CRYSTAL STRUCTURE STUDIES

ON SOME DEHYDROPEPTIDES

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF THE

REQUIREMENTS FOR THE DEGREE

of

1

3

.

-

J

5

Å

Å

1

DOCTOR OF PHILOSOPHY

of the

Council for National Academic Awards, London

by

JAMES RAFTERY

Department of Environmental Sciences

Plymouth Polytechnic

Plymouth PL4 8AA

in collaboration with

Queen Mary College, London

November 1981

PLYMOUTH POLYTECHNIC LIBRARY ----5 500059 Accn. No. RAF Class. 548.8 No X700154565 C ÷

Acknowledgements

I am grateful to my Supervisors, Drs. S. A. A. Jayaweera, J. Braven and M. B. Hursthouse for their help and advice during the course of this work.

The use of the CAD4 diffractomer, the computing facilities, and the help of the members of the Crystallography Department, at Queen Mary College, is also much appreciated.

Finally, I wish to express my thanks to Devon Local Authority for the award of a Research Assistantship and Mrs. M. Gray for the competent typing of this Thesis.

i

CONTENTS

		Page
Acknowledgements	5	i
Abstract		iv
Abbreviations		v
CHAPTER 1:	A Summary of the Theory of X-ray	
	Structure Determination.	1
1.1	Diffraction Theory	2
1.2	Data Reduction	8
1.3	Direct Methods	16
1.4	Least Squares Refinement	32
CHAPTER 2:	Dehydro-peptides: Chemistry and Occurence	40
2.1	Introduction	41
. 2.2	Preparation	44
2.3	Biosynthesis and Biochemistry	53
CHAPTER 3:	Experimental Details	60
3.1	Preparations	61
3.2	Data Collection Computing and	
	Miscellaneous Details	66
CHAPTER 4:	N-Acetyl Dehydro-Alanine	70
4.1	Introduction	71
4.2	N-Acetyl Dehydro-alanine	73
CHAPTER 5:	N-Acetyl-L-Ala-L-Ala-OEt and	
	N-Acetyl-L-Ala-L-Ser-OEt	88
5.1	Introduction	89
5.2	N-Acetyl-L-Ala-L-Ala-OEt	91
5.3	N-Acetyl-L-Ala-LSer-OEt	110
5 <i>1</i>	Structure Solution Strategies	130

Pag	e
	-

CHAPTER 6:		N-Acetyl Dehydro-phenylanine and	
		N-Acetyl Dehydro-phenylalanyl-L-Proline	137
	6.1	Introduction	138
	6.2	N-Acetyl Dehydro-phenylalanine	140
	6.3	N-Acetyl Dehydro-phenylalanyl-L-Proline	160
APPENDICES			
	Α.	Probability Aspects of Structure	
		Solution	vii
	в	Peptide Chemistry	xiii
		B.l Synthesis	xiv
		B.2 Racaemisation	xx
	c.	Other Structures Attempted	xxvi
	D.	Recent Developments	xxx
	E.	Structure Factor Tables	xxxvi
Table	e E.1	N-Acetyl Dehydro-Alanine	xxxvii
Table	e E.2	N-Acetyl-L-Ala-L-Ser-OEt	. xlvii
Table	e E.3	N-Acetyl Dehydro-Phenylalanyl-L-Proline	Ļvi
Table	E.4	N-Acetyl-L-Ala-L-Ala-OEt	lxviii
mable.		N-Acetyl Dehydro-phenylalanine	exxvi

Abstract

Crystal Structure Studies on Some Dehydropeptides James Raftery

The aim of the project was the preparation and structural determination of dehydro-peptides.

To provide a benchmark for the dehydro-peptide structures two N-acetyl dipeptide structures were studied, N-Ac-L-Ala-L-Ser-OEt (Space group C2, a = 15.283 Å, b = 9.822 Å, c = 9.710 Å, β = 112.94^O) which was solved in a straight forward manner and N-Ac-L-Ala-L-Ala-OEt (Space group C2, a = 14.6557 Å, b = 9.9312 Å, c = 9.631 Å, β = 112.501^O) which was eventually solved using a starting set selected by a technique based on negative quartets.

Then N-acetyl dehydro-amino acids were investigated and the structures of N-Ac- Δ Ala-OH (Space group P2₁/n, a = 3.9414 Å, b = 10.1558 Å, c = 14.8975 Å, β = 91.64°) and N-Ac- Δ Phe-OH (Space group P2₁/a, a = 18.2815, b = 6.0807, c = 11.4019, β = 105.94°) resolved.

Next N-acetyl dehydro-peptides were considered. Various methods of preparation were attempted:

<u>Carbodiimide coupling</u>. This was in the main unsuccessful, though interestingly when the coupling of N-Ac- Δ Phe-OH with an amino acid ester was attempted using D.C.C.I./HOBT the unsaturated oxazolone was formed in good yield.

 β elimination. The tosyl derivatives of dipeptides containing serine were prepared. The products were non-crystalline.

Bergmann synthesis. A wide range of N-acetyl dehydro-peptides were synthesised using this technique. Although in the main crystalline, the crystals were usually unsuitable for structural determination. However, one compound, N-Ac- Δ Phe-L-Pro-OH (Space group C2, a = 18.516 Å, b = 9.515 Å, c = 10.538 Å, β = 120.102°) was found to form suitable crystals.

ABBREVIATIONS

. 1	
Bessel function of argument Z and order m	- I _m (Z)
Σf ⁿ j	- σ _n
$2\sigma_3\sigma_2^{-3/2} \mathbf{E}_{\mathbf{h}}\mathbf{E}_{\mathbf{h}-\mathbf{h}},\mathbf{E}_{\mathbf{h}} $	-κ _{hh} '
$2\sigma_{4}\sigma_{2}^{-2} E_{h}E_{k}E_{\ell}E_{-h-k-\ell} $	- B
Probability distribution of Ω with fixed A	- Ρ(Ω A)
Expection value of Ω with fixed A	- E(Ω A)
Average value of F	- <f`> h h</f`>
Ungerade	– u
Gerade	- g
A determinant	- ∆ *
$ \mathbf{F}_0 - \mathbf{F}_c $	- ∆ *
The torsion angle $\Theta(C_{i-1}, N_i, C_i^{\alpha}, C_i)^2$	- ¢ *
The torsion angle $\Theta(N_i, C_i^{\alpha}, \dot{C}_i, N_{i+1})^2$	- Ψ.* i
The torsion angle $\Theta(C_{i},C_{i},N_{i+1},C_{i+1})^{2}$	- ω _i *
Di-isopropyl ethylamine	- D.I.E.A
Dicyclohexylcarbodimide	- D.C.C.I
Dicyclohexylamine	- D.C.H.A
Dicyclohexylurea	- D.C.U.
Diazabicyclo [5.4.0] undec-5-ene	- D.B.U.
Phenyl	´- Φ*
Acetic anhydride	- Ac ₂ O
Toluene sulphonyl(tosyl)	- Ts
Benzyloxycarbonyl	- Z

v

ABBREVIATIONS continued

l-hydroxybenzotriazole	– HOBT
Tertiary butyl hypochlorite	- tert-BHC

Standard contractions are used for amino acids. The dehydro- (Δ^*) prefix signifies an α,β double bond.

"If an abbreviation has more than one meaning, it will be used only where the context makes the intended meaning clear.

CHAPTER ONE

A SUMMARY OF THE THEORY OF

X-RAY STRUCTURE DETERMINATION

1.1 Diffraction Theory

It can be shown from diffraction theory³ that the amplitude of X-ray scattering, in a direction represented by hkl, from a three dimensional regularly repeating motif is

1.1.1
$$F_{hkl} = V \iiint \rho(x,y,z) e^{2\pi i \underline{r} \cdot \underline{r}} dx dy dz$$

 $\rho(x,y,z)$ - electron density at (x,y,z)
 $\underline{r} = x\underline{a} + y\underline{b} + z\underline{c}$
 $\underline{a}, \underline{b}, \underline{c}$ - the vectors defining the unit cell
 $\underline{r}^* = h\underline{a}^* + k\underline{b}^* + l\underline{c}^*$

 $\underline{a}^* = \underline{b} \times \underline{c} / \underline{a} \cdot \underline{b} \times \underline{c}$ with similar definitions for $\underline{b}^* \underline{c}^*$, the three vectors defining the reciprocal lattice (r.l.) cell. It is computationally more convenient to write 1.1.1 as

1.1.2
$$F_{hkl} = \sum_{j}^{N} f_{j} e^{2\pi i (hx_{j}+ky_{j}+lz_{j})}$$

 (x_{j}, y_{j}, z_{j}) are the coordinates of the jth atoms centre

 f_j is an atomic scattering factor, representing the scattering ability of the jth atom, being, because of finite atomic size, only the same as the atomic number Z when $\sin\theta/\lambda = 0$. As 1.1.1 is the Fourier transformation of $\rho(x,y,z)$, the desired electron density function

1.1.3
$$\rho(x,y,z) = 1/V \sum_{h \neq l} \sum_{k \neq l} F_{hkl} e^{-2\pi i (hx+ky+lz)}$$

A comparison of 1.1.3 with a three dimensional Fourier series⁴ reveals it to be just a Fourier series with Fourier coefficients F_{hkl}/V . Representing F_{hkl} on an Argand diagram reveals the difficulty in calculating 1.1.3 (Figure 1.1.1)

 $F_{hkl} = A + iB$ (Generally, but for centrosymmetric

structures B=0)

 $\mathbf{F}_{hkl} = |\mathbf{F}_{hkl}| e^{2\pi i \alpha} hk^{l}$ with $\alpha_{hkl} = Tan^{-1} B/A$



Figure 1.1.1. An Argand diagram showing the structure factor F_{hkl} and its relationship to the atomic scattering factor f_i .

The parameter actually measured in X-ray diffraction is effectively $|F_{hk\ell}|$ so rewriting in terms of $|F_{hk\ell}|$ and $\alpha_{hk\ell}$, the phase angle,

1.1.4 $\rho(x,y,z) = 1/V \Sigma \Sigma \Sigma |F_{hkl}|e^{-2\pi i (hx+ky+lz-\alpha_{hkl})}$ a form that lays bare the central problem in structure determination, the phase problem.

The exponent of 1.1.4 has the form of a plane, so that each term is a function having constant electron density on the plane (hkl), the electron density varying sinusoidally along the normal to the plane, with a maximum magnitude $|F_{hkl}|/V$ (electrons per unit volume) at a distance α_{hkl} from the origin.

As α_{hkl} is experimentally unknown so is the position of the maximum, though in the case of centrosymmetric structures with α as 0 or π , the problem reduces to whether the maximum or minimum is at the origin. The Reciprocal Lattice

A very convenient way of visualising the diffraction conditions is by means of the Ewald sphere (Figure 1.1.2).

The sphere is of unit radius and diffraction only occurs when a point (of the reciprocal lattice) is on the surface of the sphere.

 $d^* = 2 \sin\theta$

From comparison with the Bragg equation, $n\lambda = 2d \sin \theta$ one sees that

$$n\lambda/d_{hkl} = d*_{hkl}$$

'n' being the diffraction order so

 $d_{hkl}^{*} = d_{nh nk nl}^{*}/n$

To clarify the relationship between the real and reciprocal lattice (r.l.) consider the set of real planes (hkl), with the usual meaning that the plane intercepts <u>a</u>, <u>b</u>, <u>c</u> at (l/h, l/k, l/l) (h,k,l being relatively prime). All the planes so defined (real/rational planes) contain lattice points, and conversely, no lattice points are



Figure 1.1.2. Reciprocal space and its relationship to diffraction

in real space.

located away from such a plane.

Consider the reciprocal lattice vector

$$\underline{\underline{r}}_{\alpha\beta\gamma}^{*} = \underline{\alpha}\underline{\underline{a}}^{*} + \beta \underline{\underline{b}}^{*} + \gamma \underline{\underline{c}}^{*}$$

If this vector is perpendicular to the vectors (-1/h,0,1/l) and (-1/h,1/k,0), then the scaler products are zero

$$((1/-h)\underline{a} + (1/k)\underline{b}) \cdot (\underline{\alpha}\underline{a}^{*} + \beta\underline{b}^{*} + \gamma\underline{c}^{*}) = 0$$
$$((-1/h)\underline{a} + (1/k)\underline{c}) \cdot (\underline{\alpha}\underline{a}^{*} + \beta\underline{b}^{*} + \gamma\underline{c}^{*}) = 0$$
$$\underline{\alpha}/h = \beta/k = \gamma/k$$

An obvious solution is $\alpha = nh$, $\beta = nk$, $\gamma = nl$, n integer, so that associated with the plane (hkl) are the points $\frac{r}{-nh}$ nk nl, with the values of n corresponding to the orders of reflexion from (hkl).

Examining the Bragg equation it is obvious that the maximum value of $n < 2d_{hkl}/\lambda$.

In terms of the Ewald sphere the only r.l points that can be observed lie within 2 r.l.units of the origin, the so called limiting sphere.

This introduces one of the problems of X-ray crystallography, namely series termination effects.

As the r.l. points correspond to the coefficients of a threedimensional Fourier series, a series formed from diffraction data is truncated. This is equivalent to multiplying the reciprocal lattice by a three-dimensional top hat function; by the convolution theorem this results in $\rho(x,y,z)$ being multiplied by

 $(\sin(4\pi s) - 2\pi s \cos 4\pi s)/2\pi^2 s^3$

$$s = (x^2 + y^2 + z^2)^{\frac{1}{2}}$$

producing so called diffraction ripples, peaks and troughs around atoms, which make it difficult to pick out from an electron density map, light atoms in the proximity of a heavy atom. However, if, using

known atomic positions to calculate F_C , a Fourier series is computed with coefficients $|F_O| - |F_C|$, a difference synthesis is produced

$$\rho(\mathbf{x},\mathbf{y},\mathbf{z}) = \Sigma\Sigma\Sigma(|\mathbf{F}_0| - |\mathbf{F}_C|)e^{-2\pi i(\mathbf{h}\mathbf{x}+\mathbf{k}\mathbf{y}+\mathbf{l}\mathbf{z}-\mathbf{\alpha}\mathbf{h}\mathbf{k}\mathbf{l})}$$

Since Fourier series calculated using both F_{C} and F_{O} as coefficients produce diffraction ripples, the difference will almost eliminate the ripples, indeed all common features, leaving, hopefully, the light atoms revealed.

In practice, though refinement using only high order reflexions has been found to give accurate bond lengths⁵, the light atoms are often located geometrically or increasingly, from neutron diffraction data, thereby avoiding problems with the hydrogen scattering factor⁶. Before the structure factors can be used in calculations they must first be found from the measured intensities, which, in the case of a diffractometer, are proportional to the total count (corrected with background readings):

 $K = \frac{A_{hkl}}{hkl} \frac{L_{hkl}}{hkl} \frac{p_{hkl}}{F_{hkl}} |F_{hkl}|^2 V = \frac{E_{hkl}}{\omega} / I_{o} = \text{Integrated reflexion}$ K = constant

e = primary extinction correction

A = absorption and secondary extinction correction

L_{hkl} = Lorentz factor

p_{hkl} = polarisation factor

I = beam intensity

The integrated reflexion is used as the reciprocal lattice points have finite volume, not only due to finite crystal size, but also thermal effects and the non-monochromatic nature of the X-ray beam.

If the Bragg equation is differentiated

$$\lambda = 2d_{hkl} \sin\theta$$
$$d\lambda = 2d_{hkl} \cos\theta d\theta$$
$$d\theta = d \lambda/2d \cos\theta$$

one sees, as well as a lengthening in the direction of the origin, that there exists an exceptionally large range d0 over which the Bragg equation is satisfied at $0 = 90^{\circ}$; furthermore, the incident beam is never ideally collimated and so the incident beam impinges on the crystal over a range of incident angles.

Earlier it had been assumed that the nuclear positions were constant. In fact, thermal agitation causes each atom to execute vibrations about the equilibrium position. If 'u' is taken to indicate the displacement of the atom, then the magnitude of the thermal motion can

be expressed by $\langle u^2 \rangle$ the mean square displacement.

Generally the field of restoring forces and hence the mean square displacement varies with direction and can be shown to mark out a quadric surface; regarding the restoring forces as linear, this surface is an ellipsoid. Unless restricted by the symmetry of the atoms location the mutually perpendicular axes of the ellipsoid are not parallel with the axes of the unit cell.

If the atoms of a crystal were motionless, then each atom of a set of translation equivalent atoms would be fixed relative to one another. Under such circumstances every atom on a particular plane scatters exactly in-phase with all other atoms on that plane and exactly 2π out of phase with the atoms on the next plane. As a result all atoms of the translation equivalent set scatter in phase with an amplitude corresponding to f_i .

With thermal motion the atoms are somewhat displaced from their planes and this spoils the in-phase relations of their combined scattering; consequently the amplitude is reduced by a factor of the form e^{-W} , the temperature factor.

The temperature factor for isotropic temperature vibrations is

$$T_{hk\ell} = e^{-2\pi^2 \langle u^2 \rangle / d_{hk\ell}^2}$$

= $e^{-2\pi^2 \langle u^2 \rangle (2\sin\theta/\lambda)^2}$
= $e^{-B(\sin\theta/\lambda)^2}$
= $e^{-2\pi^2 \langle u^2 \rangle r^* hk\ell^2}$
= $e^{-2\pi^2 \langle u^2 \rangle r^* hk\ell^2}$
 $r^*_{hk\ell} = 1/d$

The most general form of the temperature factor is then 1.2.1 $T_{hkl} = e^{-2\pi^2 (U_{11}h^2a^{*2}+U_{22}k^2b^{*2}+U_{33}l^2c^{*2}+2U_{12}hka^*b^*+2U_{23}klb^*c^*+2U_{13}hla^*c^*)}$ Other forms commonly used have

$$2\pi^{2}U_{11}a^{*2} = \frac{1}{4}B_{11}a^{*2} = \beta_{11}$$
$$4\pi^{2}U_{12}a^{*}b^{*} = \frac{1}{2}B_{12}a^{*}b^{*} = \beta_{12}$$

The exponent of 1.2.1 has the form of an ellipsoid referred to an arbitrary coordinate system

$$Ax^2 + By^2 + Cz^2 + 2Dxy + 2Eyz + Fzx = a$$

with

$$A = 2\pi^2 U_{11}$$

and

 $h^2a^{*2} = x^2$ etc.

which can readily be converted to the neater form

$$x^2/a^2 + y^2/b^2 + z^2/c^2 = 1$$

where a, b and c are the three orthogonal principal axes of the ellipsoid; the directions of these axes are, in general, oblique to each of the crystal axes, so a complete description includes the angles between the ellipsoid axes and the crystal axes.

After the temperature factor correction 1.1.2 becomes

$$F_{hkl} = \Sigma g_i e^{2\pi i (hx_i + ky_i + lz_i)}$$

$$g_i = T_{hkl}f_i$$

Lorentz Factor

As the r.l points have volume they spend a finite time on the Ewald sphere, a time proportional to the velocity normal to the sphere. Considering the zero level as a simple case (Figure 1.2.1)

$$OP = 2\sin\theta$$
$$v_n = v\cos\theta$$
$$= OP.\omega\cos\theta$$
$$= 2\sin\theta\cos\theta\omega$$
$$\omega/v_n = 1/\sin^2\theta = L_{hkl}$$

For other geometries the Lorentz factor will take different forms, the



Figure 1.2.1. A diagrammatic rationalisation of the Lorentz correction.

various expressions being found in Vol. II of International Tables. Polarisation Factor

The need for P_{hkl} , the polarisation correction, can be understood by considering the randomly polarised incident X-ray beam I_i , which notionally consists of two equal plane polarised mutually perpendicular beams. The beam in the plane presents no problems; however, the beam whose plane of polarisation is perpendicular to the reflecting plane will be reduced in intensity. The scattered beam can be thought of as being produced by electrons moving up and down the electric vector of the incident beam. The component of the electron motion in the direction of the electric vector of the reflected beam I_s is proportional to cos 20. Hence:

$$I_{i} = A_{x}^{2} + A_{y}^{2}$$

$$A_{x} = A_{y} = (I_{i}/2)^{\frac{1}{2}}$$

$$I_{s} = K I_{i}(\frac{1}{2} + \frac{1}{2}\cos^{2}2\theta)$$

$$I_{s} \propto (1 + \cos^{2}\theta)/2 = p_{hkl}$$

If the beam has been monochromated by reflexion from a crystal plane, the p_{hkl} has to be modified to allow for the partial polarisation of the incident beam.

This arises from the scattered beam being $\pi/2$ out of phase with the incident beam and scattering from a parallel plane, resulting in interference effects.

It is difficult to allow for, if only because the calculation of $tanh (nq)/nq(e_{hkl})$, requires the determination of 'n', the average number of perfectly parallel planes within the crystallites, which is

difficult. However, as $\lim tanh(x)/x = 1$ $x \neq 0$

if 'n' can be made small, by for example breaking up the crystallites with a sudden temperature drop, then e_{hkl} will approach unity; generally, however, the correction is small and ignored.

Secondary Extinction

This results from the diminution in energy of an incident beam due to scattering; it means that the incident beam is less intense on internal planes than on superficial planes; it behaves rather like absorption and the correction is of a similar form

$$I_o = I_e^{-2gI}c$$
 $I_c = corrected intensity$

The 'g' can be found by expanding the expression and neglecting higher terms

$$I_{c}/I_{o} = 1 + 2gI_{c}$$
.

A plot of I_c/I_c v I_c will give g, assumed to be a constant for a particular crystal, as the gradient.

In practice, because of temperature factor and atomic scattering factor dependence upon $\sin\theta/\lambda$, secondary extinction only effects significantly $\sin\theta/\lambda < 0.2$ reflexions, which can be detected by looking for $I_c > I_o$ by an amount greater than the residual index, R. <u>Absorption</u>

The intensity, I_T of a beam passing through a thickness 'x' is given by

$$I_{T} = I_{i}e^{-\mu_{\lambda}x}$$

 μ_{λ} = the linear absorption coefficient.

Knowing the size and shape of a crystal, preferably spherical or cylindrical, and with μ_{λ} being a computable quantity, the attenuation

can be calculated for a given reflexion.

To obviate the need for correction, at least for organic structures, highly penetrating radiations such as Mo, may be used, especially in diffractometry. Where intensities are recorded on film this is less true because of inefficiency.

Point Atom Structures

It has been shown previously that F_{hkl} decreases with $\sin\theta/\lambda$; this introduces serious restrictions when inequalities and probabilities are used, where the magnitudes of the reflexions are of paramount importance.

In an attempt to circumvent this limitation F_{hkl} is transformed into the form it would have had if the structure consisted of point, stationary atoms.

$$F_{hkl} = \Sigma g_{j} e^{2\pi i (hx_{j}+ky_{j}+lz_{j})}$$
$$g_{i} = T_{hkl} f_{j}$$

For one atomic type F = gE

E being the sum of the exponential terms. With point atoms g = Z, Z being the atomic number.

$$F_{point}/F_{real} = Z E/E f T_{hkl}$$
$$F_{point} = Z/f T_{hkl}F_{real}$$

As however, generally, crystals contain more than one type of atom, some average value of Z/g, such as $\Sigma Z_i / \Sigma g_i$ must be used.

Noting that $\Sigma Z_{j} = F_{j}$, an early point atom function was the unitary structure factor

$$U_{hkl} = (F_{hkl})_{point}/F_{ooo} = (F_{hkl})_{real}/\Sigma_{gj}$$

However, for an accurate comparison of intensities allowance must be made for the enhancement in intensity for certain classes of re-

flexion⁷.

For this, and other reasons 'E', a normalised structure factor was introduced⁸.

$$E^{2}_{hkl} = U^{2}_{hkl} / \overline{U}^{2}$$

 \overline{U}^2 is the average over all reflexions, excepting systematically absent reflexions, for a particular set of reflexions.

Equivalently

1.2.2 $E_{hkl}^2 = |F_{hkl}|^2 / \epsilon \Sigma g_j^2$

with ε compensating for the enhancement mentioned earlier.

Note that the denominator in 1.2.2 is the local average intensity:

$$F_{hkl}|^{2} = F_{hkl}F_{hkl}^{*}$$
$$= \Sigma g_{j}^{2} + 2 \Sigma \Sigma g_{k}g_{j} \cos(\underline{h} \cdot (\underline{r}_{j} - \underline{r}_{k}))$$

If reflexions within a shell centred on the origin, so $\sin\theta/\lambda$ is constant, are averaged, then the second term on the right hand side vanishes as it is as likely to be plus as minus.

The temperature factors are initially either specified from experience or estimated from a Wilson plot⁹.

Wilson Plot^{10,11}

As shown above, at constant $\sin\theta/\lambda$

$$|\mathbf{F}_{hkl}|^2 = \Sigma g^2_{i}$$

If equal isotropic temperature factors are assumed this equals

 $e^{-2B(\sin\theta/\lambda)}\Sigma f^2$

As measured intensities are naturally on a relative scale

$$I_{rel} = K I_{abs} = K e^{-2B(\sin\theta/\lambda)^2} \Sigma f_{i}^2$$
$$I_{rel}/\Sigma f_{i}^2 = K e^{-2B(\sin\theta/\lambda)^2}$$
$$\ln(I_{rel}/\Sigma f_{i}^2) = \ln K - 2B(\sin\theta/\lambda)^2$$

Plotting $\ln(I_{rel}/\Sigma f_i^2)$ against $(\sin\theta/\lambda)^2$ gives -2B as the slope and ln K when $(\sin\theta/\lambda)^2 = 0$.

1.3 Direct Methods

This is the generic term given to methods which use intensities directly to solve the structure.

The growth in the importance of direct methods can be dated from 1948^{12} , when classical inequalities were used to derive relationships between reflexions, although earlier work had been done¹³.

For example, with centrosymmetric structures, the unitary structure factor U_{hkl} takes the form:

$$2\Sigma n_j \cos 2\pi (hx_j + ky_j + \ell z_j) n_j = g_j / \Sigma g_j$$

Adopting the notation $U_h = U_{hkl} = 2\Sigma n_j \cos(h)$

$$U_{k} = U_{h'k'\ell'} = 2\Sigma n_{j} \cos(k)$$

 $U_{h} + U_{k} = 2 \sum_{j} (\cos(h) + \cos(k)) = 4 \sum_{j} \cos((h + k)/2) \cos((h - k)/2)$ Using the Cauchy inequality $(\sum_{i} b_{i})^{2} < (\sum_{i} a_{i}^{2}) (\sum_{i} b_{i}^{2})$

$$(U_{h} + U_{k})^{2} < 16 \Sigma n_{j} \cos^{2}((h + k)/2) \Sigma n_{j} \cos^{2}((h - k)/2)$$

 $< 16 \Sigma n_{j} \frac{1}{2} (1 + \cos(h + k)) \Sigma n_{j} \frac{1}{2} (1 + \cos(h - k))$
 $< (1 + U_{h+k}) (1 + U_{h-k})$

Similarly

$$(u_{h} - u_{k})^{2} < (1 - u_{h+k}) (1 - u_{h-k})$$

These relationships have recently found use in symmorphic space 14 15 16 groups ' ' based on the fact that if

$$|\mathbf{U}_{h}|, |\mathbf{U}_{h+k}|, |\mathbf{U}_{h+k}|$$
 are large, $|\mathbf{U}_{k}| = 0$

and

$$|u_{h}|^{2} > (1 - |u_{h+k}|) (1 - |u_{h-k}|)$$

then

$$S(h + k)S(h - k) = -1.$$

More general inequalities, based on $\rho(x,y,z)$ being non-negative

(which the Harker-Kasper inequalities implicitly assumed) are the Karle-Hauptmann Determinants¹⁷.

1.3.1
$$\begin{bmatrix} E_{000} & E_{-h_1} & E_{-h_2} & \cdots & E_{-h_n} \\ E_{h_1} & E_{000} & E_{h_1-h_2} & \cdots & E_{h_1-h_n} \\ E_{h_2} & E_{h_2-h_1} & E_{000} & \cdots & E_{h_2-h_n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ E_{h_n} & E_{h_n-h_1} & E_{h_n-h_2} & \cdots & E_{000} \end{bmatrix} > 0$$

However, the use of inequalities has been restricted mainly to relatively small molecules¹⁸.

In more recent years, probability methods have predominated, with the normalised structure factor, E, being used almost universally. If equal temperature factors and profiles for f_i are assumed then

$$f_{j}/\sigma_{2}^{l_{2}} = \phi_{j} = Z_{j}/(\Sigma Z_{j}^{2})^{l_{2}} = N^{-l_{2}} \text{ for an equal atom structure}$$
$$E_{hkl} = \Sigma \phi_{j} e^{2\pi i (hxj+kyj+lzj)}$$

Consider the product E(h')E(h-h') for an equal atom structure

$$E(h')E(h-h') = N^{-1} \Sigma e^{2\pi i \underline{h} \cdot \underline{r} \underline{j}} + N^{-1} \Sigma \Sigma e^{2\pi i (\underline{h}' \cdot \underline{r} \underline{j} + (h-h') \cdot \underline{r} \underline{k})}$$

Averaging over h', the double summation, assuming random atomic positions, is zero.

1.3.2
$$E(h) = N^{\frac{1}{2}} < E(h^{\prime}) E(h-h^{\prime}) >_{h^{\prime}}$$

It can be shown that for an unequal atom structure 1.3.2 takes the form

1.3.3 E(h) = $\sigma_3^{-1}\sigma_2^{3/2} < E(h')E(h-h')_{h'}$

Superficially 1.3.3 appears useless since it is necessary to know the magnitude and phase of all 'E's to determine one. It was pointed out^{19} however, that where E_{hkl} is large, the series must lean strongly

in one direction and this direction will tend to be dominated by products between large E's.

For centrosymmetric structures 1.3.3 is adequate, but for noncentrosymmetric structures, where the values of the phases must be determined, rather than the sign, 1.3.3 must be modified:

1.3.4 $E_h = |E_h| \cos\phi_h + i|E_h| \sin\phi_h$

Equating real and imaginary components

$$\begin{aligned} |\mathbf{E}_{h}| \cos\phi_{h} &= \sigma_{3}^{-1}\sigma_{2}^{3/2} < |\mathbf{E}_{h}| |\mathbf{E}_{h-h}| \cos(\phi_{h} + \phi_{h-h}) >_{h} = C \\ |\mathbf{E}_{h}| \sin\phi_{h} &= \sigma_{3}^{-1}\sigma_{2}^{3/2} < |\mathbf{E}_{h}| |\mathbf{E}_{h-h}| \sin(\phi_{h} + \phi_{h-h}) >_{h} = S \end{aligned}$$

$$1.3.5 \quad \operatorname{Tan} \phi_{h} = S/C \end{aligned}$$

The tangent formula, 1.3.5 is only an equality of the summation ranges over all h'; if there is only one term 1.3.5 reduces to

1.3.6 $\phi_h \approx \phi_{h'} + \phi_{h-h'}$

1.3.5 may be improved by the use of weights²⁰.

$$B_{h} = \langle W_{h}, W_{h-h}, |E_{h}, ||E_{h-h}, |\cos(\phi_{h}, + \phi_{h-h})\rangle_{h},$$
$$T_{h} = \langle W_{h}, W_{h-h}, |E_{h}, ||E_{h-h}, |\sin(\phi_{h}, + \phi_{h-h})\rangle_{h},$$

1.3.7 Tan $\phi_h = T_h/B_h$

 $W_h = 0.2 \alpha_h \qquad 5 > \alpha_h$

but these weights have been criticised²¹.

1.3.7 is up to four times faster than 1.3.5, the weights being based on a commonly used measure of variance of the indicated phase, 22 \tilde{h}

∝_h > 5

1.3.8
$$\alpha_h = 2\sigma_3\sigma_2^{-3/2} |E_h| (T_h^2 + B_h^2)^{\frac{1}{2}}$$

If there is only one contributor, 1.3.8 becomes

1.3.9
$$\alpha_{h} = \kappa_{hh'} = 2\sigma_{3}\sigma_{2}^{-3/2} |E_{h}E_{h'}E_{h-h'}|$$

Table 1.3.1

Statistical Properties of E

	Centric	Acentric ²³
P(E)	$(2\pi)^{-\frac{1}{2}} e^{- \mathbf{E} ^2/2}$	E e ^{- E 2}
< E >	$(2/\pi)^{\frac{1}{2}} = .798$	$\pi^{\frac{1}{2}}/2 = .886$
<e></e>	0	0
< E ² >	1	1
< 2 ² -1>	· • •	0
<(E ² -1)>	$(2\pi e)^{\frac{1}{2}} = .968$	2/e = .735
<(E ² -1) ² >	2	1
<(E ² -1) ³ >	8	2
% E >1	31.7	36.8
% E >2	4.6	1.8
% E >3	. 3	.1

The question as to whether a crystal is centrosymmetric or not can often be resolved by the statistical properties shown in Table 1.3.1 (though non-crystallographic symmetry can cause confusion^{24,25,26}). Other information, such as the presence of unusually high Patterson overlap can also be gleaned (e.g. $from < (|E_h^2|-1)^2 > or$ $< (|E_h^2|-1)^3 >)^{27}$.

The Starting Set of Phases

Formulae such as 1.3.6 are of little use however, unless some phases are known; happily some phases, generally three, may be allocated specific values, to fix the origin, and, if appropriate for the space group, an addition phase to fix the enantiomorph. Before continuing a few definitions may be helpful:

<u>A structure invariant</u> is a reflexion, or group of reflexions, whose phase, or sum of whose phases, is determined solely by the crystal structure, i.e. is independent of origin.

Equivalent points are points which are (geometrically) related to all symmetry elements in the same way.

<u>A permissable origin</u> is one related in such a way to the symmetry elements, that full advantage is taken of space group symmetry.

<u>A seminvariant</u> is a reflexion, or group of reflexions, whose phase, or the sum of whose phases, is independent of the choice of origin within the equivalence class (the equivalence class being implicitly selected by the functional form of the structure factor); obviously structure invariants are structure seminvariants.

The intensity invariants are a subset of the structure invariants whose phases, as a consequence of space group symmetry, are either O or π .

In International Tables, the forms of the structure invariants and seminvariants are indicated by ω_{I} and ω_{s} , the invariant, and seminvariant, moduli respectively, with a reflexion 'h' being an invariant or seminvariant according as 'h' is divisible by ω :

> $\omega_{j} = 0$ implies $h_{j} = 0$ $\omega_{j} \neq 0$ implies h_{j}/ω_{j} is integer

To clarify matters consider P2:

(202) = $\omega_{I} = \omega_{S}$ (as there is only one equivalence class) A = 2cos 2 π (hx + lz)cos 2 π (ky)

 $B = 2\cos 2\pi (hx + lz) \sin 2\pi (ky)$

Moving from (0,y,0) as origin, to for example $(\frac{1}{2}, y, \frac{1}{2})$ requires, if A and B are to be unchanged that h, l = g(gerade) and k = 0. Origin Specification²⁸

To select the origin of a unit cell, the phases of, generally, three reflexions are specified

1.3.10
$$\sum_{i=1}^{3} h_{i} X + k_{i} Y + \ell_{i} Z = K_{i} + n_{i} (0 < K_{i} < 1)$$

1.3.10, the planes of equal phase for the three reflexions, solve to give:

L.3.11
$$X = \Delta^{-1}$$
 $\begin{vmatrix} k_1 & \ell_1 & K_1 + n_1 \\ k_2 & \ell_2 & K_2 + n_2 \\ k_3 & \ell_3 & K_3 + n_3 \end{vmatrix}$
= Δ^{-1} $\begin{vmatrix} k_1 & \ell_1 & K_1 \\ k_2 & \ell_2 & K_2 \\ k_3 & \ell_3 & K_3 \end{vmatrix}$ + $\begin{vmatrix} k_1 & \ell_1 & n_1 \\ k_2 & \ell_2 & n_2 \\ k_3 & \ell_3 & K_3 \end{vmatrix}$

$$\Delta = \begin{vmatrix} h_1 & k_1 & \ell_1 \\ h_2 & k_2 & \ell_2 \\ h_3 & k_3 & \ell_3 \end{vmatrix}$$

Y and Z are of a similar form.

Since the same origin is required for each unit cell, $\Delta = + 1$ (symmetry allow less restrictive conditions). An equivalent result can be obtained by considering the effect of a shift of origin. The phase ψ ' of E, with respect to the new origin, related to the old by r_0 is given by:

$$E = \Sigma \phi_{i} e^{2\pi i \underline{h} \cdot \underline{r}_{i}}$$
$$E' = \Sigma \phi_{i} e^{2\pi i \underline{h} \cdot (\underline{r}_{i} - \underline{r}_{o})}$$

1.3.12 $\psi' + 2\pi m = \psi - 2\pi h.r_0$

If $\underline{h} = \sum_{i=1}^{n} \underline{h}_{i}$ where \underline{a}_{i} are rational and \underline{h}_{i} are independent vectors then 1.3.12 gives:

1.3.13
$$2\pi m + \psi' = \psi - 2\pi (\Sigma (a_{i-i} \cdot r_{o})).$$

= $\psi + \Sigma a_{i} (\psi'_{i} - \psi_{i} + 2\pi m_{i})$

It is clear that unless a, are integer, i.e. as

1.3.14
$$\Delta^{-1} = a_1 / \begin{vmatrix} h_2 & h_3 & h \\ k_2 & k_3 & k \\ k_2 & k_3 & k \end{vmatrix} = -a_2 / \begin{vmatrix} h_1 & h_3 & h \\ h_1 & h_3 & h \end{vmatrix} = a_3 / \begin{vmatrix} h_1 & h_2 & h \\ h_1 & h_2 & h \\ k_1 & k_3 & k \end{vmatrix}$$

that the base reflexions form a primitive set $(\Delta = + 1)$, ψ has values differing by $2\pi/D$ ($\Delta = GD$, G is the greatest common divisor (g.c.d.) of the determinants).

To define the ambiguous origin more closely ψ may be given one of its D values and adjoined to the basic set; the g.c.d., G, of the non-

zero 3 x 3 subdeterminants of the matrix

$$\begin{array}{cccc} h_1 & k_1 & \ell_1 \\ h_2 & k_2 & \ell_2 \\ h_3 & k_3 & \ell_3 \\ h_4 & k_4 & \ell_4 \end{array}$$

not all of whose 3×3 subdeterminants vanish, will give the number of permitted origins. If G > 1 the process is continued until one finally attains a primitive set (a matrix the g.c.d. of whose 3×3 subdeterminants is unity).

 $\begin{array}{cccc} h_1 & k_1 & \ell_1 \\ h_2 & k_2 & \ell_2 \\ h_3 & k_3 & \ell_3 \\ \cdot & \cdot & \cdot \\ h_n & k_n & \ell_n \end{array}$

As an example consider P2 again. A convenient origin specifying set is $h_g \circ l_u$, $h_u \circ l_g$, $h_g k l_g (k \neq 0)$. The first two reflexions have two possible values (o or π) either of which may be chosen; the third reflexion may be assigned an arbitrary phase. It can readily be shown that for the triple to be primitive $k_3 = +1$ and $h_1 l_2 - h_2 l_1 = +1$.

While an origin specifying set, as above, can often be put in a simple general form by making indices zero, this has the disadvantage of reducing the number of distinct symmetry related reflexions. For example, the number of known phases is reduced in the P2 case from twelve, with general reflexions, to eight.

Specification of Enantiomorph.

The question only arises, of course, in non-centrosymmetric space groups. If one enantiomorph has co-ordinates (x_j, y_j, z_j) j = 1, N, then the other enantiomorph has co-ordinates $(\bar{x}_j, \bar{y}_j, \bar{z}_j)$,

j = 1, N (except for Fdd2).

As sin $2\pi(hx + ky + lz)$ is an odd function, the enantiomorph is specified by restricting a structure (sem)invariant to one side of the real axis, in the complex plane.

