

**STRUCTURAL STUDIES ON
PEPTIDES, ALKALOIDS
AND THEIR COMPLEXES**

STEVEN S.B. GLOVER

A Thesis presented for the degree of Doctor of Philosophy.

University of Edinburgh, 1984



DECLARATION

This thesis is the original composition of the author's work, unless otherwise stated, and it has not been submitted for any other degree.

LECTURES AND COURSES ATTENDED

In the course of this work the following lectures and courses were attended.

Natural Products	5 lectures	Prof. A. I. Scott
Biosynthesis	5 lectures	Dr. T. J. Simpson
Current Topics In Organic Chemistry	15 lectures	Members of the Organic Staff
Bio-organic Chemistry	5 lectures	Invited Speakers
Medicinal Chemistry	5 lectures	Dr. R. M. Paton
Synthesis and mode of action of β -lactam anti-biotics	5 lectures	Speakers from Glaxo Research
Chemistry of the Photographic process	5 lectures	Speakers from Kodak
Pulse sequences and Applications in NMR	5 lectures	Dr. G. A. Morris
X-ray crystallography	10 lectures	Dr. R. O. Gould Dr. M. D. Walkinshaw Dr. A. J. Welch
One week course in FORTRAN77		Staff of E.R.C.C.

In addition, the following group seminars have been attended:

Bio-organic group weekly meetings	One year
Crystallography	Two years
Organic Chemistry evening seminars	Two years

& departmental colloquia over a period of three years.

ACKNOWLEDGEMENTS

I would like to thank Dr. R. L. Baxter, Dr. R. O. Gould and Professor A. I. Scott for their help, encouragement and patient instruction while I was under their supervision. I would also like to thank Dr. M. D. Walkinshaw, who has been a supervisor in all but name, for all the guidance he has given me both in practical technique and the many discussions of the results presented here.

I would like to thank Miss Rosemary Kelly for the crystals of N-acetyl methionine and strychnine nitrate and Dr. E. M. Gordon, who prepared both the two disulphide compounds whilst working for Professor Scott in 1973.

The N.M.R spectra discussed in the disulphide chapters were measured by Dr. A. Boyd and Dr. D. Reid, to both of whom I would like to express my thanks.

During the period in which this work was carried out, I was supported by a post-graduate studentship from the Department of Education for Northern Ireland, for which I am very grateful.

I would also like to thank the staff of the E.R.R.C. advisory service for their assistance and advice on the use of the proportional spacing version of the LAYOUT text formatting program and document printer with which this thesis was compiled.

Finally, I must thank my wife Jennifer for every kind of immeasurable assistance while I was doing this work. In a very real sense this thesis is as much the result of her work as it is of mine.

**TO JENNIFER
AND TO ANTARA DAWN**

ABSTRACT

This dissertation reports eight crystal structure determinations carried out on amino acids, peptides, alkaloids and their complexes. Two of the compounds (N-phenylacetyl cysteinyl penicillamine cyclic disulphide methyl ester and bis(N-phenylacetyl) cystinyl bis(valine methyl ester)) were examined as part of a project on model compounds for penicillin biosynthesis, while the others (N-acetyl methionine, DL-glutamic acid monohydrate, strychnine, strychnine nitrate, brucine ethanolate dihydrate and strychnine: N-acetyl tyrosine tetrahydrate) were examined during a project on molecular recognition.

The first chapter contains a brief description of the monocyclic intermediate hypothesis of penicillin biosynthesis and the attempts to verify it, along with a discussion of molecular recognition phenomena ranging from substrate: receptor interactions of biological molecules to the resolution of racemic mixtures of amino acids by complexation with chiral bases.

The second chapter is a description of the methods used in determining X-ray structures, from the initial photographs and data collection, through the solution and refinement processes to the production of the results.

The structure of N-phenylacetyl cysteinyl penicillamine cyclic disulphide methyl ester is described in chapter 3. This compound was refined to $R = 0.0511$ on the basis of 1259 data. The conformation of the ring is compared with that of two other cyclic cysteine derivatives and the conformation of the disulphide linkage is compared with those of a selection of cystine derivatives from the Cambridge Crystallographic Database. The chapter also includes a discussion of an nmr solution conformation study carried out on the cyclic disulphide.

Chapter 4 is a description of the structure of bis(N-phenylacetyl) cystinyl bis(valine methyl ester). The asymmetric unit of this compound contains 4 independent half molecules, one of which is disordered. The

structure was refined to $R = 0.0912$ based on 2201 data. The conformation is compared with that of the antibiotic tetra N-desmethyl triostin A (TANDEM), which also contains an intramolecularly hydrogen bonded cystinyl valinyl fragment. This chapter also contains a description of a solution study on the disulphide.

The next chapter describes the structure of N-acetyl methionine, which was refined to $R = 0.0465$ on the basis of 947 data. The conformation of the methionine side chain is compared with that of 20 other methionine residues from the Cambridge database.

Chapter 6 describes the structure of DL-glutamic acid monohydrate, which was refined to $R = 0.0376$ with 1154 independent reflections. Once again the Cambridge database was used to provide a comparison of side chain conformation, with 24 independent glutamate residues found.

The next two chapters describe the structures of strychnine and strychnine nitrate. The free base refined to $R = 0.0370$ based on 1195 data, while the salt refined to $R = 0.0555$ based on 1076 reflections. The structures are compared with each other, and in the case of the nitrate salt, with a range of other strychnine salts. The chapter on the nitrate salt also includes a description of the calculation of an "idealised" strychnine cation from all the coordinates measured to date, which could be used for the rotation searching of Patterson vector maps of other strychnine structures. Listings of both the programs used are included in this chapter.

Chapter 9 describes the structure of brucine ethanolate dihydrate, which may be considered as an alcohol inclusion complex. This structure refined to $R = 0.0387$ with 1684 data. The structure is compared with the strychnine free base.

The last chapter discusses the molecular complex of strychnine with N-acetyl tyrosine. In this complex the tyrosine moiety is disordered over two sites of equal occupancy and the water molecules are also disordered. However, it was possible to refine the structure to $R = 0.104$ based on 1162 data by using rigid groups and distance constraints for the disordered molecule. The structure is compared with that of

a strychnine: N-benzoyl alanine complex.

CONTENTS

Chapter 1	Introduction	1
	Cystine Peptides	
	Molecular Recognition	4
Chapter 2	Methods used in X-ray Structural Determination	18
	List of Programs	29
Chapter 3	N-Phenylacetyl Cysteinyl Penicillamine Methyl Ester Cyclic Disulphide	32
	Introduction	
	Preparation and Crystallisation	
	Crystal Data	34
	Data Collection and Reduction	35
	Solution and Refinement	
	Description of Structure	37
	Comparison With Other Structures	50
	Solution Conformation Study	52
	Experimental	63
	Key to Compound Codes	66
	Fractional Coordinates	68
	Thermal Vibration Parameters	70
Chapter 4	Bis(N-Phenylacetyl) L-cystinyl bis(L-valine Methyl Ester)	71
	Introduction	
	Preparation and Crystallisation	72
	Crystal Data	

Space Group Determination	72
Data Collection and Reduction	73
Solution and Refinement	
Description of structure	75
Comparison with the cyclic octadepsipeptide	87
TANDEM	
Solution Study on the Dimeric Disulphide	91
Experimental	94
Fractional Coordinates	96
Thermal Vibration Parameters	100
 Chapter 5	
N-acetyl Methionine	101
 Introduction	
Crystallisation	
Crystal Data	
Data Collection and Reduction	102
Solution and Refinement	
Structure Description	103
Comparison With Other Structures	110
Key to Compounds	112
Fractional Coordinates	115
Thermal Vibration Parameters	116
 Chapter 6	
DL-Glutamic Acid Monohydrate	117
 Introduction	
Crystallisation	
Crystal Data	
Data Collection and Reduction	118
Solution and Refinement	
Description of Structure	119
Comparison With Other Structures	127
Key to Compounds	129
Fractional Coordinates	133
Thermal Vibration Parameters	134

Chapter 7	Strychnine	135
	Introduction	
	Crystallisation	136
	Crystal Data	
	Space Group Determination	
	Data Collection and Reduction	137
	Solution and Refinement	
	Description of Structure	
	Comparison With the Strychnine Salts	149
	Fractional Coordinates	150
	Thermal Vibration Parameters	152
Chapter 8	Strychnine Nitrate	153
	Introduction	
	Crystallisation	
	Crystal Data	154
	Data Collection and Reduction	
	Solution and Refinement	
	Structure Description	155
	Calculation of an Ideal Strychnine Cation	171
	Table of Bond Lengths	175
	Table of Bond Angles	176
	Table of Torsion Angles	177
	Fractional Coordinates	179
	Thermal Vibration Parameters	181
Chapter 9	Brucine Ethanolate Dihydrate	182
	Introduction	
	Crystallisation	
	Crystal Data	183
	Data Collection and Reduction	
	Solution and Refinement	
	Structure Description	184
	Comparison With Strychnine	198
	Fractional Coordinates	199

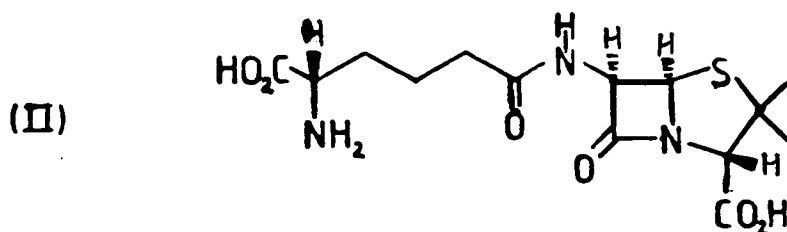
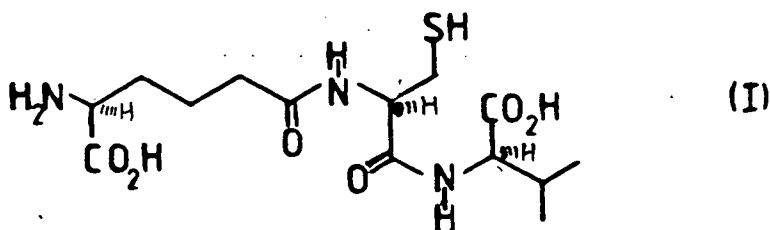
Thermal Vibration Parameters	201
Chapter 10 Strychnine: N-acetyl Tyrosine Tetrahydrate	202
Introduction	
Crystallisation	203
Crystal Data	204
Data Collection and Reduction	
Solution and Refinement	
Description of Structure	207
Fractional Coordinates	222
Thermal Vibration Parameters	224
References	225
Appendix Structure Factor Tables	234

CHAPTER ONE
INTRODUCTION

INTRODUCTION

Cystine Peptides

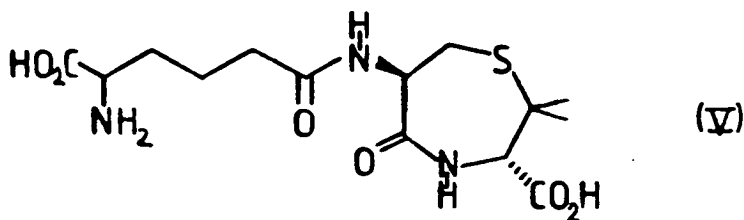
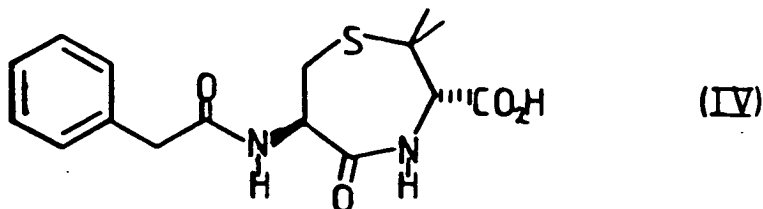
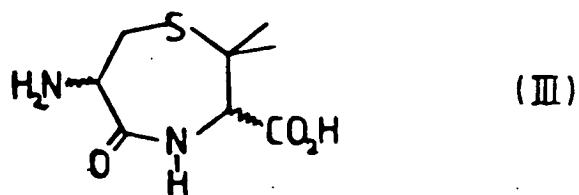
Although it has been known for nearly twenty five years that the penicillins are derived from the tripeptide L- α -aminoadipyl L-cysteinyl D-valine (I) (Arnstein *et al.*, 1960), the exact mechanism of the cyclisation of the tripeptide to the bicyclic isopenicillin N ring system (II) is unknown. Several proposals have been made for the existence of a cyclic peptide or thiazepine derivative, either of which could possibly be oxidised to the bicyclic system.



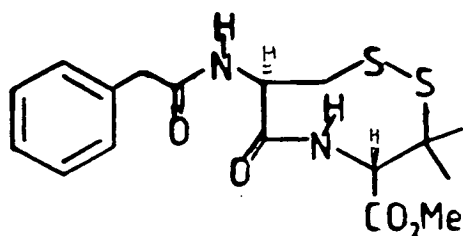
The first attempt to verify this hypothesis was in 1958 when it was reported by Arnstein that a cyclic peptide (III), possibly formed *in vivo* by an attack by the cysteine sulphur on the hydroxyl function of β -hydroxy valine, did not lower the rate at which a colony of *Penicillium chrysogenum* utilised a substrate for β -lactam manufacture. This failure to validate the hypothesis did not however immediately disprove it as the peptide used was a statistical mixture of all possible stereoisomers, and there was no evidence that it was even taken into

2

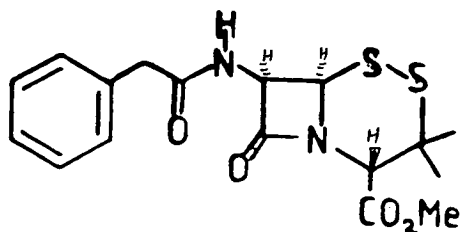
the mycelia. In 1965, van Tamelin found that the N-phenylacetyl derivative (IV) of the correct stereoisomer was not incorporated into penicillin. Again, it was considered possible that the material was not taken into the mycelia. It was also considered possible that the L- α -aminoadipyl side chain was necessary for recognition by the enzyme. It has since been shown that this is the case. A compound bearing the L- α -aminoadipyl side chain (V) has been synthesised (Wolfe, 1981) but it has since been shown to take no part in the biosynthetic pathway (Baldwin *et al*, 1981).



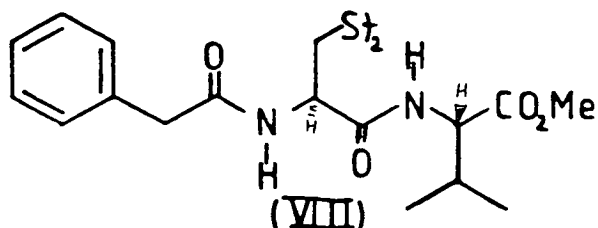
The cyclic disulphide compound (VI) was prepared as a model for the oxidation of the cyclic peptide to a bicyclic system (VII). No β -lactam product was obtained under any of the oxidation conditions used, however, which suggests that there is some form of steric hindrance to the reaction pathway.



(VI)



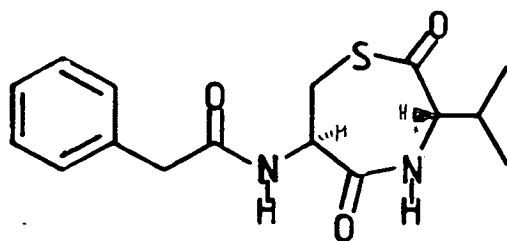
(VII)



(VIII)

The difficulty of closing the ring by oxidation depends greatly on the conformation of the peptide bond, as can be seen by examination of space-filling models of both isomers. It was therefore of importance to determine the configuration of the amide linkage, so that the probability of a successful bicyclisation could be assessed accurately.

The dimeric disulphide (VIII) was prepared during the same investigation of β -lactam biosynthesis as a synthetic intermediate in the preparation of the thiolactone (IX), which was also at one point considered to be a valid model for the cyclisation of the Arnstein tripeptide to the β -lactam system. As the disulphide (VIII) is structurally related to the cyclic compound (VI), its crystal structure is of interest from the point of view of a comparison of molecular conformation.



(IX)

Molecular Recognition

Molecular recognition is a fundamental part of a wide range of processes including catalysis, diastereomeric salt resolutions, drug receptor interactions and enzyme substrate reactions. (The latter two processes may involve similar effects, and are examples of biological recognition.)

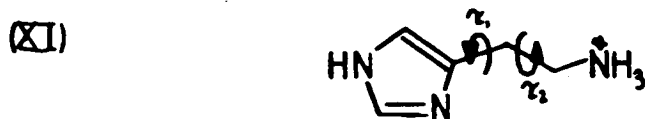
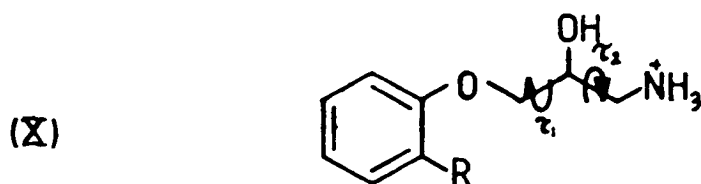
Biological recognition at the molecular level involves the binding of two mutually recognising species. In most cases, these are a macromolecule with an "active site" or *receptor* (usually a protein) and a much smaller, possibly flexible molecule or *drug* such as a substrate, inhibitor, neurotransmitter, hormone or other pharmacologically active substance. (Richards *et al*, 1975).

The ideal method for the study of biological recognition is to extract and purify the macromolecule and solve its crystal structure both in the native state, and with a variety of smaller molecules bound to the active site. An example of this approach is Lipscomb's work on zinc carboxy-peptidase (1970, 1971 & 1972), using both enzyme inhibitors and model substrates, but it has not yet been possible to extend this treatment to drug receptor complexes. In any case, this method is extremely time-consuming and not really rewarding in terms of detailed measurement of the recognition interactions as the resolution of crystal structures of macromolecules may not be sufficiently high.

In general then, the detailed structure of the receptor is unknown, as is its environment. It is possible that the substrate molecule may be required to adopt a different conformation from that adopted in the solid state or in solution. This becomes more likely when it is noted that some transmitters have more than one biological role and react with more than one specific receptor. Examples of this include acetylcholine, histamine and noradrenaline, each of which have synthetic analogues that stimulate or inhibit their different receptors to greater or lesser degrees.

As direct crystallographic examination of receptor substrate complexes is virtually impossible, the use of some form of model system is mandatory. The types of model system used can be divided into physical models, where the interactions between molecules are measured from crystal structures, thermodynamic data and the like, and theoretical methods, where molecular interactions are measured or inferred from energy calculations.

A simple example of the theoretical approach is that of Richards *et al* (1975). Working from the premise that as molecules bound into active sites are often distorted from their equilibrium conformation, standard structural methods (such as X-ray crystallography or nmr of solutions) cannot provide a full picture, they have correlated the degree of biological activity with the theoretically calculated potential energy surfaces for the phenoxy propanolamine β -adrenoceptor agonists (X) and the histamine H1 activity of a range of methyl histidines (XI).



They show that the efficacy of a drug has a linear correlation with the population of the relevant conformational form. The method is, however, incapable of indicating whether the conformation found is that of the molecule when at the active site, or if it is a transition state between the free and the bound conformations that the molecule must pass through to reach the active site.

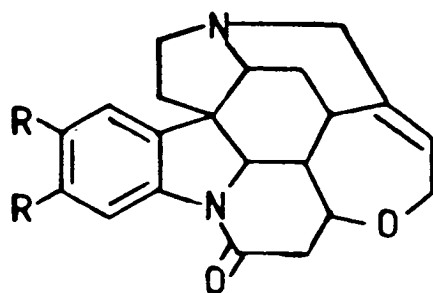
A more rigorous approach to the problem is taken by Pincus and Scheraga (1981), who examine molecular recognition and catalysis from

6

theoretical calculations on the enzyme substrate complex. This entails two stages of calculations. The first stage is the calculation of energy over all conformational space for the enzyme, substrate and bound complex so that the global minimum energy conformation can be found. The second stage of the calculation is the use of quantum mechanical calculations on the enzyme-substrate complexes produced in the first step to identify "good" or "bad" substrates. Although this method is very demanding of computer resources, it has been successfully applied to the prediction of the structures of a complex of α -chymotrypsin with N-acetyl-L-phenyl alaninamide and of lysozyme with N-acetyl glucosamine hexamer and the results compare favourably with crystal structures involving similar substrates.

At its simplest level, the physical approach to molecular recognition involves the examination of a crystal structure of a molecule (or series of related molecules) with reference to its preferred environment and configuration. The underlying assumption is that this environment will be similar to that of the molecule when bound to an active site, as almost all the interactions possible between protein residues may be found in crystal structures of amino acids and peptides, if enough structural determinations are compared.

The series of strychnine salts examined in Edinburgh show that the strychnine cation (XIIa, R = H) has a strong preference for a particular environment. Other reasons for using strychnine as a model for recognition studies are that it has a fairly rigid structure (causing it to take the same conformation in solution or on a receptor as it has in the solid state) and that it has several possible "interactive sites". These include an aromatic ring which could interact with a hydrophobic pocket in the receptor, a protonated nitrogen which could form a hydrogen bond to glutamate or aspartate residues and a carbonyl group which could act as a hydrogen bond acceptor. The alkaloid brucine (XIIb, R = OCH₃) has similar interactive capabilities, but possibly due to its larger hydrophobic region it does not have quite the same properties as strychnine either *in vivo* as a central stimulant nor *in vitro* as a resolving agent.



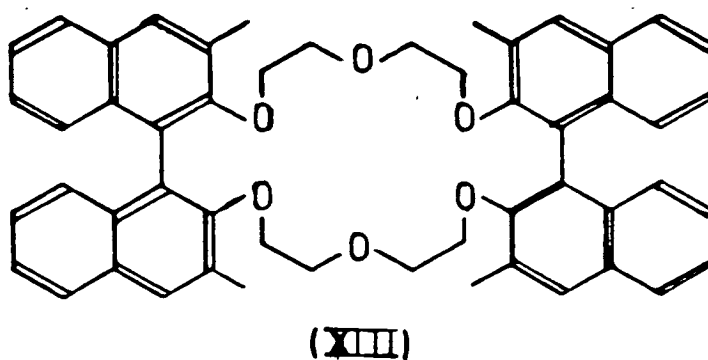
(XIIa) R=H

(XIIb) R=OCH₃

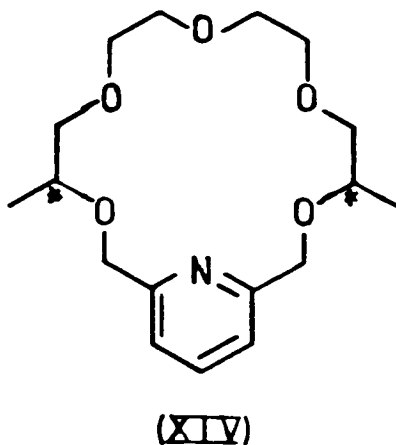
It is of interest that the relative positions of these interactive sites correspond closely to the positions of similar residues in the major loop of α -cobratoxin (Walkinshaw & Saenger, 1984), which has similar physiological effects to strychnine.

There have been several studies on host:guest complexation both in the solid state and in solution. Solution studies have been carried out on chiral recognition by crown ether derivatives or cryptands using a number of techniques to measure enantioselectivity.

In a study of chiral recognition of six amino acids and their methyl esters (alanine, methionine, valine, phenylalanine, tryptophan and phenylglycine, all as the perchlorate salt) by crown ether analogues (XIII) containing substituted dinaphthyl residues (Lingenfelter *et al.*, 1981), aqueous solutions of the racemic amino acids or esters were extracted enantioselectively with organic solvent containing the crown ether (this method is the basis of an amino ester resolving machine (Newcomb *et al.*, 1974)).



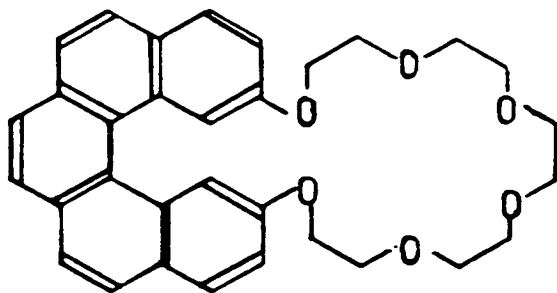
A study of chiral recognition of chiral alkylammonium salts by dimethyldioxopyridino-18-crown-6 (XIV) (Bradshaw *et al*, 1982) used the three independent methods of temperature dependent nmr, titration calorimetry and selective crystallisation to demonstrate the efficiency of their host. This report also contains the first measurement of the enthalpy and entropy of a chiral recognition reaction. The authors also note that the dimethyl substituted compound is more enantioselective than the diphenyl analogue.



It has recently been shown (Nakazaki *et al*, 1983) that pentahelicene and hexahelicene containing crown ethers were able to resolve methyl phenylglycinate and phenylethylamine hydrochlorides to about 25% optical

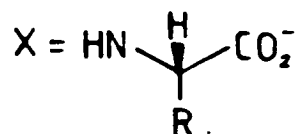
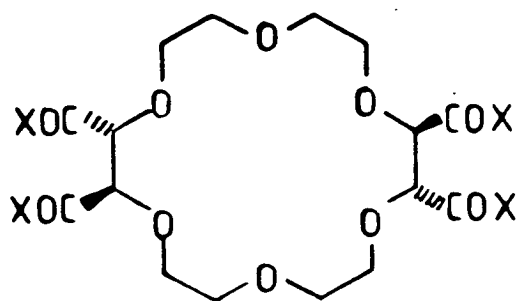
9

purity, except that the pentahelicene compound (XV) showed a 75% purity when used with the methyl phenylglycinate hydrochloride. No rationalisation of this is given, other than to say it is consistent with studies made using Corey Pauling Koltun space-filling models.



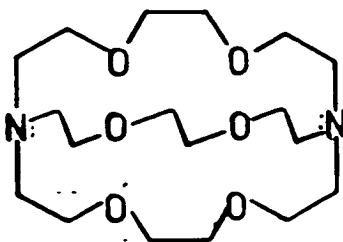
(XV)

Crown ethers bearing amino acid side chains (XVI) have been used as model receptors in a study of structural effects and substrate recognition of organic and biogenic ammonium ions (Behr et al, 1982), in which the interactions examined are divided into "central" and "lateral" discrimination. The central discrimination is derived from the size of the $C\alpha$ substituent and the number of protons on the quaternary nitrogen, and the lateral discrimination is derived from electrostatic and hydrophobic interactions between the guest molecule and the macrocycle side chains. However, the host molecule still binds potassium or ammonium cations more tightly than any of the organic ammonium ions.



(XVI)

Recently, a comparison of the receptor properties of 18-crown-6, the (2-2-2) cryptand (XVII) and α -chymotrypsin using both experimental (stability constant) and theoretical (molecular mechanics calculation) methods has been published (Wipff et al, 1983). The authors suggest that the active site of a receptor tends to mimic the solvation shell of the substrate.

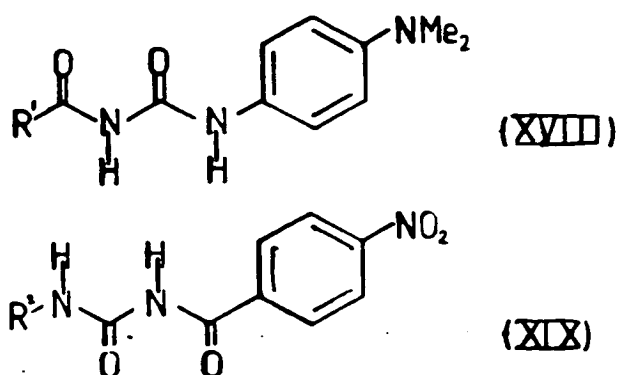


(XVII)

The design of a molecular receptor is briefly discussed in terms of the benefits of rigidity ("lock and key") and flexibility ("induced fit", allostery or cooperativity). It is also stated that the chiral discrimination by α -chymotrypsin takes place in the transition state rather than on initial binding.

Host:guest complexes in the solid state can be divided into clathrates, where one molecular species is imprisoned in closed cages formed by another species, the channel type (e.g. brucine ethanol dihydrate) where continuous channels running through a crystal accommodate the guest molecule, and the layer type (e.g. the strychnine salts) where layers of one species alternate with layers of the other. Although several clathrate compounds have been analysed by X-ray crystallography, they do not seem to have any specifically biomimetic interactions (MacNicol *et al.*, 1978).

Endo and co-workers (1982, 1983) have shown shape specific weak interactions between the alkyl groups of aromatic acylurea alkyl compounds (XVIII, XIX), both by reaction in solution to form disulphides and by crystallisation of 1:1 molecular complexes.



It might appear that the molecular association would be caused by hydrogen bonding at the acylurea sites, with bulky alkyl groups causing a decrease in association, but it is here suggested that complex formation is enhanced by the similarity of the alkyl groups rather than a simple lack of steric bulk. No mechanism for this interaction is proposed, although it is suggested that the mechanism would be useful for distinguishing between hydrophobic amino acids.

The series of complexes whose structures were determined by Ishida, Inoue and their co-workers over the last seven years under the general title of "Structural studies of the interactions between indole

12

derivatives and biologically important aromatic compounds" provides interesting examples of the type of results that may be obtained by using molecular complexes to model larger biological molecules, although it also shows some of the possible pitfalls inherent in the procedure. The main problem with preparing molecular complexes to model biological interactions is that the model compounds may have another mode of interaction available that is more favourable in the solid state than the one under examination. This may at first seem to imply that the model system is invalid, but it provides information on the relative strengths of possible interactions in the complex.

An example of this occurs in the structure of the indole-3-acetic acid: nicotinamide (1:1) complex (Ishida, et al, 1978i), which was prepared in order to investigate whether protonation of the pyridine ring to give the pyridinium ion is a necessary pre-condition for formation of a charge-transfer complex. When the structure was solved, it was seen that there were no charge-transfer or "ring stacking" interactions, but that the whole structure was held together by hydrogen bonds to form infinite layers perpendicular to the c axis. This result was taken as implying the necessity for a quaternary nitrogen in the nicotinamide moiety as a prerequisite for charge transfer, but it could also be possible that the hydrogen bonding interaction is energetically more favourable.

Examples of other biological interactions investigated in this manner are:

1) Between indole and benzene rings (Inoue et al, 1978ii; Ishida & Inoue, 1981iii) using tryptamine: phenylacetic acid (1:1) and tyramine: indoleacetic acid (1:1) complex. No charge transfer interactions.

2) Between flavins and indoles or pyrimidines (Ishida & Inoue, 1979; Inoue et al, 1980; Inoue et al, 1981; Ishida et al, 1981; Inoue et al, 1983)

These complexes all exhibit an interaction between the aromatic systems

13

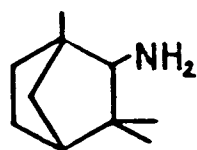
3) between nucleotides and indoles (Ishida et al, 1979i; Ishida et al, 1979ii; Ishida et al, 1980ii; Ishida et al, 1980iii; Ishida et al, 1981i; Ishida et al, 1982ii) These complexes model the interaction between proteins and nucleic acids, and again hydrogen bonding is prominent though there is some ring stacking.

There are also a number of other complexes in the series that do not quite fit this classification scheme. These are: thiamine indole-3-propionate (Inoue et al, 1982), which is a model for coenzyme: tryptophan interactions; a uracil: phenylethylamine complex (Ishida et al, 1983), which is a model for protein: nucleic acid interactions and two adenine compounds that are not strictly complexes, but discrete molecules. These are 1-[2(-(Adenin-9-yl) ethyl)-3-carbamoyl pyridinium chloride monohydrate, which is a pyridinium coenzyme model without any form of ring stacking (Ishida et al, 1982i), and a model nucleotide of 5'-deoxy -5'-adenosine acetic acid (Ishida et al, 1980i).

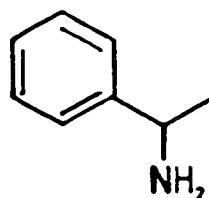
Other complexes involving nucleosides and aromatic compounds have been prepared in order to clarify the mode of action of intercalating antibiotics. The antitrypanosomal drug ethidium has been used to form a complex with 5'-iodouridylyl-(3'-5')-adenosine (Tsai et al, 1975) and the antimalarial drug proflavine has been co-crystallised with cytidylyl-(3'-5')-guanosine sulphate (Berman et al, 1979).

The diastereomeric salts of amino acids and alkaloids have been used in optical resolutions since the turn of the century (Fischer, 1899), and such complexes would be good models for drug receptor interaction as well as being suitable for a study of chiral recognition. Strychnine and brucine are the most commonly used resolving agents for N-protected amino acids (Dunn & Rockland, 1947; Greenstein, 1954 & 1961), but almost any chiral compound that can form a diastereomeric salt with the racemic amino acid may be used. Some examples of the use of resolving agents other than the indole alkaloids occur in the work of Ingersoll's group in the early 1950's and more recently, in the series of complexes prepared by Toda and his co-workers. Both groups made use of a "chain of resolutions", where each resolved compound may be used

to resolve another racemic mixture. Ingersoll used the N-acetyl derivative of the naturally occurring L-leucine to resolve crude dl- α -fenchylamine (XX) (Ingersoll & DeWitt, 1951) and then used the resulting optically pure amines for the resolution of DL-phenylalanine and DL-valine (via the N-acetyl or formyl derivatives). In the same paper (Overby & Ingersoll, 1951) the use of α -phenylethylamine (XXI) in resolution is discussed and it is mentioned that the resolved acetyl amino acids were capable of resolving dl-fenchylamine.

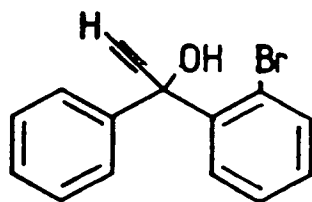


(XX)

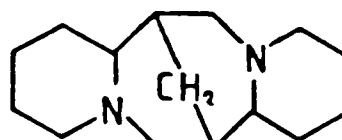


(XXI)

Toda used brucine to resolve the enantiomers of 1-(o-bromophenyl)-1-phenyl-2-propynol (XXII) (Toda et al, 1981) and determined the crystal structure of the resulting complex. It is suggested that the linearity of the acetylenyl group may be important in formation of the complex and further suggested that other racemic compounds with linear side chains (cyanohydrins, allenic or allyl alcohols) may also be resolved with brucine. The chiral alcohol was then used to resolve dl-sparteine (XXIII) (Toda et al, 1983i).



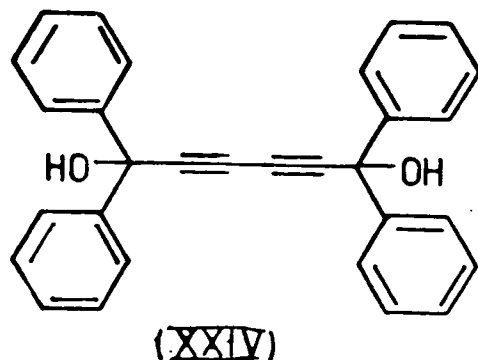
(XXII)



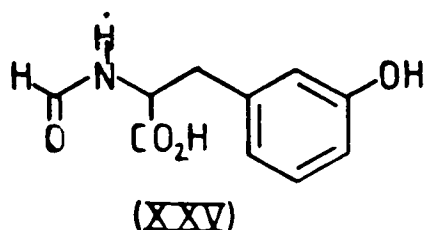
(XXIII)

It is reported that resolution by complexation with sparteine is more efficient than with brucine and furthermore, that sparteine is very much less toxic. Moreover, sparteine has an advantage over "natural" bases in that both enantiomers are available for complex formation from the same synthetic material, but d-brucine would be at best a prohibitively expensive curiosity. Toda and his colleagues carried the concept of diastereomeric salt formation a step further than simple optical resolution, when they used brucine to resolve compounds with two chiral centres (Toda *et al*, 1983ii). However, in the case of the compounds examined here (α - and β -haloacetylenic alcohols), several complexations were required to give optically pure alcohol, and the resolution did not work at all for a β -alkyl acetylenic alcohol.

They have also used a dimeric analogue of the acetylenic alcohol resolved with sparteine to resolve 3-methylcycloalkanones and 5-methyl- γ -butyrolactone (Toda *et al*, 1983iii). It is noted that an attempt at the resolution of 2-methylcyclohexanone was not successful and it is suggested that the distance between the chiral centre and the carbonyl group is crucial to the success of the resolution. Replacing bromine with chlorine does not greatly alter the resolving power of the complex, but the o-fluorophenyl derivative is much less efficient. Although no crystal structure of this complex has been published as yet, the structure of the analogous 2:1 acetone tetraphenyl-2,4-hexadiyne-1,6-diol complex is available (Toda *et al*, 1981) and it can be seen that the "wheel and axle" geometry of the diol (XXIV) that allows complexation to occur by holding the structure open. The infinite channels containing nothing other than acetone are similar in appearance to the channels containing 2-propanol in the cyclotricatechylene di-2-propanolate complex whose structure was determined by Hyatt's group (1980). The environment of the solvent in these cases is quite similar to that of the ethanol and water of the brucine complex whose structural determination is discussed below.



The procedure used for the resolution is essentially unchanged since the Fischer resolution of N-benzoylalanine, and consists of dissolving equimolar amounts of racemate and resolving agent in a suitable solvent (usually water or aqueous alcohol) by warming the suspension, filtering the resulting solution, and allowing it to slowly cool to room temperature as the less soluble diastereomeric salt is precipitated. The amino acid may be recovered from the diastereomeric salt by metathesis with a stronger base than that used in the precipitation or by acidification and extraction with a non-polar solvent. It should be noted though that the solvent chosen for the resolution may influence the relative solubilities of the diastereomeric salts obtained. Shabica and Tishler (1949) have shown that the brucine salt of N-acetyl D-tryptophan separated cleanly from dry ethanol while the residual L-isomer crystallised from methanol/ether only after concentration and charcoal filtration. Recently, the N-acetyl D-tryptophan complex with brucine was prepared by crystallisation from aqueous alcohol solution (Gould et al, 1984). This complex contains two molecules of brucine per tryptophan residue, of which only one is charged. This can be compared with the observations of Sealock and co-workers (1951), that brucine forms an ethanol insoluble complex with N-formyl D-meta tyrosine (XXV) while the same preparation carried out in aqueous solution yields a precipitate containing the L-tyrosine derivative. Presumably an attempt at resolving this compound using aqueous alcohol would not succeed as both salts could crystallise together.



Information on molecular recognition phenomena may also be obtained from crystal structures of amino acids and their derivatives by comparing the common features in environment and conformation over a range of independent crystal structure determinations containing the same amino acid residues. This can lead to an understanding of the relative magnitudes of the interactions between an amino acid and its environment. This may be considered as a form of autorecognition as the preferred environment of each part of the molecule should be of the same type in the simple amino acid as it is when the amino acid is part of the active site of a receptor protein.

CHAPTER TWO
METHODS USED IN X-RAY STRUCTURAL DETERMINATION

METHODS USED IN X-RAY STRUCTURAL DETERMINATION

The first stage in the determination of a crystal structure is the selection and mounting of a suitable crystal. The crystal should be large enough to diffract X-rays well, but small enough to remain within the X-ray beam at all times. To minimise absorption effects, the crystal should be of uniform size in all dimensions (though this is less important for crystals of compounds containing only the lighter elements, as they absorb the X-rays to a lesser extent), so the crystal may be cut to shape. A cuboidal or spherical crystal of about 0.3 mm in diameter is an ideal size. The crystal should also have plane faces without surface cracks and extinguish a beam of polarised light uniformly. A crystal fulfilling these requirements is then mounted, either by cementing it to a glass fibre, or in the case of unstable crystals, by sealing it into a glass tube with some mother liquor. The fibre or tube is then attached to a goniometer head in order to take preliminary photographs on the Weissenberg camera. These initial photographs are used to set the crystal accurately on the the goniometer head, and the crystal is then photographed again for unit cell and space group determination. An oscillation photograph and a zero-level Weissenberg photograph are used for cell measurement. For a correctly set crystal, the oscillation photograph will give the length of the unit cell axis parallel to the oscillation axis in the form of an interlayer spacing. When the camera is set up to take a Weissenberg photograph, screens are positioned so that the film is exposed to only one of these layers of reflections, and the film case is set in lateral motion by means of a gear system connected to the oscillating goniometer head in order to spread out the zone of reflections into an interpretable two dimensional net. For a first level Weissenberg photograph the screens are shifted to capture the first layer only. A precession photograph, showing an undistorted zone of the reciprocal lattice, may also be examined. The local program FILMFIT was used to refine the cell dimensions using accurately measured positions of reflections from Weissenberg or precession photographs.

With the exception of the cyclic disulphide, N-acetyl methionine and

the strychnine:N-acetyl tyrosine molecular complex, it was possible to determine the space groups uniquely from the systematic absences, for all the compounds examined during this work. It was, however, possible to deduce the space groups of these three structures given that they contained chiral molecules and that cell dimensions and density could be used to calculate the number of asymmetric units likely per cell.

Data collection was carried out using an Enraf-Nonius CAD-4 four circle diffractometer controlled by a PDP-11 computer for all but one of the structures examined. The strychnine nitrate data set was collected on the STADI-2 two circle diffractometer under the control of a PDP-8 computer. The output file of either type of data set contains intensity information for each reflection and its background along with the angle at which the reflection was measured. One or more reflections are measured several times at regular intervals to provide a means of measuring drift. This provides an indication of whether the crystal is decomposing or moving partially out of the X-ray beam.

The program used for initial data processing depends on whether the data was collected on the CAD-4 or STADI-2 diffractometers, although both CADABS and STOEABS perform similar operations on the input file. First the input file is read in order to calculate the drift curve. Values for the observed intensities are calculated by subtracting the background on either side of a reflection from the peak value and scaling according to the relative time taken for each measurement. The standard deviation may be derived from the square root of the scaled peak and background measurements.

The amplitudes of the structure factors are calculated from the intensities by first applying empirically calculated corrections for absorption and drift (derived from the curves plotted earlier), and then corrections for the effects of polarisation of the diffracted beam and the Lorentz effect (these are effects of data collection geometry). The structure factor amplitude e.s.d's are calculated from the intensity e.s.d's in a similar manner. As well as listing out the reflections, structure factors and e.s.d's, the data reduction program

20

provides statistical data which can be of use in space group determination.

The next stage of the procedure depends on the compound under examination. If the compound contains a few atoms of markedly higher atomic number than the rest of the structure (e.g. sulphur in an organic compound) then the structure may be solved by the Patterson or heavy-atom method. On the other hand, when the structure consists of atoms of approximately equal atomic number (neglecting hydrogen) then solution by direct methods is more applicable.

The process of structure determination is based on the fact that the data set collected is a Fourier transform of the unit cell. To determine the positions of the unit cell contents it is necessary to have two pieces of information for each structure factor. These are the amplitude and the relative phase of the diffracted beam producing the reflection. The amplitude information may be obtained from the intensity, but the phase information is impossible to record. It is this lack of phase information that prevents X-ray structural determination from being totally automatic.

The Patterson method avoids the need for phase information by using the squares of the structure factors and phase angles of zero in a Fourier synthesis. The resultant map does not contain information on the position of atoms relative to an origin, but only the inter-atomic vectors. The heights of the peaks on the map are proportional to the product of the atomic numbers of the atoms they connect, so the highest non-origin peaks should be those connecting the heavy atoms. As the map also contains vectors connecting symmetry related atoms, it may be possible to deduce a set of coordinates for the heavy atoms relative to an origin. A difference map can then be calculated using the phases calculated from the coordinates of the heavy atoms.

The structure of the cysteinyl penicillamine cyclic disulphide and of N-acetyl methionine were solved in this manner using the program SHELX to calculate the map, but the Patterson map calculated for the

21

dimeric disulphide contained so many vectors that the sulphur-sulphur vectors were hidden.

The program SHELX was also used to solve the structure of DL Glutamic acid by a direct method. It can be shown that for a centrosymmetric structure, all the phase angles are either zero or π radians. It can further be shown that for any three strong reflections H1, H2 and H3 with indices such that $H1 + H2 = H3$, the products of the cosines of their phases is usually positive. In principle, phases can therefore be determined for all the strong reflections if a suitable starting set of phases is chosen. In practice, several starting sets are chosen and the most consistent sets of phases calculated are used to calculate the electron density map.

For the non-centrosymmetric case the application of direct methods is not quite as straightforward as phase angles are no longer restricted to zero or π radians. The MULTAN system of programs was used to solve the structures in this category. The MULTAN system contains five programs which perform all the necessary calculations from normalisation of the structure factors to output of a list of coordinates for further refinement.

The first program in the suite is NORMAL, which normalises the structure factors by means of a calculated expected intensity for each reflection. This expected intensity is calculated using either randomly positioned atoms in the cell (using the Wilson plot), or if the information is available, randomly positioned spherically averaged groups of atoms (with a Debye curve). The more stereochemical information that is input in this way, the easier the solution of the structure should be. If no stereochemical information is read in, then the points on the Debye curve are the same as those of the Wilson plot. A line of best fit is calculated through the Wilson or Debye curves, which is used to obtain the scale multiplier and the temperature coefficient. It is also possible to input an arbitrary temperature coefficient to the program (this is how the disordered strychnine:N-acetyl tyrosine structure was solved). The scaling factors calculated are applied to the

27

reflections in such a way that the mean value of the square of the normalised structure factor is unity for each index group. After producing a comprehensive set of intensity statistics for use in space group determination, a file is prepared containing the strongest and weakest reflections (about the top 200 and bottom 50) for input to MULTAN.

The program MULTAN is itself divided into three sub-programs: SIGMA2, CONVERGE and FASTAN. The SIGMA2 section sets up the phase relationships which will be used by the tangent formula. These are relations between triplets of reflections such that the indices of H1 minus those of H2 equal those of H3 such that the phases may be related similarly:

$$\phi(H1) = \phi(H2) + \phi(H1-H2)$$

The weight given to each of these relationships is proportional to the product of their normalised structure factors. The CONVERGE sub-program uses the space group information to determine which reflections are structure seminvariants and which phase restrictions are imposed by symmetry for use in the application of the sigma 1 formula and phase determinations. This section also chooses the reflections that define the origin and fix an enantiomorph along with a few other starting reflections. The sigma 1 relation is similar to the phase relationship defined in SIGMA2, but it applies when two of the three reflections are the same. The calculated phases are once again stated with a probability of being correct, and the most probably correct phases are used as part of the tangent formula starting set with a weight of one less than twice their probability. The phases of the starting set reflections are assigned using a "magic integer" system unless the phase is restricted, when it is allowed to take two possible values only. The magic integer principle is a means of increasing the size of the starting set of reflections without making the total number of phase permutations prohibitively large, by expressing approximate values of several phases in terms of one parameter and writing

$$\phi = mx \pmod{1}$$

where ϕ is the phase in cycles and m is one of a set of integers. This means that with a set of five magic integers, starting phases of fifteen reflections may be defined with only three parameters. The relationship between phases in magic integer notation can be expressed as

$$Hx + Ky + Lz + b = 0$$

where H, K, L are sums of magic integers. It is not usually possible to find values of the parameters x, y, z to satisfy all the possible relationships, but the best solutions can be selected for by relating phase probabilities to intensities. These sets of phases provide the multiple starting points for the tangent formula, which are evenly distributed over the possible range of phase permutations to ensure that at least one is close to the correct set.

The final sub-program FASTAN develops these starting sets of phases by using them with the relationships derived by SIGMA2. The phases are then refined using the weighted tangent formula (Germain *et al*, 1971):

$$\tan \phi(H) = \frac{\sum (W(K) W(H-K) | E(K) E(H-K) | \sin(\phi(K) + \phi(H-K)))}{\sum (W(K) W(H-K) | E(K) E(H-K) | \cos(\phi(K) + \phi(H-K)))} = \frac{T(H)}{B(H)}$$

where $W(H)$ is the weight given to the phase of the reflections. This weighting leads to an efficient propagation of phase information as poorly determined phases have little effect on the values of well determined (and hence higher weighted) phases.

After the phases for each set have been determined by refinement with the tangent formula, three parameters - the figures of merit - are calculated as an indication of which phase sets lead to valid solutions. The absolute figure of merit is a measure of internal consistency among the SIGMA2 relationships and is zero if the phases

are randomly determined and greater than unity if the relationships are wholly consistent. The second figure of merit, $\Psi(0)$ is calculated using the weakest E-value reflections and should have as low a value as possible. This figure is sensitive to molecular positions and totally independent of the tangent formula. The last figure of merit is the residual calculated for the equation

$$E(H) = S * \langle E(K) * E(H-K) \rangle$$

where 'S' is a scale factor calculated to minimise the value of the residual and $\langle X \rangle$ means 'the average value of X'. All three figures of merit are combined into one, with the option of altering the weight given to each figure. This combined figure of merit is used to rank the phase sets for the generation of the electron density maps.

The last three programs, EXPAND, FFT and SEARCH are used together to produce an electron density map and analyse it for possible molecular fragments. A complete hemisphere of reflections is generated from the input reflections by the program EXPAND, which passes the result to FFT which synthesises the Fourier map corresponding to the highest unused figure of merit. (In the MULTAN 80 system, used for the solution of Strychnine:N-acetyl tyrosine complex, these two programs are combined into EFFT 80). Unless specifically requested, the map is not explicitly printed out, but passed to the program SEARCH. This program locates the highest peaks in the output map and lists their coordinates in order of peak height. The list of peaks is analysed to detect molecular fragments and a table of inter-peak distances and angles is calculated. A fragment interpretation routine applies minimum and maximum bond lengths and angles to eliminate spurious peaks and prints out an interpretation for each significantly different subset of the fragment. Up to three molecular plots are output depending on molecular shape. If the molecular fragment found is roughly planar, a plot of the peaks on the least-squares plane is sufficient to yield a clearly recognisable representation, but if the fragment is more spherical, then the plot of peaks on the least-squares plane can be somewhat confusing, so plots of peaks on the most-squares plane and

25

the plane orthogonal to these two are also output. In the intermediate case, of a "cylindrical" molecule, the plot of peaks on the most-squares plane is omitted. Finally, a list of coordinates is output for possible phase recycling. At this stage of the determination, the program SHELX can be used to calculate a difference map for the purpose of locating the rest of the non-hydrogen atoms in the structure, or to start the refinement procedure.

It is, however, possible that the known part of the structure is insufficient to phase all the reflections well enough to locate the unknown part on a difference map. This problem may be solved by the use of the DIRDIF system. This program takes as input the coordinates of the part structure and the $F(\text{obs})$ values for all reflections, which are used for the calculation of difference structure factors and then phase extension and refinement with the weighted tangent formula. The improved phases are then output to the MULTAN system programs EFFT and SEARCH for Fourier map generation and interpretation. The resulting fragment may be used for least squares refinement, or recycled through DIRDIF to produce a larger fragment. The DIRDIF system contains four sub-programs for use when both the position and orientation of the input molecular fragment are known, and two others for use when only orientation is known (the TRADIR system). The first subprogram, ENTER, is used to prepare the data file for use later in the system and to check how well the origin and enantiomorph are fixed. Structure factors and temperature factors are then calculated by part 1 of SFANDB. After calculation of the structure factors, this subprogram calculates a scale factor based on default thermal parameters and then an overall temperature factor from a Wilson-Parthasarathy plot. This is a form of Wilson plot which has been modified to allow for the contributions from the known part-structure. It is followed by a two dimensional refinement of the temperature factors of the known and unknown part-structure. In the second part of SFANDB, the temperature and scale factors are used to calculate maximum and minimum values for the E-values of the unknown part-structure, depending on whether $F(\text{obs})$ is in or out of phase with $F(\text{cal})$. A probability of the two being in phase rather than out of

phase is calculated. At this stage, several diagnostic parameters are calculated using the new temperature and scale factors.

The subprogram DIFTAN applies the weighted tangent formula for origin and enantiomorph fixation (if necessary) and for refinement and expansion of the phase information. For any phase that changes during refinement, the most probable E-value for the unknown part-structure is recalculated. After each cycle, the average shifts of the special and general reflections are output, with a consistency factor. If there are entire parity groups of reflections that are unphased by the known atoms, or if less than 60% of the reflections have phases assigned, then symbols are assigned to the phases of some reflections. A similar procedure is used for enantiomorph fixation in the case where the known atoms in an acentric structure have a pseudo centre of symmetry. Numerical values are then assigned to the symbols by a method using weighted symbol relations.

The last part of the DIRDIF system is the subprogram EXIT, which prepares a file containing the new phase information for output to EFFT and SEARCH of the MULTAN system (or optionally to the XRAY system) in the form of Fourier coefficients for a difference map, a weighted difference map, or a full electron density map.

The next subprogram is EXPAND, which is used only as part of TRADIR, to prepare a file for the program EXTRA. The file contains the reduced symmetry, as the next subprograms (SFANDB and DIFTAN) must be run with space group P1, the expanded data set and the original space group information (used by EXTRA only).

The program EXTRA (only used as part of TRADIR) calculates the structure factors of a symmetry related fragment by applying the relevant symmetry operators to the structure factors of the input fragment. These structure factors are combined with those of the unknown part-structure and a translation function is produced by the Fourier transform of the result. This translation function can be interpreted to obtain a shift vector which may be applied to the

input fragment in a run of DIRDIF.

The program SHELX is used for the last stages of structural determination. The molecular fragment found by DIRDIF (or by MULTAN, if that fragment was large enough) is used to phase a difference map based on all observed data (it is possible to vary the criteria for deciding if a reflection is observed). This difference map should indicate if any atoms are missing, or wrongly positioned. In some cases, the hydrogen atoms will be visible at this stage and these may be input to the next run of SHELX either as independent atoms or constrained to be in idealised positions relative to their environment. The thermal parameters of all non-hydrogen atoms may be either isotropically or anisotropically calculated, depending on the available data, but they are usually isotropic in the initial stages of refinement. Unless they are allowed to vary independently the hydrogen thermal parameters are usually either fixed to the same value (which may or may not refine) or they may be derived from the thermal parameter of the atom they are attached to. There are a number of mechanisms for constraining parameters during least squares refinement which may be of great use if the data does not permit all the parameters to vary independently, or if the structure is disordered. At its simplest, the constraint mechanism involves fixing a parameter so that it does not vary (e.g. for origin fixation in $P1$ or $P2_1$, or to hold all occupancy factors at unity), or it may be used to constrain a group of atoms to maintain a specific geometry (either as a rigid group or by explicitly listing interatomic distances and permissible e.s.d's). Another use of the constraint mechanism is when an atom is disordered over two sites and the total occupancy must be equal to that of a non-disordered atom (it is necessary in this situation to constrain the thermal parameters), as in the strychnine: N-acetyl tyrosine structure.

Another useful feature of the SHELX program is the facility for printing out the entire electron density map and not just the plot of peaks on the least squares plane. This map may be used as input to a transformation and interpolation program, FMAPO, which produces an output file suitable for contouring and detailed examination for the

location of atomic positions in disordered regions. This is essential, as it is possible for the two possible sites of a disordered atom to be so close to each other that the overlap region appears higher in electron density than the atomic centres, which is a virtually impossible situation for the peak search routines of SHELX to cope with.

Although it is possible to "refine" a structure by repetitively generating Fourier maps and manually altering the file of atomic parameters to improve the fit between the model and the structure, the use of least squares refinement is more efficient as it directly minimises the difference between $F(\text{obs})$ and $F(\text{cal})$. If each of the data points involved in the minimisation is assigned a weight varying with the precision of measurement, and there are many more observations than parameters, the quantity to be minimised may be expressed as:

$$Q = \sum_H w_H (|F_{oH}| - |F_{cH}|)^2$$

which is summed over all observed data. To minimise this figure, it is necessary to find the point at which the derivative of Q with respect to each parameter is zero, leading to the same number of independent simultaneous equations as there are parameters. This would imply that a solution to these equations is easily determinable as there are usually 5-10 observations (reflections) for each parameter. The method of least squares, however, applies only to linear equations, and the above differential equations are anything but linear. The solution to this problem is to use the shift in the parameters and not the parameters themselves as variables. This method requires a trial structure to be reasonably close to the final structure, so that a Taylor series expanded from the initial parameters will lead to a better set. A further assumption is that the trial structure is sufficiently close to the final structure to permit the use of the first derivative term only, leading to the relation

$$|F_o| - |F_c| = \sum \frac{\delta |F_c| \cdot \Delta p_i}{\delta p_i}$$

for all parameters P_i . As this method involves several approximations, several iterations are needed before final convergence is achieved. At each end of each cycle, the R factors are output, along with a list of parameters (with e.s.d's and shift/e.s.d.). The difference map is then calculated and the peaks plotted. The final part of the output shows the number and variance of the reflections in each angular zone and by parity group. These may be used to determine if a weighting scheme is necessary.

After the final convergence is reached and the structure is fully determined SHELX may be used to print out a list of reflections with $F(\text{obs})$ and $F(\text{cal})$ values and the program PAPER used to format the SHELX output into suitable tables of coordinates and thermal parameters with e.s.d's for eventual publication. The molecular geometry may be fully explored using the program CALC to produce lists of bond-lengths, angles, torsion angles and contact distances as well as planes through the molecule (and their dihedral angles) or to fit one molecule or fragment to another. Finally, the effects of thermal vibration or of molecular packing may be shown graphically by means of the programs ORTEP and PLUTO.

List of Programs

DIRDIF

Direct Methods for Difference Structures. An Automatic Procedure for Phase Extension and Refinement of Difference Structure Factors.

P. T. Beurskens, W. P. Bosman, H. M. Doesburg, R. O. Gould, Th. E. M. van den Hark, P. A. J. Prick, J. H. Noordik, G. Beurskens and V. Parthasarathy.

Technical report 1981/2, Crystallography lab. Toerooiveld, 6525 ED, Nijmegen, Netherlands

MULTAN-77

A System of Computer Programs for the Automatic Solution of Crystal Structures from the X-ray Diffraction Data

P. Main, L. Lessinger, M. M. Woolfson, G. Germain and J-P. Declerq, 1977.

SHELX-76

A Program for Crystal Structure Solution.

G. Sheldrick, University of Cambridge, 1976.

OR TEP-II

A Fortran Thermal Ellipsoid Plot Program for Crystal Structure Illustrations.

C. K. Johnson, Oak Ridge National Laboratory, Tennessee, 1976.

PLUTO

Molecular Plotting Program

W. D. S. Motherwell, University of Cambridge, 1976.

In addition, the following 'local' programs were used:

FILMFIT: a program for refining cell constants from film data

CADABS & STOEABS: programs for the reduction of diffractometer collected data

TABLES: a program for formatting tables of bonds, angles and torsion angles with their e.s.d's

PAPER: a program for formatting tables of atomic coordinates and thermal parameters with their e.s.d's

All the above written by Dr. R. O. Gould.

FMAPO: a program for the interpolation and scaling of SHELX electron density output for line-printer listing and contouring

By Dr. M. M. Harding

CALC: a program for molecular geometry

By R. O. Gould and P. Taylor

A NOTE ON STEREOCHEMISTRY

Of the eight crystal structures presented here, seven are chiral compounds crystallising in acentric space groups. As it is not possible to determine the absolute configuration of a compound crystallographically without the use of anomalous dispersion measurements, it is necessary to decide the configuration of the compounds on the basis of chemical information and by reference to the structures of known compounds.

For the cyclic disulphide, the stereochemistry of the penicillamine used in synthesis is known from optical rotation measurements, indicating that the compound contains l-cysteine and d-penicillamine. This indirectly allows determination of the stereochemistry of the dimeric disulphide, as it was prepared from the same batch of l-cystine.

The N-acetyl methionine was supplied as the l-isomer, and was assumed to be so as it formed an insoluble complex with brucine, but not with strychnine (Greenstein, 1954).

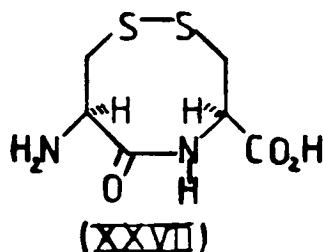
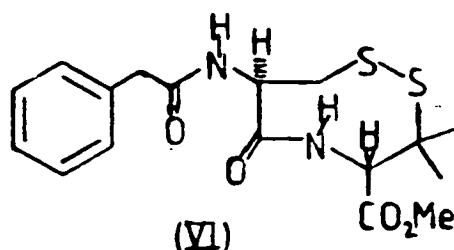
The remaining structures all contain the naturally produced strychnine molecular skeleton, and their configurations are in agreement with that found by anomalous dispersion methods on strychnine hydrobromide (Peerdeman, 1956). Even without this study, it would have been possible to determine the absolute configuration of strychnine from the strychnine tyrosine complex if the absolute configuration of the tyrosine used was known.

CHAPTER THREE
N-PHENYLACETYL CYSTEINYL PENICILLAMINE METHYL ESTER
CYCLIC DISULPHIDE

3,3-DIMETHYL-4-D-CARBOMETHOXY-6-OXO-7-L-PHENYLACETYLAMINO PERHYDRO-1,2,5-DITHIAZOCINE

Introduction

Although this cyclic disulphide (VI) was first prepared as a biosynthetic model for β -lactam formation (as described above), the structure is of interest in its own right as the smallest ring system containing both a disulphide bridge and a *trans* amide linkage. There is a similar system, *cis* L-cysteinyll cysteine disulphide (XXVII), but this compound contains a *cis* amide linkage.

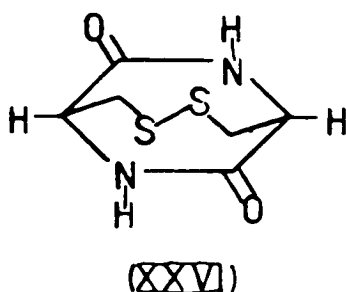


Energy calculations carried out on the *cis* compound (Chandrasekharan & Balasubramanian, 1969) have indicated the possibility of two distinct conformations of this compound, differing only in the sense of the torsion angle χ^3 (C-S-S-C). The conformation is referred to as P-helical when the angle χ^3 is positive (or *circa* Plus 90°) and as M-helical when the angle is negative (Minus 90°). In neither of these conformations is the amide linkage planar and some of the intra-annular angles are distorted, as the potential function used for the energy minimisation included the effects of changes in non-bonded interactions, electrostatic interactions, dihedral angle distortion and variations in bond angles, while keeping bond lengths constant at standard values. Both the predicted conformations have amide linkages distorted by over

ten degrees, but in opposite directions to the sign of the disulphide torsion.

Conformation I	$\Delta\omega = -12^\circ$	$E = -3.55 \text{ kcal mol}^{-1}$	P-helical
Conformation II	$\Delta\omega = 14^\circ$	$E = -2.92 \text{ kcal mol}^{-1}$	M-helical

Conformation I with its smaller distortion in amide planarity is therefore the expected conformation, but the difference is not extreme and it may be possible for either conformation to exist, depending on the size of the barrier between the states. Later work by Chandrasekharan and Mitra (1977) on the conformation of the disulphide bridge in *cyclo* L-cystine (cystine diketopiperazine or cys-DKP) (XXVI) again predicts two possible conformations of the disulphide bridge with two-fold rotational symmetry in each case.



The M-helical conformation is the more stable ($-9.2 \text{ kcal mol}^{-1}$ as against $-8.9 \text{ kcal mol}^{-1}$ for the P-helical structure) and there are no other differences in gross conformation. At first glance, this would appear to contradict the work of Ottnad and co-workers (1975) who studied the conformation of cys-DKP in DMSO solution and found it to contain a P-helical disulphide bridge, but the crystal structure of the acetic acid solvate of cys-DKP (Mez, 1974) shows the bridge to be M-helical.

More recently, the crystal structures of two derivatives (XXVI, XXVII) of *cyclo* L-cysteinyl L-cysteine have been determined (Capasso et al, 1977; Ashida et al, 1977), showing both to be P-helical but to have the

34

opposite distortion at the amide linkage. These compounds will be discussed in more detail below. The seeming contradiction in the work of Ottnad and Mez may be resolved if it is assumed that the helicity of the disulphide bridge may change depending on the molecular environment, even in a rigid molecule such as cys-DKP. However, a solution conformation study on the *cyclo* L-cysteinyl L-cysteine (Capasso *et al*, 1979) presented no real evidence for either configuration of the bridge to the exclusion of the other. This being the case, it was decided that an investigation of the solution conformation of the cysteinyl penicillamine cyclic disulphide would indicate whether or not its disulphide bridge was able to alter its configuration in solution. The results of this nmr study are also presented below.

Preparation and Crystallisation

The compound was prepared and crystallised by Dr. E. M. Gordon, according to the procedure outlined in the experimental section below. The crystals were grown by dissolving the compound in a slight excess of chloroform, adding 40/60 petrol ether until crystallisation began and then leaving the solution at room temperature overnight.

Crystal Data

Formula $C_{17}H_{22}N_2O_4S_2$ $M = 382.47$ $F(000) = 404$
Space group $P1$ Int. Tab. No. 1
Cell dimensions $a = 9.240(4)$ $b = 9.642(4)$ $c = 11.317(4)$ Å
 $\alpha = 94.97(5)$ $\beta = 98.85(5)$ $\gamma = 110.37(5)^\circ$
Volume = 923.2 cubic Angstroms $Z = 2$
 $D_c = 1.376 \text{ gcm}^{-3}$ $D_m = 1.348 \text{ gcm}^{-3}$
Radiation $Mo K\alpha$ $\lambda = 0.71069$ Angstrom $\mu = 2.99 \text{ cm}^{-1}$
Final $R = 0.0511$ based on 1259 independent reflections
 h range 0 to 10 k range -11 to 10 l range -13 to 13
 θ max = 25° $\text{Sin}(\theta \text{ max})/\lambda = 0.5947$
3282 reflections measured
3246 after merging
1987 considered unobserved ($F < 6\sigma(F)$)
Parameters refined 290
Max shift/ σ of last cycle 0.186
Final difference Fourier max. 0.3444 eA^{-3} min. -0.2407 eA^{-3}
Weight = $1.3833/(\sigma^2(F) + 0.000363F^2)$

Data Collection and Reduction

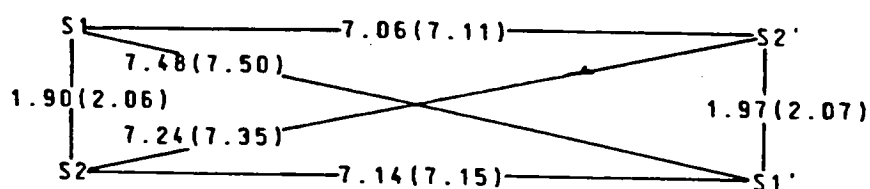
The data set was collected on the CAD4, and corrected for Lorentz and polarisation effects using the program CADABS.

Solution and Refinement

The structure was solved from the Patterson map produced by the Program SHELX. The top four non-origin peaks in the map were interpreted as being consistent with a centrosymmetric arrangement of four sulphur atoms in a parallelogram, such that two of the vectors were double weighted. These four peaks, and the next two highest peaks are tabulated below.

No.	Height	x/A	y/B	z/C	length
1	231	0.6804	0.4704	0.2990	7.105
2	191	0.1066	0.1603	0.9042	2.023
3	111	0.2224	0.3822	0.7866	7.287
4	105	0.5626	0.2994	0.3950	7.490
5	92	0.7922	0.4243	0.1865	5.486
6	88	0.8693	0.4777	0.1040	5.237

The diagram below shows the four sulphur-sulphur vectors and their final lengths.



The coordinates derived from these vectors were used as input for the program SHELX, but the resulting difference map showed only a part of the two molecules, and recycling this fragment through SHELX did not lead to any further improvement in the difference map. This situation, where the only part of the structure known is the position of a few heavy atoms around a pseudo-inversion centre, is an almost ideal case for the use of the DIRDIF procedure.

The first run of DIRDIF produced a map containing a fragment of one molecule, which was used as input to a new run. The new map was then examined for a new input fragment for a further run of the procedure. The diagnostics produced by the program are tabulated below.

Input Fragment Test	S ₄	S ₂ C ₅ N	S ₂ C ₉ N ₂ O ₃
R	0.63	0.66	0.60
BH	4.50	3.04	3.41
BL	3.70	4.96	5.39
oC(EH)	2.4*	5.4	5.31
(EL)	0.6*	2.3	2.18
oA(EH)	6.4	1.9*	0.68
(EL)	2.5	1.0*	1.91
Consistency			
cycle 1	0.30(332)	0.26(400)	0.27(399)
2	0.21(526)	0.17(688)	0.19(828)
3	0.22(562)	0.21(687)	0.26(841)

In this table BH and BL are the overall temperature factors for the known and unknown part-structures respectively.* The next set of figures relate to the distribution function of the expectation values of E calculated from both centric and acentric expressions. The change in these values on going from a centric to an acentric input fragment is marked thus '*'. The last part of the table shows the progress of the weighted tangent phase extension and refinement, with the figures in brackets being the number of reflections in the basis set for each cycle.

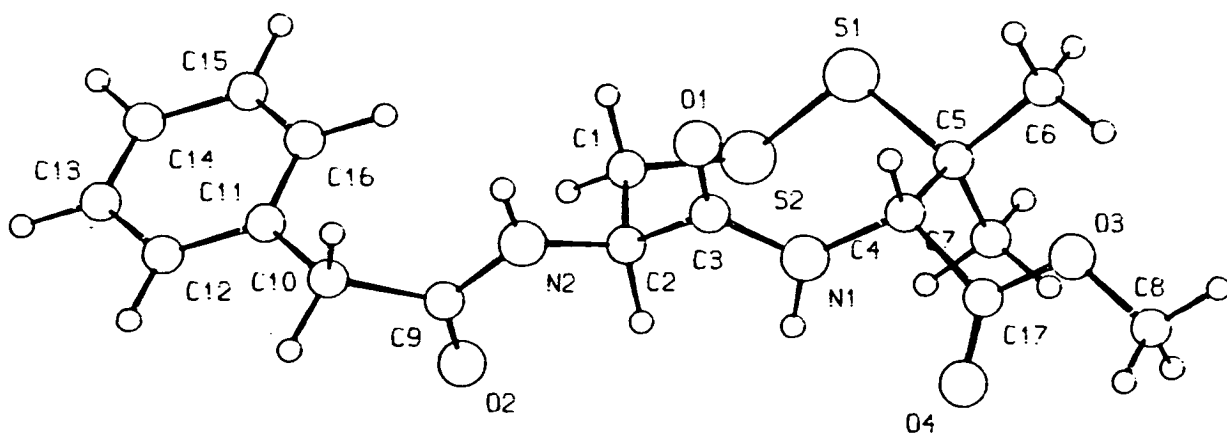
After another couple of cycles through DIRDIF, the structure was complete enough to use the program SHELX for refinement. The progress of refinement is summarised in the table below.

*
In this table, the letters H and L indicate the expressions involving the known fragment of the molecule (or Heavy atoms) and the unknown fragment (or the Light atoms).

R	Comment
0.2554	Almost two complete molecules input (2 carbons missing)
0.2435	All of one molecule located, and all of the other but for a phenyl ring
0.1406	All non-H atoms found, isotropic thermal parameters tied to a free variable
0.0947	All thermal parameters freed
0.0943	Some H atoms located in the difference map H atoms input in calculated positions
0.0714	Sulphur atoms with anisotropic thermal parameters and phenyl rings constrained to ideal geometry
0.0644	Amide oxygens anisotropic
0.0592	Nitrogen atoms anisotropic
0.0567	Convergence
0.0569	Thermal parameters of H atoms 'riding' on those of atoms they bond to, aromatic rings no longer constrained to ideal geometry
0.0546	Absolute stereochemistry corrected
0.0511	Final convergence*

Description of Structure

The unit cell contains two independent molecules arranged head to tail (causing the pseudocentric arrangement of the four sulphur atoms) and connected by hydrogen bonding into an infinite ribbon parallel to the y axis, in a manner analogous to the β -pleated sheet found in most polypeptides and proteins. The numbering scheme used in this section is illustrated on the diagram below.



*All hydrogens were constrained to remain at calculated positions, with their thermal parameters 'riding' on those of the atom to which they are bonded

The following tables show all unconstrained bond lengths, angles and torsion angles for both molecules with the addition of some torsion angles involving hydrogen which are referred to in the section on solution conformation below.

