will be interacting most strongly, and this interaction is particularly dominant when there is overlap of the charge clouds, because of the exponential nature of the resulting repulsion. Thus, the sum of the van der Waals radii of two atoms should correspond to a separation where the intermolecular interaction between the atoms is sufficiently strong that the intermolecular forces from other interactions cannot move the atoms closer. Hence, we can define a van der Waals surface by an equipotential of the dominant interaction, and use the separations at which the H H, C H and C C interactions have the same specified energy to define an effective van der Waals radius. To be precise, we define $R_e(\Omega)$ as the largest centre-of-mass separation at which there is an intersite contact with energy E_{con}. The case of $E_{con} = 0$ corresponds to a H H separation of 2·910 Å, a C C separation of 3·466 Å and a C H separation of 3·190 Å, giving effective radii $r_H \simeq 1{\cdot}455$ Å, $r_C \simeq 1{\cdot}733$ Å, and a tiny error in adding the effective radii for the C H contact.

Table 5. *Comparison of different van der Waals surfaces*

Expansion of R, the centre-of-mass separation, at point of contact between surfaces defined by the van der Waals radii:

	Bondi	$\frac{1}{2}\kappa_{000}$	$E_{con} = 0$ kJ mol^{-1}	$E_{con} = 0{\cdot}5$ kJ mol^{-1}
r_C (Å)	1·7	0·0	≃1·733	≃1·666
r_H (Å)	1·2	1·4249	≃1·455	≃1·246
Coefficient of				
Z_{000} (Å)	3·8035	4·2314	4·2951	3·8793
Z_{303} (Å)	0·2833	0·3381	0·3362	0·3123
Z_{330} (Å)	−0·0293	−0·0297	−0·0296	−0·0325
Z_{404} (Å)	−0·1256	−0·0824	−0·0842	−0·1109
Z_{440} (Å)	−0·0142	−0·0249	−0·0240	−0·0195
Z_{332} (Å)	0·0311	0·0263	0·0252	0·0321
Z_{334} (Å)	−0·0142	−0·0160	−0·0159	−0·0207
Z_{336} (Å)	−0·0775	−0·0953	−0·0904	−0·0809
R.m.s. error	0·0575	0·0863	0·0841	0·0711
R.m.s. value	3·8166	4·2420	4·3056	3·8916
Max. value	4·439	4·889	4·950	4·533
Min. value	3·4	3·362	3·466	3·331
Number of C C contacts	28	0	1	11
Number of C H contacts	146	0	14	91
Number of H H contacts	282	456	441	354

Table 5 compares the expansion of the centre-of-mass separations corresponding to van der Waals contacts for four different definitions of the van der Waals surfaces, which can be contrasted with the expansion of the minimum-energy separation $\rho(\Omega)$ (Table 2). The anisotropic coefficients are qualitatively similar, with Z_{303}, Z_{404} and Z_{336} being the largest, showing that it is a reasonable crude approximation to map out the potential surface by considering the closest contacts between molecules. However, at a more quantitative level, there are significant differences. The similarity in the expansion of $R_v(\Omega)$ for the Bondi radii and $R_e(\Omega)$ for $E_{con} = 0{\cdot}5$ kJ mol^{-1}, and also for $R_v(\Omega)$ defined by κ_{000} and $R_e(\Omega)$ for $E_{con} = 0$, can be attributed to the similarity in the radii, emphasizing that van der Waals radii can be considered as energy contours. The expansions are different for these two sets of radii because they correspond to qualitatively different shapes, and hence have a different proportion of C C, C H and H H closest contacts. The carbon radii for the Bondi and $E_{con} = 0{\cdot}5$ kJ mol^{-1} sets are relatively large, so that the carbon sphere protrudes above the plane which is tangential to three hydrogen spheres. The carbon sphere is below this plane for the radii defined by $E_{con} = 0$. The hydrogen-only surface is a superposition of spheres with a dimple at the centre of the face. The latter two shapes correctly predict that the

molecules can get closest for the global minimum orientation. This success in predicting a global minimum which has six equivalent closest H H contacts, occurs despite the fact that the van der Waals model does not take account of the number of close contacts, probably because small centre-of-mass separations are inevitably associated with multiple contacts for methane molecules. However, the surfaces with a protruding carbon sphere give rise to many orientations with C C as the closest contact, and so cannot predict the dimer structure.