However, in some space groups single phase (sem)invariants are also intensity invariants (phases are 0 or π) and therefore unsuitable for enantiomorph specification. Multi-phase invariants can be derived by considering a linear combination of phases with coefficients A_h . If the origin is moved by \underline{r}_0 then:

> $\phi_{new} = \phi_{old} - 2\pi (\underline{h}, \underline{r}_{o})$ $\Sigma A_{h} \phi_{h(new)} = \Sigma A_{h} \phi_{h(old)} - 2\pi (\Sigma A_{h} \underline{h}) \cdot \underline{r}_{o}$

1.3.15 $\Sigma A_{h_{-}}^{h} = 0$

and the value is independent of origin, i.e., is a structure invariant. As an example consider C2.

The equivalent positions referred to a primitive $cell^{29,30}$ are

	x		1	1	0		х	· .
i	Y	Y =		ī	0		У	
	Z	Primitive	0	0	ī		z	Centred
(x,y,z) $(x + y, x)$			к –	у,	ī)	()	X,Y,2	2)
$(\overline{x}, y, \overline{z})$ $(\overline{x} + y,$		- x -	Y,	z)	()	, , , , , ,	Z)	

The seminvariant modulus ω_s = (2 O 2). In the primitive cell, as

h' = h + kk' = h - k $\ell' = \overline{\ell}$

the seminvariant vector is (h - k, l) and ω_s is (0,2).

A convenient origin defining set (O.D.S.) is

$$h_1 + h_2 = 0 \mod(\omega_c)$$

implying, in this case, h_1 , h_2 , and l_2 , l_2 , have the same parity, then the linear combination

$$\phi_{h_2} + \phi_{h_2} + \phi_{2h02l}$$
 (2h = h₁ + h₂, 2l = l₁ + l₂)

is a structure seminvariant (generalising, $\Sigma A_h h = 0 \mod(\omega_s)$ implies $\Sigma A_h \phi_h$ is a structure seminvariant) which suggests, as $\phi_{2h02\ell}$ is a structure seminvariant, that $\phi_{h_1} + \phi_{h_2}$ is as well. The enantiomorph for C2 can be specified by a $0 < \phi_{h_1} + \phi_{h_2} < \pi$ restriction.

A systematic method has been developed of isolating a large set of reflexions which specify enantiomorph strongly, using cosine invariants; it relies upon formulae for calculating approximately, from intensity data alone $\cos(\phi_h + \phi_k + \phi_\ell)$ (h + k + $\ell = 0$), these formulae being particularly useful when the cosine, due to space group symmetry must be $\overline{+}$ 1.

Even when the cosines must be $\overline{+}$ 1, the value may be ambiguous, so checks on internal consistency are desirable.

Consider:

$$\phi_1 = \phi_{h_1} + \phi_{h_2} + \phi_{h_3}$$

$$\phi_2 = \phi_{-h_1} + \phi_{-k} + \phi_{h_1+k}$$

$$\phi_3 = \phi_{-h_3} + \phi_k + \phi_{h_3-k}$$

$$\phi_4 = \phi_{-h_1-k} + \phi_{-h_2} + \phi_{-h_2+k}$$

With $\Sigma h_i = 0$, $\Sigma \phi_i = 0$ is a structure invariant, so if, for example $\cos \phi_i = 1$ i = 1,3, then $\cos \phi_4$ must also be 1.

By using these quadrupoles, by requiring consistency between the formulae, and, especially where from space group symmetry the cosine must be $\overline{+}$ 1, using only a percentage (the most strongly indicated) of the expected number of a sign, error may be minimised.

To illustrate how these cosine invariants are used consider $P2_1^{31}$:

By space group symmetry

 $\cos\phi_{\overline{hkl}} = (-1)^k \cos\phi_{\overline{hkl}}$

Reflexions hkl, with k fixed, if $|E|_{02k0}$ is not too large, will take values distributed over 0 to 2π .

With the origin chosen so 0 2k 0 = 0

 $k = 2n \quad \cos(\phi_{h\bar{k}\ell} + \phi_{\bar{h}\bar{k}\bar{\ell}} + \phi_{02k0}) = \cos 2\phi_{h\bar{k}\ell} \approx \pm 1$ according as $\phi_{h\bar{k}\ell} \approx 0, \pi \text{ or } \pm \pi/2$ respectively $k = 2n + 1 = -\cos 2\phi_{h\bar{k}\ell} \approx \pm 1 \text{ according as}$ $\phi_{h\bar{k}\ell} \approx 0, \pi \text{ or } \pm \pi/2$ respectively.

The basic idea is to find an integer k and two clases of re-

|E|_{02k0} is moderately large
 Any two phases in class 1 differ by 0 or π
 Any two phases in class 11 differ by 0 or π
 Any phase in class 1 differs from any in class 11 by approximately π/2

5) Every [E] in class 1 and 11 be large

Assuming 1) and 5) are satisfied, the classes are created by placing in class 1 cosine invariants having high values (presumably

+1) and in class 11 those having small values (presumably -1) to produce:

Class l

 $\phi_{n\bar{k}l} = \frac{1}{2}\phi_{02k0} \quad \text{or} \quad \frac{1}{2}\phi_{02k0} + \pi \qquad k = 2n$ $= \frac{1}{2}\phi_{02k0} + \pi/2 \qquad k = 2n + 1$

Class 11

·..

 $\phi_{h\bar{k}l} = \frac{1}{2}\phi_{02k0} + \frac{\pi}{2}\pi$ k = 2n = $\frac{1}{2}\phi_{02k0}$ or $\frac{1}{2}\phi_{02k0} + \pi$ k = 2n + 1

To ensure 2) and 3) pertain, only phases ϕ_{hkl} , $\phi_{h'kl}$, which interact strongly are retained, i.e. have large

$$A = 2/N^{\frac{1}{2}} \left| E_{hk\ell} E_{h'k\ell'} E_{h+h'o\ell+\ell'} \right|$$

and large (calculated) values (= +1) for at least one of the following cosines

$$\cos \left(\phi_{\bar{\mathbf{h}} \bar{\mathbf{k}} \bar{\boldsymbol{\ell}}}^{\dagger} + \phi_{\bar{\mathbf{h}}} \cdot \bar{\mathbf{k}} \bar{\boldsymbol{\ell}}}^{\dagger} + \phi_{\bar{\mathbf{h}} + \bar{\mathbf{k}} \bar{\mathbf{k}} \bar{\mathbf{k}}}^{\dagger} + \phi_{\bar{\mathbf{h}} + \bar{\mathbf$$

In a similar manner 4) is verified by retaining those phases which interact weakly with the other class and have small calculated cosines (\approx 0).

A second method uses a generalisation of the tangent formula. It can be shown (Appendix A) that

If \underline{k} ranges over an arbitrary set of reciprocal space then it can be considered as composed of subsets within which A is constant. Therefore,

1.3.16
$$\langle |\mathbf{E}_{\mathbf{k}}\mathbf{E}_{\mathbf{h}-\mathbf{k}}| \sin(\phi_{\mathbf{h}} + \phi_{\mathbf{h}-\mathbf{k}} + \phi_{\mathbf{k}}) \rangle_{\mathbf{k}} = 0$$

With

$$C = \langle |E_{k}E_{h-k}| \cos(\phi_{k} + \phi_{h-k}) \rangle_{k}$$

$$S = \langle |E_{k}E_{h-k}| \sin(\phi_{k} + \phi_{h-k}) \rangle_{k}$$

1.3.16 can be put in the form

 $C \sin \phi_h - S \cos \phi_h = 0$

$$\operatorname{Tan}_{h} = S/C$$

While ϕ_h will satisfy 1.3.16 it will not necessarily satisfy

1.3.17
$$<\cos(\phi_{h} + \phi_{h-k} + \phi_{k})|_{A>} = I_{1}(A)/I_{O}(A)$$

To incorporate both 1.3.16 and 1.3.17 a function Φ is defined thus:

$$\Phi = \langle w_{s} \sin^{2}(\phi_{h} + \phi_{h-k} + \phi_{k}) + w_{c} [\cos(\phi_{h}^{-} + \phi_{h-k} + \phi_{k}) - I_{1}(A)/I_{o}(A)]^{2} \rangle_{k}$$

W_s and w_c are the reciprocals of the variances of $\sin(\phi_{h-k} + \phi_{\bar{h}} + \phi_{\bar{k}})$ and $\cos(\phi_{\bar{h}} + \phi_{\bar{k}} + \phi_{h-\bar{k}})$ respectively, and I₁(A) and I₀(A) are modified Bessel functions.

If Φ is plotted over the interval O to 2π , then for enantiomorph sensitive reflexions (assuming more than one term in the summation) two equal minima will be found, preferably well defined, at approximately + $\pi/2$.

Starting Set Selection

The starting set, which contains reflexions that specify origin, and, if appropriate, enantiomorph, as well as reflexions that take a selection of values (or a symbolic value), has many possible members.

To ensure a good phase development, members of the starting set clearly should have large |E| values and a high degree of interaction with other strong reflexions, requirements met by reflexions with a high $\alpha_{\rm h}$.
As there is no phase information initially, an approximation for α_{h} must be used:

$$<\alpha_{h}^{2}>^{32} = \Sigma \kappa_{hh}^{2} + \Sigma \kappa_{hh} \kappa_{hh} (I_{1}(\kappa_{hh})/I_{0}(\kappa_{hh}))$$

 $\times (I_{1}(\kappa_{hh})/I_{0}(\kappa_{hh}))$

 $I_1(\kappa)/I_0(\kappa) \simeq 0.5658\kappa - 0.1304\kappa^2 + 0.0106\kappa^3$

(with a maximum error of 4% for $5 > \kappa$).

If one calculates $\langle a_h^2 \rangle$ for all reflexions, removes from the data set the reflexion with the lowest value, modifies $\langle a_h^2 \rangle$ to allow for the removal, repeats the process, all the time listing eliminated reflexions and $\langle a_h^2 \rangle$ values, until the process terminates, then the result is a list that 'converges' onto the reflexions that interact most strongly with the data set; a consequence of the listing is that $\langle a_h^2 \rangle$ is calculated only from reflexions lower down the list; a low value for $\langle a_h^2 \rangle$ means that the phase is not well indicated by the reflexions below; in the extreme case of $\langle a_h^2 \rangle = 0$ there is a break in the phase expansion necessitating the reflexion being given a selection of values ($\pm \pi/4$, $\pm 3\pi/4$).

However, a large number of multisolution reflexions is clearly very expensive in computing time; hence the development of the method of magic integers³³.

This is based on a set $\{n\}$ of carefully chosen integers and the proposition that a value x(o<x<1) exists that satisfies approximately for n phases ϕ_i , the following equations:

 $\phi_{i} = n_{i} \times \text{mod} (1) \qquad i = 1, n$ where ϕ_{i} are phases expressed in cycles.
For example:

 $\phi_1 = 2 \times \mod (1) \qquad \phi_5 = 2 \times \mod (1)$ $\phi_2 = 3 \times \mod (1) \qquad \phi_6 = 3 \times \mod (1)$ $\phi_3 = 7 \times \mod (1) \qquad \phi_7 = 7 \times \mod (1) .$ $\phi_4 = 13 \times \mod (1) \qquad \phi_8 = 13 \times \mod (1)$

If there are Σ_2 relationships between these phases (and known phases) . then a convenient notation is

Cos $2\pi(Hx + Ky + a) \stackrel{1}{=} 1$

(will tend to be close to, but less than, one).

Finding the maximum (s) of the function

 $\psi(\mathbf{x}, \mathbf{y}) = \Sigma \mathbf{w}_{i} \cos(\mathbf{H}_{i}\mathbf{x} + \mathbf{K}_{i}\mathbf{y} + \mathbf{a}_{i})$ $\mathbf{w}_{i} = \text{weight of } \Sigma_{2} \text{ interaction, } |\mathbf{E}_{h}\mathbf{E}_{h-k}\mathbf{E}_{k}|$

by evaluation of ψ along x and y gives the required phases, which can then be refined independently³⁴, the magic integer link being broken, using as a criterion the maximisation of ψ .

Using the sequence 4 6 7 a r.m.s. (root mean square) error of 26.6° for the phases is obtained compared with a r.m.s. error for random phases of $\Pi/\sqrt{3}$ (103.9°).

Figures of Merit (F.O.M.'s)

Since starting sets usually contain multisolution reflexions, various figures of merit have been developed to isolate the correct phase expansion.

One of the earliest was the ψ_0 test³⁵. It is based on the Sayre equation and uses reflexions of high magnitude $|E_{h-k}|$, $|E_k|$ in calculating reflexions of low magnitude E_h .

$$\Psi_{o} = \Sigma \left| \Sigma E_{k} E_{h-k} \right|$$

Another F.O.M. based on the Sayre equation is R_{K}^{22} For reflexions with known phase the calculated structure factor is

$$E_{h \text{ calc}} = c\Sigma E_{h-k} E_{k} = |E_{k}| e^{i\varphi_{k}} \text{ etc.}$$

As the summation is only over known phases, $c \neq \sigma_3^{-1}\sigma_2^{3/2}$, but instead is obtained by requiring

$$\Sigma (|E_{h}|_{calc})^{2} = \Sigma (|E_{h}|_{obs})^{2}$$
$$R_{K} = \Sigma ||E_{h}|_{obs} - |E_{h}|_{calc} |/\Sigma |E_{h}|_{obs}$$

F.O.M.'s based on the self consistancy of the phase set, such as Z $(= \Sigma \alpha_n^2)$, ³⁶ have some use, but should be avoided with certain types of space groups, especially symmorphic space groups, where the most favourable F.O.M.'s will correspond to trivial solutions, such as all $\phi_h = 0$.

An interesting attempt at an absolute figure of merit is represented by NQEST⁴⁰. It is based on the four phase invariants, the negative quartets.

If, $h + k + \ell + m = 0$, $|E_h^{\dagger}|, |E_k^{\dagger}|, |E_{\ell}^{\dagger}|, |E_m^{\dagger}|$ and $B(2|E_hE_kE_{\ell}E_m^{\dagger}|/N)$ are larg (e.g. > 1.7 and > 1.0 respectively), and $|E_{h+k}^{\dagger}|, |E_{h+\ell}^{\dagger}|, |E_{\ell+k}^{\dagger}|$ are small (e.g. < 0.7) then

> $\cos(\phi_{h} + \phi_{k} + \phi_{\ell} + \phi_{m}) \simeq -1$ and NQEST = $\Sigma B \cos(\phi_{h} + \phi_{k} + \phi_{\ell} + \phi_{m})/\Sigma B$.

Phases are derived by tangent expansion from the S.S.P.'s⁴¹. An ideal value would, of course, be -1.

A similar F.O.M. is H.K.C.^{14,16,42}. Other F.O.M.'s proposed include those based on a Σ_1 criterion⁴³ and a more general Sayre equation⁴⁴.

1.4 Least Squares

Consider an experiment in which we make m measurements g_{i} . Each measurement is subject to an unknown error e_{i} . These are assumed to be uncorrelated.

It is desired to obtain from these measurements the best estimates of the values of n parameters x_j , related to the observed quantities by m linear equations, with known coefficients d_{ij} , henceforth referred to as the observational equations:

 $g_{1} = d_{11} x_{1} + d_{12} x_{2} + \dots d_{1n} x_{n} + e_{1}$ $g_{2} = d_{21} x_{1} + d_{22} x_{2} + \dots d_{2n} x_{n} + e_{2}$. $g_{m} = d_{m1} x + d_{m2} x_{2} + \dots d_{mn} x_{n} + e_{m}$

In matrix form

DX + E = G D = Design matrix of coefficients d_{ij} X = Matrix of parameters x_j E = Matrix of random errors e_m G = Matrix of measurements g_i

A necessary, but not sufficient condition for any progress to be made is that $m \ge n$.

To obtain the best (minimum variance) estimates for x_j the function $M = \sum_{i} (g_i - \sum_{i} x_j)^2$ where w_i is the weight for an observation, is minimised.

In matrix notation

$$M = (G-DX)^T W (G-DX)$$

with w_{ii} , the elements of the m x m diagonal matrix W, giving a measure of the reliability of the observation g_i ; ideally w_{ii} is in-

versely proportional to the variance σ_{i}^{2} of the observation. Upon minimising M, i.e. the weighted squared differences between the observed values and the calculated values from DX, if the weights are correctly chosen, minimum variance estimates can be obtained. To minimise M, n partial derivatives $\partial M/\partial x_{i}$ are found.

$$\partial M/\partial x_j = 2\Sigma w_i d_{ij} (g_i - \Sigma d_{ij} x_j) = 0$$

 $j = 1, n$

The normal equations are obtained by rearranging these n equations

In matrix notation

$$M = (G-DX)^{T} W (G-DX)$$

= $G^{T}WG - (DW)^{T} W G - G^{T}W DX + (DX)^{T}WDX$

Differentiating

$$O = O - D^{T}WG - D^{T}WG + 2D^{T}WDX$$

Rearranging as before we get the normal equations in the form

$$D^{T} WDX = D^{T}W G = NX$$

In order that N have an inverse, the matrix D must be of rank n. As N is an nxn matrix, having as elements the coefficients on the right hand side of 1.4.1, it is clear that the matrix is symmetric, with diagonal terms being necessarily positive, and of high magnitude, relative to off-diagonal terms.

 N^{-1} has the useful property of being proportional to C, the covariance matrix, whose diagonal terms are the variances of the parameter estimates; the off-diagonal terms are generally much smaller and represent correlation between parameters.

 $\rho_{jk} = C_{jk} / \sqrt{C_{jj}C_{kk}}$ with ρ_{jk} the correlation between the parameters x_j, x_k .

The way least squares are utilised in crystallography is to use a linear approximation. By Taylors theorem of the mean

$$f(x_{1} + \Delta x_{1}, x_{2} + \Delta x_{2} \dots x_{n} + \Delta x_{n}) = f(x_{1}, x_{2} \dots x_{n}) + (\Delta x_{1} \partial \partial x_{1} + \Delta x_{2} \partial \partial x_{2} \dots \Delta x_{n} \partial \partial x_{n})$$

$$x f(x_{1}, x_{2}, \dots x_{n}) \dots + 1/n! (\Delta x_{1} \partial \partial x_{1} + \Delta x_{2} \partial \partial x_{2} \dots \Delta x_{n} \partial \partial x_{n})^{n}$$

$$x f(x_{1}, x_{2}, \dots x_{n}) + R_{n}$$

Neglecting all except the first two terms

$$f(x_1 + \Delta x_1, x_2 + \Delta x_2, \dots, x_n + \Delta x_n) - f(x_1, x_2, \dots, x_n) \approx (\Delta x_1 \partial / \partial x_1 + \Delta x_2 \partial / \partial x_2 \dots + \Delta x_n \partial / \partial x_n) f(x_1, x_2, \dots, x_n)$$

If $f(x_1 + \Delta x_1, x_2 + \Delta x_2, \dots, x_n + \Delta x_n) \equiv |F_o|$, the observed structure factor, and $f(x_1, x_2, \dots, x_n) \equiv |F_c|$, the calculated structure factor, then the observational equations, after weighting, take the form

$$w^{l_2}Dx = w^{l_2}G$$

$$w^{l_2}G_i = \sqrt{w_i}\Delta_i$$

$$x_j = \Delta x_j$$

$$w^{l_2}D_{ij} = \sqrt{w_i}\partial|F_c|i/\partial x_j$$

so that after minimising M, the normal equations are

$$NX = D^{T}WDX = D^{T}WG$$

$$(D^{T}WD)_{jk} = \Sigma w_{i} \partial |F_{c}|_{i} / \partial x_{j} \partial |F_{c}|_{i} / \partial x_{k}$$

$$(D^{T}WG)_{j} = \Sigma w_{i} \Delta_{i} \partial |F_{c}|_{i} / \partial x_{j}$$

W are estimated from counting statistics, with an additional comi ponent for instrumental instability. Assuming the weights are of the form $\sigma^2 / \sigma_i^2(F_i)$ and there is a random, normal distribution of errors $\Sigma w_i \Delta_i^2 / (m-n) (= S^2)$ will have a $\chi^2_{(m-n)} / (m-n)$ distribution⁴⁶. The expectation value of a $\chi^2_{(m-n)}$ distributed function is (m-n). Ideally then, with correct weighting, $\sigma^2 = 1$ and S², sometimes called the variance (estimated) of observations of unit weight, should also equal one. However, as for observations of small magnitude $\sigma^2(|F|)$ may be of the same order as |F|, clearly estimates for $\sigma^2(F)$ will tend to be unreliable, leading to S values that differ from unit. Analysis of S² in terms of sin θ / λ , h,k,l or in any other systematic manner is a useful check for systematic error. S should have a nearly constant value⁴⁷.

The variance of the parameters is given by

$$\sigma^{2}(x_{j}) = S^{2}(N^{-1})$$

while the covariance between parameters x_i , x_k can be estimated as

S² (N⁻¹) jk

K, the scale factor, when refined presents a few problems when off-diagonal terms of $D^{T}W$ D are neglected to save storage, as it. correlates with the temperature factor^{45,48}.

As mentioned earlier, for the inverse of N(n x n) to exist, it must be of rank n. Singular normal matrixes arise from treating as independent parameters that are not; for example in a symmorphic space group such as P2 the y coordinate of one atom must be fixed to specify the origin; if it is not then one y coordinate will be linearly dependent upon the others and hence cause singularity. Another instance arises where an atom has a four-fold axis so that

 $U_{11} = U_{22}$ and $U_{12} = U_{23} = U_{13} = 0$.

These difficulties can be obviated by using parameters that are distinct from, though they may be made equal to, one or more atomic

parameters. In such a case

$$dw = dx_1 = dx_2...dx_n$$
$$d|F_c|/dw = \partial|F_c|/\partial x_1 + \partial|F_c|/\partial x_2... + \partial|F_c|/\partial x_n$$

For structures, such as proteins, where the number of reflexions, due to disorder, only exceeds the number of atomic parameters by a relatively small margin, conventional refinement is clearly inappropriate.

Although a priori calculation of the structure is conceivable 49,50 , for larger structures the computing effort becomes prohibitive. To improve the atomic coordinates therefore, a slightly different technique must be used 51 . One approach tried was to assume 'ideal' bond lengths and angles (planar trans amide groups in proteins, for example) and then, by a least squares procedure, fit this idealised structure to the X-ray co-ordinates. Finally, an energy minimisation procedure was applied $^{51-55}$.

A more elegant method is to incorporate distance and angle restraints between atoms in the observational equations; as all the information about the structure is considered, refinement to a chemically reasonable steroechemistry is very rapid, even if only a relatively small number of reflexions are used ^{56,57}.

References

- 'Theory of Bessel Functions', G. N. Watson, Cambridge University Press (1966) p.77.
- IUPAC IUB. Commission on Biochemical Nomenclature, Biochem. 9, p.3471 (1970).
- 'The Crystalline State' II 'The Optical Properties of the Diffraction of X-rays', R. W. James, G. Bell & Sons (1967).
- 4. W. L. Bragg, Proc. Roy. Soc. (A), 123, p.537 (1929).
- 5. J. Almlöf and T. Otterson, Acta Cryst. A35, p.137 (1979).
- R. F. Steward, E. R. Davidson and W. T. Simpson, J. Chem. Phys., 42, p.3175 (1965).
- 7. A. J. C. Wilson, Acta Cryst., 2, p.318 (1949).
- 8. J. Karle and H. Hauptman, Acta Cryst., 9, p.635 (1956).
- 9. J. Karle and H. Hauptman, Acta Cryst., 11, p.757 (1958).
- 10. A. J. C. Wilson, Nat., 150, p.151 (1942).
- 11. R. Rothbauer, Acta Cryst., <u>A34</u>, p.528 (1978).
- 12. D. Harker and J. S. Kasper, Acta Cryst., 1, p.70 (1948).
- 13. K. Banerjee, Proc. Roy. Soc. (A), 141, p.188 (1933).
- 14. H. Schenk, Acta Cryst., A30, p.477 (1974).
- 15. H. Schenk, Acta Cryst., <u>A29</u>, p.480 (1973).
- 16. H. Schenk and J. G. H. De Jung, Acta Cryst., A29, p.31 (1973).
- 17. J. Karle and H. Hauptman, Acta Cryst., 3, p.181 (1950).
- 18. E. W. Hughes, Acta Cryst., <u>2</u>, p.34 (1949).
- 19. D. Sayre, Acta Cryst., 5, p.60 (1952).
- 20. G. Germain, P. Main and M. M. Woolfson, Acta Cryst., <u>A27</u>, p.368 (1971).
- 21. S. E. Hull and M. J. Irwin, Acta Cryst., A34, p.863 (1978).
- 22. J. Karle and I. L. Karle, Acta Cryst., 21, p.849 (1966).
- 23. A. J. C. Wilson, Acta Cryst., 3, p.258 (1950).
- 24. D. Rogers and A. J. C. Wilson, Acta Cryst., 6, p.439 (1953).
- 25. F. H. Herbstein and F. R. L. Schoening, Acta Cryst., <u>10</u>, p.657 (1957).

26.	H. Lipson and M. M. Woolfson, Acta Cryst., <u>5</u> , p.680 (1952).
27.	H. Hauptman, Acta Cryst., <u>17</u> , p.1421 (1964).
28.	H. Hauptman, 'Crystal Structure Determination; The Role of the Cosine Seminvariants', New York, London, Plenum (1972).
29.	J. Karle and H. Hauptman, Acta Cryst., <u>14</u> , p.217 (1961).
30.	H. Hauptman and J. Karle, Acta Cryst., <u>12</u> , p.93 (1959).
31.	H. Hauptman and W. L. Duex, Acta Cryst., <u>A28</u> , p.393 (1972).
32.	G. Germain, P. Main, M. M. Wolfson, <u>B26</u> , p.274 (1970).
33.	P. S. White and M. M. Woolfson, Acta Cryst., <u>A31</u> , p.53 (1975).
34.	P. Main, Acta Cryst., <u>A33</u> , p.750 (1977).
35.	W. Cochran and A. S. Douglas, Proc. Roy. Soc. (A), <u>243</u> , p.281, (1957).
36.	G. Germain and M. M. Woolfson, Acta Cryst., <u>B25</u> , p.91 (1968).
37.	R. M. Moon and C. G. Schull, Acta Cryst., <u>17</u> , p.805 (1964).
38.	W. H. Zachariasen, Acta Cryst., <u>23</u> , p.558 (1967).
39.	P. Seiler and J. D. Dunitz, Acta Cryst., <u>A34</u> , p. <u>329</u> (1978).
40.	H. Hauptman, Acta Cryst., <u>A30</u> , p.822 (1974).
41.	G. T. de Titta, J.W. Edmunds, D. A. Langs and H. Hauptman, Acta Cryst., A31, p.472 (1975).
42.	H. Schenk, Acta Cryst., <u>A31</u> , p.259 (1975).
43.	A. R. Overbeek and H. Schenk, Proc. Kon. Ned. Akad. Wet., <u>B79</u> , p.341 (1976).
44.	M. M. Woolfson, Acta Cryst., <u>11</u> , p.393 (1958).
45.	'Computing Methods in Crystallography', Ed. Rollet, Pergamon Press.
46.	International Tables for X-ray Crystallography IV, p.293, Kynoch Press, Birmingham.
47.	A. J. C. Wilson, Acta Cryst., <u>A32</u> , p.994 (1976).
48.	'Statistics in Physical Science', W. C. Hamilton, Ronald Press (1964).
49.	J. A. Reader and P. W. G. Smith, J. Chem. Soc., p.3479 (1965).

.

- 50. W. B. Gratzer, W. Rhodes and G. D. Fasman, Biopolymers, <u>1</u>, p. 319 (1963).
- 51. P. K. Warme and H. A. Scheraga, Biochem., <u>13</u>, p.757 (1974).
- 52. N. Go, M. Go and H. A. Scheraga, Proc. Nat. Acad. Sci. US, 59, p.1030 (1968).
- 53. N. Go, P. N. Lewis and H. A. Scheraga, Macromol., <u>3</u>, p.628 (1970).
- 54. N. Go, and H. A. Scheraga, J. Chem. Phys., <u>51</u>, p.4751 (1969).
- 55. P. N. Lewis, F. A. Momany and H. A. Scheraga, Isr. J. Chem., <u>11</u>, p.121 (1973).
- 56. A. Jack and M. Levitt, Acta Cryst., <u>A34</u>, p.931 (1978).
- 57. L. G. Hoard and C. E. Nordman, Acta Cryst., A35, p.1010 (1979).

CHAPTER TWO

DEHYDRO-PEPTIDES: CHEMISTRY AND OCCURRENCE

2.1 Introduction

In recent years the number of biochemically active small molecules discovered has increased greatly.

An especially interesting group are the dehydro-peptides. The biological existence of dehydro-peptides, while having been suspected for many years, was strongly indicated by the discovery of a dehydro-peptidase¹ in 1932. Dehydro-peptidase activity was subsequently found in all plant and animal tissues studied^{2,3}.

These dehydro-peptidases were purified and separated into dehydropeptidase I and II. It was subsequently shown that neither was specific for dehydro-peptides, but operated on saturated systems as well, being an amino acid acylase and amino peptidase respectively. Their dehydropeptidase activity must be considered adventitious; a specific dehydropeptidase has not been shown to exist (at least in higher organisms).

Over the years several dehydro-peptides were isolated⁴⁻⁶ but it was not until dehydro-amino acid residues were identified in more familiar compounds, such as nisin that widespread interest began to be shown. Table 2.1.1 lists a selection of the dehydro-peptides and derivatives that have been found.

Whether dehydro-peptides have a significant role in mammalian biochemistry is unclear, though dehydro-alanine is believed to be at the active site of rat (liver) L-His ammonia lyase⁷. Also dehydroamino acids have an ephemeral existence during the degradation of hydroxy and sulphur containing amino acids (by dehydrases and desulphydrases respectively), while other amino acids can go via the imino acid (under the influence of L-oxidases, though interestingly D-oxidases of greater activity are also present), which is tautomeric with the dehydro-amino acid:



Knowledge of the general structural characteristics of these compounds, relatively unexplored^{8,9}, to facilitate relation of biochemical activity to structure, is clearly desirable. To this end, the structures of a series of N-acyl (acetyl) dehydro-amino acids and peptides were studied and compared with peptides. N-acyl derivatives allowed study of the interaction of the amide system with the unsaturated carbon as well as providing an uncharged derivative (usually essential for stable N-terminal dehydro-amino acid peptides).

Table 2.1.1

Peptide

Nisin¹⁰ Yeast phenylalanine · ammonia lyase^{11,12} Alternaria Mali Toxin¹³

Stendomycin¹⁴ Berninamycin¹⁵

Oestreogrysin A¹⁶

Albonoursin¹⁷

Telomycin¹⁸

Tentoxin¹⁹

Dehydro-amino Acid

Alanine

Butyrine

Proline

Leucine } cyclic dipeptide

Tryptophan

N-methyl Phenylalanine

2.2 Preparation

The first dehydro-peptides were prepared by Erlenmeyer²⁰ by the hydrolysis of unsaturated oxazolones, themselves prepared from aromatic aldehydes and glycine or acylated glycine (Figure 2.2.1, R=CH₃).

Though widely used, the method achieved its greatest flowering in the hands of Bergmann and co-workers. They extended dehydro-peptides at the carboxyl end by opening the oxazolone with an aqueous solution of a sodium salt of an amino acid or with an amino acid ester in a nonaqueous solvent²¹⁻²³ (Figure 2.2.1). Furthermore by generalising, using for example, as the acylated glycine derivative, N-acetyl Δ Phe-Gly-OH (Figure 2.2.1, R=G), an oxazolone can be formed, which can then yeild a peptide with two (or more) dehydro-phenylalanyl residues.

Peptides containing dehydro-aliphatic amino acid residues cannot be practically prepared by the original Erlenmeyer procedure, as the aliphatic aldehydes required contain acidic hydrogens. However, it is practicable if the reaction is split so acetic anhydride is absent when the aldehydes/ketones are added²⁴. Also, by the use of a related reaction the use of aldehydes can be avoided altogether. Thus, if N-chloroacetyl-leucine is warmed with acetic anhydride (especially in the presence of pyridine), the unsaturated azlactone is formed, by elmination of hydrogen chloride and water (Figure 2.2.2(a)).

For biochemical studies it is often desirable that the substrate has both free amino and carboxyl groups; the first representative of this class of compound, Gly Δ Phe-OH, was prepared by careful reaction of N-chloroacetyl-phenylserine with acetic anhydride, to form, via β elmination, 2-chloromethyl-4-benzylidene-oxazolin-5-one, which, after hydrolysis, was reacted with ammonia (Figure 2.2.2(b))²⁵⁻²⁷; by removing the N blocking group it is even possible to obtain the free amino group



Figure 2.2.1. The Erlenmeyer synthesis; R^1 is a suitable group such as CH₃ or CH₃C(O)N(H)C(CH ϕ) (G); 1 and 2 are Ac₂O/NaOAc and OH respectively while 3 is an amino acid salt or ester.



(b)







Figure 2.2.2. Variations on the Erlenmeyer reaction; 1 is Ac₂O/pyridine, 2 Ac₂O, 3 D.C.C.I./T.H.F./5^O C and 4 D.D.Q.; Pep is the amino end of a peptide chain and R is appropriate for the amino acid residue. on at least some unsaturated amino acid residues $^{28-30}$.

A recent variation on the Erlenmeyer procedure, essentially involves formation of the saturated oxazolone with D.C.I.I.³¹ and oxidation³² with D.D.Q. (2,3-dichloro-5,6-dicyano-1,4 benzoquinone) to the unsaturated oxazolone (see Figure 2.2.2(c)).

Using the observation that acetamide and pyruvic acid, heated under reduced pressure, lead to the formation of a mixture of N-acetyl Δ Ala-OH (α acetamido-acrylic acid) and 2,2-diacetamido propanoic³³ acid (Figure 2.2.3), a convenient procedure for preparing N-acetyl Δ Ala-OH (and derivatives), involving heating excess pyruvic acid with acetamide (or derivatives) and removing water azeotropically has evolved³⁴. Heating 2,2-acetamido propanoic acid (conveniently prepared by the Böttinger synthesis³⁵ (Figure 2.2.3)) also results in the same product.

A rather similar reaction involves mixing the appropriate nitrile (though only if it has an α halogen) with an excess of pyruvic acid and saturating the mixture with dry HCL (Figure 2.2.4(a)), though hydrolysis of the nitrile becomes more significant as R becomes larger².

An elegant general method for producing dehydro-amino acid esters is shown in Figure 2.2.4(b). It involves N-chlorination by tertiary butyl hypochlorite (tert-B.H.C.), dehydrochlorination to the corresponding imino acid ester and tautomerising to the thermodynamically more favoured, α,β dehydro-amino acid³⁶⁻³⁹.

Within the last twenty years interest in protein sequence determination has prompted study of dehydro-peptide formation by β elimination from 0 derivatives of serve residues (Figure 2.2.5)⁴⁰⁻⁴³.

An elegant use of this reaction was in a study of the role of serine in chymotrypsin⁴⁴; previous studies had sought to show its necessity by forming an O derivative, but this was open to the objec-



Figure 2.2.3. Dehydro-alanine formation: 1 and 2 consume $RCONH_2$ and 2 $RCONH_2$ respectively (A is RCO-); 1 is favoured when the water is removed; 3, 4 and 5 require glacial acetic acid, chilled H_2SO_4 and RCN/H_2O respectively.





Figure 2.2.4. 1, 2, 3 and 4 require dry $HC1/-5^{\circ}$ C, tert-B.H.C., D.B.U. and EC1 followed by NH_3 respectively.



Figure 2.2.5(a) β elimination pathways. (b) Thermolytic formation of dehydro-peptides; 1 is TsCl/pyridine/-5^O C while 2 and 3 are [O] and $\Delta/P\phi_3$ respectively. tion that it could be a steric effect; forming the dehydro-alanyl (inactive) derivative demonstrated conclusively that the hydroxyl group was essential.

However, β elimination may not be the only reaction, e.g. O-tosyl derivatives, amongst other by-products^{45,46}, can give rise to aziridines⁴⁷ (stereospecifically) (Figure 2.2.5(a)), especially when R is CH₃ (or presumably an alkyl group generally) and/or R₃ is NH-X (a general amide). Other leaving groups can also give rise to undesirable by-products^{48,49}.

For these and other reasons a method that requires only mild, neutral conditions has recently been developed⁵⁰. It relies upon thermolytic elimination from β alkyl sulphinyl derivatives (Figure 2.2.5(b)), themselves prepared from the readily available amino acid sulphides⁵¹. The conversion, by oxidation, however, introduces the serious drawback that cysteine, cystine and tryptophan are oxidised irreversably. The temperature required for thermolysis was also sometimes rather high (140° C); to obviate this difficulty use was made of the observation that elimination is reversible; by adding thiophiles, such as triphenyl phosphine, the reaction temperature is lowered considerably. As it has been shown that aromatic sulphenic acids are better leaving groups than alkyl sulphenic acids⁵², replacement of the S-benzyl or S-methyl groups used by S-aromatic groups should allow elimination at any reasonable temperature.

The 'obvious' method of dehydro-peptide synthesis, by incorporation of dehydro-amino acids using standard coupling techniques has a number of drawbacks:

The conditions for the removal of carboxyl blocking groups would be inimical to the integrity of the dehydro-peptide; the use of hydrogenolysis for blocking group removal is totally precluded in subsequent steps and, until recently, the preparation of dehydro-amino acids with

free amino groups was not practicable^{28,53,54}. The propensity of dehydro-peptides to react with various reagents⁵⁵⁻⁵⁸ makes desirable the introduction of the double bond as late in the synthesis as possible. Finally, and most importantly, the carboxyl group may be difficult or impossible to activate in a manner that produces coupling⁵⁹.

The coupling may be accomplished however, if the N-acyl group or stereo/chemical factors (proline) make azlactone formation difficult^{60,61}.

2.3 Biosynthesis and biochemistry

The biosynthesis of dehydro-peptides obviously cannot proceed by the usual mechanism for protein synthesis, with the same applying to proteins containing D amino acids.

A proposed mechanism 62 that produces both is shown below:



This mechanism produces α epimerisation naturally, in conformation with the observation that where there is more than one chiral centre, inversion generally only occurs at the α carbon⁶³ (though allo-derivatives have been isolated).⁶⁴ Additionally it is clear that derivatives without an amide hydrogen, and N-methyl amino acids occur frequently in antibiotics, could not follow this mechanism and indeed, such derivatives with a D configuration have not so far been found. Although a dehydro-peptide, tentoxin, with a methylated amide group exists, there is no obvious reason why methylation should not have occured after synthesis of II. The isolation of a tryptophan side chain α,β -oxidase provides support for this route⁶³.

Imines or dehydro-peptides provide possibly a plausable intermediate in the synthesis of the large number of exotic amino acids that have, in recent years been isolated from various fungal, plant and bacterial sources. For example, hydration of I would produce the α hydroxyalanine and α hydroxyvaline systems found in ergotamine and ergocristine⁶⁶. These and similar reactions could of course occur

intramolecularly and it has been proposed that the formation of a number of compounds proceeds in this manner (Table 2.3.1). However the generality of this proposal has been restricted by observations showing that the D-Val residue in penicillin is not formed via an α,β dehydrovaline⁶⁷.

Chemists have shown interest in such reactions nevertheless, as there is presently no direct or stereospecific way to introduce substituents into a peptide. Incorporation of substituted amino acids by standard coupling techniques suffers from the disadvantage that substituted amino acids are generally synthesised by the difficult and laborious process of condensing fragments, which themselves often have a lengthy preparation. The introduction of double bonds along the carbon chain would enable direct^{14,58,68} and possibly stereospecific derivatization^{69,70}. By varying the reaction conditions the position of substitution can be made α or β^{71} . More adventurous reactions such as photochemical cycloadditions are also possible⁷². Additionally, unsaturated analogues of biologically active peptides may themselves show desirable characteristics⁷³.