S(1) - S(2)	2.058(6)	C(2) - C(3)	1.515(20)
S(1) - C(5)	1.861(15)	C(4) - C(5)	1.565(21)
S(2) - C(1)	1.814(16)	C(4) - C(17)	1.562(22)
O(1) - C(3)	1.252(18)	C(5) - C(6)	1.569(23)
O(2) - C(9)	1.195(19)	C(5) - C(7)	1.523(22)
O(3) - C(8)	1.427(19)	C(9) - C(10)	1.581(23)
O(3) - C(17)	1.365(19)	C(10) - C(11)	1.482(21)
O(4) - C(17)	1.161(20)	C(11) - C(12)	1.418(22)
N(1) - C(3)	1.367(19)	C(11) - C(16)	1.344(21)
N(1) - C(4)	1.459(19)	C(12) - C(13)	1.41(3)
N(2) - C(2)	1.465(19)	C(13) - C(14)	1.30(3)
N(2) - C(9)	1.364(20)	C(14) - C(15)	1.41(3)
C(1) - C(2)	1.500(20)	C(15) - C(16)	1.34(3)

S(1') - S(2')	2.066(8)	C(2') - C(3')	1.529(19)
S(1') - C(5')	1.855(16)	C(4') - C(5')	1.478(21)
S(2') - C(1')	1.823(16)	C(4') - C(17')	1.515(23)
O(1') - C(3')	1.207(18)	C(5') - C(6')	1.538(22)
O(2') - C(9')	1.217(20)	C(5') - C(7')	1.549(21)
O(3') - C(8')	1.446(21)	C(9') - C(10')	1.551(22)
O(3') - C(17')	1.332(21)	C(10') - C(11')	1.476(22)
O(4') - C(17')	1.199(21)	C(11') - C(12')	1.438(25)
N(1') - C(3')	1.385(19)	C(11') - C(16')	1.348(24)
N(1') - C(4')	1.455(19)	C(12') - C(13')	1.37(3)
N(2') - C(2')	1.475(18)	C(13') - C(14')	1.27(3)
N(2') - C(9')	1.354(21)	C(14') - C(15')	1.39(3)
C(1') - C(2')	1.479(20)	C(15') - C(16')	1.40(3)

S(2)-S(1)-C(5)	110.0(5)	C(4)-C(5)-C(6)	109.5(12)
S(1)-S(2)-C(1)	108.5(5)	C(4)-C(5)-C(7)	111.4(12)
C(8)-O(3)-C(17)	115.1(12)	C(6)-C(5)-C(7)	111.9(13)
C(3)-N(1)-C(4)	124.0(12)	O(2)-C(9)-N(2)	121.5(15)
C(2)-N(2)-C(9)	120.2(12)	O(2)-C(9)-C(10)	121.8(14)
S(2)-C(1)-C(2)	115.4(10)	N(2)-C(9)-C(10)	116.8(13)
N(2)-C(2)-C(1)	107.2(11)	C(9)-C(10)-C(11)	111.9(13)
N(2)-C(2)-C(3)	113.4(11)	C(10)-C(11)-C(12)	121.0(13)
C(1)-C(2)-C(3)	107.0(11)	C(10)-C(11)-C(16)	120.5(14)
O(1)-C(3)-N(1)	123.0(13)	C(12)-C(11)-C(16)	118.5(14)
O(1)-C(3)-C(2)	122.2(13)	C(11)-C(12)-C(13)	121.5(15)
N(1)-C(3)-C(2)	114.3(12)	C(12)-C(13)-C(14)	116.7(17)
N(1)-C(4)-C(5)	109.2(12)	C(13)-C(14)-C(15)	122.1(17)
N(1)-C(4)-C(17)	10.0(12)	C(14)-C(15)-C(16)	121.7(18)
C(5)-C(4)-C(17)	106.7(12)	C(11)-C(16)-C(15)	119.4(16)
S(1)-C(5)-C(4)	106.5(10)	O(3)-C(17)-O(4)	126.9(15)
S(1)-C(5)-C(6)	104.9(10)	O(3)-C(17)-C(4)	110.8(13)
S(1)-C(5)-C(7)	112.4(11)	O(4)-C(17)-C(4)	122.3(15)

S(2')-S(1')-C(5')	110.1(6)	C(4')-C(5')-C(6')	112.4(13)
S(1')-S(2')-C(1')	107.1(6)	C(4')-C(5')-C(7')	113.7(12)
C(8')-O(3')-C(17')	115.7(13)	C(6')-C(5')-C(7')	109.2(12)
C(3')-N(1')-C(4')	122.4(12)	O(2')-C(9')-N(2')	121.5(15)
C(2')-N(2')-C(9')	120.7(12)	O(2')-C(9')-C(10')	121.5(14)
S(2')-C(1')-C(2')	116.0(11)	N(2')-C(9')-C(10')	117.0(14)
N(2')-C(2')-C(1')	106.6(11)	C(9')-C(10')-C(11')	112.6(13)
N(2')-C(2')-C(3')	114.1(11)	C(10')-C(11')-C(12')	120.4(14)
C(1')-C(2')-C(3')	107.5(11)	C(10')-C(11')-C(16')	119.7(15)
O(1')-C(3')-N(1')	124.1(13)	C(12')-C(11')-C(16')	119.8(15)
O(1')-C(3')-C(2')	121.7(13)	C(11')-C(12')-C(13')	120.9(17)
N(1')-C(3')-C(2')	113.5(12)	C(12')-C(13')-C(14')	118.7(20)
N(1')-C(4')-C(5')	110.8(12)	C(13')-C(14')-C(15')	122.6(21)
N(1')-C(4')-C(17')	110.0(12)	C(14')-C(15')-C(16')	121.6(19)
C(5')-C(4')-C(17')	109.4(12)	C(11')-C(16')-C(15')	116.1(17)
S(1')-C(5')-C(4')	106.9(10)	O(3')-C(17')-O(4')	123.9(16)
S(1')-C(5')-C(6')	104.0(10)	O(3')-C(17')-C(4')	113.1(14)
S(1')-C(5')-C(7')	110.2(10)	O(4')-C(17')-C(4')	122.8(15)

4

C(5) - S(1) - S(2) - C(1)	107.7(7)
S(2) - S(1) - C(5) - C(4)	-81.8(10)
S(2) - S(1) - C(5) - C(6)	162.2(9)
S(2) - S(1) - C(5) - C(7)	40.4(12)
S(1) - S(2) - C(1) - C(2)	-76.2(11)
C(8) - O(3) -C(17) - O(4)	8.3(23)
C(8) - O(3) -C(17) - C(4)	-173.0(12)
C(4) - N(1) - C(3) - O(1)	-19.0(22)
C(4) - N(1) - C(3) - C(2)	152.9(13)
C(3) - N(1) - C(4) - C(5)	-98.5(15)
C(3) - N(1) - C(4) -C(17)	144.7(13)
C(9) - N(2) - C(2) - C(1)	84.6(16)
C(9) - N(2) - C(2) - C(3)	-157.5(13)
C(2) - N(2) - C(9) - O(2)	10.6(22)
C(2) - N(2) - C(9) -C(10)	-168.2(12)
S(2) - C(1) - C(2) - N(2)	180.0(9)
S(2) - C(1) - C(2) - C(3)	58.0(14)
N(2) - C(2) - C(3) - O(1)	-38.2(19)
N(2) - C(2) - C(3) - N(1)	149.9(12)
C(1) - C(2) - C(3) - O(1)	79.8(17)
C(1) - C(2) - C(3) - N(1)	-92.1(14)
N(1) - C(4) - C(5) - S(1)	51.8(13)
N(1) - C(4) - C(5) - C(6)	164.7(12)
N(1) - C(4) - C(5) - C(7)	-71.1(15)
C(17) - C(4) - C(5) - S(1)	170.7(10)
C(17) - C(4) - C(5) - C(6)	-76.4(15)
C(17) - C(4) - C(5) - C(7)	47.8(16)
N(1) - C(4) -C(17) - O(3)	-172.8(12)
N(1) - C(4) -C(17) - O(4)	6.0(21)
C(5) - C(4) -C(17) - O(3)	68.9(15)
C(5) - C(4) -C(17) - O(4)	-112.3(17)
O(2) - C(9) -C(10) -C(11)	-73.2(19)
N(2) - C(9) -C(10) -C(11)	105.6(16)
C(9) -C(10) -C(11) -C(12)	133.0(15)
C(9) -C(10) -C(11) -C(16)	-48.1(19)
C(10) -C(11) -C(12) -C(13)	-180.0(15)
C(16) -C(11) -C(12) -C(13)	1.0(24)
C(10) -C(11) -C(16) -C(15)	179.0(16)
C(12) -C(11) -C(16) -C(15)	-2.1(24)
C(11) -C(12) -C(13) -C(14)	-1.3(26)
C(12) -C(13) -C(14) -C(15)	2.7(27)
C(13) -C(14) -C(15) -C(16)	-4 (3)
C(14) -C(15) -C(16) -C(11)	4 (3)

C(5') -S(1') -S(2') -C(1') 105.0(8)
 S(2') -S(1') -C(5') -C(4') -80.8(10)
 S(2') -S(1') -C(5') -C(6') 160.1(9)
 S(2') -S(1') -C(5') -C(7') 43.1(11)
 S(1') -S(2') -C(1') -C(2') -80.2(11)
 C(8') -O(3') -C(17')-O(4') -1.7(24)
 C(8') -O(3') -C(17')-C(4') -176.9(13)
 C(4') -N(1') -C(3') -O(1') -22.5(22)
 C(4') -N(1') -C(3') -C(2') 148.4(13)
 C(3') -N(1') -C(4') -C(5') -104.6(15)
 C(3') -N(1') -C(4') -C(17') 134.3(14)
 C(9') -N(2') -C(2') -C(1') 82.7(16)
 C(9') -N(2') -C(2') -C(3') -158.8(13)
 C(2') -N(2') -C(9') -O(2') 7.2(23)
 C(2') -N(2') -C(9') -C(10') -171.9(12)
 S(2') -C(1') -C(2') -N(2') -175.2(10)
 S(2') -C(1') -C(2') -C(3') 62.0(14)
 N(2') -C(2') -C(3') -O(1') -35.1(19)
 N(2') -C(2') -C(3') -N(1') 153.8(12)
 C(1') -C(2') -C(3') -O(1') 82.9(16)
 C(1') -C(2') -C(3') -N(1') -88.2(14)
 N(1') -C(4') -C(5') -S(1') 55.6(13)
 N(1') -C(4') -C(5') -C(6') 169.1(12)
 N(1') -C(4') -C(5') -C(7') -66.3(16)
 C(17')-C(4') -C(5') -S(1') 177.1(10)
 C(17')-C(4') -C(5') -C(6') -69.4(16)
 C(17')-C(4') -C(5') -C(7') 55.2(16)
 N(1') -C(4') -C(17')-O(3') -155.0(13)
 N(1') -C(4') -C(17')-O(4') 29.7(22)
 C(5') -C(4') -C(17')-O(3') 83.0(16)
 C(5') -C(4') -C(17')-O(4') -92.2(19)
 O(2') -C(9') -C(10')-C(11') -64.0(20)
 N(2') -C(9') -C(10')-C(11') 115.1(16)
 C(9') -C(10')-C(11')-C(12') 133.7(16)
 C(9') -C(10')-C(11')-C(16') -49.8(20)
 C(10')-C(11')-C(12')-C(13') -178.6(17)
 C(16')-C(11')-C(12')-C(13') 4.9(27)
 C(10')-C(11')-C(16')-C(15') 177.9(16)
 C(12')-C(11')-C(16')-C(15') -5.6(25)
 C(11')-C(12')-C(13')-C(14') -4 (3)
 C(12')-C(13')-C(14')-C(15') 3 (3)
 C(13')-C(14')-C(15')-C(16') -4 (3)
 C(14')-C(15')-C(16')-C(11') 5 (3)

H(1A) - N(1) - C(4) -H(041) -158 (9)
 H(2A) - N(2) - C(2) -H(021) 164 (10)
 H(011)- C(1) - C(2) -H(021) 59.7(20)
 H(012)- C(1) - C(2) -H(021) 178.0(16)
 H(1'A)-N(1') -C(4') -H(41') -137
 H(2'A)-N(2') -C(2') -H(21') 121 (9)
 H(11')-C(1') -C(2') -H(21') 64.6(20)
 H(12')-C(1') -C(2') -H(21') -177.3(16)

The two molecules are very similar in conformation, with no significant differences in bond lengths or angles. The largest difference in bond length occurs in the C4-C5 bond, but it is not significant.

C(4)-C(5) 1.565(21)Å C(4')-C(5') 1.478(21) 0.087Å difference

There are no significant differences in torsion angle around the dithiazocine ring, but the amide linkage is distorted significantly from planarity.

	mol. 1	mol. 2	Ideal
C(2)-C(3)-N(1)-C(4)	152.9(13)	148.4(13)	180.00

This can be explained on grounds of ring strain, as this is the smallest ring to contain a *trans* amide linkage. The smallest lactam to have a *trans* amide linkage in the solid state is caprylolactam (C₈H₁₅NO), with a torsion angle of 148.4° in a nine membered ring (Dunitz and Winkler, 1975).

The only significant conformational difference between the two molecules lies in the carbomethoxy group.

Torsion	mol. 1	mol. 2
C(5)-C(4)-C(17)-O(4)	-112.3(17)	-92.2(19)
C(5)-C(4)-C(17)-O(3)	68.9(15)	83.0(16)
N(1)-C(4)-C(17)-O(3)	-172.8(12)	-155.0(13)
N(1)-C(4)-C(17)-O(4)	6.0(21)	29.7(13)

43

C(4)-C(17)-O(3)-C(8)	-173.0(12)	-176.9(13)
O(4)-C(17)-O(3)-C(8)	8.3(23)	-1.7(24)

Intermolecular Contacts

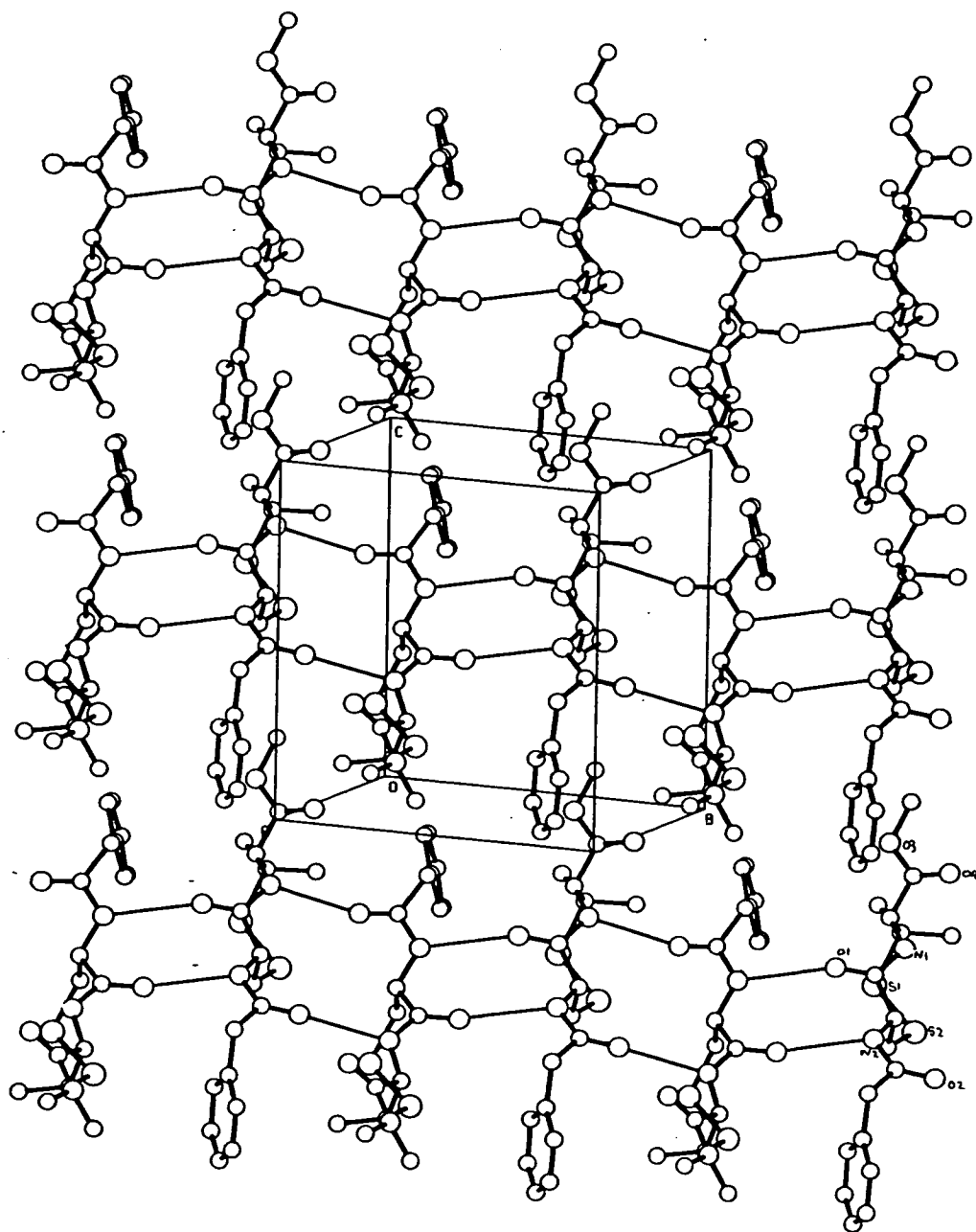
The tables below show all significant intermolecular contacts. They are organised into three tables, on the following basis:

- table 1) All contacts not involving hydrogen
- table 2) Contacts involving well-defined hydrogen atoms such as methene, methine or trigonal hydrogens
- table 3) Contacts involving methyl group hydrogens, (which are not necessarily in the position calculated for them by SHELX as the methyl groups may rotate)

The most important of these contacts are the oxygen-nitrogen hydrogen bonds listed in the first table, which can be seen in the three packing diagrams below.

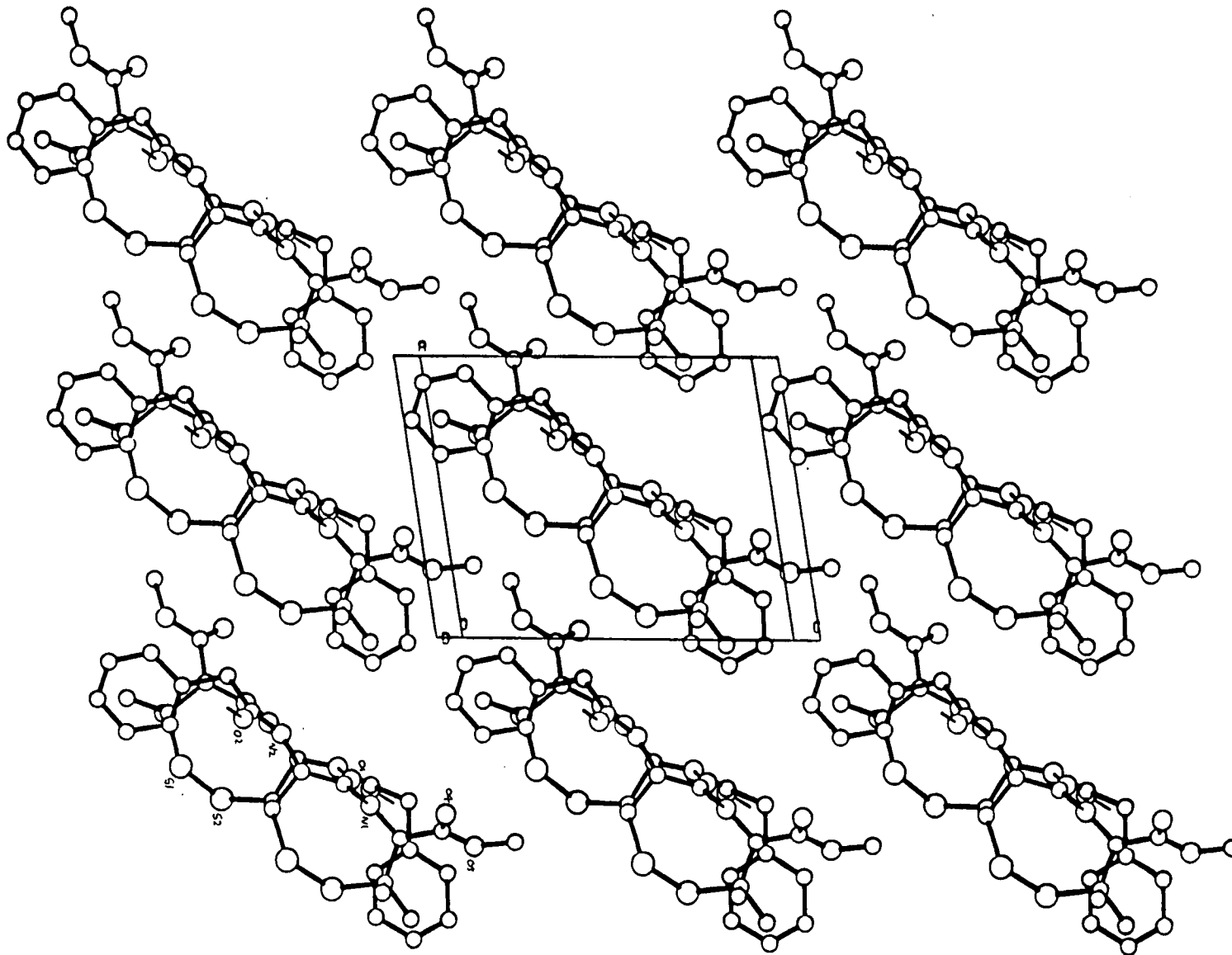
**N-PHENYLACETYL CYSTEINYL VALINE METHYL ESTER
CYCLIC DISULPHIDE**

PROJECTION DOWN THE X- AXIS



***N*-PHENYLACETYL CYSTEINYL VALINE METHYL ESTER
CYCLIC DISULPHIDE**

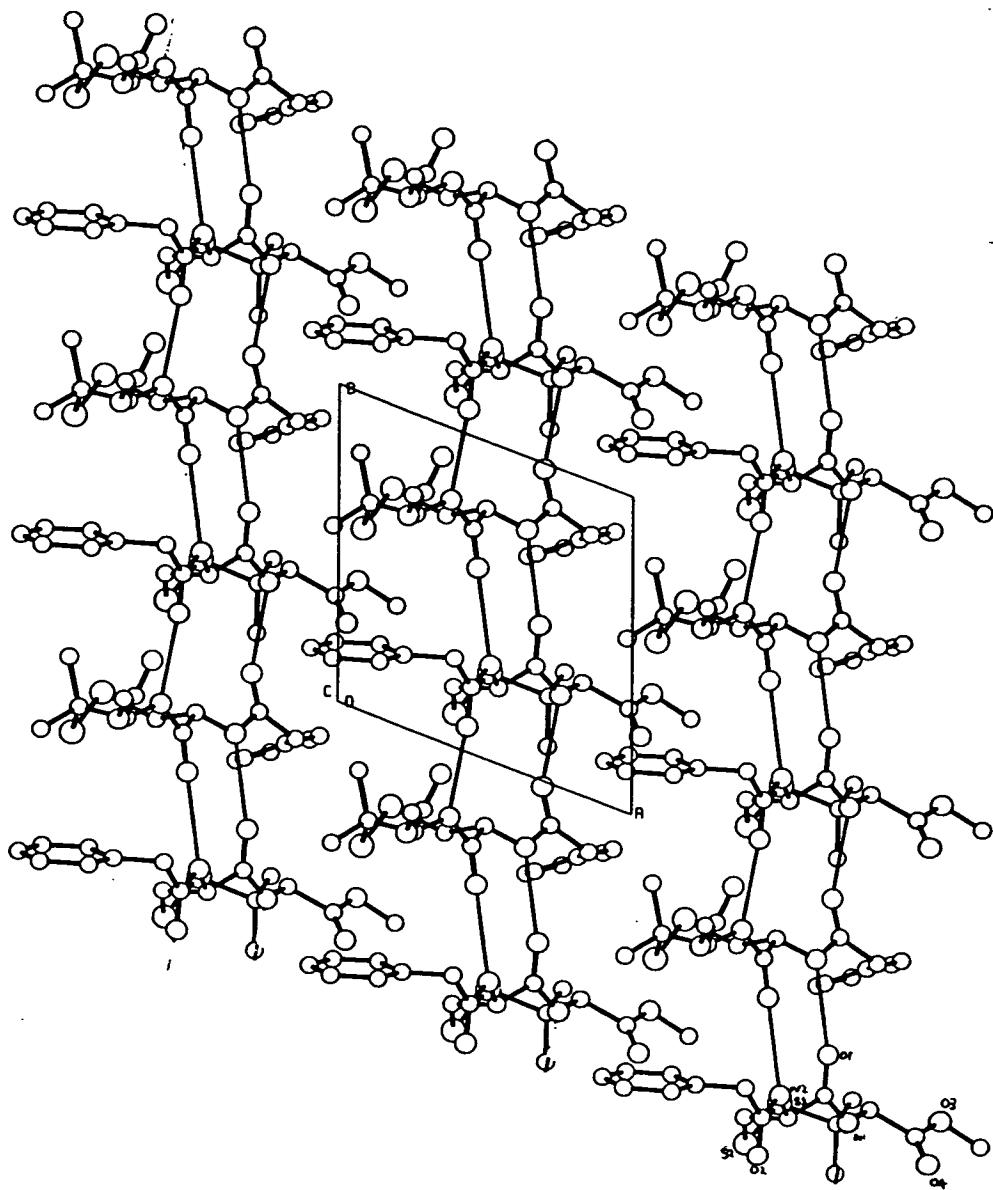
PROJECTION DOWN THE Y-AXIS



4

**N-PHENYLACETYL CYSTEINYL VALINE METHYL ESTER
CYCLIC DISULPHIDE**

PROJECTION DOWN THE Z-AXIS



S2	4.482 (9)	S2'	(X, -1 +Y, Z)
S1	3.938 (12)	04'	(X, -1 +Y, -1 +Z)
S2	3.649 (12)	02'	(X, -1 +Y, Z)
S2	3.837 (13)	04'	(X, -1 +Y, -1 +Z)
S1'	3.479 (12)	01	(-1 +X, Y, Z)
S1'	3.609 (13)	04	(-1 +X, Y, Z)
S2'	3.721 (13)	02	(X, 1 +Y, Z)
S2'	3.899 (13)	03	(-1 +X, Y, Z)
S1'	3.522 (13)	N1	(-1 +X, Y, Z)
S1	3.410 (15)	C8	(-1 +X, Y, Z)
S1	3.857 (18)	C16'	(X, Y, Z)
S2	3.752 (16)	C8	(-1 +X, Y, Z)
S2	3.444 (18)	C8'	(X, -1 +Y, -1 +Z)
S1'	3.620 (16)	C3	(-1 +X, Y, Z)
S1'	3.658 (16)	C4	(-1 +X, Y, Z)
S1'	3.816 (21)	C15	(X, Y, Z)
S1'	3.791 (17)	C16	(X, Y, Z)
S1'	3.611 (17)	C17	(-1 +X, Y, Z)
S2'	3.931 (18)	C17	(-1 +X, Y, Z)
S2'	3.506 (16)	C10'	(-1 +X, Y, Z)
01	2.949 (16)	N2'	(X, Y, Z)
02	2.915 (18)	N1'	(X, -1 +Y, Z)
01'	3.001 (16)	N2	(X, Y, Z)
02'	2.919 (16)	N1	(X, 1 +Y, Z)
01	3.362 (18)	C10'	(X, Y, Z)
02	3.337 (17)	C2'	(X, -1 +Y, Z)
03	3.443 (21)	C8'	(1 +X, Y, -1 +Z)
01'	3.372 (19)	C10	(X, Y, Z)
02'	3.343 (17)	C2	(X, 1 +Y, Z)
02'	3.333 (23)	C15	(1 +X, 1 +Y, Z)
04'	3.491 (20)	C7	(X, 1 +Y, 1 +Z)

table 1

S1	3.207 (16)	H102	(X,Y,-1 +Z)
S1	3.241 (17)	H121	(X,Y,-1 +Z)
S2	3.553 (16)	H11'	(X,-1 +Y,Z)
S2	3.033 (20)	H2'1	(-1 +X,-1 +Y,Z)
S1'	3.304 (16)	H041	(-1 +X,Y,Z)
S1'	3.529 (21)	H151	(X,Y,Z)
S1'	3.524 (17)	H161	(X,Y,Z)
S1'	3.294 (16)	H0'1	(-1 +X,Y,Z)
S2'	3.386 (16)	H011	(X,1 +Y,Z)
S2'	3.405 (16)	H041	(-1 +X,Y,Z)
S2'	2.930 (16)	H0'1	(-1 +X,Y,Z)
S2'	3.180 (16)	H0'2	(-1 +X,Y,Z)
01'	2.44 (14)	H2'A	(X,Y,Z)
01	2.478 (18)	H0'1	(X,Y,Z)
02	2.989 (21)	H141	(1 +X,Y,Z)
02	2.29 (16)	H1'A	(X,-1 +Y,Z)
02	2.412 (17)	H21'	(X,-1 +Y,Z)
04	2.818 (23)	H151	(1 +X,Y,Z)
04	2.999 (20)	H161	(1 +X,Y,Z)
01'	2.01 (14)	H2A	(X,Y,Z)
01'	2.488 (19)	H101	(X,Y,Z)
02'	1.93 (13)	H1A	(X,1 +Y,Z)
02'	2.446 (17)	H021	(X,1 +Y,Z)
02'	2.524 (23)	H151	(1 +X,1 +Y,Z)
03'	2.890 (25)	H5'1	(X,Y,1 +Z)
N1	2.843 (23)	H151	(1 +X,Y,Z)
C8	2.991 (21)	H012	(1 +X,Y,Z)
C13	2.995 (29)	H3'1	(-1 +X,-1 +Y,1 +Z)
C8'	2.887 (23)	H11'	(X,Y,1 +Z)
C13'	2.904 (29)	H131	(1 +X,1 +Y,-1 +Z)
H1A	2.15 (13)	H151	(1 +X,Y,Z)
H021	2.378 (19)	H21'	(X,-1 +Y,Z)

table 2

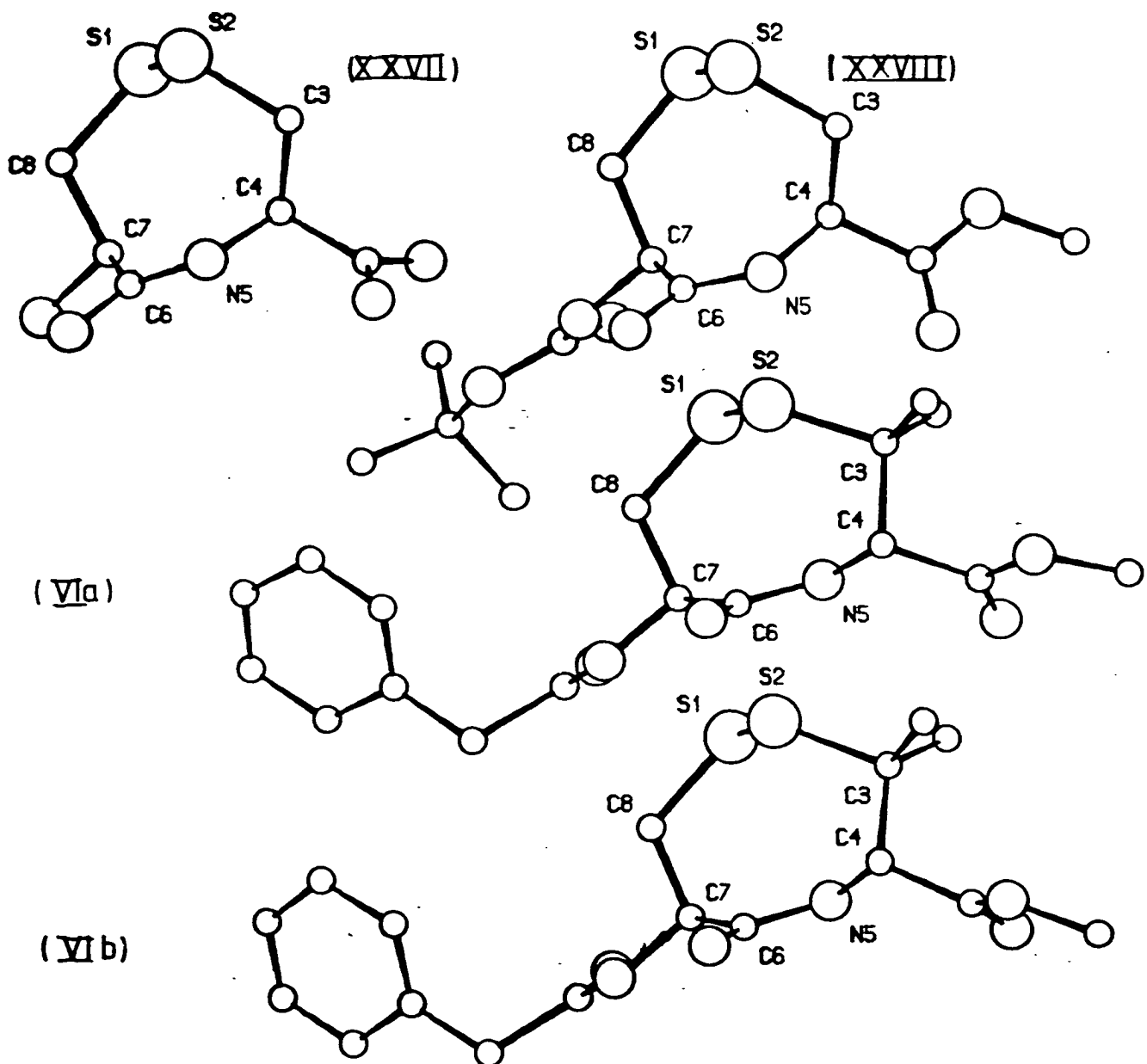
S1	3.335 (15)	H081	(-1 +X,Y,Z)
S1	3.162 (15)	H082	(-1 +X,Y,Z)
S1	3.152 (15)	H083	(-1 +X,Y,Z)
S1	3.477 (18)	H6'1	(X,Y,Z)
S2	3.150 (16)	H082	(-1 +X,Y,Z)
S2	3.414 (16)	H083	(-1 +X,Y,Z)
S2	2.935 (18)	H82'	(X,-1 +Y,-1 +Z)
S2	2.995 (18)	H83'	(X,-1 +Y,-1 +Z)
S2'	3.432 (19)	H81'	(X,Y,-1 +Z)
O2	2.645 (19)	H71'	(X,-1 +Y,Z)
O3	2.434 (21)	H81'	(1 +X,Y,-1 +Z)
O2'	2.605 (19)	H072	(X,1 +Y,Z)
O3'	2.813 (19)	H081	(-1 +X,Y,1 +Z)
O4'	2.551 (20)	H071	(X,1 +Y,1 +Z)
C8	2.732 (23)	H81'	(1 +X,Y,-1 +Z)
C10	2.961 (23)	H062	(X,Y,1 +Z)
C12	2.953 (23)	H72'	(X,-1 +Y,Z)
C14	2.745 (24)	H63'	(X,Y,Z)
C15	2.676 (26)	H63'	(X,Y,Z)
C8'	2.928 (23)	H081	(-1 +X,Y,1 +Z)
C12'	2.957 (25)	H073	(X,1 +Y,Z)
C12'	2.818 (26)	H82'	(1 +X,Y,-1 +Z)
C13'	2.988 (27)	H073	(X,1 +Y,Z)
C13'	2.798 (28)	H82'	(1 +X,Y,-1 +Z)
C14'	2.908 (27)	H061	(X,Y,Z)
C15'	2.659 (28)	H061	(X,Y,Z)
C16'	2.903 (24)	H061	(X,Y,Z)
H012	2.126 (21)	H083	(-1 +X,Y,Z)
H062	2.187 (23)	H102	(X,Y,-1 +Z)
H081	2.333 (23)	H81'	(1 +X,Y,-1 +Z)
H083	2.312 (22)	H12'	(1 +X,Y,Z)
H102	2.187 (23)	H062	(X,Y,1 +Z)

table 3

Comparison With Other Structures

There are only two crystal structures in the literature of cyclic disulphides containing an amide linkage. These are L-cysteinylcysteine cyclic disulphide (XXVII) (Capasso et al, 1977) and N-tert-Butyloxycarbonyl cysteinylcysteine methyl ester cyclic disulphide (XXVIII) (Ashida et al, 1977).

As can be seen from the diagrams, the heterocyclic rings in these compounds have almost exactly the same conformation. They are however significantly different from cyclic disulphide (VI), as they have the 'L' configuration at both chiral centres. This means that a *cis* amide linkage is necessary in order to allow the ring substituents to take up the energetically more favourable equatorial positions.



A comparison of the torsion angles around the ring is of interest, as it shows the essential similarity of conformation at the sulphur bridge in both ring systems.

Torsion	VIa	VIb	XXVI	XXVII
S1-S2-C3-C4	-81.8	-80.8	-48.43	-48.63
S2-C3-C4-N5	51.8	55.6	-54.06	-59.43
C3-C4-N5-C6	-98.5	-104.6	112.44	101.85
C4-N5-C6-C7	152.9	148.4	-7.23	10.80
N5-C6-C7-C8	-92.1	-88.2	-84.2	-89.42
C6-C7-C8-S1	58.0	62.0	87.12	80.30
C7-C8-S1-S2	-76.2	-80.2	-80.66	-89.42
C8-S1-S2-C3	107.7	105.2	94.99	95.68

It is also noticeable that the amide linkages in (XXVII) and (XXVIII) are a lot less strained than the *trans* amide in (VI) as the two C α 's do not need to be so far apart.

It is also useful to compare the sulphur bridges with those of other cystine type compounds. In general, disulphide bridges can be divided into two classes on the basis of the sign of their torsion angles. The type I bridge has C-C-S-S and C-S-S-C torsion angles of approximately + 90°, whilst the type II bridge has angles of about - 90°, and hence opposite chirality.

In both these conformations the distance between the C α 's is a maximum, compared with the few cases in which the bridge conformation is not of these types, for example in the bicyclic cyclo-L-cystine diketopiperazine (Varughese *et al*, 1981; Mez, 1973) or the sulphur bridges in proteins, where they are constrained to be closer for other reasons.

The table below shows the parameters of the sulphur bridges from a selection of cystine derivatives.



COMPOUND CODE	S-S	S-CB	S'SCB	SCBCA	CSSC	SSCC
CLCYST10	2.019	1.818	103.86 103.73	114.97 114.07	-93.53	66.22 61.63
CYLCYS	2.003	1.834 1.787	105.17 104.65	116.33 115.48	-91.44	63.11 67.22
CYSMEC	2.045	1.808 1.805	103.08 99.99	113.61 114.89	-84.44	-77.37 -79.19
CYSTBR01	2.024	1.864	103.89	111.97	-81.34	-88.93
CYSTCL02	2.038	1.801	103.83	115.84	-82.55	-88.45
CYSTIN10	2.044	1.820 1.804	100.08 102.33	112.65 114.62	-79.75	-69.99 -82.35
DGLYCH01	2.037	1.824	103.43	112.75	-84.43	-95.22
LCSTIM	2.051	1.757	102.87	119.09	-81.41	-94.56
LCYSCC	2.040	1.818 1.832	99.52 103.26	113.32 114.07	-79.93	-70.79 -80.62
LCYSTI10	2.032	1.820	104.51	116.21	73.67	81.64
LCYSTI11	2.042	1.800 1.830	104.17 105.74	116.79 115.72	69.32	66.93 75.14
PENICS10	2.049	1.865 1.866	105.40 105.44	109.95 106.83	114.71	67.11 69.61

A key to compound codes giving names of compounds and references will be found at the end of the chapter.

Solution Conformation study

The solid state structure of the cyclic disulphide (VI) shows two independent molecules of similar conformation in the unit cell. To determine whether this similarity was due to packing constraints, or to the rigidity of the molecule itself, a study of the solution

conformation was carried out by nuclear magnetic resonance.

The first step in an nmr study of molecular configuration is to assign each signal to the relevant proton. This can be done by means of irradiation techniques if the complexity of the spectrum warrants it, but in this case it was possible to assign almost all the resonances on the basis of the information contained in one spectrum run in deuteriochloroform at ambient temperature. The table below summarises the assignment of the spectrum. (the two methyl resonances were assigned on the basis of their n.o.e's to C(4)-H and N(1)-H)

Shift/ppm	Assignment	
1.435	C(6)H ₃ singlet	
1.447	C(7)H ₃ singlet	
2.842	A of ABMX system C(1)H ₂	$^2J_{AB} = 14.04 \text{ Hz}$
3.350	B of ABMX system	$^3J_{AM} = 11.19 \text{ Hz}$ $^3J_{BM} = 5.15 \text{ Hz}$
3.570	C(10)H ₂ singlet	
3.741	C(8)H ₃ singlet	
4.555	M of ABMX system C(2)H	$^3J_{MX} = 7.02 \text{ Hz}$
4.890	C(4)H ₂ doublet	$^3J = 11.10 \text{ Hz}$
6.386	X of ABMX system N(2)	
6.642	N(1)H doublet	
7.297	Aromatic multiplet (C(12)-C(16))	

The coupling constants in the above table were of use in the assignment of the spectrum because it is known that the vicinal coupling constant depends on the dihedral angle between the protons involved. This relationship was first quantified for the H-C-C-H moiety (Karplus, 1959; Karplus, 1963; Barfield and Karplus, 1969) as an equation of the form:

$$^3J_{HH} = A + B \cos \theta + C \cos 2\theta$$

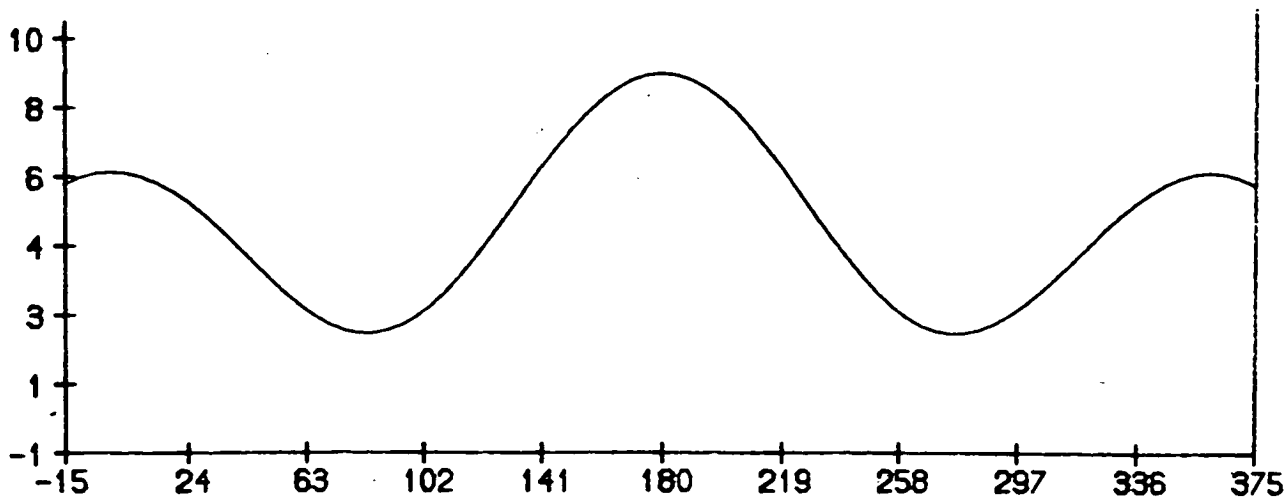
later rewritten as:

$$^3J_{HH} = A \cos^2 \theta + B \cos \theta + C$$

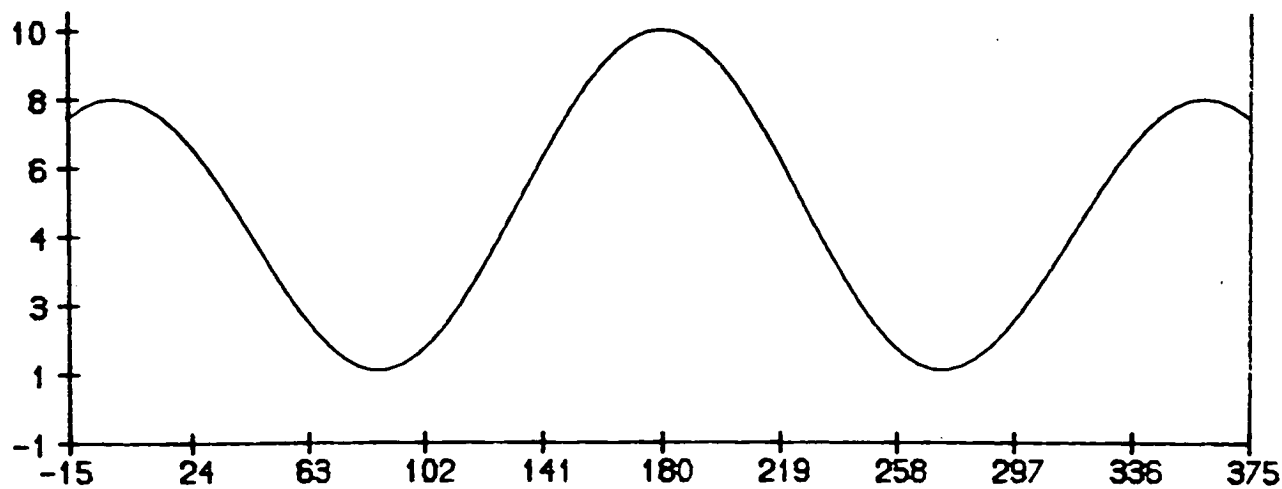
or:

$${}^3J_{\text{HH}} = A \cos^2\theta + B \cos\theta + C \sin^2\theta$$

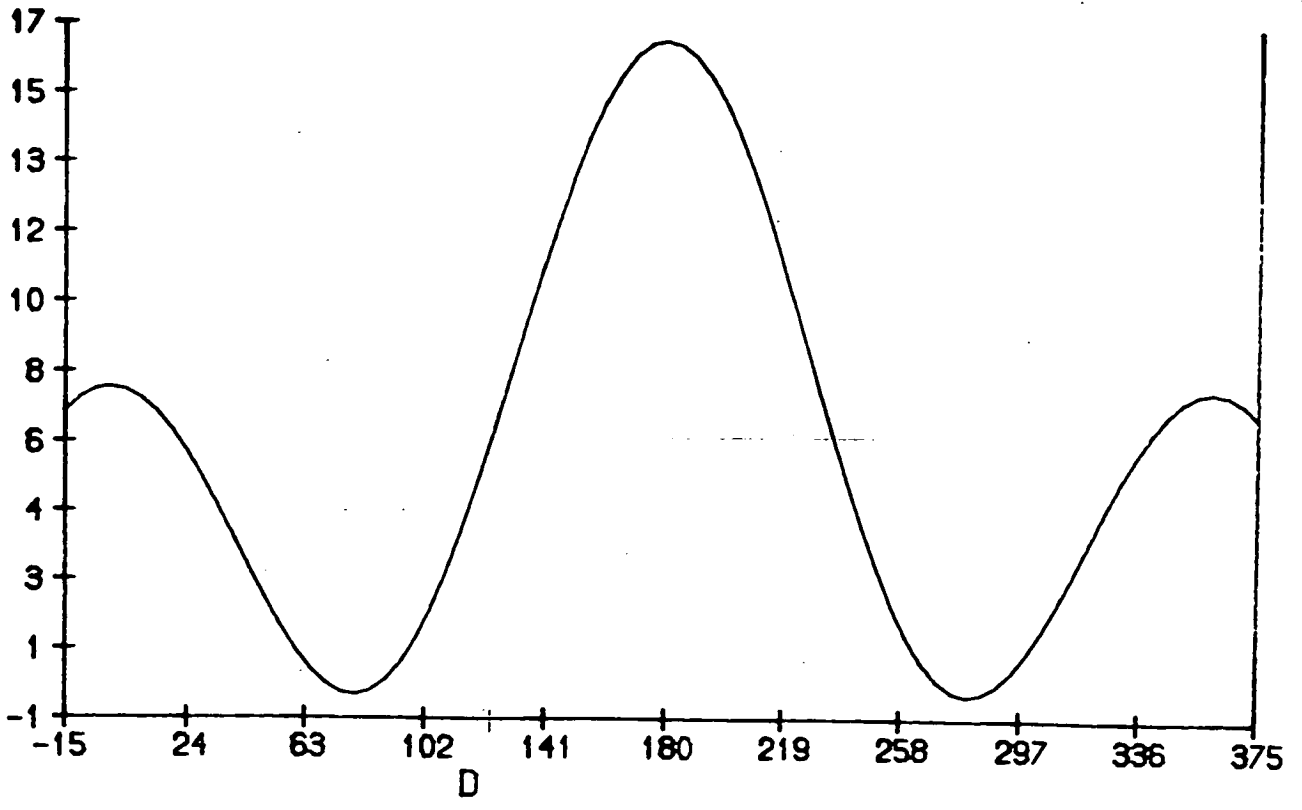
This equation has been used for the H-N-C α -H system by changing the values used for the Karplus constants (Bystrov et al, 1969; Ramachandran et al, 1971; Barfield and Gearhart, 1973; and DeMarco et al, 1978). The graphs below show how coupling constant varies with dihedral angle for each set of parameters used in the Karplus relationship.



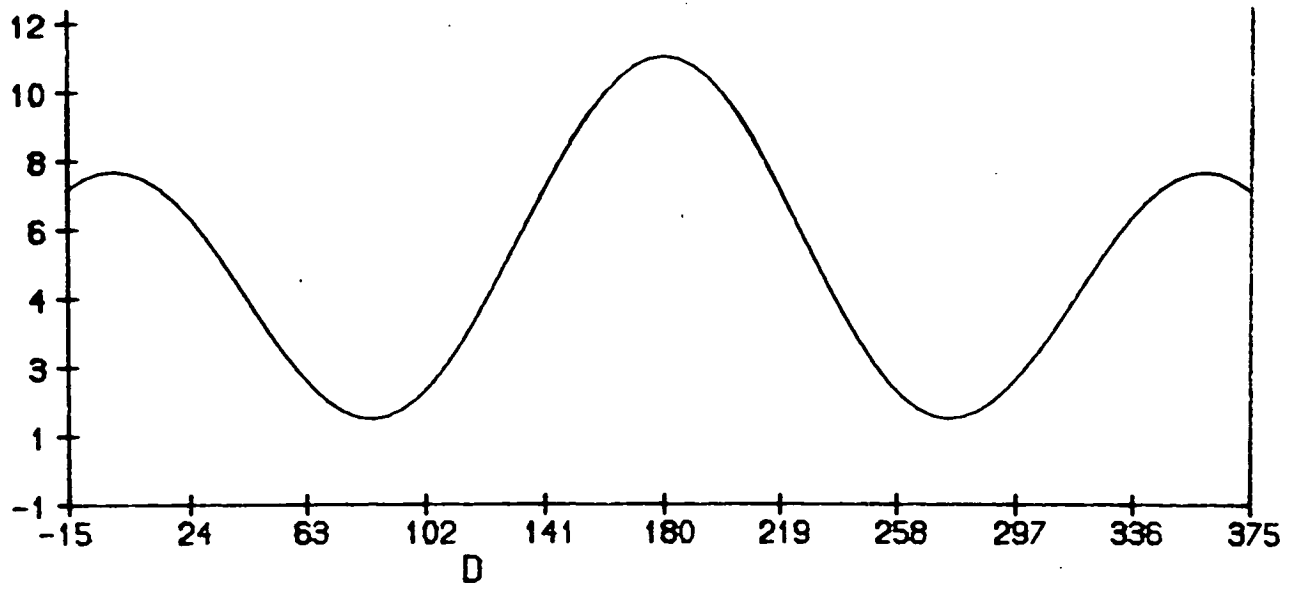
DEMARCO et al



BYSTROV et al



BARFIELD et al



RAMACHANDRAN et al

The method that Bystrov and his co-workers used to calculate values for the constants in the Karplus equation, was to use infra-red spectroscopy to find the conformational states of a range of alanine dipeptides in solution, which could then be correlated with nmr spectra of the same solutions. Each coupling constant measured was taken as an average of the values for the different conformations possible at that site, and the equation below derived.

$${}^3J_{\text{HH}} = (8.9 \pm 0.9) \cos^2 \theta - (0.9 \pm 0.9) \cos \theta + (0.9 \pm 0.9) \sin^2 \theta$$

Ramachandran's group used a similar method, correlating nmr spectra with conformational energy calculations for a range of model compounds. A least squares procedure was then used to calculate the values of the constants in the equation below

$${}^3J_{\text{HH}} = 7.9 \cos^2 \theta - 1.5 \cos \theta + 1.3 \sin^2 \theta$$

This equation was then tested by applying it both to simple model peptides (N-acetylamino acid N-methylamides) and to three cyclic oligopeptide systems. Good agreement was found for the model peptides when their calculated conformations were compared with the coupling constants, unless the side chain extended beyond the C β . In the case of valine, isoleucine, phenylalanine and tyrosine compounds, the observed coupling constants were larger than those expected from the calculation. As the values derived for leucine, methionine and tryptophan fitted the calculated conformational distribution, it was considered that the disagreements were the result of interactions not taken into account in the conformational calculations. The results obtained by checking solution nmr data for the cyclic oligopeptides against independently determined structural parameters were considered to be more satisfactory. The peptides used were ferrichrome, for which an nmr spectrum (Llinas *et al*, 1970) and an X-ray structure (Zalkin *et al*, 1966) have been published, valinomycin, for which the nmr data (Ivanov *et al*, 1969; Ohnishi and Urry, 1969) were compared with conformational calculations (Stern *et al*, 1968) and gramicidin S-A, for which the only structural information available was derived from the nmr data in that paper.

The curve presented by Barfield and Gearhart was derived for the model compound N-methylacetamide by means of self-consistent perturbation theory in the semi-empirical INDO (intermediate neglect of differential overlap) approximation of self-consistent-field molecular orbital theory. Values were calculated for the coupling constant at 30° intervals for both *cis* and *trans* orientations of the amide bond, leading to the equations below, which correspond to within 0.05 Hz of the calculated values:

$$\text{cis } {}^3J_{\text{HH}} = 11.27\cos^2\theta - 4.32\cos\theta + 0.01 \text{ Hz}$$

$$\text{trans } {}^3J_{\text{HH}} = 12.06\cos^2\theta - 4.48\cos\theta + 0.01 \text{ Hz}$$

The method used by DeMarco and Llinas is the most empirical, owing nothing to molecular energy calculations, but being derived directly from a comparison of nmr data with the results of crystallographic analysis (Zalkin *et al*, 1966; Norrestam *et al*, 1975). The use of two independent structures of ferrichrome peptides allows a range of dihedral angles to be associated with each coupling constant, leading to a more accurate set of parameters. The equation they derive (DeMarco *et al*, 1978b) is given by:

$${}^3J_{\text{HH}} = (5.41 \pm 0.02)\cos^2\theta - (1.27 \pm 0.15)\cos\theta + (2.17 \pm 0.54) \text{ Hz}$$

In the same paper, they suggest that seriously nonplanar amide linkages may have higher coupling constants than would ordinarily be calculated by this formula. In a later paper (DeMarco and Llinas, 1980) they present three similar curves derived for different solvents and show that solvent effects on peptides in solution are mainly of conformational origin and are relatively minor. The same group have also produced a set of constants for a Karplus curve to describe the non-amide protons in the ferrichrome series of peptides. (Llinas *et al*, 1978i)

$${}^3J_{\text{HH}} = (9.5 \pm 0.3)\cos^2\theta - (1.6 \pm 0.2)\cos\theta + (1.8 \pm 0.6) \text{ Hz}$$

These equations, all of which are soundly based in theory, give expected values for the H-C(4)-N(1)-H coupling constant in the cyclic disulphide

ranging from 5.7 Hz (Bystrov et al, 1969) to 12.4 Hz (Barfield and Gearhart, 1973), making it quite clear that the curves only give a rough guide to molecular conformation in most cases. It is however possible to deduce that the ring amide linkage in (VI) is *trans*, while the exocyclic amide is best treated as either rotating freely about the bond C(2)-N(2), or having three possible conformations about this bond (3J of 7 Hz can be rationalised as an average of two *gauche* and one *trans* rotamer - $J = 5$ Hz for *gauche* and 11 Hz for *trans*).

In a similar fashion the conformation about the C(1)-C(2) bond can be deduced from the coupling constants, showing that the upfield methene proton must be *trans* to the methine proton.

Useful structural information may also be obtained by comparing spectra of a compound run in two different solvents (Llinas and Klein, 1975; Williams et al, 1982). When the spectrum of the cyclic disulphide (VI) is measured in d_6 -DMSO (perdeutero dimethyl sulphoxide), the most striking change was the downfield shift of the amide proton resonances by over 2 ppm. This is caused by these protons hydrogen bonding to the solvent, indicating that they are exposed to the medium and are involved in neither inter- nor intramolecular hydrogen bonding. The changes in the spectrum on going from $CDCl_3$ to d_6 -DMSO are tabulated below.

Resonance	$CDCl_3$	d_6 -DMSO	$\Delta\delta$ /ppm
C_6H_5	7.297	7.246	-0.051
N(2)-H	6.624	8.808	2.165
N(1)-H	6.387	8.396	2.010
C(4)-H	4.890	4.755	-0.135
C(2)-H	4.555	4.677	0.122
C(8)H ₃	3.741	3.687	-0.054
C(10)H ₂	3.570	3.310	-0.260
C(1)H(<i>gauche</i>)	3.350	3.137	-0.213
C(1)H(<i>trans</i>)	2.842	2.901	0.059

The effects of variation in temperature on the amide proton resonances provide confirmatory evidence for the solvent exposure of

these protons. It has been shown (Llinas and Klein, 1975) that the resonance of a hydrogen bonding proton exposed to the solvent shifts upfield on warming as hydrogen bonding to the solvent is thermally disrupted.

peak	CDCl ₃	d ₆ -DMSO
N(2)-H	-7.5(8)	-6.21(6)
N(1)-H	-0.1(11)	-3.83(8)

temperature coefficients in ppm/°C * 10³

The table above shows the temperature coefficients of the amide protons in d₆-DMSO and in CDCl₃. These are calculated by measuring the chemical shift of the amide protons at various temperatures and then calculating a best-fit line through the plot of chemical shift versus temperature. In this case the temperature range covered was 20 - 80°C in d₆-DMSO and -40 - 49°C in CDCl₃, rising in 30°C increments in both cases. As would be expected, the coefficients for DMSO show hydrogen bonding to the solution, and the coefficient for the penicillamine amide (N(1)-H) in chloroform shows no hydrogen bonding. It is possible that at low temperatures the compound exists in chloroform solution as an hydrogen bonded dimer, as the coefficient for the cysteine amide (N(2)-H) does seem to show a loss of hydrogen bonding on warming.

The final method used to examine the solution conformation of the cyclic disulphide (VI) was nuclear Overhauser effect difference spectroscopy, which can be used to obtain information on the relative proximities of protons (Williams et al, 1982). A brief description of the theory of the nuclear Overhauser effect is as follows (Sanders and Mersh, 1982).

In the nmr experiment, excess spin population is induced to move between energy levels by an appropriate frequency of electromagnetic radiation, this is followed by a radiationless return to equilibrium, which requires fluctuating magnetic fields of appropriate frequencies. This is

spin-lattice relaxation. The main sources of this relaxation for diamagnetic organic molecules are the magnetic dipoles of protons in the same molecule (dipole-dipole interactions). The efficiency of this relaxation depends on the magnitude and frequency of the generated fluctuating magnetic fields. This in turn depends on the length and correlation time (or tumbling time) τ_c of the vector connecting the nuclei. Nuclear Overhauser effects, and their growth or decay rates, are essentially measures of the strength of dipole-dipole interactions between the spins involved. This is why they depend on interatomic distance and correlation time, giving them their power as structural probes. It can be demonstrated that the n.O.e. is a kinetic effect, as it can take some time to build up after the start of the irradiation causing it, and it persists for some time when the irradiating frequency is switched off. It can also be shown that for a rigid molecule tumbling isotropically, the rate at which the n.O.e. builds up and decays is inversely proportional to the sixth power of the internuclear distance. The constant of proportionality is related to the correlation time for any given system. There are a few specific occasions when the picture is not quite this simple. These are:

- 1) Nuclei with close proton neighbours are relaxed quickly and when observed in an n.O.e. experiment will reveal little long-range information; what long-range information is available will take a while to develop. Methyl and methene protons are a good example of this situation.
- 2) Nuclei without nearby neighbours obtain relaxation and n.O.e.'s from distant protons. This causes them to relax slowly, and when observed in an n.O.e. experiment they show large, slowly growing n.O.e.'s from distant protons.

As a consequence of these points, a proton equidistant to a methine and a methene proton will in general give a larger n.O.e. to the former. Methyl groups have even smaller n.O.e.'s than those of methene groups. The most reliable measure of distance is to observe n.O.e.'s at a methine proton whilst irradiating at a methyl or methene proton and not *vice versa*. The single most important application of the n.O.e. is

its ability to make through-space connections between nuclei independently of chemical bonding considerations, but there are cases where the n.O.e is too small to be visible with conventional methods. In these circumstances, use of difference spectroscopy allows the measurement of n.O.e's down to a fraction of one percent.

The usual methodology for a difference n.O.e experiment is as follows:

- 1) A data file is prepared containing the frequencies of all resonances to be irradiated.
- 2) A subfile is created for each irradiating frequency (and possibly also an extra one for the control spectrum, which is not irradiated).
- 3) The pulse sequence is started with a slight delay.
- 4) A weak 'tickling' decoupling power is switched on at one of the irradiating frequencies.
- 5) One 'scan' or accumulation of data is collected.
- 6) The above sequence is repeated for the next irradiating frequency.
- 7) When a scan has been collected into each subfile, the next scan is recorded at each frequency of irradiation.
- 8) The entire sequence is then repeated until each subfile contains ca 1000 scans.

The reason for collecting the spectral information in this piecemeal, scan by scan, manner is that it minimises perturbation effects caused by continual or prolonged irradiation at the same frequency. When the control spectrum is subtracted from each irradiated spectrum, the resultant trace shows only the changes caused by irradiation at that frequency.

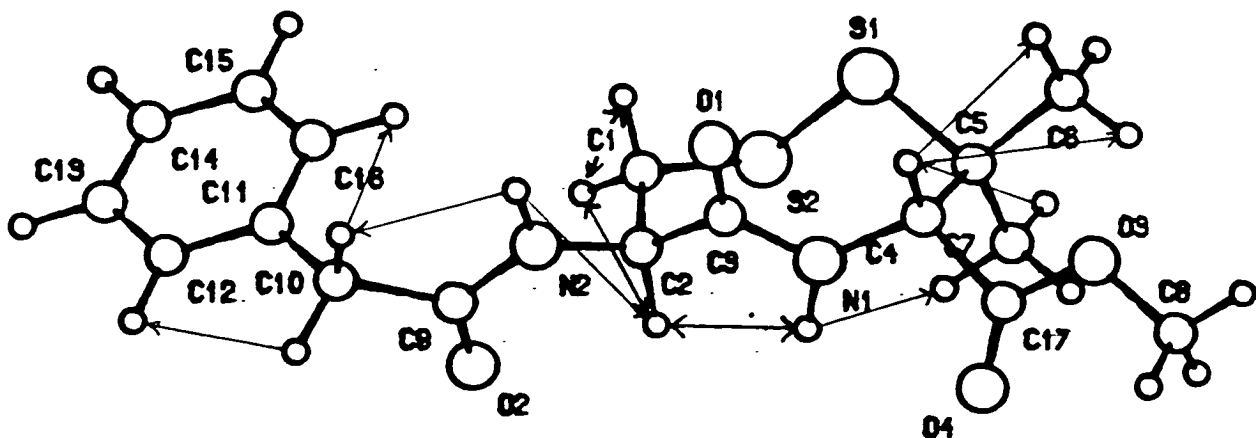
When this pulse sequence was carried out on the cyclic disulphide (VI), the following n.O.e's were observed.

Irradiated peak

Effect

N(2)-H	Small effect on C(2)-H and C(10)H ₂
N(1)-H	Larger effect on C(2)-H and C(7)H ₃ (small)
C(2)-H	N(1)-H and pro-S proton of C(1)H ₂
C(4)-H	Slight effect on C(6)H ₃
C(1)-H trans	Effect on pro-S proton of C(1)H ₂
C(1)-H gauche	Effect on pro-R proton of C(1)H ₂ and on C(2)-H
C(10)H ₂	Slight effect on aromatic protons and C(1)H ₂
C(8)H ₂	None
C(7)H ₃ or C(6)H ₃	slight effect on C(4)-H

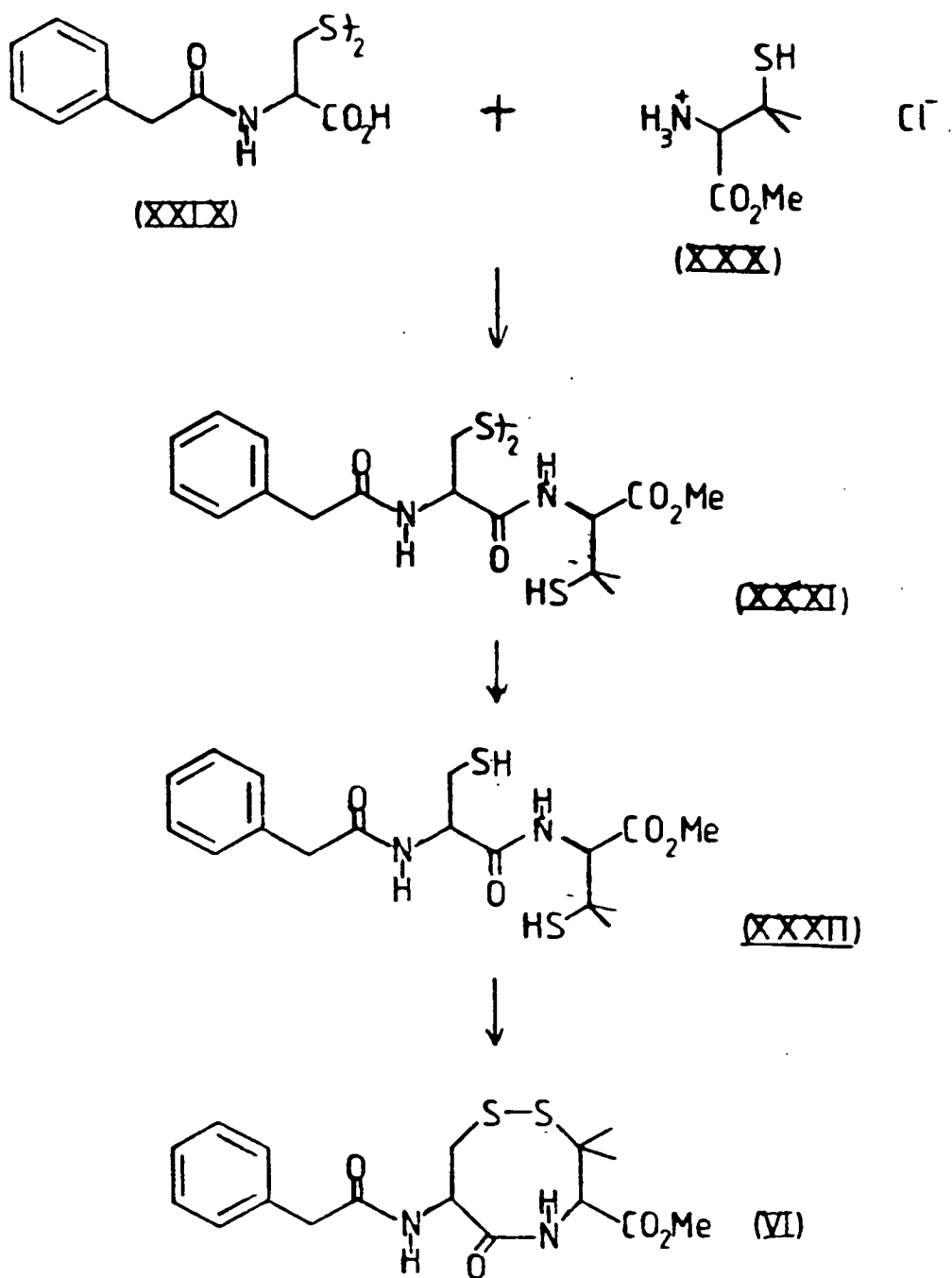
All these results are consistent with the solid state structure, and it may be stated with confidence that the cyclic disulphide (VI) exists in solution as monomers of one conformation only. The diagram below shows the measured n.o.e.'s graphically.



Experimental

(Synthesis of the cyclic disulphide (VI) as carried by E. M. Gordon, 1973)

The course of this synthesis is illustrated in the following diagram.



Nmr spectra were measured on Perkin Elmer EM 360, Bruker WM 300 or WM 200 spectrometers as deuteriochloroform solutions using deuteriochloroform as an internal standard for the FT machines, or tetramethylsilane for the CW machine. Optical rotations were recorded on a Mettler 141 polarimeter. Thin layer chromatography (t.l.c.) was carried out on 0.25mm layers of Merck GF₂₅₄ silica. Melting points are uncorrected. Organic solvents were purified and dried by established procedures, and organic extracts typically dried over MgSO₄.

D-Penicillamine methyl ester hydrochloride (XXX) was prepared from D-Penicillamine (Aldrich Chemical Co. Inc.) by the standard method (Clarke, Johnson and Robinson, 1949) m.p. 178-180°C (lit. 180-182); $[\alpha]_D -10.39^\circ$ (c 2.H₂O) (lit. $[\alpha]_D -7.50^\circ$, $[\alpha] -10.82^\circ$)

N,N'-bisphenylacetyl-L-cystine (XXIX) was prepared by the method of Foldi (1954) and had m.p. 171°C (lit 170-171°C)

N,N'-bisphenylacetyl-L-cystinyl-bis(D-penicillamine methyl ester) (XXXI)

To a solution of the acid (XXIX)(2.38g, 5mM) and D-penicillamine methyl ester hydrochloride (XXX) (1.63g, 10mM) in tetrahydrofuran (140 cm³), triethylamine (1.5 cm³) and N-ethyloxycarbonyl-2-ethyloxy-1,2-dihydroquinoline (EEDQ) (2.6g, 10.5mM) in tetrahydrofuran (40 cm³) were added sequentially and the solution stirred at room temperature for 24hrs. The solution was evaporated and the residue dissolved in ethyl acetate (50 cm³), washed sequentially with aqueous HCl (1M, 20 cm³ * 2), sat. aqueous NaHCO₃ (20 cm³ * 2) and water (20 cm³ * 2) before being dried and evaporated. The residue was chromatographed on 2 * 200 * 200 mm silica plates, eluting with benzene:ethyl acetate (2:1) to afford crude (XXXI) which was crystallised from CHCl₃:hexanes (1.91g, 2.45mM) mp 208°C

3,3-dimethyl-4-D-carbomethoxy-6-oxo-7-L-phenylacetyl-amino
perhydro-1,2,5-dithiazocine (VI)

The dipeptide (XXXI) (1.00g, 1.28mM) was dissolved in MeOH:aq. 5M HCl (120 cm³, 5:1) at 0°C and Zn dust (10g) added to the cooled, stirred liquid over a period of 10 min. After a further 30 min. the solution was filtered, concentrated *in vacuo* to 30 cm³ and poured into cold aqueous NaOH. The basic solution was washed with CHCl₃ (50 cm³) then taken to pH 1 with conc. HCl, and the acidic solution extracted with CHCl₃ (100 cm³ * 3). Evaporation of the CHCl₃ afforded free dithiol dipeptide (XXXII) as a colourless oil (0.97g).

