J. Indian Inst. Sci. 62 (C), Dec. 1980, Pp. 83–98

The crystal structure of N-acetyl-L-prolyl-D-alanyl-methylamide hemi hydrate by the method of packing analysis on contact criteria

S. RAMAPRASAD*
Molecular Biophysics Unit, Indian Institute of Science, Bangalore 560 012, India.

Received on July 24, 1980.

Abstract.

The structure determination of the tripeptide N-acetyl-L-prolyl-D-alanyl-methylamide by the method of packing analysis is described. The molecular geometry deduced from conformational analysis was used to orient and position the molecules in the unit cell and various trial crystal structures were generated and refined against the X-ray data in a systematic way. Out of the eight reasonable packing arrangements, one refined to an R-value of 0·30 by the rigid body refinement procedure, when 50 low-angle reflections were used. This model was further refined by a series of conventional least-square procedures gradually increasing the number of reflections at each stage to a final R value of 0·062.

Key words : Contact criteria, trial model, packing analysis, rigid body refinement, least-squares refinement.

1. Introduction

The theory of biopolymer conformation and the stability of crystal structures are so interrelated[1] that potential energy functions which have been successfully applied to protein conformation can equally be exploited for solving crystal structures of simple peptides and other molecules whose shapes are otherwise known. Such techniques are very well documented. For example, in this laboratory, the potential functions were used to locate the position and orientation of rigid molecules of known geometry like benzene and sulphur dioxide[2] in their respective crystal structures. A similar packing method has been used by Kitaigorodsky[3] and Williams[4]. In these analyses, the packing energy is generally minimised. On the other hand, the packing analysis can be based just on the set of contact distance criteria originally proposed by Ramachandran and his colleagues to predict successfully the allowed ranges of dihedral angles (ϕ, ψ) for

* Present address : Department of Chemistry, University of California, Davis 95616, USA.

a system of two linked peptides[5]. This procedure bypasses the evaluation of potential energy and thus saves the corresponding computation time involved in such calculation. Recently this new technique in conjunction with the X-ray intensity data has been tested in deriving the crystal structure of cyclo (Gly-L-Tyr-Gly)$_2$. $2H_2O$, in the space group P1[6]. As part of the conformational studies on N-acetyl-LPro-D-Ala-Methylamide, the approximate shape of the tripeptide molecule was known from the results of spectroscopic studies and empirical energy calculations. The crystal data such as space group and unit cell dimensions were also available from preliminary X-ray investigations. Hence this method of crystal structure determination was undertaken. The principles and details of the method are described below.

2. Method

2.1. *Contact search*

When the shape of the molecule is known, it can be treated as a single rigid entity for generating the trial crystal structures. The orientation and the position of the molecule in the unit cell can be defined by means of the three Eulerian angles θ_1, θ_2, θ_3[7] and the translation parameters u, v and w, which define the fractional coordinates of the origin of a local frame of reference fixed in the molecule. For the sake of convenience, we shall restrict ourselves to the orthorhombic space groups of cell dimensions a, b and c. The orientation of the molecule, with respect to the unit cell frame of reference, is varied systematically by stepping through the three Eulerian angles. The transformation of the molecular frame of reference (x_L, y_L, z_L) to the crystal frame of reference (x, y, z) can be then represented by

$$\begin{pmatrix} ax \\ by \\ cz \end{pmatrix} = R \begin{pmatrix} x_L \\ y_L \\ z_L \end{pmatrix} + \begin{pmatrix} au \\ bv \\ cw \end{pmatrix} \tag{1}$$

where x, y, z are the fractional coordinates and

$$R = \begin{pmatrix} \cos\theta_1 \cos\theta_3 - \sin\theta_1 \cos\theta_2 \cos\theta_3 & -\cos\theta_1 \sin\theta_3 - \sin\theta_1 \cos\theta_2 \cos\theta_3 & \sin\theta_1 \sin\theta_2 \\ \sin\theta_1 \cos\theta_3 + \cos\theta_1 \cos\theta_2 \sin\theta_3 & -\sin\theta_1 \sin\theta_3 + \cos\theta_1 \cos\theta_2 \cos\theta_3 & -\cos\theta_1 \sin\theta_2 \\ \sin\theta_2 \sin\theta_3 & \sin\theta_2 \cos\theta_3 & \cos\theta_2 \end{pmatrix} \tag{2}$$

For each orientation, the three translation parameters are also varied systematically in any desired range. The symmetry operations of the space group are then applied to the asymmetric unit coordinates (x, y, z) to generate the trial models of crystal packing. However, in reality, many of them can be judiciously rejected in terms of

short contacts between neighbouring molecules. An acceptable packing arrangement is one in which the intermolecular contacts are generally above the threshold distances.