In studies of the complement system it was found that peptides containing the Phe-Tyr sequence inhibited an enzyme (EAC 1423) responsible for the mobilisation of the complement response⁷⁴. In a search for more effective inhibitors a number of dehydro-peptides were tried of which two, N-AcA-3-(2-R)Ala-Tyr-OH (R is thienyl or furyl) were active, but in the opposite sense expected; there was a dose dependent, delayed (reaching a peak 48 hours after treatment of the mice) activation (measured by in vitro cancer cell destruction) of the murine macrophages presumably mediated by an allosteric effect on the enzymes responsible for complement activation (attempts at in vitro activation of the macrophages were unsuccessful)⁷⁵.



Antibiotics and their putative precursors

Dehydro-amino acid residue



Serine



Cysteine



Arginine

Antibiotic

Oestreorgrysin A¹⁶ Griseoviridin⁷⁸

Micrococcin⁷⁶ Thiostrepton⁷⁷



Capreomycin $(R^1 = R^2 = H)^{79,80}$ Viomycin $(R^1 = OH, R^2 = H)^{81}$ Stendomycin $(R^1 = H, R^2 = CH_3)$

References

1.	Μ.	Bergmann	and	H.	Schleich,	z.	Physiol.	Chem.,	205,	p.65	(1932).	•
----	----	----------	-----	----	-----------	----	----------	--------	------	------	---------	---

- 2. V. E. Price and J. P. Greenstein, J. Biol. Chem., 171, p.477 (1947).
- C. E. Carter and J. P. Greenstein, J. Natl. Cancer Inst., 7, p.51 (1946).
- 4. B. Heinemann, M. A. Kaplan, R. D. Muir and I. R. Hooper, Antibiot. and Chemoth., 3, p.1239 (1953).
- R. Brown, C. Kelly and S. E. Wibberley, J. Org. Chem., <u>30</u>, p.277 (1965).
- 6. N. N. Gerber, J. Org. Chem., <u>32</u>, p.4055 (1967).
- 7. I. L. Givot and R. H. Abeles, J. Biol. Chem., 245, p.3271 (1970).
- O. Pieroni, G. Montagnoli, A. Fissi, S. Merlino and F. Ciardelli, J.A.C.S., <u>97</u>, p.6820 (1975).
- 9. O. Pieroni, A. Fissi, S. Merlino and F. Ciardelli, Isr. J. Chem., 15, p.22 (1977).
- 10. E. Gross and L. J. Morell, J.A.C.S., 89, p.2791 (1967).
- 11. D. S. Hodgins, J. Biol. Chem., 246, p.2977 (1971).
- 12. K. R. Hansen and E. A. Havir, Arch. Biochem. Biophys., <u>141</u>, p.1 (1970).
- T. Okuno, Y. Ishita, A. Sugawara, Y. Mori, K. Sawai and T. Matsumoto, Tetrahedron Lett., p.335 (1975).
- 14. M. Bodanszky, J. Izdelski, I. Muramatsu and A. Bodanszky, Paper read at the European Peptide Symposium, Orsay, France (1968).
- 15. J. M. Liesch and J. M. Rinehart, Jr., J.A.C.S., <u>99</u>, p.1645 (1977).
- 16. G. R. Delpierre, F. W. Eastwood, G. E. Gream, D. G. I. Kingston, P. S. Sarin, Lord Todd and D. H. Williams, J. Chem. Soc.(c), p.1653 (1966).
- A. S. Khokholov and G. B. Lokshin, Tetrahedron Letters, p.1881 (1963).
- J. C. Heehan, D. Mania, S. Nakmura, J. A. Stock and K. Maeda, J.A.C.S., 90, p.462 (1968).
- 19. W. L. Meyer, G. E. Templeton, C. I. Grable, R. Jones, L. F. Kuyper, R. B. Lewis, C. W. Sigel and S. H. Woodhead, J.A.C.S., <u>97</u>, p.3802 (1975).

20. E. Erlenmeyer and E. Früstück, Ann., 284, p.48 (1895).

- 21. M. Bergmann and K. Grafe, Z. Physiol. Chem., <u>187</u>, p.196 (1930).
- M. Bergmann, V. Schmidt and A. Miekeley, Z. Physiol. Chem., <u>187</u>, p.264 (1930).
- 23. M. Bergmann, F. Stern and C. Witte, Ann., 449, p.277 (1926).
- 24. M. Crawford and T. Little, J.C.S., p.729 (1959).
- 25. M. Bergmann, A. Miekeley and E. Kann, Z. Physiol. Chem., <u>146</u>, p.247 (1928).
- 26. M. Bergmann and F. Stern, Ann., <u>448</u>, p.20 (1926).
- 27. D. G. Doherty, J. E. Tietzman and M. Bergmann, J. Biol. Chem., <u>147</u>, p.617 (1943).
- 28. E. G. Breitholle and C. H. Stammer, Tetrahedron Letters, p.2381 (1975).
- 29. F. Weygand and V. Glocker, Chem. Ber., <u>89</u>, p.653 (1956).
- 30. M. Weygand, W. Steglich and H. Tanner, Ann., 658, p.128 (1962).
- 31. De Los F. De Tar, R. Silverstein and E. F. Rogers, Jr., J.A.C.S., 88, p.1024 (1966).
- 32. S. Konno and C. H. Stammer, Synthesis, p.598 (1978).
- 33. M. Bergmann and K. Grafe, Z. Physiol. Chem., <u>187</u>, p.187 (1930).
- 34. T. Wieland, G. Ohnacker and W. Ziegler, Chem. Ber., 90, p.194 (1957).
- 35. C. Böttinger, Ber., 14, p.1599 (1881).
- 36. U. Schmidt and E. Öhler, Angew. Chem. Int. Ed. Engl., <u>16</u>, p.327 (1977).
- 37. H. Poisel, Chem. Ber., <u>110</u>, p.948 (1977).
- 38. H. Poisel and U. Schmidt, Chem. Ber., <u>108</u>, p.2547 (1975).
- 39. H. Poisel and U. Schmidt, Chem. Ber., <u>108</u>, p.2917 (1975).
- 40. B. Witkop, "Advances in Protein Chemistry", <u>16</u>, (1961): Academic Press.
- 41. M. Sokolovsky, M. Wilchek and A. Patchornik, Bull. Res. Council Isr., <u>11A</u>, p.79 (1962).
- 42. A. Patchornik and M. Sokolovsky, Bull. Res. Council Isr., <u>11A</u>, p.80 (1962).

43. I. Photaki, J.A.C.S., 85, p.1123 (1963).

44. D. H. Strumeyer, W. N. White and D. E. Koshland, Jr., Proc. N.A.S., USA, <u>50</u>, p.931 (1963).

- 45. C. Zioudrou and G. L. Schmir, J.A.C.S., <u>85</u>, p.3258 (1963).
- 46. L. Benoiton, R. W. Hanson and H. N. Rydon, J. Chem. Soc., p.824, (1964).
- 47. Y. Nakagawa, T. Tsuno, K. Nakajima, M. Iwai, H. Kawai and K. Okawa, Bull. Chem. Soc. Japan, 45, p.1162 (1972).
- 48. E. Schnabel, J. Stoltefuss, H. A. Offe and E. Klauke, Liebigs Ann. Chem., <u>743</u>, p.57 (1971).
- 49. A. Srinwasan, R. W. Stephenson and R. K. Olsen, J. Org. Chem., <u>42</u>, p.2253 (1977).
- 50. D. H. Rich and J. P. Tam, J. Org. Chem., <u>42</u>, p.3815 (1977).
- 51. I. Z. Eiger and J. P. Greenstein, Arch. Biochem., 19, p.467 (1948).
- 52. B. M. Trost, T. N. Salzmann and K. Hiroi, J.A.C.S., <u>98</u>, p.4887, (1976).
- 53. F. McCapra and M. Roth, J.C.S. Chem. Commun., p.894 (1972).
- 54. C. Shin, M. Masaki and M. Ohta, J. Org. Chem., <u>32</u>, p.1860 (1967).
- 55. E. Gross, K. Noda and B. Nisula, Angew. Chem. Int. Ed. Eng., <u>12</u>, p.664 (1973).
- 56. A. L. Love and R. K. Olsen, J. Org. Chem., <u>37</u>, p.3431 (1972).
- 57. A. Patchornik and M. Sokolovsky, J.A.C.S., 86, p.1206 (1964).
- 58. R. S. Asquith and P. Carthew, Tetrahedron, 28, p.4769 (1972).
- 59. D. H. Rich, P. Bhatnagar, P. Mathiaparanam, J. E. Grant and J. P. Tam, J. Org. Chem., 43, p.296 (1978).
- 60. A. M. Felix, C. Wang, A. A. Liebman, C. M. Delaney, T. Mowles, B. A. Burghardt, A. M. Charnecki and J. Meienhofer, Int. J. Peptide Protein Res., <u>10</u>, p.299 (1977).
- G. Lucente, A. Romeo and G. Zanotti, Gazzetta Chimica Italiana, <u>102</u>, p.941 (1972).
- 62. B. W. Bycroft, Nature, 224, p.595 (1969).
- 63. M. Bodanszky and D. Perlman, Nature, <u>218</u>, p.291 (1968).
- 64. W. A. König, H. Kneifel, E. Beyer, G. Muller and H. Zahner, J. Antibiotics, 26, p.44 (1973).
- Y. Noda, K. Takai, T. Tokuyama, S. Narumiya, H. Ushiro and
 O. Hayaishi, J. Biol. Chem., <u>252</u>, p.4413 (1977).
- 66. A. Stoll and A. Hoffmann, "The Alkaloids", p.8: Academic Press, London and New York (1965).

- 67. F. C. Huang, J. A. Chan, C. J. Sih, P. Fawcett and E. P. Abraham, J.A.C.S., <u>97</u>, p.3858 (1975).
- 68. J. M. Riordan and C. H. Stammer, Tetrahedron Letters, p.4969 (1971).
- 69. E. Öhler, E. Prantz and U. Schmidt, Chem. Ber., 111, p.1058 (1978).
- 70. S. Lee, T. Kanmera, H. Aoyagi and N. Izumiya, Int. J. Pep. Protein Res., 13, p.207 (1979).
- 71. P. J. Machin and P. G. Sammes, J.C.S. Perkin I, p.698 (1974).
- 72. P. K. Sen, C. J. Veal and D. W. Young, J. Chem. Soc. Chem. Comm., p.678 (1977).
- 73. G. H. Fisher, D. I. Marborough, J. W. Ryan and A. M. Felix, Arch. Biochem. Biophys., <u>189</u>, p.81 (1978).
- 74. A. Kossarotow, W. Opitz, E. Etschenberg and U. Hadding, Biochem. J., 167, p.377 (1977).
- 75. H.-U. Schorlemmer, W. Opitz, E. Etschenberg and D. Bitter-Suermann, Cancer Res., 39, p.1847 (1979).
- 76. P. Brooks, R. J. Clark, A. T. Fuller, M. P. V. Mijovic and J. Walker, J.C.S., p.928 (1960).
- 77. M. Bodanszky, J. T. Sheehan, J. Fried, N. J. Williams and C. A. Birkheimer, J.A.C.S., 82, p.4747 (1960).
- 78. G. I. Birnbaum and S. R. Hall, J.A.C.S., <u>98</u>, p.1926 (1976).
- 79. E. B. Herr, "Antimicrobial Agents and Chemo-Therapy 1962", p.201, (1963).
- 80. E. B. Herr and M. O. Redstone, Ann. New York Acad. Sci., <u>135</u>, p.940 (1966).
- 81. B. W. Bycroft, D. Cameron, L. R. Croft, A. Hassanali-Walji, A. W. Johnson and T. Webb, Tetrahedron Letters, p.5901 (1968).

CHAPTER THREE

EXPERIMENTAL DETAILS

3.1 Preparations

The aim was to prepare crystalline peptide and dehydropeptide derivatives and determine their molecular structures by X-ray diffraction techniques; N-acetyl derivatives were favoured as it allows study of an extra amide link.

A number of preparative methods were tried:

Carbodiimide Method

The coupling of N-acetyl dehydro-alamine or N-acetyl dehydrophenylalanine with various amino acid esters was attempted.

In the case of N-acetyl dehydro-alanine (N-AcAAla-OH) the attempt was totally unsuccessful; with N-acetyl dehydro-phenylalanine (N-AcAPhe-OH), instead of the expected product, 2-methyl 4-benzylidene oxazolin-5-one was formed in appreciable yield. With hindsight the latter result is not unexpected, as, in contrast to saturated azlactones, unsaturated azlactones are, because of conjugation, rather stable.



An examination of the structures of N-Ac- Δ Phe-OH and N-Ac- Δ Ala-OH (Figure 6.2.1 and Figure 4.2.1 respectively) makes clear why azlactone formation would be slower in the latter case; while N-Ac- Δ Phe-OH has the correct stereochemistry for azlactone formation, N-Ac- Δ Ala-OH requires a 180° rotation about N4-C5 to product a suitable stereochemistry.

β elimination

Two peptides, N-Ac D, L-Ser-L-Ala-OEt and N-AcL-Ala-L-Ser-OEt, were prepared using D.C.C.I. and, in the latter case, HOBT (not being used in the former case because of the danger of lactone formation¹).

The hydroxyl group of the serine residue was then reacted to form a suitable derivative for β elimination. The leaving group chosen was the toluene sulphonate anion. In an effort to simplify isolation 'polytosyl chloride' was prepared by chlorosulphonation of a polystyrene matrix². However, handling characteristics proved unsatisfactory and so tosyl chloride chloride was used instead.

It was found in both cases β elimination occured spontaneously giving rise to non-crystalline products.

Bergmann Synthesis

The N-acetyl dehydro-peptide N-Ac Δ PheGly-OH and analogues with Gly replaced by L-Ala, L-Leu, L-Phe, L-Pro, L-Ser-OEt were prepared from 2-methyl 4-benzylidene oxazolin-5-one in the usual manner³⁻⁵; N-Ac Δ Phe Δ Phe-OH was also prepared.

Although the products were crystalline, the crystals were, with the exception of N-Ac Δ Phe-L-Pro-OH, unsuitable for single crystal diffraction. The dicyclohexyl ammonium salt of N-Ac Δ Phe-L-Leu-OH was prepared in an effort to obtain a suitable crystal, but to no avail.

Miscellaneous Methods

A method that was tried a number of times was that of Greenstein for N-chloroacetyl dehydro-alanine⁶. Despite azeotropic drying of the reactants, in every case the principal product was chloroacetamide. The preparation of the same compound was tried by a modification of Wielands method⁷, but with no success.

N-Ac Δ Leu-OMe was prepared from the imine⁸, but could not be induced

to crystallise. The free acids preparation from the corresponding azlactone was tried, but the purification of the azlactone was not achieved.

The growth of crystals, suitable for diffraction work, of Z-Gly Δ Ala-OEt, BOC-D,L-Phe Δ Ala-OMe and Tos⁻⁺H₃N D,L-Phe Δ Ala-OMe was attempted, but unsuccessfully. Tos⁻⁺H₃N D,L-Phe Δ Ala-OMe was converted to the chloride by passage through C.G.401[C1⁻], but proved to be an oil.

Infra-red spectra were recorded on a Perkin-Elmer 357 while u.v. spectra were obtained from a Perkin-Elmer 402. T.l.c. used air dried SiO₂, the solvent system being EtOAc/C₆H₆ (4:1).

Experimental

N-Ac∆Phe-L-Ser-OEt

N-Ac Δ Phe-OH (5 mmol) was dissolved in 20 mL T.H.F., D.C.I.I. and HOBT (5 mmol each) washed in with 10 mL T.H.F. and the reaction mixture cooled to 0° C with a freezing mixture.

After 1.5 hours D.C.U. (1.02 g) was filtered off. ECL.Ser-OEt (5 mmol) was added to the filtrate in 30 mL T.H.F./acetonitrile (2:1) together with diisopropylethylamine (5 mmol). The solvent was removed under reduced pressure after a further 2 hours and the resultant solid mixed with 70 mh $H_2O/EtOH$ (4:3) and filtered.

Following unsuccessful efforts to isolate the expected product from the filtrate, the $H_2O/EtOH$ insoluble residue was dissolved in chloroform, D.C.U. filtered off and the solvent removed; the solid remaining was shown by m.p. (150[°] C), n.m.r., u.v., i.r. and t.l.c. to be identical with 2-methyl 4-benzylidene oxazolin-5-one.

Yield (based on reacted reagent) was 0.46 g (50%). The desired compound was subsequently prepared in the following manner:

5 mmol of 2-methyl 4-benzylidene oxazoline-5-one (azlactone) was refluxed in acetone with HCl.L-Ser-OEt/diisopropylethylamine (5 mmol) for 5 hours. The solvent was removed, 120 mL EtOH/H₂O (1:5) added, and the mixture filtered to remove unreacted azlactone (0.25 g). T.l.c. revealed the main product at $R_f = 0.19$. The product was purified with a column (65 g Kieselgel 60 : 40 mL H₂O), the column being developed with EtOAc/C₆H₆ and the eluant being analysed by t.l.c.

 $m.p. = 2 \ 147^{\circ}$ C, i.r. (KBr) 3300, 1730, 1640 cm⁻¹

Yield = 0.33 g (29%).

N-Ac Dhe-L-Pro-OH

To a solution of proline (10 mmol) in 16.7 mL of 0.6 M NaOH was added 10 mmol of azlactone and 8 mL of acetone. The mixture was shaken until the azlactone dissolved and allowed to stand until t.l.c. indicated reaction was complete (4 hours). 11 mL of IMHCl were then added and the acetone blown off. Tarry material was removed and the solution left overnight in the fridge. The resulting crystalline precipitate was filtered off and recrystallised from EtOH/H₂O.

m.p. = $141^{\circ} - 143^{\circ}$ C, i.r. (KBr) 3200, 1730, 1660 cm⁻¹ Yield = 1.7 g (56%).

Analogues of this compound with L-Pro replaced by other amino acids were prepared in a similar manner.

N-AcL-Ala-L-Ser-OEt

10 mmol of N-AcL-Ala-OH and HCL.L-Ser-OEt were suspended in 60 mL of acetonitride at 0⁰ C. A solution of HoBT, diisopropylethylamine and D.C.C.I. (10 mmol each) in 15 mL acetonitride was then added and the reactants left stirring overnight.

After filtering off D.C.U. (2.0 g), the acetonitrile was removed under reduced pressure and the resultant oil mixed with cold water (30 mL). The precipitate of HOBT was filtered off and the filtrate passed
through a column of cation exchange resin (Amberlite C.G. 120[Na⁺]). Following removal of the water, ethyl acetate was added and NaCl filtered off. Upon removal of the solvent the product separated as an oil which solidified when left overnight in the fridge. It was recrystallised from EtOH/n-C₆E₁₄.

 $[\alpha]_{D} = -75^{\circ}$ (c = 1, H₂O), m.p. = $136^{\circ} - 138^{\circ}$ C i.r. (KBr) 3310, 3090, 1620, 1550 cm⁻¹

Yield = 2.1 g (85%)

N-AcD,L-Ser-L-Ala-OEt (but with no HoBT) and N-AcL-Ala-L-Ala-OEt were prepared in a similar manner.

N-AcL-AlaO-Tosyl-L-Ser-OEt

N-AcL-Ala-L-Ser-OEt (2 mmol) was dissolved in 5 mL of dry pyridine (stored at 3° C over 3 Å molecular sieves) and 2.1 mmol of tosyl chloride (purified by the Pelletier method¹⁰).

After 30 hours t.l.c. showed that reaction was complete and indicated that β elimination had spontaneously occurred (strong $R_f = 0$, presumably toluene sulphonic acid). However, to ensure complete reaction diethylamine was still added. The reaction mixture had water added and was passed through a bed of ion exchange resin (C.G. $120[Na^+]$). The solvent was removed under reduced pressure (pyridine forms an azeotrope with water) and the residue dissolved in acetone/water. The undissolved portion was filtered off (sodium tosylate, m.p. > 260° C) and the solvent removed to leave an oil which could not be crystallised.

u.v. 230 n.m. i.r. (NaCl disc) 3300, 1640, 1530 cm⁻¹ . Yield = 0.55 mmol (27.5%)

The preparation of N-AcO-Tosyl D,L-Ser-L-Ala-OEt proceeded in a similar manner; the product was again non-crystalline.

3.2 Data Collection, Computing and Miscellaneous Details

The unit cell parameters and space group were usually determined initially from Weissenberg photographs. Data were collected on a Nonius CAD4 four circle diffractometer using the $\omega/2\theta$ scan mode. Crystal integrity was monitored by periodically remeasuring reflexions.

The geometry of the CAD4 differs somewhat from that of the classical four circle diffractometer (Figure 3.2.1).

The goniometer head, mounted on the ϕ (PHIK) axis, is supported by the κ block. This can be rotated about the κ axis (which makes an angle \propto , nominally 50°, with the ϕ axis), while the block on which it is carried, the ω block, is itself rotatable about the ω (OMK) axis. Additionally, the detector rotates on an axis (20) which is coaxial with the ω axis.

The zero positions for κ , ω , and 20 are defined in terms of the geometry of the instrument; the point in κ rotation, where PHIK and OMK coincide, is defined as $\kappa = 0$; the zero for ω is defined as the point in ω rotation where the κ axis lies in the XZ plane and the κ block is most distant from the collimator, with a similar definition for 20. The point where the key on the goniometer head mount is parallel to the Y axis and κ and ω are zero defines $\phi = 0$.

Data consistency is estimated from

$$R' = (\Sigma N \Sigma w_{ij} (F_j - F_{ij})^2 / \Sigma (N - 1) \Sigma w_{ij} F_{ij}^2)^{\frac{1}{2}}$$

where internal summations are over equivalent reflexions and \overline{F} is the average of equivalent reflexions.

Computing was done on the I.C.L. 1900, 2980 or C.D.C. 7600. The structures were solved using either SHELX 76 or MULTAN 78. Indices used to measure progress in structure determination are defined below:



Figure 3.2.1. The CAD4 diffractometer and its geometry.

$$R = \Sigma \Delta_{i} / \Sigma |F_{o}|_{i}$$

$$R_{w} = \Sigma (w^{2} \Delta)_{i} / \Sigma (w^{2} |F_{o}|)_{i}$$

$$R_{G}^{2} = \Sigma (w \Delta^{2})_{i} / \Sigma (w |F_{o}|^{2})_{i}$$

 $R_M = R_G$ calculated with a scale factor that minimises R_G . Molecular geometry was calculated using XANADU. Least-squares planes were calculated using atoms at unit weight; hydrogen atoms were not used. The equations of the planes are given in right-handed Certesian coordinates with X_o parallel to X, Y_o parallel to ZxX* and Z_o parallel to Z.

Structures were drawn using ORTEP; hydrogen was given the dummy temperature factor of B = 1.

In the absence of units, distances are quoted in Angstroms, angles in degrees, i.r. peaks in cm^{-1} and u.v. values in n.m.

References

1.	W.	Konig	and	R.	Geiger,	Chem.	Ber.,	103,	788	(1970)	•
-											

- 2. Deutsches Patentamt Auslegescrift 1016445 (1957).
- 3. D. G. Doherty, T. E. Tietzman and M. Bergmann, J. Biol. Chem., 147, 617 (1943).
- O. Behrens, D. G. Doherty and M. Bergmann, J. Biol. Chem., <u>136</u>, 61 (1940).
- 5. J. P. Greenstein and M. Winitz, "Chemistry of the Amino Acids" II, p.827 (1961), John Wiley and Sons Inc.
- 6. V. E. Price and J. P. Greenstein, J. Biol. Chem., <u>171</u>, 477 (1947).
- 7. T. Wieland, G. Obnacker and W. Ziegler, Chem. Ber., <u>90</u>, 194 (1957).
- 8. H. Poisel and U. Schmidt, Chem. Ber., <u>108</u>, 2547 (1975).
- T. Miyazawa, T. Shimanouchi and S. Mizushima, J. Chem. Phys., 29, 611 (1958).

10. S. W. Pelletier, Chem. Ind., 1034 (1953).

CHAPTER FOUR

N-ACETYL DEHYDRO-ALANINE

4.1 Introduction

The dehydro-alanyl residue occurs with frequency in two classes of biologically active molecules: antibiotics and enzymes.

It was first detected in nisin, and later in subtilin, as part of the carboxyl terminal sequence, dehydro-alanyl lysine; significantly this feature appears essential for the antibiotic activity.¹ The necessity or otherwise of the other dehydro-alanyl residues (one each for subtilin and nisin plus a β methyl dehydro-alanyl (butyrine) residue for nisin) is unknown.

Interestingly nisin (and subtilin) contain lanthionine and β methyl lanthionine residues, which could plausably be formed by addition of cysteine to dehydro-alanyl and butyrine residues respectively, as mercapto derivatives generally add readily². A special feature of nisin activity that was predicted (and found) on this basis was antimaterial activity, as the malarial parasite is known to be inhibited by a shortage of Co-enzyme A(HS-CoA)³. Another suggested role for addition (after the intermediate production of dehydro-alanine) is as the final stage of a mechanism for replacement of β substituents⁴.

The inactivating activity of nisinase, which appears to act as a reductase or dehydro-peptidase on both nisin and subtilin⁵, suggests its use as a quick screening method for antibiotics (and other molecules?) containing dehydro-alanyl residues.

Enzymes containing dehydro-alanine are generally ammonia lyases, with dehydro-alanyl residue at the active site. A mechanism involving a Schiff base has been proposed⁶.

Following the observation of in vitro inactivation of His-ammonia lyase by nitromethane, the in vivo inactivation (in rats) was

attempted (successfully), by the same means, to model a human genetic disorder, histidinaemia⁷. However, because another degrading enzyme existed (His-pyruvate transaminase) the modelling was unsuccessful; it would seem to have shown, none-the-less, at least provisionally, that dehydro-alamine has no essential structural biochemical role in mammals.

Further details of the chemistry and biochemistry can be found in Chapter 2.

4.2 N-Acetyl Dehydro-Alenine

Experimental

The compound was prepared by Wielands method⁸ and recrystallised from methanol.

Formula - $C_5H_7NO_3$ Space group = $P2_1/n$ = 272 a = 3.941 Å b = 10.156 Å c = 14.898 Å β = 91.64° v = 596.05 Å^3 z = 4 $D_o = 1.43 \text{ Mg m}^{-3} (n-C_6H_{14}/CCl_4)$ $D_c = 1.45 \text{ Mg m}^{-3}$ $\mu = 0.078 \text{ mm}^{-1}$

Data were collected using Mo-K_{α} radiation monochromated with graphite (1.5 < θ < 27[°]); 1570 reflexions (30 were intensity controls) were measured, 1489 of which were systematically present (1283 unique, 1133, 88.3% |F| > 3 σ (|F|).

After data reduction SHELX was used to produce an E-map, from which the structural skeleton could be picked out. After three cycles of refinement with isotropic temperature factors R = 16, three cycles with anisotropic temperature factors produced R = 8.7. The hydrogens were located on a difference map and after insertion and six further cycles, R = 4.4.

An attempted weighted refinement, with the weighting parameter (g) being refined did not converage. 'g' was therefore fixed during refinement, being altered between refinements until the analysis of

variance was satisfactory. The final values are shown below.

	R	RW	$R_{M} (= R_{G})$	s ²
all F	4.8	4.9	5,9	0.043
F > 3σ(F)	4.3	4.7	5.9.	1.000

A final difference map revealed quite a noisy background with significant ripples between most of the bonded atoms.

The final atomic parameters are shown in Tables 4.2.1 and 4.2.2. $10|F_0|/10|F_c|$ tables are in Appendix E (Table E.1).

Discussion

The initial impression of a planar molecule was confirmed by least squares plane analysis (Table 4.2.3, plane 1). Although the planarity might be expected due to the resultant conjugation of the carboxylic, olefin and amide systems, it produces some very small intra-molecular non-bonded distances (Table 4.2.4).

Rather more surprising, considering the very small a axis, is that the angle between the least squares plane of the molecules normal and the a axis is as large as 35° .

A comparison of the bond lengths and angles with those of peptides (Table 4.2.8) show the angles agree within two degrees except for $\hat{C-N-C^{\alpha}(C2-N4-C5)}$. Other dehydro-amino acids (Chapter 6) still show this angle as anomalously large. Examination of the intra-molecular non-bonded distances, especially involving O3, suggest, together with the N4-C5-C6 value of 127°, the angular expansion occurs to relieve steric strain.

The bond lengths are similar except for the N-C^{α} and C-N distances which are significantly shorter and longer respectively than those in peptides⁹; the shortening of the first bond would be expected as it is an sp² - sp² link, but together with the lengthening of the second bond it suggests some delocalisation is taking place. N-Ac- Δ Phe-OH (see 6.2) shows a similar but less pronounced effect, lass which is consistent with delocalisation as the amide plane is skewed out of the plane of the olefinic system.

Fumaramic acid¹⁰, a closely related, essentially planar molecule, shows several similarities to N-Ac- Δ Ala-OH; the most noteworthy are an insensitivity of the olefinic bond length to molecular planarity, a characteristic often found¹¹⁻¹⁴, and a significant difference in carbonyl bond distances, the amide carbonyl being long (1.233 Å) than

the carboxylic (1.206 Å). The last observation is related to the first in that it implies that contributions from resonance hybrids involving the carbonyl are not significantly changed from the isolated carboxylic acid group.

An analysis of hydrogen bonding (Table 4.2.5) reveals that, although constraints are placed upon hydrogen bonding in such an inflexible molecule, it takes place to the maximum, albeit limited, extent. A detailed examination of the hydrogen bonding shows a number of interesting features; the C7-08...E4 (168°) and C7-08...N4 (172.7°) values show the oxygen lone pair to be in an sp orbital, which is fairly unusual for a carboxylic carbonyl and probably differs from the situation in N-Ac- Δ Phe-OH (see 6.2); also the C2-03...E4 (124°) and C2-03...09(120.7°) values suggest the lone pairs on the amide oxygen are in sp² orbitals which, while being usual for amides in general and agreeing with the other two dehydro-peptides (see for example 5.2 and 5.3). The bond formed by H9, while being, not unexpectedly, shorter than the bond involving E4 is, nevertheless, longer than observed for equivalent bonds in most amide-carboxylic complexes (2.50 ± 0.02 Å)¹⁵.

The carboxylic acid group is planar (Table 4.2.3, plane 4) and its parameters C7-09 (1.310 Å), C7-08 (1.206 Å), 08-C7-09 (124.8°) agree well with those expected (1.31 Å, 1.21 Å and 123° respectively)¹⁶. As the aforementioned planarity, somewhat surprisingly, even includes the hydrogen, which only deviates from the least squares plane by 0.04 Å (with an approximate e.s.d. of 0.03 Å) no accommodation for the hydrogen bond to 03 seems to have been necessary. The hydrogen is in the conformation (synplanar) which minimises the distance to the carbonyl oxygen, presumably due to an electrostatic interaction, though

packing considerations may also be important. In fact the symplanar conformation has been variously estimated as being more stable than the antiplanar by 2 Kcal/mole¹⁷ or at least 4 Kcal/mole¹⁸. This is consistent with the observation of the antiplanar form only when an intra-molecular hydrogen bond is formed.

The inflexible nature of the molecule makes a rigid body librational correction appropriate as can be seen from the excellent agreement between $(U_{ij})_{OBS}$ and $(U_{ij})_{CALC}$ (Table 4.2.6) and $R_{G}(10.7%)$.

Figure 4.2.2, a projection down the b axis, shows the molecular packing. It can be rationalised as hydrogen bonded ribbons, parallel to the a c diagonal, the ribbons being held together by hydrogen bonding to form a distorted sheet, and sheets cohering by Van der Waals forces.

The carboxylic carbonyl group C7-08 is antiplanar to the olefinic bond C5-C6 (see figure 4.2.1). A comparison, assuming similar bond lengths and angles, with the non-bonded distances of the synplanar conformation 09...N4 (2.52 Å) 08...C6 (2.79 Å) provides at least a partial rationalisation. While being a very useful rule-of-thumb for saturated acids, producing the 'correct' synplanar conformation (for unbranched acids), in unsaturated (α,β) acids steric factors are often ambiguous and it has been suggested, not altogether convincingly, using the bend bond model for double bonds¹⁹, that the antiplanar conformation will then be found. However, for less simple molecules more distant parts can influence the conformation (see 6.2).



Figure 4.2.1. ORTEP drawing of the molecular structure of N-AcAAla-OH with 50% probability ellipsoids; the view is parallel to the axis.



· • •

Figure 4.2.2. A projection down to the b axis showing the molecular packing.

Atomic fractional co-ordinates with the e.s.d.'s in parenthesis and of the same magnitude as the final digit.

	x	У	Z
Cl	0.3274(4)	0.4872(1)	1.2393(1)
C2	0.4365(3)	0.3667(1)	1.1905(1)
03	0.6025(3)	0.2781(1)	1.2275(1)
N4	0.3472(3)	0.3608(1)	1.1023(1)
C5	0.4322(3)	0.2620(1)	1.0410(1)
C6	0.6234(4)	0.1571(1)	1.0566(1)
С7	0.2819(3)	0.2902(1)	0.9493(1)
08	0.1317(3)	0.3910(1)	0.9324(1)
09	0.3323(3)	0.1951(1)	0.8918(1)
HLA	0.254(6)	0.462(2)	1.298(2)
HIB	0.164(6)	0.534(2)	1.209(2)
HIC	0.519(8)	0.548(3)	1.250(2)
H4	0.252(4)	0.427(2)	1.080(1)
нба	0.659(4)	0.098(2)	1.007(1)
н6в	0.721(4)	0.138(2)	1.114(1)
Н9	0.251(6)	0.216(2)	0.838(2)

Bond angles (degrees). The e.s.d.'s, in parenthesis, are of the same magnitude as the final digit.

HIA-CI-HIB	111(2)
ELA-C1-HIC	106 (2)
H1B-C1-H1C	106 (2)
C2-C1-H1A	109(1)
C2-C1-H1B	113(1)
C2-C1-H1C	111(2)
C1-C2-03	122.4(1)
C1-C2-N4	116.0(1)
03-C2-N4	121.5(1)
H4-N4-C2	117(1)
H4-N4-C5	116(1)
C2-N4-C5	127.1(1)
N4-C5-C6	127.4(1)
N4-C5-C7	111.1(1)
C6-C5-C7	121.4(1)
С5-С6-Н6А	117.7(9)
С5-С6-Н6В	123(1)
H6A-C6-H6B	119(1)
c5-c7-08	122.1(1)
C5-C7-09	113.1(1)
08–C7 –O 9	124.8(1)
C7-09-H9	111(1)

Deviations from least squares planes (in $\stackrel{0}{A}$)

	_ 1	2	· 3	4
C1	0.014		0.002	
C2	-0.023		-0.005	
03	-0.039		0.002	
N4	-0.018	0.002	0.002	
C5	0.016	-0.007		0.000
C6	0.086	0.003		
С7	-0.004	0.002		0.000
80	0.048			0.000
09	-0.081		•	0.000

i	a	b _i	° i	ďi
l	0.8409	-0.4787	-0.2526	-3.4359
2	0.8167	-0.5172	-0.2558	-3.5244
3	0.8509	-0.4589	-0.2558	-3.4620
4	0.8568	-0.4301	-0.2845	-3.5638

-				
(<	Van	der Waals	+ 0.35	A)
	C1		2.46	
	H11	3N4	2.49	
	Hll	3H4	2.24	
	03	C5	2.844	
	03	C6	2.830	
	03	Н6в	2.28	
	C6	C2	3.022	
	C7	E4	2.40	
	08	<u>H4</u>	2.27	
	08	N4	2.665	
	08	н9	2.32	
	09	C6	2.708	
	09	H6A	2.34	

Intra-molecular non-bonded distances

Table 4.2.5

Hydrogen bonding

Donor	Acceptor	Angle	Distance	Distance	Position of
Х-Н	Y	ХНХ _О	HY	XY	Acceptor
N4-H4	08	159	2.4	3.18	(-x,1-y,2-z)
09-н9	03	168	1.7	2.60	(x-12,12-y,z-12)

Observed and calculated U_{ij} (upper and lower line respectively). Hydrogen U_{ij} were not analysed. $R_{G} = 10.7$ %.

	U ₁₁	U ₁₂	U ₁₃	U ₂₂	U ₂₃	U ₃₃
C1	0.0478 0.0495	-0.0011 -0.0009	-0.0090 -0.0100	0.0495 0.0505	0.0046 0.0050	0.0300 0.0296
C2	0.0341 0.0396	0.0073 0.0040	-0.0072 -0.0081	0.0419	-0.0030 -0.0018	0.0246 0.0255
03	0.0639 0.0615	-0.0065 -0.0053	-0.0165 -0.0154	0.0503 0.0500	-0.0023 -0.0020	0.0279 0.0280
N4	0.0404 0.0292	-0.0002 0.0077	-0.0103 -0.0058	0.0372 0.0358	-0.0011 -0.0033	0.0249 0.0249
C5	0.0344 0.0348	0.0069 0.0047	-0.0074 -0.0072	0.0373 0.0366	-0.0028 -0.0027	0.0250 0.0259
C6	0.0478 0.0493	-0.0014 -0.0041	-0.0105 -0.0122	0.0434 0.0418	-0.0010 0.0001	0.0337 0.0321
C7	0.0380 0.0500	0.0042 -0.0019	-0.0076 -0.0118	0.0401 0.0432	-0.0006 0.0001	0.0251 0.0267
08	0.0649 0.0634	-0.0105 -0.0083	-0.0162 -0.0171	0.0468 0.0474	-0.0007 -0.0001	0.0319 0.0290
09	0.0765	-0.0141 -0.0110	-0.0202 -0.0172	0.0538 0.0525	0.0088	0.0294 0.0308
HIA	0.084					
HlB	0.074		-			
HlC	0.090					
Н4	0.037					
нба	0.043					
н6в	0.056					
н9	0.082					

Librational corrections and corrected bond lengths (in Angstroms)

.

Cl-BlA	0.005	0.973
C1-H1B	0.004	0.919
C1-H1C	0.005	0.990
C2-C1	0,004	1.498
C2-O3	0.005	1.239
C2-N4	0.005	1.357
N4-84	0.004	0.839
C5-N4	0.004	1.407
C5-C6	0.005	1.327
C6-H6A	0.003	0.965
С6-н6в	0.004	0.945
C7-C5	0.006	1.507
C7-08	0.005	1.210
C7-09	0.003	1.313
09-н9	0.004	0.880

Table 4.2.8

.

	Peptide	N-Ac-∆Ala-OH
C ^{°−} −C (0)	1.51	1.493
C=0	1.24	1.233
C (O) –N	1.32	1.352
N-C ^{°°}	1.45 .	1.403
c [°] -Ĉ (0) -0	120.5	122.4
C^{α} - $\hat{C}(O)$ -N	116	116.0
0-ĉ-n	123.5	121.5
C (O) −Ñ−C [∝]	122	127.1
N-Ĉ ⁻ -C(O)	111	111.1

REFERENCES

- E. Gross, J. L. Morell and L. C. Craig, Proc. Nat. Acad. Sci. U.S.A., <u>62</u>, p.952 (1969).
- J. P. Greenstein and M. Winitz, The Chemistry of the Amino Acids (Wiley), II, p.865 (1960).
- 3. E. Gross and J. L. Morell, J.A.C.S., 89, p.791 (1967).
- W. A. Newton, Y. Morino and E. E. Snell, J. Biol. Chem., <u>240</u>, p.1211 (1965).
- 5. B. Jarvis, J. Gen. Microbiol., 47, p.33 (1967).
- K. R. Hanson and E. A. Havir, Arch. Biochem. Biophys., <u>141</u>, p.1, (1970).
- 7. I. L. Givot and R. H. Abeles, J. Biol. Chem., 245, p.3271 (1970).
- 8. T. Wieland, G. Ohnacker and W. Ziegler, Chem. Ber., 90, p.194 (1957).
- 9. R. E. Marsh and J. Donohue, Advan. Protein Chem., 22, p.235 (1967).
- V. Benghiat, H. W. Kaufman, L. Leiserowitz and G. M. J. Schmidt, J.C.S. Perkin II, p.1758 (1972).
- 11. S. E. Filippakis, L. Leiserowitz, D. Rabinovich and G. M. J. Schmidt, J.C.S. Perkin II, p.1750 (1972).
- S. E. Filippakis, L. Leiserowitz, and G. M. J. Schmidt, J.C.S.(B), p.305 (1967).
- O. Bastiansen, L. Hedberg and K. Hedberg, J. Chem. Phys., <u>27</u>, p.1311 (1957).
- D. P. Shoemaker, H. Kindler, W. G. Sly and R. C. Srivastava, J.A.C.S., 87, p.482 (1965).
- 15. L. Leiserowitz and F. Nader, Acta Cryst. B33, p.2719 (1977).
- 16. T. Miyazawa and K. S. Pitzer, J. Chem. Phys., 30, p.1676 (1959).
- 17. L. Leiserowitz, Acta Cryst. B32, p.775 (1976).
- 18. D. R. Lide, Ann. Rev. Phys. Chem., 15, p.225 (1964).
- 19. H. Einspahr and J. Donohue, Acta Cryst., B29, p.1875 (1973).