The dithiol (XXXII) was then dissolved in MeOH (70 cm³), iodine (0.64g) added, the solution heated under reflux for 2hrs, then concentrated *in vacuo* to 20 cm³ and ethyl acetate (50 cm³) added. The resultant solution was washed sequentially with sat. aqueous Na₂S₂O₃ (50 cm³ * 2), aqueous HCl (1M, 50 cm³), aqueous NaOH (1M, 50 cm³) and water (50 cm³). Evaporation of the ethyl acetate layer and crystallisation of the residue from CHCl₃/petrol afforded the cyclic disulphide (VI) (0.60g, 1.57mM) m.p.210°C transition, remelting at 218-220° C. (found C 53.37; H 5.81; N 7.29; S 16.61. C₁₇H₂₂N₂O₂S₂ requires C 53.40; H 5.80; N 7.33; S 16.73)

¹Hnmr see text; MS (EI) m/e 382 (M⁺, exact mass 382.1021; required 382.1022 for C₁₇H₂₂N₂O₂S₂)

Key to compound codes

- CLCYST10 cyclo-L-cystine
- K. I. Varughese, C. T. Lu, G. Kartha, Int. J. Pept. Protein Res., 18, 88, 1981
- CYLCYS cyclo-L-cystine acetic acid solvate
- H.-C. Mez, Cryst. Struct. Commun., 3, 657, 1974
- CYSMEC L-cystine dimethyl ester dihydrochloride monohydrate
- B. K. Vijayalakshmi, R. Srinivasan, Acta Cryst., B31, 993, 1975
- CYSTBR01 L-cystine dihydrobromide
- J. Peterson, L. K. Steinrauf, L. H. Jensen, Acta Cryst., 13, 104, 1960
- CYSTCL02 L-cystine dihydrochloride
(Neutron study)
- D. D. Jones, I. Bernal, M. N. Frey, T. F. Koetzle, Acta Cryst., B30, 1220, 1974
- CYSTIN10 L-cystine dihydrobromide dihydrate
- R. E. Rosenfield Jnr., R. Parthasarathy, Acta Cryst., B31, 816, 1975
- DGLYCH01 N, N'-diglycyl-L-cystine dihydrate
(Re-refinement)
- W. C. Stallings Jnr., J. Donohue, Acta Cryst., B32, 1916, 1976
- LCSTIM L-cystine diamide dihydrochloride
- M. O. Chaney, L. K. Steinrauf, Acta Cryst., B24, 1564, 1968

67

LCYSCC L-cystine dihydrochloride dihydrate
 (Copper(II) doped)

S. Kominami, P. Riesz, T. Akiyama, J. V. Silverton, J. Phys. Chem., 80,
203, 1976

LCYSTI10 L-cystine
 (Hexagonal form)

B. M. Oughton, P. M. Harrison, Acta Cryst., 12, 396, 1959

LCYSTI11 L-cystine
 (Tetragonal form)

M. O. Chaney, L. K. Steinrauf, Acta Cryst., B30, 711, 1974

PENICS10 D-penicillamine disulphide dihydrochloride
 (Absolute configuration)

R. E. Rosenfield Jnr., R. Parthasarathy, Acta Cryst., B31, 462, 1975

Table Fractional Coordinates of Atoms with Standard Deviations

	x	y	z	U _{eq}
S(1)	0.52140(0)	0.27580(0)	0.17070(0)	0.0562
S(2)	0.4055(6)	0.1064(6)	0.2599(5)	0.0564
O(1)	0.6872(12)	0.4844(11)	0.4510(9)	0.0477
O(2)	0.4466(13)	0.0798(11)	0.6796(11)	0.0723
O(3)	1.0667(11)	0.4041(11)	0.2033(10)	0.0584
O(4)	1.0290(12)	0.2567(12)	0.3493(10)	0.0677
N(1)	0.7577(12)	0.2898(12)	0.3825(10)	0.0335
N(2)	0.5234(14)	0.3003(13)	0.6128(11)	0.0379
C(1)	0.4000(16)	0.1884(16)	0.4088(13)	0.0418
C(2)	0.5514(14)	0.2379(14)	0.4996(12)	0.0284
C(3)	0.6752(16)	0.3506(16)	0.4472(13)	0.0357
C(4)	0.8302(16)	0.3548(16)	0.2845(13)	0.0355
C(5)	0.7192(16)	0.2723(16)	0.1605(13)	0.0454
C(6)	0.7715(19)	0.3670(17)	0.0580(14)	0.0622
C(7)	0.7162(18)	0.1138(16)	0.1328(14)	0.0538
C(8)	1.2034(16)	0.3739(16)	0.1866(14)	0.0528
C(9)	0.4601(16)	0.2083(18)	0.6916(13)	0.0315
C(10)	0.4026(17)	0.2792(17)	0.7977(14)	0.0520
C(11)	0.2286(15)	0.2284(15)	0.7770(12)	0.0369
C(12)	0.1462(19)	0.1779(17)	0.8703(16)	0.0603
C(13)	-0.0196(20)	0.1293(18)	0.8517(17)	0.0671
C(14)	-0.0923(20)	0.1360(17)	0.7457(15)	0.0571
C(15)	-0.0116(22)	0.1830(20)	0.6521(18)	0.0733
C(16)	0.1456(18)	0.2316(16)	0.6694(14)	0.0526
C(17)	0.9880(17)	0.3290(17)	0.2849(15)	0.0351
S(1')	0.0853(6)	0.5756(6)	0.5681(5)	0.0649
S(2')	0.1872(7)	0.7368(7)	0.4659(5)	0.0776
O(1')	0.4798(12)	0.5918(10)	0.6672(9)	0.0557
O(2')	0.7000(12)	0.9811(11)	0.4206(10)	0.0537
O(3')	0.2426(13)	0.6723(12)	0.9931(9)	0.0607
O(4')	0.3569(14)	0.8986(11)	0.9428(10)	0.0743
N(1')	0.3862(15)	0.7732(12)	0.7240(13)	0.0412
N(2')	0.6445(14)	0.7709(14)	0.5047(11)	0.0380
C(1')	0.3715(16)	0.7186(16)	0.4439(13)	0.0478
C(2')	0.5069(14)	0.7878(14)	0.5461(11)	0.0298
C(3')	0.4641(15)	0.7091(16)	0.6545(12)	0.0301
C(4')	0.2695(15)	0.6860(15)	0.7884(12)	0.0318
C(5')	0.1086(17)	0.6634(16)	0.7258(14)	0.0454
C(6')	-0.0189(17)	0.5476(17)	0.7765(14)	0.0591
C(7')	0.0743(17)	0.8097(15)	0.7264(13)	0.0489
C(8')	0.2596(19)	0.7421(18)	1.1154(15)	0.0661
C(9')	0.7268(18)	0.8689(18)	0.4383(14)	0.0436
C(10')	0.8550(16)	0.8304(17)	0.3842(13)	0.0420
C(11')	0.8185(17)	0.8017(16)	0.2507(14)	0.0429
C(12')	0.9374(21)	0.8688(19)	0.1828(17)	0.0710
C(13')	0.904(3)	0.8455(22)	0.0589(19)	0.0917
C(14')	0.7673(23)	0.7571(20)	0.0046(19)	0.0830
C(15')	0.6483(25)	0.6887(23)	0.0659(18)	0.0856
C(16')	0.6752(19)	0.7057(18)	0.1923(16)	0.0674
C(17')	0.2987(18)	0.7651(20)	0.9163(16)	0.0483

Table Fractional Coordinates of Atoms with Standard Deviations

	x	y	z	Ueq
H(1A)	0.7570	0.1924	0.4130	0.0500
H(2A)	0.5141	0.3921	0.6189	0.0500
H(011)	0.3108	0.1062	0.4439	0.0436
H(012)	0.3677	0.2847	0.3986	0.0436
H(021)	0.5922	0.1485	0.5188	0.0565
H(041)	0.8498	0.4727	0.2992	0.0483
H(061)	0.7721	0.4783	0.0812	0.0537
H(062)	0.6902	0.3142	-0.0265	0.0537
H(063)	0.8883	0.3739	0.0489	0.0537
H(071)	0.6386	0.0602	0.0469	0.0465
H(072)	0.6742	0.0510	0.2031	0.0465
H(073)	0.8336	0.1179	0.1287	0.0465
H(081)	1.2572	0.4383	0.1208	0.0547
H(082)	1.1700	0.2561	0.1547	0.0547
H(083)	1.2864	0.4042	0.2715	0.0547
H(101)	0.4507	0.3996	0.8042	0.0494
H(102)	0.4447	0.2472	0.8815	0.0494
H(121)	0.2114	0.1758	0.9575	0.0501
H(131)	-0.0811	0.0908	0.9234	0.0699
H(141)	-0.2191	0.0996	0.7302	0.0631
H(151)	-0.0769	0.1851	0.5649	0.0700
H(161)	0.2071	0.2701	0.5977	0.0511
H(1'A)	0.4273	0.8476	0.7230	0.0500
H(2'A)	0.6741	0.7130	0.5505	0.0500
H(11')	0.4035	0.7696	0.3658	0.0367
H(12')	0.3510	0.6007	0.4267	0.0367
H(21')	0.5353	0.9057	0.5744	0.0462
H(41')	0.3082	0.6060	0.8081	0.0339
H(61')	-0.1340	0.5350	0.7280	0.0561
H(62')	-0.0082	0.5850	0.8713	0.0561
H(63')	-0.0040	0.4413	0.7659	0.0561
H(71')	0.1635	0.8913	0.6908	0.0521
H(72')	0.0756	0.8531	0.8180	0.0521
H(73')	-0.0399	0.7869	0.6710	0.0521
H(81')	0.2107	0.6571	1.1696	0.0484
H(82')	0.1979	0.8188	1.1144	0.0484
H(83')	0.3829	0.8024	1.1535	0.0484
H(0'1)	0.8644	0.7315	0.4172	0.0487
H(0'2)	0.9661	0.9230	0.4137	0.0487
H(2'1)	1.0534	0.9417	0.2306	0.0492
H(3'1)	0.9938	0.8980	0.0090	0.0723
H(4'1)	0.7412	0.7366	-0.0931	0.0692
H(5'1)	0.5323	0.6158	0.0182	0.0750
H(6'1)	0.5853	0.6532	0.2422	0.0497

Thermal Vibration Parameters with Standard Deviations

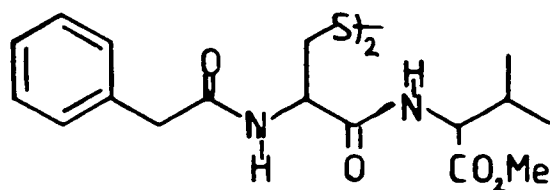
	U11	U22	U33	U23	U13	U12
S(1)	0.0393(26)	0.0747(33)	0.0482(31)	0.0092(26)	0.0056(24)	0.0334(25)
S(2)	0.0419(28)	0.0503(29)	0.0608(34)	-0.0143(24)	0.0063(23)	0.0077(24)
O(1)	0.0588(65)	0.0294(59)	0.0503(66)	0.0017(50)	0.0285(55)	0.0206(49)
O(2)	0.0890(83)	0.0360(64)	0.0913(93)	0.0297(63)	0.0470(74)	0.0371(58)
O(3)	0.0384(60)	0.0602(71)	0.0745(79)	0.0166(62)	0.0259(59)	0.0307(54)
O(4)	0.0569(67)	0.0829(81)	0.0585(73)	0.0163(65)	0.0096(57)	0.0467(63)
N(1)	0.0275(62)	0.0309(70)	0.0431(78)	0.0125(61)	0.0141(59)	0.0236(56)
N(2)	0.0473(73)	0.0203(66)	0.0415(81)	0.0067(64)	0.0182(65)	0.0139(62)
S(1')	0.0439(26)	0.0762(36)	0.0535(33)	-0.0249(27)	0.0039(23)	0.0092(26)
S(2')	0.0537(31)	0.1216(44)	0.0492(35)	0.0124(33)	0.0042(26)	0.0533(31)
O(1')	0.0795(75)	0.0245(56)	0.0651(76)	0.0198(51)	0.0282(61)	0.0415(54)
O(2')	0.0625(65)	0.0314(58)	0.0646(77)	0.0169(54)	0.0255(60)	0.0272(50)
O(3')	0.0900(79)	0.0562(69)	0.0260(63)	0.0098(57)	0.0223(59)	0.0304(60)
O(4')	0.1141(102)	0.0327(69)	0.0483(75)	-0.0047(56)	0.0008(67)	0.0034(64)
N(1')	0.0453(85)	0.0242(74)	0.0469(82)	0.0016(75)	0.0143(71)	0.0119(68)
N(2')	0.0328(68)	0.0429(80)	0.0373(79)	0.0174(64)	0.0061(63)	0.0276(59)

CHAPTER FOUR
BIS-(N-PHENYLACETYL) CYSTINYL BIS(VALINE METHYL ESTER)

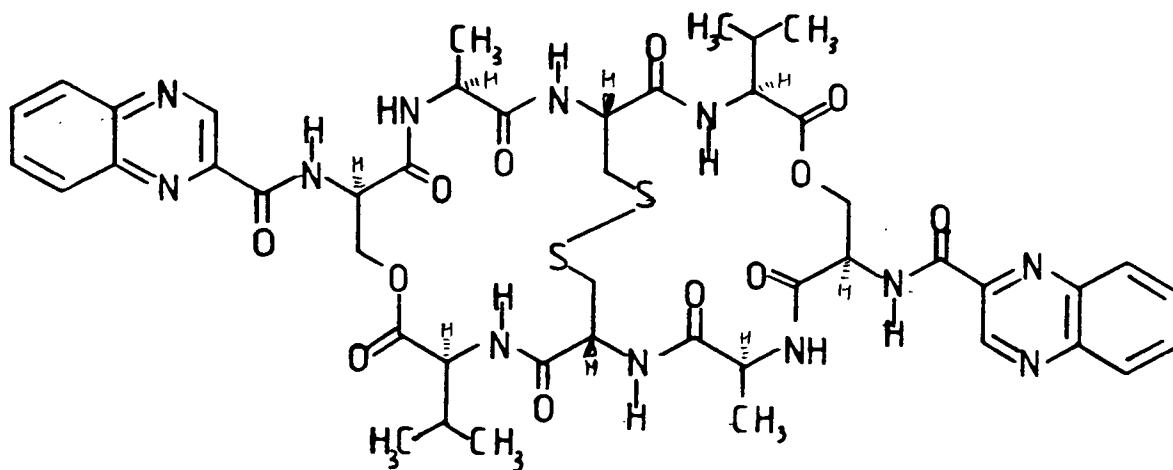
BIS(N-PHENYLACETYL) L-CYSTINYL BIS(L-VALINE METHYL ESTER)

Introduction

This compound (VIII) was a synthetic intermediate prepared during a study on the biosynthesis of β -lactam antibiotics. It is of structural interest because of its similarity to the cysteinyl penicillamine cyclic disulphide described previously. The central portion of this molecule is also similar in conformation to the anti-tumour drug Triostin and its synthetic analogue TANDEM (XXXIV) or tetra N-desmethyl Triostin A (Williams et al, 1982), although this compound is not expected to have any such activity, as Triostin functions by intercalating two quinoxaline chromophores into the minor groove of the DNA helix between the adenine and thymine bases (Viswamitra et al, 1981).



(VIII)



(XXXIV)

Preparation and crystallisation

The compound was prepared by Dr. E. M. Gordon, according to the procedure described in the experimental section below. The crystal used for both preliminary photographs and data collection was grown from a solution of the compound in a slight excess of chloroform, to which enough petrol-ether was added to cause a slight turbidity of the solution. The solution was then left for two days at room temperature.

Crystal Data

Formula $C_{34}H_{46}N_4O_8S_2$ Mol.Wt. = 702.84 $F(000) = 5984$

Space Group $P4_3 2_1 2$ Int. Tab. No. 96
 Cell dimensions $a = 21.534(5)$ $c = 34.507(12)$
 $D_c = 1.167 \text{ gcm}^{-3}$ $D_m = 1.166 \text{ gcm}^{-3}$
 $Z = 16$ Vol. = 16001.3 cubic Angstroms
 Radiation $\text{MoK}\alpha$ $\lambda = 0.71069$ Angstroms $\mu = 1.73 \text{ cm}^{-1}$
 Final R = 0.0912 based on 2201 data
 $h_{\text{max}} = 16$ $k_{\text{max}} = 23$ $l_{\text{max}} = 37$

$\theta_{\text{max}} = 22.5^\circ$ $\text{Sin}(\theta_{\text{max}}) = 0.3867$ $\text{Sin}(\theta_{\text{max}})/\lambda = 0.5385$

Reflections measured 5934 (36 systematically absent)

After merging 5877

Unobserved 3676 ($F < 4\sigma(F)$)

No. parameters refined = 396

Max. Shift/ σ in last cycle = 0.452

Final difference Fourier Max. = 0.3847 eA^{-3} Min. = -0.2655 eA^{-3}

Space Group Determination

The space group was determined to be one of the enantiomorphous pair $P4_3 2_1 2$ or $P4_1 2_1 2$ from the systematic absences on the preliminary photographs (conditions for reflection: $00l, l=4n$; $h00, h=2n$). These space groups differ in the sense of rotation of the 4-fold screw axis and it is only possible to distinguish them when the absolute configuration of the compound is known. In this case the choice was made on the basis of the known configuration of the cystinyl α carbon.

Data Collection and Reduction

The data set was collected on the CAD4 and the program CADABS was used to correct the data for Lorentz and polarisation effects.

Structure Solution and Refinement

Due to the large number of atoms in the asymmetric unit (96 non-hydrogen atoms in four half-molecules, of which 4 were sulphur atoms), it was not possible to solve the structure by means of the Patterson map. Accordingly, the program NORMAL in the MULTAN suite of programs was used to produce a set of normalised structure factors. An automatic run of the MULTAN system was unsuccessful in solving the structure, but normalisation of the data with the number of atoms in the unit cell set to a small fraction of the actual number, revealed the positions of the four sulphur atoms. This may be because a lower value for the total electron density in the cell leads to an upgrading of all phase relationships and an emphasis on the contributions from heavy atoms. The coordinates of these sulphur atoms were used as a starting point for the application of the DIRDIF procedure. Although the first run of the program led to a recognisable molecular fragment, there was not quite enough information available from this to allow the program SHELX to improve significantly on it. Subsequently, the fragment was recycled thrice more through DIRDIF before a fragment large enough for use in SHELX was obtained. The diagnostics produced for the first and the last two of these runs are tabulated below.

Test	S ₄	S ₄ O ₄ N ₄ C ₁₃	S ₄ O ₆ N ₈ C ₃₂
R	0.74	0.44	0.38
BH	7.0	4.4	5.6
BL	8.4	4.9	5.7
σC(EH)*	4.0*	4.2	4.9
(EL)	2.7*	2.5	0.8
σA(EH)	8.4	1.6*	2.07
(EL)	4.0	0.8*	0.9

CONSISTENCY

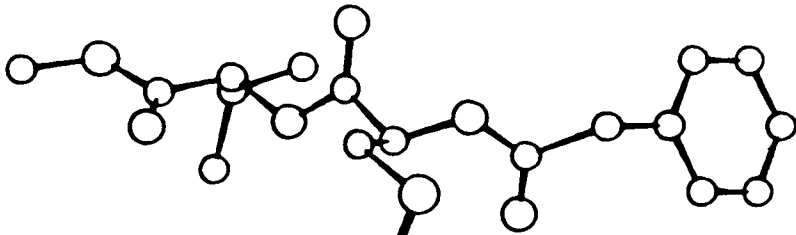
CYCLE 1	0.15(352)	0.38(87)	0.40(91)
2	0.21(738)	0.31(139)	0.32(144)
3	0.13(734)	0.31(138)	0.31(152)

Refinement of the structure progressed slowly, due to the disorder in one molecule (a phenylacetyl group occupying two possible sites) and the high temperature factors of the atoms in each molecule furthest from the region held in position by the sulphur bridge and hydrogen bonding systems. The final model has anisotropic thermal parameters for sulphur, nitrogen and the amide oxygen atoms, all carbon atoms and the ester oxygens having isotropic temperature factors. All hydrogen atoms are in calculated positions with temperature factors constrained to be equal. The phenyl rings are rigid groups, idealised to fit the positions of the rings found in the difference maps. The occupancy ratio for the sites of the disordered ring is 0.54 : 0.46 when the temperature factors are constrained to be equal for both rings. The methyl ester group on one of the molecules (molecule 4) is very poorly defined without having another possible location visible in the difference map.

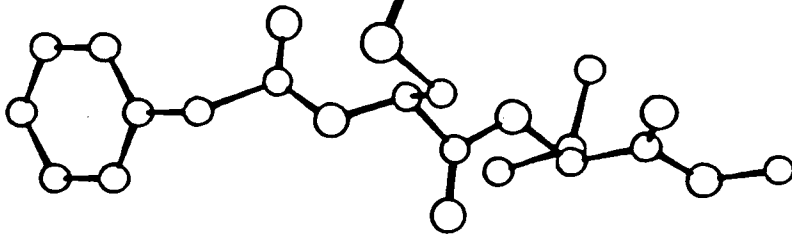
* For an explanation of the notation used in this table, please see p36

Description of Structure

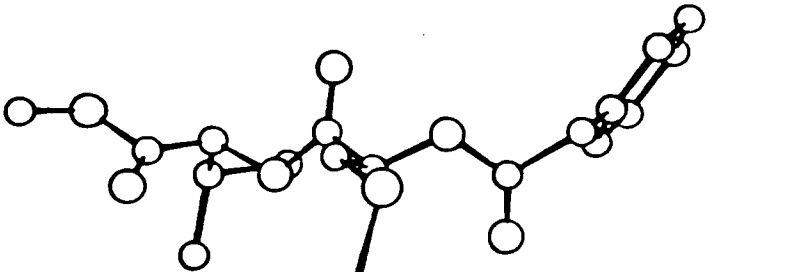
The unit cell contains 16 molecules, distributed over eight asymmetric units; with one full molecule in a general position, and two half molecules lying across crystallographic dyads in each asymmetric unit. These molecules are illustrated below.



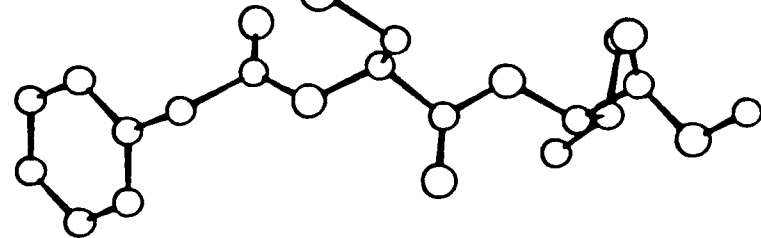
Molecule 1



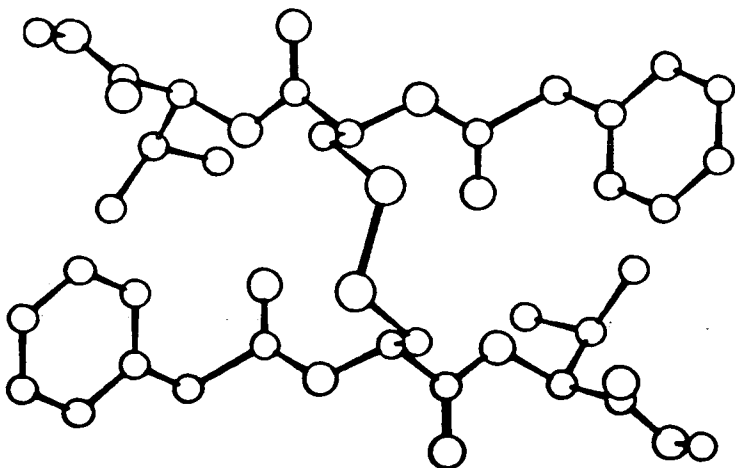
Molecule 2



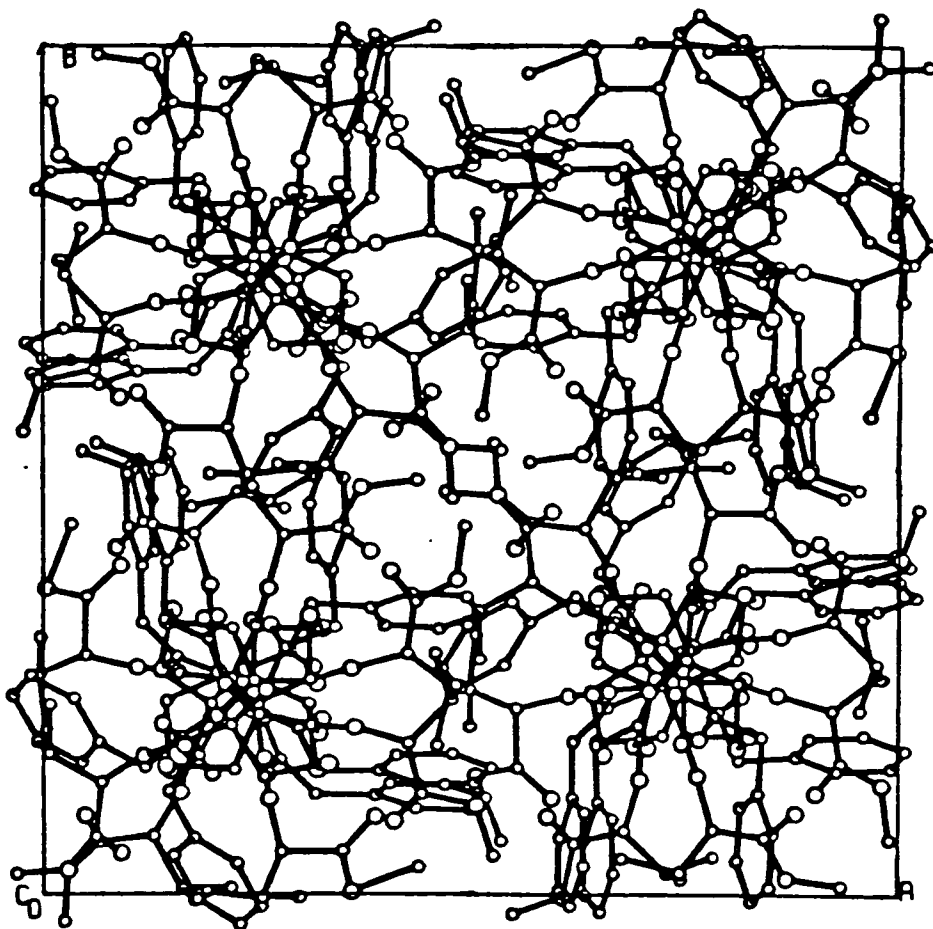
Molecule 3



Molecule 4



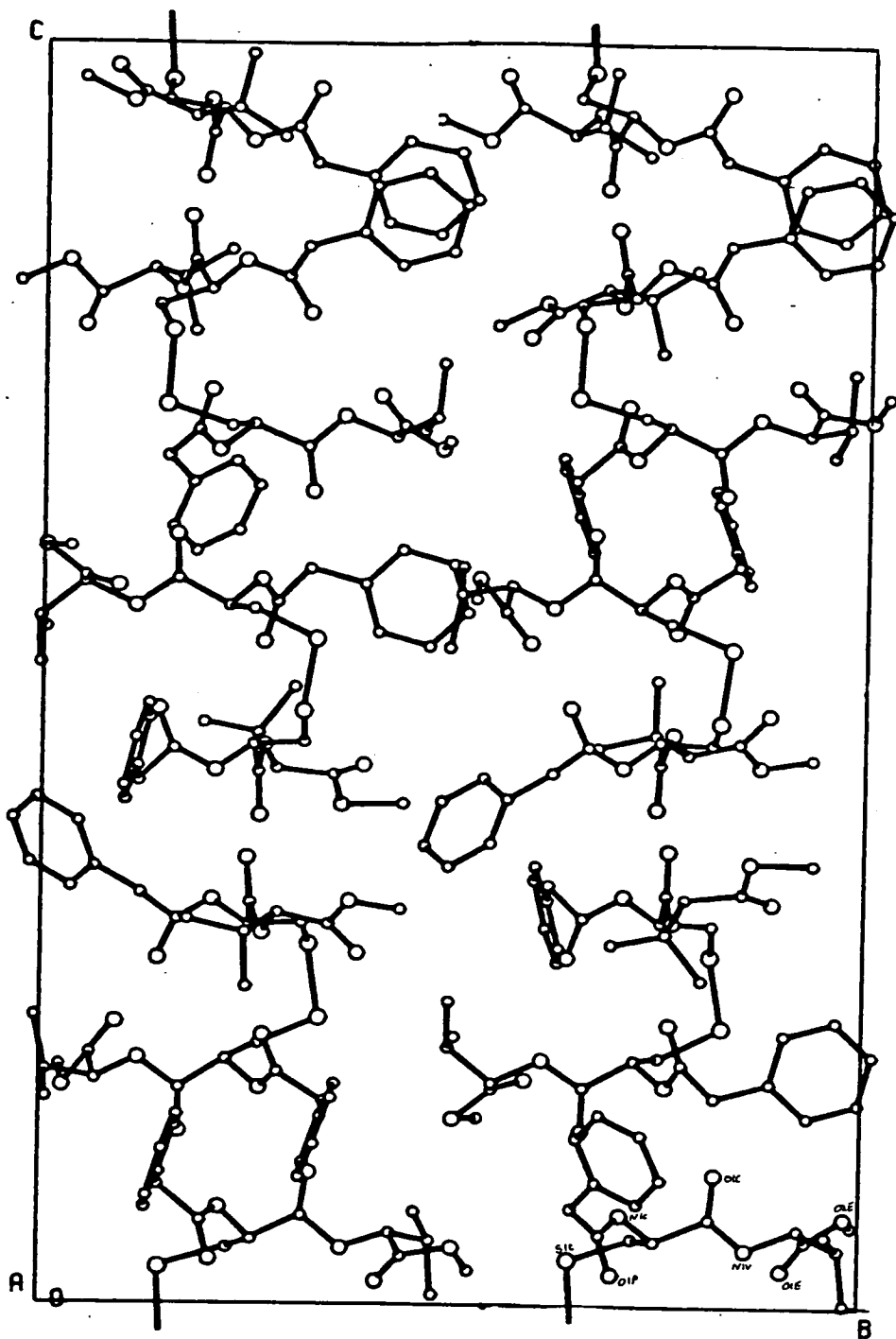
The molecule in the general position is disordered about one of its two phenylacetyl groups. Hydrogen bonding is both inter- and intramolecular, and occurs parallel to the z-axis in a manner analogous to the β -pleated sheet, which is the most common hydrogen bonding pattern for peptides. As a consequence of the twofold rotation axes at $z = 0$ and $z = 1/2$, the hydrogen bonding continues as an infinite ribbon as can be seen in the projection down half the x-axis. This helical arrangement occurs four times in the cell due to the 4_3 screw axis. Each column of molecules is well defined at the centre, giving way to a looser arrangement of the molecules at the periphery of each quadrant. This shows up strikingly in the projection down the z-axis, which has a surprising amount of clear space for a projection down 35 angstroms.



PROJECTION DOWN THE Z AXIS

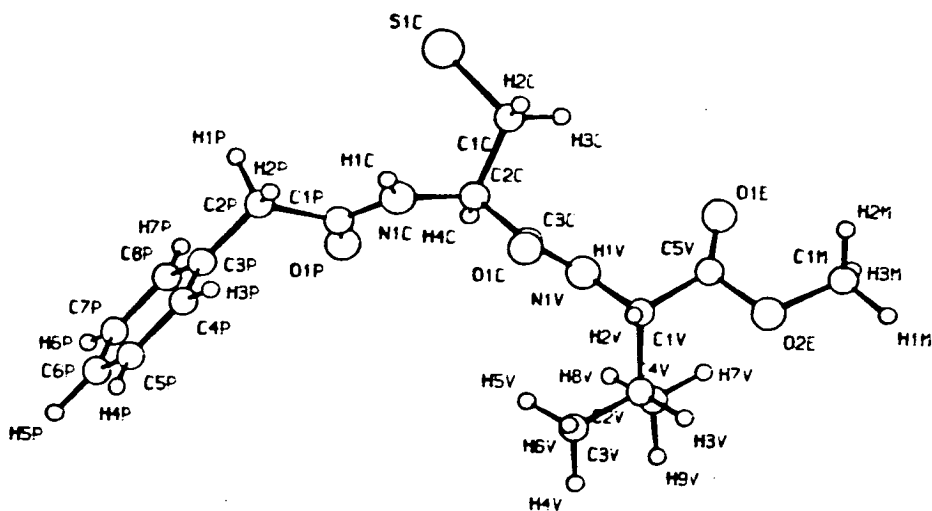
BIS-(N-PHENYLACETYL) CYSTINYL BIS-(VALINE METHYL ESTER)

BIS-(N-PHENYLACETYL) CYSTINYL BIS-(VALINE METHYL ESTER)



HALF PROJECTION DOWN THE X AXIS

Almost all the bond lengths are within their usual ranges, except that the valine C α to C' distance is significantly different between molecule 2 (1.57(3) Angstrom) and molecule 3 (1.36(3) Angstrom). All bond angles are well within the usual ranges except for the ether oxygen angle in molecule 4, which is 7.6 e.s.d.'s from the mean (85.6(27)° versus 106°). The diagram below shows the numbering scheme used in these tables.



The amide groups are all somewhat twisted out of planarity, as can be seen from the deviations from the mean plane and torsion angles below (the extra value in the column for molecule 3 is due to the alternate position of the phenylacetyl group).

Amide bond	M1	M2	M3	M4
Phen Cys	0.06	0.09	0.05 0.08	0.03
Cys Val	0.04	0.06	0.08	0.04

RMS deviation from mean plane of C-CO-N-C
(in Angstrom units)

Torsions	M1	M2	M3	M4
C2C-N1C-C1P-C2P	-169.5(14)	-162.8(14)	-177.5(16)	-178.5(16)
			-160.6(19)	
C2C-N1C-C1P-O1P	1.7(27)	14.8(28)	12 (4)	10 (3)
			22 (5)	
C1V-N1V-C3C-C2C	173.7(14)	169.2(15)	164.9(16)	172.9(17)
C1V-N1V-C3C-O1C	-3.6(25)	-6.8(28)	-5 (3)	-2 (3)

This deviation is probably due to the hydrogen bonding pattern, and allows for the rotation between each molecule and its neighbours on the z-axis. There are a number of conformational differences between the phenylacetyl groups on the different molecules. Half molecules 1 and 2 have similar torsions here, but the major occupancy ring on half molecule 3 points away from the main body of the molecule by about 45° compared with the first two, and it is also twisted a similar amount about the methene to carbonyl bond. The minor occupancy ring is closer in orientation to those of half molecules 1 and 2, but it still differs by more than 30° in torsion about the methene to carbonyl bond. In half molecule 4, the conformation of the phenylacetyl group is totally different, with the torsions being of the opposite sense to those in the rest of the structure, as can be seen from examination of the molecular plots. At the other end of the chain, the valine residue adopts a similar conformation in all molecules, except for molecule 4, where it is twisted *circa* 30° about the N-C α bond compared to the rest of the molecules. The only important difference in conformation of the ester group occurs in molecule 2, where the torsion angle C α -C'-O-C is *cis* rather than *trans*. The most significant contacts are those involved in hydrogen bonding either within or between molecules, although there are also a number of oxygen to carbon contact distances less than 3.5 Angstroms. There is one nitrogen to nitrogen and one carbon carbon contact less than 3.5 Angstroms, both between molecules 1 and 2. There are 10 contacts involving sulphur atoms less than 4.0 Angstroms and 11 sulphur to hydrogen contacts less than 3.5 Angstroms. There are numerous contacts between hydrogen and carbon, nitrogen or oxygen atoms, of which the most significant are those involved in hydrogen bonding or

80

the intramolecular close contacts between the cystine C α protons.

LABEL	M1	M2	M3	M4	MX
S1C-S1C	2.019(7)	2.045(7)		2.019(9)	
S1C-C1C	1.853(17)	1.767(17)	1.795(17)	1.834(21)	
O1C-C3C	1.217(21)	1.210(23)	1.226(24)	1.21(3)	
O1P-C1P	1.214(21)	1.176(22)	1.21(4)	1.22(3)	1.21(4)
O1E-C5V	1.166(25)	1.15(3)	1.26(3)	1.21(4)	
O2E-C5V	1.37(3)	1.42(3)	1.34(3)	1.30(4)	
O2E-C1M	1.47(3)	1.37(4)	1.36(3)	1.42(5)	
N1C-C2C	1.427(20)	1.445(20)	1.489(20)	1.455(23)	
N1C-C1P	1.304(22)	1.301(23)	1.32(4)	1.26(3)	1.32(4)
N1V-C3C	1.368(22)	1.361(23)	1.32(3)	1.32(3)	
N1V-C1V	1.463(22)	1.432(24)	1.53(3)	1.51(3)	
C1C-C2C	1.557(23)	1.569(22)	1.501(23)	1.48(3)	
C2C-C3C	1.486(23)	1.531(25)	1.57(3)	1.50(3)	
C1V-C2V	1.52(3)	1.48(3)	1.47(3)	1.57(4)	
C1V-C5V	1.48(3)	1.57(3)	1.36(3)	1.47(4)	
C2V-C3V	1.46(3)	1.46(3)	1.59(3)	1.47(4)	
C2V-C4V	1.52(3)	1.51(3)	1.51(3)	1.52(4)	
C1P-C2P	1.56(3)	1.54(3)	1.53(4)	1.59(3)	1.53(4)
C2P-C3P	1.492(18)	1.532(18)	1.520(14)	1.476(22)	1.526(15)

The table above shows all the unconstrained bond distances in the structure. The fifth column shows the bond distances in the minor site of the phenylacetyl group of molecule 3.

LABEL	M1	M2	M3	M4	MX
S1C-S1C-C1C	103.8(6)	105.0(6)	105.5(6)	108.5(7)	
C5V-O2E-C1M	113.3(16)	108.5(22)	116.6(21)	85.6(27)	
C2C-N1C-C1P	122.5(14)	123.0(14)	118.6(19)	122.2(16)	117.5(20)
C3C-N1V-C1V	120.2(14)	120.1(15)	118.8(16)	121.2(17)	
S1C-C1C-C2C	111.8(11)	112.6(11)	116.2(12)	113.8(14)	
N1C-C2C-C1C	111.3(13)	111.7(13)	108.5(13)	112.0(15)	
N1C-C2C-C3C	112.4(13)	112.1(13)	108.1(13)	108.5(15)	
C1C-C2C-C3C	104.2(13)	106.6(13)	108.9(14)	109.7(16)	
O1C-C3C-N1V	121.1(16)	124.7(17)	125.5(18)	121.6(19)	
O1C-C3C-C2C	124.9(16)	121.4(16)	120.4(17)	121.5(19)	
N1V-C3C-C2C	113.9(14)	113.9(15)	113.2(16)	116.7(17)	
N1V-C1V-C2V	110.8(15)	113.0(17)	111.3(17)	108.4(19)	
N1V-C1V-C5V	106.8(15)	106.9(17)	105.6(19)	105.1(21)	
C2V-C1V-C5V	116.0(16)	107.9(18)	111.2(21)	113.9(24)	
C1V-C2V-C3V	111.0(18)	109.3(19)	107.7(19)	108.0(24)	
C1V-C2V-C4V	111.9(17)	115.9(20)	115.7(19)	107.4(24)	
C3V-C2V-C4V	111.9(18)	107.5(20)	105.3(19)	120.6(27)	
O1E-C5V-O2E	122.1(19)	124.2(24)	112.3(23)	123 (3)	
O1E-C5V-C1V	130.0(19)	126.5(23)	135.7(25)	129 (3)	
O2F-C5V-C1V	107.8(16)	105.1(19)	112.0(22)	107.6(28)	
O1P-C1P-N1C	125.4(16)	125.9(17)	125 (3)	126.4(21)	125 (3)
O1P-C1P-C2P	121.3(16)	124.3(17)	122.4(27)	115.8(19)	123 (3)
N1C-C1P-C2P	112.7(15)	109.7(15)	112.1(25)	117.3(19)	112.1(28)
C1P-C2P-C3P	117.1(13)	115.3(13)	110.2(16)	112.7(16)	111.9(17)
C2P-C3P-C4P	120.8(7)	118.5(7)	112.6(5)	119.3(9)	124.4(6)
C2P-C3P-C8P	119.2(7)	121.5(7)	127.4(5)	120.6(9)	115.6(6)

This table gives a comparison of all the bond angles in the structure. It can easily be seen that the methyl ester on molecule 4 is very badly distorted. The table of comparative torsion angles is shown overleaf.

C1C-S1C-S1C-C1C	97.4 (8)		99.0 (8)		99.6 (10)	
S1C-S1C-C1C-C2C	-96.5 (11)	-100.0 (11)	-84.9 (12)	-89.2 (14)		
C1M-O2E-C5V-O1E	3 (3)	-22 (4)	4 (3)	1 (4)		
C1M-O2E-C5V-C1V	179.7 (16)	179.8 (21)	-174.4 (22)	-174 (3)		
C1P-N1C-C2C-C1C	123.2 (16)	109.4 (17)	122.4 (21)	117.5 (20)	118.4 (23)	
C1P-N1C-C2C-C3C	-120.3 (17)	-131.1 (17)	-119.7 (22)	-121.2 (20)	-123.6 (23)	
C2C-N1C-C1P-O1P	1.7 (27)	15 (3)	12 (4)	10 (3)	22 (5)	
C2C-N1C-C1P-C2P	-169.5 (14)	-162.8 (14)	-177.5 (16)	-178.5 (16)	-160.6 (19)	
C1V-N1V-C3C-O1C	-3.6 (25)	-7 (3)	-5 (3)	-2 (3)		
C1V-N1V-C3C-C2C	173.7 (14)	169.2 (15)	164.9 (16)	172.9 (17)		
C3C-N1V-C1V-C2V	118.4 (17)	121.8 (19)	123.9 (20)	155.1 (20)		
C3C-N1V-C1V-C5V	-114.4 (17)	-119.7 (19)	-115.2 (21)	-82.7 (25)		
S1C-C1C-C2C-N1C	-49.3 (15)	-48.6 (16)	-57.6 (16)	-59.5 (18)		
S1C-C1C-C2C-C3C	-170.7 (11)	-171.3 (11)	-175.1 (12)	180.0 (13)		
N1C-C2C-C3C-O1C	-36.3 (23)	-29.0 (23)	-35.7 (22)	-43.3 (25)		
N1C-C2C-C3C-N1V	146.5 (14)	154.8 (14)	154.4 (15)	141.9 (17)		
C1C-C2C-C3C-O1C	84.3 (20)	93.4 (19)	82.1 (21)	79.4 (24)		
C1C-C2C-C3C-N1V	-92.8 (16)	-82.7 (17)	-87.9 (18)	-95.4 (21)		
N1V-C1V-C2V-C3V	-63.9 (21)	-66.6 (23)	-59.2 (22)	-59.2 (27)		
N1V-C1V-C2V-C4V	61.9 (21)	54.9 (25)	58.2 (25)	72.3 (26)		
C5V-C1V-C2V-C3V	174.2 (17)	175.4 (19)	-176.8 (20)	-175.8 (26)		
C5V-C1V-C2V-C4V	-60.1 (23)	-63.0 (25)	-59.4 (27)	-44 (3)		
N1V-C1V-C5V-O1E	6 (3)	-2 (3)	-11 (4)	-6 (5)		
N1V-C1V-C5V-O2E	-171.3 (14)	155.9 (17)	167.2 (19)	169.0 (24)		
C2V-C1V-C5V-O1E	129.7 (23)	120 (3)	110 (3)	113 (4)		
C2V-C1V-C5V-O2E	-47.2 (22)	-82.3 (22)	-71.9 (27)	-72 (3)		
O1P-C1P-C2P-C3P	50.0 (22)	48.6 (24)	72 (3)	-80.2 (23)	84 (4)	
N1C-C1P-C2P-C3P	-138.4 (14)	-133.8 (14)	-98.5 (24)	107.2 (21)	-93.7 (27)	
C1P-C2P-C3P-C4P	107.6 (13)	100.0 (13)	154.4 (15)	-153.7 (12)	100.8 (18)	
C1P-C2P-C3P-C8P	-72.5 (14)	-78.4 (14)	-24.6 (18)	29.8 (19)	-78.9 (18)	
C2P-C3P-C4P-C5P	-180.0 (8)	-178.4 (8)	-179.1 (6)	-176.5 (10)	-179.7 (7)	
C2P-C3P-C8P-C7P	180.0 (8)	178.4 (8)	179.0 (7)	176.5 (10)	179.7 (6)	

Intramolecular Contacts

01P1	2.893	(18)	N1V1	(Y,X,1 -Z)
01P2	2.961	(20)	N1V3	(X,Y,Z)
01P3	2.865	(18)	N1V2	(X,Y,Z)
01P4	2.893	(21)	N1V4	(Y,X,1 -Z)
01P1	3.405	(20)	C2C1	(Y,X,1 -Z)
01E1	3.465	(14)	C8P1	(Y,X,1 -Z)
01P2	3.337	(20)	C2C3	(X,Y,Z)
01P3	3.256	(20)	C2C2	(X,Y,Z)
01P4	3.364	(23)	C2C4	(Y,X,1 -Z)
01P4	3.46	(4)	C4V4	(Y,X,1 -Z)

Intermolecular Contacts

S1C1	3.795	(13)	01C2	(1/2+X,1 1/2-Y,1 1/4-Z)
S1C1	3.852	(19)	02E3	(Y,1 +X,1 -Z)
S1C3	3.488	(16)	02E1	(Y,X,1 -Z)
S1C3	3.781	(15)	01C4	(1/2-X,1/2+Y,3/4-Z)
S1C1	3.927	(5)	C7P2	(1 +X,Y,Z)
S1C2	3.60	(3)	C4V4	(Y,1 +X,1 -Z)
S1C3	3.919	(23)	C2V1	(Y,X,1 -Z)
S1C3	3.989	(20)	C4V1	(Y,X,1 -Z)
S1C3	3.896	(25)	C1M1	(Y,X,1 -Z)
S1C4	3.542	(27)	C4V2	(Y,X,1 -Z)
01C1	2.897	(17)	N1C2	(1/2+X,1 1/2-Y,1 1/4-Z)
01C2	2.817	(17)	N1C1	(-1/2+X,1 1/2-Y,1 1/4-Z)
01C3	2.874	(20)	N1C4	(1/2-X,1/2+Y,3/4-Z)
01C4	2.765	(19)	N1C3	(1/2-X,-1/2+Y,3/4-Z)
01C1	3.344	(21)	C2P2	(1/2+X,1 1/2-Y,1 1/4-Z)
01C2	3.496	(20)	C1C1	(-1/2+X,1 1/2-Y,1 1/4-Z)
01C4	3.455	(22)	C1C3	(1/2-X,-1/2+Y,3/4-Z)
N1C1	3.477	(18)	N1C2	(1/2+X,1 1/2-Y,1 1/4-Z)
C4P1	3.463		C4P2	(1/2+X,1 1/2-Y,1 1/4-Z)

Intramolecular Hydrogen Contacts

O1P1	2.662	(20)	H4C1	(Y,X,1 -Z)
O1P2	2.486	(20)	H4C3	(X,Y,Z)
O1P3	1.812	(18)	H1V2	(X,Y,Z)
O1P3	2.71	(3)	H8V2	(X,Y,Z)
O1P4	1.834	(21)	H1V4	(Y,X,1 -Z)
O1P4	2.43	(4)	H9V4	(Y,X,1 -Z)
O1P1	1.849	(18)	H1V1	(Y,X,1 -Z)
O1P1	2.904	(23)	H8V1	(Y,X,1 -Z)
O1E1	2.399	(14)	H7P1	(Y,X,1 -Z)
O1P2	1.904	(20)	H1V3	(X,Y,Z)
O1P2	2.843	(26)	H8V3	(X,Y,Z)
O1E2	2.999	(23)	H2PX	(X,Y,Z)
O1P3	2.434	(20)	H4C2	(X,Y,Z)
O1E3	2.644	(17)	H7P2	(X,Y,Z)
O1P4	2.514	(23)	H4C4	(Y,X,1 -Z)
C2C1	2.920	(22)	H4C1	(Y,X,1 -Z)
C2C2	2.875	(23)	H4C3	(X,Y,Z)
C7P2	2.993	(23)	H7V3	(X,Y,Z)
C4V2	2.709	(26)	H7PX	(X,Y,Z)
C2C3	2.941	(23)	H4C2	(X,Y,Z)
C1P3	2.87	(4)	H1V2	(X,Y,Z)
C1PX	2.85	(4)	H1V2	(X,Y,Z)
C2C4	2.957	(24)	H4C4	(Y,X,1 -Z)
H4C1	1.913	(22)	H4C1	(Y,X,1 -Z)
H4C2	1.914	(23)	H4C3	(X,Y,Z)
H7V2	2.076	(26)	H7PX	(X,Y,Z)
H8V2	2.447	(26)	H7P3	(X,Y,Z)
H4C4	1.932	(24)	H4C4	(Y,X,1 -Z)

Intermolecular Hydrogen Contacts

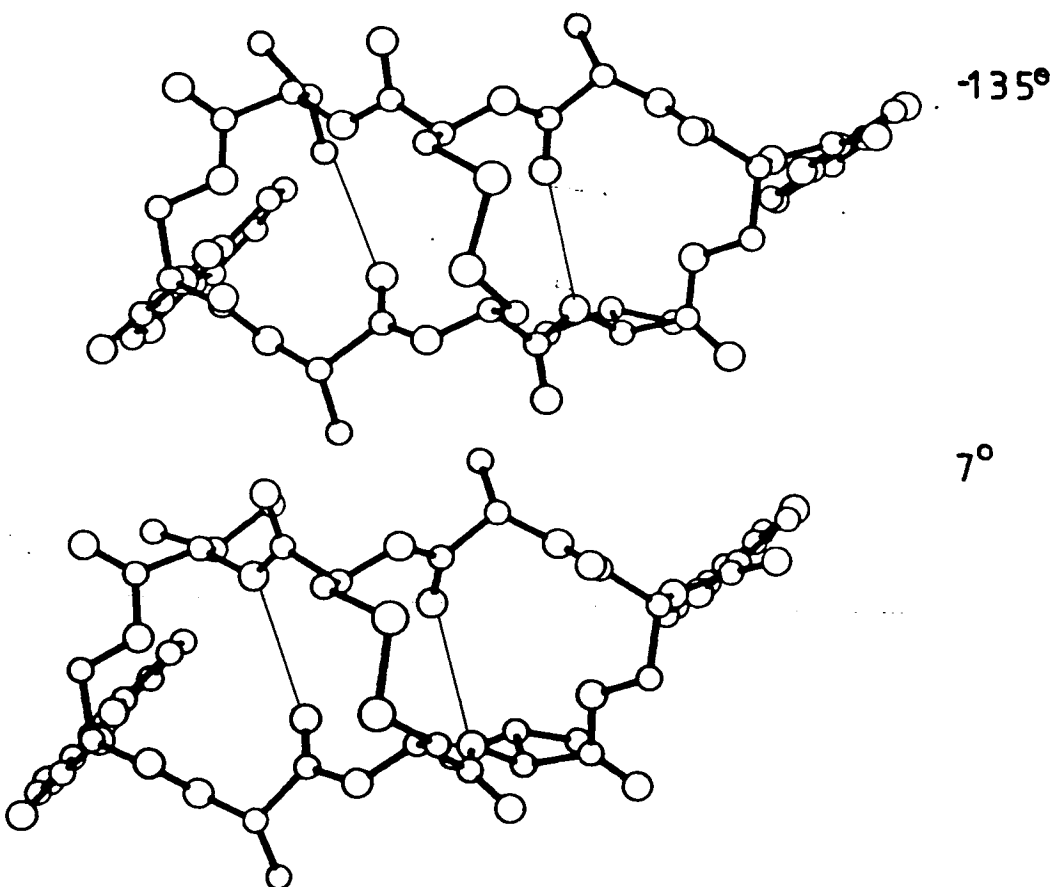
S1C1	2.923	(5)	H6P2	(1 +X,Y,Z)
S1C1	3.482	(23)	H3V3	(Y,1 +X,1 -Z)
S1C2	3.099	(5)	H6P1	(X,Y,Z)
S1C2	3.31	(3)	H3V4	(Y,1 +X,1 -Z)
S1C2	3.16	(3)	H7V4	(Y,1 +X,1 -Z)
S1C2	3.25	(3)	H8V4	(Y,1 +X,1 -Z)
S1C3	2.982	(23)	H3V1	(Y,X,1 -Z)
S1C3	3.418	(25)	H1M1	(Y,X,1 -Z)
S1C	3.393	(3)	H7V4	(Y,1 +X,1 -Z)
S1C4	2.612	(27)	H9V2	(Y,X,1 -Z)
S1C4	3.139	(6)	H6PX	(X,Y,Z)
O1C1	2.700	(11)	H5P1	(1/2+X,1 1/2-Y,1 1/4-Z)
O1C1	1.878	(17)	H1C2	(1/2+X,1 1/2-Y,1 1/4-Z)
O1C1	2.381	(21)	H2P2	(1/2+X,1 1/2-Y,1 1/4-Z)
O1P1	2.967	(12)	H3P2	(1/2+X,1 1/2-Y,1 1/4-Z)
O1C2	1.805	(17)	H1C1	(-1/2+X,1 1/2-Y,1 1/4-Z)
O1C2	2.808	(21)	H2P1	(-1/2+X,1 1/2-Y,1 1/4-Z)
O1P2	2.940	(12)	H3P1	(-1/2+X,1 1/2-Y,1 1/4-Z)
O1E2	2.71	(4)	H3M2	(Y,X,1 -Z)
O1C3	1.871	(20)	H1C4	(1/2-X,1/2+Y,3/4-Z)
O1E3	2.916	(17)	H4P3	(1/2-X,1/2+Y,3/4-Z)
O1C4	2.344	(19)	H1C3	(1/2-X,-1/2+Y,3/4-Z)
O1C4	2.732	(20)	H2P3	(1/2-X,-1/2+Y,3/4-Z)
O1C2	2.904	(20)	H2C1	(-1/2+X,1 1/2-Y,1 1/4-Z)
O1C2	2.563	(11)	H5P2	(1/2+X,1 1/2-Y,1 1/4-Z)
O1C3	2.824	(26)	H1P4	(1/2-X,1/2+Y,3/4-Z)
O2E3	2.93	(3)	H6V2	(1/2-Y,1/2+X,-1/4+Z)
O1C4	2.786	(22)	H2C3	(1/2-X,-1/2+Y,3/4-Z)
O1C4	2.663	(21)	H1PX	(1/2-X,-1/2+Y,3/4-Z)
O2E4	2.68	(3)	H2V1	(1/2-Y,-1/2+X,-1/4+Z)

Intermolecular Hydrogen Contacts - continued

N1C1	2.686	(18)	H1C2	(1/2+X,1 1/2-Y,1 1/4-Z)
N1C2	2.691	(18)	H1C1	(-1/2+X,1 1/2-Y,1 1/4-Z)
N1C3	2.805	(19)	H1C4	(1/2-X,1/2+Y,3/4-Z)
C3C1	2.731	(22)	H1C2	(1/2+X,1 1/2-Y,1 1/4-Z)
C3C1	2.970	(25)	H2P2	(1/2+X,1 1/2-Y,1 1/4-Z)
C1P1	2.842	(18)	H3P2	(1/2+X,1 1/2-Y,1 1/4-Z)
C4P1	2.845	(0)	H4P4	(1/2+Y,1 1/2-X,1/4+Z)
C3P1	2.788	(0)	H4P4	(1/2+Y,1 1/2-X,1/4+Z)
C3C2	2.693	(23)	H1C1	(-1/2+X,1 1/2-Y,1 1/4-Z)
C3C2	2.988	(19)	H5P2	(1/2+X,1 1/2-Y,1 1/4-Z)
C7P2	2.886	(28)	H1M3	(-1 +Y,1 +X,1 -Z)
C1C3	2.931	(16)	H5P3	(1/2-X,1/2+Y,3/4-Z)
C3C3	2.863	(24)	H1C4	(1/2-X,1/2+Y,3/4-Z)
C4V3	2.906	(23)	H4P1	(-1/2+X,1 1/2-Y,1 1/4-Z)
C6P3	2.952	(22)	H2P4	(1/2-X,1/2+Y,3/4-Z)
C5PX	2.990	(22)	H2P4	(1/2-X,1/2+Y,3/4-Z)
C3C4	2.957	(21)	H3PX	(1/2-X,-1/2+Y,3/4-Z)
C4P4	2.701	(25)	H4V3	(1/2-X,-1/2+Y,3/4-Z)
C5P4	2.856	(24)	H2M1	(Y,-1 +X,1 -Z)
C6P4	2.628	(24)	H2M1	(Y,-1 +X,1 -Z)
C3V4	2.91	(4)	H1P2	(-1 +Y,X,1 -Z)
H1C1	2.060	(18)	H1C2	(1/2+X,1 1/2-Y,1 1/4-Z)
H3C1	2.412	(29)	H9V3	(1 +X,Y,Z)
H4P1	2.429	(23)	H9V	(1/2+X,1 1/2-Y,1 1/4-Z)
H6V1	2.433	(3)	H2V4	(1/2+Y,1/2-X,1/4+Z)
H9V1	2.293	(25)	H3C2	(Y,X,1 -Z)
H1P2	2.14	(4)	H3V4	(Y,1 +X,1 -Z)
H6P2	2.415	(28)	H1M3	(-1 +Y,1 +X,1 -Z)
H6V2	2.37	(3)	H2V3	(-1/2+Y,1/2-X,1/4+Z)
H1C3	2.237	(19)	H1C4	(1/2-X,1/2+Y,3/4-Z)
H2C3	2.210	(16)	H5P3	(1/2-X,1/2+Y,3/4-Z)
H4V3	2.225	(25)	H3P4	(1/2-X,1/2+Y,3/4-Z)
H2M3	2.460	(28)	H4P3	(1/2-X,1/2+Y,3/4-Z)
H6PX	2.469	(20)	H2C4	(X,Y,Z)

Comparison with the cyclic octadepsipeptide TANDEM

A particularly interesting feature of this structure is the sulphur bridge connecting two peptide chains hydrogen bonded in a manner analogous to the β -sheet. Although there are more than a thousand crystal structures of peptides on file with the Cambridge Crystallographic Database, only about twenty of these also contain sulphur bridges. Most of these compounds are simple cystine derivatives with no intramolecular hydrogen bonding, such as the compounds discussed in the previous chapter. After an exhaustive search of the database, in both the interactive and the batch modes, one other compound was found to contain a similar hydrogen bonded sulphur bridge. This compound, known as TANDEM (an abbreviation for des-N-tetramethyl triostin A) (XXXIV), is a cyclic octadepsipeptide antibiotic with a strong binding affinity for the alternating sequence of adenine and thymine. It is suggested (Viswamitra *et al.*, 1981) that the selectivity shown in binding of a factor of 7,500 in favour of poly(dA-dT) against poly(dG-dC) is explainable in terms of the backbone conformation, which is controlled by the intramolecular hydrogen bonds. The atomic coordinates from the two independent crystal structures at -135° and 7° C (Hossain *et al.*, 1981) were used to produce the diagrams below, in which each molecule is projected onto the best plane of its peptide backbone.



78

The tables below give a numerical comparison of the central portion of the cystinyl valine dimer with the equivalent portion of the antibiotic. Standard deviations quoted for each figure are based on the unweighted mean of the relevant parameters and do not reflect the e.s.d's involved as the e.s.d's for the antibiotic were unavailable. In the case of the torsion angles, the standard deviation calculated is not quoted if it is greater than 10° , as this is taken to indicate a genuine difference between the structures and not merely a minor difference in conformation.

As with the cyclic disulphide discussed in the previous chapter, the torsion angles at the sulphur bridge belong to neither to class I nor to class II, but to the group of those compounds that are geometrically constrained to have the two $C\alpha$'s closer to one another than is consistent with a type I or II structure. This would seem to indicate that the gain in conformational stability from intramolecular hydrogen bonding is more than enough to offset the slight steric crowding of the disulphide bridge. With the exception of the sulphur bridge, the side chains have different conformations entirely in the two compounds. However, the conformation at the amide bond is approximately the same in each compound.

bond	M1	M2	M3	M4	T1	T2	T3	T4	mean	es.d
B1C-H1C	2.02	2.05	2.05	2.02	2.04	2.04	2.01	2.01	2.03	0.02
B1C-C1C	1.85	1.74	1.74	1.83	1.82	1.80	1.81	1.82	1.82	0.02
B1C'-C1C'	1.85	1.77	1.77	1.83	1.80	1.82	1.82	1.81	1.81	0.03
U1C-C3C	1.22	1.21	1.23	1.21	1.24	1.23	1.26	1.20	1.22	0.02
O1F-C3V	1.17	1.15	1.26	1.21	1.24	1.19	1.15	1.21	1.20	0.04
O2E-C1M	1.47	1.37	1.36	1.42	1.45	1.44	1.41	1.38	1.41	0.04
O2E-C3V	1.37	1.42	1.34	1.30	1.35	1.37	1.29	1.30	1.34	0.04
O1P-C1P	1.21	1.18	1.21	1.22	1.25	1.20	1.26	1.25	1.22	0.03
N1C-C2C	1.43	1.45	1.49	1.45	1.44	1.46	1.43	1.43	1.45	0.02
N1C-C1P	1.30	1.30	1.32	1.26	1.34	1.36	1.32	1.26	1.31	0.03
N1V-C3C	1.37	1.34	1.32	1.32	1.33	1.36	1.35	1.34	1.34	0.02
N1V-C1V	1.46	1.47	1.53	1.51	1.47	1.44	1.46	1.40	1.46	0.04
C1C-C2C	1.56	1.57	1.50	1.48	1.52	1.50	1.49	1.52	1.52	0.03
C2C-C3C	1.49	1.53	1.57	1.50	1.52	1.52	1.39	1.50	1.50	0.05
C1P-C2P	1.56	1.54	1.53	1.59	1.49	1.56	1.47	1.61	1.54	0.04
C1V-C2V	1.52	1.48	1.47	1.57	1.54	1.55	1.64	1.52	1.54	0.05
C1V-C3V	1.48	1.57	1.36	1.47	1.46	1.51	1.54	1.51	1.49	0.06
C2V-C3V	1.46	1.44	1.59	1.47	1.54	1.56	1.42	1.52	1.50	0.05
C2V-C4V	1.52	1.51	1.51	1.52	1.52	1.57	1.59	1.61	1.54	0.04

angle	M1	M2	M3	M4	T1	T2	T3	T4	mean	es.d
B1C'-B1C-C1C	103.84	104.98	104.98	108.46	105.85	105.90	105.65	106.83	105.81	1.29
B1C-H1C'-C1C'	103.84	105.46	105.46	108.46	105.90	105.85	106.83	105.65	105.93	1.23
C1M-O2E-C3V	113.30	108.55	116.57	85.61	117.41	117.17	116.63	116.20		
C2C-N1C-C1P	122.54	123.04	118.62	122.17	123.84	122.53	123.50	122.07	122.89	1.51
C3C-N1V-C1V	120.23	120.14	118.76	121.21	123.79	119.32	127.39	125.80	122.08	2.99
B1C-C1C-C2C	111.81	112.57	116.23	113.84	116.31	114.92	115.21	115.70	114.57	1.58
N1C-C2C-C1C	111.27	111.65	108.53	112.04	110.81	110.00	112.85	108.74	110.74	1.45
N1C-C2C-C3C	112.44	112.13	108.12	108.48	107.97	107.60	109.05	109.26	109.38	1.75
C1C-C2C-C3C	104.24	106.56	108.91	109.73	109.71	111.55	108.50	109.62	108.60	2.11
U1C-C3C-N1V	121.13	124.66	125.47	121.62	121.23	123.16	111.29	119.10	120.96	4.13
O1C-C3C-C2C	124.91	121.36	120.43	121.46	121.31	121.05	124.31	122.95	122.22	1.54
N1V-C3C-C2C	113.90	113.85	113.24	116.70	117.43	115.65	124.39	117.49	116.98	3.34
O1P-C1P-N1C	123.37	123.90	124.75	126.44	121.63	124.43	122.20	129.08	124.98	2.22
O1P-C1P-C2P	121.30	124.32	122.36	115.80	122.33	123.02	120.89	115.24	120.66	3.12
N1C-C1P-C2P	112.74	109.74	112.11	117.27	116.03	112.39	116.91	115.51	114.09	2.54
N1V-C1V-C2V	110.79	113.00	111.31	108.40	111.66	110.15	109.61	111.27	110.77	1.31
N1V-C1V-C3V	106.83	106.85	105.65	105.08	111.67	110.93	110.79	108.08	108.24	2.40
C2V-C1V-C3V	116.00	107.90	111.24	113.92	116.25	112.48	109.92	113.11	112.85	2.82
C1V-C2V-C3V	111.01	109.29	107.65	107.98	109.40	108.35	125.68	110.72	111.26	5.57
C1V-C2V-C4V	111.87	115.92	115.70	107.35	111.15	112.26	98.97	111.17	110.55	5.07
C3V-C2V-C4V	111.89	107.47	105.27	120.64	112.49	111.10	123.11	114.64	113.33	5.68
O1E-C3V-O2E	122.09	124.24	112.27	123.33	120.99	122.92	124.14	124.41	121.80	3.77
O1E-C3V-C1V	130.04	126.49	135.69	128.86	127.70	127.35	120.97	124.23	127.66	4.03
O2E-C3V-C1V	107.80	105.10	112.01	107.57	110.85	109.47	114.89	111.26	109.87	2.86

torsion	M1	M2	M3	M4	T1	T2	T3	T4	mean	esd
C1C-B1C-B1C'-C1C'	97.37	99.01	99.01	99.65	101.71	101.71	99.69	99.69	99.73	1.35
B1C'-S1C-C1C-C2C	-96.51	-100.02	-84.91	-89.21	-87.85	-84.23	-99.88	-79.05	-90.21	7.29
C1M-O2E-C5V-O1C	2.48	-21.99	4.06	0.87	-10.24	-10.30	-0.67	-3.85	-3.20	8.18
C1M-O2E-C5V-C1V	179.70	179.79	-174.35	-174.05	176.94	175.21	179.91	177.60	180.09	3.63
C1P-N1C-C2C-C1C	123.19	109.40	122.35	117.53	147.36	141.90	134.89	148.64		
C1P-N1C-C2C-C3C	-120.30	-131.09	-119.66	-121.19	-92.47	-96.39	-104.47	-91.74		
C2C-N1C-C1P-O1P	1.70	14.79	12.35	9.83	-4.52	-2.65	-9.92	-6.36	1.90	8.72
C2C-N1C-C1P-C2P	-169.51	-162.78	-177.54	-178.54	174.34	172.95	170.12	168.68	179.71	9.50
C1V-N1V-C3C-O1C	-3.61	-6.80	-4.48	-1.81	2.28	0.00	8.94	5.02	-0.06	4.91
C1V-N1V-C3C-C2C	173.67	169.18	164.89	172.90	-179.52	175.83	-172.22	177.50	175.28	6.54
C3C-N1V-C1V-C2V	118.42	121.80	123.90	129.13	126.17	179.40	106.96	126.49		
C3C-N1V-C1V-C5V	-114.37	-119.68	-115.23	-85.72	-101.78	-55.40	-131.56	-106.20		
B1C-C1C-C2C-N1C	-49.27	-48.58	-57.60	-59.50	-53.85	-60.98	-52.02	-63.96	-55.72	5.28
B1C-C1C-C2C-C3C	-170.69	-171.33	-175.10	179.96	-172.98	179.70	-172.98	176.65	184.16	4.45
N1C-C2C-C3C-O1C	-36.31	-29.04	-35.66	-43.33	-66.71	-94.32	-47.10	-53.83		
N1C-C2C-C3C-N1V	146.52	154.82	154.38	141.94	115.10	89.86	134.21	133.99		
C1C-C2C-C3C-O1C	84.34	93.40	82.09	79.37	54.15	26.41	76.18	65.24		
C1C-C2C-C3C-N1V	-92.83	-82.73	-87.87	-95.37	-124.05	-149.40	-102.51	-106.94		
N1V-C1V-C2V-C3V	-63.86	-66.63	-59.22	-59.18	-65.83	-67.07	50.78	-64.21		
N1V-C1V-C2V-C4V	61.91	54.94	58.17	72.33	169.34	169.86	-166.65	167.15		
C5V-C1V-C2V-C3V	174.16	175.45	-176.75	-175.77	164.47	168.62	-71.21	172.41		
C5V-C1V-C2V-C4V	-60.08	-62.96	-59.37	-44.25	39.64	45.55	71.36	43.77		
N1V-C1V-C5V-O1C	5.68	-1.64	-10.70	-5.54	141.40	141.40	137.82	127.21		
N1V-C1V-C5V-O2E	-171.25	155.93	167.20	169.01	-46.38	-44.41	-42.76	-56.23		
C2V-C1V-C5V-O1E	129.74	120.15	110.22	112.96	-88.90	-94.71	-100.88	-107.73		
C2V-C1V-C5V-O2E	-47.19	-82.27	-71.88	-72.49	83.31	79.47	78.54	68.83		

Solution Study on the Dimeric Disulphide

In the crystalline state the central portion of the molecule is held in a rigid conformation by the disulphide bridge and the intramolecular hydrogen bonds while the periphery of the molecule is not so constrained and thus has a far higher degree of thermal motion. In order to determine whether this was also the case in solution, it was decided to examine the conformation of the molecule by the method of nuclear magnetic resonance. The results of such a study could be compared with those obtained in the study on TANDEM (Williams *et al.*, 1982) and in the study of the cyclic disulphide discussed in the previous chapter. It was expected that the study would show the molecule to be symmetrical on the nmr timescale with a well resolved ABMX system (the cystine methylene, methine and amide protons) and a similarly well resolved A_3B_3 MRX system (the valine methyls, methines and amide protons). It was further expected that the coupling constants for these systems would indicate a rigid central portion of the molecule held together by hydrogen bonding and free rotation about the bonds in the valine and phenylacetyl group regions of the spectrum. Accordingly, the spectrum of the compound was measured at 360 MHz in deuteriochloroform solution (It was felt that using perdeutero dimethyl sulphoxide would have severely lessened the probability of intramolecular hydrogen bonding) and the results are tabulated below.

Shift/ppm(m)	Assignment		
0.823d	C γ 2 of valine	$^3J_{MA} = 6.9$ Hz	(A $_3$ B $_3$ MR system)
0.849d	C γ 1 of valine	$^3J_{MB} = 6.9$ Hz	" "
2.094dqq	C β of valine	$^3J_{RM} = 5.3$ Hz	" "
		$^3J_{MA} = 6.8$ Hz	
		$^3J_{MB} = 6.8$ Hz	
2.725) m	Cystine C β poorly resolved AB		
2.881)	of ABMX system		
3.585s	Phenylacetyl methylene		
3.677s	Methyl ester		
4.379dd	C α of valine	$^3J_{RX} = 8.3$ Hz	(A $_3$ B $_3$ MR system)
		$^3J_{MR} = 5.1$ Hz	
4.755ddd	C α of cystine	$^3J_{MX} = 7.7$ Hz	(ABMX system)
		$^3J_{MA} = 5.2$ Hz	
		$^3J_{MB} = 6.4$ Hz	
6.862d	cystine amide	$^3J_{MX} = 7.9$ Hz	" "
7.181d	valine amide	$^3J_{RX} = 8.3$ Hz	(A $_3$ B $_3$ MR system)
7.255m	aromatic ring		

Where

s = singlet

d = doublet

t = triplet

q = quartet

and m = unresolved multiplet

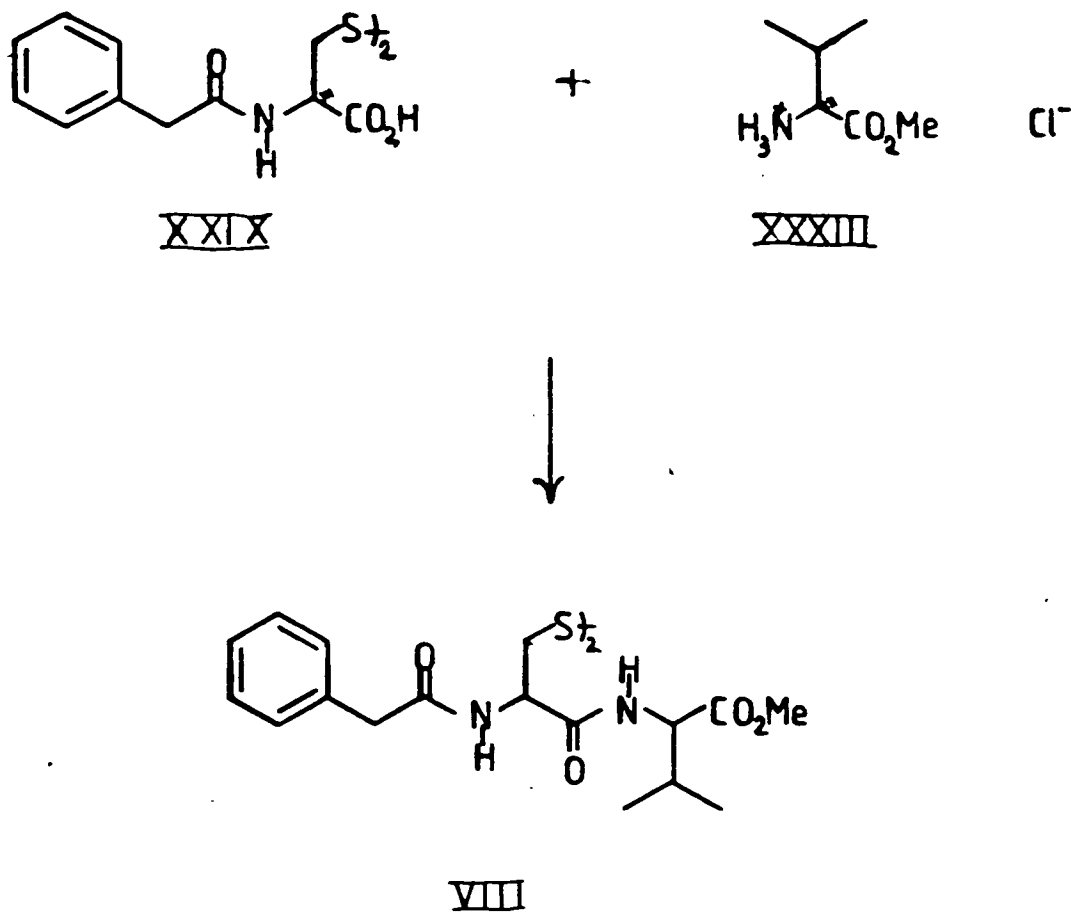
with dd meaning "doublet of doublets" etc.

As the cystine methylene protons are so poorly resolved, it is unlikely that the molecule takes up one conformation about the disulphide bridge and the values of all measurable coupling constants ($^3J = 7 \pm 2$ Hz) lend credence to the suggestion that the molecule is freely rotating about all possible bonds and hence does not exist in its intramolecular hydrogen bonded state. More evidence for this is obtained from the results of a nuclear Overhauser effect study of the molecule which was almost totally negative, showing only a slight interaction between the *ortho*- protons of the phenyl ring and the adjacent methylene group. It was impossible to assign the valine methyl resonances on the basis of n.O.e as was done for the cyclic disulphide. A reason for the difference in conformation between the crystalline compound and the compound in solution might lie in the close contact between the

cystine C(a) protons (see above) and the unusual conformation of the cystine bridge. Presumably the combination of hydrogen bonding interactions with the crystal packing forces is strong enough to cause this close contact and to prevent the disulphide linkage from taking up either a type I or a type II conformation, but in solution the hydrogen bonds alone are not strong enough to hold this region of the molecule together. If the molecule is undergoing free rotation, it would be expected that the molecule would appear to have two-fold symmetry on the nmr timescale with the exception of the cystine methylene protons which may only have a symmetrical arrangement in the non-helical type of bridge conformation (as in the crystalline state), leading to the superposition in the spectrum of at least two possible AB subsystems of the ABMX spin system and hence to the poor resolution and concomitant lack of measurable coupling constant that is observed in this spectrum.

Experimental

(Synthesis of the dimeric disulphide (VIII) as carried by E. M. Gordon, 1973) The scheme below shows the compounds involved in this synthesis



Organic solvents were purified and dried by established procedures, and organic extracts typically dried over Na₂SO₄. Chemicals were purchased from the Aldrich Chemical Co. Inc.

N,N'-bis(phenylacetyl)-L-cystine (XXIX) was prepared by the method of Foldi (Foldi, 1954) and had m.p. 171°C (lit 170-171°C)

Preparation of L-valine methyl ester hydrochloride (XXXIII)

To a slurry of L-valine (100g, 0.854mol) in absolute methanol (600ml) held at a temperature between -5°C and 0°C , thionyl chloride (130g, 0.96mol, 77ml) was added dropwise with vigorous stirring. After addition, the mixture was allowed to rise slowly to room temperature for two hours. The mixture was again cooled and another equal portion of thionyl chloride added dropwise. The mixture was then left stirring overnight, rising to room temperature.

The mixture was then filtered, and the methanol evaporated off at reduced pressure. The residue was dissolved in chloroform (400ml), filtered and again evaporated to dryness under reduced pressure, yielding an off-white crystalline solid (141g, 0.84mol).

Preparation of bis(N-phenylacetyl) L-cystinyl bis(L-valine methyl ester) (VIII)

Finely powdered L-valine methyl ester hydrochloride (XXXIII) (16.8g, 0.1mol) was mixed with bis(N-phenylacetyl) L-cystine (XXIX) (23.8g, 0.05mol) and EEDQ (24.7g, 0.1mol) in dry tetrahydrofuran (600ml). The mixture was stirred vigorously and N-methyl morpholine (10.8ml, 0.1mol) was injected in one portion. The mixture was then stirred overnight. The THF was removed by evaporation at reduced pressure and the residue was dissolved in ethyl acetate (50ml). The solution was washed twice with 1N HCl, once with 1N NaOH and once with saturated brine, before being dried over Na_2SO_4 and filtered. The solvent was removed under reduced pressure, and the residue was crystallised from hot chloroform and hexane (17.3g, 24.6mmol, fine white crystals, melting at $166^{\circ}\text{--}168^{\circ}\text{C}$. Analysis (expected for $\text{C}_{34}\text{H}_{46}\text{N}_4\text{O}_8\text{S}_2$, found): C (58.103, 58.06%); N (7.972, 7.92%); S (9.123, 9.12%); H (6.597, 6.37%). For ^1H nmr see text.