More quantitatively, the interatomic contacts between the reference molecule (say at x, y, z) and its neighbours are computed and compared with the corresponding extreme limits[8]. If the contacts (d) are lower than the extreme limits ($d_{e.l.}$) the differences are summed to yield the parameter $\triangle$ given by

$$\triangle = \Sigma\,(d_{e.l.} - d) \qquad (3)$$

over all short contacts

This provides a measure for comparing competing models.

The contact search program (for space group P1), written for an IBM 360/44 computer with 32 K words memory by Dr. Shamala of this laboratory, was modified for the space group $P2_12_12_1$.

The essential steps of the program are:

(1) The atomic coordinates (x, y, z) of the reference molecule for any orientation (θ_1, θ_2, θ_3) and position (u, v, w) in the unit cell are first calculated using eqn. (1).

(2) All the four molecules for the space group $P2_12_12_1$ are generated using the symmetry elements.

(3) The intermolecular contacts between the atoms in the reference molecule and those in the neighbouring molecules (n_m) are evaluated and compared with the corresponding extreme limits. The value of $\triangle$ is also evaluated.

(4) If the trial model satisfies the selection criterion for $\triangle$, then those contacts, below the extreme limits are listed. Otherwise, the control is transferred to step (1) to produce the model corresponding to the next grid point in the six-dimensional space.

(5) To select those models, which possess NH...O intermolecular hydrogen bonds, the donor and acceptor groups and the limiting values of hydrogen bond length and bond angle are specified in the input. In such a case, the allowed models are further checked for intermolecular hydrogen bonds. If found possible, the related parameters are printed by the program.

2.2. *Rigid body structure factor least-squares (RBLS) refinement*

The method of refining only the rigid body parameters against the X-ray intensity data should offer the following advantages over the usual individual atom refinement.

1. The entire molecule is made to shift as a single entity, thereby avoiding improbable values of the geometrical and thermal parameters.

2. The number of parameters to be refined are drastically reduced and the same applies to the order of the matrix to be computed and inverted.

The packing model was refined in two stages. In the first stage, the trial model was refined by the rigid body least-squares procedure. In the second stage, the model so obtained was refined by the conventional least-squares procedure.

The principles of the method and the mathematical procedure of rigid body least-squares program are presented here. A flow chart of the program is given in Fig. 3. Similar constrained refinements of crystal structures of organic molecules have been successfully carried out by Damiani *et al*[12], Williams[13], Pawley[14] and Rao[15]. The method described in this paper is quite general and concerned with the entire molecule being treated as a rigid body for the purpose of refinements.

The fractional coordinates so generated are then used to compute the structure factors, corresponding to the observed reflections, using the following equation.

$$F_c = K \cdot F_{c_0} e^{-B \sin^2 \theta/\lambda^2} \tag{4}$$

where B is the thermal parameter, and K the scale factor. F_{c_0} is given by

$$F_{c_0} = \sqrt{A_{hkl}^2 + B_{hkl}^2} \tag{5}$$

where A_{hkl} and B_{hkl} represent the real and imaginary parts of the structure factor for a given hkl reflection. The quantity minimised during the process of refinement is given by

$$R_s = \sum_{i=1}^{M} W_i (\,|\,F_0\,|_i - |\,F_c\,|_i)^2 \tag{6}$$

where i may range over all or a portion of the intensity data as the case may be.

The normal equations which can be used to compute the shifts in the parameters, to minimise the quantity R_s, are given in the matrix notation as

$$Ah = C \tag{7}$$

where A is the derivative matrix given by

$$A_{nm} = \sum_i w_i \frac{\partial (F_c)_i}{\partial p_n} \cdot \frac{\partial (F_c)_i}{\partial p_m} \tag{8}$$

w_i's are the weights associated with particular reflections. h_n = desired shift in the n^{th} parameter.