CHAPTER FIVE

N-ACETYL-L-ALA-L-ALA-OET

AND

•.

. .

.

N-ACETYL-L-ALA-L-SER-OET

5.1 Introduction

Because of their biological importance, peptides and proteins have been an area of intensive study.

As the peptide linkage, at least to a first approximation, is planar, the conformation of a peptide chain may be described by the angle between the planes $C_{i-1}^{\alpha}C_{i-1}O_{i-1}N_iC_1^{\alpha}$ and $C_i^{\alpha}C_iO_iN_{i+1}C_{i+1}^{\alpha}$ i.e. $\Phi(C_{i-1}N_iC_i^{\alpha}C_i)$ and $\Psi(N_iC_i^{\alpha}C_iN_{i+1})$. Not all pairs of these angles are possible (due to steric clash) or, if possible, likely; certain combinations of Φ and Ψ are specially favoured as they produce conformations stabilised by well directed hydrogen bonds, in the general case, between amide hydrogen and oxygen; two of the more commonly encountered are the α helix (3.6 residues per repeat unit) and the anti-parallel pleated sheet (two residues per repeat unit), so-called because adjacent hydrogen bonded chains are $N \rightarrow C$, $C \rightarrow N$, $N \rightarrow C$ etc., and the sheet produced has 'folds' in it.

Peptide conformations have been actively studied by model building and computer simulation and islands of stability established^{1,2}. These are usually presented in the form of a Ramachandran diagram, which is a plot of Φ against Ψ contoured to show potential energy values.

Naturally if premises such as peptide bond planarity or Van der Waals radii are incorrect the 'islands' of stability move, but the Ramachandran plot allows a convenient initial analysis of a conformation.

The relationships between Φ , Ψ , n (the number of residues per repeat unit, $n = 360^{\circ}/\Theta^{\circ}$) and d, the distance per residue in the axial direction, assuming average values for the peptide bond, are shown below.

5.1.1. $\cos(\theta/2) = 0.817 \sin((\Phi + \Psi)/2) + 0.045 \sin((\Phi - \Psi)/2)$ 5.1.2. $d \sin(\theta/2) = -2.967 \cos((\Phi + \Psi)/2) - 0.664 \cos((\Phi - \Psi)/2)$. The structure of N-Ac-L-Ala-L-Ala-OEt was determined to compare it with

 $N-A_C-L-Ala\Delta Ala-OEt$. However, as $N-Ac(L-Ala)_2^{OEt}$ suffers from domain anisotropy, which has effects on bond lengths, temperature factors, etc., the structure of N-AcL-Ala-L-Ser-OEt was also determined in the hope (mistaken) that it would be free from such an effect.

Details of preparative methods and chemistry are to be found in Appendix B.

Experimental

The compound was prepared by the carbodiimide method³ and recrystallised from acetone/n-Hexane.

Formula - $C_{10}H_{18}N_2O_4$ Space group = C2 F(000) = 496 a = 14.656 Å b = 9.931 Å c = 9.631 Å β = 112.501° Z = 4 v = 1295.06 Å³ D_o = 1.18 Mg m⁻³ (CCl₄/cyclo-C₆H₁₂) D_c = 1.18 Mg m⁻³

Data were collected on a CAD4 diffractometer using Mo-K_a radiation monochromated with graphite $(1.5^{\circ} < 0 < 27^{\circ})$. 1609 reflexions were measured, 1499 of which were unique $(1031, 69\$ |F|>3\sigma(|F|))$. Attempts to solve the structure using SHELX were unsuccessful, probably due to the lack of weight given to Ψ_{o} in the figure of merit. The structure was subsequently solved with the assistance of Dr. G. M. Sheldrick, using his private program. The structure was then refined, using SHELX until R = 8.6, $R_{M} = R_{G} = 7.0$ (all reflexions). As some of the bond angles, bond lengths and temperature factors were highly variable, domain anisotropy was suspected. Compensation for the effect was achieved by a refinement procedure incorporated in SHELX. Actually it was found that suppression of OO1, O20, $\overline{112}$ and $\overline{202}$ (selected in the

light of the $|F_0|/|F_c|$ analysis) was equally effective.

Weighted refinement then commenced; 'g' was varied until the analysis of variance was satisfactory.

The final indices are shown below:

	R	R _W	$R_{G} = R_{M}$	s²
all F	8.0	6.5	7.3	0.037
F > 2σ(F)	5.1	5.8	7.1	1.498
	• · •	• • •		

The final atomic positions are listed in Table 5.2.1 and Table 5.2.2. $10|F_0|/10|F_c|$ are listed in Appendix E (Table E.4).

Discussion

The peptide groups have the trans disposition that, with the notable exception of proline (which, when associated with reverse open turns, is often in the cis form), is usually found. The energy difference between cis and trans is actually quite small (~ 2.8 Kcal/mole)⁴ but cis is sterically unfavourable in the context of an otherwise trans chain^{5,6}. Least squares plane analysis (Table 5.2.2) shows that only one of the peptide groups (plane 7, $\omega_1 = 180.0^{\circ}$) has the planarity expected; the other has the chiral carbon (C6) deviating from the plane (plane 5, $\omega_2 = 172.0^{\circ}$). The estimated standard deviation (e.s.d.) in the direction of the plane normal may be approximated by the mean of the e.s.d.'s in the directions of the cell axes; 'exact' calculation would require knowledge of the correlation between the axes, which is not readily available; the deviation of C6, at 0.168 A is significant at the 3\sigma level, though naturally the e.s.d.'s represent only lower limits for standard deviations.

Table 5.2.8 compares the values of Ψ and Φ with those of related compounds; it is clear that Φ_1 and Ψ_1 roughly agree with the values for an anti-parallel pleated sheet; the number of residues per repeat unit, at 1.91, compares well with the anti-parallel pleated sheet value of 2.00.⁷

The analysis of hydrogen bonding (Table 5.2.6) shows the antiparallel effect to be produced by the two fold axis, molecules related by the two fold being hydrogen bonded; infinite ribbons, parallel to the ac plane are therefore formed, held together along the b axis by Van der Waals forces between the methyl side groups. A projection down the b axis (Figure 5.2.2) shows the molecular packing.

The hydrogen bond lengths are in the expected range for amide hydrogen bonds. The $C14-O15...H8(179^{\circ})$, $C14-O15...N8(176.2^{\circ})$ and $C9-O10...H13(161^{\circ})$, $C9-O10...N13(166.8^{\circ})$ angles are those expected from hydrogen bonding to sp lone pairs.

 Φ_2 and Ψ_2 , which aren't constrained by hydrogen bonding considerations have values close to but outside the limits of the γ information $(-85^{\circ} < \phi < -60^{\circ}, 95^{\circ} < \Psi < 160^{\circ})^8$, which is stabilised by non-bonded interaction between the amide hydrogen (H8) and the side chain (in this case a methyl group C7) and is favoured by a large side chain, T(C7, C6, N8, H8) is 31[°].

A comparison of the bond lengths and angles with those expected in a peptide⁹ (Table 5.2.9) show good agreement.

The ester group parameters 10 C6-C $\hat{4}$ -O3 (11.6 $^{\circ}$), C6-C4-O5 (126.6 $^{\circ}$), C4-O5 (1.181 Å), C4-O3 (1.343 Å) agree reasonably well with the average values for esters (methyl) of 111 $^{\circ}$, 125 $^{\circ}$, 1.20 Å, 1.34 Å. The latter bond length is, a little surprisingly, longer than the equivalent length in carboxylic acids (1.31 Å).

The ethyl (C1-C2) bond distance is very short (1.42 Å) with a high temperature factor (Table 5.2.2) for Cl, but detailed examination of the electron density map showed no evidence of another stable conformation. The distance, although short, was longer than before suppression of reflexions. The ethyl group is practically planar (Table 5.2.5, plane 1) with the carboxylate group, in contrast to the situation in N-Ac-L-Ala-L-Ser-OEt (see 5.3).

The intersheet distance, perpendicular to the ac plane is clearly 4.96 Å (b/2, the C-centring producing the halfing) compared with the 5.7 $^{011}_{A}$ expected for an antiparallel pleated sheet, indicating a somewhat denser packing.







Figure 5.2.2. A projection down the b axis showing the molecular packing.

Table 5.2.1

•

Atomic fractional co-ordinates. The e.s.d.'s, in parenthesis, are of the same magnitude as the final digit.

x	x	У	Z
Cl	-0.0413(7)	0.211(1)	-0.3310(9)
C2 ·	0.0316(5)	0.2547(8)	-0.1929(9)
03	0.0911(2)	0.1381(5)	-0.1174 (4)
C4	0.1702(3)	0.1603(5)	0.0097(4)
05	0.1917(3)	0.2695(4)	0.0599(4)
C6	0.2247(2)	0.0315(5)	0.0728(4)
C7	0.1664(3)	-0.0569(6)	0.1369(5)
N8	0.3204(2)	0.0621(0)	0.1854(3)
С9	0.3948(2)	0.1079(4)	0.1509(3)
010	0.3901 (2)	0.1141(4)	0.0213(2)
C11	0.4853(2)	0.1527(5)	0.2822(3)
C12	0.4837(4)	0.3054(6)	0.2954(6)
N13	0.5736(2)	0.1066(5)	0.2614(3)
C14	0.6542(2)	0.0686(6)	0.3762(3)
015	0.6578(2)	0.0689(6)	0.5050(2)
C16	0.7401(3)	0.0238(9)	0.3400(5)
н6	0.231(2)	-0.013(4)	-0.007(4)
H7A	0.158(4)	-0.017(6)	0.220(6)
Н7В	0.196(4)	-0.142(7)	0.164(6)
H7C	0.093(3)	-0.068(5)	0.065(5)
н8	0.333(2)	0.069(3)	0.275(4)
Hll	0.485(2)	0.112(3)	0.376(3)
H12A	0.447(4)	0.350(7)	0.319(6)

Table 5.2.1 continued

.

	x	. Y	z
H12B	0.476(3)	0.344(5)	0.194(6)
H12C	0.540(4)	0.324(5)	0.373(5)
н13	0.567(3)	0.096(4)	0.180(5)
H16A	0.783(5)	-0.043(7)	0.399(7)
H16B	0.814(6)	0.091(9)	0.401(8)
H16C	0.723(3)	0.006(5)	0.251(6)

Table 5.2.2

.

Temperature factors - e.s.d.'s, in parenthesis, are of the same magnitude as the final digit.

	U ₁₁	U ₂₂	U ₃₃	U ₂₃	U ₁₃	U ₁₂
Cl	0.163(7)	0.165(7)	0.109(5)	0.028(5)	0.010(5)	0.040(6)
C2	0.098(4)	0.101(4)	0.139(5)	0.004(3)	-0.017(4)	0.013(3)
03	0.083(2)	0.080(2)	0.095(2)	-0.002(2)	-0.002(2)	-0.003(2)
C4	0.058(2)	0.072(2)	0.070(2)	-0.009(2)	0.026(2)	-0.013(2)
05	0.101(2)	0.072(2)	0.109(2)	-0.016(2)	0.018(2)	-0.012(2)
C6	0.056(2)	0.071(2)	0.053(2)	-0.012(1)	0.027(1)	-0.012(1)
C7	0.075(3)	0.086(3)	0.074(2)	0.001(2)	0.036(2)	-0.022(2)
N8	0.055(1)	0.084(2)	0.040(1)	-0.006(1)	0.024(1)	-0.008(1)
C9	0.051(2)	0.074(2)	0.043(1)	0.002(1)	0.025(1)	0.001(1)
010	0.059(1)	0.122(2)	0.040(1)	-0.001(1)	0.025(1)	-0.011(1)
C11	0.050(2)	0.079(2)	0.038(1)	0.001(1)	0.022(1)	-0.001(1)
C12 .	0.072(3)	0.096(3)	0.082(3)	-0.021(3)	0.028(2)	-0.006(2)
N13	0.053(1)	0.110(2)	0.036(1)	0.001(1)	0.024(1)	-0.001(1)
C14	0.054(2)	0.118(3)	0.040(1)	-0.005(2)	0.019(1)	0.004(2)
015	0.080(2)	0.210(4)	0.038(1)	0.009(2)	0.022(1)	0.040(2)
C16	0.057(2)	0.187(6)	0.052(2)	-0.009(3)	0.017(2)	0.023(3)
H6	0.05(1)					
Н7А	0.09(1)					
Н7В	0.10(2)					
H7C	0.09(1)	•				

H8 0.04(1)

H11 0.05(1)

- H12A 0.10(2)
- H12B 0.08(1)
- H12C 0.08(1)
- H13 0.06(1)
- E16A 0.11(2)
- H16B 0.15(2)
- H16C 0.08(1)
Bond lengths (in Angstroms) - e.s.d.'s, in parenthesis, are of the same magnitude as the final digit.

C1-C2	1.42(1)	C9-C11	1.508(4)
C2-03	1,464(7)	C11-C12	1.523(7)
03-C4	1.343(4)	C11-N13	1,456(4)
C4-05	1.181(5)	C11-H11	0.99(3)
C4-06	1,508(6)	C12-H12A	0.80(7)
C6 , C7	1.512(5)	С12-Н12В	1.02(5)
C6-N8	1.439(4)	C12-H12C	0,89(5)
С6-н6	0.92(4)	N13-H13	0.76(4)
C7-H7A	0.95(6)	N13-C14	1,327(4)
с7-н7в	0.93(6)	C14-015	1.221(4)
с7-н7с	1.04(5)	C14-C16	1.498(5)
N8-H8	0.82(3)	C16-H16A	0.94(7)
N8-C9	1.335 (4)	C16-H16B	1.22(8)
C9-010	1.225(3)	C16-H16C	0.82(5)

Bond angles (degrees) - e.s.d.'s, in parenthesis, are of the same magnitude as the final digit.

C1-C2-O3	108.3(6)
C2-03-C4	117.7(4)
03 - C4-05	121.8(4)
03-C4-C6	111.6(3)
05-C4-C6	126.6(3)
С4-С6-Н6	106 (2)
C4-C6-C7	111.2(3)
C4-C6-N8	109.8(3)
H6-C6-C7	109 (2)
H6-C6-N8	111(2)
С6-С7-Н7А	112(3)
С6-С7-Н7В	111(3)
C6-C7-H7C	113(3)
н7а-с7-н7в	110(5)
Н7А-С7-Н7С	100(4)
Н7В-С7-Н7С	110(4)
C6-N8-H8	125 (2)
H8-N8-C9	112(2)
C6-N8-C9	122.5(2)
N8-C9-010	122.6(3)
N8-C9-C11	115.6(2)
010-C9-C11	121.8(2)
С9-С11-Н11	110(2)
C9-C11-C12	109.1(3)
C9-C11-N13	109.6(2)

ø

Table 5.2.4 continued

H11-C11-C12	109 (2)
H11-C11-N13	108 (2)
C11-C12-C12A	127 (4)
C11-C12-C12B	107 (3)
C11-C12-C12C	103(3)
H12A-C12-H12B	103(5)
H12B-C12-H12C	99 (5)
H12B-C12-H12C	117 (4)
C11-N13-H13	114(3)
C11-N13-C14	121.9(2)
H13-N13-C14	123(3)
N13-C14-015	121.7(3)
N13-C14-C16	116.7(3)
015-C14-C16	121.6(3)
C14-C16-H16A	120(4)
C14-C16-H16B	113(4)
C14-C16-H16C	111(3)
Н16А-С16-Н16В	80(5)
H16A-C16-H16C	111 (6)
Н16В-С16-Н16С	119(5)

Deviations (in Angstroms) from the least squares plane. Asterisked atoms are not used in the least squares calculation.

		-	•			· · · ·	
	1	2	3	4	5	6	7
Cl	-0.046						
C2	0.054						
03	0.025	-0.001					
C4	-0.013	0.004				-0.058	
05	-0.021	-0.002				0.034	
C6		-0.001	0.013	0.050	0.168*	0.048	
C7			·.		<u> </u>	<u> </u>	
N8			-0.028	-0.064	0.001	-0.024	<u> </u>
C9			0.030	-0.016	-0.002		
010			-0.015	-0.008	0.001		
C11	,			0.038	0.001		0.007*
C12							
N13							-0.001
C14							0.003
015							-0.001
C16							-0.001

.

Table 5.2.5 continued

Equations of least squares planes

	a i	b _i	c _i	d i
1	-0.5923	-0.1157	0.7974	-2.229
2	-0.5629	-0.1279	0,8165	-2.2082
3	-0.2967	0.9497	0.1002	-0.6751
4	-0.3335	0.9410	0.0583	-0.8030
5	-0.3731	0.9256	0.0637	-1.0493
6	-0.4123	-0.1072	0.9047	-1.8416
7	0.3414	0.9397	-0.0198	3.6612

Hydrogen bonding

Donor	Acceptor	Angle	Distance	Distance	Position of
X-H	Y	XHYO	ΗY	XY	Acceptor
N13-H13	010	155	2.3	2.97	(1-x, y, -z)
N8-H8	015	171	2.1	2.88	(l-x, y, l-z)

Intra-molecular non-bonded close contacts

(< Van der Waals + 0.35Å)

03Н6	2.43	010C6	2.778
o3c7	2.982	010H13	2,45
05C2	2,661	010N13	2.797
05N8	2,742	N13H12C	2.54
H8C11	2.36	N13H16C	2,44
H8H11	2.11	C14H11	2.51
N8Hll	2.46	015Cl1	2.745
C9C4	3.088	015Hll	2.39
С9Н13	2.43	Cl6H13	2.52

Comparison of the conformational angles with similar compounds

		¢ 1	¢2	фз	Ψ1	Ψ2	ψз
Anti-parallel pl	eated sheet ¹²		-142		145		
β	-Poly-L-Ala ¹³		-139		135		
. (L-Ala) ₂ .EC1 ¹⁴		-153		157	163	
	$(L-Ala)_2^{15}$		- 113		165	103	
	Molecule A ¹⁶		-146	-147	153	146	172
(L-AIA) 3	Molecule B		-156	- 160	162	150	144
N-Acetyl	(L-Ala) ₂ -OEt	-145	-74		138	166	

	lst Amide	2nd Amide	Expected Value
с [∞] -С(О)	1.498	1.508	1.51
C=0	1.221	1.225	1.24
C-N	1.327	1.335	1.32
N-C ^{°°}	1.456	1.439	1.45
c [∞] −c ^β	1.523	1.512	1.53
c [∞] -ĉ(o)-0	121.6	121.8	120.5
C[°]- Ĉ (0) -N	116.7	115.6	116
0Ĉ-N	121.7	122.6	123.5
C−Ñ−C [∝]	121.9	122.5	122
N-Ĉ [*] -C(O)	109.6	109.8	111

5.3 N-Acetyl-L-Ala-L-Ser-OEt

Experimental

The compound was prepared by the carbodiimide method³ and recrystallised from EtOH/Et₂O.

> Formula $C_{10}H_{18}N_2O_5$ a = 15.283 Å b = 9.822 Å c = 9.710 Å β = 112.94° Z = 4 V = 1342.29 Å³ D_o = 1.22 Mg m⁻³ (CCl₄/C₆H₁₄) D_c = 1.22 Mg m⁻³

Space group C2

 $\mu = 0.062 \text{ mm}^{-1}$

F(COO) = 528

Data were collected using Mo-K_a radiation monochromated with graphite (1.5 < θ < 30[°]). 2210 reflexions were measured (44 of which were intensity controls), 2067 being unique (1386, 67% $|\mathbf{F}| > 3 \sigma(|\mathbf{F}|)$).

After data reduction, a convergence map was calculated and used to select a starting set. Tangent expansion using SHELX with $E \ge 1.2$ and examination of E-maps with higher figures of merit (RA) failed to reveal the structure.

Consequently a two pronged approach was used:

A. N-Acetyl (L-Ala)2-OEt

The compound is almost isomorphic with the desired compound. A rough index of how closely they are related is given by:

 $ISOM = N\Sigma \left(\left| E \right|_{1}^{2} - \left| E \right|_{1}^{2} \right)^{2} / n$

= o for 'perfect' isomorphism

= 1 for a random relationship

 $N = (2(1 - m^2))^{-1}$

m = 0.798 for centrosymmetric end

= 0.886 for non-centrosymmetric crystals

For N-Ac-L-Alu-L-Ser-OEt and N-Ac-L-Alu-L-Ma-OEt

ISOM = 0.427

Nine atomic positions from N-Ac-L-Ala-L-Ala-OEt were used to phase the reflexions (one atom having a fixed y coordinate), the atoms being chosen so that their positions hopefully wouldn't be altered too much by distortions introduced by the hydroxyl group (such as hydrogen bonding), and a difference map calculated. At this point R = 52%, $R_G =$ $R_M = 55$ %. Atoms revealed in the map were inserted and the process continued until the entire molecule had been found. Least squares refinement then continued isotropically until R = 25%, $R_G = 21$ %, $R_M = 17$ % and thereafter anisotropically. At R = 13%, $R_M = R_G = 11$ %, hydrogen atoms were inserted.

After weighting (g = 0.003) and with all reflexions R = 10%, $R_W = 9\% = R_M = R_G$. This was after a domain anisotropy correction when an F_O/F_C analysis showed evidence of extinction. As well as the reduction in peaks close to atoms, the C-C ethyl bond length 'improved' from 1.38A to 1.43A. A detailed examination of the electron density map didn't reveal any splitting of the Cl peak. A refinement using atoms with variable occupancy was not successful.

The final atomic parameters are shown in Table 5.3.1 and Table 5.3.2. Structure factor tables are in Appendix E (Table E.2).

B. MULTAN 78

By giving the Ψ_{o} component of FOM a high weight (1.3) the structure was solved. This contrast with N-Ac(L-Ala)₂OEt can be explained by noting the beneficial effects of an increase in data. For a fixed number of strongest reflexions (300) and Σ_{2} interactions an increase (from 1499 to 2067) in the number of reflexions leads to an increase in E_{min} , 1.41 for N-Ac-L-Ala-L-Ser-OEt compared with 1.24 for N-Ac(L-Ala)₂OEt and κ_{min} , 1.50 for N-Ac-L-Ala-L-Ser-OEt and 1.39 for N-Ac(L-Ala)₂OEt. This means that any fragmentation of the data set is reduced and a better phase development obtained.

Discussion

While both amide groups have the usual trans configuration, only one of the amide groups shows the expected planarity (Table 5.3.5, plane 1) with $\omega_1 = 179^{\circ}$; the other amide group has the chiral carbon (C8) deviating significantly from the plane (plane 3, $\omega_2 = 173^{\circ}$). Although the resonance stabilisation energy associated with a planar amide group is almost 20 Kcal/mole¹⁷, there is only a small energy penalty for limited rotations about N_i - C_i. Both theoretical and experimental studies performed on amides and other peptide model compounds suggest that rotations of up to 15° are energetically allowable¹⁸⁻²⁰.

Theoretical modelling of the peptide system with ω as a variable indicates that there is an increased conformational flexibility of the peptide chain, with the gain in stability from non-bonded, electrostatic and hydrogen bonding more than compensating for the energy expenditure required for peptide bond non-planarity. The effect is greater with larger side chains²¹.

The values of Φ_1 and Ψ_1 (Table 5.3.8) approach quite closely those for an anti-parallel pleated sheet, as can be seen by computing 'n', the number of residues per repeat unit; the value of 1.94 is quite close to the sheet value of 2.0, bearing in mind that mean bond lengths and angles were assumed in the calculation. The position given by Φ_1 and Ψ_1 on a Ramachandran plot (as it is usually drawn) is to the right of the n = 2 line and corresponds to a righthand twist (about an axis along the peptide chain) to the sheet.

It has been argued²², with some experimental support²³, from a survey of Ψ , Φ values in lysozyme, carboxypeptidase A, α -chymotrypsin, myoglobin and ribonuclease S, that a sheet with a right hand twist is more stable than either a 'perfect' anti-parallel sheet or one with a left

handed twist. The basis of the argument is as follows:

By definition the stablest state has the lowest free energy G:

G = H - TS

As

$$S = k \log \Omega$$

where Ω is the number of a priori equally probably microstates of the assembly, and assuming that all ϕ, Ψ values within the allowed region are equally likely, then, because the n = 2 line splits the allowed region so that the area to the left is less than the area to the right, peptides with a right hand twist have a higher entropy and, other things being equal, a greater stability than sheets with no twist or a left hand twist.

Since sheets with a right-hand twist are relatively pliable (especially near the centre of the allowed region) they are able to modify their environments locally to achieve the tight packing that 'good' packing in a crystal or protein generally entails.

The hydrogen bonding analysis (Table 5.3.6) shows bonding between molecules related by the two fold axis; a projection down the b axis (Figure 5.3.2) shows the effect of this on the molecular packing; infinite sheets parallel to the a c plane, are formed, being held together along the b axis by Van der Waals forces between the $-CH_3$ and $-CH_2OH$ groups.

The value expected of the inter-sheet distance for an anti-parallel pleated sheet is provided by silk, which can be considered as (Gly-Ser-Gly-Ala-Gly-Ala)_n. The sheets are held together by Van der Waals forces between the side chains; Gly and Ala/Ser alternately (along the sheet normal) pack together so that there are two inter-sheet distances, one for Gly (3.5 Å) and the other for Ala/Ser (5.7 Å)¹¹. The inter-sheet distance perpendicular to the a c plane, at 4.911 Å (b/2), is therefore considerably smaller than 5.7 Å or even, surprisingly, the equivalent N-Ac (L-Ala)₂-OEt distance (4.96 Å).

The values of Φ_2 and Ψ_2 , which are not constrained by hydrogen bonding considerations, have values close to of the γ conformation, (T(C7,C8,N9,H9)=-33) which was indicated by a shoulder in the i.r. (KBr) at 3440 cm⁻¹ (v(N-H)). Overall, Table 5.3.8 shows that substitution of Ser for Ala produces relatively little change in the Φ and Ψ angles.

Table 5.3.9 shows the bond lengths and angles conform quote closely to the weighted average for small peptides⁹. The parameters of the ester group C5-O3 (1.319 Å), C5-O4 (1.195 Å), C8-C5-O4 (125.2^O), and C8-C5-O3 (111.8^O) compared quite well with the expected values of 1.34 Å, 1.20 Å, 125° and 111° respectively¹⁰.

A more detailed examination of the hydrogen bonding, particularly the C15-O16...H9 (174°), C15-O16...N9 (177.2°) and C10-O11...H14 (161°), C10-O11...N14 (167.0°) angles show the amide hydrogens donating to sp lone pairs. As O16 is also an acceptor for H6, the second hydrogen bond would be expected to interact with the lone pair in the 2p orbital implying a C15-O16...H6 angle of 90°. The observed value is 116° with C15-O16...O6 at 120.4° .









Atomic fractional co-ordinates; the e.s.d.'s, in parenthesis, are of the same magnitude as the final digit.

·	x	У	Z
Cl	0.017(1)	0.233(1)	-0.313(1)
C2	0.0331(4)	0.2358(9)	-0.1594(8)
03	0.1007(2)	0.1248(7)	-0.0880(4)
04	0.2040(3)	0.2695(6)	0.0647(5)
C5	0.1815(2)	0.1557(7)	0.0233(4)
06	0.1675(3)	0.0104(8)	0.2699(4)
C7	0.1888(3)	-0.0594(7)	0.1622(5)
C8	0.2392(2)	0.0322(6)	0.0877(4)
N9	0.3331(2)	0.0678(6)	0.1954(3)
C10	0.4023(2)	0.1087(6)	0.1546(3)
011	0.3953(2)	0.1105(6)	0.0245(2)
C12	0.4920(2)	0.1568(7)	0.2820(4)
C13	0.4902(4)	0.3129(8)	0.2925(8)
N14	0.5744(2)	0.1082(7)	0.2573(3)
C15	0.6518(2)	0.0654 (0)	0.3686(4)
016	0.6588(2)	0.0635(7)	0.4991(3)
C17	0.7342(3)	0.0210(9)	0.3283(5)
Н6	0.209(5)	0.025(7)	0.342(8)
H7A	0.226(2)	-0.160(4)	0.202(4)
Н7В	0.127(4)	-0.095(5)	0.095(5)
E 8	0.238(3)	-0.021(5)	0.003(5)
Н9	0.340(2)	0.076(3)	0.280(4)
H12	0.492(3)	0.118(4)	0.369(5)

Table 5.3.1 continued

	x	У	Z
H13A	0.487(4)	0.364(7)	0.182(7)
H13B	0.510(4)	0.356(6)	0.376(7)
H13C	0.445(7)	0.34(1)	0.33(1)
H14	0.568(3)	0.111(4)	0.162(5)
H17A	0.7237(3)	0.0339(9)	0.2125(5)
H17B	0.7542(3)	-0,0830(9)	0.3617(5)
H17C	0.7898(3)	0.0886(9)	0.3971(5)

119

۰.

•

.

Temperature factors. The e.s.d.'s, in parenthesis, are of the same magnitude as the final digit.

	U ₁₁	U ₂₂	U ₃₃	U ₂₃	U ₁₃	U ₁₂
Cl	0.20(1)	0.160(8)	0.088(5)	0.015(5)	0.034(6)	0.059(8)
C2	0.071(3)	0.117(5)	0.103(5)	0.004(4)	-0.003(3)	0.025(3)
03 [·]	0.061(2)	0.082(2)	0.083(2)	-0.011(2)	-0.007(2)	0.000(2)
04	0.084(3)	0.068(2)	0.108(3)	-0.005(2)	0.010(2)	-0.014(2)
C5	0.048(2)	0.067(2)	0.049(2)	-0.005(2)	0.015(2)	-0.013(2)
06	0.057(2)	0.177(4)	0.061(2)	-0.005(2)	0.027(2)	-0.040(2)
C7	0.059(2)	0.082(3)	0.064(2)	-0.002(2)	0.017(2)	-0.037(2)
C8	0.047(2)	0.062(2)	0.043(2)	-0.011(2)	0.020(1)	-0.012(1)
N9	0.041(1)	0.078(2)	0.033(1)	-0.005(1)	0.017(1)	-0.013(1)
C10	0.042(1)	0.059(2)	0.037(2)	-0.002(2)	0.020(1)	-0.007(1)
011	0.054(1)	0.123(3)	0.036(1)	-0.001(1)	0.023(1)	-0.017(2)
C12	0.039(1)	0.076(2)	0.038(2)	0.000(2)	0.021(1)	-0.007(2)
C13	0.061(3)	0.079(3)	0.096(4)	-0.018(3)	0.020(3)	-0.016(2)
N14	0.040(1)	0.101(2)	0.034(1)	-0.001(2)	0.018(1)	-0.003(2)
C15	0.042(2)	0.111(3)	0.038(2)	-0.002(2)	0.014(1)	0.011(2)
016	0.059(2)	0.172(4)	0.036(1)	0.010(2)	0.017(1)	0.031(2)
C17	0.053(2)	0.165(5)	0.055(2)	-0.001(3)	0.024(2)	0.025(3)
H6	0.09(2)					
H7A	0.04(1)					
Н7в	0.07(1)					
н8	0.05(1)					

H9 0.027(7)

.

.

.

120

•

U 1	1
------------	---

H12 0.05(1)

- H13A 0.10(2)
- H13B 0.08(1)
- H13C 0.13(3)
- H14 0.05(1)
- H17A 0.17(3)
- H17B 0.24(5)
- H17C 0.16(4)

Bond lengths (in Angstroms) - e.s.d.'s, in parenthesis, are of the same magnitude as the final digit.

C1-C2	1.42(1)	C10-C12	1.520(5)
C2-03	1.476(7)	C12-H12	0.92(4)
03 - C5	1.322(5)	C12-C13	1.538(8)
04–C5	1.192(6)	C13-H13A	1.16(7)
06-н6	0.75(7)	C13-H13B	0.97(6)
06-C7	1.390(6)	C13-H13C	0.9(1)
С7-Н7А	1.13(4)	N14-C12	1.451(4)
С7-Н7В	0.99(5)	N14-H14	0.90(4)
C7-C8	1.536(5)	N14-C15	1.323(4)
С8-н8	0.96(5)	C15-016	1.229(4)
C8-C5	1.486(6)	C15-C17	1.521(5)
C8-N9	1.451(4)	Cl7-Hl7A	1.08(0)
N9-H9	0.79(3)	С17-Н17В	1.08(0)
N9-C10	1.327(4)	C17-H17C	1.08(0)
C10-011	1,226(4)		

Bond angles (degrees) - e.s.d.'s, in parenthesis, are of the same magnitude as the final digit.

C1-C2-O3	106.4(7)	С10-С12-Н12	107(3)
C2-03-C5	118.3(4)	C10-C12-C13	109.3(4)
03 - C5-04	123.1(4)	C10-C12-N14	109.2(3)
03-C5-C8	111.6(3)	H12-C12-C13	110(3)
04-C5-C8	125.3(3)	H12-C12-N14	109 (2)
H6-06-C7	116(5)	C13-C12-N14	112.2(3)
06-С7-Н7А	114(2)	С12-С13-Н1-ЗА	111 (3)
06-С7-Н7В	103(3)	С12-С13-Н13В	118(3)
06-C7-C8	111.8(4)	С12-С13-Н13С	113(6)
Н7А-С7-Н7В	98(3)	Н13А-С13-Н13В	109 (5)
С8-С7-Н7А	,114 (2)	H13A-C13-H13C	115(7)
С8-С7-Н7В	115(3)	н13в-с13-н13с	89(7)
С5-С8-Н8	106(3)	C12-N14-H14	114(3)
C5-C8-C7	110.7(3)	C12-N14-C15	121.8(3)
C5-C8-N9	111.4(3)	C15-N14-H14	124(2)
С7-С8-Н8	104(3)	N14-C15-O16	122.5(3)
C7-C8-N9	109.8(3)	N14-C15-C17	116.6(3)
N9-C8-H8	116(3)	016-C15-C17	120.9(3)
C8-N9-H9	118(2)	С15-С17-Н17А	115.6(2)
C8-N9-C10	122.5(3)	С15-С17-Н17В	111.1(3)
С10-N9-Н9	118(2)	С15-С17-Н17С	101.3(3)
N9-C10-O11	123.5(3)	Н17А-С17-Н17 В	109.5(0)
N9-C10-C12	115.0(2)	H17A-C17-H17C	109.5(0)
011-C10-C12	121.5(3)	Н178-С17-Н17С	109.5(0)

Deviations (in Angstroms) from the least squares planes.

Asterisked atoms are not used in the least squares calculation.

			 	· · · · · · · ·	
	1	2	3	4	
c17	0.002				
016	0.003		. <i>.</i> .		
C15	-0.008				
N14	0.003				
C12	0.038*	0.002		· ·	
011	· ·	0.002	-0.011		
C10		-0.006	0.022		
		0.002	-0.020		
C8		0.139*	0.009	0.000	
C5				-0.001	
04	·			0.000	
03				0.000	
C2			· · ·	0.064*	
Cl				-1.087*	

Table 5.3.5 continued

	a _.	b _i	° _i	ďi
1	0.3529	0.9356	-0.0047	3.8486
2	-0.3625	0.9313	0.0366	-1.0842
3	-0.2971	0.9527	0.0649	-0.7439
4	-0.4442	-0.0888	0.8915	-2.0329

Equations of the least squares planes

Hydrogen bonding

Donor	Acceptor	Angle	Distance	Distance	Position of
Х-Н	¥	хну ^о	Н-Ү	х-х	Acceptor
N9-H9	016	167	2.1	2.94	(1-x,y,1-z)
06-н6	016	165	2.0	2.78	(1-x,y,1-z)
N14-H14	011	159	2.1	2.95	(1-x,y,-z)

Intra-molecular contact distances (< Van der Waals + 0.35Å)

c204	2.690	011C8	2.793
C5C1	3.342	OllHl4	2.44
c506	2.866	011N14	2.789
c5c10	3.141	С12Н9	2.44
c703	2.905	C12016	2.755
С789	2.53	H12H9	2.18
H7AH6	2.34	н12н13С	2.33
H803	2.40	H12C15	2.49
E8E7B	2.32	н12016	2.42
N904	2.735	C13C15	3.337
N906	2.944	H14C10	2.50
N9H12	2.41		

Conformational	angles	(in degree	s)	
	¢ 1	¢2	Ψı	Ψ2
N-Ac-L-Ala-L-Ser-OEt	-142	-78	140	172
N-Ac-L-Ala-L-Ala-OEt	-145	-74	138	166

Significant	bond	lengths	and	angles	(in	Angstroms	and	degrees
respectively).								

	First Amide	Second Amide	Expected Value
C–C (O)	1.521	1.520	1.51
C=0	1.229	1.226	1.24
C-N	1.323	1.327	1.32
N-C ^{°°}	1.451	1.451	1.45
c ^{°°} -c ^β	1.538	1.536	1.53
c−ĉ (0) –0	120.9	121.5	120.5
C-Ĉ (0) -N	116.6	115.0	116
O−Ĉ−N	122.5	123.5	123.5
C−Ñ−C ີ	121.8	122.5	122
N−Ĉ [∞] −C (O)	109.2	111.4	111

5.4 Structure Solution Strategies

The problems raised by the initial failure to solve the structure of N-Ac-(L-Ala)₂OEt provoked the study of the constitution of a good starting set.

A starting set may be selected very satisfactorily by a convergence map, which isolates those reflexions that interact most strongly with the data set. However, it can happen, when these reflexions are used, that subsets of reflexions, poorly linked with the rest of the data set, are not satisfactorily phased.

To obviate that possibility, representative reflexions from the subsets are required in the starting set, i.e. what might be called core reflexions, that interact minimally with the main data set, while being well indicated by some subset. The latter requirement is produced by a high $\langle \alpha^2 \rangle$, while to satisfy the second involves looking for reflexions which are badly determined by Σ_2 relationships with some strong reflexions.

The obvious way is to search through Ψ_{o} reflexions (defined to be the strong reflexions of a triplet with two strong and one weak reflexion) looking for reflexions that are present in a number of relationships. However, it is by no means clear, once a list of such reflexions has been prepared, what criterion to use for ordering, whether number of Ψ_{o} interactions, or the magnitudes of the strong reflexions, or a mixture of both. For this reason, it was decided to look at a restricted group of reflexions $E_{-h-k-\ell}$ which have a Ψ_{o} interaction with three other reflexions E_{h} , E_{k} , E_{ℓ} (5.4.1 and 5.4.2). This will be recognised as the definition of a negative quartet (N.Q.)²⁴.

5.4.1 $|E|_{h}, |E|_{k}, |E|_{\ell}, |E|_{-h-k-\ell} > EMIN$ 5.4.2 $|E|_{h+k}, |E|_{h+\ell}, |E|_{k+\ell} < EMAX$

The other three reflexions naturally have a similar relationship.

N.Q.s are of course used in the NQEST test precisely because the phases are largely independent of the phase expansion process, i.e. they are not strongly linked by Σ_2 . The use of N.Q. reflexions in the starting set would be expected to remove some of this independence as has been found when negative quartet relationships are used in the phase expansion (NQEST becomes a less discriminating F.O.M.)²⁵.