Table Fractional Coordinates of Atoms with Standard Deviations

	x	y	z	Ueq
S(1C1)	0.85003(25)	0.85438(22)	0.52920(15)	0.0812
N(1C1)	0.7526(6)	0.7837(5)	0.5765(4)	0.0511
O(1C1)	0.8112(5)	0.6747(5)	0.6047(3)	0.0721
O(1P1)	0.6725(5)	0.7939(6)	0.5357(4)	0.0787
N(1V1)	0.8043(7)	0.6348(6)	0.5445(4)	0.0701
C(1C1)	0.8543(8)	0.7721(7)	0.5446(5)	0.0604
C(2C1)	0.7888(7)	0.7442(7)	0.5521(5)	0.0486
C(3C1)	0.8015(7)	0.6829(8)	0.5703(5)	0.0483
C(1P1)	0.6986(9)	0.8054(8)	0.5661(5)	0.0449
C(2P1)	0.6739(8)	0.8567(8)	0.5941(5)	0.0797
C(4P1)	0.58751(0)	0.83613(0)	0.64120(0)	0.1163
C(5P1)	0.52441(0)	0.83439(0)	0.65027(0)	0.1404
C(6P1)	0.48043(0)	0.85108(0)	0.62247(0)	0.1535
C(7P1)	0.49956(0)	0.86951(0)	0.58560(0)	0.1406
C(8P1)	0.56267(0)	0.87125(0)	0.57653(0)	0.0939
C(3P1)	0.60664(0)	0.85456(0)	0.60432(0)	0.0664
C(1V1)	0.8227(8)	0.5728(8)	0.5575(5)	0.0607
C(2V1)	0.7705(10)	0.5262(10)	0.5509(6)	0.1017
C(3V1)	0.7163(10)	0.5417(9)	0.5743(6)	0.1055
C(4V1)	0.7544(9)	0.5199(9)	0.5083(6)	0.1003
C(5V1)	0.8825(9)	0.5585(10)	0.5383(6)	0.0776
O(1E1)	0.9140(6)	0.5897(7)	0.5188(4)	0.1041
O(2E1)	0.8962(7)	0.4973(7)	0.5454(4)	0.1196
C(1M1)	0.9548(11)	0.4769(11)	0.5277(7)	0.1483
S(1C4)	0.3283(3)	0.3131(3)	0.47153(20)	0.1214
O(1C4)	0.1597(7)	0.2614(8)	0.3882(4)	0.1230
O(1P4)	0.2649(8)	0.1402(6)	0.4731(4)	0.1242
N(1C4)	0.2604(7)	0.2073(7)	0.4244(4)	0.0779
N(1V4)	0.1118(6)	0.2668(7)	0.4450(4)	0.0792
C(1C4)	0.2533(10)	0.3143(9)	0.4463(6)	0.0969
C(2C4)	0.2225(8)	0.2532(8)	0.4444(5)	0.0616
C(3C4)	0.1620(10)	0.2592(9)	0.4234(7)	0.0721
C(1P4)	0.2802(9)	0.1590(11)	0.4411(7)	0.0758
C(2P4)	0.3225(10)	0.1134(11)	0.4164(6)	0.1162
C(4P4)	0.43409(0)	0.10031(0)	0.40089(0)	0.1441
C(5P4)	0.49640(0)	0.09900(0)	0.41193(0)	0.1477
C(6P4)	0.51305(0)	0.11307(0)	0.45000(0)	0.1744
C(7P4)	0.46739(0)	0.12844(0)	0.47703(0)	0.1804
C(8P4)	0.40509(0)	0.12974(0)	0.46599(0)	0.1492
C(3P4)	0.38844(0)	0.11568(0)	0.42792(0)	0.1283
C(1V4)	0.0494(11)	0.2807(12)	0.4269(7)	0.1120
C(2V4)	-0.0029(15)	0.2583(15)	0.4556(9)	0.1699
C(3V4)	0.0054(15)	0.1914(15)	0.4621(8)	0.2088
C(4V4)	-0.0038(15)	0.3034(15)	0.4895(9)	0.2225
C(5V4)	0.0495(18)	0.3480(17)	0.4200(10)	0.1645
O(1E4)	0.0881(12)	0.3862(13)	0.4292(7)	0.2150
O(2E4)	0.0019(12)	0.3608(12)	0.3984(7)	0.2166

Table Fractional Coordinates of Atoms with Standard Deviations

	x	y	z	Ueq
S(1C2)	0.34669(23)	0.83575(23)	0.52261(15)	0.0746
N(1C2)	0.2602(6)	0.7646(6)	0.5775(4)	0.0545
O(1C2)	0.3259(6)	0.6686(5)	0.6128(3)	0.0761
O(1P2)	0.1821(5)	0.7678(6)	0.5362(4)	0.0764
N(1V2)	0.3385(7)	0.6191(6)	0.5549(4)	0.0604
C(1C2)	0.3622(7)	0.7609(7)	0.5433(5)	0.0553
C(2C2)	0.3014(7)	0.7255(7)	0.5552(5)	0.0439
C(3C2)	0.3220(8)	0.6679(9)	0.5778(6)	0.0599
C(1P2)	0.2079(9)	0.7857(8)	0.5641(5)	0.0437
C(2P2)	0.1856(8)	0.8403(8)	0.5892(5)	0.0797
C(4P2)	0.10151(0)	0.81722(0)	0.63774(0)	0.0911
C(5P2)	0.03937(0)	0.81331(0)	0.64889(0)	0.1145
C(6P2)	-0.00738(0)	0.83047(0)	0.62303(0)	0.1118
C(7P2)	0.00801(0)	0.85153(0)	0.58603(0)	0.1210
C(8P2)	0.07016(0)	0.85544(0)	0.57488(0)	0.0887
C(3P2)	0.11691(0)	0.83828(0)	0.60073(0)	0.0713
C(1V2)	0.3687(9)	0.5662(10)	0.5715(6)	0.0813
C(2V2)	0.3344(11)	0.5076(11)	0.5650(7)	0.1041
C(3V2)	0.2760(11)	0.5096(12)	0.5862(7)	0.1396
C(4V2)	0.3194(12)	0.4924(13)	0.5233(7)	0.1687
O(1E2)	0.4525(7)	0.5917(8)	0.5268(5)	0.1368
O(2E2)	0.4713(9)	0.5271(8)	0.5777(6)	0.1591
C(5V2)	0.4335(12)	0.5603(12)	0.5511(8)	0.1119
C(1M2)	0.5292(15)	0.5199(16)	0.5617(9)	0.2552
S(1C3)	0.35363(23)	0.82302(25)	0.46403(16)	0.0836
N(1C3)	0.2895(6)	0.7076(6)	0.4272(4)	0.0579
O(1C3)	0.1779(5)	0.7470(7)	0.3952(4)	0.1056
O(1P3)	0.2982(6)	0.6409(6)	0.4770(3)	0.1049
N(1V3)	0.1383(7)	0.7663(8)	0.4549(4)	0.0983
C(1C3)	0.2765(7)	0.8148(7)	0.4474(5)	0.0563
C(2C3)	0.2492(7)	0.7507(7)	0.4497(5)	0.0558
C(3C3)	0.1833(9)	0.7513(8)	0.4304(6)	0.0598
C(1V3)	0.0749(10)	0.7845(10)	0.4385(6)	0.0925
C(2V3)	0.0252(11)	0.7452(10)	0.4545(7)	0.1035
C(3V3)	0.0408(12)	0.6751(11)	0.4436(7)	0.1591
C(4V3)	0.0189(11)	0.7459(11)	0.4980(7)	0.1303
O(1E3)	0.0937(8)	0.8831(8)	0.4703(5)	0.1438
O(2E3)	0.0191(8)	0.8709(9)	0.4292(5)	0.1407
C(5V3)	0.0667(12)	0.8450(13)	0.4484(7)	0.1206
C(1M3)	0.0104(12)	0.9327(13)	0.4345(8)	0.1739
C(1P3)	0.3137(20)	0.6590(16)	0.4452(8)	0.0956
C(2P3)	0.3561(6)	0.6209(6)	0.4185(5)	0.0956
C(4P3)	0.34512(0)	0.54460(0)	0.36728(0)	0.0956
C(5P3)	0.31785(0)	0.49429(0)	0.34834(0)	0.0956
C(6P3)	0.26657(0)	0.46459(0)	0.36468(0)	0.0956
C(7P3)	0.24256(0)	0.48521(0)	0.39995(0)	0.0956
C(8P3)	0.26984(0)	0.53553(0)	0.41889(0)	0.0956
C(3P3)	0.32112(0)	0.56522(0)	0.40256(0)	0.0956
C(1PX)	0.3161(21)	0.6615(20)	0.4463(8)	0.0956
C(2PX)	0.3718(7)	0.6347(7)	0.4243(5)	0.0956
C(4PX)	0.34372(0)	0.58174(0)	0.35943(0)	0.0956
C(5PX)	0.32688(0)	0.52824(0)	0.33920(0)	0.0956
C(6PX)	0.31988(0)	0.47210(0)	0.35889(0)	0.0956
C(7PX)	0.32973(0)	0.46946(0)	0.39881(0)	0.0956
C(8PX)	0.34657(0)	0.52296(0)	0.41904(0)	0.0956
C(3PX)	0.35356(0)	0.57910(0)	0.39935(0)	0.0956

Table Fractional Coordinates of Atoms with Standard Deviations

	x	y	z	Ueq
H(1C1)	0.7706	0.7961	0.6047	0.1984
H(1V1)	0.7924	0.6425	0.5145	0.1984
H(2C1)	0.8812	0.7692	0.5710	0.1984
H(3C1)	0.8770	0.7455	0.5221	0.1984
H(4C1)	0.7618	0.7394	0.5259	0.1984
H(1P1)	0.6626	0.9012	0.5807	0.1984
H(2P1)	0.7001	0.8533	0.6207	0.1984
H(3P1)	0.6216	0.8232	0.6627	0.1984
H(4P1)	0.5096	0.8201	0.6788	0.1984
H(5P1)	0.4316	0.8497	0.6295	0.1984
H(6P1)	0.4655	0.8824	0.5641	0.1984
H(7P1)	0.5775	0.8855	0.5480	0.1984
H(2V1)	0.8306	0.5706	0.5883	0.1984
H(3V1)	0.7870	0.4814	0.5605	0.1984
H(4V1)	0.6802	0.5080	0.5693	0.1984
H(5V1)	0.6997	0.5872	0.5662	0.1984
H(6V1)	0.7287	0.5416	0.6046	0.1984
H(7V1)	0.7955	0.5083	0.4921	0.1984
H(8V1)	0.7358	0.5633	0.4978	0.1984
H(9V1)	0.7202	0.4836	0.5046	0.1984
H(1M1)	0.9629	0.4287	0.5342	0.1984
H(2M1)	0.9927	0.5044	0.5389	0.1984
H(3M1)	0.9517	0.4830	0.4967	0.1984
H(1C4)	0.2720	0.2147	0.3943	0.1984
H(1V4)	0.1153	0.2629	0.4761	0.1984
H(2C4)	0.2610	0.3305	0.4171	0.1984
H(3C4)	0.2229	0.3462	0.4613	0.1984
H(4C4)	0.2154	0.2377	0.4738	0.1984
H(1P4)	0.3187	0.1259	0.3862	0.1984
H(2P4)	0.3057	0.0666	0.4204	0.1984
H(3P4)	0.4212	0.0894	0.3714	0.1984
H(4P4)	0.5317	0.0871	0.3910	0.1984
H(5P4)	0.5613	0.1121	0.4585	0.1984
H(6P4)	0.4803	0.1393	0.5065	0.1984
H(7P4)	0.3697	0.1416	0.4869	0.1984
H(2V4)	0.0402	0.2572	0.3998	0.1984
H(3V4)	-0.0493	0.2603	0.4438	0.1984
H(4V4)	-0.0305	0.1759	0.4815	0.1984
H(5V4)	0.0503	0.1835	0.4751	0.1984
H(6V4)	0.0023	0.1660	0.4352	0.1984
H(7V4)	-0.0394	0.2894	0.5097	0.1984
H(8V4)	-0.0135	0.3497	0.4793	0.1984
H(9V4)	0.0407	0.3028	0.5039	0.1984

Table Fractional Coordinates of Atoms with Standard Deviations

	x	y	z	Ueq
H(1C2)	0. 2736	0. 7767	0. 6067	0. 1984
H(1V2)	0. 3285	0. 6205	0. 5243	0. 1984
H(2C2)	0. 3906	0. 7670	0. 5688	0. 1984
H(3C2)	0. 3871	0. 7332	0. 5223	0. 1984
H(4C2)	0. 2751	0. 7122	0. 5298	0. 1984
H(1P2)	0. 1935	0. 8827	0. 5732	0. 1984
H(2P2)	0. 2128	0. 8408	0. 6154	0. 1984
H(3P2)	0. 1377	0. 8039	0. 6578	0. 1984
H(4P2)	0. 0274	0. 7970	0. 6775	0. 1984
H(5P2)	-0. 0555	0. 8274	0. 6317	0. 1984
H(6P2)	-0. 0282	0. 8648	0. 5660	0. 1984
H(7P2)	0. 0821	0. 8717	0. 5462	0. 1984
H(2V2)	0. 3724	0. 5730	0. 6024	0. 1984
H(3V2)	0. 3650	0. 4715	0. 5753	0. 1984
H(4V2)	0. 2508	0. 4670	0. 5816	0. 1984
H(5V2)	0. 2485	0. 5482	0. 5760	0. 1984
H(6V2)	0. 2855	0. 5151	0. 6167	0. 1984
H(7V2)	0. 3619	0. 4906	0. 5069	0. 1984
H(8V2)	0. 2895	0. 5278	0. 5113	0. 1984
H(9V2)	0. 2964	0. 4480	0. 5218	0. 1984
H(1M2)	0. 5591	0. 4949	0. 5813	0. 1984
H(2M2)	0. 5482	0. 5655	0. 5565	0. 1984
H(3M2)	0. 5259	0. 4948	0. 5347	0. 1984
H(1C3)	0. 2687	0. 7462	0. 4120	0. 1984
H(1V3)	0. 1464	0. 7654	0. 4857	0. 1984
H(2C3)	0. 2753	0. 8291	0. 4173	0. 1984
H(3C3)	0. 2474	0. 8454	0. 4643	0. 1984
H(4C3)	0. 2462	0. 7363	0. 4796	0. 1984
H(2V3)	0. 0733	0. 7786	0. 4074	0. 1984
H(3V3)	-0. 0172	0. 7632	0. 4421	0. 1984
H(4V3)	0. 0460	0. 6708	0. 4126	0. 1984
H(5V3)	0. 0034	0. 6456	0. 4533	0. 1984
H(6V3)	0. 0834	0. 6613	0. 4576	0. 1984
H(7V3)	0. 0084	0. 7924	0. 5075	0. 1984
H(8V3)	0. 0620	0. 7308	0. 5110	0. 1984
H(9V3)	-0. 0181	0. 7151	0. 5066	0. 1984
H(1M3)	-0. 0288	0. 9479	0. 4176	0. 1984
H(2M3)	0. 0514	0. 9576	0. 4253	0. 1984
H(3M3)	0. 0017	0. 9418	0. 4647	0. 1984
H(1P3)	0. 3313	0. 6050	0. 4438	0. 1500
H(2P3)	0. 3276	0. 6537	0. 4028	0. 1500
H(3P3)	0. 3848	0. 5676	0. 3546	0. 1500
H(4P3)	0. 3364	0. 4783	0. 3210	0. 1500
H(5P3)	0. 2455	0. 4256	0. 3500	0. 1500
H(6P3)	0. 2029	0. 4622	0. 4126	0. 1500
H(7P3)	0. 2513	0. 5515	0. 4462	0. 1500
H(1PX)	0. 3841	0. 6798	0. 4129	0. 1500
H(2PX)	0. 3877	0. 6310	0. 4539	0. 1500
H(3PX)	0. 3491	0. 6252	0. 3442	0. 1500
H(4PX)	0. 3193	0. 5303	0. 3083	0. 1500
H(5PX)	0. 3068	0. 4307	0. 3432	0. 1500
H(6PX)	0. 3243	0. 4260	0. 4141	0. 1500
H(7PX)	0. 3542	0. 5209	0. 4499	0. 1500

Thermal Vibration Parameters with Standard Deviations

	U or U11	U22	U33	U23	U13	U12
S(1C1)	0.1111(46)	0.0644(37)	0.0681(34)	0.0004(30)	-0.0176(34)	-0.0242(31)
N(1C1)	0.0558(103)	0.0482(91)	0.0494(98)	-0.0237(76)	-0.0086(79)	0.0111(76)
O(1C1)	0.1175(108)	0.0675(86)	0.0313(74)	-0.0036(66)	-0.0233(70)	0.0152(74)
O(1P1)	0.0631(88)	0.1092(106)	0.0640(94)	-0.0003(80)	-0.0036(77)	0.0183(73)
N(1V1)	0.1337(144)	0.0301(91)	0.0465(98)	0.0049(80)	-0.0080(95)	0.0135(85)
S(1C2)	0.0727(39)	0.0694(38)	0.0818(39)	0.0045(32)	-0.0055(32)	0.0010(28)
N(1C2)	0.0721(115)	0.0563(102)	0.0352(89)	-0.0045(76)	0.0116(84)	0.0123(85)
O(1C2)	0.1209(113)	0.0965(104)	0.0109(70)	0.0049(66)	-0.0056(69)	0.0333(80)
O(1P2)	0.0530(87)	0.1079(105)	0.0684(97)	-0.0202(83)	-0.0253(77)	0.0138(72)
N(1V2)	0.0981(125)	0.0542(102)	0.0289(88)	0.0091(82)	0.0118(88)	0.0147(90)
S(1C3)	0.0672(38)	0.0971(44)	0.0864(42)	0.0106(34)	0.0132(33)	-0.0065(31)
N(1C3)	0.0608(102)	0.0687(108)	0.0444(93)	-0.0235(82)	-0.0026(80)	0.0101(76)
O(1C3)	0.0633(92)	0.1975(156)	0.0560(101)	-0.0061(97)	0.0016(78)	0.0119(93)
O(1P3)	0.1696(138)	0.0877(103)	0.0575(98)	-0.0059(81)	-0.0224(96)	0.0343(94)
N(1V3)	0.0496(110)	0.1847(185)	0.0607(117)	-0.0398(117)	-0.0076(96)	0.0103(113)
S(1C4)	0.0747(44)	0.1350(58)	0.1543(59)	0.0280(51)	-0.0070(44)	-0.0228(37)
O(1C4)	0.1226(130)	0.2113(171)	0.0350(86)	0.0051(97)	-0.0074(90)	0.0600(115)
O(1P4)	0.2275(175)	0.0675(99)	0.0775(108)	0.0380(84)	0.0601(114)	0.0536(104)
N(1C4)	0.1028(133)	0.0911(128)	0.0399(98)	0.0216(95)	0.0225(89)	0.0523(106)
N(1V4)	0.0358(99)	0.1556(162)	0.0460(104)	-0.0164(101)	-0.0168(86)	0.0116(98)

CHAPTER FIVE
N-ACETYL METHIONINE

***N*-ACETYL METHIONINE**

Introduction

N-Acetyl amino acids are useful model compounds for the peptide linkage in the shorter peptides. They have an advantage over simpler compounds like acetamide or N-methyl formamide in that they can participate more fully in the network of hydrogen bonds characteristic of amino acid and peptide structures. Methionine is itself an essential amino acid for human growth. Its main function (apart from as a 'structural' amino acid in proteins) is as a source of methyl groups as in S-adenosyl methionine. Although methionine is found in one of the enkephalin peptides, it is frequently replaced in this role by leucine - this would imply that the sole function of methionine here is to fill a hydrophobic pocket in the receptor. The crystals used in this determination were grown in the course of an attempted preparation of a molecular complex of acetylcholine and the amino acid. (These crystals were prepared by Miss R. Kelly as part of an Honours project on amino-acid:alkaloid complexes.)

Crystallisation

The crystals were grown from an equimolar solution of N-acetyl methionine and acetylcholine chloride in water, to which had been added some acetone as an aid to crystallisation.

Crystal Data

Formula $C_7H_{13}O_3NS$ Mol. Wt. 191.23 $F(000) = 204$

Space group $P2_1$ Int. Tab. No. 4

Cell dimensions $a = 6.002(3)$ $b = 9.2499(13)$ $c = 9.1737(9)$ Å
 $\beta = 91.234(20)^\circ$

$V = 508.71 \text{ \AA}^3$ $Z = 2$ $D_c = 1.24 \text{ gcm}^{-3}$

Radiation $\text{MoK}\alpha$ $\lambda = 0.71069 \text{ \AA}$ $\mu = 2.77 \text{ cm}^{-1}$

Final $R = 0.0465$ based on 840 data

$h_{\text{max}} = 7$ $k_{\text{max}} = 11$ $l_{\text{max}} = 10$

$\theta_{\text{max}} = 25^\circ$ $\text{Sin}\theta_{\text{max}} = 0.4226$ $\text{Sin}\theta/\lambda = 0.5947$

Reflections measured 1011 (5 systematic absences)
 After merging 947 (R = 0.0326)
 Unobserved reflections 107 (F<4σ(F))
 Intensity standards 1 5 3 2 3 5
 0.993 < drift curve < 1.007
 115 parameters refined
 Max shift/σ for final cycle 0.005
 Final difference Fourier max. 0.1967 min. -0.3283 e Å⁻³
 Weighting scheme :weight = 10.3571/(σ²(F)+0.00059F²)

Data collection and reduction

The data set was collected on the CAD4 and corrected for Lorentz and polarisation effects using the program CADABS. The space group was determined from the systematic absences (reflection conditions, 0k0, k=2n) and the chirality of the compound to be P2₁.

Solution and refinement

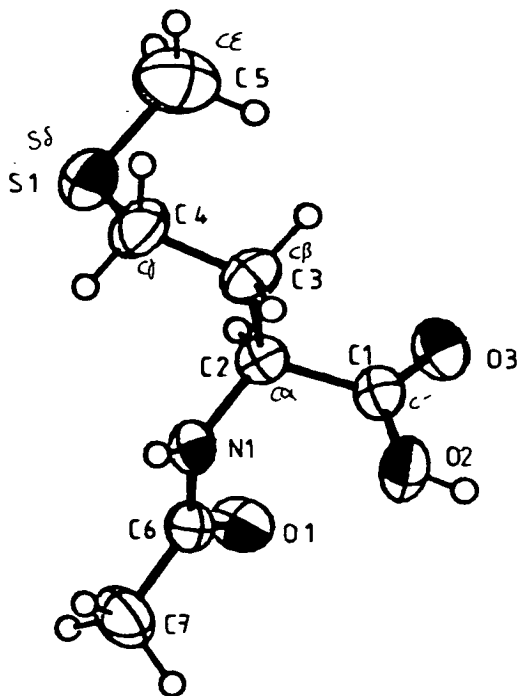
The structure was solved by location of the sulphur sulphur vector in the sharpened Patterson map calculated by the program SHELX-76. The progress of the refinement is shown below.

R	Model
0.4574	One Sulphur atom only
0.1271	All non-Hydrogen atoms included with isotropic thermal parameters
0.1203	Hydrogens input in calculated positions with constrained geometry and thermal parameters
0.0775	All non-Hydrogen atoms with anisotropic thermal parameters
0.0707	O-H and N-H hydrogens input (not constrained)
0.0539	Three reflections with deviation greater than 6σ omitted and weighting scheme applied
0.0536	Converged
	Absolute stereochemistry corrected
0.0465	Final Convergence

All hydrogens on carbon atoms were constrained to calculated positions, with all their thermal parameters constrained to be equal

Structure Description

The numbering scheme used for this molecule is shown in the diagram below.



As can be seen from the tables of molecular geometry below, all the bond distances, angles and dihedral angles are within the usual ranges for their types.

C(1) - C(2)	1.525(6)
C(1) - O(2)	1.311(6)
C(1) - O(3)	1.202(6)
C(2) - C(3)	1.519(6)
C(2) - N(1)	1.458(5)
C(3) - C(4)	1.532(8)
S(1) - C(4)	1.807(6)
S(1) - C(5)	1.766(8)
C(6) - C(7)	1.510(8)
C(6) - O(1)	1.239(6)
C(6) - N(1)	1.324(6)

C(4) - S(1) - C(5) 102.5(3)
 C(2) - C(1) - O(3) 122.2(4)
 O(2) - C(1) - O(3) 124.4(4)
 C(2) - C(1) - O(2) 113.2(4)
 C(1) - C(2) - C(3) 108.8(3)
 C(1) - C(2) - N(1) 111.5(3)
 C(3) - C(2) - N(1) 110.9(3)
 C(2) - C(3) - C(4) 115.1(4)
 S(1) - C(4) - C(3) 115.3(4)
 O(1) - C(6) - N(1) 120.0(4)
 C(7) - C(6) - O(1) 121.6(4)
 C(2) - N(1) - C(6) 121.0(4)
 C(7) - C(6) - N(1) 118.4(4)

C(5) - S(1) - C(4) - C(3) -63.5(5)
 O(2) - C(1) - C(2) - C(3) 88.7(4)
 O(2) - C(1) - C(2) - N(1) -33.9(5)
 O(3) - C(1) - C(2) - C(3) -87.0(5)
 O(3) - C(1) - C(2) - N(1) 150.4(4)
 C(1) - C(2) - C(3) - C(4) 174.9(4)
 N(1) - C(2) - C(3) - C(4) -62.1(5)
 C(1) - C(2) - N(1) - C(6) -59.0(5)
 C(3) - C(2) - N(1) - C(6) 179.6(4)
 C(2) - C(3) - C(4) - S(1) -54.7(5)
 C(7) - C(6) - N(1) - C(2) 175.8(4)
 O(1) - C(6) - N(1) - C(2) -3.5(6)

The bond distances and angles involving the positionally constrained hydrogen atoms are not listed, but the table below gives the figures for the unconstrained hydrogens of the amide and carboxylic acid functional groups. Both of the bond distances are shorter than usual, but not significantly so, and the angles are well within the usual ranges.

O(2) -H(002) 0.80(8)
 N(1) -H(0N1) 0.87(7)

C(1) - O(2) -H(002) 120(6)
 C(2) - N(1) -H(0N1) 123(5)
 C(6) - N(1) -H(0N1) 116(5)

C(2) - C(1) - O(2) -H(002) -179(7)
 O(3) - C(1) - O(2) -H(002) -4(7)
 C(1) - C(2) - N(1) -H(0N1) 113(6)
 H(021)- C(2) - N(1) -H(0N1) -128(6)
 C(3) - C(2) - N(1) -H(0N1) -8(6)
 C(7) - C(6) - N(1) -H(0N1) 3(5)
 O(1) - C(6) - N(1) -H(0N1) -176(5)

Molecular Packing

The hydrogen bonding scheme for N-acetyl methionine consists of parallel sheets of molecules perpendicular to the c axis. The hydrogen bonded contacts are listed below.

O2	2.560 (5)	O1	(1 -X, 1/2+Y, -Z)
O3	2.909 (5)	N1	(1 +X, Y, Z)
H002	1.78 (8)	O1	(1 -X, 1/2+Y, -Z)
O3	2.04 (7)	H0N1	(1 +X, Y, Z)

These hydrogen bonding contacts may be seen clearly on all three unit cell axis projections (shown on the next three pages after this section). On the projection down the a axis they can be seen to form a staggered sheet connecting the carboxyl hydrogens to the amide carbonyls and on the projection down the b axis the amide hydrogens can be seen hydrogen bonding to the carboxyl carbonyls. Between these sheets of hydrophilic atoms lie the hydrophobic regions containing the side chains arranged head to tail about the screw axis. The close contacts to sulphur in this region can be seen in the table below.

C1	3.960 (4)	S1	(1 -X, 1/2+Y, -1 -Z)
C3	3.856 (5)	S1	(-X, 1/2+Y, -1 -Z)
C4	3.937 (6)	S1	(-X, 1/2+Y, -1 -Z)
O3	3.451 (4)	S1	(1 -X, 1/2+Y, -1 -Z)
H031	2.946 (5)	S1	(-X, 1/2+Y, -1 -Z)
H032	3.427 (5)	S1	(1 -X, 1/2+Y, -1 -Z)
H072	3.062 (6)	S1	(X, Y, 1 +Z)

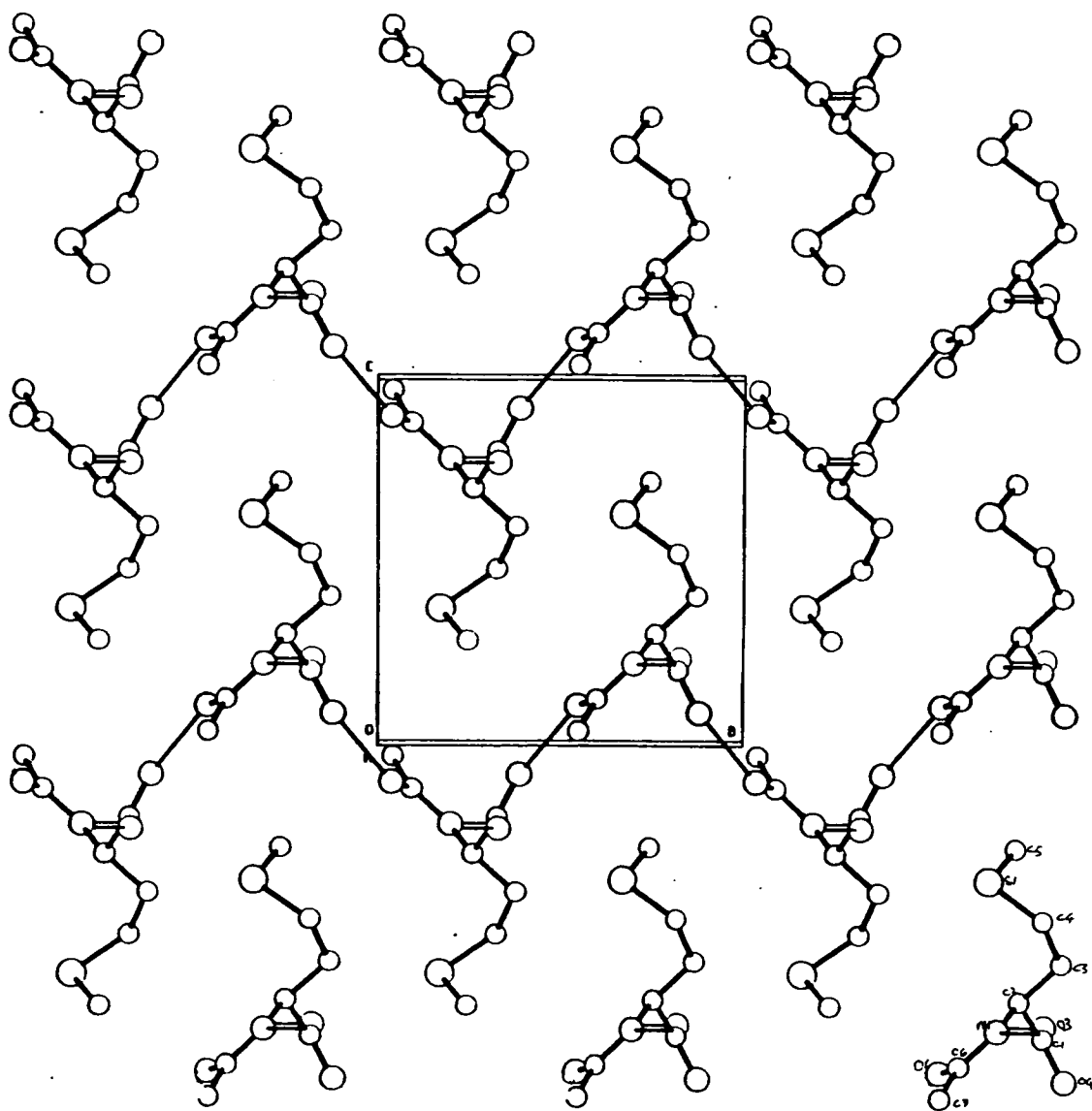
There are eight significant oxygen hydrogen contacts, but these involve

mostly methyl groups with hydrogens in calculated positions so they may not be genuine contacts.

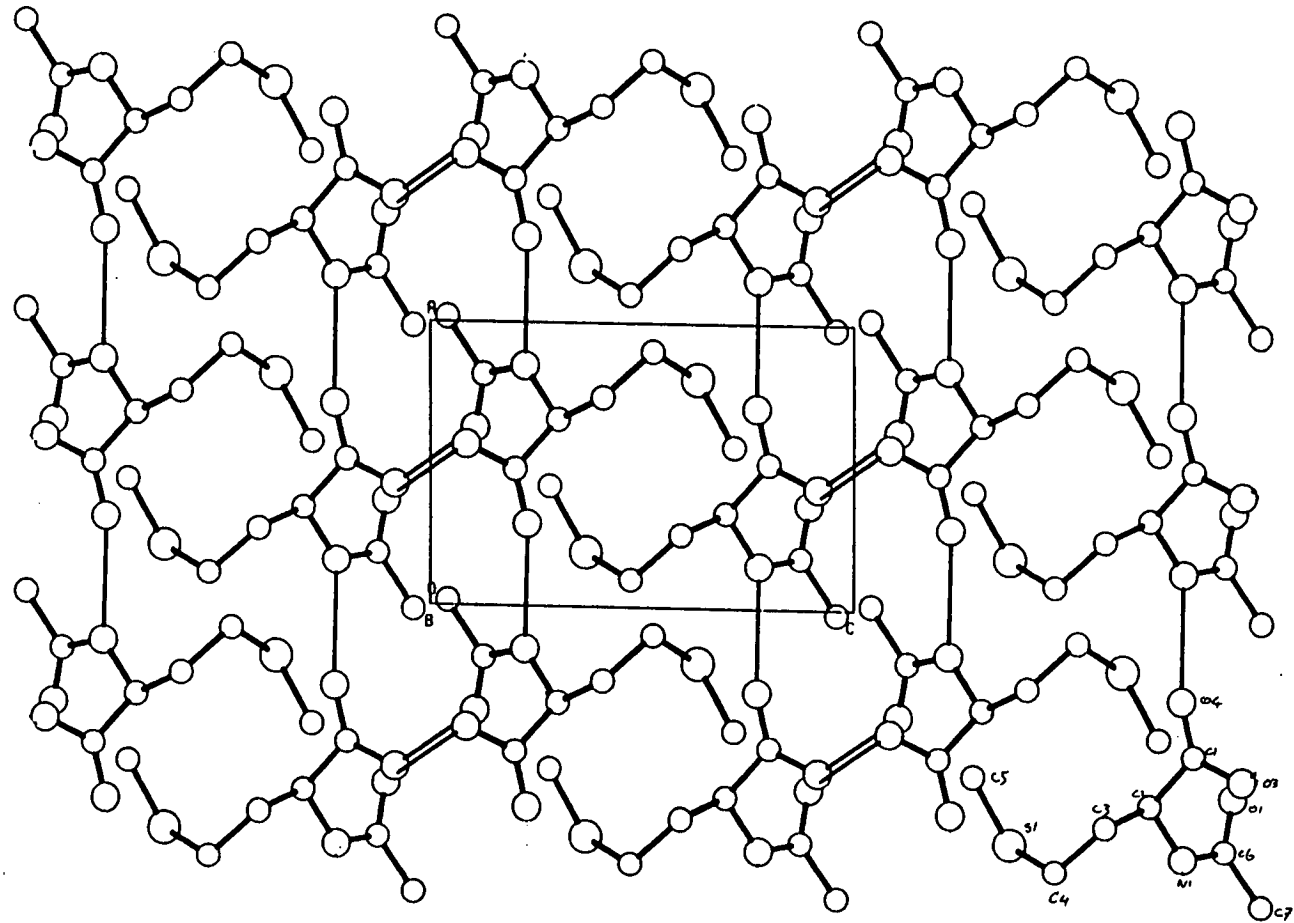
H053	2.639	(9)	O1	(1 -X, 1/2+Y, -1 -Z)
H071	2.869	(7)	O1	(-1 +X, Y, Z)
O2	2.846	(9)	H051	(1 -X, 1/2+Y, -1 -Z)
O2	2.839	(9)	H053	(X, Y, 1 +Z)
O2	2.978	(7)	H071	(-X, 1/2+Y, -Z)
O2	2.714	(7)	H073	(-X, 1/2+Y, -Z)
O3	2.745	(7)	H041	(1 +X, Y, Z)
O3	2.600	(7)	H071	(1 +X, Y, Z)

The remaining contact distances involve carbon atoms.

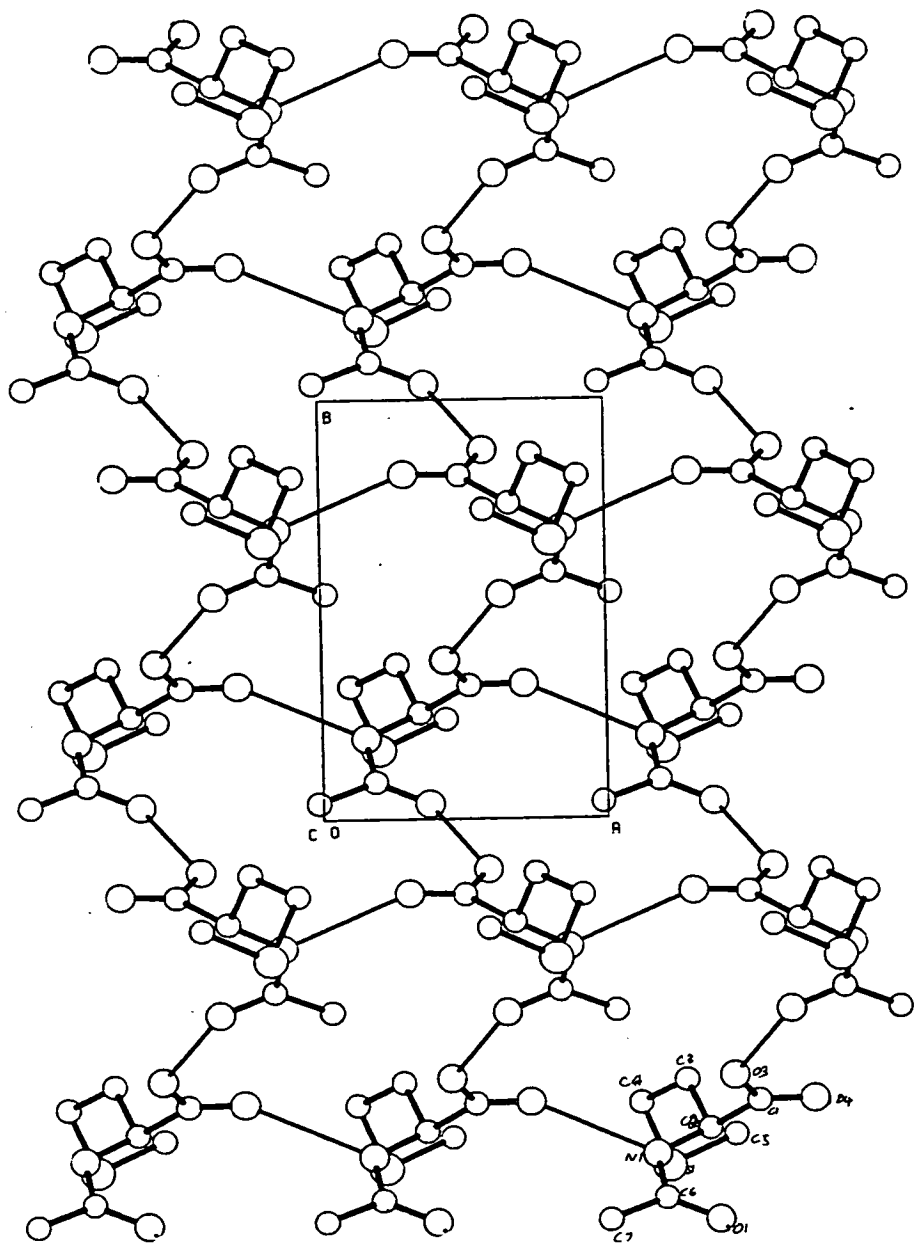
C5	3.416	(9)	O1	(1 -X, 1/2+Y, -1 -Z)
O2	3.114	(7)	C7	(-X, 1/2+Y, -Z)
N1	3.14	(7)	C1	(-1 +X, Y, Z)
H002	2.80	(8)	C6	(1 -X, 1/2+Y, -Z)



***N*-ACETYL-*L*-METHIONINE**
PROJECTION DOWN THE X AXIS



PROJECTION DOWN THE YAXIS
N-ACETYL-L-METHIONINE



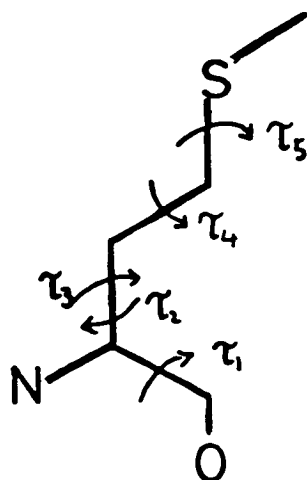
PROJECTION DOWN THE Z AXIS
***N*-ACETYL-*L*-METHIONINE**

Comparison with other structures

The table overleaf lists the torsion angles of twenty independent methionine residues from sixteen structure determinations. Bond lengths and angles are not included in this table as they are all within normal variation of the standard values. There would appear to be three types of side chain found. These may be described as follows:

- a) Fully extended ($\tau_4 = \tau_5 \approx 180^\circ$)
- b) Fully folded ($\tau_4 = \tau_5 \approx 60^\circ$)
- c) An intermediate conformation with the side chain half extended ($\tau_4 \approx 180^\circ, \tau_5 \approx \pm 60^\circ$).

The torsion angles tabulated are marked on the diagram below.



τ_1	is $\chi(O-C'-C\alpha-C\beta)$
τ_2	is $\chi(C'-C\alpha-C\beta-C\gamma)$
τ_3	is $\chi(N-C\alpha-C\beta-C\gamma)$
τ_4	is $\chi(C\alpha-C\beta-C\gamma-S\delta)$
τ_5	is $\chi(C\beta-C\gamma-S\delta-Ce)$

N-acetyl-L-methionine is an example of the second type, while N-acetyl methionine methyl ester is of the first type, but otherwise similar in conformation along the peptide backbone. Both the α and β forms of L-methionine differ in side chain conformation from the N-acetyl methionine, but they are of similar conformation about the $C'-C\alpha$ and $C\alpha-C\beta$ bonds to N-acetyl methionine. The side chain of cyclo-L-methionyl glycine trihydrate has the same conformation as that

of N-acetyl methionine.

	τ_1	τ_2	τ_3	τ_4	τ_5
ACMETM	94.20 -83.41	175.36	-63.63	177.46	-178.61
ACMTDE	70.16	62.13	-179.53	176.03	-74.52
ALAMET10	119.06 -59.96	-56.24	67.26	175.00	-173.77
AMETMA	100.66	169.51	-72.30	173.36	-176.72
BCMEGL	51.94	165.58	-73.51	175.29	174.87
BOMETM	-47.01 -19.61 163.25	-58.02 -177.67	177.62 -53.38	-60.33 -177.35	-66.45 70.67
DLMETA02	-86.35 89.46	-178.62	-59.87	177.36	68.95
DLMETA03	-87.14 89.95	-174.36	-54.35	-179.32	-174.43
FMLPLA	95.56	-58.61	61.74	172.90	-79.29
FOLMET10	108.56 -70.57	-67.88	54.37	176.10	173.64
LMETON10	105.97 -72.97 84.25 -90.70	71.49 74.29	-166.06 -165.62	174.17 73.57	179.71 73.56
METHCL	124.77 -58.56	-60.85	61.30	-178.19	68.74
METMET	70.70 114.94 -61.74	74.63 -57.65	-167.24 68.08	-168.29 -170.50	178.39 -57.65
METTS010	-74.95 102.22	-46.56	77.00	-164.11	-65.90

	τ_1	τ_2	τ_3	τ_4	τ_5
SMMETC10	-37.07 141.74	82.50	-157.15	169.60	-177.31 -72.25
GASTRN10	101.89 92.13	174.97 179.21	-67.39 -63.84	-164.48 178.09	-165.63 81.06
BIVMUO	-56.36	-69.19	58.18	71.40	52.26
MGHPHE20	61.36	71.72	-170.20	176.11	120.15

Key to compounds in the above table

ACMETM N-acetyl-methionine methyl ester

A. J. Geddes, S. Hamodrakas, B. Sheldrick, Cryst. Struct. Commun., 3, 97, 1974

ACMTDE N-acetyl-D,L-methionine-diethylamide

A. Aubry, J. Protas, M. T. Cung, M. Marraud, Acta Cryst., B35, 2634, 1979

ALAMET10 D,L-alanyl-L,D-methionine

R. E. Stenkamp, L. H. Jensen, Acta Cryst., B30, 1541, 1974

AMETMA N-acetyl-methionyl-dimethylamide

A. Aubry, M. Marraud, J. Protas, J. Neel, C. R. Acad. Sci., Ser. C, 273, 959, 1971

BCMEGL N-(T-butoxy-carbonyl)-L-methionyl-glycine-benzyl ester

T. Yamane, T. Umemura, T. Kojima, Y. Yamada, T. Ashida, Bull. Chem. Soc. Jpn., 53, 908, 1980

- BOMETM** T-butoxycarbonyl-D-methionyl-L-methionine
methyl ester
- A. Immirzi, P. Avena, M. R. Ciajolo, J. M. Becker, F. Naider, Acta Cryst.,
B34, 179, 1978
- DLMETA02** D,L-methionine
(α -form, at 333° K)
- T. Taniguchi, Y. Takaki, K. Sakurai, Bull. Chem. Soc. Jpn., **53**, 803, 1980
- DLMETA03** D,L-methionine
(β -form, at 293° K)
- T. Taniguchi, Y. Takaki, K. Sakurai, Bull. Chem. Soc. Jpn., **53**, 803, 1980
- FMLPLA** N-formyl-methionyl-leucyl-phenylalanine
- A. J. Morffew, I. Tickle, Cryst. Struct. Commun., **10**, 781, 1981
- FOLMET10** N-formyl-L-methionine
(Absolute configuration)
- C. Chen, R. Parthasarathy, Acta Cryst., **B33**, 3332, 1977
- LMETON10** L-methionine
- K. Torii, Y. Iitaka, Acta Cryst., **B29**, 2799, 1973
- METHCL** L-methionine hydrochloride
- B. Di Blasio, V. Pavone, C. Pedone, Cryst. Struct. Commun., **6**, 845, 1977
- METMET** L-methionyl-L-methionine
- R. E. Stenkamp, L. H. Jensen, Acta Cryst., **B31**, 857, 1975
- METTS010** (2S,SR)-methionine sulphoximine
(Absolute configuration)

S. Neidle, D. Rogers, J. Chem. Soc. B., 694, 1970

SMMETC10 S-methyl-L-methionine chloride hydrochloride
(Vitamin U hydrochloride)

G. Del Re, E. Gavuzzo, E. Giglio, F. Leij, F. Mazza, V. Zappia, Acta Cryst.,
B33, 3289, 1977

GASTRN10 L-tryptophanyl-L-methionyl-L-aspartyl
-L-phenylalanylamide hydrochloride methanol
diethyl ether solvate

W. B. T. Cruse, E. Egert, M. A. Viswamitra, O. Kennard, Acta Cryst., B38,
1758, 1982

BIVMUO cyclo-L-methionyl-glycine trihydrate

M. Bressan, R. Ettore, F. Marchiori, G. Valle, Int. J. Pept. Protein Res.,
19, 402, 1982

MGHPHE20 L-methionyl-L- α -glutamyl-L-histidyl
-L-phenylalanine monohydrate

G. Admiraal, A. Vos, Acta Cryst., C39, 82, 1983

Table Fractional Coordinates of Atoms with Standard Deviations

	x	y	z	U _{eq}
S(1)	0.1933(3)	-0.83453(0)	-0.63280(14)	0.0780
C(1)	0.5128(6)	-0.6803(5)	-0.1997(4)	0.0418
C(2)	0.3375(6)	-0.7462(5)	-0.3021(4)	0.0379
C(3)	0.2590(8)	-0.6296(6)	-0.4080(5)	0.0490
C(4)	0.0967(8)	-0.6824(8)	-0.5269(5)	0.0644
C(5)	0.4286(14)	-0.7610(10)	-0.7182(7)	0.0970
C(6)	0.1813(6)	-0.9104(5)	-0.1259(4)	0.0451
C(7)	-0.0165(8)	-0.9605(9)	-0.0407(6)	0.0737
O(1)	0.3694(5)	-0.9632(4)	-0.1054(4)	0.0572
O(2)	0.4276(5)	-0.6204(4)	-0.0846(4)	0.0582
O(3)	0.7074(4)	-0.6770(5)	-0.2282(4)	0.0589
N(1)	0.1504(5)	-0.8062(5)	-0.2234(4)	0.0412
H(021)	0.4111	-0.8355	-0.3602	0.0932
H(031)	0.1772	-0.5450	-0.3476	0.0932
H(032)	0.4039	-0.5848	-0.4593	0.0932
H(041)	-0.0564	-0.7146	-0.4764	0.0932
H(042)	0.0639	-0.5924	-0.5996	0.0932
H(051)	0.5023	-0.8431	-0.7858	0.0932
H(052)	0.5493	-0.7259	-0.6371	0.0932
H(053)	0.3781	-0.6688	-0.7839	0.0932
H(071)	-0.1631	-0.9007	-0.0755	0.0932
H(072)	0.0155	-0.9411	0.0730	0.0932
H(073)	-0.0433	-1.0756	-0.0586	0.0932
H(002)	0.5081	-0.5823	-0.0260	0.0932
H(0N1)	0.0161	-0.7692	-0.2291	0.0932

Thermal Vibration Parameters with Standard Deviations

	U11	U22	U33	U23	U13	U12
S(1)	0.1011(12)	0.0770(10)	0.0557(8)	-0.0102(7)	-0.0059(7)	-0.0254(8)
C(1)	0.0354(19)	0.0392(20)	0.0507(24)	0.0040(20)	-0.0052(17)	0.0049(18)
C(2)	0.0336(18)	0.0401(21)	0.0399(21)	-0.0006(17)	0.0008(16)	0.0074(16)
C(3)	0.0552(23)	0.0468(24)	0.0447(22)	-0.0009(20)	-0.0063(19)	0.0099(20)
C(4)	0.0597(26)	0.0824(35)	0.0507(27)	0.0022(29)	-0.0168(22)	0.0033(29)
C(5)	0.1199(53)	0.0906(51)	0.0811(44)	0.0013(40)	0.0290(40)	0.0043(46)
C(6)	0.0395(22)	0.0509(26)	0.0448(25)	0.0021(20)	-0.0019(18)	-0.0021(18)
C(7)	0.0467(24)	0.0957(46)	0.0785(39)	0.0254(35)	-0.0021(25)	-0.0140(30)
O(1)	0.0464(17)	0.0627(22)	0.0624(21)	0.0185(18)	0.0016(15)	0.0103(15)
O(2)	0.0385(15)	0.0751(24)	0.0611(20)	-0.0279(19)	0.0003(14)	-0.0033(16)
O(3)	0.0271(13)	0.0782(23)	0.0713(21)	-0.0001(19)	0.0045(12)	0.0046(15)
N(1)	0.0294(15)	0.0453(21)	0.0488(18)	0.0013(16)	-0.0036(14)	0.0042(15)

CHAPTER SIX
DL-GLUTAMIC ACID MONOHYDRATE

DL GLUTAMIC ACID MONOHYDRATE

Introduction

Crystals of this compound were grown from a mixture of glutamic acid and brucine in aqueous ethanol as part of an attempted preparation of brucine glutamate crystals. Although the crystals formed were obviously not those of an amino-acid:alkaloid complex, it was decided that the X-ray structure would be of value as it could provide more information on the preferred environment and conformation of the glutamate residue as it might appear in biologically important molecules. Glutamic acid plays an important role in the biosynthesis of amino acids, by acting as a source of nitrogen for transamination. It is also likely that it acts as a transmitter in the central nervous system, so a structural study on glutamic acid could possibly provide more information on which to base the rational design of inhibitors or stimulants for this system.

It should also be noted here that the crystal structure of this compound was determined previously by Ciunik and Glowiak and published by them after the refinement process here was completed (Ciunik & Glowiak, 1983).

Crystallisation

Equimolar quantities of DL-glutamic acid and brucine tetrahydrate were dissolved in the minimum of hot aqueous ethanol (50% v/v), and allowed to cool slowly to room temperature. The crystals were cuboidal in shape, rather like flattened needles.

Crystal Data

Formula $C_5H_9NO_4 \cdot H_2O$ Mol. Wt. 165.2 $F(000) = 704$
Space Group $Pbca$ Int. Tab. No. 61
Cell Dimensions $a = 9.1469(15)$ $b = 15.5228(13)$ $c = 10.6407(10)$
 $V = 1510.8 \text{ \AA}^3$ $Z = 8$ $D_c = 1.4524 \text{ gcm}^{-3}$
Radiation $MoK\alpha$ $\lambda = 0.71069 \text{ \AA}$ $\mu = 1.22 \text{ cm}^{-1}$
Final $R = 0.0376$ based on 1154 data
 $h_{\text{max}} = 11$ $k_{\text{max}} = 19$ $l_{\text{max}} = 13$

$$\theta_{\max} = 25^\circ \quad \sin\theta_{\max} = 0.4226 \quad \sin\theta_{\max} / \lambda = 0.5947$$

1979 reflections measured (283 systematic absences)

1641 data remaining after merging (R = 0.0166)

487 data considered unobserved ($F < 4\sigma(F)$)

Crystal dimensions 0.8 * 0.08 * 0.2 mm³

Intensity standards * 6 1 3 and 1 4 7

0.976 < drift curve < 1.019

135 parameters refined

Max. shift/ σ for final cycle = 0.015

Final difference Fourier map max. 0.346 eA⁻³ min. -0.248 eA⁻³

Data collection and reduction

The data was collected on the CAD4 and the program CADABS used to correct for Lorentz and polarisation effects.

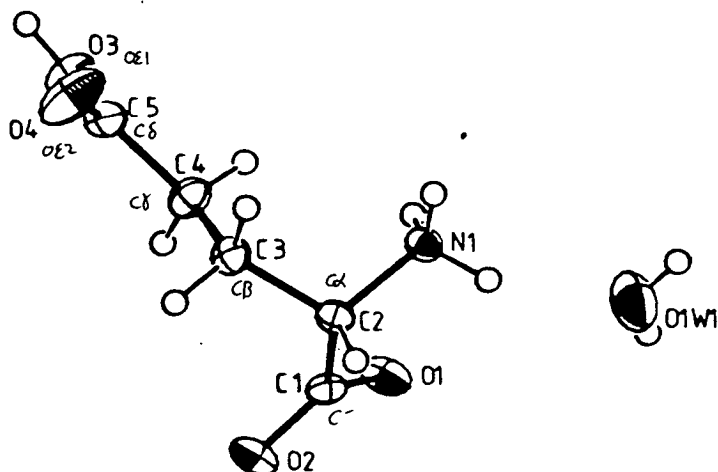
Solution and refinement

The structure was solved automatically, using the centrosymmetric structure solution routine of SHELX-76. The first difference map output showed all the non-hydrogen atoms, and the subsequent progress of refinement is described in the table below.

R	Model
0.4384	All non Hydrogen atoms input with isotropic thermal parameter of 0.05
0.2475	One cycle of least squares refinement.
0.1184	Hydrogens on carbon and nitrogen input in calculated positions with thermal parameter fixed at 0.05
0.1031	Thermal parameters for hydrogens on carbon and nitrogen tied to a free variable
0.0559	All non hydrogen atoms now with anisotropic thermal parameters
0.0474	Hydrogen atoms on oxygens found and given temperature factors fixed at 0.05
0.0495	Oxygen hydrogens thermal parameters tied to a free variable, and reflections with uneven backgrounds omitted
0.0376	Final Convergence

No positional constraints were applied to the hydrogen atoms, but the thermal parameters of the hydrogens bonded to carbon and nitrogen were constrained to be equal

Description of structure



The diagram above shows a perspective plot of the molecule, together with its numbering scheme as used in the tables below. All bond distances and angles are within the usual ranges, although one bond (C(4)-C(5)) is significantly shorter than the average carbon carbon bond length in this structure (1.496(3) angstroms vs. 1.5185 angstroms, a difference of 7.5 σ). There is a significant difference between the lengths of the two carbon oxygen bonds in the carboxylate group (5.6 σ), which may be due to one oxygen taking part in two hydrogen bonds, and the other only in one.

C(1)-O(1)	1.2421(25)	O(1)--N(1)	2.763(3)
C(1)-O(2)	1.2694(24)	O(2)--N(1)	2.827(3)
		O(2)--O(3)	2.572(3)

With the exception of one bond in the ammonium group, and the two on the water molecule, all bonds to hydrogen are within the usual range. The overall geometry of the molecule can be described in terms of two planes which intersect at C α (C(2)) with an angle of almost 90° (89.4°). The first of these planes is defined by the carboxylate group and the nitrogen atom, with a root mean square deviation from planarity of 0.014 angstroms, and the second, less well defined plane is that of the carboxyl group and the rest of the carbon skeleton, with a deviation of 0.087 angstroms from planarity.

C(1) - C(2) 1.526(3)
 C(1) - O(1) 1.2422(24)
 C(1) - O(2) 1.2679(24)
 C(2) - C(3) 1.527(3)
 C(2) - N(1) 1.4968(25)
 C(3) - C(4) 1.526(3)
 C(4) - C(5) 1.498(3)
 C(5) - O(3) 1.327(3)
 C(5) - O(4) 1.209(3)

C(2) - C(1) - O(1) 118.49(17)
 C(2) - C(1) - O(2) 115.65(17)
 O(1) - C(1) - O(2) 125.86(18)
 C(1) - C(2) - C(3) 113.52(16)
 C(1) - C(2) - N(1) 108.59(15)
 C(3) - C(2) - N(1) 110.69(15)
 C(2) - C(3) - C(4) 113.78(17)
 C(3) - C(4) - C(5) 113.28(17)
 C(4) - C(5) - O(3) 112.77(18)
 C(4) - C(5) - O(4) 125.24(20)
 O(3) - C(5) - O(4) 121.99(20)

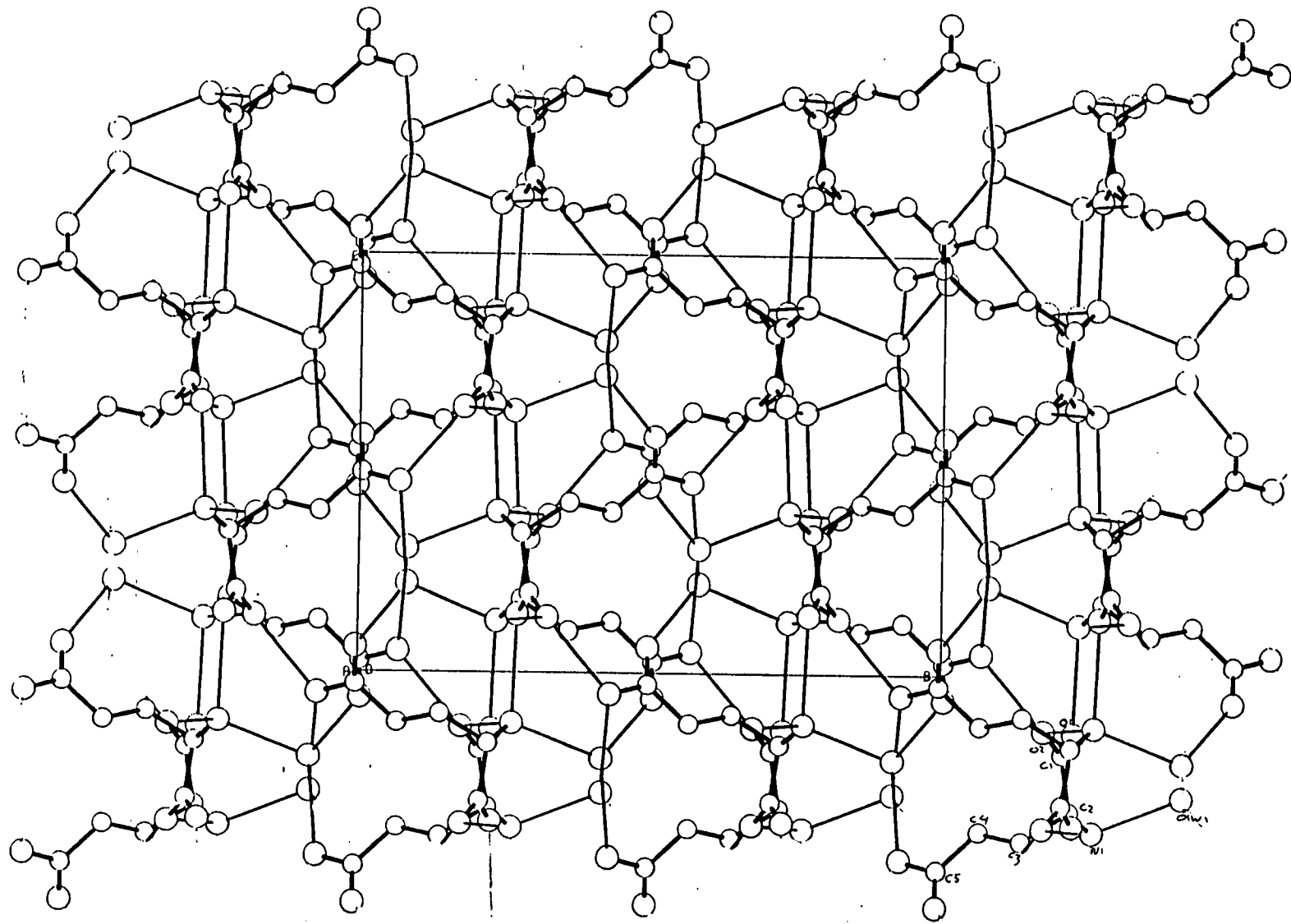
O(1) - C(1) - C(2) - C(3) 121.30(19)
 O(1) - C(1) - C(2) - N(1) -2.26(24)
 O(2) - C(1) - C(2) - C(3) -58.69(23)
 O(2) - C(1) - C(2) - N(1) 177.74(16)
 C(1) - C(2) - C(3) - C(4) -53.97(22)
 N(1) - C(2) - C(3) - C(4) 68.44(21)
 C(2) - C(3) - C(4) - C(5) -165.59(17)
 C(3) - C(4) - C(5) - O(3) -177.43(17)
 C(3) - C(4) - C(5) - O(4) 2.6(3)

C(2) -H(21) 0.979(23)
 C(3) -H(31) 0.949(24)
 C(3) -H(32) 0.967(24)
 C(4) -H(41) 0.938(24)
 C(4) -H(42) 0.949(24)
 O(3) -H(103) 0.879(24)
 N(1) -H(1N1) 0.840(25)
 N(1) -H(2N1) 0.944(24)
 N(1) -H(3N1) 0.956(24)
 O(1W1)-H(10W) 0.83(4)
 O(1W1)-H(20W) 0.78(4)

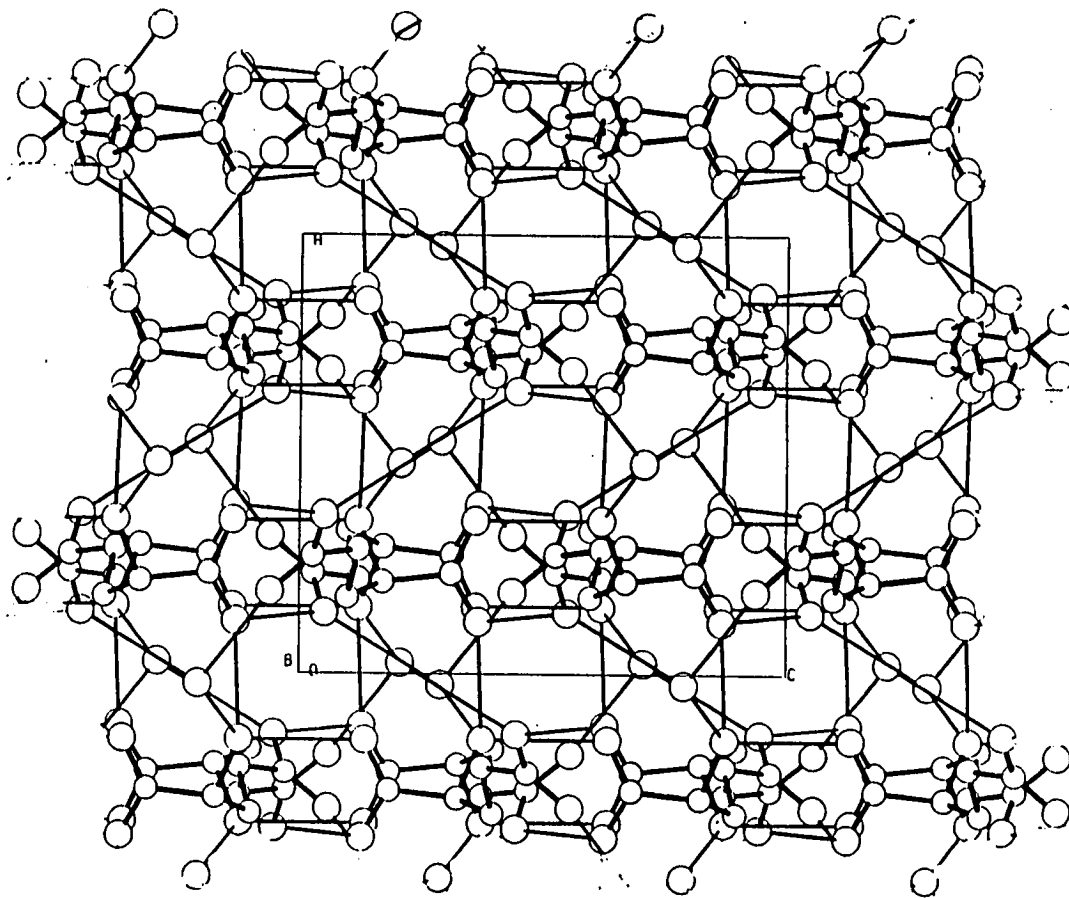
C(1) - C(2) -H(21) 106.2(14)
 C(3) - C(2) -H(21) 111.3(14)
 N(1) - C(2) -H(21) 106.2(14)
 C(2) - C(3) -H(31) 109.3(14)
 C(2) - C(3) -H(32) 106.3(14)
 C(4) - C(3) -H(31) 106.9(14)
 C(4) - C(3) -H(32) 111.2(14)
 H(31) - C(3) -H(32) 109.3(20)
 C(3) - C(4) -H(41) 110.6(14)
 C(3) - C(4) -H(42) 112.5(15)
 C(5) - C(4) -H(41) 109.1(14)
 C(5) - C(4) -H(42) 106.0(15)
 H(41) - C(4) -H(42) 104.9(21)
 C(5) - O(3) -H(103) 111.6(16)
 C(2) - N(1) -H(1N1) 108.0(17)
 C(2) - N(1) -H(2N1) 112.1(15)
 C(2) - N(1) -H(3N1) 112.7(14)
 H(1N1)- N(1) -H(2N1) 106.3(22)
 H(1N1)- N(1) -H(3N1) 110.3(22)
 H(2N1)- N(1) -H(3N1) 107.3(21)
 H(10W)-O(1W1)-H(20W) 113 (4)

O(1) - C(1) - C(2) -H(21) -116.1(14)
 O(2) - C(1) - C(2) -H(21) 63.9(14)
 C(1) - C(2) - C(3) -H(31) 65.5(15)
 C(1) - C(2) - C(3) -H(32) -176.7(15)
 N(1) - C(2) - C(3) -H(31) -172.1(15)
 N(1) - C(2) - C(3) -H(32) -54.3(15)
 H(21) - C(2) - C(3) - C(4) -173.7(15)
 H(21) - C(2) - C(3) -H(31) -54.3(21)
 H(21) - C(2) - C(3) -H(32) 63.6(21)
 C(1) - C(2) - N(1) -H(1N1) 56.1(18)
 C(1) - C(2) - N(1) -H(2N1) -60.6(16)
 C(1) - C(2) - N(1) -H(3N1) 178.2(16)
 C(3) - C(2) - N(1) -H(1N1) -69.1(18)
 C(3) - C(2) - N(1) -H(2N1) 174.2(16)
 C(3) - C(2) - N(1) -H(3N1) 53.0(16)
 H(21) - C(2) - N(1) -H(1N1) 169.9(23)
 H(21) - C(2) - N(1) -H(2N1) 53.2(21)
 H(21) - C(2) - N(1) -H(3N1) -68.0(21)
 C(2) - C(3) - C(4) -H(41) 71.6(16)
 C(2) - C(3) - C(4) -H(42) -45.4(16)
 H(31) - C(3) - C(4) - C(5) 73.7(15)
 H(31) - C(3) - C(4) -H(41) -49.2(22)
 H(31) - C(3) - C(4) -H(42) -166.1(22)
 H(32) - C(3) - C(4) - C(5) -45.6(16)
 H(32) - C(3) - C(4) -H(41) -168.5(22)
 H(32) - C(3) - C(4) -H(42) 74.6(22)
 H(41) - C(4) - C(5) - O(3) -53.8(15)
 H(41) - C(4) - C(5) - O(4) 126.2(15)
 H(42) - C(4) - C(5) - O(3) 58.7(15)
 H(42) - C(4) - C(5) - O(4) -121.2(15)
 C(4) - C(5) - O(3) -H(103) -174.0(17)
 O(4) - C(5) - O(3) -H(103) 5.9(17)

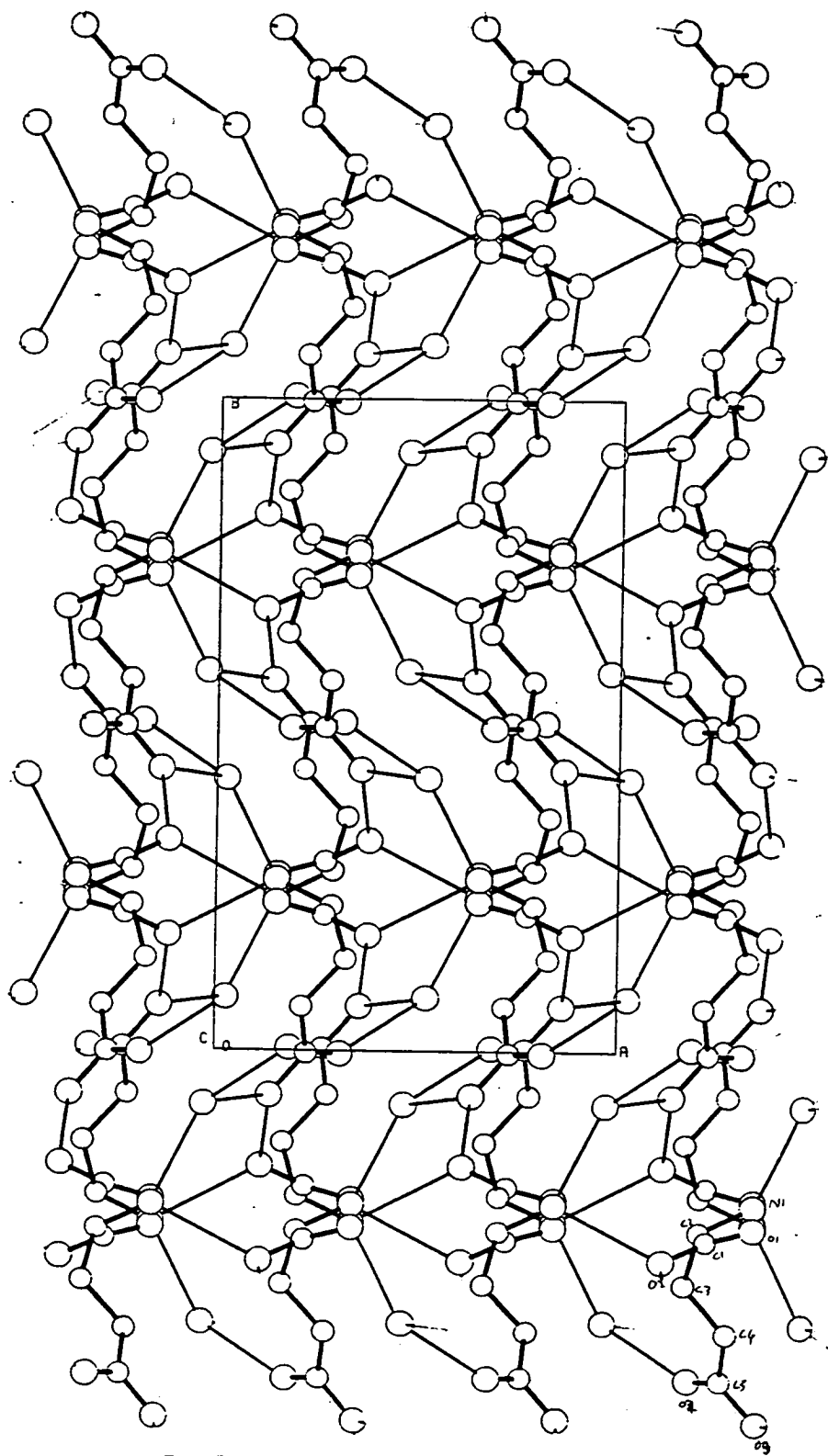
As is easily visible in the three unit cell projections overleaf, the main packing force in this structure is hydrogen bonding. Each molecule is a donor for four hydrogen bonds and an acceptor for five, with the water molecule accepting one hydrogen bond and donating protons for two (see table of distances below).



PROJECTION DOWN THE X AXIS
DL-GLUTAMIC ACID MONOHYDRATE



PROJECTION DOWN THE
Y AXIS
DL-GLUTAMIC ACID MONOHYDRATE



PROJECTION DOWN THE Z AXIS
DL-GLUTAMIC ACID MONOHYDRATE

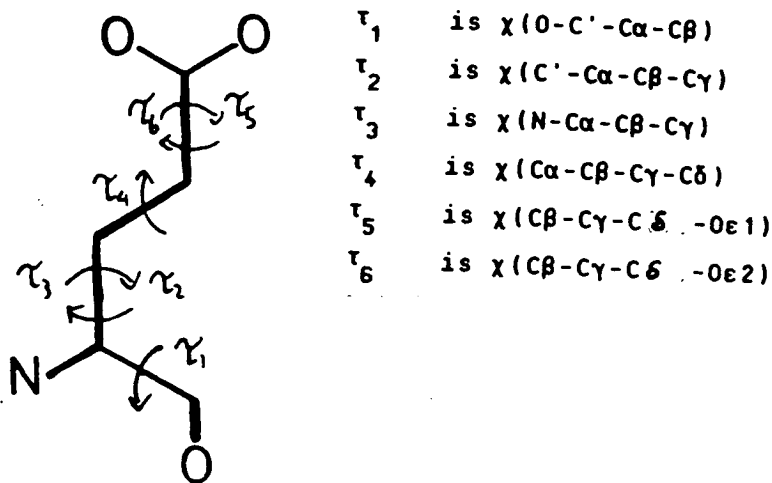
01	2.764 (2)	N1	(X, 1/2-Y, 1/2+Z)
02	2.574 (2)	O3	(1/2-X, 1 -Y, 1/2+Z)
02	2.827 (2)	N1	(1/2+X, Y, 1/2-Z)
03	2.980 (3)	O1W1	(-X, 1/2+Y, 1/2-Z)
04	2.833 (3)	O1W1	(1/2+X, 1/2-Y, -Z)
O1W1	2.817 (3)	N1	(X, Y, Z)
01	1.817 (24)	H3N1	(X, 1/2-Y, 1/2+Z)
02	2.027 (25)	H1N1	(1/2+X, Y, 1/2-Z)
02	1.703 (24)	H1O3	(1/2-X, 1 -Y, 1/2+Z)
03	2.17 (4)	H1OW	(-X, 1/2+Y, 1/2-Z)
04	2.08 (4)	H2OW	(1/2+X, 1/2-Y, -Z)
O1W1	1.918 (24)	H2N1	(X, Y, Z)

The table below shows the non-hydrogen bonding contacts.

01	3.347 (2)	C2	(-1/2+X, Y, 1/2-Z)
01	3.315 (3)	C3	(-1/2+X, Y, 1/2-Z)
02	3.336 (3)	C5	(1/2-X, 1 -Y, 1/2+Z)
02	3.300 (2)	O4	(1/2-X, 1 -Y, 1/2+Z)
03	3.437 (2)	C1	(1/2-X, 1 -Y, -1/2+Z)
03	3.309 (2)	N1	(1/2-X, 1/2+Y, Z)
03	3.497 (3)	O1W1	(1/2-X, 1/2+Y, Z)
04	3.440 (3)	O1W1	(1/2-X, 1/2+Y, Z)
O1W1	3.436 (3)	C4	(1/2-X, -1/2+Y, Z)
O1W1	3.188 (3)	C5	(1/2-X, -1/2+Y, Z)
N1	3.490 (3)	C1	(X, 1/2-Y, -1/2+Z)
01	2.697 (23)	H21	(-1/2+X, Y, 1/2-Z)
01	2.723 (24)	H31	(-1/2+X, Y, 1/2-Z)
01	2.872 (24)	H32	(X, 1/2-Y, 1/2+Z)
01	2.984 (24)	H1O3	(1/2-X, 1 -Y, 1/2+Z)
02	2.590 (24)	H42	(1/2+X, Y, 1/2-Z)
02	2.812 (24)	H3N1	(X, 1/2-Y, 1/2+Z)
03	2.825 (23)	H21	(1/2-X, 1/2+Y, Z)
03	2.979 (24)	H42	(-X, 1 -Y, -Z)
03	2.961 (24)	H2N1	(1/2-X, 1/2+Y, Z)
04	2.621 (24)	H41	(1/2-X, 1 -Y, -1/2+Z)
O1W1	2.996 (24)	H41	(1/2-X, -1/2+Y, Z)
C1	2.907 (25)	H1N1	(1/2+X, Y, 1/2-Z)
C1	2.587 (24)	H3N1	(X, 1/2-Y, 1/2+Z)
C1	2.586 (24)	H1O3	(1/2-X, 1 -Y, 1/2+Z)
H2N1	2.44 (5)	H1OW	(X, Y, Z)
H2OW	2.40 (5)	H2N1	(X, Y, Z)

Comparison with other structures

There have been crystal structure determinations carried out for several glutamic acid derivatives. The table below contains information on the conformations of 24 independent glutamate residues from twenty crystal structure determinations. As the bond lengths and angles were all within the usual ranges, only the torsion angles were tabulated.



The diagram above shows a glutamate residue with the appropriate torsion angles indicated. It is not really possible, as it is for the cystine amino acids, to divide the structures into two classes on the basis of their side chain conformation. This implies that several possible conformations of comparable energy may exist and that the actual conformation in any given crystal structure depends on the molecular environment. The only molecules of similar conformation to DL-glutamic acid in the table below are leucyl glutamic acid, glutamyl glycine and the N-terminal residue of glutamyl glutamic acid. Neither the α nor the β modifications of L-glutamic acid are of similar conformation to the DL acid, which is as expected because in the case of the racemic acid, each molecule is surrounded by molecules of the opposite hand to those in the optically active compound.