Similarly,

$$C_n = \sum_i w_i (\,|\,F_c\,|_i - |\,F_o\,|_i) \frac{\partial (F_c)_i}{\partial P_n}. \tag{9}$$

Since the rigid body constraints are to be applied, the number of parameters are reduced considerably. To set up the normal equation, which involve eight parameters,

it is necessary to evaluate the differential of the general type $\partial F_c/\partial p$ where F_c is a function of p. Thus, with respect to θ_1, the derivative is

$$\frac{\partial F_c}{\partial \theta_1} = K \cdot \frac{\partial F_{c_0}}{\partial \theta_1} e^{-B \sin^2 \theta/\lambda^2} \tag{10}$$

the factor $\partial F_{c_0}/\partial \theta_1$ can be written as

$$\frac{\partial F_{c_0}}{\partial \theta_1} = \frac{A_{hkl}}{F_{c_0}} \cdot \frac{\partial A_{hkl}}{\partial \theta_1} + \frac{B_{hkl}}{F_{c_0}} \cdot \frac{\partial B_{hkl}}{\partial \theta_1} \tag{11}$$

and

$$\frac{\partial F_c}{\partial \theta_1} = K \frac{A_{hkl}}{F_{c_0}} \cdot \frac{\partial A_{hkl}}{\partial \theta_1} + \frac{B_{hkl}}{F_{c_0}} \cdot \frac{\partial B_{hkl}}{\partial \theta_1} \tag{12}$$

For the space group $P2_12_12_1$, the quantities A_{hkl} and B_{hkl} are given by

$$A_{hkl} = \sum_{j=1}^{N} 4 \cos 2\pi\left(hx_j - \frac{h-k}{4}\right) \cos 2\pi\left(ky_j - \frac{k-1}{4}\right) \times \cos 2\pi\left(lz_j - \frac{1-h}{4}\right) \tag{13}$$

and

$$B_{hkl} = -\sum_{j=1}^{N} 4 \sin 2\pi\left(hx_j - \frac{h-k}{4}\right) \sin 2\pi\left(ky_j - \frac{k-1}{4}\right) \times \sin 2\pi\left(lz_j - \frac{l-h}{4}\right) \tag{14}$$

But A_{hkl} and B_{hkl} are functions of θ_1, θ_2 and θ_3 and hence we can write

$$\begin{aligned}\frac{\partial F_c}{\partial \theta_1} = K\Bigg[&\frac{A_{hkl}}{F_{c_0}}\Bigg(\sum_{j=1}^{N} f_j \frac{\partial (A_{hkl})_j}{\partial x_j} \cdot \frac{\partial x_j}{\partial \theta_1} \\ &+ \sum_{j=1}^{N} f_j \frac{\partial (A_{hkl})_j}{\partial y_j} \cdot \frac{\partial y_j}{\partial \theta_1} + \sum_{j=1}^{N} f_j \frac{\partial (A_{hkl})_j}{\partial z_j} \cdot \frac{\partial z_j}{\partial \theta_1}\Bigg) \\ &+ \frac{B_{hkl}}{F_{c_0}}\Bigg(\sum_{j=1}^{N} f_j \frac{\partial (B_{hkl})_j}{\partial x_j} \cdot \frac{\partial x_j}{\partial \theta_1} \\ &+ \sum_{j=1}^{N} f_j \frac{\partial (B_{hkl})_j}{\partial y_j} \cdot \frac{\partial y_j}{\partial \theta_1} + \sum_{j=1}^{N} f_j \frac{\partial (B_{hkl})_j}{\partial z_j} \cdot \frac{\partial z_j}{\partial \theta_1}\Bigg)\Bigg]\end{aligned} \tag{15}$$

The derivatives $\partial F_c/\partial \theta_2$ and $\partial F_c/\partial \theta_3$ can be written similarly.

The derivatives with respect to the translational components are given below:

$$\frac{\partial F_c}{\partial u} = K \cdot e^{-B\sin^2\theta/\lambda^2}\left[\frac{A_{hkl}}{F_{c_0}}\frac{\partial (A_{hkl})_j}{\partial x_j} \cdot \frac{\partial x_j}{\partial u} + \frac{B_{hkl}}{F_{c_0}}\frac{\partial (B_{hkl})_j}{\partial x_j} \cdot \frac{\partial x_j}{\partial u}\right] \tag{16}$$

where $\partial x_j/\partial u$ is unity.