The criterion used for ordering reflexions is the strength of the indication of the phase by N.Q. relationships. For one relationship it is proportional to

5.4.3 B =
$$[2/N]|E|_{h}|E|_{k}|E|_{\ell}|E|_{m}(Y + \Sigma(|E|_{cross}^{2} - 1))$$

where Y and n are usually 1 and 3 respectively and h + k + l + m = C. For multiple indications of phase, by a similar derivation to $\langle \alpha^2 \rangle$ (see Appendix B), the appropriate measure is

$$\langle \beta^2 \rangle = \Sigma B^2 + 2 \Sigma \Sigma B_i B_j (I_1(B_i) I_1(B_j)) / \langle I_0(B_i) I_0(B_j) \rangle$$

$$i j i$$

For a centrosymmetric structure²⁶, the appropriate measure is ΣB^2 .

The foregoing analysis rationalizes the use, in the starting set, of reflexions that occur frequently in the NQEST test, a procedure found to be of use²⁷. It would be expected that the number of N.Q.s would vary with the fragmentation of the data set, and it has indeed been noted that the number is highly structure dependent²⁴.

Program Description

Data are converted from SHELX LIST 3 format to H, K, L, |F|, ID; ID is zero if $|F| > 3 \sigma$ (|F|) otherwise one. As a necessary condition for a reflexion to be designated as strong is an ID of zero, the cutoff point can alter the reflexions chosen.

The reflexions with high $\langle \beta^2 \rangle$ are isolated by using the SIGMA2 subroutine of MULTAN 78 to provide the reflexions of the Ψ_{α} test; be-

cause symmetry is being ignored, to prevent quartet duplication²⁸, symmetry-invariant weak reflexions are ignored; in the case of C2 they are reflexions of the form O k O. For a quartet to be retained all reflexions must be distinct. If the number of quartets is greater than 250, the quartets with the highest B values are retained. The latter operations are performed by an appended segment (KWICORE) that isolates and processes the quartets.

Possible Development

Symmetry effects could be incorporated and quartets having systematically absent cross-reflexions (not lattice extinctions)²⁹.

When B > O the quartet is probably O while when B < O it is π . Thus when Y $\gtrsim \Sigma$ (1 - $|E|^2_{cross}$) the quartet is somewhere between O and π and hence potentially an enantiomorphically sensitive invariant. As lack of enantiomorphic definition is often a problem these quartets could be of value in choosing the starting set or selecting the 'best' (which is not to say correct) solution by a F.O.M. of the type defined below:

 $\Sigma W \cos (\Phi_{h} + \Phi_{k} + \Phi_{\ell} + \Phi_{m}) / \Sigma W$

A good solution would have a value near zero. Possible weights are $\langle \alpha_h \rangle \langle \alpha_k \rangle \langle \alpha_n \rangle$.

Utlilisation

In order to assess the effectiveness of the procedure, a comparison with a run of MULTAN 78, where the starting set was chosen automatically, was undertaken.

The starting set chosen, from which the structure solution could not be isolated, is shown below.

				E
Origin	11	1	8	2.892
origin	12	ο	ī	2.569
Enantiomorph	6	2	10	3.489
	6	ο	10	4.015
	12	2	ī	1.902
	·- 5	5	1	1.811
·	13	1	ī	2.659
	1	1	$\overline{2}$	1.712
	2	о	2	2.389

The number of strong reflexions was 300 ($|E| \le 1.235$) The number of Σ_2 relationships was 3500.

The number of reflexions per non-hydrogen atom (17.6) and Σ_2 relationships per reflexion (11.7) were greatly in excess of the values usually required (\sim 10 and \sim 8 respectively). Examination of the starting set shows eight out of the nine reflexions to be of the form h k $\overline{\ell}$, leaving the other half of the data set under-represented. Possibly by inputting molecular geometry and changing the balance of the starting set the structure could have been solved, but instead a starting set was selected from reflexions with high $\langle \beta^2 \rangle$ (and reasonably high $\langle \alpha^2 \rangle$).

These were isolated with $|E|_{max} \leq 0.7$ (388 reflexions, one of which was a symmetry-invariant). The number of strong reflexions (130) involved in Ψ_{o} relationships was adjusted to maximise the number of relationships while not producing so many that they needed to be sorted. The N.Q.s isolated had a $\langle \cos(\Phi_{h} + \Phi_{k} + \Phi_{l} + \Phi_{m}) \rangle$ of -0.354 and a B weighted $\langle \cos(\Phi_{h} + \Phi_{k} + \Phi_{l} + \Phi_{m}) \rangle$ (i.e. NQEST) of 0.399 based on the refined structure.

The starting set was as follows:

				Ē
Origin	7 ۲	1	10	2.618
	۲ 6	2	9	2.537
	10	2	7	2.496
	9	5	1	2.092
	4	0	2	2.492
	4	0	3	2.592
	3	1	7	2.961
Enantiomorph	13	1	ī	2.659
	1	1	ž	1.712

The final two reflexions were left for the program to choose to ensure there would be no break in the phase expansion generated by CONVERGE. The phase expansion, using the same number of strong reflexions and Σ_2 relationships as in the previous run, produced an E-map from which the structure could be found.

REFERENCES

- 1. C. Ramakrishnan and G. N. Ramachandran, Biophys. J., 5, 909 (1965).
- J. A. Schellman and S. Schellman, 'The Proteins' II, Academic Press, N.Y. (1964).
- 3. See Chapter 2 and Appendix B.
- T. Drakenberg, K. I. Dahlquist and S. Forsen, J. Phys. Chem., <u>76</u>, 2178 (1972).
- 5. G. N. Ramachandran and A. K. Mitra, J. Mol. Biol., 107, 85 (1976).
- 6. S. S. Zimmerman and H. A. Scheraga, Macromolecules, 9, 408 (1976).
- R. E. Dickerson and I. Geis, 'The Structure and Action of Proteins', W. A. Benjamin (1969).
- 8. A. W. Burgess and H. A. Scheraga, Biopolymers, 12, 2177 (1973).
- 9. R. E. Marsh and J. Donohue, Adv. Protein Chem., 22, 235 (1967).
- 10. P. W. Borthwick, Acta Cryst., B36, 628 (1980).
- 11. R. E. Marsh, R. B. Corey and L. Pauling, Biochem. Biophys. Acta, <u>16</u>, 1 (1955).
- 12. T. Miyazawa, J. Polymer Sci., 55, 215 (1961).
- 13. S. Arnott, S. D. Dover and A. Elliot, J. Mol. Biol., <u>30</u>, 201 (1967).
- 14. Y. Tokuma, T. Ashida and M. Kakudo, Acta Cryst., B25, 1367 (1969).
- R. J. Fletterick, C.-C. Tsai and R. E. Hughes, J. Phys. Chem., <u>75</u>, 918 (1971).
- 16. J. K. Fawcett and N. Camerman, Acta Cryst., B31, 658 (1975).

17. L. Pauling and R. B. Corey, Proc. Roy. Soc. London B, <u>141</u>, 21 (1953).

18. P. R. Andrews, Biopolymers, 10, 2253 (1971).

19. A. S. N. Murthy, K. G. Roa and C. N. R. Roa, J.A.C.S., <u>92</u>, 935 (1970).

- 20. M. Levitt and S. Lifson, J. Mol. Biol., <u>46</u>, 269 (1969).
- 21. W. A. Hiltner and A. J. Hopfinger, Biopolymers, 12, 1197 (1973).
- 22. C. Chothia, J. Mol. Biol., 75, 295 (1973).
- 23. F. M. Pohl, Nature New Biol., 234, 277 (1971).
- 24. G. T. De Titta, J. W. Edwards, D. A. Langs and H. Hauptman, Acta Cryst., A31, 492 (1975).

- 25. J. Gilmore, Acta Cryst., <u>A33</u>, 712 (1977).
- 26. H. Hauptman and E. A. Green, Acta Cryst., A32, 45 (1976).
- 27. J. V. Silverton, C. Kabuto and T. Akiyama, Acta Cryst.; B34, 588 (1978).
- 28. C. Giacovazzo, Acta Cryst., <u>A32</u>, 958 (1976).
- 29. B. Busetta, C. Giacovazzo, M. C. Burla, A. Nunzi,
 G. Polidori and D. Viterbo, Acta Cryst., <u>A36</u>, 68 (1980).

٠.
CHAPTER SIX

N-ACETYL DEHYDRO-PHENYLALANINE

.

AND

.

.

N-ACETYL DEHYDRO-PHENYLALANYL-L-PROLINE

-

6.1 Introduction

While dehydro-alanyl residues are constituents of a wide range of natural products, dehydro-phenylalanyl residues have, so far, been found in only two: albonoursin¹ and tentoxin². They are, remarkably, both cyclic peptides (dipeptide and tetrapeptide respectively).

The first compound, albonoursin, has been isolated from a number of organisms, where its function is not clear, but is possibly related to its sometimes present antibiotic properties; it also displays anti-tumour activity (non-dose dependent)³. Although all the geometrical isomers have been synthesised⁴, no structure-activity correlations have been reported. The isolation of a number of similar compounds (cyclic dipeptides with dehydro-tryptophan residues) has been reported, but with no details of biological activity⁵.

Tentoxin is a phytotoxin isolated from the fungus Alternaria tenuis, which, when applied to germinating seedlings, causes chlorosis; the effect is species specific and is associated with a tentoxin binding site on CF_1 (chloroplast coupling factor 1, a vital protein in A.T.P. synthesis) in sensitive species⁶.

Chemically it would be expected that dehydro-phenylalanyl derivatives would be easier to isolate than dehydro-alanyl derivatives since tantomerism with the imino form is restricted by the conjugation of the olefin bond with the phenyl ring. Indeed even derivatives with free amino groups have been synthesised⁷. In fact, β substitution of dehydroalamine generally appears to increase stability to electrophiles⁸ and similar derivatives with a range of β substituents have been synthesised⁹.

The presence of a dehydro-phenylalamine residue has been found to confer resistance to enzymolysis by both thermolysin, an endopeptidase with specificity for the amino sites of hydrophobic amino acid residues

and chymotrypsin, an endopeptidase specific for the carboxyl sites of aromatic amino acfd residues. Thus the substitution of a dehydro-phenylalanyl residue for a phenylal ine residue in an encephalin analogue (Tyr-D-Ala-Gly-Phe-Met-NH₂) produced resistance to in vivo enzymolysis as well as the bonus of increased potency $(x \ 5)^{10}$.

Further details of the chemistry can be found in Chapter 2.

6.2 N-Acetyl Dehydro-phenylalanine

Experimental

The compound was prepared by hydrolysis of 2-methyl 4-benzylidene oxazolin-5-one. Chunky prisms were grown by cooling slowly an aqueous solution.

The measured density and a molecular weight calculated on the basis of $C_{11}H_{11}NO_3$ gave Z = 4.69. Drying in vacuo over silica gel indicated eight water molecules per unit cell, confirmed later by the structural determination.

Formula = $C_{11}H_{11}NO_3 \cdot 2H_2O$ Space group = $P2_1/a$ F(000) = 512 a = 18.281 Å b = 6.080 Å c = 11.401 Å β = 105.94° Z = 4 V = 1218.75 Å³ D_0 = 1.31 Mg m⁻³ μ = 0.066 mm⁻¹

Data were collected on a CAD4 diffractometer by the $\theta/2\theta$ mode, using Mo-K_ radiation monochromated with graphite.

Of the 2589 reflexions measured (100 intensity control measurements, $1.5^{\circ} < \theta < 27^{\circ}$) 2361 were systematically present, of which 2154 were unique (1297, 60.2% $|F| > 3\sigma(|F|)$).

The structure was solved using the EEES facility of SHELX. The first E-map revealed 15 atoms (R = 41%, $R_G = 40.5$ %, $R_M = 40$ %). At R = 33%, $R_M = R_G = 37$ % a difference map revealed the two oxygens.

Refinement continued until R = 12%, $= R_G = R_M$ after which anisotropic refinement commenced. At R = 9% $= R_M = R_G$ a difference map revealed the hydrogens and refinement continued until:

$$R R_{M} = R_{G}$$
all |F| 10.5 8.6
$$|F| > 2\sigma(|F|) 5.7 5.7$$

Following weighted refinement

	R	$R_M = R_G$	RW	s ²	g
all F	10.6	7.0	6.6	0.0027	0.0015
$ \mathbf{F} > 2\sigma(\mathbf{F})$	5.8	6.8	5.7	1.65	0.0015

The high percentage of weak reflexions probably explains the high R factor with all reflexions. The final atomic parameters are listed in Table 6.2.1 and Table 6.2.2. $10|F_0|/10|F_c|$ tables are in Appendix E (Table E.5).

Discussion

The first impression, confirmed by least squares plane analysis (Table 6.2.5, plane 5), was of a plane from which the amide group skews.

Examination of the intra-molecular non-bonded distances (Table 6.2.7) reveals severe steric crowding between the amide oxygen (O3) and C5, amide nitrogen (N4) and 'hydroxyl' oxygen (O15), and the carboxylic acid carbon (C13) and the carbonyl carbon; these make it clear why the amide group isn't in the plane. Possibly an auxillary factor is to facilitate hydrogen bonding (by H1W1) to O3 and H4 to O14.

An analysis of the hydrogen bonding (6.2.6) shows how extensive it is, with all possible hydrogen bond donors donating. Hydrogen bonds involving the water molecules are clearly shorter (average value 2.72 Å) than the amide hydrogen bond (2.98 Å) as is usually found.

Despite the hydrogen positions being uncertain it is quite instructive to examine the water molecules in detail. The H-O-H values at 103° (W1) and 108° (W2), as expected, are less than $\cos^{-1}(-1/3)$.

The angles involving donating hydrogen atoms for the first water molecule (W1), H2W1...H1W2(120[°]), H2W1-W1...H2W2(110[°]), H1W1-W1...H1W2 (117[°]) and H1W1-W1...H2W2(114[°]) show the hydrogens are roughly tetrahedrally arrayed around the oxygen. Although the interaction of the donated hydrogens might be expected to alter the repulsion between lone pairs, it is still surprising that H2W2...W1...H1W2(93[°]) is less than \cos^{-1} (-1/3). The relevant angles for the second water molecule (W2), H2W2-W2...H15(128[°]) and H1W2-W2...H15(123[°]) show a similar tendency.

An equivalent analysis of the hydrogen bonding to the other donors shows some interesting features. C2-O3...HlWl(124[°]), C2-O3...Wl(125[°]) and, less convincingly C13-O14...H4(143[°]), C13-O14...N4(140.0[°]) indicate sp² hybridisation of the carbonyl oxygens. The H15-O15...Wl(91.7[°])

H15-O15...H2W1(92°), C13-O15...W1(153.1°), C13-O15...H2W1(153°) and H15-O15-C13(115°) angles indicate that H2W1 is in the plane of H15-O15-C13 as the sum of the first, third and fifth angles is 359.8° . Despite the non-ideal values, the oxygen (O15) must be presumed to be sp² hybridised.

Figure 6.2.2, a projection down the b axis, shows the molecular packing. It can be rationalised as molecular sheets, parallel to the ab plane, held together by chains of water molecules parallel to the b axis.

The phenyl group is planar (Table 6.2.5, plane 5) with an average bond length of 1.384 Å. The bond length, after a rigid body librational correction, despite the reasonably low value of $R_{G}(11.3\%)$, at 1.387 Å is considerably below the spectroscopic value for benzene (1.395 Å).

The average ring angle is 120.0° , with the angle centred on the ipso atom, i.e. the atom bonded to the substituent, C8-C7-C12 at 118.6°. The values, for a series of cinnamic acids¹¹, of the equivalent angle at 118.4° , 119.0° and 118.8° are similarly below the average value. The effect of substituents on the bond lengths and angles of the phenyl ring has been an area of active study. Most cases are adequately covered by a valence bond interpretation¹² though several other explanations exist.

If the substituent (x) is a pure σ acceptor, i.e. more electro negative than ring atoms, then the p character of the hybrid orbital donating the x is increased (as the p electron is more loosely held), producing a concomitant increase in s character for the other two hybrid orbitals. This results in a decrease in the bond length of the bonds from the ipso atom to adjacent ring atoms and an increase in the ring angle (α) centred on the ipso atom. For a σ donor the changes are in the opposite directions.

For a π donor/acceptor, by equating a presumed decrease in the

 $c_{ipso} - x$ distance with an increase in s character of the hybrid orbital a similar situation to that for a σ donor is produced resulting in similar effects¹³. Naturally the effect would be expected to be conditioned by the rotation about $c_{ipso} - x$, but although a trend is evident it isn't a rigid relationship.

The dihedral angles corresponding to Ψ and Φ in peptides are $\Phi'(C2,N4,C5,C13) = -74^{\circ}$ and $\Psi'(N4,C5,C13,O15) = -11^{\circ}$.

Table 6.2.8 shows significant bond lengths and angles compared with those in peptides¹⁴. The most obvious difference in the bond lengths is in the N-C^{α}, C-N distances, being significantly shorter and longer respectively than the corresponding peptide distance; this reflects the conjugation of the amide system with the olefinic system, even though this is reduced by the rotation about N-C(τ (C2,N4,C5,C6) = 106^o); the topic is discussed further in 6.3.

The bond angles broadly agree, with the exception of $N-C^{-}-C(O)$ which is anomalous even when compared with other dehydro-amino acids (Table 6.2.8). The intra-molecular non-bonded distances show severe N4...O15 clash, which is relieved by expansion of this angle.

The carboxylic acid group is planar, with the acid hydrogen (H15) deviating from the plane by only 0.04 Å (with an approximate e.s.d. of 0.04 Å), despite being a donor to W2. The bond formed is in fact exceptionally strong, the hydrogen being well directed towards the acceptor (O15-H15...W2 = 174°). As is usually found H15-O15 is synplanar to C13-O14. A comparison of the carboxylic acid group parameters C13-O14(1.207 Å), O15-C13-O14(123.1°), C13-O15(1.319 Å), with the expected values 1.21 Å, 123°, 1.31 Å showing good agreement¹⁵.

The disposition of the phenyl group to the carbonyl group (C13-O14) about the double bond is trans, while the double bond is

symplanar to the carbonyl group. The latter situation contrasts with the case of N-Ac- Δ Ala-OH (see 4.2) and suggests the use of steric clash between the carboxyl oxygens and the β substituents (and the bend bond model for ambiguous cases) as a criterion is limited. The molecule in fact betrays its origins, as the parent oxazolin-5-one ring is seen to be formed by O3 attack on Cl3 (Figure 6.2.1).



H1U2

1242

Figure 6.2.1. ORTEP drawing of the molecular structure of N-Ac Δ Phe-OH.2H₂O with 50% probability ellipsoids.



Figure 6.2.2. Projection down the b axis showing the molecular packing.

Atomic fractional co-ordinates with the e.s.d.'s in parenthesis and of the same magnitude as the final digit.

	x	У	Z
Cl	0.3790(3)	-0.2956(5)	0.4545(2)
C2	0.3162(2)	-0.0796(4)	0.4515(1)
03	0.3515(2)	0.0884(3)	0.4266(1)
N4	0.2186(2)	-0.0742(3)	0.4793(1)
C5	0.1524(2)	0.1234(3)	0.4794(1)
C6	0.0427(2)	0.1661(4)	0.4326(1)
C7	-0.0360(2)	0.0395(4)	0.3694(1)
C8	-0.1430(2)	0.1444(5)	0.3282(1)
C9	-0.2192(3)	0.0455(6)	0.2644(2)
C10	-0.1901 (3)	-0.1598(6)	0.2420(2)
C11	-0.0868(3)	-0.2659(5)	0.2833(2)
C12	-0.0103(3)	-0.1681(4)	0.3468(2)
C13	0.2101(2)	0.2952(4)	0.5361(1)
014	0.1725(1)	0.4813(3)	0.5342(1)
015	0.3073(2)	0.2248(3)	0.5881(1)
HIA	0.324 (4)	-0.407(8)	0.445(3)
HIB	0.427(3)	-0.301(6)	0.423(2)
ніс	0.407(4)	-0.346(7)	0.513(3)
H 4	0.196(2)	-0.201(5)	0.496(2)
H6	0.007(2)	0.309(4 <u>)</u>	0.438(1)
E8	-0.160(2)	0.292(4)	0.343(1)
Н9	-0.291(3)	0.142(5)	0.239(2)
HIO	-0.245(3)	-0.213(5)	0.199(2)
H11	-0.068(3)	-0.410(6)	0.267(2)

	x	У	Z
H12	0.056(2)	-0.249(4)	0.375(1)
H15	0.341(3)	0.320(7)	0.623(2)
HIWI	0.462(3)	0.094(5)	0.377(2)
82W1	0.571(3)	0.003(6)	0.365(2)
H1W2	0.439(3)	-0.126(7)	0.179(2)
H2W2	0.435(3)	0.020(6)	0.235(2)

Temperature factores - e.s.d.'s, in parenthesis, are of the same magnitude as the final digit.

	U ₁₁	U ₂₂	U ₃₃	U ₂₃	U ₁₃	U ₁₂
C1	0.057(2)	0.050(2)	0.069(2)	0.002(1)	0.022(2)	0.018(1)
C2	0.037(1)	0.042(1)	0.040(1)	0.002(1)	0.008(1)	0.006(1)
03	0.050(1)	0.054(1)	0.070(1)	0.015(1)	0.027(1)	0.010(1)
N4	0.040(1)	0.028(1)	0.052(1)	0.003(1)	0.015(1)	0.002(1)
C5	0.035(1)	0.033(1)	0.040(1)	0.005(1)	0.014(1)	0.003(1)
C6	0.038(1)	0.037(1)	0.041(1)	0.001(1)	0.014(1)	0.002(1)
C7	0.040(1)	0.046(1)	0.039(1)	0.000(1)	0.014(1)	-0.003(1)
C8	0.045(1)	0.058(2)	0.048(2)	-0.003(1)	0.010(1)	-0.001(1)
C9	0.053(2)	0.096(3)	0.053(2)	-0.007(2)	0.000(1)	-0.004(2)
C10	0.070(2)	0.090(3)	0.053(2)	-0.021(2)	0.010(2)	-0.028(2)
C11	0.080(2)	0.059(2)	0.059(2)	-0.019(2)	0.026(2)	-0.016(2)
C12	0.055(2)	0.048(2)	0.056(2)	-0.005(1)	0.016(1)	-0.005(1)
C13 ·	0.034(1)	0.036(1)	0.042(1)	0.006(1)	0.013(1)	0.003(1)
014	0.045(1)	0.033(1)	0.062(1)	-0.003(1)	. 0.007(1)	0.005(1)
015	0.047(1)	0.047(1)	0.048(1)	-0.001(1)	-0.004(1)	0.007(1)
Wl	0.052(1)	0.057(1)	0.052(1)	0.012(1)	0.016(1)	0.013(1)
W2	0.126(2)	0.071(2)	0.055(1)	-0.007(1)	-0.020(1)	0.038(2)
HLA	0.14(2)					
HlB	0.11(1)					. `
HIC	0.13(1)					
H4	0.07(1)					
н6	0.03(1)					
н8	0.05(1)					

Table 6.2.2 continued

U11	
-----	--

Н9	ο.	09	(1)
----	----	----	-----

	H10	0.08(1)
--	-----	---------

H11 0.08(1)

H12 0.06(1)

H15 0.11(1)

H1W1 0.08(1)

H2W1 0.09(1)

H1W2 0.10(1)

H2W2 0.09(1)

Bond lengths (in Angstroms) - e.s.d.'s, in parenthesis, are of the same magnitude as the final digit.

W1-HIW1	0.95(4)	С6-н6	0.97(2)
W1-H2W1	0.88(4)	C7-C8	1.400(3)
W2-H1W2	0.87(4)	C7-C12	1.385(4)
W2-H2W2	0.81(4)	C8-C9	1.385(4)
Cl-HlA	0.91(5)	С8-Н8	0.97(3)
C1-H1B	0.89(4)	C9-C10	1,382(5)
C1-H1C	1.07(5)	С9-Н9	1.01(3)
C1-C2	1.490(4)	C10-C11	1.373(4)
c2-03	1.230(3)	С10-н10	0.92(3)
C2-N4	1.345(3)	C11-C12	1.380(4)
N4-C5	1.419(3)	С11-Н11	0.97(3)
N4-H4	0.89(3)	C12-H12	0.93(3)
C5-C6	1.332(3)	C13-014	1.207(3)
C5-C13	1.491(3)	C13-015	1.319(3)
C6–C7	1.470(3)	015-H15	0.86(4)

Bond angles (degrees) - e.s.d.'s, in parenthesis, are of the same magnitude as the final digit.

H2W1-W1-H1W1	103 (3)
H2W2-W2-H1W2	108(3)
HIA-CI-HIB	111(4)
HIA-C1-HIC	88(3)
HIA-CI-C2 .	111(3)
H1B-C1-H1C	124(3)
H1B-C1-C2	113(3)
H1C-C1-C2	108(2)
03-C2-C1	122.9(2)
03-C2-N4	120.5(2)
C1-C2-N4	116.5(2)
C2-N4-C5	121.1(2)
C2-N4-E4	117(2)
C5-N4-H4	122(2)
N4-C5-C6	124.8(2)
N4-C5-C13	117.4(2)
C6-C5-C13	117.8(2)
C5-C6-C7	131.7(2)
С5-С6-Н6	117(1)
С7-С6-Н6	111(1)
C6-C7-C8	116.0(2)
C6-C7-C12	125.4(2)
C8-C7-C12	118.6(2)
c7-c8-c9	120.6(3)
C7-C8-H8	119(1)

С9-С8-н8	120(1)
C8-C9-C10	119.7(3)
С8-С9-Н9	112(2)
С10-С9-Н9	128(2)
C9-C10-C11	120.1(3)
C9-C10-H10	114(2)
С11-С10-Н10	126(2)
C10-C11-C12	120.5(3)
C10-C11-H11	119(2)
C12-C11-H11	121 (2)
C11-C12-C7	120.5(3)
С11-С12-Н12	118(2)
C7-C12-H12	122(2)
C5-C13-014	123.5(2)
C5-C13-015	113.4(2)
014-C13-015	123.1(2)
с13-015-н15	115(2)

Deviations (in Angstroms) from the least squares plane. Asterisked atoms are not included in the least squares calculation.

- •			•••••			
Plane	1	2	3	4	<u></u> ,5	6
Cl	0.002		1.117*			
C2	-0.008		1.217*			
03	0.003		2.263*			·
N4	0.002		0.085*			. <u> </u>
C5		-0.008	0.068			-0.091
C6		-0.019	-0.014	0.040		0.071
C7		0.015	-0.050	-0.022	0.013	<u></u>
C8			-0.032	-0.030	-0.009	
C9			0.028	0.005	-0.002	
C10			0.052	0.029	0.008	
- c11	-		0.002	0.005	-0.003	
C12			-0.054	-0.026	-0.007	•
C13	 ,	0.012	0.148*			-0.012
014		 	0.352*			-0.019
015		<u>_</u>	-0.011*	•		0.052

Table 6.2.5 continued

	a _i	b	° _i	ď
i				
1	0.5891	0.2174	0.7783	7.5979
2	0.4381	0.4096	-0.8002	5.5841
3	0.4850	0.4136	-0.7705	5.3325
4	0.5030	0.4134	-0.7590	5.2890
5	0.5147	0.4229	-0.7458	5.2356
6	-0.5333	-0.2603	0.8049	5.6749

156

.

.

Hydrogen bond lengths $(^{A})$ and angles $(^{O})$

Donor X-H	Acceptor Y	Angle XHY	Distance EY	Distance XY	Position of Acceptor
015-н15	w2	174	1.7	2.53	(x, y, z)
W2-H2W2	Wl	164	2.0	2.81	(x-1, 1,-y, 1+z)
W2-H1W2	Wl	177	1.8	2.71	(l-x, l-y, l-z)
W1-H2W1	015	178	2.0	2.86	(1-x, -y, 1-z)
W1-H1W1	03	174	1.7	2.69	(x, y, z)
N4-H4	014	170	2.1	2.98	(x, y-1, z)

Intra-molecular non-bonded distances (< Van der Waals + 0.35Å)

03 H1B	2.53
03 C5	2.708
N4 ElA	2.52
N4 C7	3.120
N4 H12	2.50
N4 015	2,680
H4 Cl	2.48
H4 H1A	2.30
H4 H12	2.36
Сб Н8	2.56
C6 014	2.794
H6 H8	2.20
H6 C8	2.46
H6 C13	2.51
H6 014	2.44
C12 N4	3.088
C12 C5	3.161
c13 c2	3.174

Bond lengths $(\stackrel{0}{A})$ and angles $(\stackrel{0}{})$ compared with those of peptides

	Peptide	N-Ac-∆Phe-OH
C ^{°°} −C (O)	1.51	1.490
C=0	1.24	1.230
C-N	1.32	1.345
N-C ^{°°}	1.45	1.419
c-ĉ (o) -o	120.5	122.9
C-Ĉ (O) -N	116	116.5
O-Ĉ-N	123.5	120.5
C−Ñ−C	122	121.1
N-Ĉ [°] -C	111	117.4

Experimental

The compound was prepared by Bergmanns method and recrystallised a number of times from methanol/water¹⁶.

Formula = $C_{16}H_{18}N_2O_4 \cdot \frac{1}{2}H_2O$ Space group = C2 F(OOO) = 636 a = 18.516 Å b = 9.515 Å c = 10.538 Å β = 120.102^O V = 1606.3 Å³ D_O = 1.28 Mg m⁻³ (CCl₄/n-C₆H₁₄) D_C = 1.29 Mg m⁻³ Z = 4 μ = 0.055 mm⁻¹

The cell parameters were determined on the diffractometer from high angle reflexions.

Data were collected using graphite monochromated Mo-K_a radiation $(1.5^{\circ} < \theta < 30^{\circ})$; 2667 reflexions (82 intensity controls) were measured 2476 being unique (82.3% (2037) $|F| > 3\sigma(|F|)$). Because of previous difficulty in solving symmorphic space groups¹⁷ using SHELX, it was decided to use MULTAN, which allowed the weights of the F.O.M. to be altered. ABSFO.M, Ψ_{o} and R_{k} were given weights of 0.7, 1.3, and 1.0 respectively. The geometry of the phenyl group was supplied for calculation of a molecular scattering factor to aid in scaling¹⁸.

The phase set with the highest F.O.M. and incidentally, the lowest Ψ_{o} , produced an E-map with eighteen atoms of the structure. Subsequent calculation was done on SHELX.

Fourier refinement revealed the missing atoms. Isotropic least

squares refinement commenced interspersed with difference map calculation. At R = 22%, R = 23%, an atom lying on the two fold axis, presumed to be the oxygen of water, was admitted for refinement.

At R = 19.58, R_m = 18.76, anisotropic refinement commenced; instead of the expected large drop in R, it stopped at R = 18.1%, R_M = 17.1%.

Eventually an error was discovered in the data processing, an inappropriate attenuation factor having been used. It is remarkable that MULTAN was able to solve the structure with the forty-six largest (|F|, admittedly, not |E|) reflexions overvalued.

The data were reprocessed and refinement continued. After four cycles of refinement R = 8.4%, $R_G = R_M = 9.6$ %. The hydrogen atoms, located from the difference map were introduced and the refinement continued to R = 6.1%, $R_M = R_G = 7.11$ %. Five reflexions, 9 7 3, 14 6 3, 0 2 1, 8 6 3, 3 7 3, were suppressed due to very large $\Delta/e.s.d$.

At this point R = 5.16%, $R_{G} = R_{M} = 4.67$ %. A weighting scheme with

 $w_{i} = (\sigma^{2}(|F|) + g(|F|^{2}|)^{-1})$

was then introduced. 'g' was not refined, as previous attempts had not converged, but varied until the analysis of variance was satisfactory. The final value of the refinement indices are shown below:

	R	RW	$R_M = R_G$	g	s ²
all F *	5.83	4.92	5.61	0.00115	0.040
allF	6.21	5.02	6.15	0.00115	.0.083
$\mathbf{F} > 2\sigma(\mathbf{F})$	4.57	4.66	5.53	0.00115	1.045

The final atomic parameters are shown in Table 6.3.1 and Table 6.3.2 The structure factors $(10|F_0|/10|F_c|)$ are listed in Appendix E (Table E.3).

except for noted suppressed reflexions.

Discussion

This, as far as is known, is the first structural determination of a dipeptide with an unsaturated residue.

Least squares analysis (Table 6.2.5) shows that the molecule consists of six planes, formed by the benzal, two amide, carboxylic acid, olefin and pyrrolidine (-Cl7) groups, with Cl7 lying out of the plane.

In all reported crystal structures of molecules containing proline residues, it has been found that four of the atoms of the pyrrolidine ring



form a plane from which the fifth, usually $C^{\gamma, 19-25}$ but occasionally $C^{\alpha 26}$ or $C^{\beta 27}$ deviates. Potential energy calculations show the pyrrolidine ring to be rather flexible, with a large number of almost iso-energetic conformations. However, two minima, one with C^{γ} displaced on the same side as C', of the plane formed by the other four atoms, and the other, higher, minimum with C^{γ} on the opposite side to C', have been indicated²⁸.

The proline residue in the present compound shows (Table 6.2.5, plane 6) $C^{\gamma}(C17)$ deviation on the opposite side to C'(C20).

The position of the phenyl group trans to the carbonyl confirms previous observations on the stereochemical course of the Bergmann synthesis²⁹.

The first amide link has a trans configuration, i.e. the nitrogen

hydrogen is trans to the carbonyl oxygen, with $\omega_1 = 166^{\circ}$. Comparison with the values for a peptide (Table 6.3.8) show the C2-N4 and N4-C5 bond lengths to be significantly longer and shorter respectively than in a peptide, reflecting conjugation with the olefin system. This conjugation will depend upon the twist about N4-C5 ($\Phi_1^{*} = -51.5^{\circ}$) as can be seen, for example, by comparing C-N in N-Ac- Δ Phe-OH(A) and N-Ac- Δ Phe-L-Pro-OH(B).

 $(C-N)_B > (C-N)_A$ (1.352 and 1.345 respectively) while $\tau(C2, N4, C5, C6)$ at 134° is greater than the equivalent torsion angle in A (105°) which is exactly the expected ordering. Other peptide structures containing dehydro-phenylalanyl residues (N-Ac(Δ Phe)_2Gly-OH²⁹ and N-Ac-(Δ Phe)_2-L-Ala-OH)³⁰ were not published in sufficient detail to permit comparison.

The second amide linke is planar and again has a trans configuration, i.e. N-C^{δ} is trans to the carbonyl oxygen, with $\omega_1 = 178.3^{\circ}$. The C-N distance is rather longer than the peptide equivalent quoted¹⁴ in Table 6.3.8, but much closer to another estimate³¹; the angles are in broad agreement. The C5-C13 at 1.51 A is a little surprising as both atoms are sp² hydridized.

The phenyl group is planar (Table 6.3.5, plane 4) with an average bond length of 1.384 Å. The value of the angle centred on the atom bonded to the rest of the molecule (C12- $\hat{C7}$ -C8) at 118.4°, is, as usual (see 6.2), below the average value (120.0°).

The styryl group is non-planar with τ (C5,C6,C7,C8) at 38°. The twist appears to produce little effect on the C6-C7 bond length (1.467A). The equivalent bond lengths in a series of cinnamic acids are 1.486, 1.469, 1.484 with torsion angles of 11.7°, 65.4 and 83.3° respectively¹¹. The bond length and angle in N-Ac- Δ Phe-OH, 1.470 and

 5° respectively, follows this lack of trend. It confirms the insensitivity of the sp² - sp² bond lengths shown in a comparison of bond lengths in planar butadienes (zero torsion angle) and non-planar cyclo-octatetraenes (torsion angle of 60°).

The carboxylic and group is planar (Table 6.3.5, plane 5) but with the hydrogen (H21) deviating from the plane by 0.13A. With an approximate e.s.d. of 0.04Å this is clearly significant. Certainly such a compromise of the carboxylic and groups planarity will be expensive in loss of resonance energy ($\propto A \sin^2 \tau$, τ being the twist about C20-021, A being approximately 15 K cal/mole), and is clearly an accommodation to allow the hydrogen to donate more easily to the water molecule. Of the conformations for the and hydrogen that maintain planarity of the group the one which minimises the distance to the carbonyl oxygen (the synplanar conformation) was the one found, suggesting an electrostatic interaction; possibly packing is also a factor. Experimentally, the syn conformation has been estimated as being more stable by around 4 Kcal/ mole, and is, except where an intra-molecular hydrogen bond is formed, the observed conformation in the crystalline phase.

The C20-O21-H21 (111°), O22-C20-O21 (123.7°), C20-O22 (1.196A) values agree reasonably with expected values¹⁵ of 110 - 114.5° (neutron determinations), 123°, 1.21A respectively, though C20-O21 (1.336) is longer than expected (1.31A).

The molecular conformation (Figure 6.3.1) may be grossly described by the conventional torsion angles of peptides, with $\Phi_1'(C2,N4,C5,C13) =$ -51.5°, $\Psi_1'(N4,C5,C13,N15) = 135.8$, $\Phi_2(C13,N15,C19,C20) = -70.2°$ and $\Psi_2(N15,C19,C20,O21) = 164.7°$. Naturally, the significance of the angles of the dehydro-phenylalanyl residue is not the same as for saturated residues as the spatial relationships differ. Comparison

with N-Ac- Δ Phe-OH($\phi' = -74^{\circ}$, $\Psi' = -11^{\circ}$) show considerable change.

Figure 6.3.2 shows a view parallel to the plane defined by the three labled atoms (the olefinic plane), while Figure 6.3.3, perpendicular to the plane, shows the less hindered side. Hydrogenation of the double bond, assuming:

- a) a similar conformation under the reaction conditions;
- b) the validity of the Prelog rule (that the substrate approach the catalyst with the least sterically hindered side);

and c) simple cis_attack

would be expected to produce an excess of N-Ac-D-Phe-L-Pro-OH over N-Ac-L-Phe-L-Pro-OH, a result actually observed³².

The justification for believing the conformation would be similar in solution lies in consideration of the intra-molecular non-bonded distances, particularly the H4...H8, H6...H12, O14...C19 distances (Table 6.3.7); these indicated a drive to co-planarity of the phenyl, olefin and amide groups which leaves little conformational freedom. Certainly, under the conditions of hydrogenation, which takes place in a protic solvent (methanol), this conformation seems closer to reality than the ring structure (formed by an intra-molecular hydrogen bond between the acidic hydrogen and amide oxygen) proposed.

There is however in N-Ac- $(\Delta Phe)_2$ -Res (Res being an amino acid residue) some evidence, from circular dichroism (C.D.) measurements, for a ring structure, as addition of urea produces significant change.

The hydrogen bond analysis (Table 6.3.6) shows the bond lengths to be within the expected range, with the amide hydrogen bond being longer (and presumably weaker) than the bonds involving the water molecule.

H23A-O23-H23B at 103° is, as expected, less than $\cos^{-1}(-1/3)$.

The H23A-O23...H21(92°) and H23B-O23...H21(119°) values deviate somewhat from the tetrahedral values. The distorsion is probably real and doesn't reflect uncertainty in the hydrogen positions, as the H23A-O23...O21(92.3°) and H23B-O23...O2(122.0°) are very similar; this would be expected if the hydrogen bonds are assumed to be linear.

The C13-O14...H4(112°) and C13-O14...N4(114.8) and, slightly less convincingly C2-O3...H23B(133°) and C2-O3...O23(137.1) values, again assuming hydrogen bond linearity (though O.3...H23A-O23 is only 162°), suggest that the lone pairs on the carbonyl oxygen are in sp² orbitals. Certainly in some carbonyl systems such as dehetopiperazine and cyanuric acid³³ there are lone pair lobes at about 110° to C=O.

A projection down to b axis (Figure 6.3.4) shows the molecules packed so as to form a layer, roughly parallel to the a b plane, held together by hydrogen bonding involving the water molecule, while the layers are held together by hydrogen bonding between the amide hydrogen (H4) and the amide oxygen (O14).



Figure 6.3.1. ORTEP drawing of the molecular structure of N-Ac Δ PheL-Pro-OH. $\frac{1}{2}$ H₂O with 50% probability ellipsoids.