CODE	τ_1	τ_2	τ_3	τ_4	τ_5	τ_6
BOFZOL	106.31 -70.80	-63.29	59.05	171.15	-139.23	37.55
MGHPHE20	-80.47	168.17	-65.11	-56.54	-30.71	152.62
AMBZGU10	147.41 -33.09	-176.53	-54.22	74.29	40.94	-142.10
ANCBGL	121.55 -56.70	-174.86	-61.70	175.92	12.85	-166.87
ARGGLU10	-70.99 107.59	-179.34	-56.83	-172.05	-2.58	-179.91
BALGLH	-78.91 99.49	179.38	59.25	62.26	-155.67	25.14
BCYTGA	88.52 -92.95 76.39 -103.89	-176.80	-51.88	-164.04	176.18	-3.16
BELCUQ	93.70 88.19 -91.53	-58.14 179.35	60.31 -57.83	-174.55 -131.06	174.20 36.02	-4.20 -145.50
BRCPDG	135.03 -44.65 -12.27 172.53	90.60 173.60	-144.72 -61.31	68.91 -65.44	-162.45 154.30	16.88 -25.39
BZGEGA	-9.47 158.65 60.24	170.52 56.48	-65.56 -178.80	171.89 -167.66	-171.81 6.31	4.45 -171.21
CAGLCL10	89.58 -89.43	67.33	-173.03	170.16	12.28	-170.84
CAGLUT10	101.86 -76.18	58.15	179.40	174.72	12.30	-168.94

CTBGLU	76.53	-162.11	77.07	81.33	-6.89	177.65
	-106.91					
CTSGLM	-143.94	-179.15	52.82	-176.18	-172.25	10.43
	39.41					
DLGLAC	77.01	-169.08	69.59	171.78	-17.65	164.12
	-103.37					
GLUGLY	108.45	-47.76	72.10	-166.70	24.33	-157.62
LGLUAC03	68.90	59.22	178.19	68.30	74.21	-104.65
	-110.56					
LGLUAC11	78.72	-171.06	-51.79	-73.10	18.80	-160.70
	-97.99					
LGLUCA	105.58	-179.57	-56.00	-160.72	15.64	-166.40
	-72.59					
LGLUTA	-76.31	171.06	-68.91	-173.13	14.95	-167.00
	101.30					
LGPYRG	-91.01	-175.03	-55.90	-65.20	-173.29	3.79
	86.05					

Key to compounds in the above table

BOFZOL α -L-leucyl-L-glutamic acid

D. S. Eggleston, D. J. Hodgson, Acta Cryst., C39, 75, 1983

MGHPHE20 L-methionyl-L- α -glutamyl-L-histidyl-L-phenylalanine monohydrate

G. Admiraal, A. Vos, Acta Cryst., C39, 82, 1983

AMBZGU10 N-(p-aminobenzoyl)-L-glutamic acid hydrochloride

C. Chatterjee, J. K. Dattagupta, N. N. Saha, Acta Cryst., B38, 2086, 1982

- ANCBGL N-carboxy- η -benzyl-L-glutamate anhydride
- H. Kanazawa, T. Kawai, Y. Ohashi, Y. Sasada, Bull. Chem. Soc. Jpn., 51, 2200, 1978
- ARGGLU10 L-arginine L-glutamate monohydrate
- T. N. Bhat, M. Vijayan, Acta Cryst., B33, 1754, 1977
- BALGLH T-butoxycarbonyl-D-alanyl-D-glutamic acid monohydrate
- O. Dideberg, J. Lamotte, L. Dupont, L. Christiaens, Acta Cryst., B37, 1150, 1981
- BCYTGA 5-bromocytosine N-tosyl-L-glutamic acid
- M. Ohki, A. Takenaka, H. Shimanouchi, Y. Sasada, Bull. Chem. Soc. Jpn., 49, 3493, 1976
- BELCUQ α -L-glutamyl-L-glutamic acid
- D. S. Eggleston, D. J. Hodgson, Acta Cryst., B38, 1216, 1982
- BRCPDG 5-bromocytosine-phthaloyl-DL-glutamic acid hemihydrate
- M. Ohki, A. Takenaka, H. Shimanouchi, Y. Sasada, Bull. Chem. Soc. Jpn., 50, 90, 1977
- BZGEGA N-benzyloxycarbonyl-(η -ethyl)-L-glutamyl-(η -ethyl)-L-glutamic acid ethyl ester
- E. Benedetti, B. Di Blasio, V. Pavone, C. Pedone, G. Germain, M. Goodman, Biopolymers, 18, 517, 1979
- CAGLCL10 Calcium L-glutamate chloride monohydrate
- H. Einspahr, G. L. Gartland, C. E. Bugg, Acta Cryst., B33, 3385, 1977

CAGLUT10 Calcium L-glutamate trihydrate

H. Einspahr, C. E. Bugg, Acta Cryst., B30, 1037, 1974

CTBGLU cytidine - N-carbenzoxyglutamic acid dihydrate

T. Hata, M. Yoshikawa, S. Sato, C. Tamura, Acta Cryst., B31, 312, 1975

CTSGLM cytosine N, N-phthaloyl-DL-glutamic acid dihydrate

A. Takenaka, M. Ohki, Y. Sasada, Bull. Chem. Soc. Jpn., 53, 2724, 1980

DLGLAC DL-glutamic acid hydrochloride

B. Dawson, Acta Cryst., 6, 81, 1953

GLUGLY α -L-glutamyl-glycine

D. S. Eggleston, E. J. Valente, D. J. Hodgson, Acta Cryst., B37, 1430, 1981

LGLUAC03 L-glutamic acid, α -form
(Neutron study)

M. S. Lehmann, A. C. Nunes, Acta Cryst., B36, 1621, 1980

LGLUAC11 L-glutamic acid, β -form
(Neutron study)

M. S. Lehmann, T. F. Koetzle, W. C. Hamilton, J. Cryst. Mol. Struct., 2, 225, 1972

LGLUCA Calcium di-L-glutamate tetrahydrate

H. Einspahr, C. E. Bugg, Acta Cryst., B35, 316, 1979

LGLUTA L-glutamic acid hydrochloride
(Neutron study)

A. Sequeira, H. Rajagopal, R. Chidambaram, Acta Cryst., B28, 2514, 1972

LGPYRG

L-glutamic acid L-pyroglutamic acid monohydrate

Z. Taira, W. H. Watson, Acta Cryst., B33, 3823, 1977

Table Fractional Coordinates of Atoms with Standard Deviations

	x	y	z	Ueq
C(1)	0.27108(21)	0.29127(12)	0.31289(18)	0.0218
C(2)	0.28919(20)	0.27690(12)	0.17183(18)	0.0201
C(3)	0.32520(22)	0.35903(13)	0.09922(20)	0.0241
C(4)	0.21697(23)	0.43228(13)	0.12160(21)	0.0272
C(5)	0.23023(23)	0.50375(14)	0.02771(18)	0.0266
O(1)	0.15284(15)	0.27155(11)	0.36258(13)	0.0336
O(2)	0.38100(15)	0.32293(10)	0.36890(14)	0.0292
O(3)	0.13978(17)	0.56863(10)	0.05101(15)	0.0333
O(4)	0.31303(22)	0.50443(12)	-0.06080(17)	0.0515
N(1)	0.15257(18)	0.23623(11)	0.12208(16)	0.0216
O(1W1)	0.02525(24)	0.08193(14)	0.20825(19)	0.0536
H(21)	0.367(3)	0.2338(15)	0.1624(22)	0.0337
H(31)	0.419(3)	0.3794(15)	0.1244(22)	0.0337
H(32)	0.328(3)	0.3430(15)	0.0114(23)	0.0337
H(41)	0.228(3)	0.4549(15)	0.2027(22)	0.0337
H(42)	0.118(3)	0.4136(15)	0.1179(22)	0.0337
H(1N1)	0.082(3)	0.2683(16)	0.1409(23)	0.0337
H(2N1)	0.134(3)	0.1824(16)	0.1600(22)	0.0337
H(3N1)	0.156(3)	0.2271(15)	0.0333(23)	0.0337
H(1O3)	0.142(3)	0.6075(16)	-0.0091(23)	0.0337
H(1OW)	-0.007(4)	0.077(3)	0.280(4)	0.0903
H(2OW)	-0.020(4)	0.057(3)	0.158(4)	0.0903

Thermal Vibration Parameters with Standard Deviations

	U11	U22	U33	U23	U13	U12
C(1)	0.0223(9)	0.0229(9)	0.0201(9)	0.0008(8)	-0.0018(8)	0.0062(8)
C(2)	0.0170(9)	0.0242(9)	0.0193(9)	-0.0011(8)	0.0001(7)	0.0017(7)
C(3)	0.0216(10)	0.0259(10)	0.0248(10)	0.0026(8)	0.0031(8)	-0.0010(8)
C(4)	0.0266(11)	0.0269(11)	0.0279(10)	0.0054(9)	0.0056(9)	0.0018(8)
C(5)	0.0262(10)	0.0279(10)	0.0258(10)	0.0034(8)	-0.0014(8)	0.0022(8)
O(1)	0.0233(7)	0.0589(11)	0.0187(7)	0.0009(7)	0.0020(6)	-0.0042(7)
O(2)	0.0209(7)	0.0367(8)	0.0300(7)	-0.0109(7)	-0.0062(6)	0.0039(6)
O(3)	0.0376(9)	0.0303(8)	0.0319(8)	0.0102(7)	0.0071(7)	0.0102(7)
O(4)	0.0609(12)	0.0515(11)	0.0421(10)	0.0191(9)	0.0266(9)	0.0214(9)
N(1)	0.0202(8)	0.0256(9)	0.0189(8)	-0.0007(7)	-0.0001(7)	-0.0009(7)
O(1W1)	0.0625(13)	0.0576(12)	0.0406(11)	-0.0005(9)	0.0035(10)	-0.0302(11)

CHAPTER SEVEN
STRYCHNINE

STRYCHNINE

Introduction

The indole alkaloid strychnine is obtained from the seeds of *strychnos nux vomica*, and in smaller yields from related species. Although it is no longer considered to be of therapeutic value (it was once used as a "tonic" stimulant and as a laxative), its mechanism of action is better understood than that of almost any other stimulant compound.

The main effect of strychnine on the central nervous system consists of the stimulation, followed by the depression, of reflexes. The drug increases the motor effects of spinal reflexes and diminishes the latent period. This leads to reflexes becoming more general and, after large doses, small sensory stimuli will send all the voluntary muscles in the body into violent convulsions. It is known that strychnine does not excite directly, but inhibits the inhibition of spinal motor neurone signals by glycine. Strychnine has an independent excitatory action on the medulla oblongata, where it stimulates the vasomotor and vagal centres. It also has an enhancing effect on all five senses, although its bitter taste tends to mask the enhancement of flavour. This same bitter taste increases the appetite. A high dose of strychnine causes restlessness and small stimuli will cause the limbs to jerk. This leads to generalised convulsions with the trunk and limbs rigidly extended and the face distorted into a rictus. Each convulsion is very painful and lasts a few minutes. After five or six convulsions, each of which is followed by a longer period of exhaustion, respiration fails to return and death by asphyxia results. It is possible to counteract the effect of strychnine after absorption and stop the convulsions by reducing sensory stimulation to a minimum and administering a general anaesthetic. Barbiturates or mephenesin (a muscle relaxant and interneuronal blocker) may be used as an antidote.

Although crystals of strychnine were examined at Edinburgh as long ago as 1944, the crystal structure was not determined as it was too complex a problem for the methods available at the time. After the

publication of the structures of strychnine hydrobromide (Robertson and Beevers, 1951) and strychnine sulphate (Bokhoven, Schoone and Bijvoet, 1951) interest in the pure base waned, although the crystal structures of some other strychnine salts have been determined.

Crystallisation

The crystals of strychnine used in this determination were grown from a solution of equimolar quantities of strychnine and N-acetyl-L-phenylalanine dissolved in the minimum of hot water and allowed to cool slowly to room temperature. In general, strychnine crystals were found in most attempted crystallisations of molecular complexes of strychnine with N-acyl amino acids. This could be because not all of the strychnine was in solution as strychninium cations.

Crystal Data

Formula $C_{21}H_{22}N_2O_2$ Mol.Wt. 334.41 $F(000)=712$

Space group $P2_12_12_1$ Int. Tab. No. 19

Cell dimensions $a = 11.267(2)$ $b = 11.89(1)$ $c = 12.105(4)$

$V = 1621.9 \text{ \AA}^3$ $Z = 4$ $D_c = 1.369 \text{ gcm}^{-3}$

Radiation $\text{MoK}\alpha$ $\lambda = 0.71069 \text{ \AA}$ $\mu = 0.83 \text{ cm}^{-1}$

Final $R = 0.0370$ based on 1195 independent data

$h_{\text{max}} = 13$ $k_{\text{max}} = 14$ $l_{\text{max}} = 14$

$\theta_{\text{max}} = 25^\circ$ $\sin\theta_{\text{max}}/\lambda = 0.5947$

Intensity control $-1 -1 4$

$0.988 < \text{drift curve} < 1.016$

1648 reflections measured

453 reflections considered unobserved ($F < 4\sigma(F)$)

Weighting scheme : $W = 0.1625/(\sigma^2(F)+0.006531(F^2))$

314 parameters refined

Final difference Fourier max. $0.1699 \text{ e}\text{\AA}^{-3}$ min. $-0.2131 \text{ e}\text{\AA}^{-3}$

Max. Shift/ σ in final cycle 0.005

Space group determination

The space group was uniquely determined to be $P2_12_12_1$ by the systematic absences ($h00$, h odd; $0k0$, k odd; $00l$, l odd).

Data collection and reduction

The data was collected on the CAD4 and corrected for Lorentz and polarisation effects by the program CADABS.

Solution and Refinement

The structure was solved automatically with the MULTAN 77 system. The first electron density map output revealed the entire structure except for one carbon atom.

The program SHELX was used for least-squares refinement, the progress of which is summarised below:

R	Model
0.2598	All non H but for one carbon
0.2385	All non H atoms included
0.0941	14 hydrogen added in calculated positions with free isotropic thermal parameters
0.0819	Another 7 hydrogen atoms added
0.0476	All hydrogen atoms included in model and all non-hydrogen atoms with anisotropic thermal parameters
0.0411	Weighting scheme applied
0.0370	Final convergence*

Description of structure

It could be claimed that strychnine has a more complex structure than any other compound with a similar number of skeletal atoms. This results in a very rigid structure with a greater than usual deviation of skeletal bond lengths and angles from the ideal values. There is a difference of 12 e.s.d.'s between the longest and the shortest carbon carbon bonds,^y and a difference of 61 e.s.d.'s between the largest and smallest angles at "tetrahedral" carbon.

C(7)-C(8) 1.556(5) A C(17)-C(18) 1.497(5) A

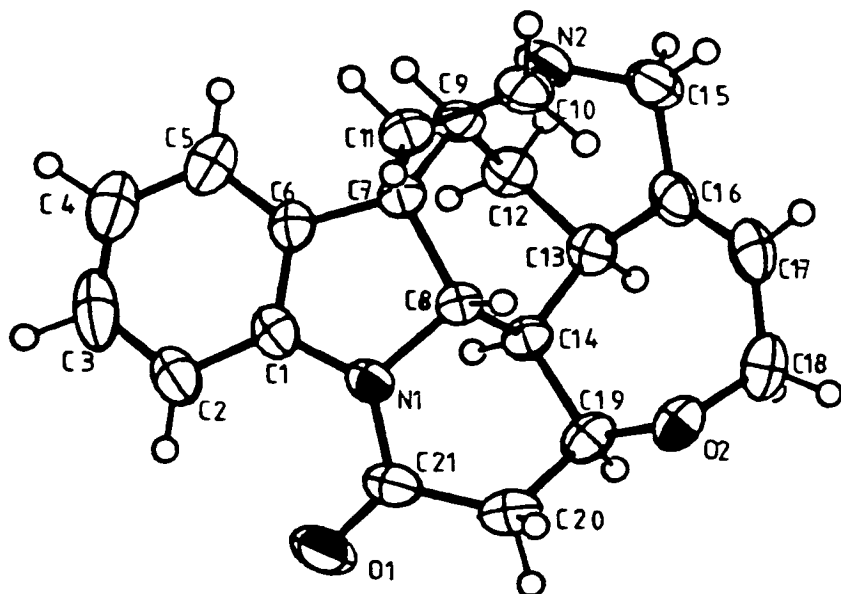
mean C-C 1.523(19) A

* No constraints were applied to the hydrogen positions or thermal parameters

* C(17)-C(18) is so short because it is a bond between an sp^2 and an sp^3 carbon atom.

C(9)-C(7)-C(11) 100.6(3)° C(13)-C(14)-C(19) 118.9(3)°

mean C-C-C 111(5)°



The tables below list the bonds, angles and torsion angles of strychnine, using the numbering system illustrated above.

C(1) - C(2)	1.395(6)	C(10) - N(2)	1.471(6)
C(1) - C(6)	1.386(5)	C(12) - C(13)	1.522(6)
C(1) - N(1)	1.408(5)	C(13) - C(14)	1.530(5)
C(2) - C(3)	1.396(8)	C(13) - C(16)	1.520(6)
C(3) - C(4)	1.375(8)	C(14) - C(19)	1.543(5)
C(4) - C(5)	1.380(7)	C(15) - C(16)	1.506(6)
C(5) - C(6)	1.386(6)	C(15) - N(2)	1.464(6)
C(6) - C(7)	1.498(5)	C(16) - C(17)	1.321(6)
C(7) - C(8)	1.556(5)	C(17) - C(18)	1.497(7)
C(7) - C(9)	1.542(5)	C(18) - O(2)	1.435(6)
C(7) - C(11)	1.545(5)	C(19) - C(20)	1.528(6)
C(8) - C(14)	1.531(5)	C(19) - O(2)	1.425(5)
C(8) - N(1)	1.489(5)	C(20) - C(21)	1.498(6)
C(9) - C(12)	1.516(6)	C(21) - N(1)	1.371(5)
C(9) - N(2)	1.493(5)	C(21) - O(1)	1.221(6)
C(10) - C(11)	1.515(6)		

C(2) - C(1) - C(6)	121.1(4)	C(12) -C(13) -C(16)	109.4(3)
C(2) - C(1) - N(1)	128.8(4)	C(14) -C(13) -C(16)	114.8(3)
C(6) - C(1) - N(1)	110.1(3)	C(8) -C(14) -C(13)	112.8(3)
C(1) - C(2) - C(3)	117.1(4)	C(8) -C(14) -C(19)	107.1(3)
C(2) - C(3) - C(4)	122.4(5)	C(13) -C(14) -C(19)	118.9(3)
C(3) - C(4) - C(5)	119.4(5)	C(16) -C(15) - N(2)	111.6(4)
C(4) - C(5) - C(6)	119.9(4)	C(13) -C(16) -C(15)	113.9(4)
C(1) - C(6) - C(5)	120.0(4)	C(13) -C(16) -C(17)	122.3(4)
C(1) - C(6) - C(7)	110.7(3)	C(15) -C(16) -C(17)	123.8(4)
C(5) - C(6) - C(7)	129.1(4)	C(16) -C(17) -C(18)	122.7(4)
C(6) - C(7) - C(8)	102.6(3)	C(17) -C(18) - O(2)	111.8(4)
C(6) - C(7) - C(9)	116.6(3)	C(14) -C(19) -C(20)	109.9(3)
C(6) - C(7) -C(11)	111.9(3)	C(14) -C(19) - O(2)	114.7(3)
C(8) - C(7) - C(9)	114.2(3)	C(20) -C(19) - O(2)	105.5(3)
C(8) - C(7) -C(11)	111.2(3)	C(19) -C(20) -C(21)	117.3(4)
C(9) - C(7) -C(11)	100.6(3)	C(20) -C(21) - N(1)	115.6(4)
C(7) - C(8) -C(14)	116.8(3)	C(20) -C(21) - O(1)	122.6(4)
C(7) - C(8) - N(1)	104.3(3)	N(1) -C(21) - O(1)	121.8(4)
C(14) - C(8) - N(1)	106.5(3)	C(1) - N(1) - C(8)	109.3(3)
C(7) - C(9) -C(12)	114.2(3)	C(1) - N(1) -C(21)	125.2(3)
C(7) - C(9) - N(2)	105.3(3)	C(8) - N(1) -C(21)	118.1(3)
C(12) - C(9) - N(2)	111.3(3)	C(9) - N(2) -C(10)	108.1(3)
C(11) -C(10) - N(2)	105.2(3)	C(9) - N(2) -C(15)	113.5(3)
C(7) -C(11) -C(10)	101.6(3)	C(10) - N(2) -C(15)	113.4(3)
C(9) -C(12) -C(13)	108.9(3)	C(18) - O(2) -C(19)	114.6(3)
C(12) -C(13) -C(14)	106.5(3)		

C(6) - C(1) - C(2) - C(3)	1.5(6)
N(1) - C(1) - C(2) - C(3)	-178.5(4)
C(2) - C(1) - C(6) - C(5)	-1.7(6)
C(2) - C(1) - C(6) - C(7)	174.0(4)
N(1) - C(1) - C(6) - C(5)	178.3(4)
N(1) - C(1) - C(6) - C(7)	-6.0(4)
C(2) - C(1) - N(1) - C(8)	174.2(4)
C(2) - C(1) - N(1) -C(21)	25.0(6)
C(6) - C(1) - N(1) - C(8)	-5.8(4)
C(6) - C(1) - N(1) -C(21)	-155.1(4)
C(1) - C(2) - C(3) - C(4)	0.1(8)
C(2) - C(3) - C(4) - C(5)	-1.5(8)
C(3) - C(4) - C(5) - C(6)	1.3(8)
C(4) - C(5) - C(6) - C(1)	0.2(7)
C(4) - C(5) - C(6) - C(7)	-174.6(4)
C(1) - C(6) - C(7) - C(8)	14.5(4)

C(1) - C(6) - C(7) - C(9) 140.1(3)
C(1) - C(6) - C(7) -C(11) -104.9(4)
C(5) - C(6) - C(7) - C(8) -170.4(4)
C(5) - C(6) - C(7) - C(9) -44.7(6)
C(5) - C(6) - C(7) -C(11) 70.3(5)
C(6) - C(7) - C(8) -C(14) 100.3(3)
C(6) - C(7) - C(8) - N(1) -16.9(3)
C(9) - C(7) - C(8) -C(14) -26.9(4)
C(9) - C(7) - C(8) - N(1) -144.1(3)
C(11) - C(7) - C(8) -C(14) -139.9(3)
C(11) - C(7) - C(8) - N(1) 102.9(3)
C(6) - C(7) - C(9) -C(12) -83.6(4)
C(6) - C(7) - C(9) - N(2) 154.0(3)
C(8) - C(7) - C(9) -C(12) 36.0(4)
C(8) - C(7) - C(9) - N(2) -86.5(4)
C(11) - C(7) - C(9) -C(12) 155.2(3)
C(11) - C(7) - C(9) - N(2) 32.8(4)
C(6) - C(7) -C(11) -C(10) -167.5(3)
C(8) - C(7) -C(11) -C(10) 78.4(4)
C(9) - C(7) -C(11) -C(10) -43.0(4)
C(7) - C(8) -C(14) -C(13) 40.1(4)
C(7) - C(8) -C(14) -C(19) 172.8(3)
N(1) - C(8) -C(14) -C(13) 156.0(3)
N(1) - C(8) -C(14) -C(19) -71.2(3)
C(7) - C(8) - N(1) - C(1) 14.5(4)
C(7) - C(8) - N(1) -C(21) 166.2(3)
C(14) - C(8) - N(1) - C(1) -109.6(3)
C(14) - C(8) - N(1) -C(21) 42.1(4)
C(7) - C(9) -C(12) -C(13) -58.3(4)
N(2) - C(9) -C(12) -C(13) 60.8(4)
C(7) - C(9) - N(2) -C(10) -9.9(4)
C(7) - C(9) - N(2) -C(15) 116.7(4)
C(12) - C(9) - N(2) -C(10) -134.2(3)
C(12) - C(9) - N(2) -C(15) -7.5(5)
N(2) -C(10) -C(11) - C(7) 38.2(4)
C(11) -C(10) - N(2) - C(9) -17.8(4)
C(11) -C(10) - N(2) -C(15) -144.6(4)
C(9) -C(12) -C(13) -C(14) 69.2(4)
C(9) -C(12) -C(13) -C(16) -55.4(4)
C(12) -C(13) -C(14) - C(8) -60.3(4)
C(12) -C(13) -C(14) -C(19) 173.1(3)
C(16) -C(13) -C(14) - C(8) 61.0(4)
C(16) -C(13) -C(14) -C(19) -65.7(5)
C(12) -C(13) -C(16) -C(15) 0.4(5)
C(12) -C(13) -C(16) -C(17) 177.8(4)

C(14) -C(13) -C(16) -C(15) -119.2(4)
 C(14) -C(13) -C(16) -C(17) 58.2(6)
 C(8) -C(14) -C(19) -C(20) 43.8(4)
 C(8) -C(14) -C(19) - O(2) -74.9(4)
 C(13) -C(14) -C(19) -C(20) 173.1(3)
 C(13) -C(14) -C(19) - O(2) 54.4(5)
 N(2) -C(15) -C(16) -C(13) 52.5(5)
 N(2) -C(15) -C(16) -C(17) -124.8(5)
 C(16) -C(15) - N(2) - C(9) -48.2(5)
 C(16) -C(15) - N(2) -C(10) 75.7(5)
 C(13) -C(16) -C(17) -C(18) -2.5(7)
 C(15) -C(16) -C(17) -C(18) 174.7(4)
 C(16) -C(17) -C(18) - O(2) -65.2(6)
 C(17) -C(18) - O(2) -C(19) 88.2(4)
 C(14) -C(19) -C(20) -C(21) 10.1(5)
 O(2) -C(19) -C(20) -C(21) 134.3(4)
 C(14) -C(19) - O(2) -C(18) -66.1(4)
 C(20) -C(19) - O(2) -C(18) 172.8(3)
 C(19) -C(20) -C(21) - N(1) -41.7(5)
 C(19) -C(20) -C(21) - O(1) 139.6(4)
 C(20) -C(21) - N(1) - C(1) 159.9(4)
 C(20) -C(21) - N(1) - C(8) 13.1(5)
 O(1) -C(21) - N(1) - C(1) -21.3(6)
 O(1) -C(21) - N(1) - C(8) -168.1(4)

The bonds, angles and torsion angles involving hydrogen are tabulated below.

C(2) -H(021)	0.97(4)	C(12) -H(122)	1.03(5)
C(3) -H(031)	0.93(6)	C(13) -H(131)	0.99(5)
C(4) -H(041)	1.05(6)	C(14) -H(141)	1.03(4)
C(5) -H(051)	0.98(5)	C(15) -H(151)	1.09(6)
C(8) -H(081)	0.98(4)	C(15) -H(152)	1.06(5)
C(9) -H(091)	1.01(3)	C(17) -H(171)	0.99(5)
C(10) -H(101)	1.04(5)	C(18) -H(181)	0.96(6)
C(10) -H(102)	1.00(5)	C(18) -H(182)	1.04(5)
C(11) -H(111)	0.96(4)	C(19) -H(191)	0.95(4)
C(11) -H(112)	0.92(4)	C(20) -H(201)	0.98(7)
C(12) -H(121)	0.93(5)	C(20) -H(202)	1.00(5)

C(1) - C(2) -H(021)	120.4(25)	H(122)-C(12) -C(13)	109.8(26)
H(021)- C(2) - C(3)	122.5(25)	C(12) -C(13) -H(131)	115 (3)
C(2) - C(3) -H(031)	117 (4)	H(131)-C(13) -C(14)	108 (3)
H(031)- C(3) - C(4)	120 (4)	H(131)-C(13) -C(16)	103 (3)
C(3) - C(4) -H(041)	117 (4)	C(8) -C(14) -H(141)	105.7(25)
H(041)- C(4) - C(5)	123 (3)	C(13) -C(14) -H(141)	100.2(25)
C(4) - C(5) -H(051)	118 (3)	H(141)-C(14) -C(19)	111.4(25)
H(051)- C(5) - C(6)	122 (3)	H(151)-C(15) -H(152)	110 (4)
C(7) - C(8) -H(081)	107.2(23)	H(151)-C(15) -C(16)	109 (3)
H(081)- C(8) -C(14)	112.2(23)	H(151)-C(15) - N(2)	109 (3)
H(081)- C(8) - N(1)	109.4(23)	H(152)-C(15) -C(16)	112.2(27)
C(7) - C(9) -H(091)	108.1(20)	H(152)-C(15) - N(2)	105.7(27)
H(091)- C(9) -C(12)	109.2(20)	C(16) -C(17) -H(171)	118 (3)
H(091)- C(9) - N(2)	108.5(20)	H(171)-C(17) -C(18)	119 (3)
H(101)-C(10) -H(102)	108 (4)	C(17) -C(18) -H(181)	113 (4)
H(101)-C(10) -C(11)	113.7(26)	C(17) -C(18) -H(182)	113.7(27)
H(101)-C(10) - N(2)	109.5(26)	H(181)-C(18) -H(182)	98 (4)
H(102)-C(10) -C(11)	107.2(27)	H(181)-C(18) - O(2)	109 (4)
H(102)-C(10) - N(2)	113.1(27)	H(182)-C(18) - O(2)	111.0(27)
C(7) -C(11) -H(111)	107.1(26)	C(14) -C(19) -H(191)	108.8(24)
C(7) -C(11) -H(112)	112.4(24)	H(191)-C(19) -C(20)	111.0(24)
C(10) -C(11) -H(111)	114.6(26)	H(191)-C(19) - O(2)	106.9(24)
C(10) -C(11) -H(112)	112.4(24)	C(19) -C(20) -H(201)	111 (4)
H(111)-C(11) -H(112)	109 (4)	C(19) -C(20) -H(202)	106 (3)
C(9) -C(12) -H(121)	100 (3)	H(201)-C(20) -H(202)	92(5)
C(9) -C(12) -H(122)	111.7(26)	H(201)-C(20) -C(21)	109 (4)
H(121)-C(12) -H(122)	111 (4)	H(202)-C(20) -C(21)	118 (3)
H(121)-C(12) -C(13)	116 (3)		

C(6) - C(1) - C(2) -H(021)	-180 (3)
N(1) - C(1) - C(2) -H(021)	1 (3)
C(1) - C(2) - C(3) -H(031)	-170 (4)
H(021)- C(2) - C(3) -H(031)	11 (5)
H(021)- C(2) - C(3) - C(4)	-179 (3)
C(2) - C(3) - C(4) -H(041)	177 (4)
H(031)- C(3) - C(4) -H(041)	-13 (6)
H(031)- C(3) - C(4) - C(5)	168 (4)
C(3) - C(4) - C(5) -H(051)	-179 (3)
H(041)- C(4) - C(5) -H(051)	3 (5)
H(041)- C(4) - C(5) - C(6)	-177 (4)
H(051)- C(5) - C(6) - C(1)	-180 (3)
H(051)- C(5) - C(6) - C(7)	5 (3)
C(6) - C(7) - C(8) -H(081)	-132.8(24)
C(9) - C(7) - C(8) -H(081)	100.0(24)

14

C(11) - C(7) - C(8) -H(081)	-13.0(24)
C(6) - C(7) - C(9) -H(091)	38.1(21)
C(8) - C(7) - C(9) -H(091)	157.7(21)
C(11) - C(7) - C(9) -H(091)	-83.1(21)
C(6) - C(7) -C(11) -H(111)	-47 (3)
C(6) - C(7) -C(11) -H(112)	72.1(26)
C(8) - C(7) -C(11) -H(111)	-161.1(27)
C(8) - C(7) -C(11) -H(112)	-42.0(26)
C(9) - C(7) -C(11) -H(111)	77.5(27)
C(9) - C(7) -C(11) -H(112)	-163.4(26)
C(7) - C(8) -C(14) -H(141)	-68.4(26)
H(081)- C(8) -C(14) -C(13)	-84.3(25)
H(081)- C(8) -C(14) -H(141)	167 (4)
H(081)- C(8) -C(14) -C(19)	48.5(25)
N(1) - C(8) -C(14) -H(141)	47.6(26)
H(081)- C(8) - N(1) - C(1)	128.9(25)
H(081)- C(8) - N(1) -C(21)	-79.4(25)
C(7) - C(9) -C(12) -H(121)	63 (3)
C(7) - C(9) -C(12) -H(122)	-180 (3)
H(091)- C(9) -C(12) -H(121)	-58 (4)
H(091)- C(9) -C(12) -H(122)	59 (4)
H(091)- C(9) -C(12) -C(13)	-179.4(21)
N(2) - C(9) -C(12) -H(121)	-178 (3)
N(2) - C(9) -C(12) -H(122)	-61 (3)
H(091)- C(9) - N(2) -C(10)	105.6(21)
H(091)- C(9) - N(2) -C(15)	-127.7(21)
H(101)-C(10) -C(11) - C(7)	-82 (3)
H(101)-C(10) -C(11) -H(111)	163 (4)
H(101)-C(10) -C(11) -H(112)	39 (4)
H(102)-C(10) -C(11) - C(7)	159 (3)
H(102)-C(10) -C(11) -H(111)	44 (4)
H(102)-C(10) -C(11) -H(112)	-81 (4)
N(2) -C(10) -C(11) -H(111)	-77 (3)
N(2) -C(10) -C(11) -H(112)	158.6(26)
H(101)-C(10) - N(2) - C(9)	104.8(27)
H(101)-C(10) - N(2) -C(15)	-22.0(27)
H(102)-C(10) - N(2) - C(9)	-135 (3)
H(102)-C(10) - N(2) -C(15)	99 (3)
C(9) -C(12) -C(13) -H(131)	-171 (3)
H(121)-C(12) -C(13) -H(131)	78 (5)
H(121)-C(12) -C(13) -C(14)	-42 (4)
H(121)-C(12) -C(13) -C(16)	-167 (4)
H(122)-C(12) -C(13) -H(131)	-49 (4)
H(122)-C(12) -C(13) -C(14)	-168.2(27)
H(122)-C(12) -C(13) -C(16)	67.1(27)

C(12) -C(13) -C(14) -H(141)	51.6(26)
H(131)-C(13) -C(14) - C(8)	176 (3)
H(131)-C(13) -C(14) -H(141)	-72 (4)
H(131)-C(13) -C(14) -C(19)	49 (3)
C(16) -C(13) -C(14) -H(141)	172.9(26)
H(131)-C(13) -C(16) -C(15)	123 (3)
H(131)-C(13) -C(16) -C(17)	-59 (3)
C(8) -C(14) -C(19) -H(191)	165.6(25)
C(13) -C(14) -C(19) -H(191)	-65.2(26)
H(141)-C(14) -C(19) -H(191)	51 (4)
H(141)-C(14) -C(19) -C(20)	-71.3(27)
H(141)-C(14) -C(19) - O(2)	170.1(27)
H(151)-C(15) -C(16) -C(13)	-67 (3)
H(151)-C(15) -C(16) -C(17)	116 (3)
H(152)-C(15) -C(16) -C(13)	171 (3)
H(152)-C(15) -C(16) -C(17)	-6 (3)
H(151)-C(15) - N(2) - C(9)	72 (3)
H(151)-C(15) - N(2) -C(10)	-165 (3)
H(152)-C(15) - N(2) - C(9)	-171 (3)
H(152)-C(15) - N(2) -C(10)	-47 (3)
C(13) -C(16) -C(17) -H(171)	-177 (3)
C(15) -C(16) -C(17) -H(171)	0 (3)
C(16) -C(17) -C(18) -H(181)	172 (4)
C(16) -C(17) -C(18) -H(182)	61 (3)
H(171)-C(17) -C(18) -H(181)	-14 (5)
H(171)-C(17) -C(18) -H(182)	-124 (5)
H(171)-C(17) -C(18) - O(2)	109 (3)
H(181)-C(18) - O(2) -C(19)	-146 (4)
H(182)-C(18) - O(2) -C(19)	-40 (3)
C(14) -C(19) -C(20) -H(201)	-117 (4)
C(14) -C(19) -C(20) -H(202)	144 (3)
H(191)-C(19) -C(20) -H(201)	123 (5)
H(191)-C(19) -C(20) -H(202)	24 (4)
H(191)-C(19) -C(20) -C(21)	-110.3(26)
O(2) -C(19) -C(20) -H(201)	8 (4)
O(2) -C(19) -C(20) -H(202)	-92 (3)
H(191)-C(19) - O(2) -C(18)	54.6(25)
H(201)-C(20) -C(21) - N(1)	86 (4)
H(201)-C(20) -C(21) - O(1)	-93 (4)
H(202)-C(20) -C(21) - N(1)	-170 (3)
H(202)-C(20) -C(21) - O(1)	11 (3)

Unit cell projections down all three axes are shown on the next three pages following this section, and they demonstrate clearly both that the molecular packing is dominated entirely by hydrophobic interactions and that there are no hydrogen bonding contacts at all. There are two carbon oxygen contacts and one carbon nitrogen contact less than 3.5 angstroms.

C(4)	O(1)	3.310	(-1-x, 1/2+y, -1/2-z)
C(12)	O(1)	3.456(6)	(1/2+x, -1/2-y, -1-z)
C(20)	N(2)	3.372(6)	(-x, -1/2-y, -1/2-z)

There are five oxygen hydrogen and one nitrogen hydrogen contacts closer than 3 angstroms.

O(1)	H(041)	2.41(6)	(-1-x, -1/2+y, -1/2-z)
O(1)	H(112)	2.79(4)	(-1/2-x, -1-y, -1/2-z)
O(1)	H(122)	2.61(5)	(-1/2+x, -1/2-y, -1-z)
O(2)	H(051)	2.87(5)	(-x, -1/2+y, -1/2-z)
O(2)	H(091)	2.98(4)	(-x, -1/2+y, -1/2-z)
N(2)	H(201)	2.53(7)	(-x, 1/2+y, -1/2-z)

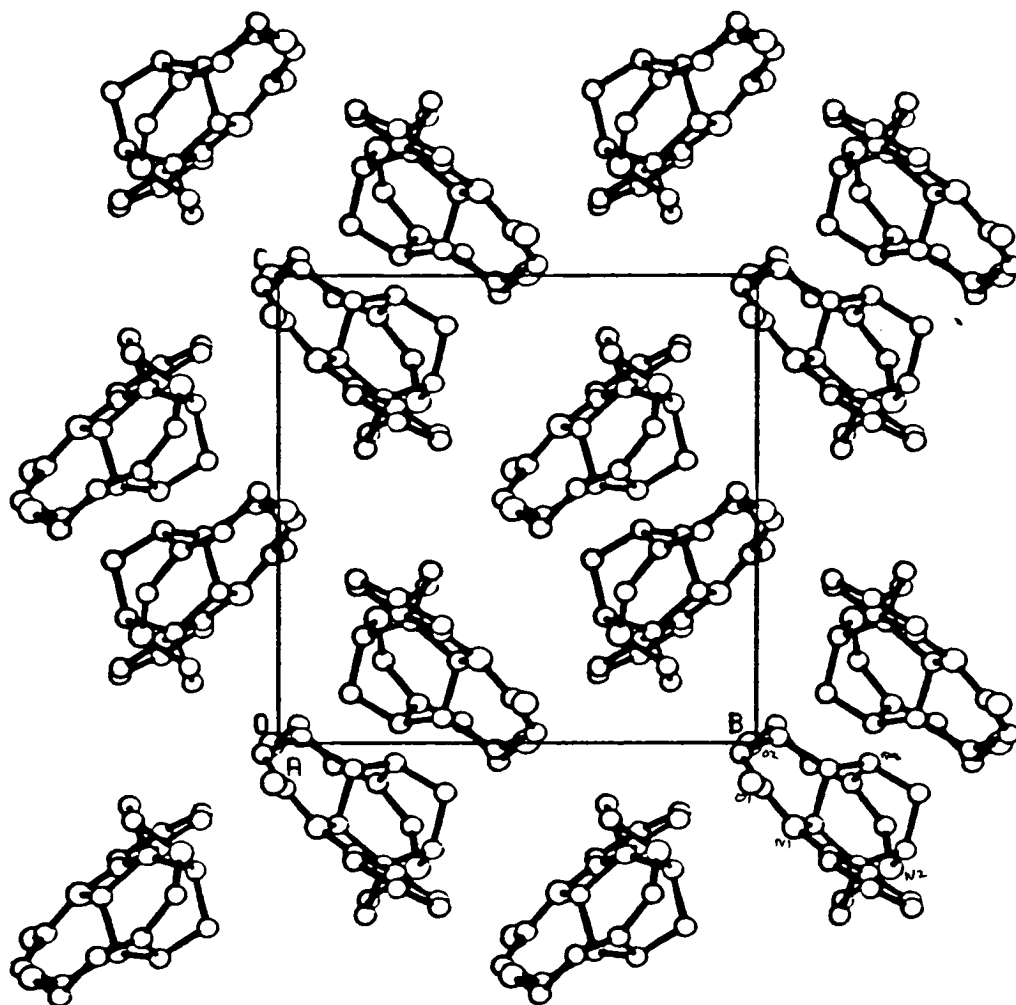
With one exception the carbon to hydrogen contacts may be divided into those involving the carbons of the aromatic ring, and those involving the double bond carbon atoms in the ether ring, which may be a result of an interaction between hydrogen atoms and a cloud of π electrons. The contacts between the aromatic ring and protons above or below the ring plane are of a type interpreted as an energetically favoured interaction between the π -cloud of the ring system and neighbouring protons when they occurred in solid benzene (Beevers, 1976).

C(1)	H(202)	2.83(5)	(-1/2-x, -1-y, 1/2+z)
C(2)	H(131)	2.83(5)	(-1/2+x, -1/2-y, -1-z)
C(2)	H(202)	2.92(5)	(-1/2-x, -1-y, 1/2+z)
C(4)	H(102)	2.99(5)	(-1/2+x, -1/2-y, -z)

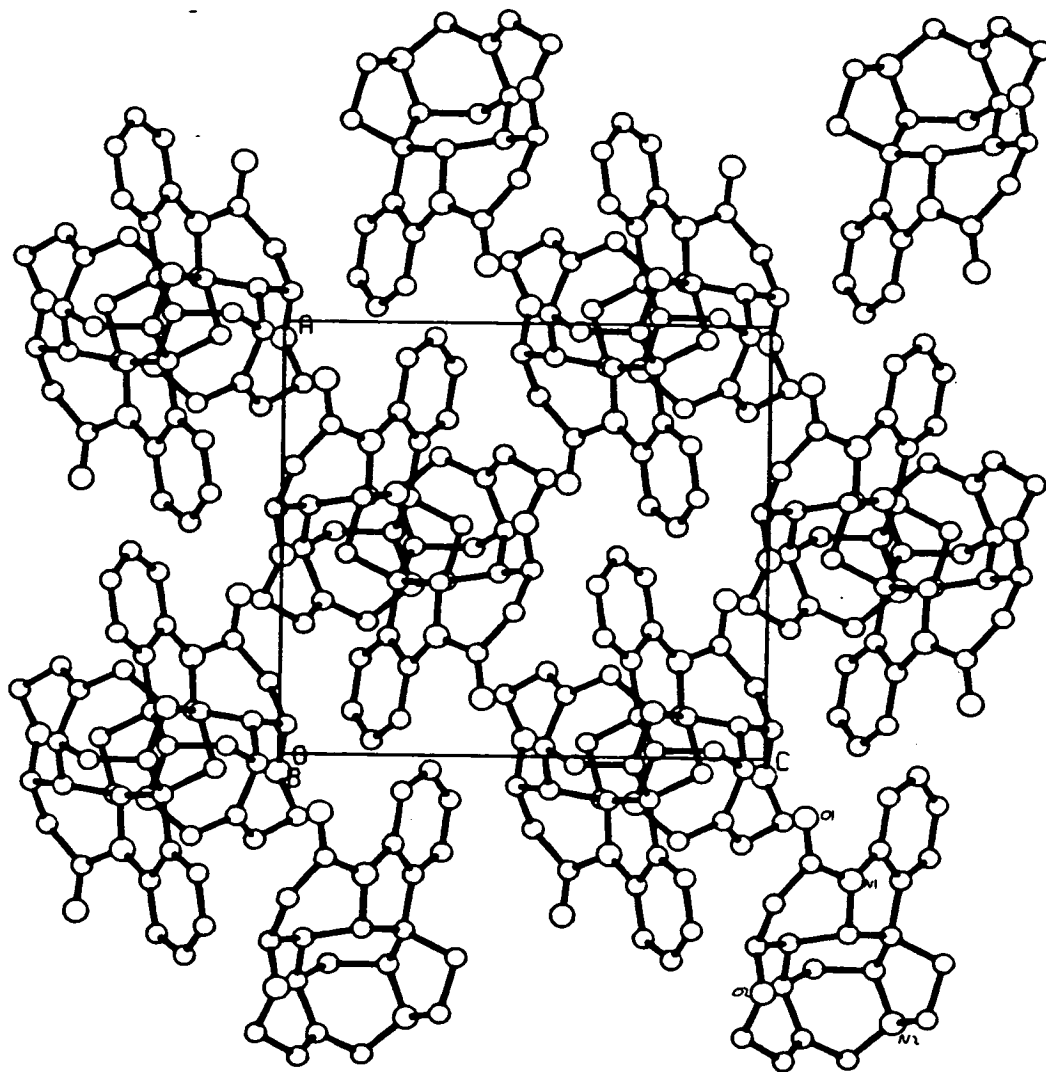
C(16)	H(141)	2.92(5)	$(1/2+x,$	$-1/2-y,$	$-1-z)$
C(17)	H(051)	2.86(5)	$(-x,$	$-1/2+y,$	$-1/2-z)$
C(17)	H(121)	2.90(5)	$(1/2+x,$	$-1/2-y,$	$-1-z)$
C(17)	H(141)	2.89(5)	$(1/2+x,$	$-1/2-y,$	$-1-z)$
C(18)	H(051)	2.88(5)	$(-x,$	$-1/2+y,$	$-1/2-z)$
C(10)	H(201)	2.84(7)	$(-x,$	$1/2+y,$	$-1/2-z)$

There are only three hydrogen to hydrogen contacts less than 2.5 Å.

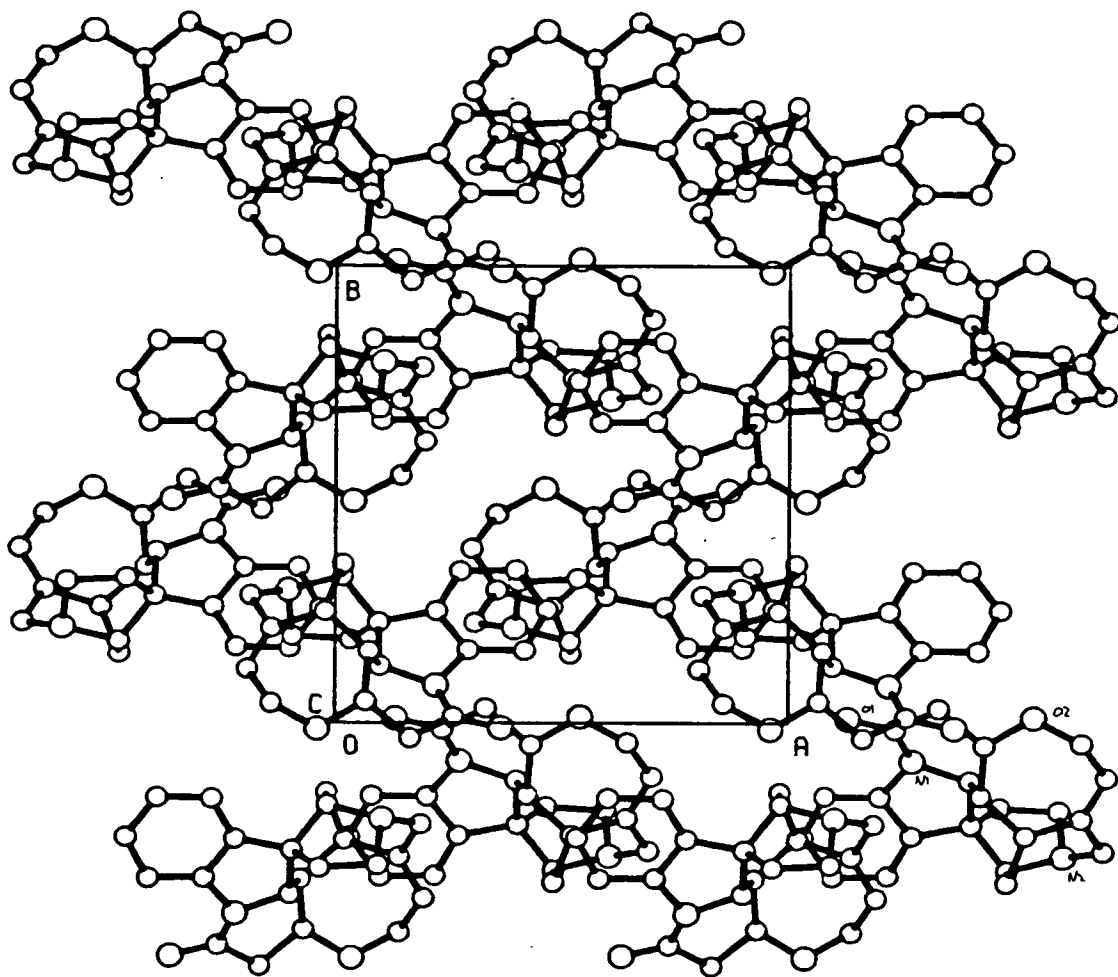
H(081)	H(091)	2.35(5)	$(-x,$	$-1/2+y,$	$-1/2-z)$
H(102)	H(201)	2.47(8)	$(-x,$	$1/2+y,$	$-1/2-z)$
H(112)	H(122)	2.36(6)	$(-x,$	$-1/2+y,$	$-1/2-z)$



PROJECTION DOWN THE X AXIS



PROJECTION DOWN THE Y AXIS



PROJECTION DOWN THE Z AXIS

Comparison with the strychnine salts

A superficial comparison of the structure of the strychnine molecule with that of the average strychninium cation reveals little in the way of systematic structural differences. This may even be taken as self-evident, given the rigid nature of the molecular skeleton, but there is one significant difference between the two molecules. The bonds to the tertiary amine nitrogen in the free base are uniformly shorter than the equivalent bonds (to the tertiary ammonium nitrogen) in the idealised cation by about 5 σ .

bond	base	ideal cation	Δ
C(9)-N(2)	1.493(5)	1.537(3)	0.044
C(10)-N(2)	1.470(6)	1.501(3)	0.031
C(15)-N(2)	1.464(6)	1.445(3)	0.031

A possible rationalisation for this difference is that the lone pair on the nitrogen in the neutral molecule is free to overlap the bonding orbitals of the adjacent carbons and thus shorten the bonds, while the equivalent electron density in the cation is used to form a dative bond to the proton. There are no other significant differences between the two molecules. The packing of the strychnine molecules in the unit cell has no similarity at all to the packing of the strychnine salts, as it is governed solely by the need to fill space as efficiently as possible untempered by any consideration of electrostatic interactions.

Table Fractional Coordinates of Atoms with Standard Deviations

	x	y	z	Ueq
C(1)	-0.2946(3)	-0.3397(3)	-0.2632(3)	0.0374
C(2)	-0.4172(4)	-0.3407(5)	-0.2471(3)	0.0526
C(3)	-0.4661(5)	-0.2507(6)	-0.1887(4)	0.0671
C(4)	-0.3982(5)	-0.1640(5)	-0.1481(4)	0.0630
C(5)	-0.2768(4)	-0.1663(4)	-0.1629(4)	0.0507
C(6)	-0.2247(4)	-0.2542(3)	-0.2203(3)	0.0360
C(7)	-0.0957(3)	-0.2788(3)	-0.2384(3)	0.0322
C(8)	-0.0992(3)	-0.3752(3)	-0.3253(3)	0.0302
C(9)	-0.0165(4)	-0.1785(3)	-0.2713(3)	0.0368
C(10)	0.0958(4)	-0.3102(4)	-0.1635(4)	0.0445
C(11)	-0.0338(4)	-0.3172(4)	-0.1310(3)	0.0382
C(12)	-0.0176(4)	-0.1515(4)	-0.3937(4)	0.0416
C(13)	0.0226(4)	-0.2549(3)	-0.4576(4)	0.0384
C(14)	-0.0761(4)	-0.3424(3)	-0.4456(3)	0.0339
C(15)	0.1912(4)	-0.2205(4)	-0.3229(4)	0.0525
C(16)	0.1434(4)	-0.2923(4)	-0.4152(3)	0.0430
C(17)	0.1979(4)	-0.3825(4)	-0.4538(4)	0.0510
C(18)	0.1437(4)	-0.4579(4)	-0.5390(4)	0.0548
C(19)	-0.0679(4)	-0.4525(3)	-0.5127(3)	0.0393
C(20)	-0.1668(4)	-0.5328(4)	-0.4776(4)	0.0469
C(21)	-0.2619(4)	-0.4867(4)	-0.4038(3)	0.0434
N(1)	-0.2234(3)	-0.41769(25)	-0.3205(3)	0.0355
N(2)	0.1052(3)	-0.2091(3)	-0.2329(3)	0.0450
O(1)	-0.3666(3)	-0.5103(3)	-0.4153(3)	0.0679
O(2)	0.03974(25)	-0.51362(24)	-0.49798(23)	0.0453

Table Fractional Coordinates of Atoms with Standard Deviations

	x	y	z	Ueq
H(021)	-0.465(4)	-0.401(3)	-0.277(3)	0.0369
H(031)	-0.545(6)	-0.258(4)	-0.168(5)	0.0809
H(041)	-0.443(6)	-0.098(5)	-0.109(5)	0.0849
H(051)	-0.230(4)	-0.105(4)	-0.132(4)	0.0535
H(081)	-0.045(3)	-0.434(3)	-0.300(3)	0.0351
H(091)	-0.044(3)	-0.111(3)	-0.229(3)	0.0193
H(101)	0.126(4)	-0.380(4)	-0.208(4)	0.0470
H(102)	0.143(4)	-0.305(4)	-0.094(4)	0.0474
H(111)	-0.057(4)	-0.266(4)	-0.074(4)	0.0426
H(112)	-0.056(4)	-0.389(3)	-0.110(3)	0.0274
H(121)	-0.097(5)	-0.132(4)	-0.402(4)	0.0621
H(122)	0.037(4)	-0.085(4)	-0.412(4)	0.0475
H(131)	0.036(4)	-0.242(4)	-0.538(4)	0.0527
H(141)	-0.148(4)	-0.295(4)	-0.469(4)	0.0455
H(151)	0.211(5)	-0.137(5)	-0.356(5)	0.0694
H(152)	0.269(5)	-0.256(4)	-0.287(4)	0.0553
H(171)	0.274(4)	-0.404(4)	-0.420(4)	0.0629
H(181)	0.198(4)	-0.513(5)	-0.567(5)	0.0840
H(182)	0.126(4)	-0.418(4)	-0.614(4)	0.0552
H(191)	-0.074(3)	-0.435(3)	-0.589(4)	0.0323
H(201)	-0.135(5)	-0.601(6)	-0.445(6)	0.0884
H(202)	-0.192(4)	-0.573(4)	-0.546(4)	0.0551

Thermal Vibration Parameters with Standard Deviations

	U or U11	U22	U33	U23	U13	U12
C(1)	0.0354(21)	0.0489(21)	0.0280(17)	0.0120(18)	0.0009(15)	-0.0031(18)
C(2)	0.0372(24)	0.0861(34)	0.0347(22)	0.0071(24)	-0.0011(17)	-0.0079(25)
C(3)	0.0354(25)	0.1237(49)	0.0423(25)	0.0178(31)	0.0071(22)	0.0245(28)
C(4)	0.0574(29)	0.0886(39)	0.0431(22)	0.0020(27)	0.0050(23)	0.0247(31)
C(5)	0.0599(28)	0.0534(26)	0.0388(21)	0.0007(21)	0.0057(22)	0.0173(24)
C(6)	0.0353(20)	0.0446(21)	0.0280(16)	0.0050(17)	0.0003(15)	0.0042(17)
C(7)	0.0356(19)	0.0312(19)	0.0298(18)	0.0008(15)	-0.0014(15)	-0.0004(16)
C(8)	0.0333(18)	0.0288(17)	0.0284(16)	0.0011(15)	-0.0021(16)	0.0004(16)
C(9)	0.0411(22)	0.0273(19)	0.0421(20)	-0.0029(18)	-0.0002(17)	-0.0075(16)
C(10)	0.0389(23)	0.0529(25)	0.0419(22)	-0.0033(21)	-0.0129(19)	0.0002(19)
C(11)	0.0454(24)	0.0383(24)	0.0311(19)	0.0012(17)	-0.0047(17)	0.0007(19)
C(12)	0.0481(24)	0.0317(19)	0.0450(23)	0.0053(19)	0.0022(21)	-0.0024(20)
C(13)	0.0423(21)	0.0387(19)	0.0341(18)	0.0065(17)	0.0045(18)	-0.0013(18)
C(14)	0.0397(21)	0.0342(19)	0.0280(18)	0.0005(16)	-0.0023(15)	-0.0004(18)
C(15)	0.0424(23)	0.0617(29)	0.0535(25)	-0.0047(24)	-0.0014(22)	-0.0165(22)
C(16)	0.0361(21)	0.0525(23)	0.0405(21)	0.0085(19)	0.0093(17)	-0.0093(20)
C(17)	0.0382(25)	0.0673(29)	0.0475(24)	0.0112(24)	0.0127(22)	-0.0004(21)
C(18)	0.0527(27)	0.0635(29)	0.0481(25)	0.0036(24)	0.0184(22)	0.0111(25)
C(19)	0.0527(24)	0.0407(21)	0.0246(18)	0.0012(16)	-0.0008(16)	0.0068(19)
C(20)	0.0592(26)	0.0382(22)	0.0433(22)	-0.0045(19)	-0.0056(20)	-0.0075(21)
C(21)	0.0469(25)	0.0452(22)	0.0382(18)	0.0011(19)	-0.0054(18)	-0.0161(18)
N(1)	0.0336(15)	0.0381(16)	0.0350(16)	0.0027(15)	-0.0016(14)	-0.0093(14)
N(2)	0.0386(18)	0.0501(18)	0.0465(19)	0.0022(17)	-0.0036(16)	-0.0120(17)
O(1)	0.0548(20)	0.0879(24)	0.0611(19)	-0.0116(20)	-0.0061(16)	-0.0342(19)
O(2)	0.0533(17)	0.0408(15)	0.0419(15)	-0.0025(14)	0.0029(14)	0.0104(13)

CHAPTER EIGHT
STRYCHNINE NITRATE

STRYCHNINE NITRATE

Introduction

Although the inorganic salts of strychnine have varying degrees of hydration and crystallise in several different space groups, the actual molecular packing in each structure determined so far is quite similar (Gould et al, 1981). This implies that the recognition taking place between the strychnine cations depends neither on the degree of solvation, nor on the counter-ion present.

To test this hypothesis, and to provide more information for a comparative study of the strychnine salts, the X-ray structure of strychnine nitrate was examined. The table below summarises the similarities between the salts already examined.

Acid	HI	HCl	HBr	0.5(H ₂ SO ₄)
a/A	7.750	7.833	7.70	7.84
b/A	16.19	32.22	33.20	34.34
c/A	7.580	7.558	7.64	7.56
$\alpha/^\circ$	90	90	90	90
$\beta/^\circ$	90	90.82	90	90
$\gamma/^\circ$	92.57	90	90	94.75
Space Group	P2 ₁	P2 ₁	P2 ₁ 2 ₁ 2 ₁	I2
Z	2	4	4	4
Hydration	1	1.5	2	2.5

Crystallisation

The crystals were grown by dissolving the salt in the minimum possible quantity of hot water, and allowing the solution to cool slowly to ambient temperature. The crystal used for diffraction was grown by Miss Rosemary Kelly as part of an Honours project on drug-receptor interactions.

Crystal Data

Formula $C_{21}H_{23}N_2O_2^+ NO_3^-$ $M = 397.41$ $F(000) = 840$

Space Group $P2_12_12_1$ Int. Tab. No. 19

Cell dimensions $a = 7.371(4)$ $b = 31.020(10)$ $c = 8.056(4)$ Å

$V = 1842.0$ Å³ $Z = 4$ $D_c = 1.433$ gcm⁻³

Radiation MoK α $\lambda = 0.71069$ Å $\mu = 0.97$ cm⁻¹

$\theta_{max} = 25^\circ$ $\sin\theta_{max} / \lambda = 0.5947$

$h_{max} = 8$ $k_{max} = 36$ $l_{max} = 8$

Final $R = 0.0555$ based on 1076 independent reflections

1865 reflections measured (1859 after sorting)

765 reflections considered unobserved ($F < 5\sigma(F)$)

332 parameters refined

Max. shift/ σ in final cycle 0.048

Difference Fourier map max. 0.2141 eÅ⁻³ min. -0.2299 eÅ⁻³

Data collection and reduction

The data was collected on the STADI-2 two circle diffractometer and the program STOEABS was used to correct for Lorentz and polarisation effects.

Solution and refinement

The structure was solved using the MULTAN77 system. The program NORMAL was used to calculate structure factors based on a Debye curve derived from the spherically averaged molecular scattering factors of a strychnine cation (from the strychnine hydrochloride structure). The electron density map revealed almost the entire strychnine cation and the resultant atomic coordinates were refined using the program SHELX-76. The table below summarises the course of the refinement.

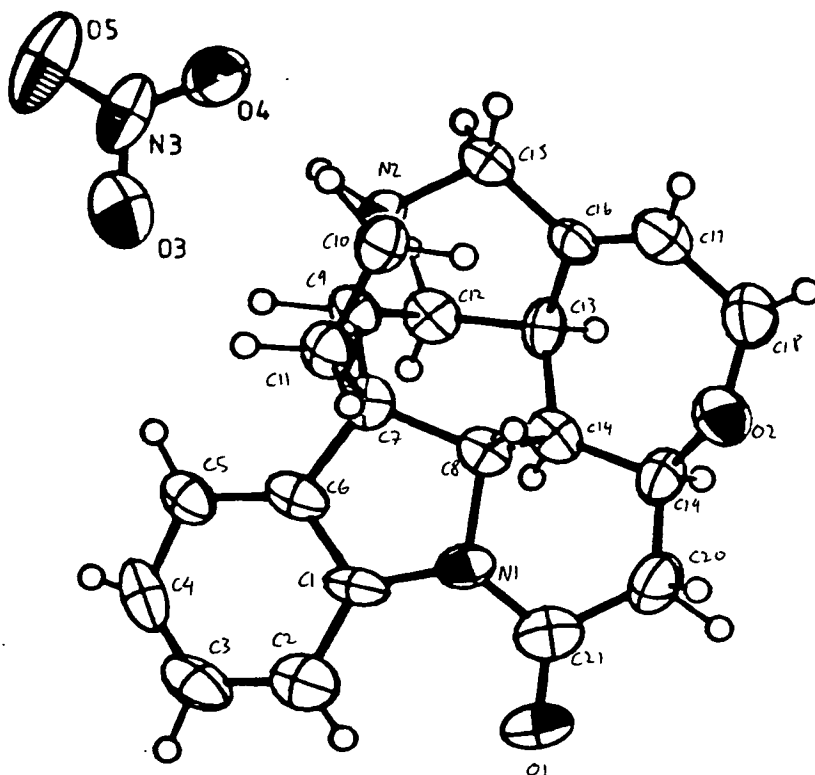
R	Model
0.3433	All but one carbon atom of cation input
0.3366	Nitrate group and other carbon atom located
0.1820	All non-Hydrogen atoms input
0.1274	Thermal parameters allowed to refine
0.1158	Hydrogens input in calculated positions
0.0989	All non-Hydrogen atoms with anisotropic thermal parameters
0.0621	Convergence and removal of constraints to H positions
0.0555	Final convergence

All hydrogen thermal parameters were constrained to be equal, but no positional restraints were employed

At one stage during the refinement, the fractional coordinates of a strychnine cation (from strychnine hydrochloride) were input to SHELX to investigate whether they would refine towards the correct coordinates. This was not found to be the case, but this may have been due to the difference in length of the long axes of the unit cells and the difference in relative position in the unit cell for the two molecules (This was approximately 1/3 of a unit cell in all directions, as can be seen from the packing diagrams on the succeeding pages).

Structure Description

There are no significant differences between the strychnine nitrate cation and the idealised strychnine cation (see below). This is almost certainly attributable to the rigid conformation of the molecule and the similarity of molecular environment over the series of strychnine salts used in the idealisation process. As would be expected, there are no significant differences between the strychnine nitrate cation and the cations in any of the other salt structures. Structural parameters for all the strychnine salts studied at Edinburgh are tabulated at the end of the section on the calculation of the idealised cation.



A comparison of the strychnine cation with the neutral molecule shows only slight differences in lengths of bonds to the protonated nitrogen and of the ether ring double bond which are not very significant and the perspective plot of the cation shows that this similarity extends to the thermal motions of the atoms.

Bond	StrH ⁺	Str	Idealised
C9 N2	1.543(11)	1.493(5)	1.53(4)
C10 N2	1.518(12)	1.470(6)	1.50(4)
C15 N2	1.498(12)	1.464(6)	1.49(3)
C16 C17	1.277(14)	1.322(6)	1.314(16)

There are also slight differences in torsion angle around the amide and ether rings, but as this is the only part of the molecule where the geometry is not rigidly constrained, it is to be expected that some small variation in this region will occur between the free base and the cation depending on the different environments in the base or in the salt.

Molecular packing

The molecular packing in all the inorganic salts of strychnine examined so far consists of bilayers of strychnine cations separated by sheets containing the counter-ions. With the exception of the strychnine sulphate sesterhydrate (Bokhoven *et al.*, 1951), the bilayers in each structure appear superficially similar, but the finer detail of the intermolecular interactions within the bilayer varies with the symmetry operators applied in each space group. The anhydrous nitrate salt and the bromide dihydrate both crystallise in $P2_12_12_1$, with alternate bilayers related to one another by a twofold screw axis, and the molecules within each bilayer are related to each other by a combination of unit cell translations and twofold screw axes. Although the iodide monohydrate and the chloride sesquihydrate both crystallise in $P2_1$, there are two strychnine cations in the asymmetric unit of the chloride salt, related by a non-crystallographic twofold screw axis. The bilayers are related by the crystallographic screw axis. In the iodide salt however, the bilayers are related only by a unit cell translation and the twofold screw axis relates molecules within the bilayer. The sulphate sesterhydrate contains molecules related by rotation and by the body centre of the lattice. This results in a distinctive shift of half a unit cell translation between the bilayers and a striking change in relative orientation within them as compared to the rest of the structures. It is particularly noticeable that the alternation of orientation of the cations between layers no longer occurs in the bilayers.

The layers of anions contain varying amounts of water of crystallisation in the different structures, ranging from none in the nitrate, to the sulphate sesterhydrate which has some solvent disorder. Where present, this water allows the formation of infinite hydrogen bonded sheets running parallel to the bilayers. This can be contrasted with the nitrate structure in which each anion interacts specifically with only one cation.

It is a measure of the stability of the bilayer packing arrangement

that these salts are so insoluble compared with the strychnine: amino acid complexes, which tend to have less close contacts between strychnine molecules, although some vestiges of the bilayer system are detectable in the strychnine:benzoylalanine complex (Gould et al, 1984).

The tables below show the close contacts found in the strychnine nitrate structure. It is immediately noticeable that there are proportionately far more contacts between the strychnine cation and the nitrate than occur between cations. This is due to the relative weakness of van der Waals forces as compared to electrostatic interactions.

There is only the one hydrogen bonded contact:

O(4)	2.695 (10)	N(2)	(x,y,z)
H(1N2)	1.82 (8)	O(4)	(x,y,z)

Other contacts to the nitrate ion are:

O(2)	3.315 (10)	O(4)	(x,y,1+z)
O(3)	3.195 (12)	N(2)	(x,y,z)
O(3)	3.194 (13)	C(9)	(x,y,z)
O(3)	3.497 (14)	C(11)	(x,y,z)
O(3)	3.476 (16)	C(18)	(-1+x,y,-1+z)
O(4)	3.455 (11)	C(9)	(x,y,z)
O(4)	3.455 (12)	C(15)	(x,y,z)
O(4)	3.241 (12)	C(15)	(2 1/2-x,-y,-1/2+z)
O(4)	3.479 (13)	C(17)	(2 1/2-x,-y,-1/2+z)
O(5)	3.156 (14)	C(10)	(x,y,-1+z)
O(5)	3.366 (14)	C(10)	(1 1/2-x,-y,-1/2+z)
O(5)	3.485 (13)	C(11)	(x,y,-1+z)
N(2)	3.385 (12)	N(3)	(x,y,z)

H(5)	2.91	(8)	O(3)	(x,y,z)
H(8)	2.96	(8)	O(5)	(x,y,1+z)
H(9)	2.31	(7)	O(3)	(x,y,z)
H(101)	2.96	(7)	O(4)	(x,y,z)
H(101)	2.53	(7)	O(5)	(1 1/2-x,-y,1/2+z)
H(102)	2.38	(8)	O(5)	(x,y,1+z)
H(112)	2.86	(7)	O(3)	(x,y,z)
H(121)	3.00	(8)	O(3)	(1+x,y,1+z)
H(152)	2.40	(7)	O(4)	(2 1/2-x,-y,1/2+z)
H(171)	2.78	(7)	O(4)	(2 1/2-x,-y,1/2+z)
H(1N2)	2.52	(8)	O(3)	(x,y,z)
H(1N2)	2.53	(8)	N(3)	(x,y,z)

The rest of the contacts are between strychnine cations in the same bilayer:

O(1)	3.396	(14)	C(20)	(-1/2+x,-1/2-y,2-z)
C(4)	3.440	(16)	C(12)	(-1+x,y,z)
H(202)	2.80	(7)	O(1)	(1/2+x,-1/2-y,2-z)
H(3)	2.87	(8)	C(1)	(-1/2+x,-1/2-y,1-z)
H(3)	2.92	(8)	C(6)	(-1/2+x,-1/2-y,1-z)
H(9)	2.97	(7)	C(20)	(x,y,-1+z)
H(111)	2.95	(8)	C(16)	(-1+x,y,z)
H(111)	2.79	(8)	C(17)	(-1+x,y,z)
H(13)	2.81	(8)	C(1)	(1+x,y,z)
H(13)	2.79	(8)	C(2)	(1+x,y,z)
H(13)	2.97	(8)	C(6)	(1+x,y,z)
H(201)	2.94	(8)	C(9)	(x,y,1+z)
H(2)	2.31	(11)	H(202)	(-1/2+x,-1/2-y,2-z)
H(9)	2.07	(11)	H(201)	(x,y,-1+z)

Comparison of molecular interactions in strychnine bilayers

The molecular interactions in these structures can be divided into four classes:

- a) Interactions within the hydrophilic sheets.
- b) Interactions between the hydrophilic sheets and the strychnine cations.
- c) Interactions between strychnine cations in the same half of a bilayer. (i.e separated by a unit cell translation)
- d) Interactions between strychnine cations on opposite sides of the bilayer.

An examination of these last two groups should provide the most useful information on the structural similarities and intermolecular interactions in the series of salts. The first two groups of interactions depend mostly on the counter-ion and degree of hydration of the salt. In general, the closest contacts within a bilayer are those involving a unit cell translation corresponding to the x direction in the nitrate salt (i.e x direction in the chloride and bromide or y direction in the iodide or sulphate). The table below shows those contacts of this type closer than 4 Å in the nitrate structure, with the equivalent distances from the other structures for comparison. It can be seen that although the precise numerical figures might change, these contacts all involve the same regions of the cation (C12-C17 to the aromatic ring and C11).

contact	NIT	CL	CL'	I	BR	SUL
C4-C12	3.439	3.476	3.490	3.481	3.569	3.678
C11-C16	3.605	3.788	3.818	3.877	3.787	3.751
C2-C13	3.627	3.637	3.607	3.600	3.523	3.892
C3-C13	3.640	3.632	3.633	3.539	3.516	3.985
C11-C17	3.669	3.791	3.903	3.859	3.767	4.013
C5-C12	3.694	3.831	3.855	3.855	4.006	3.576
C3-C12	3.712	3.694	3.648	3.673	3.683	3.734
C4-C13	3.717	3.793	3.853	3.706	3.699	4.270
C1-C13	3.734	3.778	3.745	3.759	3.749	3.660
C6-C13	3.769	3.924	3.967	3.933	3.873	3.988
C5-C13	3.774	3.929	4.011	3.864	3.953	3.968
C2-C14	3.932	4.106	4.147	4.180	4.142	4.676
C11-C15	3.956	4.043	4.008	4.055	4.155	4.166

The next table shows the contacts between the strychnine cations separated by a unit cell translation equivalent to the 'z' direction in the strychnine nitrate structure.(i.e by z in all but the bromide, which has y). Again it can be seen that the contacts involve the same general regions of the molecule. Although the exact orientation is not totally maintained, the interaction is in each case between the face of the molecule with the C4-C5 and C9-C12 bonds and that with the oxygen containing ring and carbonyl group.

contact	NIT	CL	CL'	I	BR	SUL
O2-C12	3.605	3.295	3.313	3.269	3.356	3.317
O1-C4	3.739	3.614	3.655	3.607	3.693	3.284
O2-C9	3.745	3.679	3.680	3.610	3.764	3.575
O1-C5	3.757	4.055	4.088	3.967	3.901	4.206
C12-C20	3.796	4.043	4.051	3.904	4.009	4.127
C9-C20	3.840	4.247	4.255	4.082	4.167	4.170
C5-C21	3.921	3.894	3.930	3.833	3.926	3.858
C5-C20	3.940	3.702	3.714	3.712	3.727	3.617

There are no contacts between strychnine cations involving translation in both of the short unit cell directions. For most of the structures examined, the contacts across a bilayer involve the carbonyl group and the aromatic ring. These contacts are not close enough to be considered as any form of ring stacking, but they do seem as if the non-polar end of each molecule forms a sort of hydrophobic pocket for the aromatic group across the bilayer from it.

contact	NIT	CL	CL'	I	BR	SUL
O1-O1	3.939	4.165	4.115	4.135	4.183	4.306
O1-C20	3.397	3.152	3.159	3.240	3.153	3.851
O1-C19	3.621	3.326	3.391	3.400	3.479	4.902
O1-C4	3.939	3.604	3.633	3.625	3.655	3.608
O1-C21	3.954	4.014	4.034	4.068	4.071	4.831
O1-C2	3.955	4.122	4.812	4.792	4.842	3.170
C1-C3	3.680	3.814	3.821	3.796	3.759	>5
C2-C3	3.750	3.756	3.760	3.731	3.777	4.953
C3-C6	3.838	4.442	4.469	4.417	4.435	>5
C3-C3	3.881	4.346	4.346	4.333	4.372	>5
C2-C20	3.975	4.294	4.426	4.384	4.343	3.094
C3-C4	3.993	4.914	4.935	4.878	4.770	>5

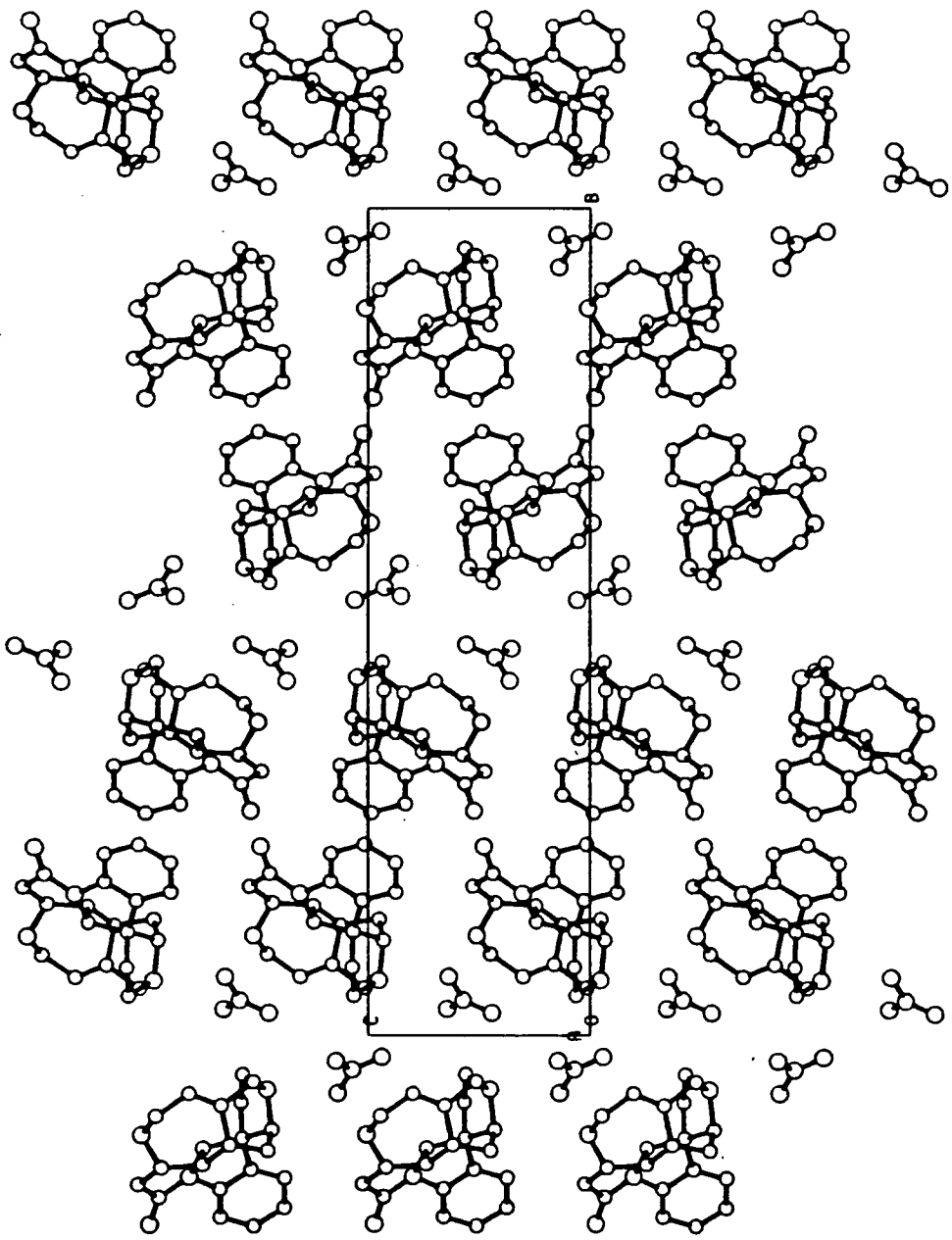
It is quite clear that the contacts listed for the strychnine sulphate cation show that it has a very different interaction at this part of the bilayer. The greatest similarities lie between the four halide structure cations, although their interactions are not extremely different from those of the nitrate. In the halide salts, the carbonyl oxygen is close to the aromatic ring, while in the nitrate, the carbonyl oxygen seems to prefer an interaction with the carbonyl carbon of a neighbouring cation. In none of the strychnine salts is there any sign of hydrogen bonding to this carbonyl group. This could be because the

nitrogen of the amide allows formation of a continuous π -system with the aromatic ring, leading to loss of electron density on the oxygen.