The derivatives $\partial x_j/\partial \theta_1$ ..., etc., and $\partial (A_{hkl})_j/\partial x_j$ ..., etc.,

take the form given below where only typical derivative is given, and similar terms can be written on the same lines.

$$\frac{\partial x_j}{\partial \theta_1} = \frac{(x_L)_j}{a}(-\sin\theta_1 \cos\theta_3 - \cos\theta_2 \cos\theta_1 \sin\theta_3) + \frac{(y_L)_j}{a}(\sin\theta_3 \sin\theta_1 - \cos\theta_2 \cos\theta_1 \cos\theta_3) + \frac{(z_L)_j}{a}(\sin\theta_2 \cos\theta_1). \tag{17}$$

Following this, the derivatives with respect to θ_2 and θ_3 can be written down. On the same lines as the one written for x_j, derivatives of y_j and z_j with respect to the variables θ_1, θ_2 and θ_3 follow. However, among them $\partial z_j/\partial \theta_1$ will be zero as it is independent of θ_1.

Typical derivatives of $(A_{hkl})_j$ and $(B_{hkl})_j$ with respect to x_j can be written as

$$\frac{\partial (A_{hkl})_j}{\partial x_j} = -8\pi h \sin 2\pi\left(hx_j - \frac{h-k}{4}\right)\cos 2\pi\left(ky_j - \frac{k-1}{4}\right) \times \cos 2\pi\left(lz_j - \frac{1-h}{4}\right) \tag{18}$$

and

$$\frac{\partial (B_{hkl})_j}{dx_j} = -8\pi l \sin 2\pi\left(hx_j - \frac{h-k}{4}\right)\sin 2\pi\left(ky_j - \frac{k-1}{4}\right) \times \cos 2\pi\left(lz_j - \frac{1-h}{4}\right) \tag{19}$$

Similar expressions with respect to the variables y_j and z_j follow.

The derivatives of F_c with respect to the thermal parameter (B) and the scale factor (K) are given below.

$$\frac{\partial F_c}{\partial K} = e^{-B\sin^2\theta/\lambda^2}\sqrt{A_{hkl}^2 + B_{hkl}^2} \tag{20}$$

$$\frac{\partial F_c}{\partial B} = K(-\sin^2\theta/\lambda^2)\sqrt{A_{hkl}^2 + B_{hkl}^2} \tag{21}$$

These expressions are used for setting up the normal equations $Ah = C$, and are then solved for the shifts (p's). Usually several iterations are necessary to complete the refinement.

3. Application of the method to the tripeptide structure

The structural formula of the molecule N-acetyl-L-prolyl-D-alanyl-methylamide is shown in Fig. 1.

The crystals of $C_{11}H_{19}N_3O_3 \cdot O \cdot 5H_2O$ were grown by evaporation technique using acetone. The crystal data are given in Table I.

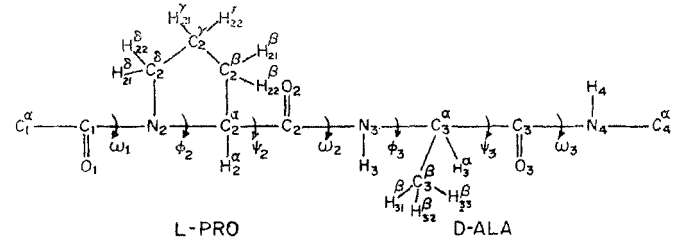

FIG. 1. Structural formula of the molecule showing the backbone dihedral angles.

Table I

Compound	N-acetyl-L-prolyl-D-alanyl methylamide	
Molecular formula	$C_{11}H_{19}N_3O_3 \cdot O \cdot 5H_2O$*	
Formula weight	250·0	
Systematic absences	$h\,O\,O$	h-odd
	$O\,k\,O$	k-odd
	$O\,O\,l$	l-odd
Space group	$P2_12_12_1$	
a	18·91 (1) Å	
b	8·900 (6) Å	
c	8·132 (2) Å	
D_m	1·214 gcm^{-3}	
D_c	1·213 gcm^{-3}	
Z	4	
μ (MoK α)	0·98 cm^{-1}	

* The presence of water was determined only in the final stages of X-ray refinement.

Out of the 975 measurable reflections, collected on a CAD-4 diffractometer, 39 were found to be below 3σ (I) level. By the method of Wilson's intensity statistics[9] the overall temperature factor was found to be 4 Å^2 and the scale factor as 0·84. Two standard reflections were measured after every fifty reflections, and there was no significant change in their intensities during the period of data collection. Intensities were then corrected for Lorentz and Polarisation factors but not for absorption.