Figure 6.3.2. A view of the molecule parallel to the plane defined by the labelled atoms.



Figure 6.3.3. View of the molecule normal to the plane defined by the labelled atoms.



Figure 6.3.4. View normal to the plane defined by the atoms, but antiparallel to that of Figure 6.3.3.



.

.



A projection down the b axis showing the molecular packing.

<u>Table 6.3.1</u>

Atomic fractional co-ordinates with e.s.d.'s in parenthesis and of the same magnitude as the final digit.

		x	У	z
C	21	0.4195(2)	0.9457(4)	0.5600(4)
c	22	0.4658(1)	0.8307(3)	0.6676(2)
c	03	0.5051(1)	0.8471(3)	0.8018(2)
N	14	0.4648(1)	0.7031(0)	0.6102(2)
. c	.5	0.5211(1)	0.5968(3)	0.6997(2)
c	:6	0.5015(1)	0.4634(3)	0.7048(2)
c	27	0.4192(1)	0.3961(3)	0.6367(2)
c	:8	0.3537(1)	0.4297(3)	0.4974(2)
c	:9	0.2769(1)	0.3644(3)	0.4431(3)
С	:10	0.2641(1)	0.2667(3)	0.5251(3)
С	11.	0.3286(2)	0.2320(3)	0.6636(3)
С	:12	0.4054(1)	0.2955(3)	0.7181(3)
С	:13	0.6110(1)	0.6446(3)	0.7828(2)
0	14	0.6401(1)	0.7100(3)	0.7177(1)
N	15	0.6568(1)	0.6086(3)	0.9241(2)
С	16	0.6257(2)	0.5560(4)	1.0202(3)
С	17	0.685 (2)	0.6185(5)	1.1667(3)
С	18	0.7691(2)	0.6185(4)	1.1686(2)
С	19	0.7440(1)	0.6566(3)	1.0098(2)
С	20	0.8000(1)	0.5847(3)	0.9642(2)
0	21	0.8722(1)	0.6541(3)	1.0170(2)
0	22	0.7843(1)	0.4793(2)	0.8934(2)
0	23	1.0000(0)	0.5227(3)	1.0000(0)
Table 6.3.1 continued

	x	У	Z
HLA	0.385(4)	0.913(8)	0.423(8)
HIB	0.384(3)	0.975(6)	0.576(5)
HIC	0.455(3)	1.018(7)	0.609(6)
H4	0.435(2)	0.694(4)	0.510(4)
H6	0.543(2)	0.406(3)	0.759(3)
H8	0.362(2)	0.493(4)	0.435(3)
Н9	0.238(2)	o.378(4)	0.357(4)
HIO	0.211(2)	0.230(3)	0.487(3)
Hll	0.326(2)	0.148(4)	0.729(4)
H12	0.458(2)	0.281(4)	0.819(3)
H16A	0.565(2)	0.582(5)	0.983(4)
H16B	0.634(2)	.0.440(5)	1.027(4)
H17A	0.689(3)	0.575(5)	1.249(5)
H17B	0.669(3)	0.721(7)	1.147(6)
H18A	0.807(2)	0.530(4)	1.209(4)
H18B	0.805(2)	0.680(4)	1.225(4)
H19	0.745(2)	0.751(3)	0.998(3)
H21	0.908(2)	0.606(4)	1.004(4)
H23	1.010(2)	0.464(4)	1.074(4)

173

Temperature factors. The e.s.d.'s, in parenthesis, are of the same magnitude as the final digit.

.

	Ull	U ₂₂	U ₃₃	U ₂₃	U ₁₃	U ₁₂
Cl	0.062(1)	0.047(1)	0.083(2)	-0.002(1)	0.035(1)	0.011(1)
C2	0.0441 (9)	0.049(1)	0.051(1)	-0.0113(9)	0.0250(8)	-0.0024(8)
03	0.084(1)	0.079(1)	0.053(1)	-0.024(1)	0.0265 (9)	0.012(1)
N4	0.0356(6)	0.0420(7)	0.0370(7)	-0.0035(7)	0.0172(6)	0.0014(6)
C5	0.0330(7)	0.0450(9)	0.0357(7)	-0.0039(7)	0.0169(6)	-0.0001(7)
C6	0.0372(8)	0.0436(9)	0.046(1)	-0.0011(8)	0.0182(7)	-0.0005(8)
C7	0.0406(8)	0.0349(8)	0.0411(9)	-0.0040(7)	0.0207(7)	-0.0020(7)
C8	0.049(1)	0.046(1)	0.0397 (9)	-0.0003(8)	0.0211(8)	-0.0101 (8)
С9	0.0443(9)	0.051(1)	0,047(1)	-0.003(1)	0.0158(8)	-0.0071(9)
C10	0.048(1)	0.055(1)	0.066(1)	-0.007(1)	0.032(1)	-0.0131(9)
C11	0.068(1)	0.051(1)	0.057(1)	0.003(1)	0.037(1)	-0.012(1)
C12	0.055(1)	0,046(1)	0.046(1)	0.0053(9)	0.0205(9)	-0.0035(9)
C13	0.0351(7)	0.0394 (8)	0.0346(7)	-0.0019(7)	0.0160(6)	-0.0018(7)
014	0.0394(6)	0.0550(8)	0.0365(6)	0.0037(6)	0.0179(5)	-0.0040(6)
N15	0.0391(7)	0.0518(9)	0.0350(7)	0.0003(7)	0.0176(6)	-0.0052(7)
C16	0.070(1)	0.075(2)	0.044(1)	-0.003(1)	0.035(1)	-0.019(1)
C17	0.081(2)	0.010(2)	0.042(1)	-0.011(1)	0.033(1)	-0.028(2)
C18	0.059(1)	0.087(2)	0.0302(9)	-0.006(1)	0.0145(8)	0.003(1)
C19	0.0402(8)	0.0415(9)	0.0339(8)	-0.0025(7)	0.0134(6)	-0.0011(7)
C20	0.0395(8)	0.0393(9)	0.0329(7)	-0.0002(7)	0.0126(6)	0.0001(7)
021	0.0410(7)	0.0540(9)	0.063(1)	-0.0129(8)	0.0216(7)	-0.0060(7)
022	0.0581(8)	0.0471(8)	0.0606(9)	-0.0152(7)	0.0288(7)	-0.0050(7)
023	0.051(1)	0.047(1)	0.040(1)	0.000(0)	0.0239(8)	0.000(0)
HLA	0.17(2)					•

174

	U ₁₁
HIB	0.10(1)
HIC	0.11(2)
84	0.07(1)
H6	0.05(1)
. H8	0.06(1)
H9	0.08(1)
HIO	0.05(1)
Hll	0.06(1)
H12	0.06(1)
H16A	0.08(1)
H16B	0.08(1)
HL'7A	0.09(1)
817B	0.12(2)
H18A	0.07(1)
H18B	0.07(1)
819	0.03(1)
H21	0.07(1)
H23	0.09(1)

Bond lengths (in $\stackrel{0}{A}$); e.s.d.'s, in parenthesis, are of the same magnitude as the final digit.

	Cl-HlA	1,29(7)	C11-C12	1.378(3)
	Cl-H1B	0.81(6)	C12-H12	1.03(3)
	C1-H1C	0.91(6)	C13-014	1.231(2)
	C1-C2	1.498(4)	C13-N15	1.337(2)
	C2-03	1.234(3)	N15-C19	1.472(2)
•	C2-N4	1.352(3)	N15-C16	1.481(3)
	N4-H4	0.92(3)	C16-H16A	1.02(4)
	N4-C5	1.419(3)	C16-H16B	1.11(4)
	C5-C13	1.511(2)	C16-C17	1.500(4)
	C5-C6	1.328(3)	C17-H17A	0.93(5)
	С6-н6	0.88(3)	C17-H17B	1.01(6)
	C6-C7	1.467(3)	C17-C18	1.536(4)
	C7-C12	1.392(3)	C18-H18A	1.04(4)
	C7–C8	1.393(3)	C18-H18B	0.86(4)
	C8-H8	0.96(3)	C18-C19	1.538(3)
	C8-C9	1.386(3)	С19-н19	0.91(3)
	С9-Н9	0.84(4)	C19-C20	1.508(3)
	C9-C10	1.369(4)	C20-022	1.196(2)
	C10-H10	0.93(3)	C20-021	1.336(2)
	C10-C11	1.385(4)	021-H21	0.87(4)
	C11-H11	1.08(4)	023-н23	0.90(4)

.

Bond angles (degrees). The e.s.d.'s, in parenthesis, are of the same magnitude as the final digit.

H23A-023-H23B	103(5)
H21-021-C20	111(2)
021-C20-022	123.7(2)
021-C20-C19	110.3(2)
022-C20-C19	126.0(2)
C20-C19-C18	111.0(2)
С20-С19-Н19	110(2)
С18-С19-Н19	112(2)
N15-C19-H19	108(2)
C20-C19-N15	112.5(2)
С18-С19-№15	103.6(2)
С19-С18-Н18А	114(2)
С19-С18-Н18в	108(2)
H18A-C18-H18B	98(3)
C19-C18-C17	103.4(2)
C17-C18-H18A	118(2)
C17-C18-H18B	115(2)
С18-С17-Н17А	113(3)
С18-С17-Н17В	102(3)
H17A-C17-H17B	119(4)
C18-C17-C16	103.5(2)
C16-C17-H17A	117(3)
С16-С17-Н17В	100(3)
C17-C16-N15	103.2(2)
С17-С16-Н16А	114(2)

177

Table 6.3.4	continued
С17-С16-Н16В	109 (2)
Н16А-С16-Н16В	111 (3)
N15-C16-H16A	113(2)
N15-C16-H16B	107 (2)
C19-N15-C16	111.7(2)
C19-N15-C13	119.4(2)
C16-N15-C13	126.9(2)
N15-C13-014	122.6(2)
N15-C13-C5	117.9(2)
014-C13-C5	119.4(2)
C13-C5-N4	113.5(2)
c13-c5-c6	120.5(2)
C6-C5-N4	125.7(2)
С7-С6-Н6	114(2)
C7-C6-C5	129.4(2)
С5-С6-Н6	117(2)
C6-C7-C8	123.9(2)
C6-C7-C12	117.7(2)
C12-C7-C8	118.4(2)
c7-c8-c9	120.2(2)
С7-С8-Н8	121(2)
С9-С8-Н8	119(2)
C8-C9-C10	120.7(2)
С8-С9-н9	121(3)
С10-С9-Н9	118(3)
C9-C10-C11	119.8(2)
С9-С10-Н10	118(2)

C11-C10-H10	122(2)
c10-c11-c12	120.0(2)
C10-C11-H11	124(2)
C12-C11-H11	116(2)
c11-c12-c7	121.0(2)
C7-C12-H12	111(2)
С11-С12-Н12	128(2)
C5-N4-C2	120.1(2)
C5-N4-H4	121(2)
C2-N4-H4	118(2)
N4-C2-C1	116.3(2)
N4-C2-O3	119.6(2)
c1-c2-03	124.0(2)
C2-C1-HIA	116(3)
C2-C1-H1B	108(4)
C2-C1-H1C	98(3)·
HIA-CI-HIB	109(4)
HIA-CI-HIC	129(5)
H1B-C1-H1C	93 (5)

Deviations from least squares planes (in Angstroms). Atoms with asterisks are not used in the least squares plane calculation.

	1	2	3	4	5	6	7	8
Cl	0.002	3.229*						
C2	-0.008	2.663*				·		
03	0.003	3.246*						
N4	0.003	1.422*					-0.010	-0.012
C5	0.290*	0.690*	-0.004				0.030	0.026
C6		0.021					-0.011	-0.030
C7		-0.014		0.004				0.014
C8		-0.010		0.000				
C9		0.001		-0.003				
C10		0.013	•	0.002				
C11		0.006		0.003				
C12.		-0.016	•.	-0.005				
C13		0.610*	0.013				-0.009	0.002
014		0.317*	-0.005					
N15		0.812*	-0.005			-0.028		
C16						0.017		
C17						0.571*		
C18						-0.016		
C19			0.050*		0.002	0.026		
C20					-0.006	-1.133*		
021					0.002			
022					0.003			

180

÷

-

Table 6.3.5 continued

.

	a _i	b,	c,	đ,
i	_	-	Ŧ	1
1	0.8229	0.2340	- 0.5178	6.5953
2	-0.2997	0.7179	0,6283	2.4774
3	-0.2699	0.8542	0.4444	3.7278
4	-0.2903	o.7224	0.6275	2.5373
5	0.3940	-0.5491	0.7370	4.0144
6	-0.2804	0.9181	0.2800	3.4136
7	-0.1306	0.2435	0.9611	2.6979
8	-0.1228	0.2527	0.9597	2.8166

,

Hydrogen bonding - bonding expected by symmetry is not tabulated

Donor	Acceptor	Angle	Distance	Distance	Position of
Х-н	Y	XHY	HY	XY	Donor
N4-H4	014	170	2.1	3.00	(-l-x,y,-l-z)
021-H21	023	170	1.9	2.76	(x, y, z)
023-H23A	03	162	1.8	2.71	(-x-½,y-½,-z)

(< Van der Waals	+ 0.35Å)
H1C03	2.39
C2C13	2.926
03C13	2.827
N4C7	3.092
N4014	2.852
E4C1	2.50
H4H8	2.25
H4Hla	2.28
C503	2.690
C5C8	3.170
C5C16	2.957
C6H12	2.47
C6C16	3.072
H6H12	2.31
H6Cl3	2.55
н9н8	2.29
H9H10	2.20
014C19	2.738
N15C6	2.975
N15H17B	2.49
N15022	2.820
H16AH17B	2.26
H17BH18B	2.25
C19H17B	2.53
H19H18B	2.18

Close intra-molecular non-bonded contacts

t

Table 6.3.7 continued

819021	2.44
C20C13	3.085
C20014	3.037
H21022	2.32

Comparison of the important bond lengths $(\stackrel{0}{A})$ and angles $(\stackrel{0}{})$				
	Peptide	lst Amide group	2nd Amide group	N-Ac-∆Phe-OH
C-C(0)	1.51	1.498	1.511	1,490
C=0	1.24	1.234	1.231	1.230
C-N .	1.32	1.352	1.337	1.345
N-C ^c	1.45	1.419	1.472	1.419
c-ĉ (o) -o	120.5	124.0	119.4	122.9
с-ĉ (о) -n	116.0	116.3	117.9	116.5
0-Ĉ-N	123.5	119.6	122.6	120,5
C−Ñ−C [∞]	122	120.1	119.4	121.1
$N = \hat{c}^{\ast} = C(\Omega)$	111	113.5	112.5	117.4

REFERENCES

- C. Shin, Y. Chigira, M. Masaki and M. Ohta, Tetrahedron Lett., p.4601 (1967).
- 2. W. L. Meyer, G. E. Templeton, C. I. Grable, R. Jones, L. F. Kuyper, R. B. Lewis, C. W. Sigel and S. H. Woodhead, J.A.C.S., 97, p.3802 (1975).
- 3. K. Fukushima, K. Yazawa and T. Arai, J. Antibiotics, <u>26</u>, p.175 (1973).
- 4. C. Shin, M. Hayakawa, K. Mikami and J. Yoshimura, Tetrahedron Lett., p.863 (1977).
- 5. A. Dossena, R. Marchelli and A. Pochini, J.C.S. Chem. Comm., p.771 (1974).
- M. Koncewicz, P. Mathiaparanam, T. F. Uchytil, L. Sparapano,
 J. Tam, D. H. Rich and R. D. Durbin, Biochem. Biophys. Res. Commun.,
 53, p.653 (1973).
- 7. E. G. Breitholle and C. H. Stammer, Tetrahedron Lett., p.381 (1975).
- 8. D. H. Rich, P. Bhatnagar, P. Mathiaparanam, J. A. Grant and J. P. Tam, J. Org. Chem., <u>43</u>, p.296 (1978).
- 9. F. McCapra and M. Roth, J.C.S. Chem. Comm., p.894 (1972).
- M. L. English and C. H. Stammer, Biochem. Biophys. Res. Comm., 85, p.780(1978).
- 11. S. E. Filippakis, L. Leiserowitz, D. Rabinovich and G. M. J. Schmidt, J.C.S. Perkin II, p.1750 (1972).
- 12. A. D. Walsh, Discuss. Faraday Soc., <u>2</u>, p.18 (1947).
- A. Domenicana, A. Vaciago and C. A. Coulson, Acta Cryst., <u>B31</u>, p.221 (1975).
- 14. R. E. Marsh and J. Donohue, Adv. Protein Chem., 22, p.235 (1967).
- 15. L. Leiserowitz, Acta Cryst., <u>B32</u>, p.775 (1976).
- 16. See Chapter 3.
- 17. H. Schenk, Acta Cryst., A28, p.412 (1972).
- 18. P. Debrye, Ann. Phys. (Leipzig), <u>46</u>, p.809 (1915).
- 19. A. McL. Mathieson and H. K. Welsh, Acta Cryst., 5, p.599 (1952).
- 20. Y. C. Leung and R. E. Marsh, Acta Cryst., <u>11</u>, p.17 (1958).
- 21. J. Fridrichson and A. McL. Mathieson, Acta Cryst., 15, p.569 (1962).

186

- 22. T. Ueki, T. Ashida, M. Kakudo, Y. Sasada and Y. Katsube, Acta Cryst. <u>B25</u>, p.1840 (1969).
- 23. T. Matsuzaki and Y. Iitaka, Acta Cryst. <u>B27</u>, p.507 (1971).
- 24. G. Kartha and G. Ambady, Acta Cryst. <u>A28</u>, p.533 (1972).
- 25. N. Sharnala and K. Venkatesan, Cryst. Struc. Commun., 21, p.5 (1973).
- 26. Y. Mitsui, M. Tsuboi and Y. Iitaka, Act Cryst. <u>B25</u>, p.2182 (1969).
- 27. E. Benedetti, M. R. Ciajolo and A. Maisto, Acta Cryst. <u>B30</u>, p.1783 (1974).
- 28. G. N. Ramachandran, A. V. Lakshminarayanan, R. Balasubramanian and G. Tegoni, Biochem. Biophys. Acta, <u>221</u>, p.165 (1970).
- O. Pieroni, G. Montagnoli, A. Fissi, S. Merlino and F. Ciardelli, J.A.C.S., <u>97</u>, p.6820 (1975).
- O. Pieroni, A. Fissi, S. Merlino and F. Ciardelli, Israel J. Chem., <u>15</u>, p.22 (1977).
- G. N. Ramachandran, A. S. Kolaskar, C. Ramakrishnan and
 V. Sasisekharan, Biochem. Biophys. Acta, <u>359</u>, p.298 (1974).
- 32. O. Pieroni, D. Bacciola, A. Fissi, R. A. Felicioli and
 E. Balestreri, Int. J. Peptide Protein Res., <u>10</u>, p.107 (1977).
- 33. G. C. Verschoor and E. Keulen, Acta Cryst., B27, p.134 (1971).

APPENDIX A

PROBABILITY ASPECTS OF STRUCTURE SOLUTION

Application of probability theory¹ produces a very useful relationship between intensities and the triplet invariants.

A.1
$$P(\Omega | A) \approx [1/2\pi I_0(A)] \exp(A\cos\Omega)$$

$$A = 2 |E_h E_{h-k} E_k| / N^{\frac{1}{2}} (1 - |E_h|^2 / N) \approx 2 |E_h E_{h-k} E_k| / N^{\frac{1}{2}}$$

$$I_0 \text{ is a Bessel function and } \Omega = \phi_h + \phi_{-h-k} + \phi_k$$

Using A.l, the conditional expectation for a given A of:

a)
$$\phi_{h} + \phi_{k} + \phi_{-h-k} \text{ is}$$

$$E((\phi_{h} + \phi_{k} + \phi_{-h-k})|A) = [1/2\pi I_{0}(A)] \int_{0}^{2\pi} \Omega \exp(A\cos\Omega) d\Omega = 0$$
b) Sin m($\phi_{h} + \phi_{-h-k} + \phi_{k}$) is
$$E(\sin m(\phi_{h} + \phi_{-h-k} + \phi_{k})|A) = [1/2\pi I_{0}(A)]$$

$$x \int_{0}^{2\pi} \sin m\Omega \exp(A\cos\Omega) d\Omega = 0$$
c) $\cos m(\phi_{h} + \phi_{-h-k} + \phi_{k}) \text{ is}$
A.2 $E(\cos m(\phi_{h} + \phi_{-h-k} + \phi_{k})|A) = [1/2\pi I_{0}(A)]$

$$x \int_{0}^{2\pi} \cos m\Omega \exp(A\cos\Omega) d\Omega$$

As $\exp(A\cos\Omega) = I_0(A) + 2\Sigma I_m(A)\cos m\Omega$.

$$A.2 = I_{m}(A) / I_{0}(A)$$

d)
$$\cos^{2}(\phi_{h} + \phi_{k} + \phi_{-h-k})$$
 and $\sin^{2}(\phi_{h} + \phi_{k} + \phi_{-h-k})$
 $E(\cos^{2}(\phi_{h} + \phi_{k} + \phi_{-h-k})|A) = [1/2\pi I_{0}(A)]$
 $\times \int_{0}^{2\pi} \cos^{2}\Omega \exp(A\cos\Omega) d\Omega = \Lambda$

As
$$\sin^2 \Omega = 1 - \cos^2 \Omega$$

$$E(\sin^2 (\phi_h + \phi_k + \phi_{-h-k}) | A) = [1/2\pi I_0(A)]$$

$$\times \int_0^{2\pi} \exp(A\cos\Omega) d\Omega - A = 1 - A$$

$$= 1 - A$$

Using $\cos^2 \Omega = \frac{1}{2} (1 + \cos 2 \Omega)$ and $I_2(A) = I_0(A) - 2I_1(A)/A$ $\Lambda = 1 - I_1(A)/(AI_0(A))$ The variance is found using

$$\sigma^{2}(x) = E(x^{2}) - (E(x))^{2}$$

$$\sigma^{2}(\sin(\phi_{h} + \phi_{k} + \phi_{-h-k})|A)$$

$$= E(\sin^{2}(\phi_{h} + \phi_{k} + \phi_{-h-k})|A) - E(\sin(\phi_{h} + \phi_{k} + \phi_{-h-k})|A)^{2}$$

$$= E(\sin^{2}(\phi_{h} + \phi_{k} + \phi_{-h-k})) = I_{1}(A)/AI_{0}(A)$$

Similarly

$$\sigma^{2} (\cos(\phi_{h} + \phi_{k} + \phi_{-h-k}) | A)$$

= $E (\cos^{2}(\phi_{h} + \phi_{k} + \phi_{-h-k}) | A) - (E (\cos(\phi_{h} + \phi_{k} + \phi_{-h-k}) | A))^{2}$
= $\Lambda - I_{1}^{2} (A) / I_{0}^{2} (A)$

Use of these relationships has been made in selecting a good starting set of reflexions by estimating α_{h} .

$$\alpha_{h}^{2} = \Sigma \kappa_{hh}^{*} \sin (\phi_{h}^{*} + \phi_{h-h}^{*} - \phi_{h}^{*})^{2} + \Sigma \kappa_{hh}^{*} \cos (\phi_{h}^{*} + \phi_{h-h}^{*} - \phi_{h}^{*})^{2}$$

$$= \Sigma \kappa_{hh}^{2} + \Sigma \Sigma \kappa_{hh}^{*} \kappa_{hh}^{*} (\cos \phi_{h}^{*} + \phi_{h-h}^{*} - \phi_{h}^{*})$$

$$\cos (\phi_{h}^{*} + \phi_{h-h}^{*} - \phi_{h}^{*}) + \sin (\phi_{h}^{*} + \phi_{h-h}^{*} - \phi_{h}^{*})$$

$$\sin (\phi_{h}^{*} + \phi_{h-h}^{*} - \phi_{h}^{*})$$

As the trigonometric terms are unknown, expectation values are substituted to give an estimate of $\frac{a}{b}^2$

$$< \alpha_{h}^{2} > = \Sigma \kappa_{hh}^{2}, + \Sigma \Sigma \kappa_{hh}, \kappa_{hh}, (I_{1}(\kappa_{hh})/I_{0}(\kappa_{hh}))$$
$$\times (I_{1}(\kappa_{hh})/I_{0}(\kappa_{hh}))$$

The estimate can be weighted² by substituting $W_{h'}W_{h-h'}\kappa_{hh'}$ for $\kappa_{hh'}$ where $W_{h} = \tanh[\langle \alpha_{h}^{2} \rangle^{\frac{1}{2}}/2]$.

A similar relationship to A.l exists for quartets (four phase invariants):

A.3
$$P(\psi | B) \approx [1/2\pi I_0(B)] \exp B\cos(\psi)$$
$$B = 2 | E_h E_k E_{\ell} E_{-h-k-\ell} | / N$$
$$\psi = \phi_h + \phi_k + \phi_{\ell} + \phi_{-h-k-\ell}$$

However, A.3 can be drastically modified if the magnitudes of the cross products, $|E_{h+k}|$, $|E_{h+l}|$, $|E_{l+k}|$ are known. If they are large A.3 remains unchanged, but if they are (all) zero or small the relationship becomes³

A.4 $[1/2\pi I_0(2B)] \exp -2B\cos(\psi)$

The way in which expectations about the invariant can be altered just from knowledge of three reflexions is a consequence of the concept of nested neighbourhoods, a neighbourhood being a set of reflexions, relatively effective, compared with the body of the data, in determining the invariant⁴. In the case considered $|E_h|$, $|E_k|$, $|E_{l}|$, $|E_{-h-k-l}|$ constitute the first neighbourhood, which nests in the second neighbourhood $|E_h|$, $|E_k|$, $|E_l|$, $|E_{-h-k-l}|$, $|E_{h+k}|$, $|E_{l+k}|$, $|E_{l+k}|$, which nest in the third, in a similar manner³.

Utilization of A.4 produces the following results

A.5 $E(\psi | B) = \pi$

 $E(\sin(m\psi)|B) = O$

 $E(\cos(m\psi)|B) = -I_m(2B)/I_0(2B)$

If there are several indications of ϕ_{-h-k-l} by A.4, the conditional probability distribution will take the form

A.6 $N^{-1} \pi_i P(\psi_i | B_i)$

where N is a normalising term.

The exponent of A.6 can be simplified in the following manner:

A.7
$$\Sigma B_{i} \cos(\phi_{h+k+\ell} - \gamma_{i})$$
$$= \cos\phi_{h+k+\ell} \Sigma B_{i} \cos(\gamma_{i}) + \sin\phi_{h+k+\ell} \Sigma B_{i} \sin(\gamma_{i})$$
$$\gamma_{i} = \phi_{h_{i}} + \phi_{k_{i}} + \phi_{\ell_{i}}$$

Using $\langle \sin(\phi_h + \phi_k + \phi_\ell - \phi_{h+k+\ell}) \rangle = 0$

$$Tan(\delta) = \Sigma B_{i} sin(Y_{i}) / \Sigma B_{i} cos(Y_{i}) = S/C$$

$$Sin(\delta) = \Sigma B_{i} sin(Y_{i}) / \beta_{h} Cos(\delta) = \Sigma B_{i} cos(Y_{i}) / \beta_{h}$$

$$\beta_{h} = (s^{2} + C^{2})^{\frac{1}{2}}$$

A.6 becomes

A.8
$$[2\pi I_0(2\beta_h)]^{-1} \exp -2\beta_h \cos(\phi_{h+k+\ell} - \delta)$$

Higher order invariants have similar distributions to the triplets and quartets but since the strength of indication of a phase varies as $(N^{-1})^{m-2}$ where 'm' is the invariant order, they are less useful, though the idea of 'embedding' shows much promise^{5,6}.

References

1.	W. Cochran, Acta Cryst., <u>8</u> , p.473 (1955)
2.	P. J. Roberts, R. C. Pettersen, G. M. Sheldrick, N. W. Isaacs, and O. Kennard, J. Chem. Soc., Perkin Trans., 2, p.1978 (1973).
3.	H. Hauptman, Acta Cryst., <u>A31</u> , p.680 (1975).
4.	H. Hauptman, Acta Cryst., <u>A34</u> , p.525 (1978).
5.	C. Giacovazzo, Acta Cryst., <u>A33</u> , p.933 (1977).
6.	C. Giacovazzo, Acta Cryst., <u>A35</u> , p.296 (1979).

APPENDIX B

PEPTIDE CHEMISTRY

B.1 Synthesis

Since the carbodiimide method only was used, other methods will be mentioned only briefly.

Mixed Anhydride Method

The reactions involved are shown in Figure B.l.l(a). The amine base must not be too nucleophilic if racaemisation is to be minimised, but if the reaction with the chloroformate is to occur, a hindered base such as di-isopropyl ethylamine cannot be used². N-methyl morpholine has been found to be a suitable compromise.

The fact that Cl^{-} is a fairly good nucleophile, means that even if the exact quantity of amine is used, free base is present; replacing Cl^{-} by BF_{4}^{-} , a very poor nucleophile, produces a crystalline complex⁵, which reacts with minimal racaemisation².

An alternative means of anhydride preparation is represented by E.E.D.Q.³ and derivatives⁴ (Figure B.l.1(b)). E.E.D.Q. has also been incorporated in an insoluble polymer⁶.

The Azide Procedure

At one time, this was the method of choice when racaemate-free coupling was required (though racaemisation occurs under certain conditions^{7,8,9}), but faster methods are now preferred. Figure B.1.2(a) shows the coupling stages.

N-carboxyamino Acid Anhydrides (N.C.A.'s Leuch's anhydrides)

Although originally used only to generate polyamino acids¹⁰, because of the speed of reaction and lack of racaemisation, attempts have been made, with some success^{11,12}, to adapt it for general synthesis (Figure B.1.2(a))

The pH requires very careful control, as at low pH there is a tendency for the carbamate ion to lose carbon dioxide, leading to

xîv









Figure B.l.l (a)

Α

The mixed anhydride method;

for optimum yields R⁴ is isobutyl.

xv



Figure B.1.1 (b) E.E.D.Q. variation on the mixed anhydride method; because reaction 2 is slow, B (mixed anhydride) reacts with the amino component before side-reactions can occur; R is -COOEt and - the body of the molecule; 3 requires EtOCOC1/EtOH/Et₃N.

Xvi







(a)









(b)

Figure B.1.2 (a) The azide procedure; the ester (R is usually Me or Et) is converted to the hydrazide, which upon reaction with nitrous acid (2) forms the water insoluble azide. (b) The N.C.A. method; 3 requires pH %10/H₂O/^{\ooldowoodle} C.







Figure B.1.3 The carbodiimide method:- the O-acyl urea (O-A.U.) can go to the N-acyl urea (N-A.U.) if the amino component is absent. and O represent the remainder of the molecule. over-reaction, while at high pH there is significant reaction with hydroxide ion.

The D.C.C.I. Method

This method (Figure B.1.3) combines the advantages of cheap, readily available, chemicals, with not too rigorous reaction conditions and easily separated by-products.

In addition, should high optical purity be required, an additive such as HOBT^{14} (1-hydroxybenzotriazole, preferable to the earlier HOSu, which can form β alanine¹³ side products) will usually ensure it, a point of special importance when N-acetyl derivatives are being synthesised; the reaction in the 'Eintopf Methode' proceeds via the active ester; these newer active esters are superseding the older generation of active ester, which are tedious to prepare¹⁶, slow to react and often react incompletely¹⁵.

Other methods of interest include condensation by oxidation - reduction 17 and via phosphorus compounds 18 .

B.2 Racaemisation

The reactions and mechanisms that produce racaemisation during peptide synthesis have received much attention.

It has been well established that racaemisation of N-benzoylamino acid active esters is base catalysed and proceeds via oxazolone (azlactone) formation¹⁹⁻²¹. It is clear, considering the canonical forms

$$\begin{array}{c} 0 & 0^{-} \\ R' - \begin{array}{c} 0 \\ C - NH - R'' \leftrightarrow R' - \begin{array}{c} 0 \\ C = NH - R'' \end{array}$$

the carbonyl oxygen is more, or less, nucleophilic according as R' is electron donating or not. In the case of derivatives such as the urethanes, the diminished nucleophilicity of the carbonyl oxygen does not allow ring closure to form oxazolone.

Oxazolone formation is believed to result from the reactions shown in Figure B.2.1²². As well as kinetic evidence, a diminution in the racaemisation rate with N-methyl amide derivatives and imino acids such as proline support the mechanism.

The observed degree of racaemisation is very considerably effected by the ratio between the rate of oxazolone formation and racaemisation, and the rate of coupling. This is influenced greatly by the following factors:

Basicity

In addition to solvent basicity, there is the base added to neutralise the hydrochloride; ideally this base should be a considerably stronger base than the amino component while being as weak a nucleophile as possible; D.I.E.A. represents a good compromise²³. Dielectric Constant

As reactions involving charged components usually show a marked

 $\mathbf{X}\mathbf{X}$



Figure B.2.1. Racaemisation via oxazolone formation. 1 is fast and 2 is slow. C (L-oxazolone) ionises far more readily than A (L-amino acid residue) as the anion is stabilised by conjugation. D = D,L-oxazolone = D,L-amino and residue.

xxi

dependence on dielectric constant, racaemisation occurs most readily in solvents, such as water, with high dielectric constants.

Ionic Strength

This is an influence for the same reasons as the dielectric constant, and probably accounts for the 'chloride ion effect', 19,24. Nucleophilicity of Reaction Components

If a nucleophile reacts with the oxazolone rapidly, then racaemisation is prevented. The most important nucleophile during coupling is usually the amino component and, not unexpectedly, the least hindered ester is the most effective nucleophile (indicated by the degree of optical purity)²⁵.

Temperature

A low temperature is generally found to be beneficial due both to the increased stability of some reactants and a suppression of side reactions.

To minimise racaemisation during coupling clearly it is desirable that it proceed as rapidly as possible; to this end the minimum amount of solvent is used. In addition, with the D.C.C.I. method, an additive such as HOBT may be added; this has two beneficial effects:

The coupling proceeds very rapidly (via the active ester) e.g. the HOBT ester of Z-L-Phe reacts with HCl.L-Val-OMe/N-ethyl morpholime to give the dipeptide at over 90% yield in 5 minutes at $0^{\circ}C^{26}$.

HOBT is a fairly good nucleophile so it can block racaemisation at the oxazolone level (Figure B.2.1).

In derivatives such as the urethanes, which do not form oxazolones, the principal means of racaemisation is direct α -hydrogen abstraction²⁷⁻³¹, this being true even for derivatives of amino acids like cystine, well known for its tendency to racaemise, which at one

XXII

time were believed to racaemise by reversible β elimination. The effect of the side chain on the rate of racaemisation by direct proton abstraction has been determined for a series of benzoyl amino acid anilides³² (which cannot form oxazolones); the rate of racaemisation increased from value to phenylglycine in the following manner:

Val<But<Leu<Ala<Phe<Ph Gly

The rate is clearly related to the electronegativity of the side group, except for valine where the steric effect predominates.

References

٦

1

D 17

T...

.

.

_ .

1.	U. R. Vaughan UI., and R. L. USato, J.A.C.S., <u>14</u> , p.676 (1952).
2.	G. W. Anderson, 'Peptides; chemistry and biochemistry', Ed. B. Weinstein and S. Laude, Dekker, 1970.
3.	B. Belleau and G. Malek, J.A.C.S., <u>90</u> , p.1951 (1968).
4.	Y. Kiso and H. Yajima, J.C.S.(D), p. 942 (1972).
5.	J. V. Paukstelis and Moon-gue Kim, J. Org. Chem., <u>39</u> , p.1499 (1974).
6.	J. Brown and R. E. Williams, Can. J. Chem., <u>49</u> , p.3765 (1971).
7.	P. Sieber, B. Riniker, M. Brugger, B. Kamber and W. Rittel, Helv. Chim. Acta, <u>53</u> , p.2135 (1970).
8.	D. S. Kemp, Z. Bernstein and J. Rebek Jr., J.A.C.S., <u>92</u> , p.4756, (1970).
9.	G. W. Anderson, J. E. Zimmerman and F. Callahan, J.A.C.S., <u>88</u> , p.1338 (1966).
10.	E. Katchalski and M. Sela, Advan. Protein Chem., <u>13</u> , p.243 (1958).
11.	R. G. Denkewalter, D. F. Veber, F. W. Holly and R. Hirschmann, J.A.C.S., <u>91</u> , p.502 (1969).
12.	J. J. Maher, M. E. Furey and L. T. Greenberg, Tetrahedron Letters, p.1581 (1972).
13.	H. Gross and L. Bilk, Tetrahedron Letters, 24, p.6935 (1968).
14.	W. König and R. Geiger, Chem. Ber., <u>103</u> , p.788 (1970; W. König and R. Geiger, Chem. Ber., <u>103</u> , p. 2024 (1970); W. König and R. Geiger, Chem. Ber. <u>103</u> , p.2034 (1970).
15.	K. Lübke, 'Peptides 1969', Ed. E. Scoffone, North-Holland Publ., Amsterdam.
16.	J. Kovacs, G. L. Mayer, R. H. Johnson, R. E. Cover and V. R. Ghatak, J. Org. Chem., <u>35</u> , p.1810 (1970).
17.	T. Mukaiyama, R. Matsueda and M. Suzuki, Tetrahedron Letters, p.1907 (1970).
18.	G. Gawne, G. W. Kenner and R. C. Sheppard, J.A.C.S., <u>91</u> , p.5669 (1969).
19.	M. W. Williams and G. T. Young, J.C.S., p.3701 (1964).
20.	I. Antonovics and G. T. Young, J.C.S.(C), p.595 (1967).
21.	D. S. Kemp and J. Rebek, J.A.C.S., <u>92</u> , p.5792 (1970).

xxiv

- 22. D. S. Kemp and S. W. Chien, J.A.C.S., 89, p.2745 (1967).
- 23. O. Grahl-Nielson, J.C.S. Chem. Commun., p.1588 (1971)
- 24. I. Antonovics and G. T. Young, J.C.S. Chem. Commun., p.398 (1965).
- 25. M. Goodman and W. J. McGahren, Tetrahedron Letters, <u>23</u>, p.2031 (1967).
- 26. W. König and R. Geiger, 'Proceedings of the Third American Peptide Symposium', p.343, Ann Arbor Science Publ., Ann Arbor, Michigan.
- 27. B. Liberek, Tetrahedron Letters, p. 925 (1963).
- 28. B. Liberek, Tetrahedron Letters, p.1103 (1963).
- 29. B. Liberek, T. Nowicka and Z. Grzonka, Tetrahedron Letters, p.1479 (1963).
- 30. B. Liberek and Z. Grzonka, Tetrahedron Letters, p.159 (1964).
- 31. B. Liberek and A. Michalik, Acta Chim. (Budapest), 44, p.71 (1965).
- 32. M. Sato, T. Tatsuno and H. Matsuo, Chem. Pharm. Bull., <u>18</u>, p.1794 (1970).

APPENDIX C

STRUCTURES ATTEMPTED

C.1 N-Acetyl Dehydro-Phenylalanyl-L-Leucine

Experimental

The compound was prepared by the method of Bergmann¹ and crystals 'grown by cooling slowly a water/methanol/n-hexane solution; the first crystals obtained were large, off-white, hexagonal prisms and these were used for determination of the space group and cell parameters. After many recrystallisations, small pure crystals with a sharp melting point, but tiny cross section and unsatisfactory extinction under polarised light, were obtained; they were unsuitable for data collection.

Crystal characteristics

Formula $C_{17} H_{22} N_2 O_4$ a = b = 12.555 Å c = 21.711 Å $\gamma = 60^{\circ}$ $D_0 = 1.26 \text{ g.cm}^{-3}$ $D_c = 1.07 \text{ g.cm}^{-3}$ Z = 6

Space group P61

An attempt was made to collect data on the CAD 4 using an impure crystal, but the reflexions were too weak to be worth collecting.
C.2 N-Acetyl Dehydro-Phenylalanyl-L-Alanine

Experimental

The compound was prepared by the method of Bergmann² and crystals grown from water.

Data collection was attempted but not continued due to the weak nature of the reflexions.