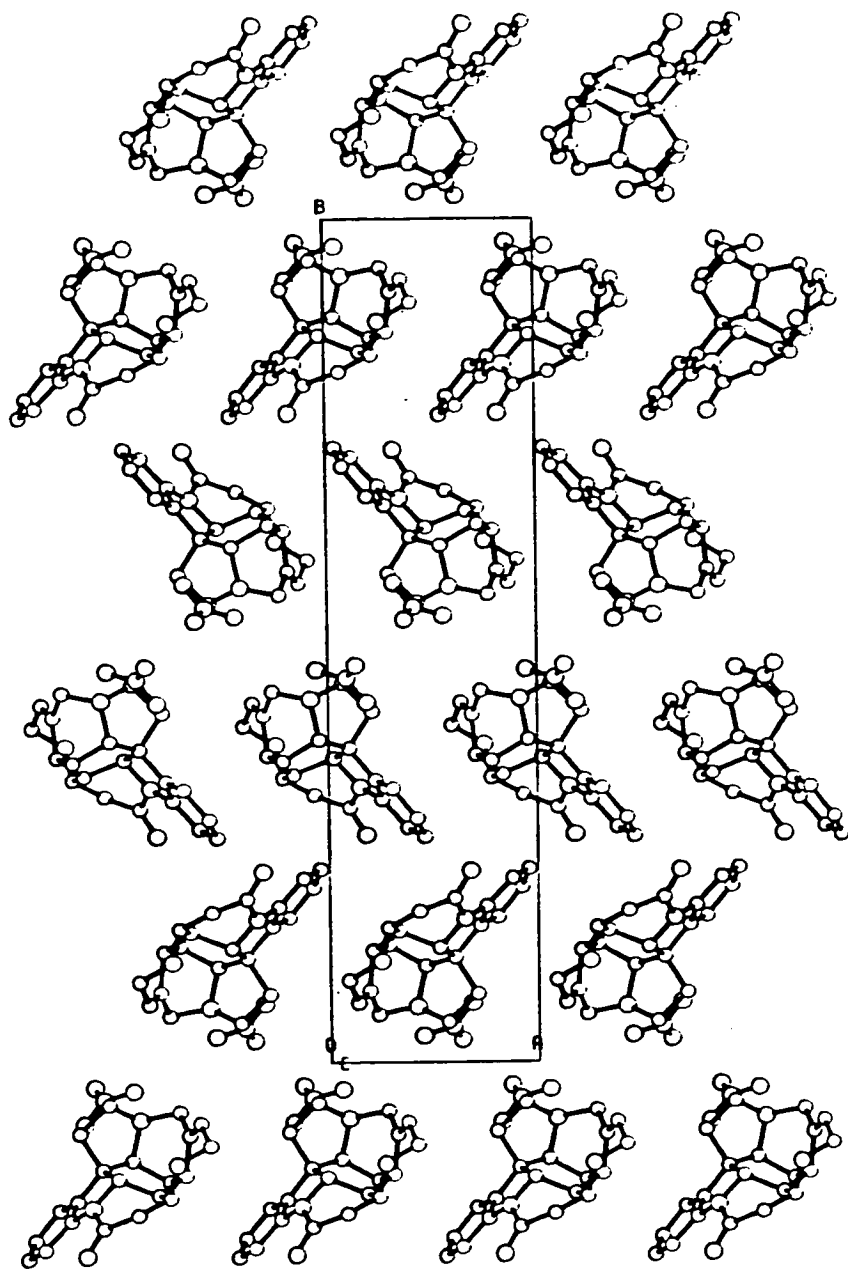
The table below shows the rest of the close contacts for the strychnine sulphate salt not listed above.

C2-C21	3.251
C4-C21	3.363
C3-C20	3.408
C3-C21	3.450
C3-C19	3.473
C2-C19	3.643
C3-C14	3.882

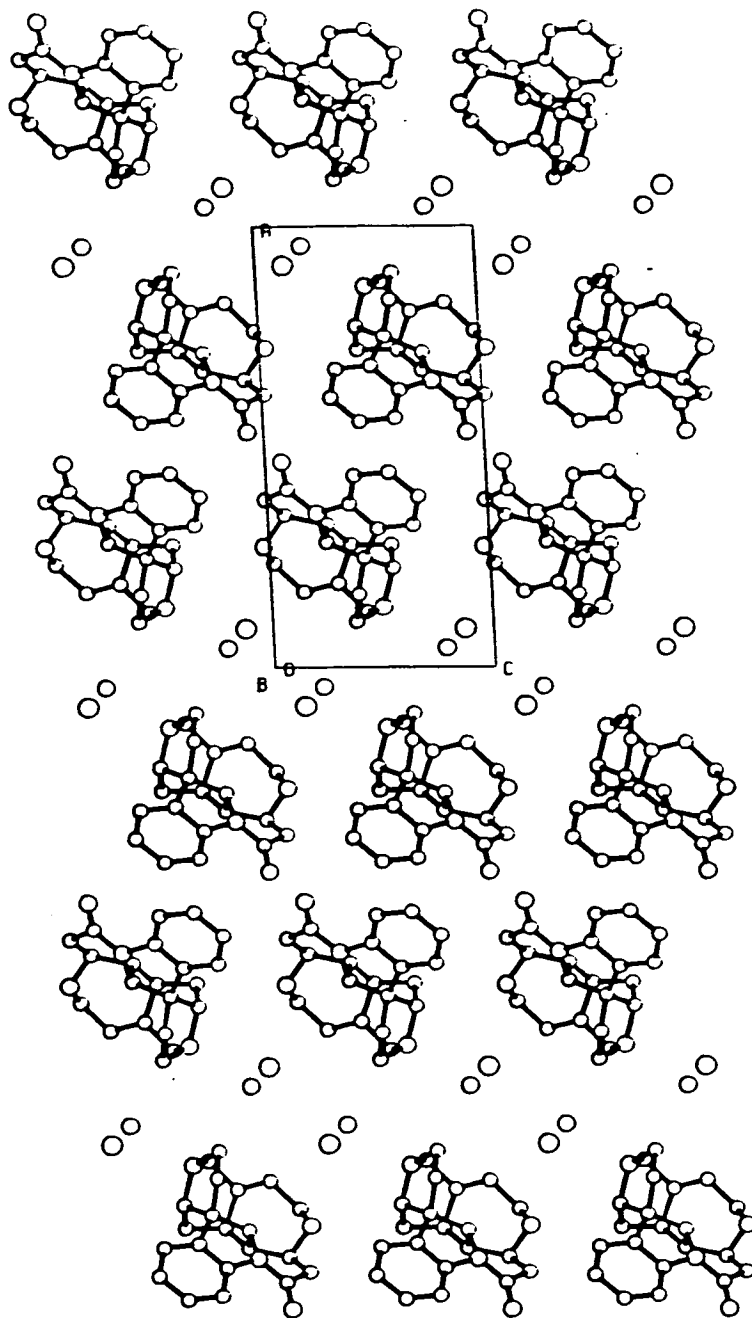
In this case it is clear that there is still an hydrophobic interaction involving the aromatic ring, but it no longer involves the carbonyl group.



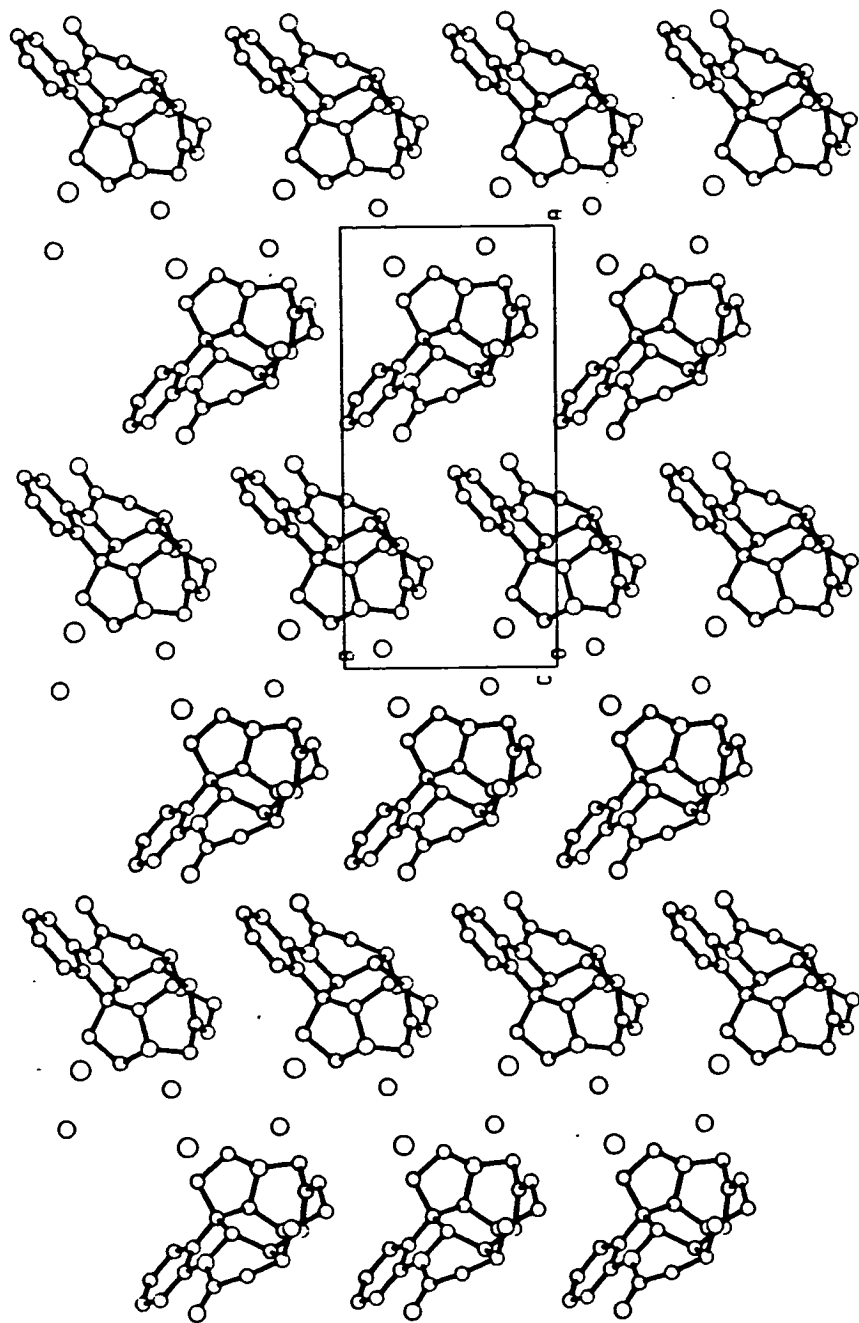
NITRATE VIEW DOWN X AXIS



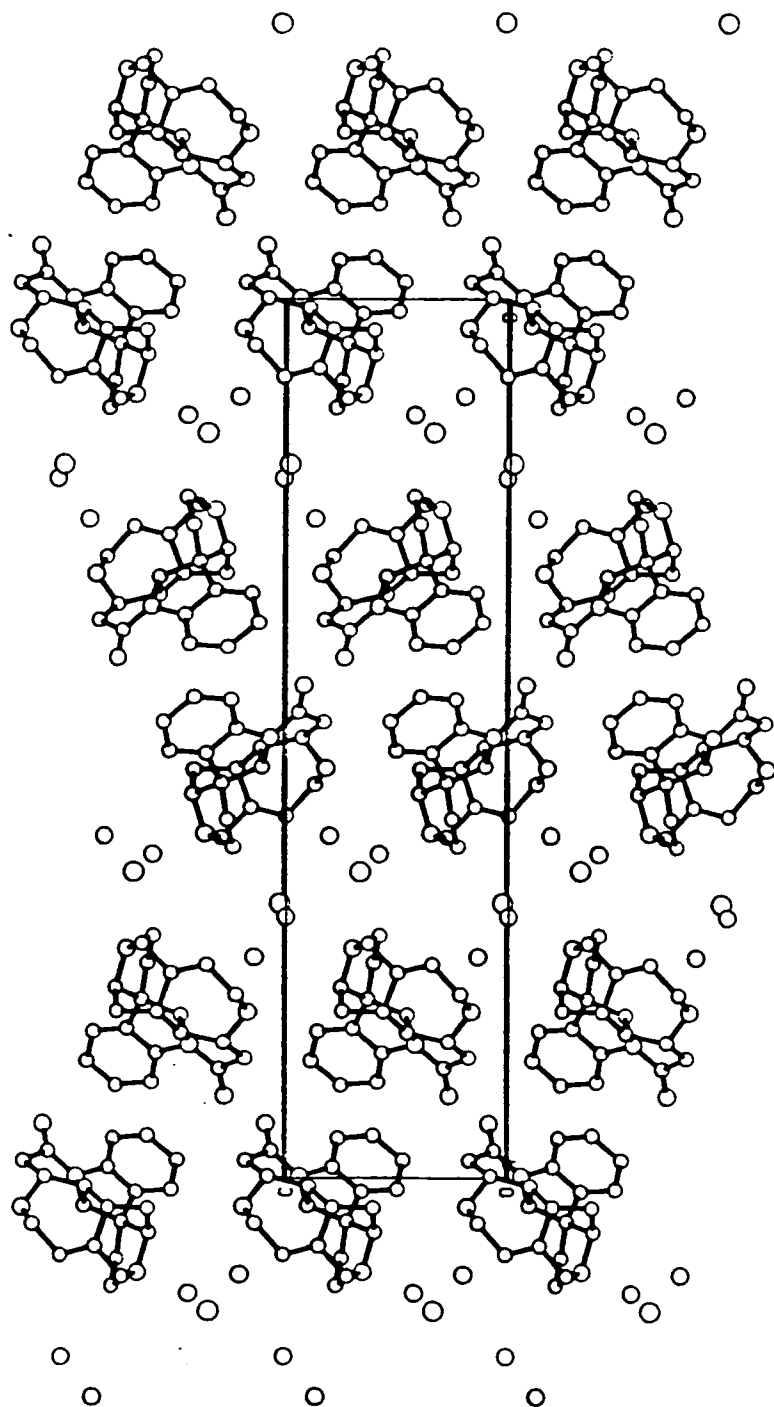
NITRATE VIEW DOWN Z AXIS



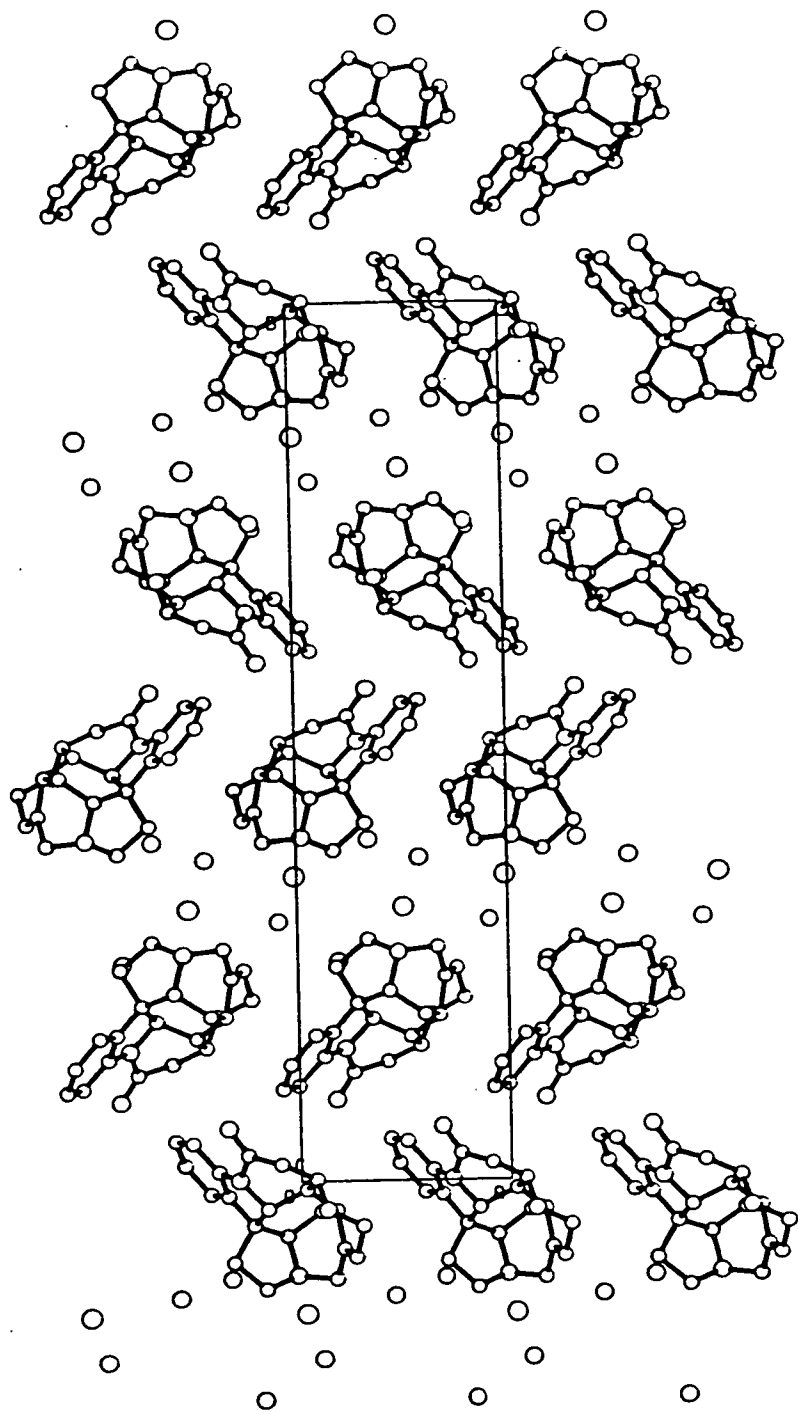
IODIDE VIEW DOWN Y AXIS



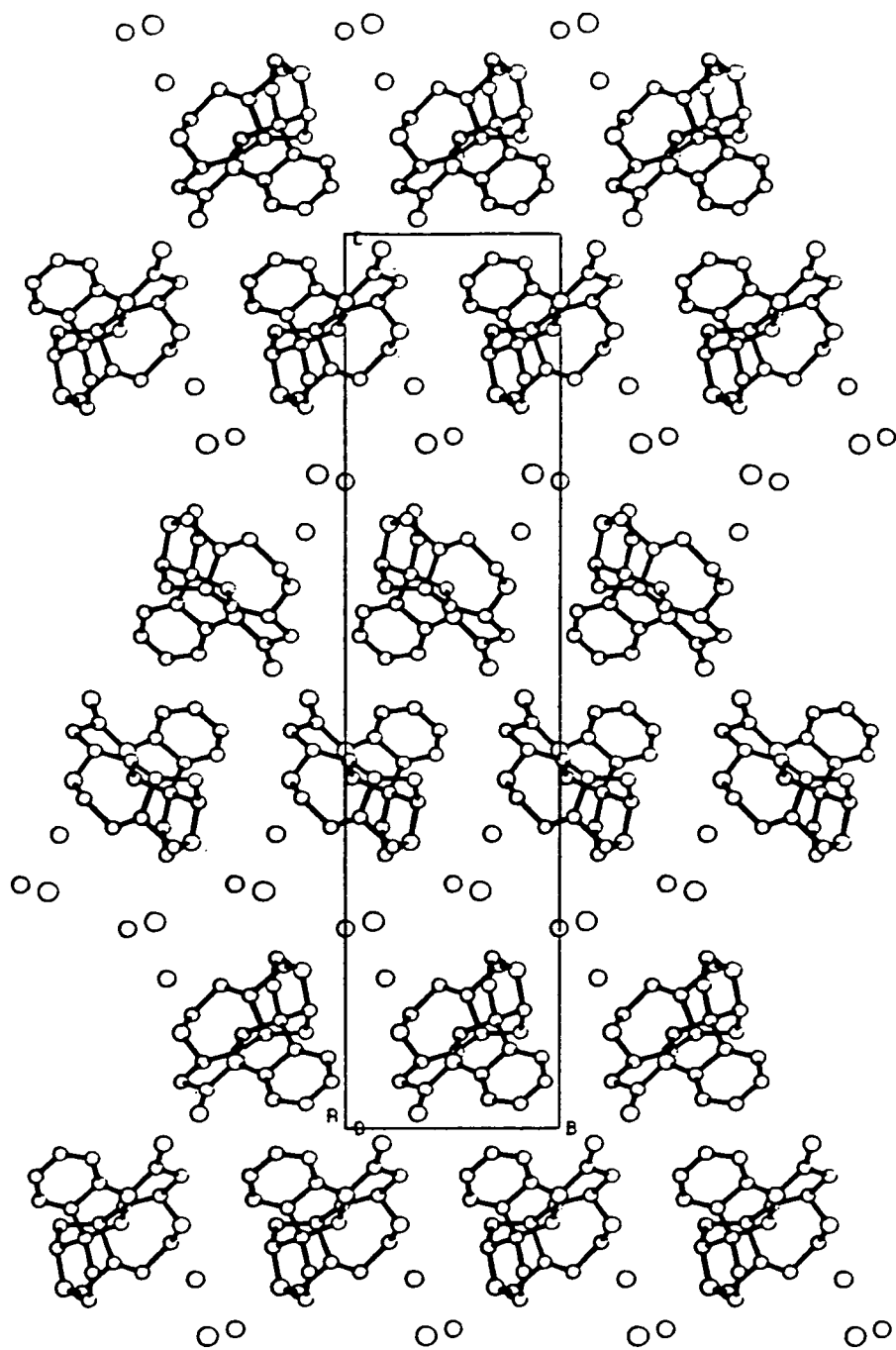
IODIDE VIEW DOWN Z AXIS



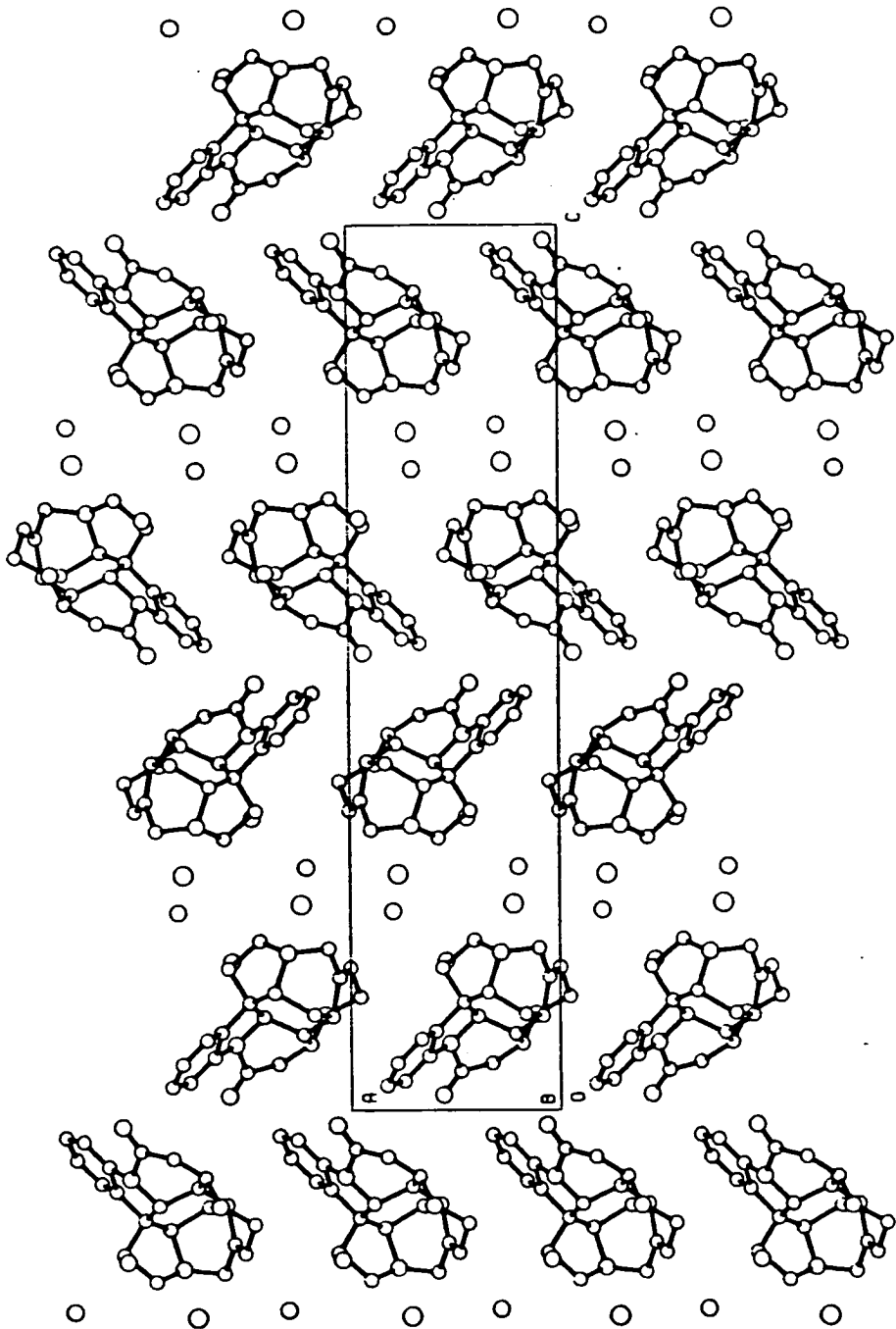
CHLORIDE VIEW DOWN X AXIS



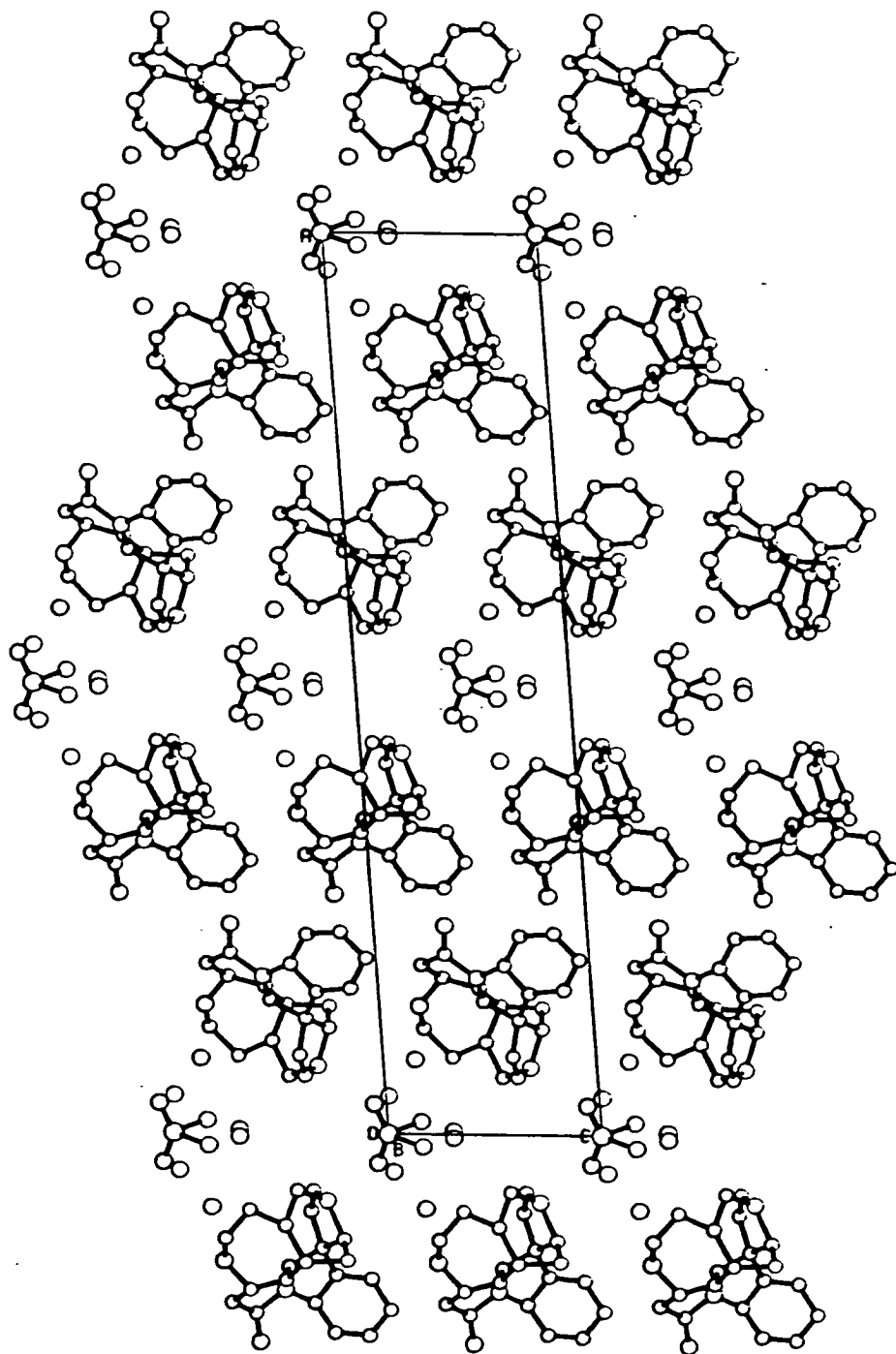
CHLORIDE VIEW DOWN Z AXIS



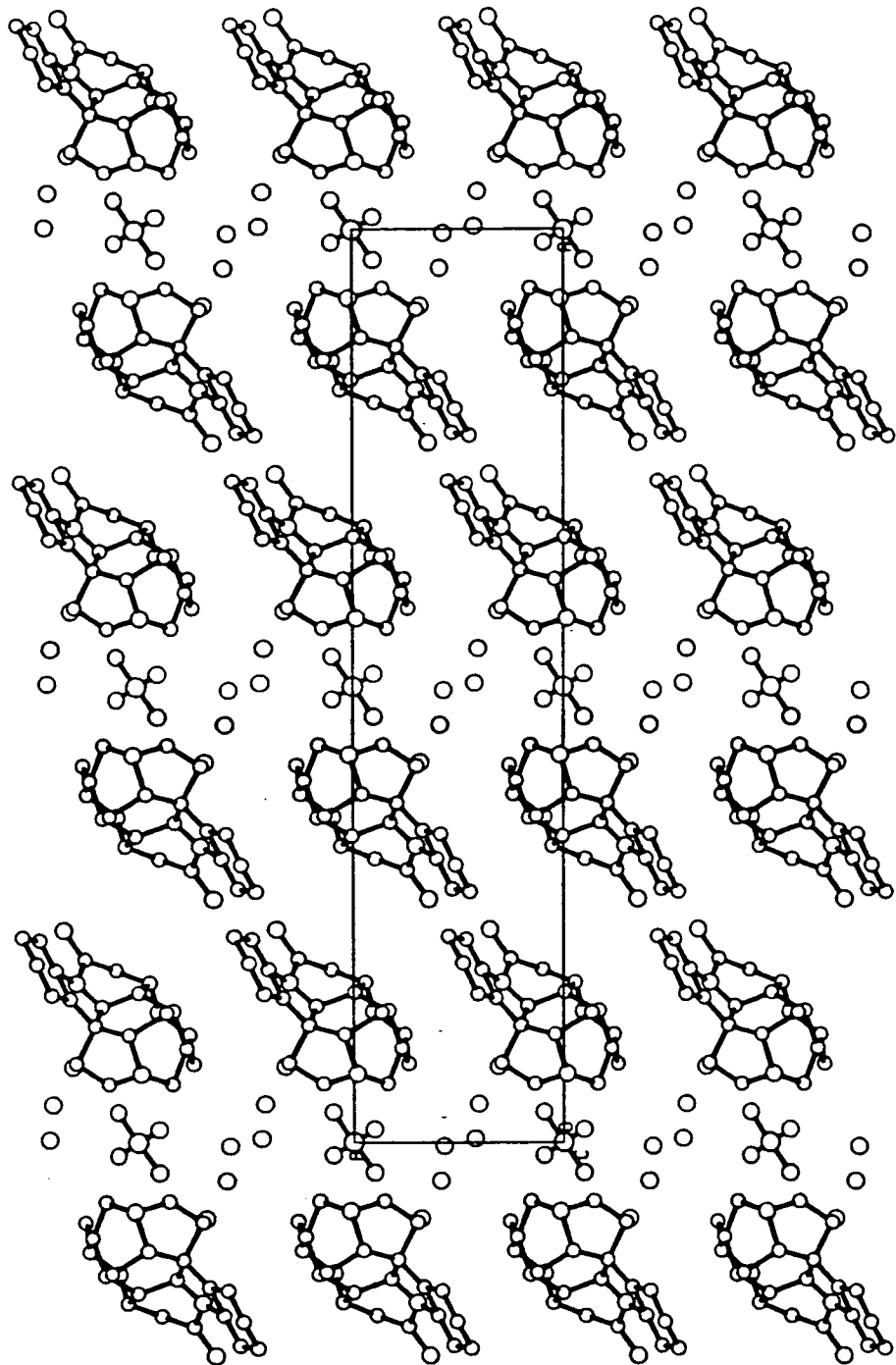
BROMIDE VIEW DOWN X AXIS



BROMIDE VIEW DOWN Y AXIS



SULPHATE VIEW DOWN Y AXIS



SULPHATE VIEW DOWN Z AXIS

Calculation of an idealised strychnine cation

As the strychnine cation is of almost identical geometry and environment in each of the salts studied so far, it was thought that an 'average strychnine' would be useful both for comparison with other strychnine containing structures, and as an aid for structure solution (the molecule being used either as a basis for the calculation of a Debye curve or for use with a rotation search of a Patterson map)

An idealised strychnine molecule, produced using the positional e.s.d.'s of the different strychnine cations whose structures are known, could possibly yield more accurate geometrical information than is normally available from a single crystal structure determination.

Two different methods were used to calculate the ideal geometry from the available sets of coordinates. The first method used was to transform the sets of coordinates to be idealised to the same coordinate system and then calculate a weighted mean of each positional parameter, which could then be used to calculate the geometry. The other method used was that of calculating a weighted mean of each bond length, angle and torsion in each of the input molecules.

Calculation of ideal molecule

The IDEAL link in the CALC program was used to produce a best fit of each input molecule onto a starting set of coordinates, which were then converted into orthogonal angstrom units and combined with their positional root mean square e.s.d.'s. The program below was then used to calculate the ideal molecule.

```

      DIMENSION F(3,10),E(10),V(10),S(3,10),S4(3,10),TOT(3),
      *RECIPV(3),SUM(3),SUM4(3),FID(3),EID2(3)
      C SETS UP ARRAYS FOR COORDINATES, ESD'S VARIANCES + THOSE USED
      C IN CALCULATING WEIGHTED MEANS
      CHARACTER*4 L1
      READ(11,*)J
      C READS IN FROM TERMINAL THE NUMBER OF MOLECULES TO AVERAGE

```

```

1   DO 10 K=1,3
    TOT(K)=0
    RECIPV(K)=0
    SUM(K)=0
    SUM4(K)=0
10  CONTINUE
C   THIS INITIALISES THE ARRAY CONTENTS TO 0 BEFORE
C   READING EACH SET OF FIGURES
    DO 20 I=1,J
C   ONCE FOR EACH MOLECULE
    RFAD(I,999)L1,(F(K,I),K=1,3),E(I)
    IF(L1.EQ.' ')STOP
C   FILE ENDING OF A BLANK LINE STOPS THE PROGRAM
    V(I)=1.0/(E(I)**2)
C   RECIPROCAL VARIANCE
    DO 30 M=1,3
C   FOR EACH OF X Y Z COORDS PER MOLECULE
    TOT(M)=TOT(M)+(F(M,I)*V(I))
C   WEIGHTED SUMS OF COORDS
    RECIPV(M)=RECIPV(M)+V(I)
C   SUMS OF RECIPROCAL VARIANCE FOR EACH MOLECULE
    FID(M)=TOT(M)/RECIPV(M)
C   WEIGHTED COORD
30  CONTINUE
20  CONTINUE
    EID1=SQRT(1.0/RECIPV(1))
C   UNWEIGHTED ESD
    DO 40 I=1,J
    DO 50 K=1,3
C   CALCULATES WEIGHTED ESD OR SIGMA2
    S(K,I)=((FID(K)-F(K,I))*V(I))**2
    SUM(K)=SUM(K)+S(K,I)
    S4(K,I)= V(I)**2
    SUM4(K)=SUM4(K)+S4(K,I)
    EID2(K)=SQRT(SUM(K)/SUM4(K))
50  CONTINUE
40  CONTINUE
    EIDISO=SQRT((EID2(1)**2+(EID2(2)**2)+(EID2(3)**2))/3)
    EID=AMAX1(EID1,EIDISO)
    WRITE(12,1000)L1,(FID(K),K=1,3),(EID2(K),K=1,3)
    *EIDISO,EID1,EID
    GO TO 1
999  FORMAT(A4,4X,4F10.5)
1000  FORMAT(' ',A4,4X,9F10.5)
END

```

$$\text{Average X coord } X_m = \frac{\sum (X_i / (\sigma_i^2))}{\sum (\sigma_i^{-2})}$$

$$\text{First e.s.d } \sigma_1 = \sqrt{1 / \sum (\sigma_i^{-2})}$$

$$\text{"Isotropic e.s.d" } \sigma_{\text{iso}} = \sqrt{((\sigma_x^2) + (\sigma_y^2) + (\sigma_z^2)) / 3}$$

$$\text{WHERE } \sigma_x = \sqrt{\sum ((X_i - X_m) / (\sigma_i^2)) / \sum (1 / (\sigma_i^2)^2)}$$

The ideal molecule with e.s.d's was read into CALC and used to produce a list of bonds, angles and torsions (column headed ID(I) in table below).

The second method involved using a modification of the above program on each molecule's ZIP output from CALC. (ZIP is the link of the CALC program that produces an output list of bonds, angles and torsion angles. The list may contain e.s.d's if requested and the e.s.d's of the atomic coordinates are supplied.)

```

DIMENSION F(10),E(10),V(10),S(10),S4(10)
CHARACTER*6 L1,L2,L3,L4
READ(11,*)J
1  TOT=0
   RECIPV=0
   SUM=0
   SUM4=0
   DO 20 I=1,J
     READ(I,999)L1,L2,L3,L4,F(I),E(I)
     IF(L1.EQ.' ')STOP
     V(I)=1.0/(E(I)**2)
     TOT=TOT+(F(I)*V(I))
     RECIPV=RECIPV+V(I)
     FID=TOT/RECIPV
20  CONTINUE
   EID1=SQRT(1.0/RECIPV)
   DO 40 I=1,J
     S(I)=((FID-F(I))*V(I))**2
     SUM=SUM+S(I)
     S4(I)= V(I)**2
     SUM4=SUM4+S4(I)
   EID2=SQRT(SUM/SUM4)
40  CONTINUE

```



```
EID=AMAX1(EID1,EID2)
WRITE(12,1000)L1,L2,L3,L4,FID,EID2,EID1,EID
GO TO 1
999 FORMAT(4A6,2F10.4)
1000 FORMAT(' ',4A6,4F10.4)
END
```

The result of this program is tabulated as ID(BAT) and has fractionally better e.s.d's than the ID(I) output. This is most noticeable around the oxepin ring, where there is most conformational difference between individual structures. As the bond lengths and angles are nearly the same in each case, the ideal parameters calculated on this basis have small weighted e.s.d's. For the parameters calculated from the averaged coordinates however, the e.s.d's are a lot larger, as there is a greater difference in actual atomic position relative to the origin of the coordinate system here than in any other part of the molecule and the average position for the atoms of this ring are closer to the molecular centre, leading to distorted, shorter bond lengths.

LABEL	CL1	CL2	I	NIT	AL	ID(1)	ID(BAT)
C1 C2	1.382(5)	1.376(5)	1.375(14)	1.401(14)	1.376(8)	1.386(13)	1.381(5)
C1 C6	1.391(5)	1.382(5)	1.402(13)	1.368(12)	1.385(7)	1.391(17)	1.386(5)
C1 N1	1.410(5)	1.406(5)	1.404(12)	1.392(11)	1.432(7)	1.410(13)	1.411(8)
C2 C3	1.399(6)	1.366(6)	1.382(16)	1.392(16)	1.377(8)	1.396(20)	1.382(15)
C3 C4	1.365(7)	1.384(7)	1.365(17)	1.328(17)	1.399(9)	1.36(3)	1.376(14)
C4 C5	1.376(6)	1.382(7)	1.364(16)	1.399(16)	1.382(9)	1.38(4)	1.382(5)
C5 C6	1.384(5)	1.383(5)	1.379(14)	1.396(13)	1.394(6)	1.38(3)	1.386(4)
C6 C7	1.495(5)	1.510(5)	1.488(12)	1.515(12)	1.512(7)	1.495(20)	1.504(8)
C7 C8	1.554(5)	1.542(5)	1.552(12)	1.553(12)	1.569(7)	1.554(15)	1.552(8)
C7 C9	1.526(5)	1.534(5)	1.524(12)	1.533(12)	1.515(7)	1.53(3)	1.527(6)
C7 C11	1.529(5)	1.527(5)	1.514(12)	1.559(13)	1.539(7)	1.53(3)	1.531(5)
C8 C14	1.533(5)	1.515(5)	1.479(12)	1.541(13)	1.531(7)	1.534(14)	1.524(10)
C8 N1	1.473(4)	1.483(4)	1.493(11)	1.483(11)	1.500(6)	1.473(12)	1.483(8)
C9 C12	1.504(5)	1.505(5)	1.476(13)	1.530(13)	1.516(8)	1.50(3)	1.506(5)
C9 N2	1.530(5)	1.543(5)	1.545(12)	1.543(11)	1.538(7)	1.53(4)	1.538(7)
C10 C11	1.512(5)	1.510(5)	1.499(13)	1.478(14)	1.528(8)	1.51(4)	1.511(6)
C10 N2	1.491(5)	1.507(5)	1.484(13)	1.518(13)	1.491(8)	1.50(4)	1.498(8)
C12 C13	1.541(5)	1.521(5)	1.540(13)	1.518(13)	1.516(8)	1.541(24)	1.529(10)
C13 C14	1.513(5)	1.521(5)	1.546(12)	1.524(13)	1.546(8)	1.513(19)	1.523(10)
C13 C16	1.512(5)	1.517(5)	1.526(12)	1.518(12)	1.510(8)	1.512(17)	1.515(3)
C14 C15	1.525(5)	1.535(5)	1.547(12)	1.512(13)	1.520(8)	1.52(3)	1.529(6)
C15 C16	1.491(5)	1.479(5)	1.473(14)	1.503(13)	1.514(9)	1.493(19)	1.490(10)
C15 N2	1.497(5)	1.490(5)	1.509(14)	1.498(12)	1.494(8)	1.49(3)	1.494(4)
C16 C17	1.314(5)	1.315(5)	1.327(14)	1.277(14)	1.315(8)	1.314(16)	1.313(4)
C17 C18	1.485(6)	1.488(6)	1.495(16)	1.488(16)	1.507(9)	1.48(3)	1.491(6)
C18 O2	1.415(5)	1.423(5)	1.412(14)	1.432(14)	1.429(8)	1.41(3)	1.421(5)
C19 C20	1.501(5)	1.523(6)	1.535(14)	1.522(16)	1.531(9)	1.50(4)	1.516(12)
C19 O2	1.426(5)	1.424(5)	1.415(12)	1.435(12)	1.440(7)	1.42(3)	1.427(5)
C20 C21	1.504(5)	1.481(6)	1.489(15)	1.512(16)	1.521(9)	1.50(4)	1.498(14)
C21 O1	1.216(5)	1.213(5)	1.194(14)	1.216(12)	1.206(7)	1.22(4)	1.212(4)
C21 N1	1.368(5)	1.373(5)	1.365(13)	1.373(12)	1.363(7)	1.37(3)	1.369(4)

LABEL	CL1	CL2	I	NIT	AL	ID(1)	ID(BAT)
C2 C1 Ce	121.2(3)	121.6(3)	121.7(9)	119.5(9)	122.0(5)	121.2(9)	121.4(4)
C2 C1 N1	122.6(3)	122.5(3)	122.4(9)	122.2(8)	127.9(5)	122.6(8)	122.38(23)
C2 C1 N1	117.3(3)	109.9(3)	109.9(8)	112.3(8)	110.1(4)	110.3(9)	110.2(3)
C1 C2 C3	117.3(4)	117.7(4)	117.3(10)	116.2(10)	117.9(5)	117.3(10)	117.6(3)
C2 C3 C4	121.9(4)	121.4(4)	121.3(11)	121.2(11)	121.3(6)	121.8(17)	121.6(3)
C3 C4 C5	120.3(4)	121.0(4)	121.6(11)	122.3(11)	120.0(6)	120.2(24)	120.6(4)
C4 C5 C6	119.5(4)	117.3(4)	116.9(10)	116.6(9)	119.0(6)	119.6(23)	118.4(11)
C1 C2 C5	119.9(3)	120.8(3)	119.1(9)	121.9(8)	119.7(5)	119.8(16)	120.3(5)
C1 C6 C7	109.6(3)	110.6(3)	110.5(8)	109.1(7)	111.4(4)	109.6(11)	110.3(6)
C5 C6 C7	130.4(3)	128.3(3)	130.0(8)	128.8(8)	128.8(5)	130.4(16)	129.2(10)
C6 C7 C8	103.4(3)	102.1(3)	102.5(7)	102.6(7)	102.1(4)	103.4(10)	102.6(7)
C6 C7 C9	115.2(3)	114.5(3)	113.3(7)	114.4(7)	116.5(4)	115.1(13)	115.0(6)
C6 C7 C11	112.6(3)	112.6(3)	114.2(7)	111.6(7)	111.4(4)	112.6(14)	112.4(4)
C8 C7 C9	113.5(3)	114.3(3)	113.4(7)	116.8(7)	113.8(4)	113.6(12)	114.0(5)
C8 C7 C11	111.1(3)	112.3(3)	111.6(7)	111.1(7)	111.2(4)	111.0(13)	111.6(6)
C9 C7 C11	101.4(3)	101.5(3)	102.2(7)	100.8(7)	102.2(4)	101.5(15)	101.58(25)
C7 C6 C14	117.3(3)	116.7(3)	118.3(7)	115.2(7)	116.6(4)	117.3(8)	116.9(4)
C7 C8 N1	104.0(3)	104.7(3)	104.3(6)	104.6(7)	105.0(4)	104.1(7)	104.5(4)
C14 C8 N1	105.4(3)	106.3(3)	106.4(7)	106.6(7)	105.7(4)	105.4(7)	105.9(4)
C7 C9 C12	115.5(3)	116.3(3)	116.8(7)	114.2(7)	116.6(5)	115.6(16)	115.9(5)
C7 C9 N2	104.9(3)	104.1(3)	104.4(7)	104.7(6)	105.4(4)	105.0(18)	104.7(4)
C12 C9 N2	110.9(3)	109.8(3)	109.9(7)	108.8(7)	110.2(4)	110.7(19)	110.2(6)
C11 C10 N2	104.3(3)	104.1(3)	104.9(7)	105.4(8)	105.1(5)	104.2(21)	104.4(3)
C7 C11 C10	103.5(3)	103.9(3)	105.3(7)	104.4(8)	102.9(4)	103.5(18)	103.7(4)
C9 C12 C13	108.2(3)	108.6(3)	109.8(8)	108.9(7)	107.8(5)	108.3(15)	108.4(3)
C12 C13 C14	106.4(3)	107.1(3)	105.4(7)	107.2(7)	106.5(4)	106.4(12)	106.6(4)
C12 C13 C16	110.1(3)	108.9(3)	107.9(7)	110.5(7)	111.1(5)	110.1(12)	109.7(7)
C14 C13 C16	114.0(3)	113.7(3)	112.8(7)	113.5(7)	113.6(4)	114.0(11)	113.74(24)
C8 C14 C13	112.0(3)	113.5(3)	113.3(7)	111.9(7)	112.4(4)	112.0(10)	112.7(7)
C8 C14 C19	107.7(3)	107.9(3)	109.3(7)	106.5(7)	108.3(4)	107.7(12)	107.9(3)
C13 C14 C19	119.1(3)	118.6(3)	116.7(7)	119.1(8)	118.9(4)	119.1(13)	118.7(3)
C16 C15 N2	110.2(3)	110.4(3)	109.3(8)	110.3(7)	109.8(5)	110.2(15)	110.19(21)
C13 C16 C15	115.4(3)	116.2(3)	116.1(8)	113.4(7)	114.4(5)	115.3(10)	115.5(6)
C13 C16 C17	123.0(3)	122.6(3)	120.8(8)	122.8(9)	123.5(5)	123.1(10)	122.8(4)
C15 C16 C17	121.5(3)	121.2(3)	123.0(9)	123.8(9)	122.0(5)	121.5(11)	121.7(4)
C16 C17 C16	121.6(4)	121.7(4)	123.5(10)	124.9(10)	121.8(6)	121.6(14)	121.9(4)
C17 C18 O2	111.2(4)	112.1(4)	110.8(9)	112.2(9)	111.4(5)	111.3(20)	111.6(4)
C14 C19 C20	110.8(3)	110.3(3)	108.3(8)	111.0(8)	110.1(5)	110.8(20)	110.4(4)
C14 C19 O2	114.5(3)	113.8(3)	113.6(7)	116.3(8)	114.8(5)	114.5(19)	114.3(5)
C20 C19 O2	103.8(3)	103.3(3)	103.0(7)	102.7(8)	104.6(5)	103.9(21)	103.7(4)
C19 C20 C21	117.0(3)	117.3(3)	118.0(9)	118.0(9)	115.6(5)	117.0(24)	117.0(5)
C20 C21 O1	123.2(4)	122.9(4)	122.7(10)	121.7(9)	123.0(5)	123. (3)	122.96(22)
C20 C21 N1	114.9(3)	114.9(3)	115.1(9)	115.1(9)	113.6(5)	114.9(20)	114.7(4)
O1 C21 N1	121.9(3)	122.2(4)	122.1(10)	123.2(9)	123.4(5)	122. (3)	122.3(4)
C18 O2 C19	116.1(3)	115.0(3)	114.5(8)	112.7(8)	115.8(4)	116.4(21)	115.4(6)
C1 N1 C8	109.6(3)	109.3(3)	108.6(7)	108.6(7)	108.7(4)	109.6(8)	109.2(3)
C1 N1 C21	125.7(3)	124.6(3)	124.7(8)	124.8(8)	124.1(4)	125.7(13)	124.9(6)
C8 N1 C21	118.5(3)	118.6(3)	118.5(8)	119.2(7)	118.9(4)	118.6(13)	118.63(17)
C9 N2 C10	107.6(3)	107.5(3)	107.8(7)	106.9(6)	107.2(4)	107.5(22)	107.48(17)
C9 N2 C15	113.9(3)	113.7(3)	113.2(7)	113.7(6)	114.1(5)	114.3(21)	113.79(17)
C10 N2 C15	112.5(3)	113.3(3)	112.9(8)	113.2(7)	113.7(5)	112.1(22)	113.1(4)

LAEEL			CL1	CL2	I	NIT	AL	ID(I)	ID(BAT)	
C6	C1	C2	C3	0 9(6)	0 1(6)	0.7(15)	4.7(15)	3.5(8)	1.0(16)	1.3(12)
N1	C1	C2	C3	181.6(4)	181.5(4)	179.9(10)	184.0(9)	182.9(5)	-178.2(11)	181.9(6)
C2	C1	C6	C5	-0.6(6)	1.3(6)	0.8(14)	-2.9(14)	-2.4(8)	-0.6(21)	-0.3(13)
C2	C1	C6	C7	175.2(3)	175.9(3)	174.8(9)	172.9(8)	174.0(5)	175.2(10)	175.1(7)
N1	C1	C6	C5	178.7(3)	180.2(3)	181.4(8)	177.7(8)	178.1(5)	178.7(16)	175.2(8)
N1	C1	C6	C7	-5.4(4)	-5.3(4)	-4.6(10)	-6.5(10)	-5.4(6)	-5.5(14)	-5.36(24)
C2	C1	N1	CE	172.8(4)	171.4(4)	171.5(9)	175.8(9)	174.9(5)	172.9(8)	172.8(12)
C2	C1	N1	C21	21.2(6)	22.4(6)	22.8(15)	26.5(15)	27.2(8)	21.1(20)	23.2(20)
C6	C1	N1	CE	-6.5(4)	-7.4(4)	-5.2(10)	-4.9(10)	-5.7(6)	-6.4(12)	-6.7(6)
C6	C1	N1	C21	-156.2(3)	-156.3(3)	-156.9(9)	-154.2(8)	-153.4(5)	-158.2(16)	-156.4(15)
C1	C2	C3	C4	0.0(7)	0.4(7)	-0.2(17)	-2.8(17)	-2.9(9)	0.0(24)	-0.6(11)
C2	C3	C4	C5	-1.0(7)	-2.4(7)	-1.7(19)	-1.0(19)	1.1(10)	-1 (4)	-1.1(12)
C3	C4	C5	C6	1.3(7)	3.7(7)	3.2(17)	2.8(17)	0.0(9)	2 (4)	2.1(14)
C4	C5	C6	C1	-0.5(6)	-3.2(6)	-2.7(15)	-0.9(14)	0.6(8)	-1 (3)	-1.4(15)
C4	C5	C6	C7	-175.3(4)	-176.6(4)	-175.3(10)	-175.7(9)	-175.2(5)	-175.4(20)	-175.8(7)
C1	C6	C7	CE	14.1(4)	14.8(4)	15.5(9)	14.3(9)	13.4(5)	14.2(13)	14.4(5)
C1	C6	C7	C9	136.6(3)	138.9(3)	138.1(8)	141.7(8)	138.1(5)	138.6(14)	138.8(4)
C1	C6	C7	C11	-105.6(3)	-105.8(3)	-105.4(9)	-104.7(8)	-105.3(5)	-105.7(16)	-105.63(22)
C5	C6	C7	CE	-170.6(4)	-171.1(4)	-171.4(10)	-170.3(9)	-170.5(5)	-170.6(19)	-170.8(3)
C5	C6	C7	C9	-46.1(5)	-47.1(5)	-46.8(13)	-42.9(12)	-45.8(8)	-46 (3)	-46.4(7)
C5	C6	C7	C11	69.5(5)	68.2(5)	67.7(12)	70.7(11)	70.8(7)	69.5(25)	69.2(9)
C6	C7	CE	C14	98.9(3)	99.0(3)	98.1(8)	100.3(8)	100.7(5)	98.9(11)	99.3(6)
C6	C7	CE	N1	-17.1(3)	-18.1(3)	-15.8(8)	-16.4(8)	-15.9(5)	-17.1(11)	-17.4(8)
C9	C7	CE	C14	-26.6(4)	-25.3(4)	-24.4(10)	-25.7(11)	-25.7(6)	-26.6(15)	-25.8(7)
C9	C7	CE	N1	-142.6(3)	-142.4(3)	-142.4(7)	-142.3(7)	-142.3(4)	-142.5(12)	-142.43(17)
C11	C7	CE	C14	-140.1(3)	-140.2(3)	-139.2(8)	-140.4(8)	-140.4(4)	-140.1(13)	-140.18(18)
C11	C7	CE	N1	103.9(3)	102.7(3)	102.8(8)	102.9(8)	103.0(4)	103.9(13)	103.2(6)
C6	C7	C9	C12	-83.4(4)	-83.9(4)	-84.3(9)	-87.8(9)	-83.6(6)	-83.4(19)	-84.0(6)
C6	C7	C9	N2	154.2(3)	155.2(3)	154.1(7)	153.3(7)	153.9(4)	154.3(16)	154.5(6)
CE	C7	C9	C12	35.6(4)	33.3(4)	32.0(10)	32.0(10)	34.9(6)	35.6(21)	34.2(12)
CE	C7	C9	N2	-86.8(3)	-87.5(3)	-85.6(8)	-86.9(8)	-87.7(5)	-86.7(18)	-87.4(5)
C11	C7	C9	C12	154.8(3)	154.5(3)	152.3(8)	152.4(7)	154.8(5)	154.7(18)	154.4(4)
C11	C7	C9	N2	32.4(3)	33.6(3)	30.7(8)	33.6(8)	32.2(5)	32.4(20)	32.8(7)
C6	C7	C11	C10	-166.6(3)	-166.8(3)	-162.8(7)	-164.5(7)	-166.7(4)	-166.6(16)	-166.4(6)
CE	C7	C11	C10	78.0(3)	78.6(3)	81.6(8)	81.7(9)	80.2(5)	78.0(18)	79.0(9)
C9	C7	C11	C10	-43.0(3)	-43.9(3)	-40.0(8)	-42.7(9)	-41.6(5)	-43.0(20)	-42.9(8)
C7	CE	C14	C13	41.0(4)	40.2(4)	40.3(10)	41.4(10)	40.0(6)	41.0(13)	40.5(4)
C7	CE	C14	C15	173.8(3)	173.8(3)	172.3(7)	173.1(7)	173.4(4)	173.8(12)	173.59(22)
N1	CE	C14	C13	156.2(3)	156.5(3)	157.2(7)	156.9(7)	156.1(4)	156.2(10)	156.37(19)
N1	CE	C14	C15	-71.0(3)	-70.0(3)	-70.8(8)	-71.4(9)	-70.4(5)	-71.0(13)	-70.6(5)
C7	CE	N1	C1	14.9(3)	16.4(3)	18.4(9)	13.6(9)	13.8(5)	14.8(10)	15.4(10)
C7	CE	N1	C21	168.9(3)	167.5(3)	168.4(8)	164.9(7)	163.4(4)	168.9(14)	167.2(16)

LABEL	CL1	CL2	I	NIT	AL	ID(I)	ID(BAT)
C14 CE N1 C1	-109.2(3)	-107.7(3)	-107.5(8)	-108.9(8)	-110.0(4)	-109.2(9)	-108.7(9)
C14 CE N1 C21	44.8(4)	43.9(4)	42.6(10)	42.5(10)	39.6(6)	44.9(15)	43.3(15)
C7 C9 C12 C13	-57.7(4)	-54.9(4)	-55.0(10)	-52.8(10)	-57.7(6)	-57.7(21)	-56.3(14)
N2 C9 C12 C13	61.4(4)	62.9(4)	63.6(9)	62.6(9)	62.3(6)	61.6(22)	62.3(8)
C7 C9 N2 C10	-10.4(4)	-11.8(4)	-11.2(9)	-13.8(8)	-10.9(5)	-10.4(25)	-11.3(7)
C7 C9 N2 C13	115.1(3)	114.9(3)	114.4(8)	111.9(7)	115.9(5)	114.7(21)	114.7(5)
C12 C9 N2 C10	-135.7(3)	-135.9(3)	-137.2(8)	-136.3(7)	-137.5(5)	-135.8(20)	-136.6(7)
C12 C9 N2 C13	-10.3(4)	-10.6(4)	-11.6(10)	-10.6(9)	-10.6(6)	-11(3)	-10.55(24)
N2 C10 C11 C7	37.0(4)	36.7(4)	33.3(9)	34.7(9)	35.3(5)	37.0(24)	36.3(7)
C11 C10 N2 C9	-16.4(4)	-15.2(4)	-13.3(9)	-13.2(9)	-15.3(6)	-16(3)	-15.4(8)
C11 C10 N2 C15	-142.7(3)	-141.7(3)	-139.0(8)	-139.1(8)	-142.3(5)	-142.8(21)	-141.8(7)
C9 C12 C13 C14	69.3(3)	66.7(4)	66.4(9)	69.1(9)	68.7(6)	69.2(17)	68.1(12)
C9 C12 C13 C16	-54.7(4)	-56.6(4)	-54.3(9)	-55.0(9)	-55.5(6)	-54.8(17)	-55.5(9)
C12 C13 C14 C8	-61.0(3)	-60.5(4)	-55.6(9)	-62.2(9)	-61.3(6)	-60.9(14)	-60.9(4)
C12 C13 C14 C15	172.2(3)	171.3(3)	172.1(7)	171.8(8)	170.7(5)	172.2(15)	171.6(5)
C16 C13 C14 C6	60.6(4)	59.8(4)	57.9(10)	59.1(10)	61.4(6)	60.6(14)	60.2(6)
C16 C13 C14 C15	-66.3(4)	-66.4(4)	-70.4(10)	-65.9(11)	-66.6(6)	-66.3(18)	-67.4(11)
C12 C13 C16 C15	-2.1(4)	-1.0(4)	-6.1(11)	-4.2(10)	-1.9(7)	-2.0(16)	-2.0(8)
C12 C13 C16 C17	179.2(4)	178.5(3)	177.0(9)	176.8(9)	180.1(6)	179.2(13)	178.8(6)
C14 C13 C16 C15	-121.6(3)	-120.3(3)	-122.1(9)	-124.7(8)	-122.0(5)	-121.5(12)	-121.4(9)
C14 C13 C16 C17	59.7(5)	59.2(5)	61.0(11)	56.4(12)	60.0(7)	59.7(16)	59.4(4)
C6 C14 C19 C20	41.0(4)	40.9(4)	42.6(9)	46.6(10)	41.5(6)	41.0(22)	41.4(7)
C8 C14 C19 C2	-76.0(4)	-74.6(4)	-71.2(9)	-70.3(10)	-76.2(5)	-76.1(20)	-75.0(10)
C13 C14 C19 C20	169.9(3)	171.7(3)	172.8(8)	174.2(8)	171.4(5)	169.9(17)	171.2(10)
C13 C14 C19 C2	52.9(4)	56.1(4)	59.0(10)	57.3(11)	53.7(7)	52.8(24)	54.8(17)
N2 C15 C16 C13	52.8(4)	52.4(4)	56.0(11)	56.0(10)	53.1(7)	52.3(18)	53.1(7)
N2 C15 C16 C17	-126.4(4)	-127.1(4)	-127.2(10)	-125.0(10)	-128.8(6)	-128.8(17)	-127.7(7)
C16 C15 N2 C9	-45.9(4)	-45.2(4)	-46.0(10)	-47.2(9)	-45.7(6)	-45.3(25)	-45.7(4)
C16 C15 N2 C10	77.0(4)	78.0(4)	76.8(10)	75.0(9)	77.7(6)	77.3(23)	77.3(6)
C13 C16 C17 C1E	-2.7(6)	-1.9(6)	-3.5(15)	-1.2(16)	-3.7(9)	-2.8(21)	-2.5(6)
C15 C16 C17 C1E	178.6(4)	177.6(4)	179.8(10)	179.9(10)	178.4(6)	178.4(16)	178.3(6)
C16 C17 C16 C2	-65.6(5)	-66.2(5)	-65.1(13)	-65.6(14)	-64.6(8)	-65.2(24)	-65.6(5)
C17 C16 C2 C15	88.2(4)	88.3(4)	89.4(10)	84.4(10)	87.1(6)	87.9(25)	87.9(6)
C14 C19 C20 C21	13.4(5)	13.6(5)	12.1(12)	4.6(13)	15.5(7)	13(3)	13.3(10)
C2 C19 C20 C21	136.8(3)	135.6(4)	132.6(9)	129.6(9)	139.4(5)	136.8(24)	136.2(13)
C14 C19 C2 C1E	-64.6(4)	-65.7(4)	-65.8(10)	-65.7(10)	-64.9(6)	-64(3)	-65.4(8)
C20 C19 C2 C1E	174.5(3)	174.7(3)	173.3(8)	172.9(8)	174.2(5)	174.7(22)	174.4(3)
C19 C20 C21 N1	-42.1(5)	-42.6(5)	-41.3(13)	-35.7(13)	-47.9(7)	-42(3)	-42.9(18)
C19 C20 C21 C1	137.4(4)	138.3(4)	137.0(11)	142.5(10)	133.2(6)	138(3)	137.3(14)
C20 C21 N1 C1	159.4(3)	156.0(3)	156.9(9)	156.2(9)	162.1(5)	159.3(17)	159.1(12)
C20 C21 N1 C6	9.9(5)	11.7(5)	12.0(13)	9.8(12)	17.4(7)	10(3)	11.9(22)
D1 C21 N1 C1	-20.2(6)	-22.8(6)	-21.5(16)	-22.1(15)	-19.1(9)	-21(4)	-21.1(14)
D1 C21 N1 C6	-169.6(3)	-169.2(3)	-166.3(10)	-168.4(9)	-163.8(5)	-170.0(25)	-168.3(17)

Table Fractional Coordinates of Atoms with Standard Deviations

	x	y	z	Ueq
C(1)	0.7460(11)	-0.18600(25)	0.6686(13)	0.0423
C(2)	0.6199(15)	-0.2198(3)	0.6665(17)	0.0566
C(3)	0.5402(14)	-0.2308(3)	0.5151(18)	0.0599
C(4)	0.5763(16)	-0.2088(4)	0.3776(17)	0.0606
C(5)	0.6931(13)	-0.1733(3)	0.3764(13)	0.0450
C(6)	0.7774(10)	-0.16296(25)	0.5268(12)	0.0405
C(7)	0.8999(11)	-0.1252(3)	0.5668(11)	0.0392
C(8)	0.9734(12)	-0.1370(3)	0.7416(12)	0.0349
C(9)	1.0412(11)	-0.11502(25)	0.4327(11)	0.0378
C(10)	0.9334(13)	-0.0482(3)	0.5638(14)	0.0486
C(11)	0.7921(12)	-0.0819(3)	0.5714(14)	0.0447
C(12)	1.2189(12)	-0.1401(3)	0.4509(12)	0.0459
C(13)	1.2930(13)	-0.1332(3)	0.6247(11)	0.0417
C(14)	1.1659(12)	-0.1564(3)	0.7443(12)	0.0414
C(15)	1.2688(12)	-0.0571(3)	0.5151(13)	0.0463
C(16)	1.3130(11)	-0.0854(3)	0.6611(12)	0.0394
C(17)	1.3651(14)	-0.0710(3)	0.8018(14)	0.0551
C(18)	1.4095(17)	-0.0982(4)	0.9483(15)	0.0690
C(19)	1.2194(15)	-0.1602(3)	0.9249(12)	0.0483
C(20)	1.0672(18)	-0.1802(4)	1.0268(14)	0.0707
C(21)	0.9001(14)	-0.1958(3)	0.9360(13)	0.0534
N(1)	0.8493(10)	-0.17117(20)	0.8022(9)	0.0429
N(2)	1.0812(9)	-0.06649(20)	0.4532(10)	0.0386
O(1)	0.8205(11)	-0.22837(21)	0.9776(10)	0.0799
O(2)	1.2539(9)	-0.12056(19)	1.0110(8)	0.0588
N(3)	0.9126(15)	-0.0425(3)	0.0735(11)	0.0672
O(3)	0.8160(12)	-0.0708(3)	0.1399(13)	0.1098
O(4)	1.0521(12)	-0.03187(21)	0.1474(10)	0.0742
O(5)	0.8679(14)	-0.0270(3)	-0.0577(10)	0.1097

Table Fractional Coordinates of Atoms with Standard Deviations

	x	y	z	Ueq
H(2)	0.598(11)	-0.2335(25)	0.759(10)	0.0323
H(3)	0.453(10)	-0.2537(24)	0.524(10)	0.0323
H(4)	0.544(12)	-0.206(3)	0.294(11)	0.0323
H(5)	0.711(10)	-0.1522(22)	0.292(10)	0.0323
H(8)	0.955(10)	-0.1136(23)	0.811(9)	0.0323
H(9)	0.979(10)	-0.1192(23)	0.300(9)	0.0323
H(101)	0.903(9)	-0.0218(22)	0.486(9)	0.0323
H(102)	0.988(11)	-0.0469(22)	0.679(10)	0.0323
H(111)	0.713(11)	-0.0848(23)	0.675(10)	0.0323
H(112)	0.709(10)	-0.0802(22)	0.479(9)	0.0323
H(121)	1.193(11)	-0.1681(22)	0.428(10)	0.0323
H(122)	1.307(10)	-0.1308(21)	0.339(10)	0.0323
H(13)	1.405(10)	-0.1460(22)	0.652(10)	0.0323
H(14)	1.165(10)	-0.1845(23)	0.703(9)	0.0323
H(151)	1.350(10)	-0.0605(22)	0.410(10)	0.0323
H(152)	1.263(10)	-0.0249(21)	0.538(9)	0.0323
H(171)	1.369(10)	-0.0419(22)	0.829(9)	0.0323
H(121)	1.426(10)	-0.0756(22)	1.037(10)	0.0323
H(122)	1.514(10)	-0.1190(22)	0.906(10)	0.0323
H(19)	1.332(10)	-0.1771(21)	0.929(10)	0.0323
H(201)	1.041(11)	-0.1656(23)	1.124(11)	0.0323
H(202)	1.120(10)	-0.1992(22)	1.120(10)	0.0323
H(1N2)	1.051(11)	-0.0539(23)	0.357(10)	0.0323

Thermal Vibration Parameters with Standard Deviations

	U or U11	U22	U33	U23	U13	U12
C(1)	0.0301(52)	0.0278(44)	0.0691(67)	0.0077(45)	0.0100(51)	-0.0030(42)
C(2)	0.0422(61)	0.0436(61)	0.0839(87)	0.0043(58)	0.0023(67)	0.0105(55)
C(3)	0.0439(59)	0.0289(47)	0.1071(110)	-0.0088(63)	-0.0025(73)	-0.0011(49)
C(4)	0.0439(67)	0.0668(71)	0.0710(90)	-0.0238(71)	-0.0219(64)	0.0059(61)
C(5)	0.0299(49)	0.0396(49)	0.0654(71)	-0.0039(47)	-0.0036(51)	0.0040(48)
C(6)	0.0192(41)	0.0371(47)	0.0652(66)	-0.0045(45)	0.0113(48)	0.0040(39)
C(7)	0.0389(47)	0.0374(46)	0.0414(50)	-0.0063(44)	0.0002(46)	0.0112(43)
C(8)	0.0358(53)	0.0278(47)	0.0411(53)	-0.0109(39)	0.0066(43)	0.0030(41)
C(9)	0.0286(44)	0.0380(47)	0.0467(54)	-0.0065(46)	0.0005(44)	0.0017(41)
C(10)	0.0459(60)	0.0510(56)	0.0488(61)	0.0056(51)	-0.0076(53)	0.0078(50)
C(11)	0.0289(50)	0.0497(55)	0.0555(59)	0.0018(54)	-0.0057(49)	0.0085(47)
C(12)	0.0407(56)	0.0446(52)	0.0523(66)	-0.0071(48)	0.0034(55)	0.0059(49)
C(13)	0.0328(51)	0.0477(52)	0.0445(56)	0.0000(45)	-0.0089(45)	0.0171(47)
C(14)	0.0381(56)	0.0311(46)	0.0551(62)	-0.0137(44)	0.0025(50)	0.0072(43)
C(15)	0.0366(52)	0.0485(56)	0.0539(64)	0.0015(50)	-0.0009(52)	-0.0120(45)
C(16)	0.0269(45)	0.0408(46)	0.0505(59)	-0.0005(48)	-0.0004(45)	-0.0059(44)
C(17)	0.0479(60)	0.0566(58)	0.0608(71)	-0.0072(57)	0.0036(56)	-0.0114(59)
C(18)	0.0712(79)	0.0870(88)	0.0489(72)	-0.0026(62)	-0.0104(70)	-0.0171(71)
C(19)	0.0584(63)	0.0513(58)	0.0351(53)	0.0016(46)	-0.0058(54)	0.0037(55)
C(20)	0.1031(95)	0.0697(77)	0.0393(67)	0.0083(58)	-0.0161(72)	-0.0146(74)
C(21)	0.0622(65)	0.0370(50)	0.0612(66)	-0.0025(55)	0.0163(62)	0.0088(52)
N(1)	0.0539(50)	0.0263(35)	0.0486(49)	0.0036(35)	0.0042(42)	0.0004(36)
N(2)	0.0381(41)	0.0345(39)	0.0432(49)	0.0118(36)	-0.0074(41)	0.0017(36)
O(1)	0.0985(56)	0.0603(39)	0.0807(56)	0.0327(40)	-0.0015(53)	-0.0259(44)
O(2)	0.0663(45)	0.0551(38)	0.0549(42)	-0.0089(33)	-0.0041(39)	-0.0097(39)
N(3)	0.0852(79)	0.0751(67)	0.0414(57)	0.0001(51)	-0.0103(58)	0.0364(65)
O(3)	0.0763(61)	0.1391(80)	0.1141(82)	0.0281(68)	-0.0331(61)	-0.0291(64)
O(4)	0.1054(64)	0.0573(44)	0.0599(49)	0.0074(39)	-0.0114(52)	-0.0118(47)
O(5)	0.1519(90)	0.1298(72)	0.0475(50)	0.0058(52)	-0.0141(60)	0.0630(69)

CHAPTER NINE
BRUCINE ETHANOLATE DIHYDRATE

BRUCINE ETHANOL DIHYDRATE

Introduction

Brucine is an indole alkaloid of the strychnine family isolated from the seeds of *strychnos nux vomica*. The ring system in brucine is identical to that of strychnine both in connectivity and in stereochemistry, with the only difference between the molecules being the presence of two methoxy groups on the aromatic ring. This alteration results in a great difference in physiological effect between strychnine and brucine. Although the main effect of brucine is a paralysis caused by a blockage of the neuromuscular junction, it does have strychnine-like effects on the central nervous system when it is present in sufficiently high concentration. This would imply that the addition of the methoxy groups does not completely prevent binding of brucine to the strychnine receptor. Perhaps the most important use of brucine and strychnine is in the optical resolution of optically active acids or alcohols (Fischer, 1899; Toda and Tanaka, 1981) by selective crystallisation of the diastereomeric salts and subsequent alkaline treatment to afford the desired enantiomer as the free acid. Although strychnine and brucine have the same configuration at all of their chiral centres, there is a distinct difference in the relative solubilities of their D and L amino acid salts. In the majority of cases the brucine salt of the D amino acid is far less soluble than the corresponding L amino acid salt, while the opposite tends to hold true for the strychnine salts (Greenstein and Winitz, 1961). As well as providing information on the mechanisms involved in chiral recognition, study of the solid state structures of these complexes provide a useful model system for the examination of drug receptor interactions.