3.1. *Generation of the molecular model*

The results from NMR studies and empirical energy calculations were used in deducing the molecular model (manuscript in preparation). NMR results indicated the presence of $4 \rightarrow 1$ internal N→ H...O hydrogen bond between the C-terminal NH and the N-terminal CO. The set of conformational angles (ϕ_2, ψ_2) and (ϕ_3, ψ_3) were estimated to be (−70°, 90°) and (90°, 30°) respectively. Using these values of dihedral angles and the standard bond length and bond angle data for the *trans* planar peptide unit the molecular model was generated. The pyrrolidine ring was maintained in the C^{γ}-*endo* puckering and the required geometrical parameters for this were adopted from the work of Ramachandran *et al*[10].

3.2. *Packing the molecule in the unit cell*

The origin of the molecular reference frame is chosen as shown in Fig. 2, to be approximately at the centroid of the molecule, namely, at the intersection of $C_2^{\alpha} - C_3^{\alpha}$ and perpendicular to this line from C_2^{α}. OC_3^{α} and OC_2^{α} define the X and Y-axes of a right handed coordinate system. Thus the coordinates expressed in the molecular

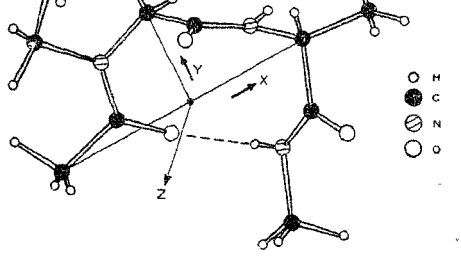

FIG. 2. The molecular geometry and the local coordinate frame of reference.

frame of reference correspond to $\theta_1 = \theta_2 = \theta_3 = 0$ and $u = v = w = 0$ with reference to the crystallographic axes. For any other set, the atomic coordinates of all the atoms are generated by making use of eqn.(1).

The interaction of only those molecules which are within a sphere of 12 Å from the centre of the reference molecule were considered to be significant. For the space group $P2_12_12_1$, the number of independent molecules relevant for the contact search are twenty six and their positions are specified in Table II.

Trial models were generated from the initial values ($\theta_1 = \theta_2 = \theta_3 = 0$, $u = v = w = 0$) by varying the two sets initially at coarse intervals of 60° and 0·8 Å and later at closer intervals of 15° and 0·4 Å respectively. The angle θ_1 was varied in the range 0° to 360° while θ_2 and θ_3 were varied from 0° to 180°. For each set of these values,

Table II

Equivalent positions of molecules generated in the packing analysis

Molecule	Fractional coordinates		
1	x,	y,	z
2	$1+x$,	y,	z
3	x,	$1+y$,	z
4	x,	y,	$1+z$
5	x,	$1+y$,	$1+z$
6	$\frac{1}{2}-x$,	$-y$,	$\frac{1}{2}+z$
7	$\frac{1}{2}-x$,	$1-y$,	$\frac{1}{2}+z$
8	$-\frac{1}{2}-x$,	$-y$,	$\frac{1}{2}+z$
9	$-\frac{1}{2}-x$,	$1-y$	$\frac{1}{2}+z$
10	$-\frac{1}{2}-x$,	$-1-y$,	$\frac{1}{2}+z$
11	$\frac{1}{2}-x$,	$-1-y$,	$\frac{1}{2}+z$
12	$\frac{1}{2}+x$,	$\frac{1}{2}-y$,	$-z$
13	$\frac{1}{2}+x$,	$\frac{1}{2}-y$,	$1-z$
14	$\frac{1}{2}+x$,	$\frac{1}{2}-y$,	$-1-z$
15	$\frac{1}{2}+x$,	$\frac{3}{2}-y$,	$-z$
16	$\frac{1}{2}+x$,	$\frac{3}{2}-y$,	$1-z$
17	$\frac{1}{2}+x$,	$\frac{3}{2}-y$,	$-1-z$
18	$\frac{1}{2}+x$,	$-\frac{1}{2}-y$,	$-z$
19	$\frac{1}{2}+x$,	$-\frac{1}{2}-y$,	$-1-z$
20	$\frac{1}{2}+x$,	$-\frac{1}{2}-y$,	$1-z$
21	$-x$,	$\frac{1}{2}+y$,	$\frac{1}{2}-z$
22	$-x$,	$\frac{1}{2}+y$,	$\frac{3}{2}-z$
23	$-x$,	$\frac{1}{2}+y$	$-\frac{1}{2}-z$
24	$1-x$,	$\frac{1}{2}+y$	$\frac{1}{2}-z$
25	$1-x$,	$\frac{1}{2}+y$	$\frac{3}{2}-z$
26	$1-x$,	$\frac{1}{2}+y$,	$-\frac{1}{2}-z$

the translational parameters were varied from 0 to 1/2. These regions are adequate for the contact analysis since the unit cell has 222 symmetry. As the interval of 60° is quite large, the intermediate regions 30°, 90°, etc., were subsequently scanned using the same interval. All those sterically allowed models which were relatively free from short contacts, $\triangle$ being the indicator, were grouped into eight categories. These are listed in Table III and the various short contacts are given in Table IV.