Other compounds that proved unsuitable are mentioned in Chapter 3.

References

 O. K. Behrens, D. G. Doherty and M. Bergmann, J. Biological Chem., <u>136</u>, p.61 (1940).

•

 D. G. Doherty, J. E. Tietzman and M. Bergmann, J. Biological Chem., <u>147</u>, p.617 (1943).

• 1

APPENDIX D

RECENT DEVELOPMENTS

Recent Developments

The last couple of years have seen the emergence of components of the likely successor to MULTAN, MAGLIN (MAGic integers and LINear equations)^{1,2}.

Magic integers have already been introduced into later versions of MULTAN, but recent progress promises to make them an even more efficient procedure³; the use of linear equations in this way is new.

In cycles Σ_2 relationships take the form:

$$\phi_{1i} + \phi_{2i} + \phi_{3i} + b_i = n_i$$

where 'n' is integer.

As the number of Σ_2 relationships greatly exceed the number of phases least squares solution suggests itself.

Putting the κ weighted Σ_2 relationships in matrix form gives

$$A\phi = C$$

which has a least squares solution

$$\phi = (A^{T}A)^{-1}A^{T}C$$

The original technique was to use estimates of the phases to solve for 'n', make 'n' the nearest integer, and substitute in C (after κ weighting) to obtain new estimates for the phases, the only change from one phase set to the next being in the column vector C. This last point is important as matrix inversion is very time consuming (increasing as the cube of the matrix order), though the matrix order can be reduced by keeping constant certain phases (recognisable from the elements of $(A^TA)^{-1}A^T)$ which are stable from one cycle to the next.

The question arises of what to do, if for example, 'n' is calculated as 1.5 or, for that matter, values clustered around a midpoint. An explicit weighting to reflect the uncertainty would mean A changing and consequent matrix inversion, every cycle. However, if the

xxxi

example is written thus:

$$\phi_{1i} + \phi_{2i} + \phi_{3i} + b_i = n_i = n_i^1 + \alpha_i$$

with $\alpha = 0.5$, the equation is effectively eliminated from the least squares, while $\alpha = 0$ will tend to decrease ϕ_{ji} j = 1,3. De facto weighting can be achieved by a function $f(\alpha)$ such that f(0.5) = .5and f(0) = 0. $4\alpha^3$ has been found to be suitable. The example now takes the form

$$\kappa (\phi_{1i} + \phi_{2i} + \phi_{3i}) = \kappa (n'_i - b_i + f_i(\alpha)) = C_i$$

so that only C varies with each cycle.

Sophistications, like extrapolation of phases, produce more efficient computer use, but lack of an adequate indicator of refinement progress, lessens the effect.

The linear equation procedure, like tangent refinement, by requiring the triplet invariants to be close to zero, blurs enantiomorph discrimination^{4,5}. Definition is possible, by rather empirical, ad hoc methods or, far more satisfactorily, by restrictions on phases related by magic integers⁶.

Renewed interest in Karle-Hauptman (K-H) determinants has been prompted by the derivation of the maximum determinant rule⁷⁻¹⁰: that the most probable set of phases for the reflexions in the K-H determinant is that which maximises its value.

xxxii



i.e. given U_m and denoting by $E_1 \dots E_m$, $E_{m+1,1} \dots E_{m+1,m}$, it can be shown that

$$P(E_{1}...E_{m}|U_{m}) = (2\pi)^{-m/2}U_{m}^{-2}e^{-N} \exp(N\Delta_{m+1}/2U_{m}) \text{ for}$$

centro-symmetric structures
$$= (2\pi)^{-m}U_{m}^{-2}e^{-N} \exp(N\Delta_{m+1}/U_{m}) \text{ for}$$

non-centrosymmetric structures.

They have been used most interestingly, combined with the magic integer method, as a figure of merit. The determinant is evaluated as a function of magic integer variables, with the peaks giving values to the variables, which are then refined by a novel method^{11,12}.

On the basis of a comprehensive theoretical framework $^{13-15}$, structural information is increasingly being utilized in direct methods to solve large or difficult structures. Proteins, an area resistant to direct methods 16,17,18 hitherto, is an obvious field where progress can be expected.

Other areas of interest include a geometric interpretation of direct methods¹⁹, and the use of weak reflexions in refinement^{20,21}

xxxiii

and scaling²².

. ·

.

References

- 1. M. M. Woolfson, Acta Cryst., A33, p.219 (1977).
- R. Baggio, M. M. Woolfson, J. P. Declercqu and G. Germain, Acta Cryst., A34, p.883 (1978).
- 3. P. Main, Acta Cryst., A34, p.31 (1978).
- 4. J. Gassman and K. Zechmeister, Acta Cryst., A28, p.271 (1971).
- 5. R. Karlsson, Acta Cryst., A34, p.698 (1978).
- 6. P. Main, Acta Cryst., A34, p.826 (1978).
- 7. G. Tsoucaris, Acta Cryst., A26, p.492 (1970).
- J. J. L. Heinerman, J. Kroon and H. Krabbendam, Acta Cryst., <u>A35</u>, p.101 (1979).
- M. Knossow, C. de Rango, Y. Mauguen, M. Sarrazin and G. Tsoucaris, Acta Cryst., A33, p.119 (1977).
- C. de Rango, Y. Mauguen and G. Tsoucaris, Acta Cryst., <u>A31</u>, p.227 (1975).
- D. J. Taylor, M. M. Woolfson and P. Main, Acta Cryst., <u>A34</u>, p.870 (1978).
- 12. "Crystallographic Computing Techniques", Ed. F. R. Ahmed, Copenhagen: Munksgaard (1976).
- 13. J. J. L. Heinerman, Acta Cryst., A33, p.100 (1977).
- 14. J. Brosius, Acta Cryst., A34, p.955 (1978).
- J. J. L. Heinerman, H. Krabbendam and J. Kroon, Acta Cryst., A33, p.873 (1977).
- 16. D. Sayre, Acta Cryst., A28, p.210 (1972).
- 17. D. Sayre, Acta Cryst., A30, p.180 (1974).
- 18. A. N. Barrett and M. Zwick, Acta Cryst., A27, p.6 (1971).
- 19. J. Navaza and S. M. Silva, Acta Cryst., A35, p.266 (1979).
- L. Arnberg, S. Hovmöller and S. Westman, Acta Cryst., <u>A35</u>, p.497 (1979).
- 21. S. French and K. Wilson, Acta Cryst., A34, p.517 (1978).
- 22. I. Vickovic and D. Viterbo, Acta Cryst., A35, p.500 (1979).

APPENDIX E

STRUCTURE FACTOR TABLES

Table E.1

Observed and calculated structure factor moduli for

N-Ac Λ Ala-OH. Symmetry related reflexions showing large inconsistencies are shown below:

h k	c l	e.s.d./F	Δ/F	N
0 11	. 5	0.280	4.245	2
07	8	0.207	0.156	2
05	5 9	0.216	0.189	2
01	. 11	0.229	0.354	2

R' = 0.0096

,

1050 13FC н к 10F0 H L 10F0 10FC R ĸ t. Н ĸ 1 . F C ų. 1:50 L ĸ I. 43 355 12 1 3 5 O 337 396 84 102 181 345 394 1, 7.5 36 ŝ -13 -17 4. 1. Ó Ō Ð. ñ. · .) 8 10 12 0 0 ŏ U 10 10 -i 21. 7 7.) Q Q -15 :0 õ яż i) -14 ·.. Ő. 258 Q -12 ij. Ũ 1.5 1) -10 1 10 1 ŝ 57 58 J Õ 36 -811. 11: 18 1 1 i 15 17 Ű, θ £) 299 295 230 30 4, 537 160 370 ś Ó 5.20 ġ 31 19 247 177 177 Δ. ເ ບິ່ງ ບິ່ງ -20 -16 - 2 2 122/2223330UO ÿ G Ó 1. 5 21,9 - 9 37 12 6 £ -16 Z, E 17 . . ģ ĝ Ê -14 0 $\frac{101}{165}$ 13 15 17 19 176 176 17 51 0 -12 Ė J. -1ū 1 6 7 782 . . 31 51 -20 -18 U - Ê Ż : 1 Ö ć ń ١Ž 979 Ó 5 . 1. Ó ŏ 1757561 21 6 ŏ -16 -14 xxxvii Ũ 70× 2°2 -17 96 Ż. 247 158 191 ÷. Ũ G -15 -13 -11 21 198 -12 -10 -8 -6 1111 Ũ 97 22555352257 841 165 254 5Ť Õ U U ć 248 58 12 14 16 .-9 11527599 12527599 **. بز** . 13 . 15 . 17 40 Ũ 1 Ď -7 . - 4 - Ś Õ ł Õ -202468 100360277 883 n 2 õ G, ġ. 401 18 к Н Ũ 402 380 - 1 ភ័ ព ij, 1258 34 ŏ Ğ -19 ij, ć Ð. -17 ŝ うろろろ Õ žž 12375 Ð 4.1 -15 1.5 ٠. 92 22 34 Q 1 Ŭ Õ 17ら 1Ž 14 1124 Ō. 1Ž 14 22 -11 Ş 51 31 ij. -9 1.8 Ű. ŝ 158 99 - 7 15 606 ö D -19 -17 * * * 3.) 612 445 471 ä Ō Ô, ŽŻŻ 20 -15 Ċ Ō. 152 1558 Q 575 107 -16 \mathbf{C} Ō. ö 203 128 128 -15 -14 ċ 187 0 ð -13 94 Š -12 38 Ŷ, -11 -10 223 Ö ç U -9 272 77 11.7 - 2 ζ. 114 ģ Õ -7 -5 -3 -1 256 738 Q : - 6 o. Ö () - 4 54 62 32 ſ -2 \mathbf{G} Õ Ó ŹŚ 6:19 - 1 i.

i

Sec.

 ${\bf e}_{i}$

LIDFO K 10 F.C -1.0E OF C 79 120 113 157 122 6000000 68 28 9 21535920310 ģ сц Гр 6777 7777777 $\mathbf{\nabla}$ 7777 X X X X X X 77888888888 11118642 1 Ō 88 6 હ્યું. PITI I 889999 12222222 ģ

1250

ಹಿತ್ರಿಸು. ತ್ರಿತ್ಯೆಗಳನ್ನು ನಾಗಿತ್ತಿದ್ದರು. ಇತ್ತಿ ಬಾಗಿಕೊಳ್ಳುತ್ತಿದ್ದು ಕ್ರಿತ್ಯೆ ಕ್ರಿತ್ಯೆ ಕ್ರಿತ್ಯೆ ಕ್ರಿತ್ಯೆ ಕ್ರಿಕ್ಟ್ರೋಗ್ ಕ್ರಿತ್ರಿ ಬ್ಯಾಂಗ್ ತಿರ್ದಿತ್ರಿ ಕ್ರಾರ್ಟ್ ಕ್ರಿತ್ಯಾಗಿತ್ತು. ಇತ್ತಿ

รับร่องไปเป็นการให้เป็นการให้เป็นการให้เป็นการให้

31

11139743

ំដំ រដ្ឋ

NNNNNNNN SALAN

ガデアアプラップ

45.211.569

C 54724

いたいというとうへん

1. 16F0 41090 11 4824 9489 1005727200 8070999 80142298942401688200 11 4824 9489 1005727260 8070999 80142298942401688200 17 COOC 12 2222 UNIVERSITY OF A DESCRIPTION OF A DESCRIP I IIIIII 80,00UN780,1-1-079,00000,000,000,1-1750,870,05728870,7087 11222,1-1-079,0000,1-1750,1-1750,870,05728870,7087 11222,1-1-079,000,1-1750,1-1750,0572,0572,0572,087 5-1070-06666890193527205556005225514221125028234222 5-06210172 11223577205558005 2251422112550282342212 11 2117582578822312 913564 2127711 124 NN 12 70 00 014 222222222 1003 1003 135 с. З Ž al) ting 0000000 802453 **ニン・シャッシンシンシンシンシンシン** 100201208 ູ່ນສະວຸ<u>້ອ</u>ູ່ເປັນເຊັ່ງເຊັ່ງເຊັ່ງເຊັ່ງເຊັ່ງເຊັ່ງເຊັ່ງເ 000000000 CHACE BURNE 6202469755110 11111111110 121 159 306 275 236 1357913 74 i. fr

14257072284

191

149771651

237.570-1.567.693

5201 10F0 10FC L. H 10F0 1 **JFC** E C °Î`Ê.F∙C 1928 899 ຍັນການທີ່ຜູ້ຜູ້ຜູ້ເຄີຍເຫຼືອງທີ່ເຫຼືອງການເກັບການເຫຼືອງການ -5 -3 -1 806004409055628331624518460042345003499977300 1 1155584606350955214231522 46704582346195 335216 1 122344 1 5 335216 KINIMA COCCUDEDCO 8397485070777259141061550910558425465956154545 1-155594605352862142 5212466955589361561 4581 12555989361561 4581 124977 7689518899377553245652545600952652540017157681111 275 3111174576640427664771692707677692789097711168478-611967 579 たちとななんないと 》在 行在 13 444444446 行きないため 990 見たいののはない 13 112112 4444 2 2 2 2 2 2 29. 7 1 43444444444444444444444444444444 11400 ----13× 9 11 15 1 8 2860 10 11 -97 -53 4444444444444444 1965 ちちちちちちちちちちちちちちちち 6066667777 44444444 ちんちちょうん たたれい 12222222222 もなみ 420246 1 () 1 () 1 () ころいてい 444 66 11 -11 56 11

-1 -1 -1

10FO 1755403 893 953 123 275 257 -10 5 07497770778727845012299603381126639364795 ్ణ సనా 1007498-04252713 -647 040+82864428002201080702324167057813 741205522033921 5534080724150372 33 53 **NSN**55 4884420009896575846995577445 1286577354946894522 4885777445 13579 ນທູ່ທີ່ເຊິ່ງ ອີດເຊິ່ງ 66752710799916979630213 1313234481615685352233 15685352233 العيامية -5-3 10 12 14 -15 -13 -19 -75 7990551 10 12 -17 -15 43 46

-13 -11 -9 7777777777777 9999999

まちちもちちらうち

これた

070077077070

-11 -7 -5 -5

°0 ° 0, ° 0,

ーマンマンシン

555

3340

NON-C STANGEON

132 67 199

183

6666

136 75 21

15

د م

11.562287290 11.5621790

いたがたたたたでででででたたででたててたててたてててていたができたが、ことのためではないですが、ことのためではないですが、ことのためではないで、このではないのです。

があるがたちちちちちち

setter und state of the set of the

2.25 WN22 7.470 2.470 2.470 2.471 2.

. برای

(IF O JFC 1070 1 UFC H 10FC '1E0 10FC 1050 6.4 -15319757 55555 241277326934920914675539805417840349656005737996 11 2 1 1 1 55202735299454903&3843077490449510165882157818277 951661426115 46432351746 12242160976234121222655 1 1 1 1 024 5899079585882479589979290032010719279222605687340 221313 3245631 32 1 1 112111132 2327880246 11432 2 3 10 UDAAYA6532116829280424157576999 17 1.1 1.1 11111 1 REPERSE 26455 142265 66267 263233 21 1 243523132147551 7777777 8228 8684758616547267607855586943472545625882022 43 21 2 222235631 1112221 231 15367852622 43 2 3 - 5 - <u>3</u> - 1 221532554 1208642024 8885 4626222766571411641658881 61423122 211425 213115 6591696151405344572134344902226889412638169&7652 0 444 777 1588 22511221121 1210 1 1132111111 00000000 58885 -16 -14 -12 -10 -8 -11357 22112374122 24211 2121 135111559 680 L -1 -1 -1 ----4555555555555555555666666666666 11086420864 666666666677777777777 ・アアアアアアアアアアアアアアアアアアア -i 2845688988 2696 -2:)246807 -1197-531 XLÍV 9 9 0 e ų -153-1975 ġ ġ 986420864202468753197 868888888888888888888 320864202755686420844 19 19 19 19 19 19 78888 236118599029974 116318755 12 217 5 10111 888 7777777777777 88888 6 6 -i አባ ģ 29755 22 777 129 124 124 15 71482051 677 9 -12 -11 -86 -4 -2 000 ò 888888 883388 5999 0 0 0 1 77777 9 61-16-344 ģ.

0.7.0 156 1<u>11</u> N-864 X - 864 X O X 468 X 597 9 X 597 117642572157858540105488860857657 431616664 46 111 ASSNO 6545 778164-0725151930693275178240012978731738871 778164-07256302 62 17241766678043126778731738871 619601744355278483404533415355724421132643 งมันการที่สุดเราสุดเลาสุดเราส -15 -13 -11 6.スアアアアアアア なんないう うちょううちゃ 705X00190290751705X75531455 9 ģ. 14 NO 01 10 0 1a 10 10 10 10 õ 42 41 Û. -153-1-97-53 11 11 11 11 11 :iê 1.9

F ();

33)

たろうと

22863 1 17 17 26 29 1823 1.1.801 S 1.1.38
 11
 8

 11
 29

 11
 27

 11
 27

 11
 27

 11
 28

 11
 28

 11
 28

 11
 28

 11
 58

 11
 58

 11
 58

 11
 58

 11
 58

 11
 58

 13
 58

 14
 58
 153 1259 1 36 61 4 2 i i 255 157 375 298 288 151 21) 1 39 15 79 99 ານນາກູ 11 i P Gran τις τη χει - Μι

21

19

-10.56422

Table E.2

Observed and calculated structure factor moduli for N-Ac-L-Ala-L-Ser-OEt. Symmetry related reflexions showing large inconsistencies are shown below:

h	k	L	e.s.d./F	Δ/F	N
19	3	0	0.274	4.247	2
14	6	0	0.209	0.166	2
17	7	0	0.221	0.401	2
4	10	ο	0.215	0.161	.2
3	11	0	0.235	0.422	2
4	1⁄2	0	0.271	4.407	2

R' = 0.0063

xlviij

н

U (ı 0 Ġ. 6

Û Õ Ő ñ Ö 6 Ģ Ü Ü Ü 6 (ı Ġ Ω 6 0 6 6 0 Ô 0000000 000000

1 5 111

b К 10 1010 112 i. C ١, 6 6 ういたいない

16F0 30.58242158 1922 - 2

4475340194044296432488354599A996309 73173911361969638601445416144634111 73

10F0

10FC 44 10143734111

-3 -2 -1 ٥ -545210

H

OFC -51 -20 121 _1 -1 18 55 |36 |04 7 -1 - 62 - 15 - 37 - 18

0 1 -j 110004 -----

Ō

82 92 -1 62 64 -48 -

10FC

H 34430101037	K	L 10F0 10F0 000 F0 1050000 105000000 10500000000 1050000000000	H 1000077777	L 10F0 10FC 753 - 45 44 - 15 15 - 120 44 15 - 120 44 41 44 41	H K L 10F0 10FC 1 1 2 21 27 1 11 277 125 1107 1 11 277 1107 1 11 277 125 1107 1 11 277 1107 1107 1 11 277 1107 1107 1107 1107 1107 1107 1	H 0-000011110	L 10F0 10FC 38 - 36 3207 2-36 259 - 19 247 - 26 - 58 159 - 19 - 58 - 104 - 10	H 2101234321	10F0 10F 312 -105 109 -121 1296 -121 1205 -102 109 -105 109
1111 111 111 1111 111 111 111 111 111 1111 111 111 111 111 111 1111 111 111 111 111 111 111 1111 11 11	744444444455555555555555555555555555555		・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・	22 70447388623591822381 - 2 732532 81581 - 1 1 1 1 - 1 81352 - 1 1 1 1 1 - 1 81352 - 1 1 1 1 1 - 1 81352 - 1 1 1 1 - 1 81352 - 1 1 1 1 - 1 81352 - 1 1 1 1 1 - 1 81352 - 1 1 1 1 1 - 1 81352 - 1 1 1 1 - 1 1 - 1 8152 - 1 1 1 - 1 8152 - 1 1 1 - 1 8152 - 1 1 1 - 1 1 - 1 8152 - 1 1 1 1 - 1 8152 - 1 1 1 - 1 1 - 1 8152 - 1 1 1 - 1 - 1 1 -		1111 111111111111111111111111111111111		-0-10-11-11-11-11-11-11-11-11-11-11-11-1	101-117-2010-0000-0000-000-000-000-000-000-000-

.

xlix

•

•

.

.

1

1640 - 11 1050 10FC Ti 1686 10F0 10FC 10F0 H . **H** 10FC H FO 10FC 0 11 100000-01-0 10-00-01-0 0 250000 4 **₽**₽7 りんちんり 21 45544 778888888896666666 -32 19 37 3 34 16 35 3 432101234 8770435519300856118188146053717050 237325731417027101135929194 182761 1 1 1 1 555 0146833366571137 43521551628141 13521551628147 1137 1441 455555555 444 44 Ó 50-450-512 500-5 50-50-50-50-50 50-50-50-50-50-50 50-50-50-50-50 50-50-50-50-50 50-50-50-50 50-50-50-5 ようじんしいちょう 904066530 1945530 -11087 -43 -2 -1 222222222 -1 444 Ò 6666666 234430 444 Ó 123 50.54 Ó 1423900858047 111111 11111 10 10 10 4 6 b -135928 66 444 10 4 3333333344444444 16 ------1. 44444 444444445555 t. 7777 777 6 6 1155467551 レイヤタイマディア 4321012 1. 88888888 -118 44444 4444 123.4 3 -18 Ġ

10F0 10FC н 10FC 10FU 10FC 10F0 Ŀ H 10FC 10F0 10FC OFO -28 -41 -81 41 3 014100/202205/5/5/5/54344000/310 0400 11152 121 21122 1122 2000 211 212 213 1111 うっちょうちょうちょう 936202575385554120738757535658555412078843 31632 1177 936202575753855541207887724 843 31632 1177 1177 234 35 222758 3 2362214 822249 2336355 62933874688087228056082713238014 2474 31 1 168033667662 825258 014 21 21 21 21 102 11 555-06-00 1000011111122200001 アアアアアアアアアアアアアアアアアアアアアアアアアア -4-3-2-101 ~~~~~ 175008567659464073524952470 8041570 452420 8412220210 516 552479804192 232101221012101 5812808880907155612000284308 11 089336676621825167 014 11 07 11127 014 666 66777777888888886666660 1111 42024452101 ろんろんん 21111 210-0-0 12222222222 くうちちちちちち

λ

10F0 1610 H κ 1 OFC OFC li. 1680 1 H 10F0 10F0 H 10FC ĸ H 10F0 10FC 34321 77 5 202 180 3 66777777888888886666666666 8 94801 1321798142552031 01141 6468 94801 1321798142552031 01141 6468 89999990000011 624 43627865176010 0 2897325768574030602346465875276385 11 212122 32217271 44 5441-2652 111 344321 13 33 14 32 4444455555555556666667777778888888 こ こううちょう きょういうざい 11 11 11 Ð $= i_i$ クァァァァァァァァァァァァ 1111 147556750366 - NUMERICAN - C 01234 93 68 16 30 019205972845308213 0000 67101113 175024 019205972845308213 452-0-254 -2960 ちょうち いいちん いたうろう ひろうろ ひろん ち 354675267 1000001 **アアアアアアアアアアアアアアア88** 02436711364 61324388 02436711364 61324388 011 1364 61324388 1438549334 しきょうげん i 1 - 1032 41521250 うちちらちちちちちんんんん 232101210 33 -62 -41 000000 1 3132710158 54112653 Ó 7091 1313 1315 1358 1358 122222222 21012 1200 1 1 1 1 0 3 1 4 1 0 9 ...4 ...2 ٥

lii

202201201 110N44

888888999944 ¥

10F0

1010

1 1

ĸ

li

10FC

Н

じんじんどうしょう

0749-85065495287998534852485083242 68883741152091781-41 116 51 210

358-7610578412096070202000001141111

10FC

10F()

10F0

10FC

45555555566666667777788888899996

10FC

-10922071 -222071 -1778754

16F0 16FC 10FC h 1610 Ŀ H FŬ 10FC К H 10F0 10FC 10F0 10FC L н ĸ 1000001 ちちちちちちち ひとし しゅうアイアス たいとうしんしゅう しゅう 4182 4 10333 4652524441 6421 24444 14777777483412 928 3 49731913 6 51 4477777483412 928 3 49731913 6 51 1418214110333140922700587300579274444 66667777788889999000111111122222222 8307674464193434364119414370831107 2372 412372 12 01568530 32 61443 1 1 1 1 3333334444445555666666777888800011 0720059519204112249054120947084198 1 56 62 333 122611253024 411291915 1 **ユードビーちいっちょうろうろうろく チチチチチ**

liv

10F0 16+0 H. Ł 10F0 10FC L 10F0 1CFC Н 10F0 10FC 10F0 -1 0 1 41 32351 21146 324 7871564 8 56 52725215244543 555 66407000 41470005 73434642124031657 1 4246148851898558404 73435642124921640 1 938899246109727395 らちちょうちょうち しし 556660001 728 068415182836281 984082418004114289 25232216921429 221 332 12-5-5-5-12-22-22-24-4-4-4-5-5-101111111111 44555001 うちょうとこと 5217242155244555 1 1 1 1 1 007777000111 2222222224444 0122010 1010 222233 -222334

10FC

-252374955771

Observed and calculated structure factor moduli for N-Ac Δ Phe-L-Pro-OH. The following reflexions were omitted because of suspected extinction or multiple reflexion effects.

h	k	l	10 F ₀	10 [F _c
0	2	1	1479	1648
3	7	3	30	153
8	6	3	387	179
9	7	3	369	100
14	6	3 [.]	335	15

R' = 0.0051

10F0 15.0 н н 1 UF C 10 F C Η 1 FC L 176 102 177 174 055145.85 150887 122214 -21 -19 -15 $\{\cdot\}$ 1ó 18 251 157 257 29 49 11 11 00 00 00 ģ 3365 180 ă ÷į Ų. Ĺ Q 14 R ä 12 14 16 18 à Ö 25-() 2223630566545996131 i_1 -13 3322 78885239 536 2197 147 -197-153-1 Ð ý Ü -11 ů 1 ñ - ¢ - 7 i) Ű ģ Ű ្ល ០ Ò È ų, ā -5 -3 -1 722 664 67462 1 6748198681887 0 0 Ő 774 -9 -7 -5 -3 1241357 133 531. 174 ÷. 10 10 10 10 0.50 U 15 17 11 11 11 11 6328 1) - 058 1049 4697 1897 1611 1907 106 H 0 0 0 17 15 17 () () () () () žóś i) -• Lyii 158764778 121859 ē 17 19 21 - 20 21 ۰ĵ 11 11 1333825 ٦, 774 474 ģ Ż 21 38 47 ž 1 1 2 3569095598557 2 3 1 44147 2 3 1 44147 ć 15 17 2 14 16 18 1312 1312 153 331 170 2788849 314947 -18 Ū č -16 -2086-142 ŧ -13 14 16 -12 000 41 111 11122243964912955 11122243964912955 -i 45 34 -10 -8 -6 28 107 0 Ĩ 13 4 1 1 502520 50252520 13 4 1 1 502520 127 127 22 22 -12 -10 -8 15 17 Õ 72855556586 110 - - - - 224 118 437277 153 167 196 -10 166437410 Ó -6 Ŭ 1Ú 12 14 15 17 ġ - 4 -2 0 2 U) 18 21 9 12 14 16 0 1 1 () Ö 203 431 -21 -19 Ð Ż F Ö

124 176

40 30

10FC

18F0

н

ŝ

S.X.

12.3

i,

7,

10:0 10F0 1 DFC 10FC н Η 10F0 10 F O Y. L н 10FC 11FO 10FC н κ L H κ L -17 -15 -13 -11 -9 -7 1479 14399 1.4 - 8 184 194 235 85 19 948 1442 399 97 164 117 うちちちちちちち -1 -8 95271233772395889124 153443227124644241800 15344322211246441800 408 106 -6 10 ---2.8 138 158 119 124648 Ó ર્ક કે Q ň 55 19 ΫŹ 1 1997 1 1999 1 1999 1 1999 1222 õ 49 iŏ -18 -186 -142 -1208 295 163 257 -----23 -23 -21 -19 -17 Ż 1231 -119 -531 Ŭ 0 0 0 1 Ú 10372929 • İΰ ÿ 22043984 2211814 162 35 17 33 12 101 Ī. -15 -11 -19 -53 -1 · ! · ž 45 138 1/ -19 -17 -15 -15 -11 44 179 - 186 - 142 -6 -4 157 267 485 710 -202468024 22 55 1108.64 18142861 11 2 24 275 480 1111 1111 85 -97531 90 81 11 11 12 34 517 91 3 9 ó 101687 ँ उँ ऐ 83 169 ٤Ó 2 26556231225825 758776839627 - 17531975 - 4 9 9 ģ -2) 24 -2024685331 217 ó 9 9 22887194 1468 8827 117 125 181 62 118 -32884 -32884 10 12 14 15 17 á ç 9 -20 -18 - 5 - 3 - 1 <u>6</u> 5 -200864 ç ç ç -16 -14 - 18 - 18 - 17 18 19 яŻ, - 12 - 10 - 8 - 6 - 4 24412386 7 28 14 00 00 00 -64 á 146 46 19 117 199 297 -15 -13 -12 -10 ý 302 1.8 -2024 à -11 -97 -75 -3 100 100 100 1542) - 12) - 12) 7382 -16 -14 -12 -10 -8 49 137 152 77 738 1921 608 653 693 989 213 316 Ü 50 ? 7 34 Ű. ÿż Ő 8 Ù

-17 -11 - 7 - 5 - 7 lix -16 -14 -12 - 3 -202 12 14 18 -19

-17

-15

. .!

1,

. **.**

н. ĸ

1) 12 14

-21 -19

. -

1060 1060

-13 -7 -7 -3 - 1 ġ 17 - 18 -16 -14 -12 -10 -8 -6 -4 - ż 8 10 14 -17 -17 -13 -11 -9 -7 -5 -3 - 1 ò

н

к

10F0

2.22

-12 -12 -18 -4 -1123 -119 -531 13570.0864 -53 -3 -1 -2208 -2208 -16 -14

H

1 DF 0

10FC

н -12 -10 - 8 -42024 1位 12 14 -23 -21 -19 -17 -15 -13 -11 -9 -7 -5-3 15 17 -220 -18 -14 -12 -11 -8 -6

ĸ L

> くろうろうろうていていて ろうろろろ ろろろろろろろ ろうろうろうろうろう ちざいていていていてい

111FC

11/FO

-4 -2 0 2 21 -19 -17 -15 -13 -11 -9 -7 -5 -3 -1 -22 -20 -18 -16 -14 -10 -8 -6 -2024

н

10FC н ĸ н К L 1.∂F0 1 1 F C н 10F0 1 🖯 F C H. 10F0 10FC н 10F0 10FC 105/60/29/11 -1420-26420 2537 107 77 77 • • • • • -15 -15 -11 192 187 10 --024 18 1. Ê 12 14 4, 290 182 286 189 27 143 24 1/3 Ó -9 ä - ý - 7 - 5 12 **といういいいいいいいいいいいいいいいいいい** 1 Ü 7.6 Ű Ĥ. 745 567 5469 21 29 Ξù 3:33 U Q -3 5£3 2 2 20 2 2 2 **ż**7 - 1 956 247 174 ויטיטווטיי ייטיאר אואטאראטיאט אניי אואראטאטאט ñ -19 ろうろろろろ 130 191 28 43 -17 -15 -13 ġ (ł かいいい 56 58 19 -25 -23 47 . ó -11 4679 -20 -18 -21 -19 -9 -7 うろうろうろう 5564 113778 1190 1590 L -16 -14 -5 -3 -1 Ŝź 149 251 106 434 139 -1208-4 235 197 การเกษารารก うるいとうできょうこうできょうできょう l. 169 x۲ 1 11 3 11 5 11 7 11 1847887544500 14 19位 1173847028 198 59 11 -5 -3 -1 108 17 Q -2() 2 E いこうろういうろう 833121 568 48 ŔŻ 9 72 58 38 21 Ŷ, 31.5
90
11222222233300000

91
11222222233300000

11
11222222222

12
11222222222

12
1122222222

12
1122222222

12
112222222

12
112222222

12
112222222

12
112222222

12
112222222

12
11222222

12
11222222

12
112222222

12
112222222

12
11222222

12
112222222

12
112222222

12
11222222

12
11222222

12
112222222

12
112222222

12
11222222

12
11222222

12
11222222

12
11222222

12
1122222

12
112222

12
112222

12
112222

12
11222

12
11222

12
11222

12
11222

12
11222

12
11222

12
11222

12
11222

12 179 157 9 SECOSS91 221 221 79 ₹ Ż ŝ 50 2266822213 -22 -18 -16 -14 -12 1 ① 1252688 1252688 1Ž 1. -17 2ï -18 -16 ., 11. 266 141 53 71 -15 -13 -11 -9 -7 -10 -8 -6 -14 -12 -10 -8 -6 Ĵ. Q 17269 17264 1555 17264 31 7 l, 24 34 Q -22 13376991 149177701 I, そころがいとうないないないとう i. ý 27 106 53 190 ç. -18 - 53 15 1132 1241 342 85 67 Ĩ 205 128 23 113 -10 - 64 256 -4 L 145 77 ć -2 0 2 - 1 -12 -10 -36 -4 Ú. -219 -12 -17 -17 ó Q () -542 93 115 ちろらてて ç U 162 97 37 31 ŏ 219 597 -8 Ĺ Ó. -2 11 2 -2 322 366 Ŷ. Q **** Ü ι, è ç $\begin{pmatrix} i \\ i \end{pmatrix}$ 34 Õ 341 Ř L 362 i, 4 37

H 1000 10FC H 1.50 1 0 F C 10FC 1GF0 1CFC 10F0 H н Н 10F0 10 F C Ł 222 324 542 - 12 - 10 - 6 - 6 - 6 29 252 69 5 136 239 143 -12 -19 5.5 6 Ú 3.2 1 2 22 56 1 2 22 56 ι. 221 62 239 138 93 147 -10864202 -8 うちちちゃ ι, 2010/02/01 うどうろうろうろうででした 122 122 122 000 -6 75 147 Ÿ 11423529 - 4 -2024 7 ï -202 127 124 171 204 31 5220 31 5220 ι, 4 Å 2 4 1 G 128 172 204 ちらうりち -16 žĹ 223835101 -14 -12 ó ĬĬ -22 -22 -218 -10 -8 Q ó i. 6 70 5222 -19 -17 -15 -11 -11 <u>າວເຄັນຄາດເປັນເປັນຄາຍຄາຍເປັນເປັນເປັນເປັນ</u> 179 - 6 1 () 1. 228 304 Q 0 0 0 -16 -14 -12 29¢ 117 ö Q -3 -8 1948465 11286465 211668 437 ġ. 69 -23 -21 -19 -17 0 0 118 204 ŷ, ģ č 6 ó š ć -10 - i. ちりうちつうううちょううちょう - 9668685 24968685 1585 -2 0 -97-53 Ő Ś Ŏ -15 -13 -11 27 - 6 127 143 229 157 27 140 -4 ζ×1 -2024 ñ 378529350009 - 1 -9-7-5-3 О Ó ζή 7,2 17 8 è ü 219 12 Ű 37 159 145 127 18 Ô Ö -19 124 - - 229 - - 11 14 7 -17 -15 -13 ġ. 0 0 -16 -14 -12 -10 -8 ģ 129 185 ú 6 6 -17 115 -11 -9 87 ÝŊ -7 Źİ ŹŻ á 125.5850 51 -17 207 117 -22 -20 -18 -3 341 225 103 50 50 112 74 111 ě, 136 ŝ S -13 そうちょうちょう -11 -**1** é 17 12223522 12212522 33289 332089 122389 12289 1258 1258 1258 -16 65 161 157 -9 -14 -12 -10 169 159 47 268 268 ĭÿ - 7 69 15 6 - 5 -8 2ÿ 10 7 98 -6 -2208-1154 -9 -7 -5 -3 3.0 いいいちいい -4 -0 7 9 19 24 -2024 t) 102 50 -18 41 9 52 147 5, 104 8 8 -16 -1 ò 42 Ϋĥ

H ĸ 10FC н 19F0 1 DEC 10FC 1 OF C H 1 G F O Н 10F0 Н 10F0 10FC 1387 3012 1490 1490 -21 -21 -17 76 - 5 - 3 - 1 -13 -11 -9 -7 -5 -3 $10 \\ 10$ 138 121 79 Ğ, 69 7 133 124 69 121 322946324 58247724 ្រ 1 68 27 ş **** 29819 14 33 37 Ż 10 1Ō 404 174 いいいい -15 -13 -11 126 1. 169 **いかかかかかいがかがないたかかかかかかなかかかたたちかなかなかないがないないです** 1. 1. 57 180 214 -9 123 441 -20 -18 -16 žģ 1.1.21. 597 274 そうよう とう 24 44 1111 716243221 273854256 1111 273854256 1111 1111 20.6221985118965089 111755 1433759 111755 1433759 -14 11 1. ě. 5 ------1.1 Ĩ4 ሬ Ś 'i i 303.00 1. -23635125 Ġ 2321113 5355544178376828 2221113 5355544178376828 22 1282113 -----32 11 17 lxii 1500051336926 111111 l. いじりょうしょうしょうしょう l, 8. ć, 1277 331 2 1256307 1277 331 2 1256307 11 17
 221331
 2
 1
 120
 1
 1
 1
 1
 2
 1
 1
 1
 2
 2
 2
 2
 3
 3
 3
 3
 3
 3
 3
 3
 3
 3
 3
 3
 3
 3
 3
 3
 3
 3
 3
 3
 3
 3
 3
 3
 3
 3
 3
 3
 3
 3
 3
 3
 3
 3
 3
 3
 3
 3
 3
 3
 3
 3
 3
 3
 3
 3
 3
 3
 3
 3
 3
 3
 3
 3
 3
 3
 3
 3
 3
 3
 3
 3
 3
 3
 3
 3
 3
 3
 3
 3
 3
 3
 3
 3
 3
 3
 3
 3
 3
 3
 3
 3
 3
 3
 3
 3
 3
 3
 3
 3
 3< Ŀ 2435544086252720548 22 1222113 11 6 133 280 146 -16 -14 -12 -10 -8 198-54 У ç ちょうじょうしょうできょうですうから 1562814.654. 146126554. 15 Ğ 6 Ç 114 63 116 118 94 ò 1.1. 155870608427 1518722151 ŏ - 6 771816984377 Ś -202 11956561×10 Ġ ģ 9 -2197-15 59 51 ò 11) 12 14 -21 -19 òò らいいいいいつ ŝ ò 210 264 188 L -13 ٤. Ç -11 -17 -15 -15 -11 -9 -7 -16 -14 Ģ 1.9 -61 162 181 -7531 Ð ころにつうて -23 -21 -19 -17 -15 313 44 235 176 看究 ;) ううううう 18 04 226 32 32 227 46 231 166 i, Ð -12 -10 10 10 6 l, 1% U. 28.1

κ Ł -15 -13 ÿ ģ -11 -19 -75 -1 -1 7 ģ Q q Q 51652 54 325575151722573 911 12045492594705452209533098458 1 12 Q -16 -14208642024 10 10 Ż 1.1 1:1 相引 1332444415 3 1:) lxiii -13 -11 -75 -53 -26 -24 -22 -22 -20 Ð i] 67 150 n Ú a Ó -16 -12086420 209 1555 555 -17 214 1) Ŋ Õ 발

10F0 10FC н ĸ 10F0 1 0 F C 14F0 10 F C 15F0 10FC 10F0 10FC L н H н ĸ Ł -25319 -12 -10 -8 -6 12220860 162 1337 107 9 37 10 $-\frac{3}{1}$ うろろう 11072294 -6 14 À 136 198 - Â Ö -17 13.0 -15%-197-5 -212468 2 07962516721672738 2 1 21 11 36351243 3450397&874552&885 32213 3252 6695 1 139 44 195 61 9 x862344995205145464204008347829422122 1 11311 11312 1 11311 111 1955 1355 1888 16 3&1126966543888840863 12164593312436 22237 -1864-120 -21 -19 -17 -15 -13 7 Ser would show -1) -197531 ŋ ŻŚ ----158 56 7 ÷1 9 9 9 ÷. Ĺ **いいいいいいいいいいいいい** 1331 1386 120 -12)8642024 669 114 93 62 27 11111111111122 ġ 1) () -23 -21 1690 1690 1751 151 69 97 32 9 ູ ---914 5829 376 142522 -25 -25 -21 -19 -17 8 1052074 515 115 522 128 128 1048237562705 1217216585 ころろうできているこうでいろうでいたい 82.88 99 1346 23 23 23 23 ģ 8 ģ -13 Q 9 10 -11 911545-3281 -175319753 Ś -9 -17 -17 -15 -15 -11 -9 -75 -7-5-3-1 197 135 -220 -220 -16 -14 189 143 199 17 9 ġ 159 78 25 39 223 ģ Q Q

۰. .