Crystallisation

Brucine ethanol dihydrate crystals were grown from a solution of brucine tetrahydrate in absolute ethanol that was allowed to evaporate slowly at room temperature. The crystals were in the form of transparent cuboids which crumbled to an opaque powder after about thirty minutes exposure to air, so it was necessary to mount the

crystal used for data collection in a glass tube containing a small amount of mother liquor to prevent its decomposition.

Crystal Data

Formula $C_{23}H_{26}N_2O_4 \cdot C_2H_5OH \cdot 2H_2O$ Mol. Wt. 476.5 $F(000) = 1024$

Space group $P2_12_12_1$ Int. Tab. No. 19

Cell dimensions $a = 7.723(1)$ $b = 12.337(1)$ $c = 25.212(2)$ Å

$V = 2402.8$ Å³ $Z = 4$ $D_c = 1.317$ gcm⁻³

Radiation MoK α $\lambda = 0.71069$ Å $\mu = 0.90$ cm⁻¹

Final R = 0.0387 based on 1684 independent data

$h_{max} = 9$ $k_{max} = 14$ $l_{max} = 29$

$\theta_{max} = 25^\circ$ $\sin\theta_{max}/\lambda = 0.5947$

2 intensity controls (0 6 1 and 0 0 12)

Drift curve min. 0.981 max. 1.049

2459 reflections measured (26 systematic absences)

749 reflections considered unobserved ($F < 4\sigma(F)$)

Weighting scheme: $W = 1.5864/(\sigma^2(F) + 0.001733F^2)$

401 parameters refined

Final difference Fourier max. 0.1439 min. -0.1436 eÅ⁻³

Max. shift/ $\sigma = 0.176$

Data collection and reduction

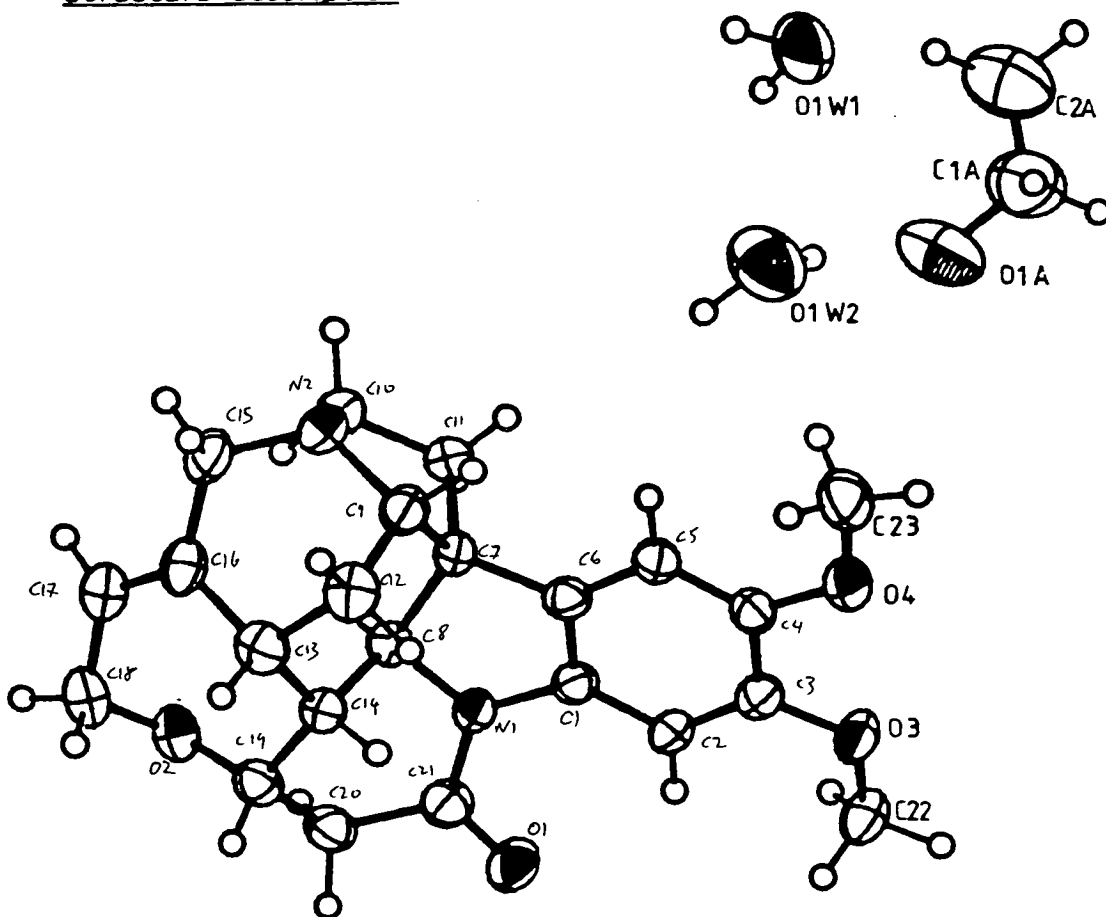
The data was collected on the CAD4 and corrected for Lorentz and polarisation effects using the program CADABS.

Solution and refinement

The structure was solved automatically using the MULTAN77 system of programs. The first electron density map to be examined contained the entire brucine skeleton with the exception of two carbon atoms. Difference Fourier maps calculated by the program SHELX were used to locate the remaining atoms and to refine the structure. The course of the least-squares refinement is summarised below.

R	Model
0.3954	$C_{21}N_2O_4$ as input fragment
0.1775	$C_{23}N_2O_4 \cdot 20 \cdot C_2O$ (all non-H atoms)
0.1057	All atoms isotropic, 14 H atoms input
0.0647	All non-H atoms anisotropic, 23 H atoms included
0.0489	33 H atoms included
0.0384	All H atoms included. H thermal parameters tied to separate free variables for brucine and for solvent molecules. Ethanol H atoms in calculated positions 793 reflections unobserved ($F < 5\sigma(F)$)
0.0401	$F < 4\sigma(F)$ 749 reflections unobserved
0.0387	Final Convergence

Structure description



The diagram above shows the additional numbering scheme used for this complex, carrying on from the scheme used to number the strychnine

*The only hydrogens to be constrained to calculated positions were those on the ethanol molecule. The thermal parameters were constrained in two sets, one for the brucine hydrogens, and one for the solvent hydrogens.

molecule and cation. All bond lengths and angles are within the expected ranges for their types, the range of carbon carbon single bonds in the brucine molecule being from 1.561(6) angstroms for C(7)-C(8) to 1.488(6) angstroms for C(17)-C(18) (a difference of 12 σ) although the carbon carbon bond in the ethanol is actually slightly shorter than the carbon oxygen bond (1.437(11) vs. 1.454(9) A). The range of values for the angles at nominally tetrahedral carbon is also rather broad (101.2(3) $^\circ$ for C(9)-C(7)-C(11) *versus* 119.3(4) for C(19)-C(20)-C(21) - a difference of almost 60 σ). The bonds from the aromatic ring to the methoxy groups are shorter than the usual carbon oxygen single bond length, possibly because of an overlap between the lone pairs on the oxygens with the π cloud of the aromatic ring. There is a pronounced alternation of bond lengths around the aromatic ring, which is likely to be caused either by the steric effect of the two oxygen substituents being in such close contact (O(3)-O(4) 2.587(4) angstroms compared with usual hydrogen bonded contact distance of 2.75 angstroms), or as a result of a possible canonical form of the structure with a double bond to one of the oxygens (this explanation also rationalises the difference between the two C-O bonds, as it would be possible for one of the two resonance hybrids to be fractionally stabilised by the amide group.). The other carbon oxygen bonds and the carbon nitrogen are within the expected ranges for their types as can be seen from the tables below.

C(1) - C(2)	1.407(6)	C(12) -C(13)	1.540(7)
C(1) - C(6)	1.371(6)	C(13) -C(14)	1.533(6)
C(1) - N(1)	1.416(5)	C(13) -C(16)	1.524(6)
C(2) - C(3)	1.375(6)	C(14) -C(19)	1.536(6)
C(3) - C(4)	1.425(6)	C(15) -C(16)	1.505(7)
C(3) - O(3)	1.353(5)	C(15) - N(2)	1.485(6)
C(4) - C(5)	1.372(6)	C(16) -C(17)	1.319(6)
C(4) - O(4)	1.374(5)	C(17) -C(18)	1.488(7)
C(5) - C(6)	1.394(6)	C(18) - O(2)	1.427(6)
C(6) - C(7)	1.512(6)	C(19) -C(20)	1.528(7)
C(7) - C(8)	1.561(6)	C(19) - O(2)	1.432(5)
C(7) - C(9)	1.550(6)	C(20) -C(21)	1.511(6)
C(7) -C(11)	1.539(6)	C(21) - N(1)	1.366(5)
C(8) -C(14)	1.522(6)	C(21) - O(1)	1.231(6)
C(8) - N(1)	1.485(5)	C(22) - O(3)	1.433(6)

C(9) - C(12)	1.508(6)	C(23) - O(4)	1.420(7)
C(9) - N(2)	1.494(6)	C(1A) - C(2A)	1.437(11)
C(10) - C(11)	1.512(7)	C(1A) - O(1A)	1.454(9)
C(10) - N(2)	1.474(6)		
C(2) - C(1) - C(6)	121.5(4)	C(12) - C(13) - C(14)	105.9(4)
C(2) - C(1) - N(1)	127.8(4)	C(12) - C(13) - C(16)	109.3(4)
C(6) - C(1) - N(1)	110.7(3)	C(14) - C(13) - C(16)	115.2(4)
C(1) - C(2) - C(3)	118.2(4)	C(8) - C(14) - C(13)	112.7(3)
C(2) - C(3) - C(4)	120.3(4)	C(8) - C(14) - C(19)	107.6(3)
C(2) - C(3) - O(3)	124.3(4)	C(13) - C(14) - C(19)	118.1(4)
C(4) - C(3) - O(3)	115.3(4)	C(16) - C(15) - N(2)	111.8(4)
C(3) - C(4) - C(5)	120.2(4)	C(13) - C(16) - C(15)	114.1(4)
C(3) - C(4) - O(4)	115.1(4)	C(13) - C(16) - C(17)	122.2(4)
C(5) - C(4) - O(4)	124.7(4)	C(15) - C(16) - C(17)	123.5(4)
C(4) - C(5) - C(6)	119.4(4)	C(16) - C(17) - C(18)	123.9(4)
C(1) - C(6) - C(5)	120.3(4)	C(17) - C(18) - O(2)	112.0(4)
C(1) - C(6) - C(7)	110.4(3)	C(14) - C(19) - C(20)	109.8(4)
C(5) - C(6) - C(7)	129.2(4)	C(14) - C(19) - O(2)	113.7(3)
C(6) - C(7) - C(8)	102.0(3)	C(20) - C(19) - O(2)	106.0(3)
C(6) - C(7) - C(9)	115.9(3)	C(19) - C(20) - C(21)	119.3(4)
C(6) - C(7) - C(11)	113.4(3)	C(20) - C(21) - N(1)	116.0(4)
C(8) - C(7) - C(9)	113.9(3)	C(20) - C(21) - O(1)	121.3(4)
C(8) - C(7) - C(11)	110.8(3)	N(1) - C(21) - O(1)	122.6(4)
C(9) - C(7) - C(11)	101.2(3)	C(1) - N(1) - C(8)	108.9(3)
C(7) - C(8) - C(14)	117.2(3)	C(1) - N(1) - C(21)	125.6(3)
C(7) - C(8) - N(1)	104.4(3)	C(8) - N(1) - C(21)	119.6(3)
C(14) - C(8) - N(1)	106.5(3)	C(9) - N(2) - C(10)	108.5(3)
C(7) - C(9) - C(12)	114.4(4)	C(9) - N(2) - C(15)	112.9(3)
C(7) - C(9) - N(2)	105.2(3)	C(10) - N(2) - C(15)	112.7(4)
C(12) - C(9) - N(2)	112.1(4)	C(18) - O(2) - C(19)	115.0(3)
C(11) - C(10) - N(2)	105.5(4)	C(3) - O(3) - C(22)	115.9(3)
C(7) - C(11) - C(10)	102.7(4)	C(4) - O(4) - C(23)	117.3(4)
C(9) - C(12) - C(13)	108.7(4)	C(2A) - C(1A) - O(1A)	112.8(6)

C(6) - C(1) - C(2) - C(3)	2.4(6)
N(1) - C(1) - C(2) - C(3)	-176.5(4)
C(2) - C(1) - C(6) - C(5)	-3.5(6)
C(2) - C(1) - C(6) - C(7)	172.9(4)
N(1) - C(1) - C(6) - C(5)	175.6(4)
N(1) - C(1) - C(6) - C(7)	-8.1(5)
C(2) - C(1) - N(1) - C(8)	174.2(4)
C(2) - C(1) - N(1) -C(21)	21.5(6)
C(6) - C(1) - N(1) - C(8)	-4.8(4)
C(6) - C(1) - N(1) -C(21)	-157.5(4)
C(1) - C(2) - C(3) - C(4)	0.9(6)
C(1) - C(2) - C(3) - O(3)	-179.7(4)
C(2) - C(3) - C(4) - C(5)	-3.2(6)
C(2) - C(3) - C(4) - O(4)	-177.1(4)
O(3) - C(3) - C(4) - C(5)	177.3(4)
O(3) - C(3) - C(4) - O(4)	-2.4(5)
C(2) - C(3) - O(3) -C(22)	4.0(6)
C(4) - C(3) - O(3) -C(22)	-176.5(4)
C(3) - C(4) - C(5) - C(6)	2.2(6)
O(4) - C(4) - C(5) - C(6)	-178.1(4)
C(3) - C(4) - O(4) -C(23)	-176.6(4)
C(5) - C(4) - O(4) -C(23)	3.7(6)
C(4) - C(5) - C(6) - C(1)	1.1(6)
C(4) - C(5) - C(6) - C(7)	-174.5(4)
C(1) - C(6) - C(7) - C(8)	16.6(4)
C(1) - C(6) - C(7) - C(9)	140.9(4)
C(1) - C(6) - C(7) -C(11)	-102.6(4)
C(5) - C(6) - C(7) - C(8)	-167.5(4)
C(5) - C(6) - C(7) - C(9)	-43.2(6)
C(5) - C(6) - C(7) -C(11)	73.3(5)
C(6) - C(7) - C(8) -C(14)	99.2(4)
C(6) - C(7) - C(8) - N(1)	-18.3(4)
C(9) - C(7) - C(8) -C(14)	-26.5(5)
C(9) - C(7) - C(8) - N(1)	-144.0(3)
C(11) - C(7) - C(8) -C(14)	-139.8(4)
C(11) - C(7) - C(8) - N(1)	102.7(4)
C(6) - C(7) - C(9) -C(12)	-82.3(5)
C(6) - C(7) - C(9) - N(2)	154.2(3)
C(8) - C(7) - C(9) -C(12)	35.6(5)
C(8) - C(7) - C(9) - N(2)	-87.8(4)
C(11) - C(7) - C(9) -C(12)	154.5(4)
C(11) - C(7) - C(9) - N(2)	31.1(4)
C(6) - C(7) -C(11) -C(10)	-165.5(4)
C(8) - C(7) -C(11) -C(10)	80.5(4)
C(9) - C(7) -C(11) -C(10)	-40.6(4)

C(7) - C(8) -C(14) -C(13)	40.5(5)
C(7) - C(8) -C(14) -C(19)	172.5(3)
N(1) - C(8) -C(14) -C(13)	156.9(3)
N(1) - C(8) -C(14) -C(19)	-71.1(4)
C(7) - C(8) - N(1) - C(1)	14.9(4)
C(7) - C(8) - N(1) -C(21)	169.5(3)
C(14) - C(8) - N(1) - C(1)	-109.7(3)
C(14) - C(8) - N(1) -C(21)	44.8(5)
C(7) - C(9) -C(12) -C(13)	-58.4(5)
N(2) - C(9) -C(12) -C(13)	61.3(5)
C(7) - C(9) - N(2) -C(10)	-9.8(4)
C(7) - C(9) - N(2) -C(15)	115.9(4)
C(12) - C(9) - N(2) -C(10)	-134.6(4)
C(12) - C(9) - N(2) -C(15)	-8.9(5)
N(2) -C(10) -C(11) - C(7)	35.8(5)
C(11) -C(10) - N(2) - C(9)	-16.3(5)
C(11) -C(10) - N(2) -C(15)	-142.1(4)
C(9) -C(12) -C(13) -C(14)	69.7(4)
C(9) -C(12) -C(13) -C(16)	-55.0(5)
C(12) -C(13) -C(14) - C(8)	-60.7(5)
C(12) -C(13) -C(14) -C(19)	172.8(4)
C(16) -C(13) -C(14) - C(8)	60.2(5)
C(16) -C(13) -C(14) -C(19)	-66.3(5)
C(12) -C(13) -C(16) -C(15)	-0.2(5)
C(12) -C(13) -C(16) -C(17)	176.0(4)
C(14) -C(13) -C(16) -C(15)	-119.3(4)
C(14) -C(13) -C(16) -C(17)	56.9(6)
C(8) -C(14) -C(19) -C(20)	46.9(5)
C(8) -C(14) -C(19) - O(2)	-71.7(4)
C(13) -C(14) -C(19) -C(20)	175.9(4)
C(13) -C(14) -C(19) - O(2)	57.3(5)
N(2) -C(15) -C(16) -C(13)	52.4(5)
N(2) -C(15) -C(16) -C(17)	-123.8(5)
C(16) -C(15) - N(2) - C(9)	-47.3(5)
C(16) -C(15) - N(2) -C(10)	76.1(5)
C(13) -C(16) -C(17) -C(18)	-2.5(7)
C(15) -C(16) -C(17) -C(18)	173.4(5)
C(16) -C(17) -C(18) - O(2)	-62.8(6)
C(17) -C(18) - O(2) -C(19)	87.5(5)
C(14) -C(19) -C(20) -C(21)	1.9(6)
O(2) -C(19) -C(20) -C(21)	125.1(4)
C(14) -C(19) - O(2) -C(18)	-69.1(5)
C(20) -C(19) - O(2) -C(18)	170.2(4)
C(19) -C(20) -C(21) - N(1)	-30.0(6)
C(19) -C(20) -C(21) - O(1)	148.6(4)

C(20) - C(21) - N(1) - C(1)	154.8(4)
C(20) - C(21) - N(1) - C(8)	4.8(5)
O(1) - C(21) - N(1) - C(1)	-23.8(6)
O(1) - C(21) - N(1) - C(8)	-173.8(4)

The next group of tables give the bonds, angles and torsion angles involving hydrogen atoms. As can be seen from the lower positional e.s.d.'s, the ethyl alcohol hydrogens were positionally constrained.

C(2) - H(021)	0.97(5)	C(20) - H(201)	1.02(5)
C(5) - H(051)	1.02(5)	C(20) - H(202)	0.92(5)
C(8) - H(081)	1.02(5)	C(22) - H(221)	1.03(5)
C(9) - H(091)	1.10(4)	C(22) - H(222)	1.00(5)
C(10) - H(101)	1.00(5)	C(22) - H(223)	1.04(4)
C(10) - H(102)	1.11(5)	C(23) - H(231)	0.94(5)
C(11) - H(111)	0.92(5)	C(23) - H(232)	0.99(5)
C(11) - H(112)	1.07(5)	C(23) - H(233)	0.95(5)
C(12) - H(121)	1.05(5)	O(1W1) - H(1W1)	0.92(11)
C(12) - H(122)	0.95(5)	O(1W1) - H(2W1)	0.96(11)
C(13) - H(131)	0.95(5)	O(1W2) - H(1W2)	0.96(11)
C(14) - H(141)	0.97(5)	O(1W2) - H(2W2)	1.03(11)
C(15) - H(151)	1.06(5)	C(1A) - H(1C1)	1.080(11)
C(15) - H(152)	0.98(5)	C(1A) - H(2C1)	1.080(11)
C(17) - H(171)	0.95(5)	C(2A) - H(1C2)	1.080(12)
C(18) - H(181)	0.99(5)	C(2A) - H(2C2)	1.080(12)
C(18) - H(182)	1.02(5)	C(2A) - H(3C2)	1.080(12)
C(19) - H(191)	1.00(5)	O(1A) - H(10A)	1.05(11)

C(1) - C(2) - H(021)	120.4(27)	C(17) - C(18) - H(182)	111.0(26)
H(021) - C(2) - C(3)	121.4(27)	H(181) - C(18) - H(182)	111 (4)
C(4) - C(5) - H(051)	121.7(26)	H(181) - C(18) - O(2)	110 (3)
H(051) - C(5) - C(6)	118.9(26)	H(182) - C(18) - O(2)	105.3(26)
C(7) - C(8) - H(081)	107.0(26)	C(14) - C(19) - H(191)	110 (3)
H(081) - C(8) - C(14)	111.0(26)	H(191) - C(19) - C(20)	105 (3)
H(081) - C(8) - N(1)	110.5(26)	H(191) - C(19) - O(2)	112 (3)
C(7) - C(9) - H(091)	107.1(23)	C(19) - C(20) - H(201)	108.4(26)
H(091) - C(9) - C(12)	108.9(23)	C(19) - C(20) - H(202)	100 (3)
H(091) - C(9) - N(2)	108.9(23)	H(201) - C(20) - H(202)	116 (4)
H(101) - C(10) - H(102)	113 (4)	H(201) - C(20) - C(21)	104.9(26)
H(101) - C(10) - C(11)	110 (3)	H(202) - C(20) - C(21)	109 (3)
H(101) - C(10) - N(2)	108 (3)	H(221) - C(22) - H(222)	106 (4)
H(102) - C(10) - C(11)	113.9(24)	H(221) - C(22) - H(223)	122 (4)
H(102) - C(10) - N(2)	106.5(23)	H(221) - C(22) - O(3)	108.6(27)

C(7) -C(11) -H(111)	108 (3)	H(222)-C(22) -H(223)	102 (4)
C(7) -C(11) -H(112)	112.4(25)	H(222)-C(22) - O(3)	109.6(27)
C(10) -C(11) -H(111)	114 (3)	H(223)-C(22) - O(3)	108.6(25)
C(10) -C(11) -H(112)	112.4(25)	H(231)-C(23) -H(232)	115 (4)
H(111)-C(11) -H(112)	107 (4)	H(231)-C(23) -H(233)	101 (4)
C(9) -C(12) -H(121)	111.8(26)	H(231)-C(23) - O(4)	113 (3)
C(9) -C(12) -H(122)	111 (3)	H(232)-C(23) -H(233)	111 (4)
H(121)-C(12) -H(122)	108 (4)	H(232)-C(23) - O(4)	110.4(27)
H(121)-C(12) -C(13)	106.5(26)	H(233)-C(23) - O(4)	106 (3)
H(122)-C(12) -C(13)	111 (3)	H(1W1)-O(1W1)-H(2W1)	81 (9)
C(12) -C(13) -H(131)	111 (3)	H(1W2)-O(1W2)-H(2W2)	110 (9)
H(131)-C(13) -C(14)	108 (3)	H(1C1)-C(1A) -H(2C1)	109.5(9)
H(131)-C(13) -C(16)	108 (3)	H(1C1)-C(1A) -C(2A)	108.6(8)
C(8) -C(14) -H(141)	102.8(27)	H(1C1)-C(1A) -O(1A)	108.6(7)
C(13) -C(14) -H(141)	107.6(27)	H(2C1)-C(1A) -C(2A)	108.6(8)
H(141)-C(14) -C(19)	106.8(27)	H(2C1)-C(1A) -O(1A)	108.6(7)
H(151)-C(15) -H(152)	99 (4)	C(1A) -C(2A) -H(1C2)	109.5(8)
H(151)-C(15) -C(16)	111.6(25)	C(1A) -C(2A) -H(2C2)	109.5(8)
H(151)-C(15) - N(2)	111.0(25)	C(1A) -C(2A) -H(3C2)	109.5(8)
H(152)-C(15) -C(16)	118 (3)	H(1C2)-C(2A) -H(2C2)	109.5(10)
H(152)-C(15) - N(2)	104 (3)	H(1C2)-C(2A) -H(3C2)	109.5(10)
C(16) -C(17) -H(171)	117 (3)	H(2C2)-C(2A) -H(3C2)	109.5(10)
H(171)-C(17) -C(18)	118 (3)	C(1A) -O(1A) -H(10A)	113 (6)
C(17) -C(18) -H(181)	108 (3)		

C(6) - C(1) - C(2) -H(021)	180 (3)
N(1) - C(1) - C(2) -H(021)	1 (3)
H(021)- C(2) - C(3) - C(4)	-177 (3)
H(021)- C(2) - C(3) - O(3)	3 (3)
C(3) - C(4) - C(5) -H(051)	-179 (3)
O(4) - C(4) - C(5) -H(051)	1 (3)
H(051)- C(5) - C(6) - C(1)	-178 (3)
H(051)- C(5) - C(6) - C(7)	7 (3)
C(6) - C(7) - C(8) -H(081)	-136 (3)
C(9) - C(7) - C(8) -H(081)	99 (3)
C(11) - C(7) - C(8) -H(081)	-14 (3)
C(6) - C(7) - C(9) -H(091)	38.4(25)
C(8) - C(7) - C(9) -H(091)	156.4(24)
C(11) - C(7) - C(9) -H(091)	-84.7(25)
C(6) - C(7) -C(11) -H(111)	-44 (3)
C(6) - C(7) -C(11) -H(112)	73.4(27)
C(8) - C(7) -C(11) -H(111)	-158 (3)
C(8) - C(7) -C(11) -H(112)	-40.6(27)
C(9) - C(7) -C(11) -H(111)	81 (3)
C(9) - C(7) -C(11) -H(112)	-161.7(27)

C(7) - C(8) -C(14) -H(141)	-75 (3)
H(081)- C(8) -C(14) -C(13)	-83 (3)
H(081)- C(8) -C(14) -H(141)	162 (4)
H(081)- C(8) -C(14) -C(19)	49 (3)
N(1) - C(8) -C(14) -H(141)	41 (3)
H(081)- C(8) - N(1) - C(1)	130 (3)
H(081)- C(8) - N(1) -C(21)	-76 (3)
C(7) - C(9) -C(12) -H(121)	59 (3)
C(7) - C(9) -C(12) -H(122)	180 (3)
H(091)- C(9) -C(12) -H(121)	-61 (4)
H(091)- C(9) -C(12) -H(122)	60 (4)
H(091)- C(9) -C(12) -C(13)	-178.1(25)
N(2) - C(9) -C(12) -H(121)	179 (3)
N(2) - C(9) -C(12) -H(122)	-61 (3)
H(091)- C(9) - N(2) -C(10)	104.7(25)
H(091)- C(9) - N(2) -C(15)	-129.5(25)
H(101)-C(10) -C(11) - C(7)	-80 (3)
H(101)-C(10) -C(11) -H(111)	164 (4)
H(101)-C(10) -C(11) -H(112)	41 (4)
H(102)-C(10) -C(11) - C(7)	152.3(26)
H(102)-C(10) -C(11) -H(111)	36 (4)
H(102)-C(10) -C(11) -H(112)	-87 (4)
N(2) -C(10) -C(11) -H(111)	-81 (3)
N(2) -C(10) -C(11) -H(112)	156.9(27)
H(101)-C(10) - N(2) - C(9)	101 (3)
H(101)-C(10) - N(2) -C(15)	-25 (3)
H(102)-C(10) - N(2) - C(9)	-137.6(24)
H(102)-C(10) - N(2) -C(15)	96.6(25)
C(9) -C(12) -C(13) -H(131)	-174 (3)
H(121)-C(12) -C(13) -H(131)	66 (4)
H(121)-C(12) -C(13) -C(14)	-50.9(27)
H(121)-C(12) -C(13) -C(16)	-175.6(27)
H(122)-C(12) -C(13) -H(131)	-51 (4)
H(122)-C(12) -C(13) -C(14)	-168 (3)
H(122)-C(12) -C(13) -C(16)	67 (3)
C(12) -C(13) -C(14) -H(141)	52 (3)
H(131)-C(13) -C(14) - C(8)	-180 (3)
H(131)-C(13) -C(14) -H(141)	-67 (4)
H(131)-C(13) -C(14) -C(19)	54 (3)
C(16) -C(13) -C(14) -H(141)	173 (3)
H(131)-C(13) -C(16) -C(15)	121 (3)
H(131)-C(13) -C(16) -C(17)	-63 (3)
C(8) -C(14) -C(19) -H(191)	162 (3)
C(13) -C(14) -C(19) -H(191)	-69 (3)
H(141)-C(14) -C(19) -H(191)	52 (4)

H(141)-C(14) -C(19) -C(20)	-63 (3)
H(141)-C(14) -C(19) - O(2)	179 (3)
H(151)-C(15) -C(16) -C(13)	-72.6(27)
H(151)-C(15) -C(16) -C(17)	111.2(27)
H(152)-C(15) -C(16) -C(13)	173 (3)
H(152)-C(15) -C(16) -C(17)	-3 (3)
H(151)-C(15) - N(2) - C(9)	78.0(27)
H(151)-C(15) - N(2) -C(10)	-158.6(27)
H(152)-C(15) - N(2) - C(9)	-176 (3)
H(152)-C(15) - N(2) -C(10)	-53 (3)
C(13) -C(16) -C(17) -H(171)	-172 (3)
C(15) -C(16) -C(17) -H(171)	4 (3)
C(16) -C(17) -C(18) -H(181)	58 (3)
C(16) -C(17) -C(18) -H(182)	179.8(27)
H(171)-C(17) -C(18) -H(181)	-133 (4)
H(171)-C(17) -C(18) -H(182)	-11 (4)
H(171)-C(17) -C(18) - O(2)	107 (3)
H(181)-C(18) - O(2) -C(19)	-32 (3)
H(182)-C(18) - O(2) -C(19)	-151.8(26)
C(14) -C(19) -C(20) -H(201)	121.7(27)
C(14) -C(19) -C(20) -H(202)	-117 (3)
H(191)-C(19) -C(20) -H(201)	4 (4)
H(191)-C(19) -C(20) -H(202)	125 (4)
H(191)-C(19) -C(20) -C(21)	-116 (3)
O(2) -C(19) -C(20) -H(201)	-115.1(27)
O(2) -C(19) -C(20) -H(202)	6 (3)
H(191)-C(19) - O(2) -C(18)	56 (3)
H(201)-C(20) -C(21) - N(1)	-151.6(27)
H(201)-C(20) -C(21) - O(1)	26.9(27)
H(202)-C(20) -C(21) - N(1)	84 (3)
H(202)-C(20) -C(21) - O(1)	-98 (3)
H(221)-C(22) - O(3) - C(3)	52 (3)
H(222)-C(22) - O(3) - C(3)	-64 (3)
H(223)-C(22) - O(3) - C(3)	-173.8(26)
H(231)-C(23) - O(4) - C(4)	-59 (3)
H(232)-C(23) - O(4) - C(4)	171 (3)
H(233)-C(23) - O(4) - C(4)	51 (3)
H(1C1)-C(1A) -C(2A) -H(1C2)	-59.5(11)
H(1C1)-C(1A) -C(2A) -H(2C2)	-179.5(9)
H(1C1)-C(1A) -C(2A) -H(3C2)	60.5(11)
H(2C1)-C(1A) -C(2A) -H(1C2)	59.5(11)
H(2C1)-C(1A) -C(2A) -H(2C2)	-60.5(11)
H(2C1)-C(1A) -C(2A) -H(3C2)	179.5(9)
O(1A) -C(1A) -C(2A) -H(1C2)	-180.0(8)
O(1A) -C(1A) -C(2A) -H(2C2)	60.0(10)

O(1A) -C(1A) -C(2A) -H(3C2)	-60.0(10)
H(1C1)-C(1A) -O(1A) -H(10A)	-73 (7)
H(2C1)-C(1A) -O(1A) -H(10A)	168 (7)
C(2A) -C(1A) -O(1A) -H(10A)	47 (7)

Hydrogen bonding and intermolecular contacts

As can be seen in the packing diagrams overleaf, there are "infinite" channels running parallel to the b axis of the crystal. These channels contain sets of solvent molecules which form short hydrogen bonded chains between the carbonyl group of one molecule and the amine lone pair on a molecule related to the first one by the screw axis parallel to z. The hydrogen bonded contacts are shown in the table below.

O1	2.875	(6)	O1W2	(1 1/2-x, 2 -y, -1/2+z)
O1W1	2.781	(6)	N2	(1/2+x, 1 1/2-y, 1 -z)
O1W1	2.736	(7)	O1A	(1 1/2+x, 1 1/2-y, 1 -z)
O1W1	2.847	(7)	O1W2	(x, y, z)
O1W2	2.800	(8)	O1A	(1 +x, 1 +y, 1 +z)
H1W1	1.89	(11)	N2	(1/2+x, 1 1/2-y, 1 -z)
H2W1	2.24	(11)	O1W2	(x, y, z)
H1W2	1.89	(11)	O1A	(1 +x, 1 +y, 1 +z)
H2W2	1.85	(11)	O1	(1 1/2-x, 2 -y, 1/2+z)
H10A	1.72	(11)	O1W1	(-1 1/2+x, 1 1/2-y, 1 -z)

There are a number of other contact distances involving oxygen, but no carbon to carbon or nitrogen contacts less than 3.5 Å. There is a possible interaction between the ether oxygen of the 7 membered ring (O(2)) and the π cloud of the aromatic ring on an adjacent brucine molecule, and another between a methoxy oxygen and the electron density on the amide nitrogen.

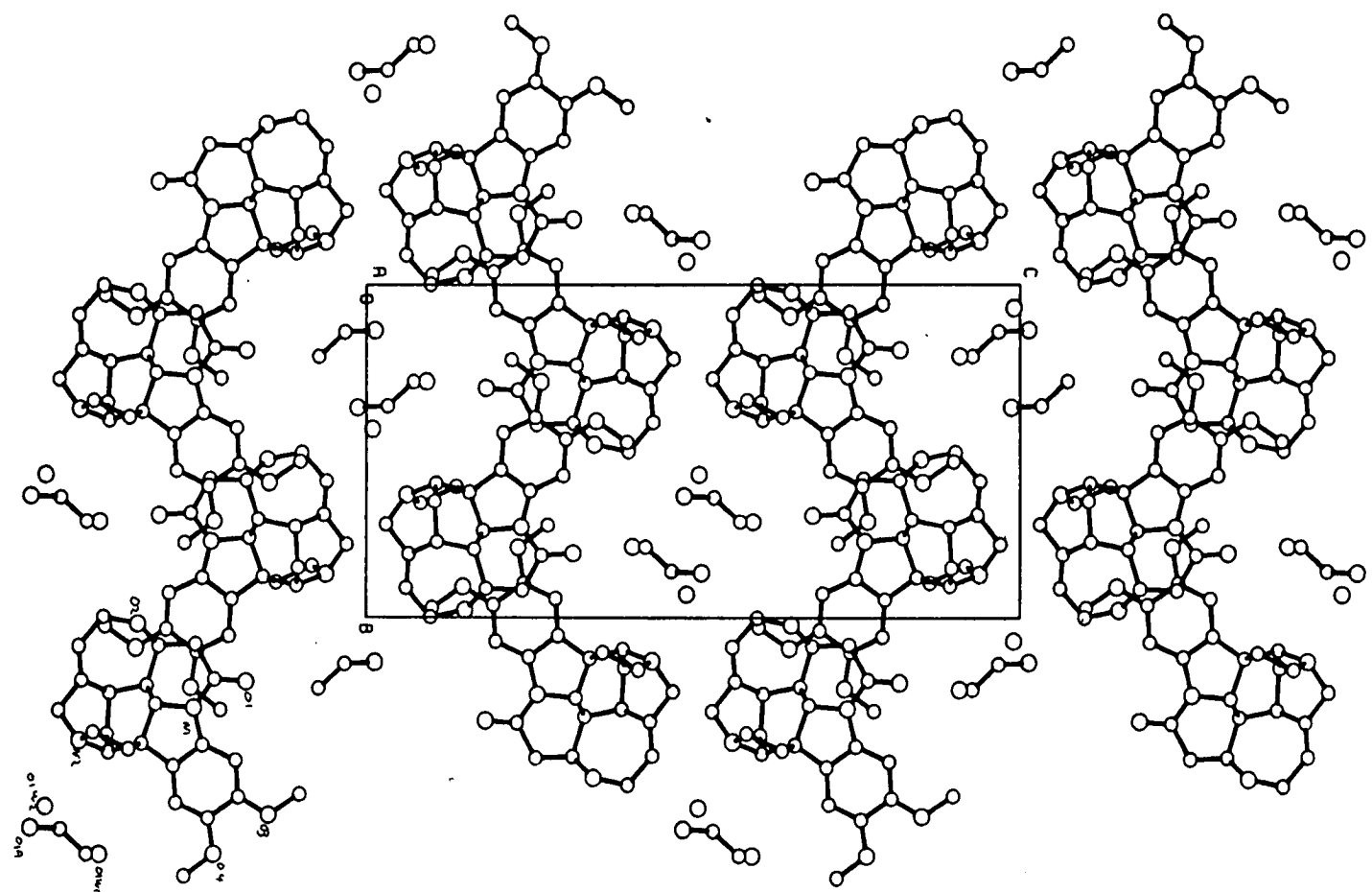
O2	3.488	(5)	C2	(1 -x, 1/2+y, 1/2-z)
O2	3.459	(5)	C3	(1 -x, 1/2+y, 1/2-z)
O2	3.483	(6)	C22	(1 -x, 1/2+y, 1/2-z)
O3	3.459	(5)	C13	(2 -x, -1/2+y, 1/2-z)
O4	3.372	(4)	N1	(1 -x, -1/2+y, 1/2-z)
O1W1	3.343	(6)	C9	(1/2+x, 1 1/2-y, 1 -z)
O1W1	3.324	(7)	C22	(2 -x, 1/2+y, 1 1/2-z)
O1A	3.213	(8)	C22	(1/2-x, 1 -y, -1/2+z)

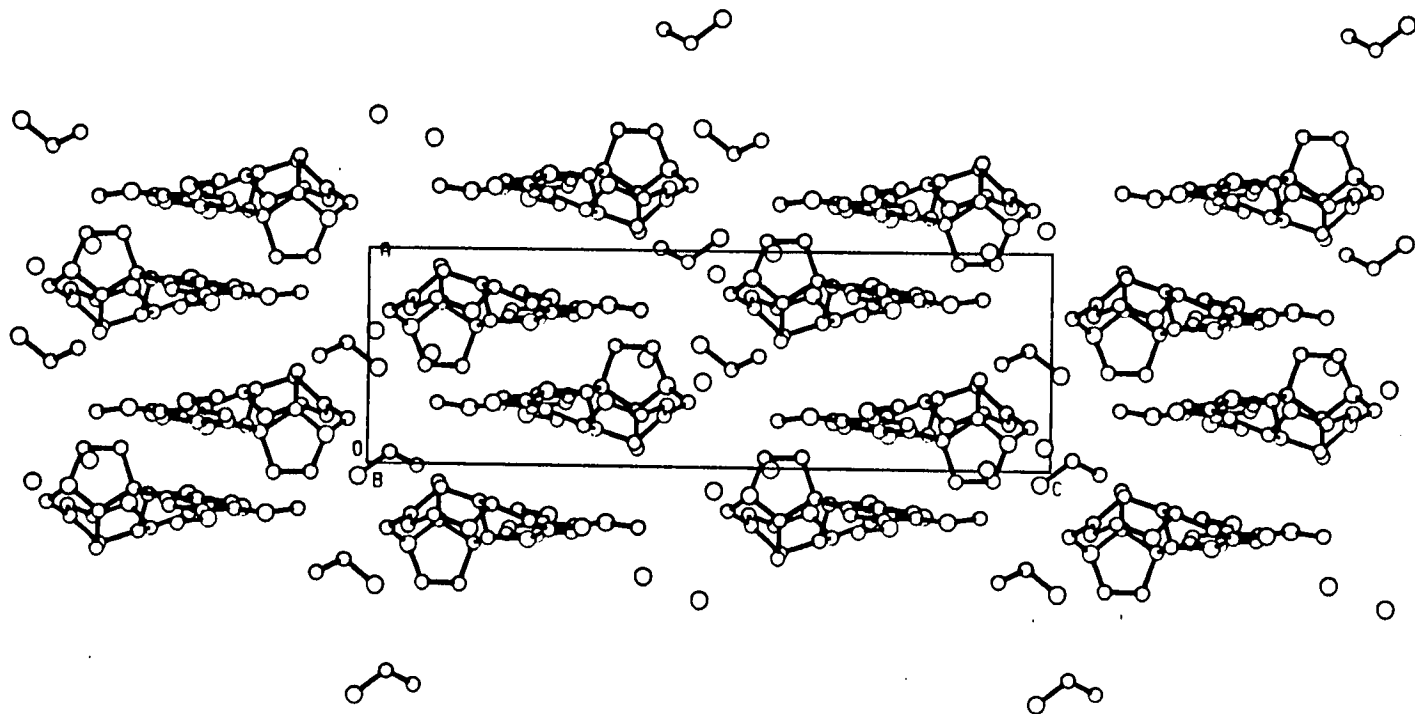
The table below shows all the hydrogen contacts less than 3 angstroms (or less than 2.5 angstroms to other hydrogens).

H051	2.99	(5)	O1W2	(-1/2+x, 1 1/2-y, 1 -z)
H081	2.79	(5)	O3	(1 -x, 1/2+y, 1/2-z)

H091	2.84	(5)	O1W1	$(-1/2+x, 1/2-y, 1-z)$
H112	2.70	(5)	O4	$(1-x, 1/2+y, 1/2-z)$
H131	2.63	(5)	O3	$(2-x, 1/2+y, 1/2-z)$
H141	2.73	(5)	O4	$(2-x, 1/2+y, 1/2-z)$
H171	2.99	(5)	O1A	$(1/2+x, 1/2-y, -z)$
H182	2.91	(5)	O1A	$(1/2+x, 1/2-y, -z)$
H191	2.74	(5)	O3	$(2-x, 1/2+y, 1/2-z)$
H221	2.67	(5)	O2	$(1-x, -1/2+y, 1/2-z)$
H221	2.74	(5)	O1A	$(1/2-x, 1-y, 1/2+z)$
H222	2.86	(5)	O1W1	$(2-x, -1/2+y, 1/2-z)$
H223	2.89	(4)	O1W1	$(2-x, -1/2+y, 1/2-z)$
H1C2	2.535	(9)	O1	$(1-x, -1/2+y, 1/2-z)$
H2C2	2.573	(10)	O1W1	$(-1+x, -1+y, -1+z)$
H3C2	2.775	(10)	O1W2	$(-1/2+x, 1/2-y, 1-z)$
H2W1	2.96	(11)	N2	$(1/2+x, 1/2-y, 1-z)$
H131	2.97	(5)	C22	$(2-x, 1/2+y, 1/2-z)$
H191	2.68	(5)	C3	$(2-x, 1/2+y, 1/2-z)$
H191	2.99	(5)	C4	$(2-x, 1/2+y, 1/2-z)$
H202	2.96	(5)	C3	$(1-x, 1/2+y, 1/2-z)$
H202	2.99	(5)	C4	$(1-x, 1/2+y, 1/2-z)$
H221	2.93	(5)	C17	$(1-x, -1/2+y, 1/2-z)$
H221	2.83	(5)	C18	$(1-x, -1/2+y, 1/2-z)$
H232	2.95	(5)	C20	$(x, -1+y, z)$
H1W1	2.62	(11)	C9	$(1/2+x, 1/2-y, 1-z)$
H1W1	2.64	(11)	C10	$(1/2+x, 1/2-y, 1-z)$
H1W1	2.89	(11)	C15	$(1/2+x, 1/2-y, 1-z)$
H1W2	2.91	(11)	C1A	$(1+x, 1+y, 1+z)$
H2W2	2.87	(11)	C21	$(1/2-x, 2-y, 1/2+z)$
H091	2.36	(12)	H1W1	$(-1/2+x, 1/2-y, 1-z)$
H102	2.45	(12)	H1W1	$(-1/2+x, 1/2-y, 1-z)$
H122	2.49	(5)	H1C1	$(1+x, y, z)$
H151	2.34	(6)	H223	$(1/2-x, 1-y, -1/2+z)$
H181	2.33	(7)	H222	$(2-x, 1/2+y, 1/2-z)$
H201	2.50	(6)	H232	$(x, 1+y, z)$
H1W1	2.37	(16)	H10A	$(1/2+x, 1/2-y, 1-z)$
H2W1	2.37	(15)	H1W2	(x, y, z)
H1W2	2.42	(15)	H10A	$(1+x, 1+y, 1+z)$

PROJECTION DOWN THE X AXIS





PROJECTION DOWN THE Y AXIS

Comparison with Strychnine

The most obvious difference between the brucine and strychnine molecules is the change in the aromatic ring caused by the methoxy groups. The range of bond lengths is fractionally larger in the brucine molecule, although the longest and shortest bonds are still in the same positions (implying that the difference in length essential for the molecular geometry to be maintained, and not merely random deviation), and the same is true of the largest and smallest angles. There is, of course, no hydrogen bond to the amine in the strychnine structure, which may explain the barely significant difference in length of the C(15)-N(2) bond between the two structures (1.485(6) angstroms in brucine and 1.464(6) angstroms in strychnine. In fact, the differences between the two structures are so small, that when one is fitted to the other, the root mean square deviation from an ideal fit is only 0.05 A. There is a somewhat larger deviation (ca. 0.15 A) in the position of the carbonyl group, almost certainly due to the hydrogen bonding to the carbonyl oxygen, which does not occur in the strychnine free base or in any of the salts and complexes examined so far.

Table Fractional Coordinates of Atoms with Standard Deviations

	x	y	z	Ueq
C(1)	0.6761(5)	0.6174(3)	0.25061(15)	0.0343
C(2)	0.6936(6)	0.5719(3)	0.30152(16)	0.0351
C(3)	0.7208(6)	0.4620(3)	0.30538(16)	0.0378
C(4)	0.7331(6)	0.3972(3)	0.25874(16)	0.0370
C(5)	0.7096(6)	0.4431(3)	0.20973(16)	0.0369
C(6)	0.6793(5)	0.5543(3)	0.20587(15)	0.0349
C(7)	0.6355(5)	0.6214(3)	0.15753(15)	0.0327
C(8)	0.6613(6)	0.7393(3)	0.17854(15)	0.0329
C(9)	0.7405(6)	0.5956(3)	0.10670(16)	0.0399
C(10)	0.4536(7)	0.6545(4)	0.08295(19)	0.0470
C(11)	0.4491(6)	0.6045(4)	0.13770(18)	0.0404
C(12)	0.9157(6)	0.6492(4)	0.10435(20)	0.0450
C(13)	0.8909(6)	0.7729(4)	0.10821(17)	0.0418
C(14)	0.8313(6)	0.7953(3)	0.16514(16)	0.0360
C(15)	0.7007(8)	0.7198(4)	0.03016(17)	0.0491
C(16)	0.7663(7)	0.8100(3)	0.06501(16)	0.0444
C(17)	0.7139(7)	0.9115(4)	0.06097(18)	0.0481
C(18)	0.7639(8)	0.9989(4)	0.09865(20)	0.0507
C(19)	0.8165(6)	0.9139(3)	0.18324(17)	0.0405
C(20)	0.7429(7)	0.9186(3)	0.23944(19)	0.0450
C(21)	0.7030(6)	0.8141(3)	0.26826(17)	0.0420
C(22)	0.7170(8)	0.4709(4)	0.39896(17)	0.0485
C(23)	0.7918(10)	0.2225(4)	0.22218(21)	0.0609
N(1)	0.6565(4)	0.72821(25)	0.23716(12)	0.0351
N(2)	0.6267(5)	0.6293(3)	0.06173(13)	0.0437
O(1)	0.7157(6)	0.80730(24)	0.31682(12)	0.0590
O(2)	0.7030(4)	0.97786(22)	0.15108(12)	0.0459
O(3)	0.7390(4)	0.40781(22)	0.35170(10)	0.0447
O(4)	0.7685(4)	0.28946(22)	0.26745(11)	0.0481
O(1W1)	1.1081(7)	1.0689(4)	0.99056(16)	0.0790
O(1W2)	1.0067(7)	1.2123(4)	0.90774(20)	0.0975
C(1A)	0.0511(11)	0.3618(6)	0.0326(3)	0.0962
C(2A)	-0.0108(11)	0.2882(7)	0.0725(3)	0.1093
O(1A)	-0.0615(8)	0.3667(4)	-0.01355(18)	0.1006

Table Fractional Coordinates of Atoms with Standard Deviations

	x	y	z	Ueq
H(021)	0.691(6)	0.618(4)	0.3328(17)	0.0477
H(051)	0.716(7)	0.398(4)	0.1759(17)	0.0477
H(081)	0.558(7)	0.784(4)	0.1655(16)	0.0477
H(091)	0.758(6)	0.507(4)	0.1053(16)	0.0477
H(101)	0.443(6)	0.735(4)	0.0856(17)	0.0477
H(102)	0.359(6)	0.619(3)	0.0549(17)	0.0477
H(111)	0.427(6)	0.531(4)	0.1381(18)	0.0477
H(112)	0.356(6)	0.642(4)	0.1629(18)	0.0477
H(121)	0.994(6)	0.627(4)	0.1365(18)	0.0477
H(122)	0.975(6)	0.631(4)	0.0726(18)	0.0477
H(131)	0.997(6)	0.810(4)	0.1029(17)	0.0477
H(141)	0.912(6)	0.760(4)	0.1887(17)	0.0477
H(151)	0.798(7)	0.691(3)	0.0040(16)	0.0477
H(152)	0.610(6)	0.737(4)	0.0040(20)	0.0477
H(171)	0.625(6)	0.926(3)	0.0360(18)	0.0477
H(181)	0.892(7)	1.004(4)	0.0988(18)	0.0477
H(182)	0.709(6)	1.071(4)	0.0881(16)	0.0477
H(191)	0.934(7)	0.947(4)	0.1861(17)	0.0477
H(201)	0.830(6)	0.958(4)	0.2631(18)	0.0477
H(202)	0.641(7)	0.955(4)	0.2329(18)	0.0477
H(221)	0.601(7)	0.512(4)	0.3965(18)	0.0477
H(222)	0.809(7)	0.528(4)	0.4008(17)	0.0477
H(223)	0.747(6)	0.423(3)	0.4315(18)	0.0477
H(231)	0.694(7)	0.222(4)	0.1999(19)	0.0477
H(232)	0.835(6)	0.150(4)	0.2328(16)	0.0477
H(233)	0.874(6)	0.259(4)	0.2004(19)	0.0477
H(1W1)	1.091(16)	1.004(8)	0.973(4)	0.1755
H(2W1)	1.004(14)	1.087(9)	0.972(4)	0.1755
H(1W2)	0.952(15)	1.259(8)	0.933(4)	0.1755
H(2W2)	0.932(15)	1.210(8)	0.874(4)	0.1755
H(1C1)	0.0599(11)	0.4420(6)	0.0497(3)	0.1755
H(2C1)	0.1779(11)	0.3356(6)	0.0200(3)	0.1755
H(1C2)	0.0777(11)	0.2880(7)	0.1057(3)	0.1755
H(2C2)	-0.0192(11)	0.2075(7)	0.0561(3)	0.1755
H(3C2)	-0.1372(11)	0.3138(7)	0.0858(3)	0.1755
H(10A)	-0.192(16)	0.376(9)	-0.004(4)	0.1755

Thermal Vibration Parameters with Standard Deviations

	U or U11	U22	U33	U23	U13	U12
C(1)	0.0343(22)	0.0352(21)	0.0333(21)	-0.0013(18)	-0.0049(18)	0.0010(19)
C(2)	0.0397(23)	0.0396(23)	0.0259(20)	0.0008(18)	-0.0011(19)	-0.0002(20)
C(3)	0.0368(24)	0.0443(24)	0.0322(22)	0.0016(18)	-0.0010(19)	-0.0008(21)
C(4)	0.0354(23)	0.0332(20)	0.0425(23)	0.0037(18)	-0.0047(19)	-0.0036(19)
C(5)	0.0427(25)	0.0335(21)	0.0345(22)	-0.0017(18)	-0.0001(20)	-0.0005(21)
C(6)	0.0351(23)	0.0392(23)	0.0303(20)	-0.0025(18)	0.0010(18)	-0.0042(19)
C(7)	0.0392(24)	0.0309(21)	0.0281(20)	0.0047(18)	-0.0004(19)	-0.0015(18)
C(8)	0.0336(24)	0.0316(22)	0.0333(22)	-0.0008(17)	0.0023(18)	0.0043(18)
C(9)	0.0518(29)	0.0355(22)	0.0325(21)	0.0005(18)	-0.0018(21)	0.0051(22)
C(10)	0.0527(30)	0.0484(29)	0.0400(26)	0.0027(22)	-0.0083(23)	-0.0008(25)
C(11)	0.0442(25)	0.0385(24)	0.0385(24)	0.0010(21)	-0.0073(20)	-0.0035(22)
C(12)	0.0471(29)	0.0456(26)	0.0423(27)	0.0001(23)	0.0110(24)	0.0091(23)
C(13)	0.0371(25)	0.0482(26)	0.0400(24)	-0.0020(21)	0.0118(21)	-0.0040(23)
C(14)	0.0390(25)	0.0333(22)	0.0357(22)	0.0017(18)	-0.0042(19)	-0.0007(20)
C(15)	0.0681(34)	0.0500(24)	0.0291(22)	0.0051(21)	0.0014(25)	-0.0001(28)
C(16)	0.0509(29)	0.0510(26)	0.0314(21)	0.0077(18)	0.0103(22)	0.0003(24)
C(17)	0.0540(30)	0.0472(27)	0.0432(26)	0.0123(21)	0.0026(24)	0.0005(25)
C(18)	0.0556(32)	0.0414(24)	0.0552(30)	0.0133(22)	0.0072(25)	-0.0042(23)
C(19)	0.0352(25)	0.0381(24)	0.0481(25)	0.0003(19)	-0.0018(20)	-0.0033(21)
C(20)	0.0482(29)	0.0338(22)	0.0531(28)	-0.0030(20)	-0.0084(24)	-0.0015(22)
C(21)	0.0488(28)	0.0374(22)	0.0399(26)	-0.0043(19)	-0.0020(22)	0.0054(21)
C(22)	0.0587(33)	0.0545(28)	0.0323(24)	0.0065(21)	0.0008(24)	0.0044(29)
C(23)	0.0913(45)	0.0360(25)	0.0554(30)	-0.0009(23)	0.0003(32)	0.0023(31)
N(1)	0.0448(22)	0.0322(18)	0.0283(17)	0.0020(15)	-0.0025(16)	-0.0008(16)
N(2)	0.0539(24)	0.0477(22)	0.0296(19)	0.0019(17)	-0.0016(18)	-0.0032(19)
D(1)	0.0969(29)	0.0470(17)	0.0330(17)	-0.0045(13)	-0.0082(18)	0.0030(21)
D(2)	0.0481(19)	0.0378(15)	0.0517(18)	0.0068(14)	0.0041(16)	0.0032(15)
D(3)	0.0566(19)	0.0452(15)	0.0323(15)	0.0083(13)	-0.0019(15)	0.0059(15)
D(4)	0.0630(21)	0.0340(15)	0.0474(17)	0.0027(13)	-0.0056(17)	0.0040(16)
D(1W1)	0.1119(36)	0.0646(24)	0.0605(24)	-0.0182(20)	-0.0127(24)	0.0092(28)
D(1W2)	0.1064(40)	0.1008(34)	0.0853(31)	-0.0009(29)	-0.0211(28)	0.0248(31)
C(1A)	0.1214(61)	0.0921(47)	0.0750(42)	0.0061(40)	0.0004(45)	0.0066(50)
C(2A)	0.1226(65)	0.1292(60)	0.0761(44)	0.0103(48)	0.0096(46)	0.0195(56)
D(1A)	0.1193(40)	0.1201(37)	0.0623(25)	0.0121(26)	-0.0003(28)	0.0272(36)

CHAPTER TEN
STRYCHNINE N-ACETYL TYROSINE TETRAHYDRATE

STRYCHNINE:N-ACETYL-L-TYROSINE TETRAHYDRATE

Introduction

This salt was prepared as part of a study of molecular recognition in amino acid:alkaloid complexes. These complexes provide information about recognition phenomena on two levels. They provide direct information on chiral recognition, as they have long been used as a means of resolving amino acids via diastereomeric salt formation and they provide indirect information on drug:receptor interaction as the alkaloids used are biologically active and would be expected to interact with a receptor protein in the same ways as they interact with an acyl amino acid. The interactions possible in this salt may be divided into hydrogen bonding and hydrophobic interactions. The strychninium cation has one possible hydrogen bond donor (the proton on the quaternary nitrogen) and one possible hydrogen bond acceptor (the carbonyl oxygen). The N-acetyl-L-tyrosine anion has two possible hydrogen bond donors (the amide and phenol hydrogens) and four possible hydrogen bond acceptors (the carboxylate group, the amide carbonyl and the phenol oxygen). There are three possible hydrophobic interactions: a stacking interaction between the aromatic rings, formation of a 'hydrophobic pocket' for one of the molecules by the packing of the other, or an interaction between the π -cloud of an aromatic ring, and the lone pair electrons of a nearby oxygen atom. A charge transfer interaction between the two aromatic systems is not impossible, but was considered to be rather unlikely.

Although tyrosine has been resolved by the method of diastereomeric salt formation on previous occasions, this is the only complex between strychnine and a tyrosine derivative to have been prepared, the earlier workers having used brucine or cinchonine. Fischer (1900) resolved N-benzoyl tyrosine by using brucine to precipitate the L-isomer and cinchonine to precipitate the D-form. Because hydrolysis of the benzoyl group required reaction conditions strong enough to epimerise the chiral centre, a resolution system using a more easily removed acyl group was sought. Both N-formyl tyrosine (Abderhalden and Sichel, 1923)

and N-acetyl tyrosine (Sealock, 1946) were resolved with brucine, the D-isomer in both cases being the first to precipitate.

It was decided to attempt the preparation of a complex of N-acetyl L-tyrosine with strychnine because this is directly comparable with the strychnine: N-benzoylalanine dihydrate complex previously determined (Gould and Walkinshaw, 1984) in terms of hydrogen bonding ability, but with the opposite arrangement of steric bulk around the C α . As most resolutions using strychnine precipitate the L-isomer whereas brucine resolutions precipitate the D-isomer almost irrespective of the size of the functional groups (Greenstein, 1961) it was thought that hydrogen bonding ability rather than molecular shape would be the determining factor in formation of a crystalline complex.

The high degree of solvation of this complex is similar to that found in the benzoyl alanine complexes, but the positional disorder in the acetyl tyrosine moiety is a feature of this complex not repeated in any of the others examined so far. It is likely that this disorder, and the concomitant difficulties involved in crystallising this complex, provide the reason why this complex was not employed in resolution.

Crystallisation

Crystals of Strychnine: N-Acetyl-L-tyrosine were prepared by dissolving equal parts (1 mM) of strychnine and acetyl tyrosine in excess (5 ml) hot aqueous ethanol (50% v/v) and allowing the solution to cool slowly to room temperature. Needle-like crystals which turned opaque and crumbled when exposed to air were produced as the solvent evaporated. The crystal used for preliminary photography and data collection was sealed in a glass tube with some mother liquor in order to prevent its decomposition in this way.

Crystal data

Formula $C_{21}H_{22}N_2O_2 \cdot C_{11}H_{13}NO_4 \cdot 4H_2O$ Mol. Wt. = 679.66 $F(000) = 672$

Space group $P2_1$ Int. Tab. No. 4

Cell dimensions $a = 16.544(2)$ $b = 7.866(3)$ $c = 15.384(2)$ A
 $\beta = 115.714(12)^\circ$

$V = 1803.8$ A³ $Z = 2$ $D_c = 1.159$ gcm⁻³ $D_m = 1.174$ gcm⁻³

Radiation MoK α $\lambda = 0.71069$ A $\mu = 0.81$ cm⁻¹

Final R = 0.1040 based on 1162 independent reflections

$h_{\text{range}} -17$ 15 $k_{\text{max}} 8$ $l_{\text{max}} 14$

$\theta_{\text{max}} = 22^\circ$ $\text{Sin}(\theta_{\text{max}})/\lambda = 0.5271$

Drift curve min. 0.991 max. 1.011

2452 reflections measured (2413 unique)

1251 reflections considered unobserved ($F < 4\sigma(F)$)

Merging R = 0.0154

351 Parameters refined

Max. shift/e.s.d in last cycle 0.632

Weighting scheme: $W = 1.2705/(\sigma^2(F) + 0.009138F^2)$

Final difference Fourier min. -0.30 eA⁻³ max. 0.353 eA⁻³

Data Collection and Reduction

The data was collected on the CAD4 automated 4-circle diffractometer and corrected for Lorentz and polarisation effects by the program CADABS.

Structure Solution and Refinement

The initial attempt at solving the structure using a totally automated run of the MULTAN system was unsuccessful. The data set was renormalised using a Debye curve calculated for spherically averaged molecules of strychnine and N-acetyl tyrosine. This changed the E-values of enough reflections to result in MULTAN choosing a new starting set of reflections, but it did not lead to a solution of the structure. The data was then renormalised on the basis of a Wilson plot calculated from twice the cell contents, but the electron density

maps output by SEARCH showed no sign of recognisable molecular fragments. A final run of the program NORMAL with the thermal parameter B input as 7.0 rather than calculated from a Wilson plot changed the E-values sufficiently for the choice of a new starting set. The Fourier map calculated from this run of MULTAN contained a recognisable strychnine molecule and a fragment of the N-acetyl-tyrosine residue. The table below shows the starting sets with final phases and figures of merit from each run of MULTAN. (the code number referred to is the position in order of decreasing E-value and the letter 'E' is used to mark enantiomorph fixing reflections) The program SHELX was used to generate a difference Fourier phased on this fragment, but it soon became obvious that the rest of the tyrosine moiety was not going to appear as a chemically reasonable fragment. The strychnine molecule was used as an input to the DIRDIF system, but no additional peaks beyond the original C₅O fragment of the aromatic ring were found in the output electron density map. This result was a strong implication that the tyrosine was positionally disordered. When dealing with such a structure it is not possible to locate atoms by means of a peak search routine applied to a difference map, as it is possible that the overlap between two adjacent disordered atoms may produce a higher peak than the 'true' positions of the atoms. The program FMAP0 was used to scale and interpolate the full difference maps output by SHELX so that they could be contoured and examined for reasonable fragments of the molecule.

Origin fixing Reflections					Other reflections in starting set					Figure of merit		
code	h	k	l	angle	code	h	k	l	angle	Abs	$\Psi(0)$	R
Run 1 B = 5.85												
4	8	3	7	0	1	0	2	0	270			
8	7	3	8	45E	18	2	1	-1	0			
9	3	2	0	<u>±45</u>	43	3	3	11	90			
					45	6	4	-1	270	1.228	2.448	37.49
Run 2 B = 6.55												
3	13	0	3	0	1	0	2	0	153E			
4	8	3	7	0	8	3	2	0	193			
113	11	0	-6	0	9	7	3	8	233			
					231	4	3	-7	313			
					371	4	2	-3	74			
					55	5	1	9	275			
					77	6	3	-2	236	1.3623	0.553	21.76
Run 3 B = 5.93												
4	8	3	7	0	1	0	2	0	11			
7	7	3	8	45E	9	3	2	0	203			
71	6	0	9	0	25	2	1	-1	34			
					261	4	3	-7	56			
					146	1	6	-10	270	1.4006	2.360	41.52
Run 4 B = 7.00 (set)												
3	8	3	7	0	8	14	2	-6	234			
60	7	0	8	0	23	7	1	8	29			
88	8	0	7	0	25	4	3	8	340			
					28	7	2	8	79E			
					47	2	1	-1	86			
					204	10	1	-13	142	1.1368	1.371	24.7

Starting with a map phased on the strychnine parameters only, it was possible to locate two probable positions for the fragment containing the phenolic oxygen, the aromatic ring and the β -carbon, as well as two fully occupied water sites. A map phased on this additional

information revealed the position of the amide side chain (again distributed between two sites). The carboxylate groups and the α -carbons were the last to be identified in the maps.

The aromatic fragment and the amide side chain were each treated as rigid groups for the purposes of refinement and distance constraints were used to hold both fragments in position relative to the $C\alpha$ and C' carbons. The carboxylate group carbon to oxygen bond distances were allowed to refine normally. A difference map phased on this model showed several peaks of electron density which may be interpreted as positionally disordered water molecules.

Description of Structure

The model used for this structure contains a non-disordered strychninium cation and an N-acetyltyrosine anion distributed over two equally occupied locations.* There are four molecules of water of crystallisation per asymmetric unit, of which two molecules are in fully occupied sites and the other two are each associated with one possible position for the tyrosine moiety and are distributed over three sites. None of the contact distances between possibly occupied sites are close enough to force the crystal structure into any form of long range order or 'super cell', although there are of course positions that may not be occupied simultaneously.

As would be expected from the rigidity of the strychninium cation, none of its bond lengths or angles are significantly different from the average strychninium cation from the strychnine salt series of structures, even though the environment of the cation is not strictly comparable. There is also no significant difference in molecular geometry between this strychnine and the strychnine in the benzoylalanine complex.

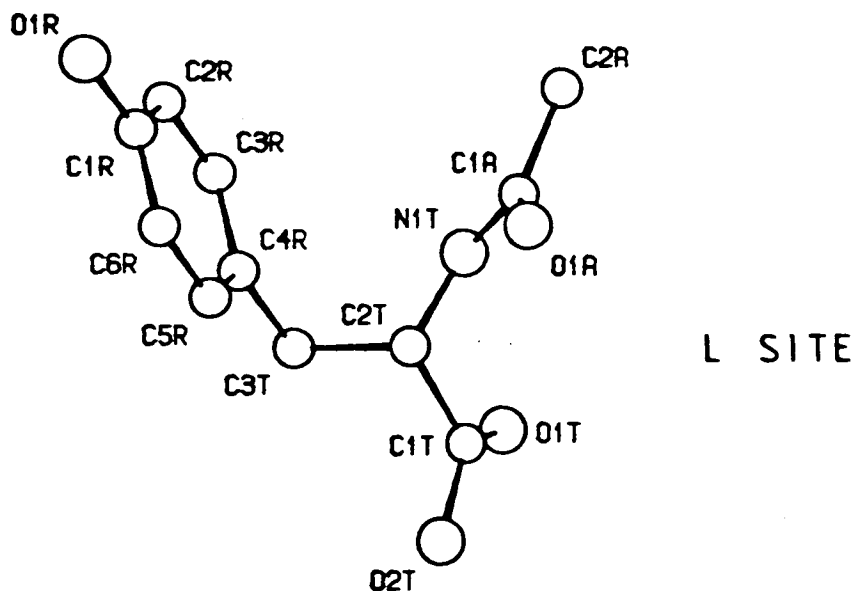
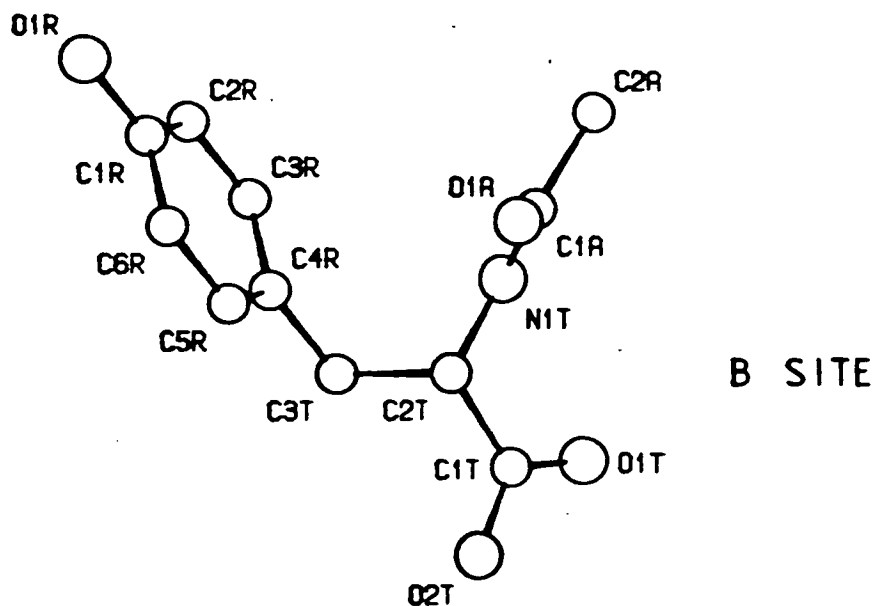
All bond lengths and angles in the N-acetyl tyrosine molecules are within three e.s.d.'s or less of their usual values, with no significant differences between the two possible molecules, but this to be

* No hydrogen atoms were positioned on the disordered acetyl tyrosine moiety, but the hydrogen atoms on the strychnine cation were totally unconstrained. No hydrogen atoms were placed on the solvent molecules.

expected because the constraints used in refinement held all bond lengths and angles, with the exception of those involving the carboxylate oxygens, close to the ideal values. There is a difference between the bond lengths of the carboxylate oxygens, but it is not quite large enough to be considered significant.

bond	B site	L site
C(1T)-O(1T)	1.20(5)	1.46(5)
C(1T)-O(2T)	1.30(5)	1.43(6)

The two tyrosine molecules are illustrated below.



Both molecules take up a similar conformation, with the only significant difference being in the two torsion angles involving O(1T) of the carboxylate group.

C(3T)-C(2T)-C(1T)-O(1T)	115.9(41)	85.9(30)
N(1T)-C(2T)-C(1T)-O(1T)	-5.6(51)	-41.7(34)

This similarity of conformation is not a result of the constraints applied during refinement, as the torsion angles which define peptide conformation (those involving the C α) were not held at any particular preconceived values. The torsion angles that were constrained were those of the aromatic ring and the amide group, which would not be expected to deviate from planarity in any case. The tables below list all bond lengths, angles and torsion angles for both the strychninium cation and the disordered tyrosine molecules, except for those fixed during refinement.