3.3. *Refinement of the tripeptide structure by RBLS procedure*

All the trial models *A* to *H* given in Table III were considered for the refinement by using fifty low angle reflections with resolution up to 2·8 Å. The six rigid body parameters and *K* and *B* defined earlier were refined against the limited X-ray data. Among the eight models, model B gave the lowest R-value of 0·30 (Table V) at the end of the third cycle of refinement. The remaining models had high R-values in the range 0·58 to 0·77 and were therefore rejected.

Model B which gave the lowest R-value was considered equivalent to the first trial crystal structure one would obtain from an E-map derived from direct methods. This trial structure was then refined against the X-ray data carefully in stages, by the conventional structure factor least squares refinement. In each stage more number of reflections from higher $\sin\theta/\lambda$ range were included in a stepwise manner, every time refin-

Table III

The six packing parameters and the value of $\triangle$ for models A to H during the contact search

$\triangle = \Sigma(d_{e.l.} - d)$, where $d_{e.l.}$ = extreme limit and d = observed value.

Model	θ_1	θ_2	θ_3	u	v	w	$\triangle$
A	90	30	150	0·15	0·40	0·40	0·00
B*	330	90	30	0·35	0·20	0·00	0·16
C	0	120	60	0·05	0·25	0·00	0·17
D	30	90	30	0·40	0·40	0·40	0·19
E	330	150	30	0·10	0·40	0·00	0·20
F	30	90	30	0·40	0·00	0·40	0·30
G	30	30	150	0·15	0·10	0·00	0·50
H	30	30	150	0·20	0·20	0·35	0·60

* This model refined to the final crystal structure.

Table IV

Short nonbonded contacts (Å) in the eight selected models. The *extreme limits* are given within parentheses

Model	Contact	Distance
A	..	..
B	C(4)...C(11)	2·84 (3·0)
C	H(N2)...O(2)	2·12 (2·2)
	H′(C4)...H(C5)	1·81 (1·9)
D	H″(C9)...C(11)	2·07 (2·2)
	C(9)...C(11)	2·94 (3·0)
E	C(9)...O(3)	2·60 (2·7)
	H′(C9)...O(3)	2·10 (2·2)
F	H′(C5)...O(1)	2·13 (2·2)
	H(C6)...C(1)	2·14 (2·2)
	C(9)...C(11)	2·94 (3·0)
	H″(C9)...C(11)	2·10 (2·2)
G	C(11)...H(C3)	2·15 (2·2)
	O(3)...O(2)	2·25 (2·7)
H	H′(C3)...H(C6)	1·70 (1·9)
	H(C4)...O(1)	1·90 (2·2)
	O(2)...C(1)	2·60 (2·7)

ing the structure over three cycles and using the refined coordinates at the end of third cycle of refinement for further studies. The details of refinement are summarized in Table VI. The gradual improvements in the distribution of bond lengths and thermal parameters may be noted. Finally, an R-value of 0·062 was achieved using all the available reflections (resolution of 0·9 Å). The final crystal structure includes half a molecule of water which was located in the difference electron density map based on the 1·2 Å data. The details are presented elsewhere (Ramaprasad and Chandrasekaran, under preparation). It is interesting to note that the trial model was devoid of any intermolecular hydrogen bond (see Table VII). In the final crystal structure NH (DAla) formed a hydrogen bond with CO(DAla) of another molecule —2·93 Å and 18·5°—while the corresponding values in the beginning were as large as 3·8 Å and 67·3°.

