A CRYSTALLOGRAPHIC STUDY OF SOME BIOLOGICALLY IMPORTANT COMPOUNDS

by

Philip Richard Lowe

A thesis presented for the degree of

DOCTOR OF PHILOSOPHY

in the

University of Aston in Birmingham

THE UNIVERSITY OF ASTON IN BIRMINGHAM
A CRYSTALLOGRAPHIC STUDY OF SOME BIOLOGICALLY IMPORTANT COMPOUNDS

Name: PHILIP RICHARD LOWE
Degree: DOCTOR OF PHILOSOPHY
Year: 1984

Summary

Single crystal X-ray structure determinations are reported for eleven compounds, each of which is either biologically active or potentially biologically important. The compounds fall into four distinct classes:-

1. Stable salts of benzylpenicillin
2. Analogues of folic acid and folate antagonists
3. Nucleoside base analogues
4. Novel bicyclic antitumour agents.

A common theme throughout the investigation is the nature, importance and possible implications of the hydrogen bonds formed in each of the compounds; the biological importance and/or medicinal interest in each of the above classes is discussed and related to the structures of the individual compounds.

Procaine penicillin G and benethamine penicillin G are stable salts of benzylpenicillin, administered as repository drugs in aqueous crystalline suspensions. Their stability to aqueous environments is related to their crystal structures.

Because of the importance of folic acid as a target for chemotherapy a number of analogues and antagonists have been synthesised at Aston. The structures of two such compounds, one a 2:4-diaminopyrimidine, the other a 2-amino-4-oxopyrimidine are discussed in Chapter 4 paying particular interest to base pair type interactions, interactions with counter ions and solvent molecules and the geometry of the pyrimidine ring. The structure of a sulphonamide is also reported in an attempt to correlate this with its microbiological properties.

The structures of the nucleoside base analogues 6-phenyl-5-azauracil, 6-benzyl-5-azauracil and 6-methylisocytosine have been determined in order to confirm: a) the tautomeric form in each case, b) to determine the nature of hydrogen bond interactions involved particularly the type of base pair formation which is dependent upon the order of proton donors and acceptors on the heterocyclic ring, c) to determine in the case of the diketo compounds whether the two carbonyl stretching frequencies observed in the infrared spectrum are due to non-equivalence, or vibrational coupling.

The structures of mitozolomide (a novel bicyclic antitumour agent developed at Aston University) and two of its analogues are discussed, not only to provide unequivocal proof of the structures but also to provide additional information which may assist in determining the possible modes of action. The structures are compared with those of known similar antitumour agents, eg. DTIC and DCTIC (see list of abbreviations). The crystallographic software at Aston and Manchester (UMRCC) is reviewed.

KEY WORDS: BENZYLPENICILLIN   NUCLEOSIDE
                 FOLATE   MITOZOLOMIDE

Acknowledgements

To Jean,Scott and Adrian

List of Contents

Abbreviations

AIC	azidoimidazocarboxamide
APPG	aqueous procaine penicillin G
BCNU	1,3-bis(2-chloroethyl)-1-nitrosourea
BPG	benethamine penicillin G
CCNU	(2-chloroethyl)-3-cyclohexyl-1-nitrosourea
CCRG	Cancer Chemotherapy Research Group
CCRG 81010	patent on mitozolomide
DCTIC	5-(3,3-chloroethyltriazenyl)imidazole-4-carboxamide
DHFR	dihydrofolate reductase
DNA	deoxyribonucleic acid
DMSO	dimethylsulphoxide
DTIC	5-(3,3-dimethyltriazenyl)imidazole-4-carboxamide
dTMP	deoxythymidylic acid monophosphate
dUMP	deoxyuridilic acid monophosphate
FAH_2	dihydrofolate
FAH_4	tetrahydrofolate
MTIC	5-(3-methyltriazenyl)-imidazole-4-carboxamide
MTX	methotrexate
NADP	nicotinamide adenine dinucleotide
NCI	National Cancer Institute
NPPHCl	nitropiperidinopyrimethamine hydrochloride
PABA	p-aminobenzoic acid
RNA	ribonucleic acid
SHMT	serinehydroxymethyltransferase
TS	thymidylate synthetase
UMRCC	University of Manchester Regional Computer Centre