C1-C2	1.37(3)	C10-N2	1.46(3)
C1-C6	1.372(23)	C12-C13	1.48(3)
C1-N1	1.394(20)	C13-C14	1.499(23)
C2-C3	1.31(3)	C13-C16	1.57(3)
C3-C4	1.46(4)	C14-C19	1.584(24)
C4-C5	1.40(4)	C15-C16	1.48(3)
C5-C6	1.44(3)	C15-N2	1.43(3)
C6-C7	1.481(25)	C16-C17	1.35(3)
C7-C8	1.574(24)	C17-C18	1.57(3)
C7-C9	1.52(3)	C18-O2	1.42(3)
C7-C11	1.518(25)	C19-C20	1.46(3)
C8-C14	1.512(23)	C19-O2	1.438(22)
C8-N1	1.502(22)	C20-C21	1.50(3)
C9-C12	1.55(3)	C21-N1	1.420(23)
C9-N2	1.58(3)	C21-O1	1.206(23)
C10-C11	1.56(3)		

B

I

C1T-O1T	1.20(5)	1.46(5)
C1T-O2T	1.30(5)	1.43(6)

C2-C1-C6	125.7(16)	C12-C13-C16	108.9(15)
C2-C1-N1	126.1(15)	C14-C13-C16	115.6(14)

C6-C1-N1	108.1(14)	C8-C14-C13	112.6(13)
C1-C2-C3	116.9(20)	C8-C14-C19	106.2(13)
C2-C3-C4	122.6(24)	C13-C14-C19	121.3(13)
C3-C4-C5	120.1(25)	C16-C15-N2	110.4(17)
C4-C5-C6	115.7(21)	C13-C16-C15	113.3(17)
C1-C6-C5	118.9(16)	C13-C16-C17	122.4(19)
C1-C6-C7	113.0(15)	C15-C16-C17	124.3(20)
C5-C6-C7	128.0(17)	C16-C17-C18	118.5(19)
C6-C7-C8	102.6(14)	C17-C18-02	111.7(17)
C6-C7-C9	115.3(15)	C14-C19-C20	110.4(15)
C6-C7-C11	112.8(15)	C14-C19-02	110.3(14)
C8-C7-C9	112.8(14)	C20-C19-02	107.8(15)
C8-C7-C11	111.4(14)	C19-C20-C21	118.9(16)
C9-C7-C11	102.3(14)	C20-C21-N1	116.4(15)
C7-C8-C14	117.5(14)	C20-C21-01	123.5(17)
C7-C8-N1	101.8(13)	N1-C21-01	119.9(16)
C14-C8-N1	107.1(13)	C1-N1-C8	111.6(12)
C7-C9-C12	116.2(16)	C1-N1-C21	126.8(14)
C7-C9-N2	105.1(14)	C8-N1-C21	114.9(13)
C12-C9-N2	109.3(15)	C9-N2-C10	107.0(15)
C11-C10-N2	105.2(16)	C9-N2-C15	111.6(15)
C7-C11-C10	103.7(15)	C10-N2-C15	117.5(16)
C9-C12-C13	107.3(15)	C18-02-C19	113.9(14)
C12-C13-C14	108.4(14)		

B

L

C4R-C3T-C2T	112.6(16)	105.2(16)
C2T-N1T-C1A	119.1(21)	121.3(19)
C3T-C2T-N1T	109.4(19)	113.1(18)
C3T-C2T-C1T	112.3(24)	111.6(21)
N1T-C2T-C1T	109.1(24)	111.6(22)
C2T-C1T-01T	122 (4)	109 (3)
C2T-C1T-02T	116 (4)	122 (3)
01T-C1T-02T	122 (4)	123 (4)

C6-C1-C2-C3	-2 (3)
N1-C1-C2-C3	177.6(19)
C2-C1-C6-C5	-1.1(28)
C2-C1-C6-C7	175.9(17)
N1-C1-C6-C5	179.6(16)
N1-C1-C6-C7	-3.4(19)
C2-C1-N1-C8	172.2(17)
C2-C1-N1-C21	22.9(26)

C6-C1-N1-C8	-8.5(18)
C6-C1-N1-C21	-157.9(16)
C1-C2-C3-C4	4 (4)
C2-C3-C4-C5	-4 (4)
C3-C4-C5-C6	1 (4)
C4-C5-C6-C1	1 (3)
C4-C5-C6-C7	-175.4(21)
C1-C6-C7-C8	12.8(19)
C1-C6-C7-C9	135.8(16)
C1-C6-C7-C11	-107.2(17)
C5-C6-C7-C8	-170.5(18)
C5-C6-C7-C9	-47.5(26)
C5-C6-C7-C11	69.5(24)
C6-C7-C8-C14	100.6(16)
C6-C7-C8-N1	-16.0(16)
C9-C7-C8-C14	-24.1(21)
C9-C7-C8-N1	-140.7(14)
C11-C7-C8-C14	-138.5(15)
C11-C7-C8-N1	104.9(15)
C6-C7-C9-C12	-84.6(20)
C6-C7-C9-N2	154.4(15)
C8-C7-C9-C12	32.8(22)
C8-C7-C9-N2	-88.2(17)
C11-C7-C9-C12	152.6(16)
C11-C7-C9-N2	31.6(17)
C6-C7-C11-C10	-165.4(15)
C8-C7-C11-C10	79.8(17)
C9-C7-C11-C10	-41.0(17)
C7-C8-C14-C13	38.7(20)
C7-C8-C14-C19	173.6(14)
N1-C8-C14-C13	152.3(13)
N1-C8-C14-C19	-72.8(15)
C7-C8-N1-C1	15.6(16)
C7-C8-N1-C21	168.9(14)
C14-C8-N1-C1	-108.3(14)
C14-C8-N1-C21	45.0(18)
C7-C9-C12-C13	-56.2(21)
N2-C9-C12-C13	62.5(18)
C7-C9-N2-C10	-10.3(19)
C7-C9-N2-C15	119.5(17)
C12-C9-N2-C10	-135.7(16)
C12-C9-N2-C15	-5.9(21)
N2-C10-C11-C7	35.2(18)
C11-C10-N2-C9	-14.9(19)
C11-C10-N2-C15	-141.3(17)

C9-C12-C13-C14	69.1(18)
C9-C12-C13-C16	-57.5(19)
C12-C13-C14-C8	-62.0(18)
C12-C13-C14-C19	170.7(15)
C16-C13-C14-C8	60.6(19)
C16-C13-C14-C19	-66.7(20)
C12-C13-C16-C15	-0.3(23)
C12-C13-C16-C17	-179.9(20)
C14-C13-C16-C15	-122.6(18)
C14-C13-C16-C17	57.9(26)
C8-C14-C19-C20	44.2(18)
C8-C14-C19-02	-74.8(16)
C13-C14-C19-C20	174.3(15)
C13-C14-C19-02	55.3(20)
N2-C15-C16-C13	58.5(22)
N2-C15-C16-C17	-122.0(23)
C16-C15-N2-C9	-52.7(21)
C16-C15-N2-C10	71.4(22)
C13-C16-C17-C18	-3 (3)
C15-C16-C17-C18	177.4(19)
C16-C17-C18-02	-66.4(25)
C17-C18-02-C19	95.4(19)
C14-C19-C20-C21	9.7(23)
02-C19-C20-C21	130.2(17)
C14-C19-02-C18	-68.9(18)
C20-C19-02-C18	170.6(16)
C19-C20-C21-N1	-39.8(24)
C19-C20-C21-01	145.6(19)
C20-C21-N1-C1	158.6(15)
C20-C21-N1-C8	10.1(21)
01-C21-N1-C1	-26.6(26)
01-C21-N1-C8	-175.1(16)

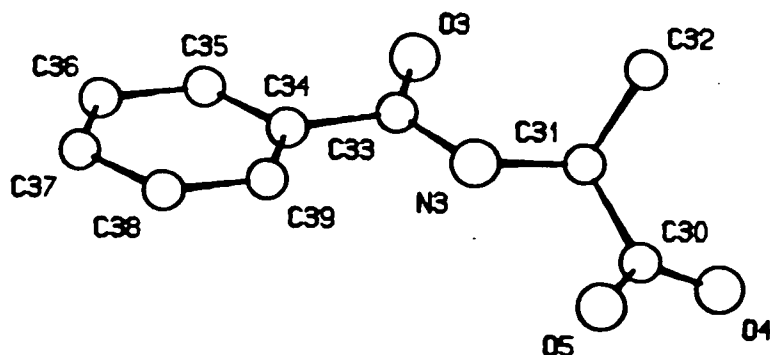
B

L

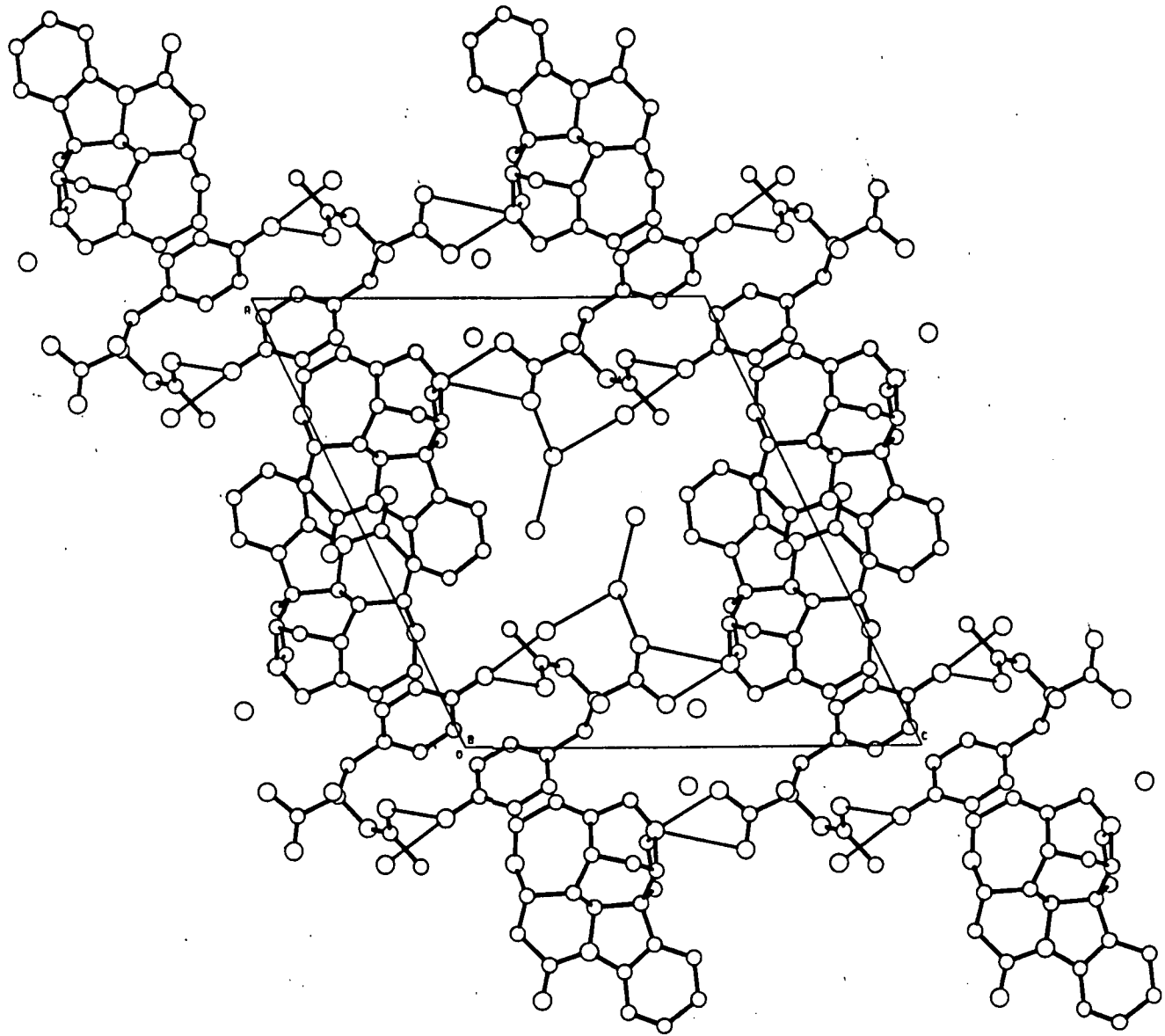
C2T-C3T-C4R-C3R	99.5(21)	100.5(20)
C2T-C3T-C4R-C5R	-80.4(22)	-79.5(21)
C4R-C3T-C2T-N1T	-67.3(22)	-73.2(21)
C4R-C3T-C2T-C1T	171.4(24)	160.0(21)
C3T-C2T-N1T-C1A	115.7(24)	128.1(21)
C1T-C2T-N1T-C1A	-121 (3)	-105.1(27)
C2T-N1T-C1A-01A	-3 (4)	5 (3)
C2T-N1T-C1A-C2A	178.6(21)	-169.4(20)
C3T-C2T-C1T-01T	116 (4)	86 (3)
C3T-C2T-C1T-02T	-64 (4)	-67 (4)

N1T-C2T-C1T-01T	-6 (5)	-42 (3)
N1T-C2T-C1T-02T	175 (3)	166 (3)

The next four pages show projections down the b axis for two possible arrangements of the strychnine tyrosine complex and projections down the a and c axes of the strychnine alanine complex. The numbering scheme used for the benzoyl alanine molecule is shown below.

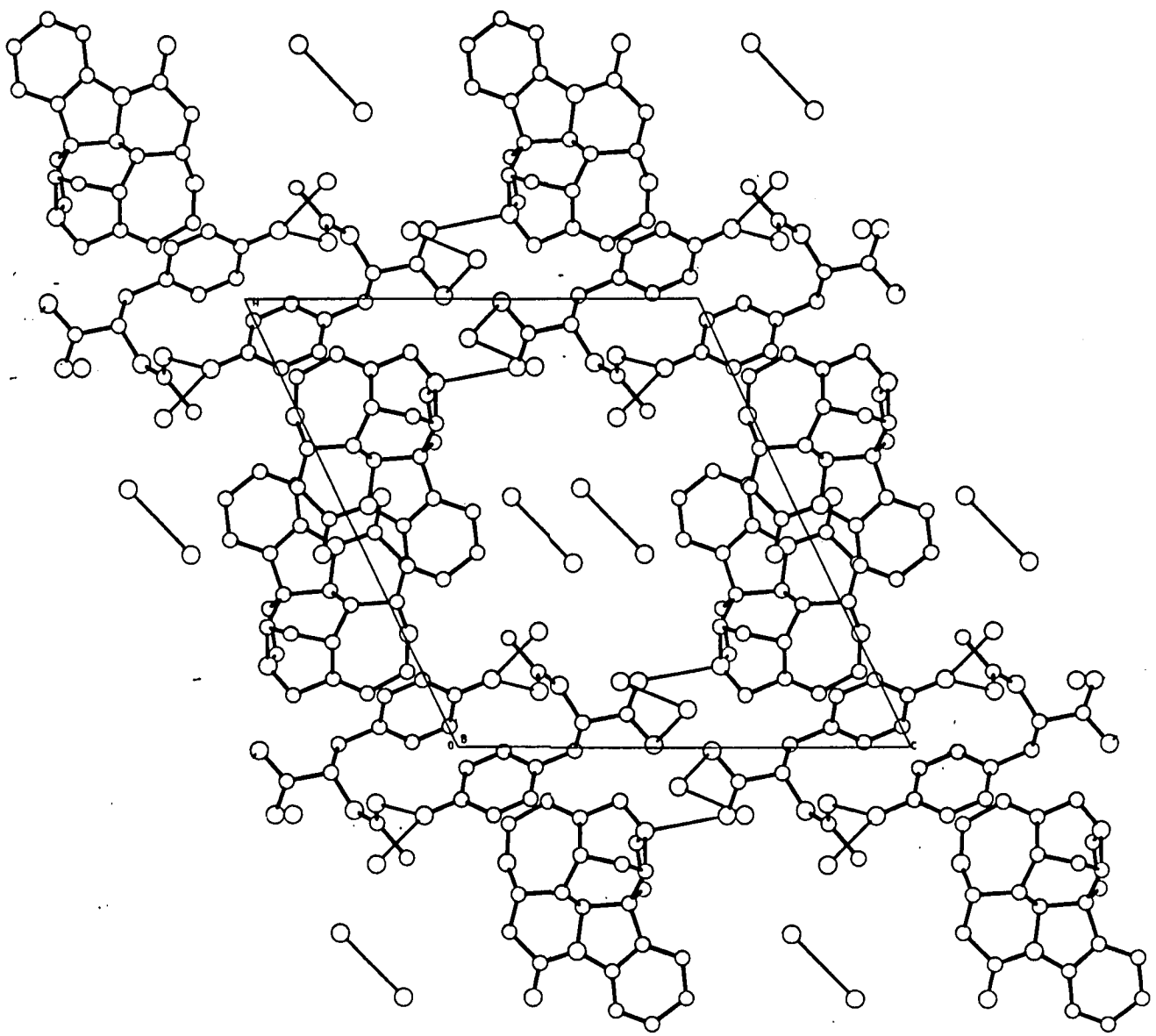


**STRYCHNINE : N-ACETYL-L-TYROSINE
TETRAHYDRATE**

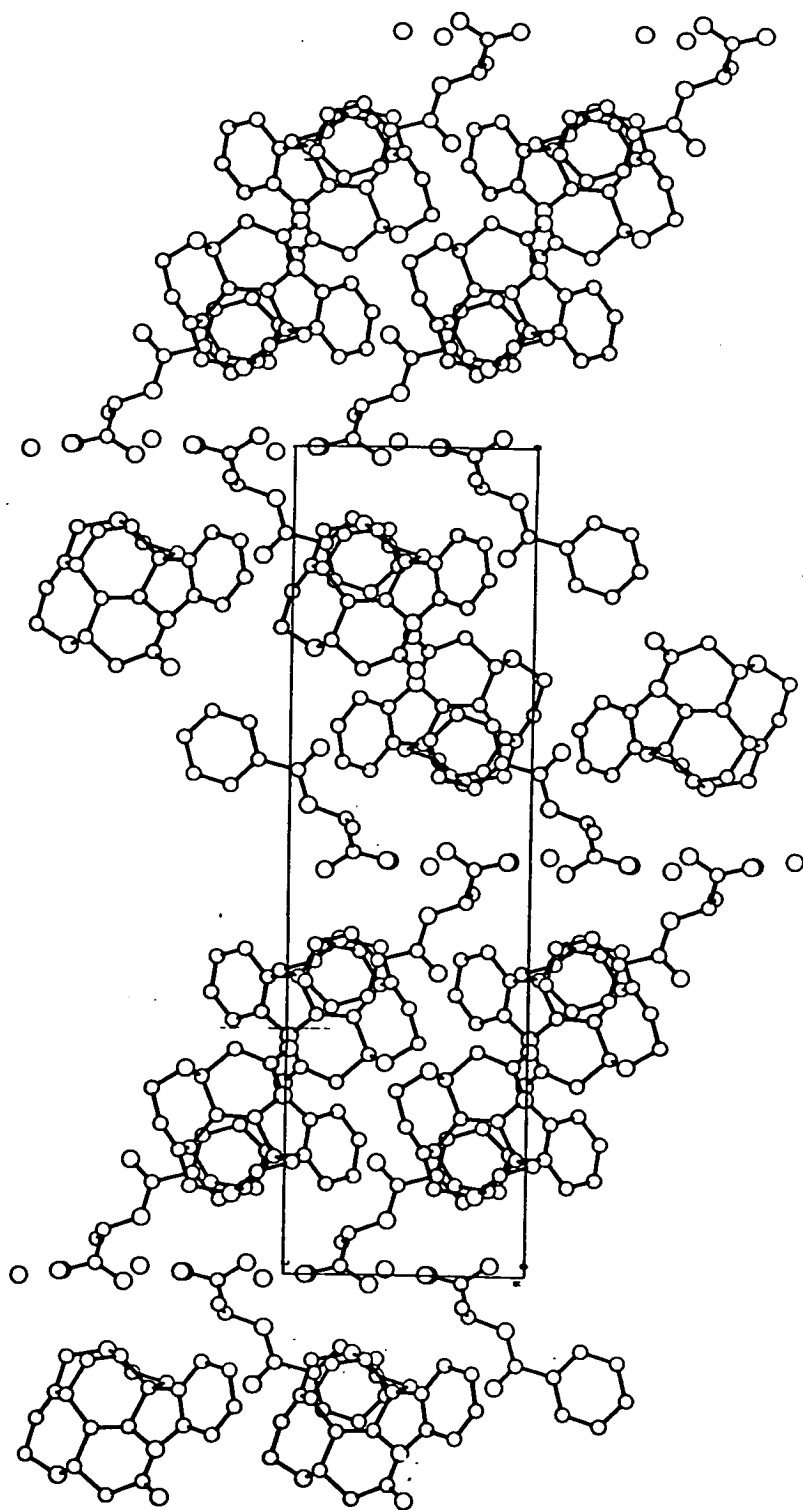


PROJECTION DOWN Y AXIS FOR 'B' SITE

**STRYCHNINE : N-ACETYL-L-TYROSINE
TETRAHYDRATE**

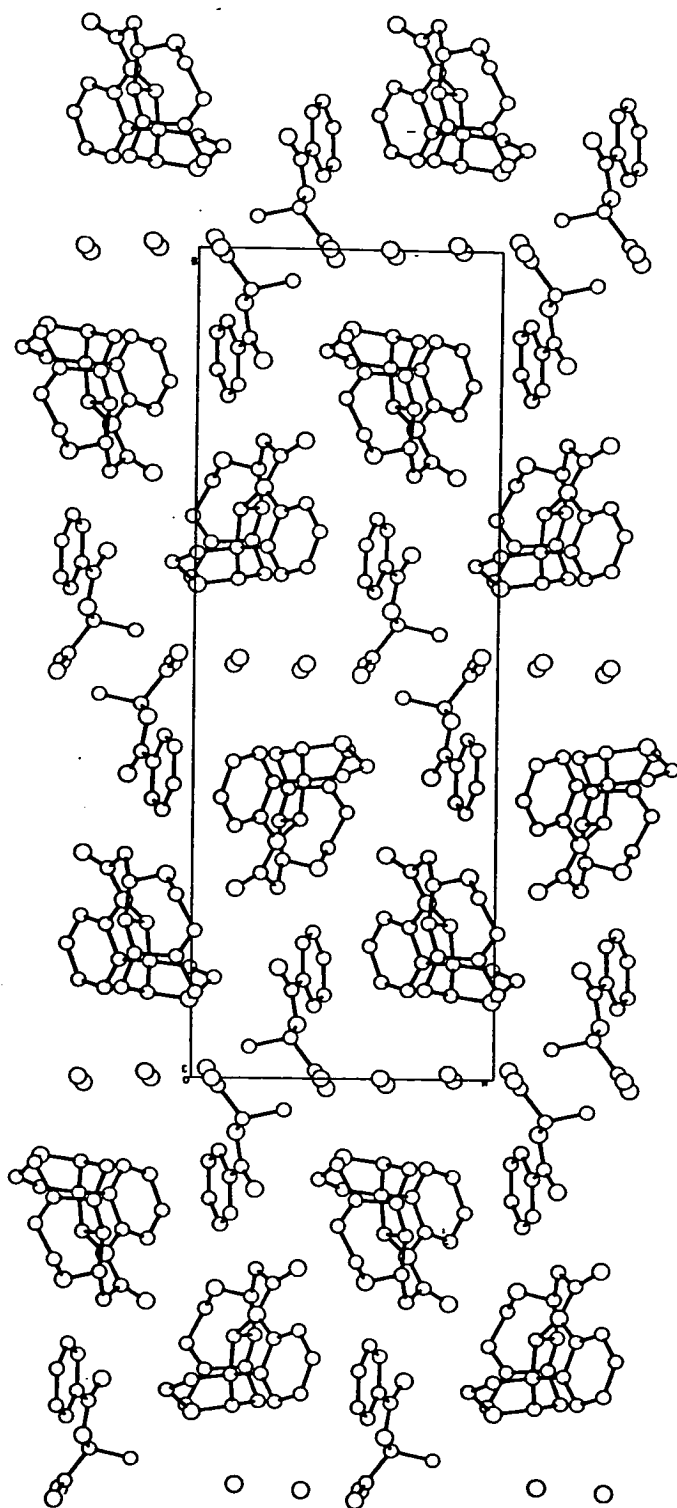


PROJECTION DOWN Y AXIS FOR 'L' SITE



PROJECTION DOWN THE X AXIS

**STRYCHNINE : N-BENZOYL-L-ALANINE
DIHYDRATE**



PROJECTION DOWN THE Z AXIS

**STRYCHNINE : N-BENZOYL-L-ALANINE
DIHYDRATE**

The only close contact between adjacent strychnine molecules involves the carbonyl oxygen O(1) and the methene carbon of the pyrrole ring C(11)

C11 3.393 (23) O1 (1 -x, -1/2+y, -z)

There are eleven contacts closer than 3.5 angstroms between the strychnine and the tyrosine molecules. It is noticeable that the aromatic residue of the amino acid makes contact with the oxepin ring, the carboxylate group with the immediate surroundings of the quaternary nitrogen and the phenol oxygen comes close to the aromatic ring of the strychnine.

C3	3.49 (3)	O1RB	(1 +x, y, z)
C9	3.29 (3)	O1TB	(x, y, -1 +z)
C9	3.48 (4)	O2TB	(x, y, -1 +z)
C9	3.40 (4)	O1TL	(x, y, -1 +z)
C10	3.38 (4)	O2TB	(x, y, -1 +z)
C1	3.470 (5)	O2TL	(-x, -1/2+y, -z)
C15	3.07 (4)	O2TB	(x, y, -1 +z)
C17	3.463 (28)	C5RL	(-x, -1/2+y, -z)
C18	3.47 (3)	C3RI	(-x, 1/2+y, -z)
N2	3.26 (5)	C1TB	(x, y, -1 +z)
O2	3.352 (22)	C6RI	(-x, -1/2+y, -z)

These may be compared with the equivalent close contacts to strychnine in the strychnine: N-benzoylalanine complex. There is a similar contact between O(1) and C(11) and again it can be seen that the carboxylate group is in close contact with the protonated nitrogen (C30 is the C', O4 and O5 are the carboxylate oxygens and O3 is the benzoyl group oxygen). The benzoyl oxygen, which does not take part in hydrogen bonding, has three close contacts to hydrophobic regions of the strychnine molecule.

O2	3.408 (7)	C2	(1/2+x, 1/2-y, -z)
O1	3.497 (7)	C8	(-1/2+x, 1/2-y, -z)
O1	3.450 (8)	C10	(-1/2+x, 1/2-y, -z)
O1	3.329 (7)	C11	(-1/2+x, 1/2-y, -z)

C30	3.385 (7)	N2	(1 -x, -1/2+y, 1/2-z)
O3	3.427 (7)	C3	(1/2+x, 1/2-y, 1 -z)
O3	3.420 (7)	C17	(-1/2+x, 1/2-y, -z)
O3	3.263 (7)	C20	(x, y, 1 +z)
O5	3.453 (7)	C9	(1 -x, -1/2+y, 1/2-z)
O5	3.365 (7)	C10	(1 -x, -1/2+y, 1/2-z)
O5	3.475 (8)	C15	(1 -x, -1/2+y, 1/2-z)
O4	3.230 (8)	C15	(1 -x, -1/2+y, 1/2-z)

The close contacts between tyrosine molecules are mostly between the aromatic ring and nearby oxygen atoms and there is one close contact between two aromatic rings. The carbon atoms of one of the amide side chains have close contact with the phenolic oxygens, but this may be a consequence of the amide oxygen: aromatic ring interaction, and not an actual interaction in itself.

C3RB	3.376 (26)	O1RB	(-x, -1/2+y, -z)
C1RL	3.204 (28)	O1AL	(-x, -1/2+y, -z)
C2RL	3.306 (28)	O1AL	(-x, -1/2+y, -z)
C1AL	3.152 (28)	O1RL	(-x, 1/2+y, -z)
C2AL	3.221 (28)	O1RL	(-x, 1/2+y, -z)
C1RB	3.477 (28)	O1AL	(-x, -1/2+y, -z)
C2RB	3.434 (28)	O1AL	(-x, -1/2+y, -z)
C3RB	3.466 (26)	C1RL	(-x, -1/2+y, -z)
C3RB	3.204 (26)	O1RL	(-x, -1/2+y, -z)
O1RB	3.381 (28)	C1AL	(-x, -1/2+y, -z)
O1RB	3.273 (28)	C2AL	(-x, -1/2+y, -z)
O1AB	3.38 (3)	C1RL	(-x, 1/2+y, -z)
O1AB	3.40 (3)	C2RI	(-x, 1/2+y, -z)
C1AB	3.46 (3)	O1RL	(-x, 1/2+y, -z)

The model used to describe this structure permits the existence of two self-consistent hydrogen bonding schemes, each of which connect the strychninium cation and the fully occupied water positions to one of the disordered tyrosine molecules via a different set of partially occupied water molecules. There are also a number of interconnections between the two systems, allowing adjacent asymmetric units to have either possible amino acid position occupied, without causing any long-range ordering of the cells. There are only four pairs of water

sites that may not be occupied simultaneously (marked '*'), allowing a very large number of possible hydrogen bonding schemes to be drawn up, depending on which combination of disordered molecules is used.

N2	3.16	(3)	01TB	(x, y, -1 +z)
N2	2.56	(3)	02TB	(x, y, -1 +z)
N1TB	2.58	(4)	01W2	(x, y, z)
01TB	2.99	(7)	01WB	(1 -x, -1/2+y, 1 -z)
02TB	2.86	(4)	01W1	(x, y, z)
01RB	2.89	(3)	01AB	(-x, -1/2+y, -z)
01RB	2.98	(3)	01W2	(-x, 1/2+y, -z)
01AB	2.69	(7)	02WB	(x, y, z)
01W1	3.38	(7)	02WB	(-x, -1/2+y, 1 -z)
01W2	2.83	(7)	01WB	(1 -x, -1/2+y, 1 -z)
01W2	3.44	(7)	02WB	(x, -1 +y, z)
01WB	2.55	(9)	03WB	(1 -x, -1/2+y, 1 -z)
N2	2.86	(4)	01TL	(x, y, -1 +z)
N1TL	3.27	(3)	01W2	(x, y, z)
02TL	2.35	(6)	01W1	(x, y, z)
02TL	2.77	(6)	01W1	(-x, 1/2+y, 1 -z)
01RL	2.318	(28)	01AL	(-x, -1/2+y, -z)
01RL	2.72	(3)	01W2	(-x, 1/2+y, -z)
01AL	3.18	(7)	01WL	(x, y, z)
01W1	2.58	(7)	01WL	(-x, -1/2+y, 1 -z)
02WL	3.18	(9)	03WL	(x, 1 +y, z)
01TB	3.49	(7)	01WL	(x, -1 +y, z)
01RB	2.664	(28)	01AL	(-x, -1/2+y, -z)
01RL	2.60	(3)	01AB	(-x, -1/2+y, -z)
01AB	3.46	(7)	01WL	(x, y, z)
01RL	3.50	(7)	02WB	(-x, -1/2+y, -z)
01AL	* 2.30	(7)	02WB	(x, y, z)
01WB	3.20	(9)	01WL	(1 -x, -1/2+y, 1 -z)
01WB	* 2.14	(9)	02WL	(x, -1 +y, z)
03WB	* 1.55	(9)	02WL	(1 -x, -1/2+y, 1 -z)
03WB	* 1.85	(9)	03WL	(1 -x, 1/2+y, 1 -z)

The hydrogen bonding scheme in the strychnine: benzoylalanine complex is comparatively simple, involving only two water molecules and three other atoms (the quaternary nitrogen and the carboxylate group of the alanine) in a 'tighter', more rigid three dimensional network.

05	2.670 (6)	N2	(1 -x, -1/2+y, 1/2-z)
04	3.423 (7)	N2	(1 -x, -1/2+y, 1/2-z)
05	2.836 (7)	091W	(1/2+x, 1/2-y, -z)
04	2.817 (9)	092W	(1 -x, -1/2+y, 1/2-z)
091W	2.785 (10)	092W	(x, y, z)
091W	2.881 (10)	092W	(1/2-x, 1 -y, 1/2+z)

Table Fractional Coordinates of Atoms with Standard Deviations

	x	y	z	Ueq
C(1)	0.5002(11)	-0.06837(0)	-0.1137(12)	0.0810
C(2)	0.5909(12)	-0.0472(23)	-0.0804(16)	0.0866
C(3)	0.6212(13)	-0.022(3)	-0.1448(18)	0.1208
C(4)	0.5621(19)	-0.005(3)	-0.2476(23)	0.1640
C(5)	0.4690(12)	-0.027(3)	-0.2813(14)	0.1123
C(6)	0.4385(11)	-0.0617(23)	-0.2086(11)	0.0755
C(7)	0.3462(10)	-0.102(3)	-0.2225(9)	0.0738
C(8)	0.3548(10)	-0.0943(24)	-0.1166(10)	0.0721
C(9)	0.2736(10)	0.015(3)	-0.2912(12)	0.0866
C(10)	0.2112(11)	-0.270(3)	-0.3004(14)	0.0915
C(11)	0.3144(10)	-0.2751(24)	-0.2677(11)	0.0760
C(12)	0.2580(11)	0.182(3)	-0.2474(12)	0.0815
C(13)	0.2385(10)	0.1353(23)	-0.1656(10)	0.0701
C(14)	0.3224(9)	0.0651(20)	-0.0866(10)	0.0569
C(15)	0.1179(12)	-0.020(3)	-0.3076(13)	0.0919
C(16)	0.1551(11)	0.014(3)	-0.2030(16)	0.1016
C(17)	0.1213(12)	-0.050(3)	-0.1442(15)	0.1055
C(18)	0.1706(13)	-0.008(3)	-0.0335(15)	0.1135
C(19)	0.3287(12)	0.0279(23)	0.0173(10)	0.0753
C(20)	0.4139(12)	-0.0556(24)	0.0778(12)	0.0901
C(21)	0.4848(12)	-0.0668(22)	0.0419(12)	0.0731
N(1)	0.4547(7)	-0.1025(18)	-0.0578(9)	0.0642
N(2)	0.1847(9)	-0.095(3)	-0.3311(9)	0.0930
O(1)	0.5641(9)	-0.0591(21)	0.0940(8)	0.1177
O(2)	0.2574(7)	-0.0837(18)	0.0105(7)	0.0872
H(2)	0.6339(12)	-0.0477(23)	-0.0038(16)	0.2771
H(3)	0.6927(13)	-0.012(3)	-0.1212(18)	0.1714
H(4)	0.5884(19)	0.016(3)	-0.2998(23)	0.2324
H(5)	0.4221(12)	-0.013(3)	-0.3562(14)	0.0811
H(8)	0.3144(10)	-0.1948(24)	-0.1080(10)	0.0893
H(9)	0.2916(10)	0.055(3)	-0.3477(12)	0.0255
H(101)	0.1953(11)	-0.302(3)	-0.2413(14)	0.0474
H(102)	0.1771(11)	-0.356(3)	-0.3596(14)	0.0429
H(111)	0.3439(10)	-0.3753(24)	-0.2154(11)	0.0590
H(112)	0.3299(10)	-0.2958(24)	-0.3283(11)	0.0575
H(121)	0.2042(11)	0.256(3)	-0.3009(12)	0.1611
H(122)	0.3192(11)	0.256(3)	-0.2205(12)	0.0055
H(13)	0.2172(10)	0.2416(23)	-0.1362(10)	0.1352
H(14)	0.3659(9)	0.1728(20)	-0.0774(10)	0.0636
H(151)	0.0639(12)	-0.111(3)	-0.3254(13)	0.0775
H(152)	0.0922(12)	0.095(3)	-0.3479(13)	0.0648
H(17)	0.0631(12)	-0.131(3)	-0.1730(15)	1.0535
H(181)	0.1303(13)	-0.056(3)	0.0005(15)	2.7469
H(182)	0.1780(13)	0.128(3)	-0.0234(15)	0.1406
H(19)	0.3204(12)	0.1469(23)	0.0475(10)	0.0765
H(201)	0.3991(12)	-0.1843(24)	0.0907(12)	0.1091
H(202)	0.4431(12)	0.0128(24)	0.1453(12)	0.1493
H(1N2)	0.1503(9)	-0.094(3)	-0.4089(9)	1.4424

Table Fractional Coordinates of Atoms with Standard Deviations.

	x	y	z	Ueq
C(3TB)	0.0379(12)	-0.184(3)	0.2789(9)	0.0857
C(1RB)	-0.1106(12)	-0.057(3)	-0.0217(9)	0.0736
C(2RB)	-0.0392(12)	-0.172(3)	0.0071(9)	0.0607
C(3RB)	0.0091(12)	-0.213(3)	0.1047(9)	0.0663
C(4RB)	-0.0142(12)	-0.139(3)	0.1735(9)	0.0987
C(5RB)	-0.0655(12)	-0.025(3)	0.1445(9)	0.0651
C(6RB)	-0.1338(12)	0.016(3)	0.0468(9)	0.0858
O(1RB)	-0.1584(12)	-0.017(3)	-0.1183(9)	0.0739
C(2TB)	0.1079(14)	-0.049(3)	0.3349(16)	0.0803
C(3TL)	-0.0070(12)	-0.131(3)	0.2591(9)	0.0732
C(1RL)	-0.1168(12)	-0.042(3)	-0.0471(9)	0.0906
C(2RL)	-0.0468(12)	-0.158(3)	-0.0034(9)	0.1075
C(3RL)	-0.0111(12)	-0.187(3)	0.0960(9)	0.0752
C(4RL)	-0.0454(12)	-0.100(3)	0.1517(9)	0.0631
C(5RL)	-0.1153(12)	0.017(3)	0.1080(9)	0.0731
C(6RL)	-0.1510(12)	0.046(3)	0.0085(9)	0.0943
O(1RL)	-0.1521(12)	-0.013(3)	-0.1456(9)	0.0949
C(2TL)	0.0565(13)	0.018(3)	0.3054(16)	0.0834
N(1TB)	0.1791(13)	-0.047(3)	0.3012(16)	0.0891
O(1AB)	0.1357(13)	0.224(3)	0.2396(16)	0.1275
C(1AB)	0.1901(13)	0.094(3)	0.2569(16)	0.1500
C(2AB)	0.2659(13)	0.097(3)	0.2239(16)	0.1414
C(1TB)	0.151(3)	-0.080(8)	0.4441(14)	0.1500
O(1TB)	0.2289(17)	-0.112(4)	0.4892(17)	0.1274
O(2TB)	0.0967(18)	-0.072(4)	0.4855(18)	0.1225
N(1TL)	0.1401(12)	0.008(3)	0.2935(15)	0.0802
O(1AL)	0.1248(12)	0.284(3)	0.2389(15)	0.1087
C(1AL)	0.1691(12)	0.142(3)	0.2568(15)	0.0746
C(2AL)	0.2467(12)	0.126(3)	0.2274(15)	0.1703
C(1TL)	0.0754(23)	0.041(6)	0.4111(19)	0.1500
O(1TL)	0.1496(22)	-0.069(5)	0.4706(17)	0.1163
O(2TL)	0.007(3)	0.092(7)	0.438(3)	0.2208
O(1W1)	-0.0860(14)	-0.125(4)	0.4498(17)	0.2739
O(1W2)	0.2609(15)	-0.326(4)	0.3008(17)	0.2642
O(1WB)	0.649(3)	0.100(9)	0.500(4)	0.2000
O(2WB)	0.098(4)	0.463(9)	0.341(4)	0.2000
O(3WB)	0.515(4)	0.547(9)	0.616(4)	0.2000
O(1WL)	0.149(4)	0.472(9)	0.431(4)	0.2000
O(2WL)	0.564(4)	0.943(9)	0.382(4)	0.2000
O(3WL)	0.416(3)	0.048(9)	0.452(4)	0.2000

Thermal Vibration Parameters with Standard Deviations

	U11	U22	U33	U23	U13	U12
C(1)	0.0886(130)	0.0748(134)	0.0733(112)	-0.0041(105)	0.0585(118)	0.0201(115)
C(2)	0.0683(129)	0.0619(127)	0.1075(151)	0.0029(114)	0.0349(112)	0.0043(106)
C(3)	0.0774(135)	0.1430(230)	0.1184(170)	0.0147(174)	0.0410(137)	0.0401(146)
C(4)	0.1560(219)	0.1148(200)	0.2136(288)	-0.0095(203)	0.1461(227)	0.0085(187)
C(5)	0.0716(114)	0.1416(201)	0.1115(148)	0.0094(138)	0.0571(113)	0.0204(131)
C(6)	0.0763(107)	0.0826(135)	0.0545(104)	0.0010(104)	0.0322(89)	0.0207(109)
C(7)	0.0675(109)	0.1083(163)	0.0229(79)	0.0096(107)	-0.0033(69)	0.0186(119)
C(8)	0.0690(105)	0.0580(117)	0.0680(102)	0.0002(98)	0.0195(85)	-0.0046(102)
C(9)	0.0698(120)	0.1031(162)	0.0615(108)	0.0095(116)	0.0097(97)	-0.0040(119)
C(10)	0.0705(124)	0.0873(165)	0.0914(137)	-0.0399(132)	0.0230(106)	-0.0001(118)
C(11)	0.0749(120)	0.0744(138)	0.0529(100)	-0.0345(109)	0.0073(82)	0.0017(105)
C(12)	0.0597(107)	0.0760(135)	0.0866(123)	0.0046(114)	0.0221(104)	0.0044(98)
C(13)	0.0585(99)	0.0857(133)	0.0539(100)	-0.0166(98)	0.0262(90)	-0.0066(98)
C(14)	0.0447(83)	0.0462(110)	0.0678(101)	-0.0110(83)	0.0265(76)	-0.0054(81)
C(15)	0.0634(109)	0.1041(180)	0.0732(125)	-0.0400(128)	-0.0058(88)	-0.0199(120)
C(16)	0.0694(111)	0.0853(157)	0.1253(175)	-0.0268(138)	0.0379(123)	0.0207(116)
C(17)	0.0962(129)	0.0700(138)	0.1394(164)	-0.0036(137)	0.0819(134)	0.0122(116)
C(18)	0.0957(142)	0.0931(186)	0.1309(185)	-0.0149(141)	0.0590(131)	0.0098(131)
C(19)	0.1107(134)	0.0373(95)	0.0564(103)	-0.0135(89)	0.0322(102)	0.0015(110)
C(20)	0.1063(140)	0.0539(125)	0.0772(120)	0.0199(110)	0.0180(108)	-0.0062(117)
C(21)	0.0798(118)	0.0531(114)	0.0684(120)	0.0137(101)	0.0307(103)	0.0262(105)
N(1)	0.0593(74)	0.0438(84)	0.0635(89)	0.0025(75)	0.0045(70)	0.0018(68)
N(2)	0.0684(94)	0.1209(155)	0.0612(82)	-0.0141(100)	0.0030(74)	0.0227(110)
O(1)	0.0901(86)	0.1293(128)	0.0861(86)	-0.0206(91)	-0.0126(71)	0.0319(96)
O(2)	0.0832(74)	0.0845(87)	0.0830(69)	-0.0073(75)	0.0513(62)	-0.0186(86)

REFERENCES

REFERENCES

- Abderhalden, E. & Sichel, H., (1923) *Z. Physiol. Chem.* **131**, 277
- Aberhart, D. J., (1977) *Tetrahedron*, **33**, 1545
- Arnstein, H. R. V. & Clubb, M. E., (1958) *Biochem. J.* **68**, 528
- Ashida, T., Hata, Y., Matsuura, Y., Tanaka, N. & Kakudo, M., (1977) *Acta Cryst.* **B33**, 3561-4
- Baldwin, J. E. & Wan, T. S., (1981) *Tetrahedron*, **31**, 1589-95
- Baldwin, J. E., Bahudur, G., Wan, T., Jung, M., Abraham, E. P., Huddleston, J. A. & White, R. L., (1981) *J. Chem. Soc. Chem. Commun.* **1981**, 1146-7
- Barfield, M. & Karplus, M., (1969) *J. Am. Chem. Soc.* **91**, 1-10
- Barfield, M. & Gearhart, H. L., (1973) *J. Am. Chem. Soc.* **95**(2), 641-3
- Beevers, C. A., (1976) *Proc. 3rd European Crystallographic Meeting, Zurich*, 257.
- Beevers, C. A. & Robertson, J. H., (1951) *Acta Cryst.* **4**, 270-5
- Behr, J.-P., Lehn, J.-M. & Vierling, P., (1982) *Helv. Chim. Acta*, **65**(6), 1853-67
- Berg, C. P., (1936) *J. Biol. Chem.* **115**, 9-15
- Berman, H. M., Stallings, W., Carrell, H. L., Glusker, J. P., Neidle, S., Taylor, G. & Achari, A., (1979) *Biopolymers*, **18**, 2405

Bokhoven, C., Schoone, J. C. & Bijvoet, J. M., (1951) *Acta Cryst.* **4**, 275-80

Bradshaw, J. S., Jones, B. A., Davidson, R. B., Christensen, J. J., Lamb, J. D., Izatt, R. M., Morin, F. G. & Grant, D. M., (1982) *J. Org. Chem.* **47**, 3362-4

Bystrov, V. F., Portnova, S. L., Tsetlin, V. I., Ivanov, V. T. & Ovchinnikov, Yu. A., (1969) *Tetrahedron*, **25**, 493-515

Capasso, S., Mattia, C. & Mazzarella, L., (1977) *Acta Cryst.* **B33**, 2080-83

Capasso, S., Mazzarella, L. & Tancredi, T., (1979) *Biopolymers*, **18**, 1555-8

Carey, F. A. & Kuehne, M. E., (1982) *J. Org. Chem.* **47**(20), 3811-5

Chandrasekaran, R. & Mitra, A. K., (1977) *Int. J. Peptide Protein Res.* **10**, 235-9

Chandrasekaran, R. & Balasubramanian, R., (1969) *Biochim. Biophys. Acta*, **188**, 1-9

Ciunik, Z. & Glowiak, T., (1983) *Acta Cryst.* **C39**, 1271-3

Clarke, H. T., Johnson, J. R. & Robinson, R., (Eds.), (1949) "The Chemistry of Penicillins", Princeton Univ. Press, Princeton, New Jersey, 472 et seq.,

Cram, D. J., Lingenfelter, D. S. & Helgeson, R. C., (1981) *J. Org. Chem.* **46**, 393-406

Davis, D. G., Live, D. H., Agosta, W. C. & Cowburn, D., (1982) *Org. Mag. Reson.* **19**(4), 211-5

Dawson, B., (1953) *Acta Cryst.* **6**, 81-7

- De Marco, A. & Llinas, M., (1980) *J. Magn. Reson.* **39**, 253-62
- DeWitt, H. D. & Ingersoll, A. W., (1951) *J. Am. Chem. Soc.* **73**, 5782-3
- Dobashi, A. & Hara, S., (1983) *Tetrahedron Letters*, **24**(14), 1509-10
- Dunitz, J. D. & Winkler, F. K., (1975) *Acta Cryst.* **B31**, 251-63
- Dunn, M. S. & Rockland, R. B., (1947) *Adv. Protein Chem.* **3**, 335-9
- Endo, T. Miyazawa, K., Endo, M., Uchida, A., Ohashi, Y. & Sasada, Y., (1982) *Chem. Lett.* **1982**, 1989-92
- Endo, T., Tasai, H., Miyazawa, K., Endo, M., Kato, K., Uchida, A., Ohashi, Y. & Sasada, Y., (1983) *J. Chem. Soc. Chem. Commun.* **1983**(11), 636-8
- Fischer, E., (1900) *Ber.* **33**, 3638
- Foldi, Z., (1954) *Acta Chim. Acad. Sci. Hung.* **5**, 187 (from *Chemical Abstracts*, **50**, 981i)
- Fraga, S., (1983) *J. Mol. Struct.*, **94**, 251-60 THEOCHEM
- Gorman, A. A., Hamblet, I., Smith, K. & Standen, M. C., (1984) *Tetrahedron Letters*, **25**(5), 581-4
- Gould, R. O., Cleasby, A., Moulden, N. & Walkinshaw, M. D., (1981) *Acta Cryst.* **A37**, C72 03. 3-11
- Gould, R. O., Taylor, P. & Walkinshaw, M. D., (1984) Manuscript in preparation
- Greenstein, J. P. & Winitz, M., (1961) "Chemistry of the Amino Acids", **1**, John Wiley (New York and London), 715 et seq.
- Greenstein, J. P., (1954) *Adv. Protein Chem.* **9**, 129-202

Hall, D. L. & Sanders, J. K. M., (1980) *J. Am. Chem. Soc.* **102**(18), 5703-11

Hossain, M. B., Van der Helm, D., Olsen, R. K., Jones P. G., Sheldrick, G. M., Egert, E. T., Kennard, O., Waring, M. J., Viswamitra, M
J. Am. Chem. Soc. **104**, 3401

Hyatt, J. A., Duesler, E. N., Curtin, D. Y. & Paul, K., (1980) *J. Org. Chem.* **43**, 5074-9

Ingersoll, A. W. & DeWitt, H. D., (1951) *J. Am. Chem. Soc.* **73**, 3360-2

Inoue, M. Shibata, M. & Ishida, T., (1980) *Biochem. Biophys. Res. Comm.* **93**(2), 415-9

Inoue, M., Hurano, H., Sugiyama, K. & Ishida, T., (1982) *Biochem. Biophys. Res. Comm.* **108**(2), 604-9

Inoue, M., Okuda, Y., Ishida, T. & Nakagaki, M., (1983) *Arch. Biochem. Biophys.* **227**(1), 52-70

Inoue, M., Sakaki, T., Fujiwara, T., & Tomita, K.-I., (1978ii) *Bull. Chem. Soc. Jpn.* **51**(4), 1123-7

Inoue, M., Sakaki, T., Fujiwara, T. & Tomita, K.-I., (1978i) *Bull. Chem. Soc. Jpn.* **51**(4), 1118-22

Inoue, M., Shibata, M., Kondo, Y.-I. & Ishida, T., (1981) *Biochemistry*, **20**(10), 2936-45

Isenberg, N. & Grdinic, M., (1972) *J. Chem. Ed.* **49**(6), 392-7

Ishida, T. & Inoue, M., (1981) *Acta Cryst.* **B37**, 2117-2120

Ishida, T. & Inoue, M., (1979) *J. Chem. Soc. Chem. Commun.* 358-360

Ishida, T. & Inoue, M., (1981) *Biochem. Biophys. Res. Comm.* **99**(1), 149-54

- Ishida, T., Inoue, M. & Tomita, K.-I., (1979) *Acta Cryst.* **B35**, 1642-8
- Ishida, T., Inoue, M., Senda, S. & Tomita K.-I., (1979) *Bull. Chem. Soc. Jpn.* **52**(10), 2953-8
- Ishida, T., Inoue, M., Ota, A., Kurihara, T., (1980) *J. Chem. Soc. Chem. Commun.* **1980**, 1074-5
- Ishida, T., Mitoguchi, S., Miyamoto, Y., Tomita, K.-I. & Inoue, M., (1980) *Biochim. Biophys. Acta*, **609**, 158-72
- Ishida, T., Miyamoto, Y., Nakamura, A. & Inoue, M., (1982) *Acta Cryst.* **B38**, 192-5
- Ishida, T., Tomita, K.-I., Inoue, M., (1980) *Arch. Biochem. Biophys.* **200**(2), 492-502
- Ishida, T., Usami, H., Inoue, M., Yamagata, Y. & Tomita, K.-I., (1982) *Biochem. Biophys. Res. Comm.* **107**(2), 746-751
- Ishida, T., Yamamoto, K. & Inoue, M., (1983) *Bull. Chem. Soc. Jpn.* **56**, 955-956
- Ishida, T., Yanagitani, T., Nakagawa, Y. & Inoue, M., (1981) *Biochem. Biophys. Res. Comm.* **99**(4), 1179-84
- Ivanov, V. T., Laine, I. A., Abdulaev, N. D., Senyavina, L. B., Popov, E. M., Ovchinnikov, Yu. A. & Shemyakin, M. M., (1969) *Biochem. Biophys. Res. Comm.* **34**, 803
- Karplus, M., (1959) *J. Chem. Phys.* **30**, 11
- Karplus, M., (1963) *J. Am. Chem. Soc.* **85**, 2870-1
- Kearley, F. J. & Ingersoll, A. W., (1951) *J. Am. Chem. Soc.* **73**, 5783-5

Kennard, O., Cruse, W. B. T., Viswamitra, M. A., Sheldrick, G. M. & Jones, P. G., (1980) *Am. Cryst. Assoc. Ser.* 2, 7, 19

Lipscomb, W. N., (1970) *Accts. Chem. Res.* 3, 81

Lipscomb, W. N. & Quioco, F. A., (1971) *Adv. Protein Chem.* 25, 1

Lipscomb, W. N., (1972) *Chem. Soc. Rev.* 3, 319

Llinas, M. & Klein, M. P., (1975) *J. Am. Chem. Soc.* 97, 4731-7

Llinas, M., De Marco, A. & Wuthrich, K., (1978) *Biopolymers*, 17, 637-50

Llinas, M., De Marco, A. & Wuthrich, K., (1978) *Biopolymers*, 17, 617-36

Llinas, M., Klein, M. P. & Neilands, J. B., (1970) *J. Mol. Biol.* 52, 399-414

MacNicol, D. D., McKendrick, J. J. & Wilson, D. R., (1978) *Chem. Soc. Rev.* 7, 65-87

Nakazaki, M., Yamamoto, K., Ikeda, T., Kitsuki, T. & Okamoto, Y., (1983) *J. Chem. Soc. Chem. Comm.* 1983, 787-8

Nardelli, M., Bigoli, F., Lanfranchi, M., Leporati, E. & Pellinghelli, M. A., (1981) *Acta Cryst.* B37, 1258-65

Newcomb, M., Helgeson, R. C. & Cram, D. J., (1974) *J. Am. Chem. Soc.* 96, 7367-9

Norrestam, R., Stensland, B. & Branden, C. I., (1975) *J. Mol. Biol.* 99, 501-6

Ohnishi, M. & Urry, D. W., (1969) *Biochem. Biophys. Res. Comm.* 36, 194

Ottvad, M., Hartter, P. & Jung, G., (1975) *Hoppe-Seyler's Z. Physiol. Chem.* 356, 1011-25

- Overby, L. R. & Ingersoll, A. W., (1951) *J. Am. Chem. Soc.* **73**, 3363-6
- Panijpan, B., (1977) *J. Chem. Ed.* **54**(11), 670-2
- Peerderman, A. F., (1956) *Acta Cryst.* **9**, 824
- Ptak, M., (1973) *Biopolymers*, **12**, 1575-89
- Queener, S. W. & Neuss, N., (1982) "Chemistry and Biology of, β -lactam Antibiotics", 3-Biochemistry 2-83 eds. Morin, R. B. & Gorman, M., (Academic Press, New York, N. Y.
- Ramachandran, G. N., Chandrasekaran, R. & Kopple, K. D., (1971) *Biopolymers*, **10**, 2113-31
- Richards, W. G., Clarkson, R. & Ganellin, C. R., (1975) *Phil. Trans. R. Soc. Lond.* **B272**, 75-85
- Robinson, R., (1947) *Nature*, **159**, 263 & *ibid*, **160**, 18
- Sanders, J. K. M. & Mersh, J. D., (1983) "Progress in NMR Spectroscopy", **15**, 353-400 (Pergamon Press, London)
- Scheraga, H. A. & Pincus, M. R., (1981) *Acc. Chem. Res.* **14**, 299-306
- Scott, A. I. & Gordon, E. M. G., (1973) Unpublished work, 1973
- Sealock, R. R., (1946) *J. Biol. Chem.* **166**, 1
- Sealock, R. R., Speeter, M. E. & Schweet, R. S., (1951) *J. Am. Chem. Soc.* **73**, 5386-88 19
- Shabica, A. C. & Tishler, M., (1949) *J. Am. Chem. Soc.* **71**, 3251
- Stern, A., Gibbons, W. A. & Craig, L. C., (1968) *Proc. Natl. Acad. Sci. U. S. A.*, **61**, 734

Stoddart, J. F., (1978) *Chem. Soc. Rev.* **8**, 85-151

Toda, F. & Tanaka, K., (1981) *Tetrahedron Letters*, **22**(46), 4669-72

Toda, F. & Tanaka, K., (1983) *J. Am. Chem. Soc.* **105**, 5151-2

Toda, F., Tanaka, K. & Mori, K., (1983) *Chem. Lett.* 827-30

Toda, F., Tanaka, K., Ueda, H., Oshima, T., (1983) *J. Chem. Soc. Chem. Commun.* **1983**, 743-4

Toda, F., Ward, D. L. & Hart, H., (1981) *Tetrahedron Letters*, **22**(39), 3865-8

Tonelli, A. E., (1978) *Biopolymers*, **17**, 1175-9

Toniolo, C., (1980) *CRC Crit. Rev. Biochem.* 1-44 (October)

Tsai, C.-C., Jain, S. C. & Sobell, H. M., (1975) *Phil. Trans. R. Soc. Lond. B*, **272**, 137-46

Urry, D. W., Venkatachalam, C. M., Khaled, M. A. & Sugano, H., (1981) *J. Am. Chem. Soc.* **103**, 2372-9

van Tamelin, E. E., Wagner, E. R., Sjoberg, B., Thelin, H. & Nathorst-Westfelt, L., (1965) *Tetrahedron Letters*, **4**, 281-6

Viswamitra, M. A., Kennard, O., Cruse, W. B. T., Egert, E., Sheldrick, G. M., Jones, P. G., Waring, M. J., Wakelin, L. P. G. & Olsen, *Nature*, **289**, 817-9

Walkinshaw, M. D. & Saenger, W., (1984) *J. Eur. Biochem.* submitted for publication

Williams, D. H., Hyde, E., Kalman, J. R., Reid, D. G. & Olsen, R. K., (1982) *J. Chem. Soc. Perkin I*, **1982**, 1041-8

Wipff, G., Kollman, P. A. & Lehn, J. M., (1983) *J. Mol. Struct.* **93**, 153-65
THEOCHEM

Wolfe, S., Bowers, R. J., Hasan, S. K. & Kazmaier, P. M., (1981) *Can. J. Chem.* **59**, 406-21

Woodward, R. B., Brehm, W. J. & Nelson, A. L., (1947) *J. Am. Chem. Soc.* **69**, 2250

Yang, C.-H., Brown, J. N., Kopple, K. D., (1981) *J. Am. Chem. Soc.* **103**(7), 1715-9

Zalkin, A., Forrester, J. D. & Templeton, D. H., (1966) *J. Am. Chem. Soc.* **88**, 1810-4

**APPENDIX
STRUCTURE FACTOR TABLES**

OBSERVED AND CALCULATED STRUCTURE FACTORS FOR CYCLIC SULPHIDE

PAGE 1

Table with 5 columns of structure factor data (H, K, L, IOFD, IOFC) for cyclic sulphide. Each column contains a series of values corresponding to different reflections.

OBSERVED AND CALCULATED STRUCTURE FACTORS FOR CYCLIC SULPHIDE

PAGE 3

Table with 5 columns of structure factor data (H, K, L, IOFD, IOFC) for cyclic sulphide. Each column contains a series of values corresponding to different reflections.

OBSERVED AND CALCULATED STRUCTURE FACTORS FOR CYCLIC SULPHIDE

PAGE 2

Table with 5 columns of structure factor data (H, K, L, IOFD, IOFC) for cyclic sulphide. Each column contains a series of values corresponding to different reflections.

OBSERVED AND CALCULATED STRUCTURE FACTORS FOR CYCLIC SULPHIDE

PAGE 4

Table with 5 columns of structure factor data (H, K, L, IOFD, IOFC) for cyclic sulphide. Each column contains a series of values corresponding to different reflections.

234

OBSERVED AND CALCULATED STRUCTURE FACTORS FOR DIMERIC DISULPHIDE

Table with 5 columns of data for observed and calculated structure factors for dimeric disulphide. Headers include H, K, L, IOFD, IOFC.

OBSERVED AND CALCULATED STRUCTURE FACTORS FOR DIMERIC DISULPHIDE

Table with 5 columns of data for observed and calculated structure factors for dimeric disulphide. Headers include H, K, L, IOFD, IOFC.

OBSERVED AND CALCULATED STRUCTURE FACTORS FOR DIMERIC DISULPHIDE

Table with 5 columns of data for observed and calculated structure factors for dimeric disulphide. Headers include H, K, L, IOFD, IOFC.

OBSERVED AND CALCULATED STRUCTURE FACTORS FOR DIMERIC DISULPHIDE

Table with 5 columns of data for observed and calculated structure factors for dimeric disulphide. Headers include H, K, L, IOFD, IOFC.

237

OBSERVED AND CALCULATED STRUCTURE FACTORS FOR DIMERIC DISULPHIDE

Table with 12 columns: H, K, L, IOFD, IOFC. It contains two main data blocks, each with 12 columns of values.

OBSERVED AND CALCULATED STRUCTURE FACTORS FOR DIMERIC DISULPHIDE

Table with 12 columns: H, K, L, IOFD, IOFC. It contains two main data blocks, each with 12 columns of values.

OBSERVED AND CALCULATED STRUCTURE FACTORS FOR DIMERIC DISULPHIDE

Table with 12 columns: H, K, L, IOFD, IOFC. It contains two main data blocks, each with 12 columns of values.

OBSERVED AND CALCULATED STRUCTURE FACTORS FOR DIMERIC DISULPHIDE

Table with 12 columns: H, K, L, IOFD, IOFC. It contains two main data blocks, each with 12 columns of values.

OBSERVED AND CALCULATED STRUCTURE FACTORS FOR

N-ACETYL METHIONINE

PAGE 1

Table with 10 columns: H, K, L, IOFD, IOFC, H, K, L, IOFD, IOFC, H, K, L, IOFD, IOFC, H, K, L, IOFD, IOFC. Contains observed and calculated structure factors for N-ACETYL METHIONINE.

OBSERVED AND CALCULATED STRUCTURE FACTORS FOR

N-ACETYL METHIONINE

PAGE 3

Table with 10 columns: H, K, L, IOFD, IOFC, H, K, L, IOFD, IOFC, H, K, L, IOFD, IOFC, H, K, L, IOFD, IOFC. Contains observed and calculated structure factors for N-ACETYL METHIONINE.

OBSERVED AND CALCULATED STRUCTURE FACTORS FOR

N-ACETYL METHIONINE

PAGE 2

Table with 10 columns: H, K, L, IOFD, IOFC, H, K, L, IOFD, IOFC, H, K, L, IOFD, IOFC, H, K, L, IOFD, IOFC. Contains observed and calculated structure factors for N-ACETYL METHIONINE.

OBSERVED AND CALCULATED STRUCTURE FACTORS FOR

N-ACETYL METHIONINE

PAGE 4

Table with 10 columns: H, K, L, IOFD, IOFC, H, K, L, IOFD, IOFC, H, K, L, IOFD, IOFC, H, K, L, IOFD, IOFC. Contains observed and calculated structure factors for N-ACETYL METHIONINE.

OBSERVED AND CALCULATED STRUCTURE FACTORS FOR N-ACETYL METHIONINE

PAGE 9

H		K		L		10FD		10FC		H		K		L		10FD		10FC		H		K		L		10FD		10FC	
1	1	7	69	71	-4	5	7	30	29	4	0	8	17	16	1	4	8	41	41	1	3	9	35	32					
2	1	7	87	86	-3	5	7	41	36	-4	1	8	15	30	2	4	8	26	25	2	3	9	36	33					
3	1	7	27	31	-2	5	7	35	37	-3	1	8	12	10	3	4	8	49	49	3	3	9	13	15					
4	1	7	47	47	-1	5	7	60	37	-2	1	8	38	43	-3	3	8	21	22	-3	4	9	15	19					
-5	2	7	23	25	0	5	7	13	10	-1	1	8	93	93	-2	5	8	17	19	-2	4	9	34	33					
-4	2	7	19	22	1	5	7	86	87	0	1	8	12	8	-1	5	8	27	27	-1	4	9	35	33					
-3	2	7	45	54	2	5	7	25	23	1	1	8	65	61	1	5	8	33	33	0	4	9	13	15					
-2	2	7	81	83	3	5	7	16	19	2	1	8	90	88	2	5	8	36	37	1	4	9	21	19					
-1	2	7	106	104	4	5	7	48	47	3	1	8	20	22	-2	6	8	32	29	-2	5	9	18	16					
0	2	7	67	63	-3	6	7	37	34	4	1	8	64	58	-1	6	8	14	11	-1	5	9	24	24					
1	2	7	34	31	-2	6	7	39	28	-4	2	8	28	27	0	6	8	31	27	0	5	9	44	42					
2	2	7	74	79	-1	6	7	27	28	-3	2	8	34	36	1	6	8	29	29	1	5	9	20	21					
3	2	7	60	61	0	6	7	93	91	-2	2	8	15	19	2	6	8	32	33	2	5	9	28	28					
-3	3	7	17	16	1	6	7	42	42	-1	2	8	27	30	-1	7	8	38	34	-1	6	9	17	17					
-4	3	7	12	6	2	6	7	39	37	0	2	8	87	83	1	7	8	23	18	-2	0	10	34	34					
-3	3	7	52	56	3	6	7	23	25	1	2	8	83	82	-2	0	9	12	13	-1	0	10	92	91					
-2	3	7	117	116	-3	7	7	25	23	2	2	8	41	44	-1	0	9	70	67	0	0	10	23	26					
-1	3	7	101	99	-2	7	7	31	31	3	2	8	51	49	0	0	9	16	17	1	0	10	15	15					
0	3	7	76	77	-1	7	7	36	36	4	2	8	28	31	1	0	9	90	88	2	0	10	61	56					
1	3	7	103	101	0	7	7	24	26	-4	3	8	24	25	-3	1	9	17	23	-2	1	10	35	34					
2	3	7	42	45	1	7	7	45	46	-3	3	8	25	21	-2	1	9	36	38	-1	1	10	32	32					
3	3	7	35	38	2	7	7	37	40	-2	3	8	19	21	0	1	9	93	89	0	1	10	75	72					
4	3	7	44	43	-1	8	7	31	19	-1	3	8	48	50	1	1	9	38	34	1	1	10	28	24					
-4	4	7	33	27	0	8	7	45	47	0	3	8	19	13	2	1	9	26	26	-1	2	10	39	39					
-3	4	7	28	33	-2	0	8	18	41	1	3	8	96	97	-3	2	9	17	20	0	2	10	33	30					
-2	4	7	28	29	-3	0	8	12	15	2	3	8	35	34	-2	2	9	21	17	1	2	10	47	45					
-1	4	7	91	93	-2	0	8	91	90	3	3	8	26	24	-1	2	9	68	61	2	2	10	39	37					
0	4	7	64	65	-1	0	8	19	18	4	3	8	46	44	1	2	9	20	21	-1	3	10	22	23					
1	4	7	40	39	0	0	8	59	62	-3	4	8	22	23	2	2	9	22	20	0	3	10	24	23					
2	4	7	74	72	1	0	8	86	85	-2	4	8	40	38	-2	3	9	44	42	1	3	10	32	31					
3	4	7	68	62	2	0	8	50	46	-1	4	8	14	11	-1	3	9	29	32	0	4	10	22	21					
4	4	7	14	16	3	0	8	91	92	0	4	8	51	50	0	3	9	14	12	1	4	10	37	34					

OBSERVED AND CALCULATED STRUCTURE FACTORS FOR dl Glutamic Acid

Table with 12 columns: H, K, L, I0FO, IOFC, H, K, L, I0FO, IOFC, H, K, L, I0FO, IOFC, H, K, L, I0FO, IOFC, H, K, L, I0FO, IOFC. Contains multiple rows of numerical data.

OBSERVED AND CALCULATED STRUCTURE FACTORS FOR dl Glutamic Acid

Table with 12 columns: H, K, L, I0FO, IOFC, H, K, L, I0FO, IOFC, H, K, L, I0FO, IOFC, H, K, L, I0FO, IOFC, H, K, L, I0FO, IOFC, H, K, L, I0FO, IOFC. Contains multiple rows of numerical data.

OBSERVED AND CALCULATED STRUCTURE FACTORS FOR dl Glutamic Acid

Table with 12 columns: H, K, L, I0FO, IOFC, H, K, L, I0FO, IOFC, H, K, L, I0FO, IOFC, H, K, L, I0FO, IOFC, H, K, L, I0FO, IOFC, H, K, L, I0FO, IOFC. Contains multiple rows of numerical data.

OBSERVED AND CALCULATED STRUCTURE FACTORS FOR dl Glutamic Acid

Table with 12 columns: H, K, L, I0FO, IOFC, H, K, L, I0FO, IOFC, H, K, L, I0FO, IOFC, H, K, L, I0FO, IOFC, H, K, L, I0FO, IOFC, H, K, L, I0FO, IOFC. Contains multiple rows of numerical data.

OBSERVED AND CALCULATED STRUCTURE FACTORS FOR Strychnine

Table with 12 columns: H, K, L, 10FO, 10FC and 3 rows of data.

PAGE 1

OBSERVED AND CALCULATED STRUCTURE FACTORS FOR Strychnine

Table with 12 columns: H, K, L, 10FO, 10FC and 14 rows of data.

PAGE 3

OBSERVED AND CALCULATED STRUCTURE FACTORS FOR Strychnine

Table with 12 columns: H, K, L, 10FO, 10FC and 24 rows of data.

PAGE 2

OBSERVED AND CALCULATED STRUCTURE FACTORS FOR Strychnine

Table with 12 columns: H, K, L, 10FO, 10FC and 24 rows of data.

PAGE 4

944

OBSERVED AND CALCULATED STRUCTURE FACTORS FOR Strychnine

Table with columns H, K, L, IOFD, IOFC and multiple rows of numerical data representing structure factors.

OBSERVED AND CALCULATED STRUCTURE FACTORS FOR Strychnine

Table with columns H, K, L, IOFD, IOFC and multiple rows of numerical data representing structure factors.

OBSERVED AND CALCULATED STRUCTURE FACTORS FOR Strychnine

Table with columns H, K, L, IOFD, IOFC and multiple rows of numerical data representing structure factors.

OBSERVED AND CALCULATED STRUCTURE FACTORS FOR Strychnine

Table with columns H, K, L, IOFD, IOFC and multiple rows of numerical data representing structure factors.

OBSERVED AND CALCULATED STRUCTURE FACTORS FOR BRYCHNINE NITRATE

Table with 5 columns of H, K, L, IOFO, IOFC values for Brychnine Nitrate. The table is organized into five groups of columns, each with 5 sub-columns.

PAGE 5

OBSERVED AND CALCULATED STRUCTURE FACTORS FOR BRYCHNINE NITRATE

Table with 5 columns of H, K, L, IOFO, IOFC values for Brychnine Nitrate. The table is organized into five groups of columns, each with 5 sub-columns.

PAGE 7

OBSERVED AND CALCULATED STRUCTURE FACTORS FOR BRYCHNINE NITRATE

Table with 5 columns of H, K, L, IOFO, IOFC values for Brychnine Nitrate. The table is organized into five groups of columns, each with 5 sub-columns.

PAGE 6

OBSERVED AND CALCULATED STRUCTURE FACTORS FOR Brucine Free Base Ethanol/water solvate

PAGE 9

H K L 10FO 10FC		H K L 10FO 10FC		H K L 10FO 10FC		H K L 10FO 10FC		H K L 10FO 10FC	
0 7 17 230 232	0 3 18 167 190	2 10 18 77 67	6 6 19 92 98	3 3 20 201 204					
2 7 17 144 142	1 3 18 203 204	1 0 19 61 70	0 7 19 118 117	4 3 20 174 182					
3 7 17 84 99	2 3 18 90 91	2 0 19 38 54	2 7 19 87 78	0 4 20 100 106					
5 7 17 75 68	3 3 18 132 130	3 0 19 116 114	3 7 19 91 81	1 4 20 118 121					
1 8 17 98 86	4 3 18 88 85	5 0 19 240 239	0 8 19 117 123	2 4 20 83 95					
2 8 17 122 121	5 3 18 80 81	6 0 19 119 121	1 8 19 99 96	3 4 20 81 78					
3 8 17 94 84	6 3 18 90 96	7 0 19 103 106	2 8 19 79 80	0 5 20 357 359					
4 8 17 131 133	7 3 18 109 107	1 1 19 132 126	3 8 19 48 55	1 5 20 179 188					
5 8 17 45 41	0 4 18 90 95	3 1 19 58 49	3 9 19 104 96	2 5 20 224 228					
1 9 17 64 57	1 4 18 49 96	4 1 19 168 171	0 10 19 137 137	3 5 20 170 170					
2 9 17 129 124	2 4 18 115 120	5 1 19 62 62	1 10 19 84 77	0 6 20 67 79					
3 9 17 90 90	3 4 18 164 161	6 1 19 104 97	0 11 19 111 110	2 6 20 89 99					
4 9 17 109 105	4 4 18 35 49	2 2 19 219 222	1 11 19 32 46	5 6 20 52 62					
0 10 17 111 103	5 4 18 72 73	3 2 19 167 163	0 0 20 300 302	1 7 20 92 95					
1 10 17 101 99	0 5 18 273 281	4 2 19 168 149	1 0 20 466 490	3 7 20 101 113					
2 10 17 36 32	1 5 18 103 99	6 2 19 71 69	2 0 20 237 220	4 7 20 109 103					
0 11 17 70 65	2 5 18 204 206	0 3 19 71 79	3 0 20 81 75	0 8 20 89 77					
2 11 17 94 84	3 5 18 75 75	1 3 19 61 53	5 0 20 143 139	1 8 20 121 131					
1 0 18 115 108	1 6 18 237 227	2 3 19 149 147	6 0 20 85 83	2 8 20 64 70					
2 0 18 77 75	2 6 18 90 82	3 3 19 104 107	0 1 20 165 165	3 8 20 113 115					
3 0 18 70 78	3 6 18 67 67	5 3 19 166 169	1 1 20 313 317	4 8 20 61 60					
4 0 18 180 172	4 6 18 59 48	0 4 19 71 82	1 2 20 208 216	0 9 20 82 92					
5 0 18 91 83	0 7 18 57 46	1 4 19 298 304	3 1 20 75 80	1 9 20 79 74					
1 1 18 75 78	1 7 18 168 175	2 4 19 267 269	5 1 20 46 47	1 10 20 93 90					
3 1 18 196 210	2 7 18 169 167	3 4 19 156 142	0 2 20 524 531	2 10 20 64 59					
4 1 18 77 72	3 7 18 132 130	4 4 19 157 144	1 2 20 48 48	1 0 21 81 79					
5 1 18 79 75	4 7 18 136 150	6 4 19 93 79	2 2 20 367 379	2 0 21 83 77					
7 1 18 102 94	5 7 18 95 90	0 5 19 219 222	3 2 20 86 90	3 0 21 226 223					
0 2 18 173 170	4 8 18 90 90	2 5 19 92 99	4 2 20 187 183	5 0 21 130 124					
1 2 18 126 130	5 8 18 96 91	3 5 19 157 150	5 2 20 58 64	6 0 21 94 93					
2 2 18 143 148	0 9 18 72 80	4 5 19 129 125	6 2 20 59 73	1 1 21 96 89					
4 2 18 79 70	1 9 18 69 74	5 5 19 72 68	0 3 20 261 272	2 1 21 195 203					
5 2 18 113 107	2 9 18 59 58	0 6 19 101 103	1 3 20 95 98	3 1 21 172 176					
7 2 18 90 87	1 10 18 66 65	1 6 19 304 307	2 3 20 200 197	4 1 21 73 73					

OBSERVED AND CALCULATED STRUCTURE FACTORS FOR Brucine Free Base Ethanol/water solvate

PAGE 10

H K L 10FO 10FC		H K L 10FO 10FC		H K L 10FO 10FC		H K L 10FO 10FC		H K L 10FO 10FC	
5 1 21 104 97	2 7 21 48 54	3 7 22 89 87	3 8 23 71 59	1 0 26 84 89					
6 1 21 62 57	4 8 21 75 74	1 8 22 96 100	0 0 24 344 361	4 0 26 47 37					
1 2 21 233 225	3 9 21 88 89	1 0 23 89 84	2 0 24 240 248	0 2 26 72 69					
2 2 21 116 127	0 0 22 132 127	3 0 23 81 74	3 0 24 84 91	1 2 26 43 43					
3 2 21 281 277	1 0 22 88 81	0 1 23 264 277	4 0 24 108 109	2 2 26 65 52					
4 2 21 152 161	2 0 22 105 106	1 1 23 75 76	0 1 24 87 98	0 3 26 102 114					
5 2 21 140 140	3 0 22 219 230	2 1 23 188 192	1 1 24 88 85	0 4 26 46 57					
6 2 21 90 87	4 1 22 288 301	4 1 23 75 82	2 1 24 60 90	3 4 26 91 44					
1 3 21 131 134	1 1 22 248 247	5 1 23 114 115	4 1 24 81 85	2 0 27 79 67					
2 3 21 161 157	2 1 22 243 242	0 2 23 133 135	5 1 24 58 55	0 1 27 98 40					
4 3 21 147 148	4 1 22 232 224	1 2 23 162 149	2 2 24 85 88	1 1 27 125 117					
5 3 21 104 105	0 2 22 137 150	2 2 23 177 179	3 2 24 59 51	1 2 27 97 86					
6 3 21 82 78	1 2 22 84 81	3 2 23 80 78	5 2 24 75 64	3 2 27 74 69					
0 4 21 80 64	2 2 22 159 152	5 2 23 95 76	2 3 24 49 62	0 3 27 72 75					
1 4 21 178 175	3 2 22 149 157	0 3 23 106 126	3 3 24 122 120	2 3 27 81 77					
2 4 21 123 129	4 2 22 57 63	2 3 23 174 168	4 3 24 97 102	3 3 27 44 41					
3 4 21 162 171	2 3 22 303 306	3 3 23 151 146	4 4 24 65 60	2 4 27 81 60					
4 4 21 76 71	3 3 22 103 95	4 3 23 103 110	0 5 24 50 50	2 5 27 68 62					
6 4 21 62 72	4 3 22 150 146	5 3 23 68 60	0 7 24 106 104	0 0 28 73 69					
0 5 21 144 149	5 3 22 92 89	3 4 23 82 84	1 8 24 57 57	1 0 28 151 152					
1 5 21 99 91	0 4 22 153 158	4 4 23 108 100	1 0 25 126 125	3 0 28 59 64					
2 5 21 158 169	1 4 22 91 86	0 5 23 116 114	3 0 25 79 73	2 1 28 57 49					
3 5 21 95 90	2 4 22 103 111	1 5 23 85 70	1 1 25 67 87	3 1 28 57 54					
4 5 21 61 62	3 4 22 69 73	2 5 23 69 72	1 2 25 48 54	1 2 28 67 73					
0 6 21 201 196	1 5 22 153 163	4 5 23 91 91	0 2 25 114 113	1 3 28 55 54					
1 6 21 139 140	2 5 22 53 58	0 6 23 120 121	2 2 25 62 57	2 3 28 66 64					
3 6 21 82 74	4 5 22 47 40	1 6 23 83 69	4 2 25 76 68	0 1 29 148 148					
4 6 21 104 105	5 5 22 70 58	3 6 23 81 88	1 4 25 49 44	1 1 29 57 52					
5 6 21 73 74	2 6 22 86 85	4 6 23 66 67	4 4 25 62 61	2 1 29 87 83					
0 7 21 113 110	4 6 22 59 43	0 7 23 54 50	3 5 25 49 21	1 3 29 59 60					
1 7 21 79 81	2 7 22 109 110	2 8 23 56 61	3 6 25 51 49						

OBSERVED AND CALCULATED STRUCTURE FACTORS FOR StrYchnine Tyrosine fixed water

PAGE 8

OBSERVED AND CALCULATED STRUCTURE FACTORS FOR StrYchnine Tyrosine fixed water

Table with 10 columns (H, K, L, IOFC, IOFC) and multiple rows of data points for various reflections.

Table with 10 columns (H, K, L, IOFC, IOFC) and multiple rows of data points for various reflections.

OBSERVED AND CALCULATED STRUCTURE FACTORS FOR StrYchnine Tyrosine fixed water

PAGE 4

Table with 10 columns (H, K, L, IOFC, IOFC) and multiple rows of data points for various reflections.