_ **f**

יג 1 1

Ŭ
н	K. L	10F0	107C	H	κι	10F0	1 JEC	· • •	κĹ	10F0	1GFC	н	к	L	14 E O	1() F C	н	к	L 1	1 () F O	1080	
80219753197531135791086420864208642086420864207531975319753197531135	<u> </u>	2767894449550840890751999298989671291790556134093 27678 93525112582 5 25 788949389671291790556134093 1 1151 5 25 7889493989671291790556134093	2 616 23587369573969142827228070803502735241975923 1 11 11 1 1231477695847413553 5790 304 4 1 11 11 11 1 1231477228070803502735241975923	57642086420246319753113864268642086420256468025 1111	990000000000001111111112222000000000000	2131 714576424131 535261331355522155831564189.23	1277371457532374114 2252111232579 1427525835888 12773257532374114 2252111232579 1427525835888 151 1 1 8811	319753197531975311757913420006420006420024680025319753197 22111111-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-	11111111111111111111122222222222222222	746881484210968049405870046878176484582548668155 10 57850600468049405870046878176484582548668155 1 111111111 1 1 2	1 1 21035784859952526368094600412877942426707225579155 2103578485995536368094600412877942426707225579155		ろうろちごろごろろんん んんんんんんんんんんんんんんかいちょうちょうちょうちょうちょう	~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~	1 1 1 1 2 2 121111 32742565566158620557 587318246217013176847893085969955566158620557 1 1 2 2 121111	1 1 1 5 6 3 6 5 8 0 7 2 1 4 9 4 1 3 0 3 0 0 5 1 5 3 7 4 2 6 6 3 5 5 0 3 0 3 7 7 4 2 3 2 5 5 8 2 8 6 7 2 1 4 9 4 1 3 0 3 0 0 5 1 5 3 7 4 2 6 6 3 5 5 0 3 0 3 7 7 4 2 3 2 5 5 8 2 8 6 2 0 0 3 5 9 6 3 3 6 0 4 0 1 5 5 1 1 8 4 1 8 1 5 2 2 5 5 8 2 8 6 2 0 0 3 5 9 6 3 3 6 0 4 0 1 5 5 1 1 8 4 1 8 1 5 2 1 5 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1		36666666666666665777777777777777898888888888	~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~	258074398576191405780558185896970259960945651237 1111 12 1 1 1 1 12 1	861269990057010836291667158851379957515103362406 1111 11 11 11 10836291667158851379957515103362406	
·																						

.

.

· · ·

1)FC 1370 1050 10F0 10FC ٠H 10F0 1 'I F C Η 10F0 1UFC 16F0 н н -9 -7 -5 10 10 70 13 - 3 - 1 11 40 49 40 49 9 -19 -17 -15 -15 -15 -17 -15 -17 -75 -75 -11 11 11 11 11 11 53 61 67 8 64 64 17 21 71 17 8 42146711512 121313437 52530164 380746286076362 12131313437 52530164 38246444437333 1 555555555555566 1) 1) 1) 1) 1) 1) 77 36 13 11 100 100 100 100 100 100 $\frac{11}{111}$ 13542086420 722086420 11120 1120 1 10 1 6161 7 84389.273353339227 99323423823303879 21 1 1 1 1 511744323 45324734234 3 94 10 832 6151325251454 63655646 716 735359 73 73 16 777 - 1 $\frac{18}{40}$ 11 ġ 1.0 Ŕ 7243 - 11454337141242454431515713331531 62 -18 -16 ġ. 35 279 66 6 ġ, ۵Ň -14 1 12472768342 10 10 10 3.83 16 ~12 24827839788978234204068363450429349 16 14 129 6 235096894819210972221858890898193449 2428928135228922739944 423832205779 1 8 1 10 8 1 10 10 10 10 10 10 10 10 10 413400879 113400879 258999 50 16 1 9545108231 108231 ģ 9 61 323544 Ó 10 134652562454677525684 1 1 9 10 9 10 9 11 1710575748 115775748 NX% -0 **11** 0 11 14 9日11 11100日 111111 11111 11111 11111 11111 50 41 563354567899779 1 254 16322 12531 41 43132562314 5 ġ 18289525445 11 11 11 11 11 11 11 -8 22867717 17 - Ĭ 11 11 11 11 11 -2024 11 11 11 'n 11 Ù. -2319-17531 11 11 11 59 1544 173554 1 11 12222222222 13 39 11 18 11 144227 3845 144227 3845 1845 1278922 11 1 47 39222475 11 6 -20 -18 -16 -14 -12 -10 11 -17 -15 -13 4 12 437 11 752644 1 11 11 44 11 11 11 91 1 -9 -7 -5 1 11 1 11 1 11 153 69 -11 28 Ū -9 -7 -21 5 ģ 11 6

1

Ŋ,

1

• • •

1 UF C 1UFC 10F0 10FC 10F0 H I. H н Ł 1 G E 0 н 10F0 19FC H 11.FO 1010 1 Ł 178 27 38 74 -21 -19 -17 -15 -11 54 11 52 10 -2024 57345 39 36 6 35 50 4 99 71357521250246331548153962396760 71357521257607724607822118611674 1 1 0 29 23 18 7. 4 L 80 599232740 13740 13740 1414 104 1() 321236432 10 ò Ŀ 10 10 10 10 ģ -22 15 9 37 9 9 2222222222222222222 1. Ò 4 ý 285738 4 13657355262152464 İŎ 5 Ÿ 53727943804 1227943804 1227486504 -18 6 8 8 åğ 18 1328301 9 63 9 435178802289 5688194624554145029084554664158 11111 11111 11 1 1. Q -16 10 l, 36 3036 ģ -14 Ó 10 9999 ģ 10 10 10 10 -23 -21 -19 າປະການການການການການການການ -12 - 9 - 7 10 -108--42 Q 388971281351 3261158884741327 1 21 1 21 1 21 1 21 1 Ś 10 10 1 ģ 257617026290905907 17344078211266 42 1 -17 - 3 - 1 9 Ġ ģ 10 531611514219 1 99 -15 -13 -11 10 10 10 10 66 49 31 99999999999999 Ŀ 88389 ġ 1 737211698212597671558723315976 11 1 11 1 9 9 9 ģ 9 10 357 10 10 10 -9 37 6 34 くうさいううう うちょううう うちょう ゆん - 7 ģ 63 81 24 -231975319753197531 1 ģ - 5 - 3 - 1 Ò 99999 NNO:NNNNNNNNNNNNNnmmmmmmmmmmmm 9 Q 10 10 10 10 10 10 10 10 10 10 384991067579184066 1 115343532116 28516084505822855565577687 9 9 - 8 Ś 10 0 999999999 ġ 352 27 -5 -3 -1 うろうろうろうろ ģ ġ 885449 5 6 -21 -19 -17 Ó 44 999 -220-18-16 1010 8584783641279100059 8115783569157273558 1 1 21 1 9 9 6 10 10 ģ 666 ġ 9 9 -14 10 333 25 10 10 10 10 10 10 10 10 9 3100 130 28 132 18 13 -6 100 100 100 000 000 -12 66 10 9 167478248076565822037 1 17759248076565822037 ġ ς -10 -8 10 -5 -3 3 ġ 40 - 8 6 Q ģ ろろう £ 9 6 9 6 9 88141253408618584 141253408618584 12467769 1 6 6 ģ ģ 7286135628698 136055628698 1315436932 6666666 73849368258258788 1 31 11 1 31 11 ģ ģ 6 898350945570141 9 Ê Q 3 0000000000000 10 10 19 2434554355 9 č 7 ģ 99999 6 -21 -19 -17 10 ģ Q ÷ģ 7777 キャート 10 10 10 10 10 10 10 10 Ċ -7-5-3-1 -2 0 2 0 10 10 10 10 10 10 10 6 -15 -13 -11 22 Q 6 ý 40 ちんらち Q 819359 36 50 57 77 99999 9 6 7 6 12 10 1 35 - 22 -19 -17 -15 -13 -11 31 28 49 -9-7-5--1-1 ģ 7 10 141 137 Ĉ 1 10 4 7 9 ええい 7 6 10 1.1 7 10 -6 137 34 1 99 $\frac{1}{7}$ -2š -23 10 10 37 Q 57 58 1 16 10 93 10 ..

lxv1

1070 10FC 1 3FO 1 3FC 10F0 10FC 10F0 10FC L н L H 10F0 11 FC L н L н κ 1 $\begin{array}{c} 0 \\ 144 \\ 0 \\ 144 \\ 144 \\ 144 \\ 144 \\ 144 \\ 144 \\ 11 \\ 144 \\ 144 \\ 11$ 52930560100820406 129305312820406 12821668 129724315796398 17422631 1 452 253337213 253337213 44121352437413 34365111153113 27363 3232590487 ろころろろでできょうかん 51352042105750 17713361 52552177224626 4555555556666600 34265221153113 -142086971531978 -117531978 -12()&675319742() -115319742() 11111111111111 7461234242233 1222222222222 1 1 14 1 14 1 14 2 14 2 14 (, (, 14 -16 14 66

• 1 . :

ς.

Table E.4

Observed and calculated structure factor moduli for N-Ac-L-Ala-L-Ala-OEt. Symmetry related reflexions showing large inconsistencies are shown below:

h k l	e.s.d./F	∆/F	N
4 10 0	0.202	0.285	2
R' = 0.0146			

The following reflexions were omitted because of suspected extinction:

h	k	l	10 F ₀	10 F _C
0	2	ο	1384	1646
-2	о	2	1245	1352
-1	1	2	. 895	1000
ο	0	1	766	. 865

lxix

,

8 Ő 0 ð Ō Ō ğ ñ 0000 ņ Ŭ 0 0 0 0 ö U) 9 Ú 0 n

178 7

н

ň Ō 0 ð ŏ Ð Ô Ú Ü 0 0 1 1

10F0

10FC

L

K

0

0 Q

ñ

ā

Ò

Ŭ 0

ā

6

-42224 Ś 10 1357 くこくろうろ

H

ĸ

L

10F0

10FC

g Ũ Ű ŏ 0 0 0 Ũ Ü 0

15 -164 -120 -3 -64 -20 24 - 1 11 13 15 -14 -12

5 102 145 -153 -179 -753 1 357 ġ

13 6

-11 55 5556

in the Ā 4

55555

333

- 7 - 5 - 3 - 1 357

Н

κ

Ł

10F0

1<u>29</u> 57 78

10FC

Н

-10 -8 -6 -4 -2

63 102 - 124 - 119 - 75 - 31

91320864207468

10 12 - 19 - 75 - 31 - 13

κ

6

6

6

6

6 6

6 7

7

8888

883

ý

9999999

L 10F0

2 2 33141257753305054121118168064231228002932

19

9

10 F C

64738432040 14738432040 1466721554

2419536

10F0 10FC н К н 10F0 19FC 10F0 10 FC Н 10F0 10FC 10F0 10FC Ł Н -11 -9 -7 115 295 -14 -12 -10 -5 -6 70 77 62 154 81

 12
 46351
 1
 12249151
 1
 1212450

 12
 46351
 1
 12249151
 1
 1212450

 12
 46351
 1
 12249151
 1
 1212450

 - 7 184256224 18556224 21) 21) いいいいいいいい こここことというという 26 12 15 17 -7 -5 -3 -1 - 5 - 3 1 ŹĨ 179 -9 12246174941327174191097 1 2349251 -10 -4202 357 21 34 49 25 12 25 25 12 11 37 269 12 12 12 12 -4 36 7 152 159 119 59 119 51 41 Ö ģ Ü เพลงพับของการระบบการระบบการระบบการระบบการระบบการระบบการระบบการระบบการระบบการระบบการระบบการระบบการระบบการระบบการ 1 21311 21351 224311900168771273851 1 21311 1 21311 1 21 1 21311 1 21311 1 21 1 21 311 1 21 311 1 21 -1103642024 ð 1 Ü ŝ 1() 7 035451576561222350175895964023933 151112 62 514523971086311241 そうかいごうちかかかいかい ろうざいてきたいですのですのですが、 125-60 590 113 515 152 - 1 474152121610512521221111212113 202 62 653 Õ ò 278 120 17 34 8 -13 113 107 84 6() 0000 -6 -2024 72 11 12 12 11 -11 -9 -7 -5 -1 15 Ó õ 131358572897 1113971497 14 - 17 - 15 ģ n 1001 1061 5289 29 ģ 134 368 1133 1955 1108 -1420 -128642 -128642 -16 -13 -14 -12 - i ī ģ -9 -7 -5 -3 -10 -3 Ū 137 331 1107 1326 326 385 108 108 108 Ö 5 970 2 70 1 79 6 7 7 7 7 7 0 1 7 9 7 -6 6 6 ŭ n 2 ġ 22 28 8 3 5 28 28 25 ie 2 () 10 Û 145210 200 Ū ið ó 122 18 12 30 12 -15 -15 -11 20 31 48 115 14 -17 - 18 - 16 - 14 13 15 -16 22 105 -15 27 īž - 12

. -

lxx

.

lxxi -12 -10 -8 -6 -4 -2 14 -17 -15 -13

<u>.</u>-

-10 - 6 -2024680747531 11247531 -1153197 15

H

ĸ

くろうないでのうちろうないのです **おおろろろろろろろろろろろろろろろろろろ**ろろ

L 10F0 ろろうろ

24232823111

-142 -108 -4

-11 -9 -7 -5 -3 -1

Н on an in stroom

うちうちちちちちち うちょうちょう

うていていていたい

10F0

10FC

н

ĸ

L

10F0

1UFC

Н

κ

L

-2024 -13 -11 - 9 - 7 - 6

うちょうちょうろう ろうろうろうろうろうろう

- 11408642024 102 -175 -153 -153 -197 -5

- 2 0 2 4 1Ŭ 12 -17 -15 -11 -11 -75 -75 -1

うろうろうろう

Ō Õ Ġ L

10FC

-3-1-1-3-57

н

K

10FC 324

1211 13122335535605100252144908028

10F0 10 F C Н 16F0 10FC Н 10F0 10 F C Н 1670 1: FO 33 20 33918559 **ຑຑຑຑຑຑຑຑຑຑຑຑຑຑຑຑຑຑຑຑຑຑຑຑຑຑຑຑ** Κ L 10F0 10FC Ħ 9999 37 17 1357086420247531186420864202468 μ K 7 9 412211716275833111335681111552801932674623941599 **ヘ**᠄シンタレをごろうたちからいです。 32 154 555666 30×9823824527711043052241089594535262275367689 11441133773676019717112665011132952122275367689 111111 2 11 2122222705297941 311 3124356755297941 **らいこういうべいいいいいいいいいいいいいいいん チャイナイ イ** 02468 112113122211705435753777212618417028565092362083277 2532200880271537537613249671570516971114234972 21 8121 21 8122 210 12121211221111111247462333627613143901854084719241223 46753113364202542024850275319753113579 10 *ᡧまれたれた*たいのどうででです。 いっていいちょうでんたい いちまんでいい いちょうちょう ちょうしつ 54231252933356325225386729799 14231252971839554211 195111 11136179758344 224478111153 1124386321 211 234 266 041323550718395648121 961124820122112421111111 041323550718395648121 96112482012211242111111 ちるちちちアニアノファアアアアアア いっちょうつ 444564444444444444444444444 10 11 11 16420864202468053197531135794208642024 1400178516538741109884082994130414323068560094 10 . . 10 10 10 10 10 10 10 11 2112111121 23745221552761323390293013643 3 127531975311357915420364202468075319753197531135 11 12:070 lxxii 3332495704705050838008063435 21 2121 2121 2121 : 111 4 44447444444444444444444 1 1 1 452 12111252 11 11 111 · · · 1Ŏ 4444445535555555555555 1 -175319753119753119753113579113579118 1 1 i 1 1 · · · , 11111 1 1 11 12222222222 68197531 -197531 -197531 11112 29 189 10 10 10 10 10 4444 5 5 11 16 12 18 19 113 · · ·

÷

÷

÷

19F0 1070 H 1úF0 10FC 10F0 10FC 10F0 10FC H 10F0 1ÚFC κ Н н κ Ł 118 78 45 73 127 -1 25 200-0882287486056880159846659011185800051-11175428265574860568801598465590011185800051-186828274800005117 123473623144352312234343073446461522672916726220 22212121111 とういん いってい とう こうい 66566 5 -108 -64 -2024 77777 373268207078519341998888897 46977243431 88643121134551 1 - ś 14 12 10 - 5 3533420253364203642086420846888783197531 777777 6 1357 6 153 <u>63214435552649020801777229211747277344239488</u> 6 66566666666666 34294100426404996357814719190502287752 11 126 2316428443313935442234932 2350338 2 2 2 2 31 122 21 -1753-1-531 6 - 13 - 11 - 97 - 7 - 53 - 1 211 36221163893434318945441 2839281385233 2 2 2 2 51 122 2 2 2 2 51 122 2 2 2 51 122 777777 135208642024197531 1197571 うろう うとう うくくとう しょう 56.06.06 6.077777 9642086420246653197531 44 6 444 3373304743805420440213 113452 1221 1248362 2 2 2 44 77785806588066899999 ちんしょん -12 -10 444555555555555566 1 13864264208642 7777 6 6 -1642086 222222222 7777777 いっしつ いち -15 -13 -11 -9 -7 0.045 -14 -12 É 12 29

н 16

txxiii

	н	ĸ	L	10F0	1070	н	ĸ	L	10F0	10FC	H	κ	L	10F0	10 F C	н	κ	L	1070	19FC	н	κL	10F0	10 F C
Xxxix	U24687551975311357642036420246875319753113576420	000001111111111111222222222222222222222	ຑ <i>ຠຬ</i> ຑຬຬຑຎຎຎຬຬຎຎຬຎຎຬຎຎຬຎຎຎຎຎຎຎຎຎຎຎຎຎຎຎຎຎຎຎ	474262531591947192165664557155784691547575321354 15912144848494911121226106283863212501248875321854 11 1 1 1 2 1 2 2 6 1 0 6 2 8 8 6 8 2 1 2 5 0 1 2 4 8 8 7 5 8 2 1 8 5 9 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	1059 11447485012 22226100333963112501248375311354 11 12 22226100333963112501248375311354		4444444455555555555556666666666777777777	ຠຎຠຏຒຬຎຒຎຬຎຬຬຬຎຨຎຏຬຬຎຬຬຬຬຬຬຬຬຬຏຬຎຏຬຎຬຬຏຬຎຬຬ	636315441751516654232241335113123211322720678325 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	\$4\$\$1543175152\$654233242325113114172312911926966406 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1		99000000000000000000000000000000000000	ଌଊଡ଼ଡ଼ଡ଼ଡ଼ଡ଼ଢ଼ଢ଼ଡ଼	216163577200426248%492%0160202120032717578357139761 1111184657342027361990536715545357139761 117345357142	215173216353123303617753130991119724908024629194427 1151735121123856573420272539999435715545357289761		44444444444555555555566666667777778388800000000000	11111111111 100000006666666666666666666	1847756725251721251542825298748889773382952760575923	18477667234537139236145363222533221223		11111111111112222222223333333333344444444	92251922232476053022142158495111234815111145364211 1111 1111 1111 1111 1111 1111 1111	028416298525457711321585951133749251 45364133 2224212 7255457711321585951133749251 45364133 111

•

•

. •

-

-

•

. ·

. 1

.

1

, :

- 1

. 1

•

. . .

%xxiy

• '

γxxγ

10F0 10FC

22224 131322

122547961854 3 21

ĸ L

10FC

87723674673 231421673

1 U F O

126312223241221

10F0 10FC

31 37

35235133

139180396828 16125221

550000112222 0001112222

186256193126

10FC

127783242533

1070

10F0 10FC

143222222336

1222222222

-12086420 -1086420 -119

133 1321332

Table E.5

Observed and calculated structure factor moduli for N-Ac Δ Phe-OH. Symmetry related reflexions showing large inconsistencies are shown below:

h	k	l	e.s.d./F	∆/F	Ν.
16	3	ο	0.213	0.109	2

R' = 0.0114

10F0 10FC 10F0 10FC H H H 10F0 10FC 1/CF0 10FC 10F0 10FC ĸ Н н 104 293 738 10 297 291 145 291 -293 -142 11 0 1 107 8 0. Ŭ 276721577 572 200 -573 13 139 -16 129 -188 -79 -64 3333 ٠A 1 ā -70 -70 -42 254 100 154 64 17 8 260 44 396 496 -153 -142 Ö 36 382 512 170 139 -99 678 118 56 158 67 ろろろろろろ 73 6 ∵1 89 58 89 19 -7Ū 85 -49 6 245 1 -6 6 123 123 123 107 107 17 128 -89 -76 -24 Ū - 10-42269046556150 - 10-42269046556150 - 11-72269046556150 34 59 175 ĝ £Ì 00000000 -10 1 6 Ö 10 -91 -28 100 37 -14 6 68124025516547683 1 - 11-5817273299 - 11-5817273299 - 13-1 - 13 6 17 18 ğ 11 0 Q 6 1 10 31 21 Ű 8 122109876543210123 103 10 6 1166 21 432 324 312 312 ğ - - 43 2 -11 - - 13 3 1 -- - 21212 - - - - - - - - - - - - - 21212 9 11 1 6 4303142233843432271311 11 10 1 2 3 105 139 -11 -10 12 13 12 3 Ĩ4 2341212 - 16828 - 45227 - 178 - 178 1 6 Ō 1 6 Õ 55 188 111 78 Ò -ÿ 174 318 1084 11544 322 111 1 6 1 - 8 7Ó 6 Õ ŏ 79 -7 6 38 4 -6 -5 6 187 187 187 187 187 187 341 181 141 11 56 58 6 Ô Ö - 4 -2-10 15 42 Ö -2 -1 á 8 ą 00000 0 6 19 178 Ż 302 120 416 474 267 34 10 192 145 Ŷ ć -27 -6 113 Ö -139 19 110 6 84741222 121271471 23574805 1545 123 12 -405 469 249 -31 47 60 31 166 102 109 725 48 20 6 -56 107 299 -167 2222 000000 6 4 1 5 Ĩ4 ١Ċ 6 4 -13 222222 6 4 6 5 -12 ÷Ö 73 -21 17 Ž ě 7 īŚ ŝ -113 9 ٦Ĉ L -11 70 50 145 32 -6 67 51 -30 -56 17 -91 -10 -9 :Ö 0 23 8 8 Ő ĝ 149 ğ -8 -11 0 - 158 - 158 - 158 - 186 Ö -13 -12 -11 30 20 36 1ġ _9 84 117 89 109 392 -39 10 112 - 12 - 10 - 109 - 7 4 ٠O 80 5 -2 -3 10 000000 61 -8 -7 39946729297 5 - C 98 59 -457 471 Š -91 1 1 30 38 ۱Ĉ 11 12 1 -72 -108 -11 1 2.0 Ś -3 -2 -1 j. 65 -6 148 -144 -78 371 281 565 73 154 54 103 Ő 1429 8809 25362 18372 18372 194 35 590 324 982 333 -3 Ó -117 57 -117 -219 -61 -52 335 277 Ü ĕ ŏ -2 -1 -90 -82 -333 -988 67 178 Q 1177 6077 6077 657 3298 2Ŭ5 51 498 144 -189 -6 -5 00000 ٠Ĉ 240 39 75 106 107 107 145 337 499 201 176 -49 į ろろろろろろろ Ó 281 146 - 39 - 107 - 105 104 -491 -4 -45 -148 -739 -22 364 -3-2-1 -51 501 -195 -146 23 5 ٦, 639 Ö 1 -180 6 33 Ö **4** 5 -5 37 361 10 -186 8 Ũ

6

lxxvii

10FC 10F0 н Н 10F0 1 ÚF C 1UFC 10F0 1.(F 0 10F0 10FC Н Н 10FC Н 175 1901 2900 2503 2503 15 16 144 95 19 28 -128 -44 124 436 -317 -1098--654 C -9 -8 -7 -104 -135 -145 -58 00 C 6 7 8 2222 1 7601792786545757829897311174999578340210 1543424222233191136814522263410428432423 107 135 151 6789 いちちろうろうのできょう 3333 -58 -111 30 -140 -88 10 4 11 ō -6 -5 - 4 61 ·13 ·12 -91 10 15 -38 61 11 うろうろろ - 36 - 46 - 103 - 4 165 - 28 -12 -12 -11 -10 -8 -7 ζż 3 -3 -2 -1 14 212 241 -28 72 74 112 76 116 -3 -2 -1 ·10 -194 26 14 31 17 9 ž -8795 -1775761957 -11305197495738930 -14127350195738930 -4127240 210 218 -8 -7 177 Ó 191 104 185 יורורור -69 29 96 403 Ó うろろろ -6 -5 -4 123 128 275 157 80 182 123 -92 -95 -557 -557 -144 147814040405629171812099 32141 1 113302379 555555555555 12752441 1262441 1262941 1264941 1264941 1264941 1264941 1264941 1264941 1264941 1264941 1264941 1264941 1264941 1264941 127941 127951 127941 1279510 1279510 12795100 4255324295130 -48 ž -52 -52 139 11 ------3 -2 558 414 אנאנאואואוא 4567 シュリ 6 7 うろろ - 1 62887 1987 1997 109 -120 -1652 -177 -172 -172 -172 -175 -1204 () - 39 - 55 - 16 -144 -66 192 -275 -44 72 111 ğ うろうろうろう 10 11 ã 567 10 lxxvii 4567 11 -11 -10 16 -66 11 12210987654 -7 45678 -9 ĩğ -2 1 161 -174 1772925172506352324192632 232 3 31221 12632 3 6 145 145 160 20 105 -151 -168 -8 -7 -6 -5 -4 33 44 31 666666666 109 -107 10 68 16 -61 4 うりりょうろう õ 11 -6 16 -51 197 -329 -219 12 31 10 できりしょう 4 1039 -81 -838 -3 -2 -1 11 4 61 80 68 -11 Ó -3 -1() -9 -8 -7 -10987654321 ž 3 Ó 6 -328 -328 -37 -37 -37 -37 <u>?</u>1 6 123456 6 0 152856417 3 8756417 2 8576417 Ð 649 1393 1586 1258 1986 188 188 3 6 2222 318 4 3333 333 31 37 -2101-232 220 -12369325 43222211 ふろろろろろろろ -216 789 NNNNNNN 140 Ó Ú 1 Ż 41 22222 531 1Ú -9 -8 -7 2.00 9 251 ġ. 10 11 -12 -11 234 2345 1 -64 5555 88 36 24 460 118 26 () -11 -10 39 84 13 1 8 Ś 1 1 -6

Н K 10F0 10FC L H 10F0 1UFC H 10F0 10FC 1 CFO 1ÜFC 10F0 н H 10FC L -1226 -106 -106 -26 -157 -104 203 116 30 17 17 -173 -123 -37 -1 -2 211 111 254 717 -251 114 244 -691 -179 12 14 112 236447313047744343151771113742508059139 1534 2 2771113742508059139 235 16 16 17 -11 -10 -9 -8 -7 Q -224 1625 3180 3150 460 ż 1 Ó -127 -276 -176 -178 -178 -178 -178 -178 -153 -453 -453 127 202 311 36 10233769011113959 7 -6 88 27 94 89 7 -5 -4-3-2-1 54235 111 111 111 1 1213 5 112 1 1 1213 5 112 1 1 1213 5 112 1 1 1213 5 112 1 1 1213 5 112 -130 -33 116 - 9 -35130 -5430 -1720440 -17204440 -12454545 -145545 -4153777 -4153777 1088 1021 2054 1009 1609 1609 - 13 - 12 - 11 - 10 - 242 265 265 - 142 - 143 - 143 - 143 - 143 - 143 - 143 - 143 - 113 - 11 - 43695 - 76723 - 1113556 - 113556 - 135215 - 195 -121098765432 -1 111 11 1222423768311112224234142 Q 116 386 114 173 -13 -13 -12 -11 -27 -44 117 -150 -170 1208 -208 -208 -388 -388 87 ñ 1675111521122450204874 1 1 111 -10 -9 -8 --39 -63 77 -161) 112 111 -51763333725401744317 2 -- -- -11 1- 1 -74 -75 -83 -65-52-1 -61 -127 -46 137 50 50 108 -27 -107 -9 -13 -12 -11 -10 -13 -77 -24 -498 ğ -13 -108 -6 20 -47037 -1228 -1228 -2228 -2285 -23567 ころろろろろろろろろろ -7654321 -4 5 -4 -174 - 14 -13 -12 -28 -24 -64

lxxxi

10F0.10FC 1.CF 0 1 (FO 10FC 1 [] F C 10F0 10FC 10F0 10FC Н ĸ H L Н κ L Н Κ L Н L L 22944 2944 2256 -10 -119 -103 -22 -102 -66 214 27 17 -12 -11 -10 -9 10() 36 111 2028744221937144277325552332211512256631331511767 22 1 211 22850 4609 88 11164 ----うちちちちちちちち 6 6 6 6 シンシンシンシンシンシン 6 6 6 39 170 -8 -7 146 86 Ō 159 159 159 29 29 29 216 6 - 46 - 169 - 188 - 188 - 188 - 188 - 188 - 100 -1 -2339 -2939 -2939 632 1369 -114 19 109 191 227 50 111 104 6 1886189701 218789701 うろうろうろうろう Ģ Ò -61 119 -7654321 213 213 -61 -99 177 25 . 31 Ó - 686399 - 2736686985223384982180726 - 2736618869852233884982180726 - 11999180726 - 43702 55359-7235227 - 1 112 - 1 44 130 20 19 -11 -10 -9 -8 -7 Ĺ 41 -6 -5 -4 - 4977106427598014 - 2276014 - 2276014 - 13 - 12 - 12 - 11 - 10 - 8 - 7 ģ Ó Ĺ 191 288 815 877 ż İ, 7 23 517 54 127 115 601 437 . 1 Ó 98 - 593 427 -9 -8 -7 -54 -161 166 504 251 -11 -258 56 Ó - 6

1 .

•

. . . lxxx

.

.

.

.

10F0 10FC н K L н 10F0 10FC 10F0 10FC ĸ Н L H 1. C F O 10FC 10F0 L H 10FC -564 -367 309 596 375 -54-32-1 56 37 84 39 123456789 -403751233 -13751233 6 0007110012172846248357636123 2223622224284163631100161857 1 1 31224234 -35 11 888 - 64878035498306391 - 1125498306391 - 125498306391 12 34567 1277U1301059886 1353333886 777 7 3116 666 -87 7 90 ĨŽ À 1 192542592110261807260070009179368478058 1925425921026180726514797550807009179368478058 1712222211797550807009179368478058 -106 -142 -61 108 140 Ò 6 6 6 1 8 1 ğ 13 89 23432109876543210123 -11 -10 -9 -8 -7 89 118 56 24 245 -13 -59 334 -94 -102 -102 70 10 11 -11 102109876543210 808888 -6 -5 -4 19 44 104 106 1222 70 70 4 18 224 108 4 ----26578141690707075375783592876296970 -121932556-615112243556296970 -1511243556296970 -115335 -154 -101 307 888888 37 4 1652212071817080992159 2311112 1 Ĺ 44 214 44 -44 -74 ž 296 137 8424511548946854551055559908 114332234310114493824162212 11 1 rxxx. 49 888 -18 100 -792 11285 -1057 -50 -10 -50 -16 5678987654321 77777777 8 888 いいいいいいいいいいいいいいいいい 825289122425541 117743482541 119 888 567 6789 89 888 5555555555555566 77777777 1210987654321012 88888 Q 1032109876543210 10 -12 -11 -11 -9 ころろろろろろろ -21 -174 -12-35 -22894 -12-35 -22894 -12-35 -22894 -12-35 888 Ü 97 128 17 108 27 38 11 57 20 012345 8 777777 1 103 682 311 ろうろうろうろう -----6 - 84 888888 -5432 -61 126 140 61 43 67 77777 6666 103 19 -7 -6 -68 1 Ŏ 1 44

1 (I F O 10FC 16F0 10FC 10F0 10FC 10F 0 10FC κ 10F0 10FC ĸ H ĸ L Н Н Н κ н Ł L L -29 -27 -41 158 -5 -4 -3 -2 -1 28 88 106 -20 -87 114 25 49 -21 -124 -84 32 20 -186 <u>9</u>. -25 -10 -1 -5 -4 ろうろうろうろう 13959 215 215 Ξñ -3 -2 -1 ģ -67 ģ 18 -------6 6 -667 - 12 - 11 - 10 - 9 - 7 - 6 -126 -117 185 103 67 19 114 ģ -4321 -4365 -3329 -3329 -120 -120 -120 h -103 67 44 10 -10 Ġ 5 5 -10 -10 -10 -18 -36 9 19 -13 -12 -11 יועוא .7 -23 14 203 175 103 18 106 - 54 - 17 - 1 04 ģ 10 -233 ġ 1 ž - 192 - 174 - 174 103 322 2222 19 -10 -9 אואואואו S うりつりん - 5 - 4 - 8 66 -13 27 -387 -177 - 7 -6 -4 1ŏ -37 -16 Ŷ, -3 -2 -1 ģ. -Š -4 78 17 -7 -6 -5 -4 14796748440561 19652 39 -49 -15928 -188894 -15246574 -152674 -10 -9 -8 -7 -6 -77 -4 ī ģ ģ. 11200 אואואואואוא ģ -6 56 -32-0-123332--11--11 - 1543 - 2722 - 1331 - 129 -3 -2 -1 ģ -20 -12 -29 18 Ó -4 j 120 75 17 242 ģ -3947 -227 -227 -234 -246 -146 -1562 -1166 6 -----10 10 ģ. ġ, 272 30 113 ģ , xxx 9 9 9 9 () オーマーマー 10 10 10 10 10 10 Ĺ 18 ñ Ë 163 18 -3 167 -7 ------158 -145 66 125 -17 10 10 72 161) 152 67 ğ 1245491217956926 146214654 -1(i -9 Z - 815 - 222 - 27645 - 195048345 - 1950483455 - 2-57 - 77 21 15 68 10 10 10 Ō 10 10 10 10 10 Ż -8 -7 ğ 2259481 2159481 Ö 118 17 174 121 126 -51 40 191 -26 -110 20 21 21 -12 -11 -10 1Õ ころろろろろろろろろろろろろろろろろろろろ -6 -18 -3 25 100 - 5 -4 -3 -2 -1 iÒ 9 9 9 9 -87-54 いいいいいいいい 71 12 11 100 -36 -128 -138 iñ -44 -102 -9 -8 ŏ -400 255 255 255 255 255 221 -7 10 10 - 145 - 183 187 72 204 ü İŬ -63 ģ 156 105 105 44 21 -3 -2 -1 -121 23 -245 -173 32 19 10 10 ż Ö 62 62 9 33 -26 -38 -3-2-1 1Ũ Õ -12 10 10 Ú. -71 - 4 - 5 - 6 0 -350 50 19 ş ģ 16 -11 -13 -12 -11 3 10 3 10 3 10 3 10 3 10 3 10 -10 -9 -8 -7 -19 29 21 27 ίŌ ģ 6Ž Ö. -11 49 -25 12 21 70 39 ī Ö İČ -17 -9 -50 127 -127 -5 -5 -4 10 10 53 29 -56 -11 -8 -7 -56 71 -67 20 19 ģ 9 9 1 1 129 42 -6 -35 ίŭ - 6 -6

lxxxiii

ť

ί,

5

۰.

10FC 10FC 10F0 10FC 10F0 10FC H κ L Н K 10F0 Η K Н κ 1 C F O H 1.0F0 10FC L -3 -2 -1 () 1 232 420 19 18 18 81) 7(i 18 11 4232919313141395901995227290617831658537962622696 1 2 31 2 13 2 28 1 2 31 2 13 2 2 2 2 -434552269 8109876543210123456709876543210123456 いち かいいいい いいん いんしんしんしん しんしんしんしん うろう 252111118138351111316222221918361171733332231114720

 +
 344321012109876543210123456782109876543210123456 0+1186104998564619255931763017 222244338111316456114460126227 11121 2 111 782109876543210123456781098765432101234567098765 66666611 4821012345687654321012345678 444 11 11 11 11 11 11 61294393110030001884111556538188819870685753467 3291394422224222528272673341382398206685753467 4444 -12 -11 444 1 1 455555555555555666111 1 11 1 1 123456789 11 111111 444 4444444 2.C 19 1111 13608925808157289 13729415808157289 15398157289 119357289 111111 445555555555555 -----11 11 11 11 1 111 11 89 11

9.xxxiv

.... :

ł

ĩ

Ó

H

κ

10FC -8615249 -25249400 -114005218 -2595218440 -395218440 -26840

 1

123457654321012321098765432101234567210987654321

Н

L

10FC - 47 - 47 26 - 12 - 16

 181
 581912680103417806268888910

 181
 5819126822225135861615656

 181
 1819126822225135861615656

 1980 - 3-7297 - 73578 64 -27 -31 -32

н

11111

L

 12141022113031110349671999486466662223936682231

 - 210622143031110349671999486466662223936688231

 - 210622143031110349671999486466662233936688231

 - 110914

 - 110914

 - 110914

 - 110914

 - 110914

 - 110914

 - 110914

 - 110914

 - 110914

 - 110914

 - 110914

 - 110914

 - 110914

 - 110914

 - 110914

 - 110914

 - 110914

 - 110914

 - 110914

 - 110914

 - 110914

 - 110914

 - 110914

 - 110914

 - 110914

 - 110914

 - 110914

 - 110914

 - 110914

 - 110914

 - 110914

 - 110914

 - 110914

 - 110914

 - 110914

 - 110914

 - 110914

 - 110914

 - 110914

 - 110914

 - 110914

 - 1109 Ó -12 -12 +35 -105 -145

1UFC

Н K

L

1ÜFC 31 76 -31 -135 112 - 438 - 438 - 438 - 238

н 345609876543210123458765432101235452101208765432

K ~~~~ 440000000000000000

1-0 F O L

		-	X	
		-	7	
		-	6	
		-	š	
		-	ĩ	
		_	Ŧ	
		_	3	
			4	
		-	Y	
			ų	
			1	
			2	
0			3	
č			Ĩ.	
\$			Ś	
1	_	1	ń	
•		<u>'</u>	ă	
		_	2	
		-	9	
•	•	-	Ç	
		-	ō.	
		-	5	
		-	4	
		-	3	
		-	Ž	
		_	ī	
			ò	
			ĭ	
			5	
			ç	
			Ş.	
			4	
			2	
		-	9	
		-	8	
		-	7	
		-	6	
		_	Š.	
		+	ź	
			_	

i

.1

0

10F0 10FC

-32-012347 -1987654321 Ò 12345098765

н

ハカバババババババイ 4 1

71133472423226333427231314133753628482951223372 71133472423226333427231314133753628482951223372

10F0 10FC

381539731457U93743U651349642661022CUUU3156494883 7-3353733 2327 154 6-4 323 116646181849522-224

146

10FC

H

L

1, C F O 62651179229907111085611470171511951060498018084 43522454581123422349444334222 322362223377322327 1

10FC 140 -32 10 -21 1 -54381688713072233384364741481712159540169093 -54381688713072233884364741481712159540169093 - 11-

H

10F0 10FC 51061115111116140554111161000391353921026444446402827 141