CHAPTER 1

Chapter 1

1.1   INTRODUCTION

Single crystal x-ray crystallography is a technique that is now used almost routinely as a method of determining the three dimensional structure and molecular dimensions of organic molecules. Well refined structures provide unambiguous information on the connectivity of organic molecules in the solid state and as such the technique is often used to confirm or refute or merely add additional information to structures proposed on the basis of synthetic pathways, or genuinely routine analytical procedures such as mass spectrometry, NMR, infrared and ultraviolet spectroscopy and CHNO microanalysis.

The statement that the technique is used almost routinely is by no means meant to imply that it is a routine procedure as are some of the other techniques mentioned. It is only the recent developments of both hardware and software that have enabled crystallography to possibly provide some information in a time period that is an order of magnitude longer than it takes to complete the conventional analytical procedures.

Current research in the field involves the development of new, and enhancement of existing procedures for solving and refining structures and applying the necessary corrections to the data. Equally important, however, is the working crystallographer who uses these methods to solve structures; each is dependent on the other. Structure determination can very broadly be split into two categories - the medium to large structures eg. carbohydrates, polypeptides, proteins or nucleic acids - and the smaller structures with, perhaps, 50 atoms or less in the asymmetric unit. The two are in many ways symbiotic, in that the substrate is equally as important as the

protein. However, since the hardware and software requirements differ somewhat between the two, dedication in one or the other seems inevitable.

Crystallography at Aston both past and present has involved the determination of the structure of small molecules, the largest of which contain approximately 40 non-hydrogen atoms per asymmetric unit. In the early stages, only film methods were available, and data collection alone could take 2-3 months or longer with the possible outcome being a difficult and disappointing refinement. However, the recent establishment of a Joint Crystallography Unit with the University of Birmingham Chemistry Department and the subsequent purchase of an Enraf-Nonius CAD4 Diffractometer has meant that both output and standard has improved significantly.

The reasons for determining the exact molecular structure of a compound may be many and varied. In the case of newly developed drugs for example, it may be important in establishing patents, eg. mitozolomide (see Chapter 6). Occasionally, conventional analytical procedures are insufficient to differentiate between isomeric forms particularly geometrical (trans and cis) stereo isomers and even rotamers as with mitozolamide. Similarly, the correct identification of tautomeric form, particularly where in the solid state the compound contains a 1:1 ratio of the two tautomers eg. isocytosine (Chapter 5) is significantly enhanced by this technique. Often, however, besides the details of connectivity and precise conformation the crystallographer is seeking to elicit additional information: In what type of inter- and intramolecular interactions does the molecule participate in the solid state, ie. hydrogen bonds either to other molecules (inter), internal hydrogen bonds (intra) or to counter ions or solvent molecules? Is the molecule planar or non-planar? Does it

readily stack and are stacking forces involved? What other conformations are available by simple rotation about single bonds and which are not possible due to physical constraint such as steric hindrance?

Possibly the most common criticism directed at the crystallographer is that his information is limited to the solid state. In compounds of pharmacological interest - the subject of this thesis - the conformation and even isomeric form that the molecule may exhibit in its biologically active form may be different. The skill, however, is in interpreting the more subtle clues from the crystal determination and postulating possible conformations and interactions in, for example, the active site of an enzyme.

Work in this area has been significantly enhanced by the use of computer graphics where the known structure of an enzyme can be displayed and the structures of possible substrates or antagonists inserted in the active site. By varying the conformation of the latter molecules and manipulating their position within possible active sites, one can postulate probable interactions. Such an extension of the crystallographer's work is invaluable in drug design, replacing to some extent the need for the synthetic organic chemist to make all possible derivatives of an active compound.

The macromolecular structure determinations and enzyme/antagonist complexes (eg. dihydrofolate reductase (DHFR)-methotrexate) have also helped to allay this criticism; unfortunately such co-crystallisation is not always possible and the resolution of the structure often precludes the fine detail. Small molecule crystallographic investigations thus complement those of the enzyme or enzyme inhibitor complex.