This study shows that the use of van der Waals radii can be remarkably successful if the optimum radii for the actual potential are used, giving a useful mapping of the anisotropy of the repulsion–dispersion potential. However, plausible radii can give misleading results. The appropriate 'radii' should be orientation dependent for atoms with non-spherical charge distributions; an analysis of close contacts in molecular crystal structures by Nyburg & Faerman (1985) has demonstrated that the effective van der Waals radius is indeed anisotropic for some heteroatoms. Thus, the assumption of spherical atoms, or inappropriate molecular shapes, may well cause significantly worse errors in the use of van der Waals radii to analyse molecular packing than the intrinsic error in modelling molecules as hard shapes.

References

AMOS, R. D. (1979). *Mol. Phys.* **38**, 33-45.

BERNE, B. J. & PECHUKAS, P. (1972). *J. Chem. Phys.* **56**, 4213-4216.

BONDI, A. (1964). *J. Phys. Chem.* **68**, 441-451.

BRINK, D. M. & SATCHLER, G. R. (1968). *Angular Momentum.* Oxford: Clarendon Press.

BUCK, U., HUISKEN, F., KOHLHASE, A., OTTEN, D. & SCHAEFER, J. (1983). *J. Chem. Phys.* **78**, 4439-4450.

BUCK, U., SCHLEUSENER, J., MALIK, D. J. & SECREST, D. (1981). *J. Chem. Phys.* **74**, 1707-1717.

BURKERT, U. & ALLINGER, N. L. (1982). *Molecular Mechanics. Am. Chem. Soc. Monogr.* No. 177.

BUSING, W. R. (1981). Report ORNL-5747. Oak Ridge National Laboratory, Tennessee.

CORNER, J. (1948). *Proc. R. Soc. London Ser. A*, **192**, 275-292.

JAMES, H. M. & KEENAN, T. A. (1959). *J. Chem. Phys.* **31**, 12-41.

KIHARA, T. (1978). *Intermolecular Forces.* New York: John Wiley.

KITAIGORODSKY, A. I. (1973). *Molecular Crystals and Molecules.* New York: Academic Press.

MAITLAND, G. C., RIGBY, M., SMITH, E. B. & WAKEHAM, W. A. (1981). *Intermolecular Forces.* Oxford: Clarendon Press.

MAKI, K., KATAOKA, Y. & YAMAMOTO, T. (1979). *J. Chem. Phys.* **70**, 655-674.

MATHISEN, H., NORMAN, N. & PEDERSEN, B. F. (1967). *Acta Chem. Scand.* **21**, 127-135.

NORMAN, N. & MATHISEN, H. (1961). *Acta Chem. Scand.* **15**, 1755-1760.

NORMAN, N. & MATHISEN, H. (1972). *Acta Chem. Scand.* **26**, 3913-3916.

NYBURG, S. C. & FAERMAN, C. H. (1985). *Acta Cryst.* **B41**, 274-279.

PACK, R. T. (1974). *J. Chem. Phys.* **60**, 633-639.

PRAGER, M., PRESS, W. & HEIDEMANN, A. (1981). *J. Chem. Phys.* **75**, 1442-1450.

PRESS, W. & HÜLLER, A. (1974). In *Anharmonic Lattices, Structural Transitions and Melting*, edited by T. RISTE, pp. 185-212. Leiden: Noordhoff.

PRICE, S. L. (1985). *Chem. Phys. Lett.* **114**, 359-369.

PRICE, S. L. & STONE, A. J. (1980). *Mol. Phys.* **40**, 805-822.

PRICE, S. L. & STONE, A. J. (1982). *Mol. Phys.* **47**, 1457-1470.

PRICE, S. L. & STONE, A. J. (1984). *Mol. Phys.* **51**, 569-583.

PRICE, S. L., STONE, A. J. & ALDERTON, M. (1984). *Mol. Phys.* **52**, 987-1001.

PROFETA, S. & ALLINGER, N. L. (1985). *J. Am. Chem. Soc.* **107**, 1907-1918.

STONE, A. J. (1978). *Mol. Phys.* **36**, 241-256.

STONE, A. J. (1979). In *The Molecular Physics of Liquid Crystals*, edited by G. R. LUCKHURST & G. W. GRAY, pp. 31-50. New York: Academic Press.

STONE, A. J. & TOUGH, R. J. A. (1984). *Chem. Phys. Lett.* **110**, 123-129.

WALMSLEY, S. H. (1977). *Chem. Phys. Lett.* **49**, 390-392.

WATANABE, H. (1966). *Operator Methods in Ligand Field Theory.* New Jersey: Prentice-Hall.

WILLIAMS, D. E. (1971). *Acta Cryst.* **A27**, 452-455.

WILLIAMS, D. E. & COX, S. R. (1984). *Acta Cryst.* **B40**, 404-417.

WILLIAMS, D. E. & STARR, T. L. (1977). *Comput. Chem.* **1**, 173-177.