Table V

Packing parameters and X-ray R-value of the eight models using fifty low-angle reflections in the rigid body least squares refinement

Model		Orientational parameters in degrees			Translational parameters in fractional units			R-value %
		θ_1	θ_2	θ_3	u	v	w	
A	Initial	90	30	150	0·15	0·40	0·40	74·0
	Final	117	35	121	0·16	0·35	0·41	67·0
B	Initial	330	90	30	0·35	0·20	0·00	58·0
	Final	341	98	32	0·37	0·22	0·07	30·0
C	Initial	0	120	60	0·05	0·25	0·00	70·5
	Final	1	124	64	0·07	0·30	0·11	58·7
D	Initial	30	90	30	0·40	0·40	0·40	70·0
	Final	22	93	34	0·40	0·41	0·45	57·8
E	Initial	330	150	30	0·10	0·40	0·00	77·0
	Final	359	135	56	0·10	0·40	0·07	53·4
F	Initial	30	90	30	0·40	0·00	0·40	73·0
	Final	28	84	30	0·40	0·02	0·50	54·5
G	Initial	30	30	150	0·15	0·10	0·00	73·0
	Final	0	29	173	0·17	0·06	0·02	68·0
H	Initial	30	30	150	0·20	0·20	0·35	71·0
	Final	22	27	146	0·20	0·20	0·32	65·5

Table VI

Summary of the various stages of conventional stepwise X-ray refinement of model B

No. of reflections	50	200	300	500	975
Average shift (in Å)	0·7	0·08	0·05	0·02	0·004
Bond length		1·0–1·8 Å	1·0–1·7 Å	1·1–1·6 Å	1·2–1·6 Å
Temperature factor	6·3 Å²	1·7–11·1 Å² (17·6*)	2·6–9·4 Å² (19·2*)	4·0–7·5 Å² (17·6*)	4·3–7·0 Å². (9·1*)
R Initial	58·0	44·0	24·9	23·3	7·1
After 3 cycles	30·0	22·0	19·5	19·7	6·2

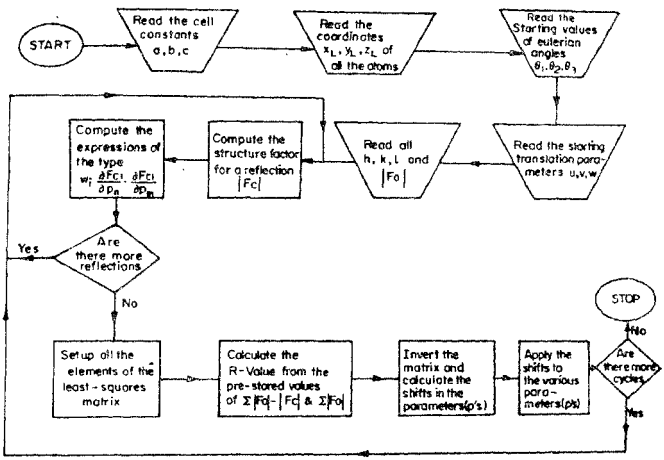

FIG. 3. Flow chart of the rigid body least squares refinement program.

Table VII

Progress of the intermolecular NH(DAla)...CO(DAla) distance r(Å) and angle θ (°) as a potential hydrogen bond during X-ray refinement of model B

Type of refinement	No. of reflections	Resolution (Å)	R-factor	r	θ
Initial	50	2·80	58·0	3·80	67·3
Rigid body	50	2·80	30·0	2·79	41·5
SFLS	200	1·68	22·0	2·77	16·9
SFLS	300	1·45	19·5	2·79	11·9
SFLS	500	1·20	19·7	2·81	7·2
SFLS	975	0·90	6·2	2·93	18·5

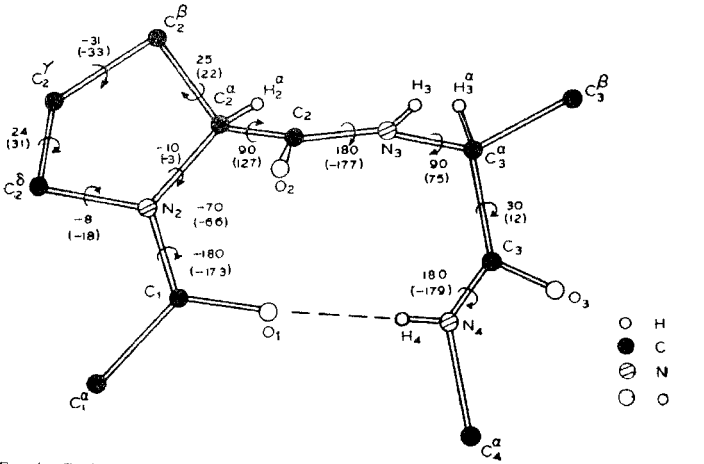

FIG. 4. Conformational angles of the molecule adopted in the contact search. The corresponding values in the final crystal structure are given within brackets.