1.2   What are the limitations?

Undoubtedly, crystallography provides information at a molecular level which is unobtainable by other conventional analytical procedures. Using high powered computing facilities, sophisticated software and precision instrumentation to mimic the action of a lens, an image of the structure is produced from which detailed molecular dimensions and geometry can be obtained. This computed image is of course only a model, and the degree to which it is either basically correct in terms of connectivity or geometry, or precise dimensionally is reflected in the difference between the observed and calculated structure amplitudes normally reported as an R factor.

Bond lengths and angles need to be interpreted with due consideration of the precision and overall accuracy of the determination and the concept of absolute configuration should not be neglected (if appropriate) where geometry is concerned.

There are, however, two further weaknesses which limit the effectiveness of x-ray crystallography in the investigation of small molecule structures from the point of view of pharmacological interest. The first of these is that nitrogen atoms may not be unambiguously distinguishable from carbon atoms; an important feature in the study of heterocyclic systems where the position of the nitrogen atoms is essential to the activity of the molecule.

Here, both routine analytical procedures (if the information is available) and bond lengths may be of considerable aid in interpreting the results. The order of atom-atom distances is C≡C> C≡N > N≡N, this inequality being dependent on approximately equal π-bonding contribution in all three cases; in heterocyclic systems this is unlikely, and bond distances may only be taken as a guide to the identification of nitrogen positions. During refinement, the

insertion of all heterocyclic atoms as carbons should result in unusually low temperature factors at the nitrogen positions; unfortunately, this effect is often masked by the variation in oscillation at different points in the heterocycle. Nevertheless, comparison of isotropic temperature factors (Uiso) for atoms at similar positions should show differences of approximately $0.02\overset{o}{A}^2$ between nitrogen and carbon atoms. The combination of these methods together with synthetic pathway usually gives a clear indication of the positions of the nitrogen atoms.

A second weakness in x-ray structure analysis is the question of hydrogen atom positions. Given high-quality intensity data hydrogen positions can be expected to appear in a difference Fourier synthesis with intensities in the range 0.3 to 0.4 $e\overset{o}{A}^{-3}$ but may well be above or below these limits. Poor quality data or large thermal parameters or indeed a combination may make hydrogen positions unobservable. This problem is important where the position of protonation is essential to understanding of the biochemical activity of the molecule. The best solution to this problem provided a satisfactory crystal can be grown and beam facilities are available, is to collect neutron diffraction intensity data, where because hydrogen has a scattering length approximately one half in magnitude to that of carbon, hydrogen atomic positions should be easily visible in a Fourier synthesis. Failure to obtain either suitable crystals or the beam facility would normally result in one placing those unobservable hydrogens in calculated positions and then attempting to refine them isotropically. This would only be done if one was sure of their existence from other analytical procedures.

1.3 The crystallography of biologically important compounds - this work

How does one decide upon which compounds to examine crystallographically, and what are the reasons for studying the structure?

From experience, it appears that there are three <u>major</u> possibilities:-

1. That the crystallographer seeks a theme and then studies a range of compounds which have characteristics that fit the general theme.

2. That with a particular known compound or group of compounds, there is an interesting but unanswered problem.

3. That for a particular new compound or group of compounds perhaps produced within the Department, the synthetic chemists and biochemists wish certain structural questions to be answered; this in some cases amounts almost to offering a service.

Under any circumstances the ways that structural problems arrive at the crystallographer's door are varied and because of this, often the overall theme to his or her work may be very broad or general, eg. 'The Crystallography of Compounds of Pharmacological Interest' or, 'The Crystallography of Transition Metal Complexes'.

In this work, the structures that have been determined cover the range of possibilities mentioned above. All of them are of biological importance, and over half are of pharmacological interest. The major theme, however, throughout the thesis is that of hydrogen bonding, particularly inter-molecular interactions, this of course being vitally important in the mode of action of so many drugs and playing such a vital part in the biological importance of many of

the molecules small or large found in the mammalian cell. Within this broad theme, the structures do fall into four logical cateogories:

1) The Penicillins

2) Analogues of folic acid and folate antagonists

3) Nucleoside base analogues

4) Novel antitumour drugs

and as such a chapter is devoted to each of these areas, each chapter having its own formal introduction and conclusions.

1.3.1   The complexed salts of benzylpenicillin

Both procaine pencillin G and benethamine penicillin G are complexed salts of benzylpenicillin and show considerable stability to aqueous environments, a feature which is used in their action as repository drugs for the slow release of benzylpenicillin. Since both drugs are used as crystalline suspensions, and administered this way by intramuscular injection it was important to determine what solid state interactions imposed such stability on these two compounds (see Chapter 3).

1.3.2   Analogues of folic acid and folate antagonists

Folic acid - vitamin Bc - (Figure 1, Chapter 4) is an important factor to mammals and micro-organisms alike, its importance being embodied in its reduced form FAH_4 (5,6,7,8-tetrahydrofolic acid) which is involved in, for example, purine and pyrimidine synthesis, these bases being essential precursors for DNA biosynthesis, and also amino acid interconversions.

Because of its obvious importance in the growth of normal cells, and also the proliferation of abnormal cells it is an important target for chemotherapy. A number of analogues of folic acid and folate antagonists have been synthesised in the Department, and the structures of two such compounds one containing the 2,4-diamino-pyrimidine moeity, the other the 2-amino-4-oxo-pyrimidines are discussed in Chapter 4. The work is part of a study on modified pyrimethamine structures as inhibitors of DHFR (Dihydrofololate reductase) with reference to the enzyme from protozoal sources.

Particular interest is paid to the nature of the hydrogen bonding, especially base pair type interactions and interactions with counter ions and solvent molecules, the site of protonation in salt formation, and the geometry and dimensions of the pyrimidine ring and its 2,4 attachments. The third structure in Chapter 4 although not an inhibitor of DHFR is a member of a group of compounds, the sulphonamides, which are known to inhibit the production of folic acid in bacteria. Its structure was studied in parallel with its microbiological properties in an attempt to correlate the two.

1.3.3 Nucleoside base analogues

Two of the compounds studied in Chapter 5, namely the 6-phenyl and 6-benzyl-1,3,5,triazin-2,4[1H,3H]-diones (trivial nomenclature 6-phenyl and 6-benzyl-5-azauracils) attracted interest for a number of reasons:-

i)   Although these have not shown much in the way of biological activity, some azauracils have shown considerable promise in cancer chemotherapy.

ii) Since the carbonyl groups in such compounds may be

implicated in base pairing and are likely to be important in binding to enzymes it was important to confirm the tautomeric forms suggested by spectroscopic analysis and to compare the intermolecular interactions with those of similar compounds.

iii)   Both compounds show two peaks in the carbonyl stretching region of the infrared spectrum. It was important to establish whether these were due to non - equivalence of the two C=O bonds, or merely vibrational coupling.

The importance of the last structure in Chapter 5 - 6-methyl-isocytosine is embodied in:-

i)   its similarity to the natural precursors of nucleic acids

ii)   its similarity to the natural substrates and inhibitors of dihydrofolate reductase.

Since it is possible for it to exist in three tautomeric forms (see Chapter 5, Figure 6) it was important to establish which of these exists in the solid state and what effect this has on its intermolecular interactions. From the arrangement of hydrogen bond donors and acceptors in and attached to the pyrimidine/triazine ring in such compounds, it is possible with some degree of accuracy to predict the nature of the interactions particularly in terms of base pairing ie. single inverse base pair, double inverse base pair or Watson and Crick base pair . Also, assuming 6-methylisocytosine existed in the keto form as spectroscopic data suggested, it appeared an ideal control compound having just a single carbonyl group to compare its C=O stretching frequency with those of the diketo compounds.

1.3.4   Novel anti-tumour drugs

Controlling cancer is currently a multifaceted attack of which chemotherapy is just one of the components used. The Department of Pharmaceutical Sciences at Aston has a large Cancer Chemotherapy Research Group, researching into the synthesis, mode of action and formulation of novel antitumour drugs.