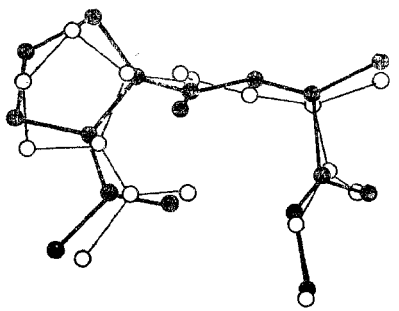

FIG. 5. Composite projection of the atoms in the tri-peptide molecule in the beginning (-●-●-) and after (-○-○-) the X-ray refinement.

In as much as the molecular shapes of both the refined and the initial trial models are grossly similar, there are some significant conformational differences as shown in Fig. 4. The closeness of the initial approximate model to the final refined model is clearly shown in Fig. 5 where a composite projection of the two models is drawn.

4. Conclusions

The method of packing analysis based on contact criteria for the crystal structure determination is applicable to molecules whose molecular shape is known before hand. The trial structure can be judiciously improved further using the method of rigid body least-squares refinement until conventional refinement of atomic coordinates can be pursued. The fact that an approximate molecular shape could be used to find the correct molecular packing position shows the utility of the packing analysis in ultimately solving the diffraction phase problem. Recently, in this laboratory, this method is being applied on other rigid molecules.

Acknowledgements

The author thanks Prof. G. N. Ramachandran for suggesting the problem and for many helpful discussions and is grateful to the Council of Scientific and Industrial Research, India, for a Senior Research Fellowship.

References

1. RAMACHANDRAN, G. N. — Molecular Forces in protein structure and crystallography, *Int. J. Protein Res.*, 1969, **1**, 5–17.
2. RAMACHANDRAN, G. N., SARATHY, K. P. AND KOLASKAR, A. S. — Interatomic potential functions and crystal packing, *Zeitschrift fur Kristallographic*, 1973, **138**, 299–312.
3. KITAIGORODSKY, A. I. — *Molecular crystals and molecules*, Wiley, New York, 1973.
4. WILLIAMS, D. E. — A method of calculating molecular crystal structures, *Acta Crsyt.* 1969, **A25**, 464–470.
5. RAMACHANDRAN, G. N., RAMAKRISHNAN, C. AND SASISEKHARAN, V. — Stereochemistry of polypeptide chain configurations, *J. Mol. Biol.* 1963, **7**, 95–99.
6. RAMACHANDRAN, G. N. AND SHAMALA, N. — Crystal structure from packing : Application to the triclinic structure of the cyclic hexapeptide cyclo $(Gly\text{-}L\text{-}Tyr\text{-}Gly)_2 \cdot 2H_2O$, *Acta Cryst.*, 1976, **A32**, 1008–1009.
7. GOLDSTEIN, H. — *Classical mechanics*, Addison Wesley, London, 1966, p. 109.
8. RAMACHANDRAN, G. N. AND SASISEKHARAN, V. — Conformation of polypeptides and proteins in *Advances in Protein Chemistry*, 1968, **23**, 283–438.
9. WILSON, A. J. C. — On the determination of absolute from relative X-ray intensities, *Nature*, 1942, **150**, 152.
10. RAMACHANDRAN, G. N., LAKSHMINARAYANAN, A. V., BALASUBRAMANIAN, R. AND TEGONI, G. — Studies on the conformation of amino acids XII. Energy calculations on prolyl residue, *Biochim. Biophys. Acta*, 1970, **221**, 165–181.
11. SHIONO, R. — S.F.L.S. Program, Private Circulation.

12. DAMIANI, A., GIGLIO, E. AND LIQUORI, A. M. Minimum residual analysis of crystal structures, *Acta Cryst.*, 1967, **23**, 681–687.

13. WILLIAMS, D. E. Crystal structure of 2,4,6-triphenyl verdazyl, *Acta Cryst.*, 1973, **B29**, 96–102.

14. PAWLEY, G. S. Constrained refinements of crystal structures and their programming, *Acta Cryst.*, 1969, **A25**, 531–535.

15. RAO, A. D. Rigid-body constraints and the least-squares refinement of crystal structures, *Acta Cryst.*, 1976, **A32**, 895–897.

The author desired that the following reply to the referee be published along with the paper for the benefit of readers who may have similar doubts. *Editor*

Referee : It would be worthwhile to know if the method would be more advantageous than the direct methods programs like MULTAN.

Author : It is difficult to make any comment one way or the other at this stage of development. One has to wait for several structure solutions from the present method for a useful comparison. However, the method could be more economical and advantageous for triclinic space groups which often is not the favourable space group for MULTAN.