The first significant success in animal trials came from mitozolomide — 8-carbamoyl-3-(2-chloroethyl)imidazo [5,1-d]-1,2,3,5-tetrazin-4(3H)-one, compound (1), Chapter 6, a novel bicyclic antitumour agent synthesised and screened at Aston. This was quickly followed by the synthesis of a number of analogues of mitozolamide most of which involved the replacement of the 2-chloroethyl group with a different side chain. Three structures are reported in this chapter, namely mitozolamide, the 3-methyl and the 3-methoxyethyl analogues. The structure of mitozolamide was determined because of its obvious importance as an antitumour agent, and to establish unambiguously its molecular structure because of commercial and clinical interest in it. The two analogues are of structural interest in that they maintain the essential features of mitozolamide with the only variant being the N(3) substituent; their activities were however substantially reduced, the 3-methyl compound still active, but the 3-methoxyethyl compound having no activity at all.

The crystal structures of each of these compounds has been studied in this work not only in order to provide unequivocal proof of structure, but also in the hope of providing additional information which may assist in determining possible modes of action.

CHAPTER 2

CHAPTER 2   CRYSTALLOGRAPHIC SOFTWARE

2.1   All of the early structural work at Aston was performed on the University's ICL 1904S computer using software implemented by Dr C H Schwalbe; this included data reduction, structure solution and least squares refinement. With the advent of the CDC7600 at UMRCC the strategy changed, such that final anisotropic refinement was performed on this computer rather than at Aston.

At this stage, the only supported crystallographic package was X-RAY 72. However, the subsequent introduction of the SWAN system greatly improved the reliability of the link between Manchester and the local computer and reduced turnaround times considerably. This, together with a considerably improved range of crystallographic packages available on the CDC7600 meant that with the exception of data reduction, all structural work from Aston could now be computed at UMRCC (see Table 2.1).

Table 2.1   Software packages used at UMRCC on the CDC7600

Programme Name	Use
MULTAN 78[1]	Normalised structure factors
(including normal)	and direct methods
EEES (SHELX)[2]	Centrosymmetric direct methods
LS (SHELX)[2]	Least squares refinement
XRAY[3]	Least squares refinement
GEOM78	Least squares planes
	Torsion Angles, Bond Angles
	and distances
Pluto 78[4]	All molecular diagrams

RETRIEV	Crystallographic literature search
CONNSER	Structure connectivity search

2.2 Data Reduction

2.2.1 FSC

The data reduction on the early structures studied, using film and two circle diffractometer data, was performed at Aston on the 1904S using the programme FSC which is based on the Hamilton Rollet and Sparkes film scaling algorithm[5]. Its purpose is to take unscaled data in the form of h, k, 1 and I and together with suitable scaling factors place the data on a common scale and output it in the form of h, k, 1, F and σF. It does this by iteratively refining scale factors input by the user for each level or sublevel of data and refines to a scale factor which has been fixed for standard reflections. The refinement procedure terminates when the ratio of the shift/parameter reaches a predetermined level. The programme follows this by averaging equivalent and identifying unobserved reflections in terms of a limit preset by the operator. Lorentz and polarisation corrections are then applied, the system having the option for both precession and equi-inclination Weissenberg methods. Also available is a spot shape correction.

Finally, weights are applied to each reflection based on the following weighting scheme:-

Weight factors:- A1, A2, B1, B2, FB

for Intensity $\leqslant$ FB then $\sqrt{wt}$ = A1 x (Intensity)$\uparrow$ B1

Intensity > FB then $\sqrt{wt}$ = A2 x (Intensity)$\uparrow$ B2

B1, B2 $\leqslant$ 0

2.2.2   **Lorentz and polarisation corrections**   -   During data reduction intensities must be corrected for factors of the geometry of crystal movement (Lorentz factor) and the polarisation of the diffracted beam.

The Lorentz factor for equi-inclination Weissenberg measurements is given by:-

$$L = \frac{\sin \Theta}{\sin 2\Theta} \times \frac{1}{\sqrt{\sin^2 \Theta - \sin^2 \mu}}$$

where μ is the equi-inclination setting angle.   For zero level reflections, and for ω-2Θ scans on an automatic diffractometer, this reduces to:-

$$L = \frac{1}{\sin 2\Theta} \qquad \text{(Stout and Jensen, 1968 - p196)}$$

The polarisation factor 'p' is independent of the method of data collection (Stout and Jensen, 1968 - p196) <u>viz</u>

$$p = \frac{1 + \cos^2 2\Theta}{2}$$

The relationship between the absolute value of the structure amplitude ($|Fhkl|$) and the background corrected intensity ($Ihkl$) can be expressed as follows:

$$Fhkl = \sqrt{\frac{K \; Ihkl}{Lp}}$$

where K is a constant of proportionality and

$$Lp^{-1} = \frac{2\sin 2\Theta}{1 + \cos^2 2\Theta}$$

for zero level Weissenberg and ω-2Θ diffractometer data.

2.3 Data reduction of diffractometer data

This was performed on the DEC20 computer at Birmingham University using a programme provided by Dr T A Hamor. The programme provides the following facilities:

a) Background subtraction (Stout and Jensen, 1968 - p456)

b) Lorentz and polarisation corrections (see above)

c) A simple linear intensity correction

d) Attenuator factor correction

e) Output provided in the form of h, k, l, Fobs and σ Fobs , where σ Fobs is calculated by the programme on the basis of counting statistics, and includes an allowance for the variation in intensity of the main beam with time.

2.3.1 Background subtraction

The method of background subtraction is described in Stout and Jensen (1968 - p456) and this may be extended to generalised counting times for peak and background as follows:-

$$Npk = N_T - r\ (N_{b1} + N_{b2})$$

where N_T = total count, Npk = peak count and N_{b1} and N_{b2} are the background counts each collected over a period $1/r$ x peak count time.

2.3.2 Standard deviation

The standard deviation due to counting statistics is calculated by:-

$$\sigma^2_{pk} = \sigma^2_T + \sigma^2_{b1} + \sigma^2_{b2}$$

where $\sigma^2_T = N_T$ and $\sigma_{b1} = r\sqrt{N_{b1}}$

The final standard deviation which includes an expression for variation in the main beam intensity can be expressed as follows:-

$$\sigma^2_I = \sigma^2_{pk} + (f\ Npk)^2$$

where f is a parameter which depends upon the constancy of the x-ray beam.

2.3.3 Intensity Correction

During data collection a small number of reflections are measured repeatedly at regular intervals throughout data collection. Should there be a monotonic decrease in intensity of these reflections indicating possible deterioration of the crystal in the x-ray beam, then their change in values at the periods measured can be used to apply a simple linear intensity correction to the measured data.

Where the data collection has proceeded without any overall trend in the intensities of the check reflections, the spread of measurements may be used to estimate the parameter f used in the beam intensity correction.

2.3.4 Normalised structure factors

Direct Methods programmes require that the structure factors be placed on an absolute scale. The magnitude of a structure factor depends not only on the degree to which atoms scatter co-operatively but also on the scattering angles, since atoms scatter less strongly at high angles. From a theoretical point of view there is an advantage in producing a structure factor which is corrected for fall off in scattering angle such that the numerical value of the structure factor is independent of its position in reciprocal space.

The principal method of achieving this is the Wilson Plot Method. Where the Lp-corrected intensities have been averaged over a particular range in θ (mean value $\langle I_\theta \rangle$) then:-

$$\langle I_\theta \rangle = K \sum_\theta \exp(-2B \sin^2\theta / \lambda^2)$$

where K is a scale factor and $\sum_\theta$ is the structure factor for a randomly distributed set of atoms evaluated at the appropriate $\sin\theta/\lambda$; B is an overall isotropic temperature factor.

Since
$$\log_e \left[\frac{\langle I_\theta \rangle}{\sum_\theta} \right] = \log_e K - 2B \sin^2\theta / \lambda^2$$

a plot of the $\log_e$ of the ratio on the left vs $\sin^2\theta / \lambda^2$ will have an intercept $\log_e K$ and a gradient $-2B$. It is essential that the averaging involved in the calculation of $\langle I_\theta \rangle$ is carried out over all points in the reciprocal lattice including those corresponding to systematic absences. The above procedures enable the scaling of the observed structure factors:-

ie. $$|F_{hkl}| = k I_{hkl}^{1/2}$$

where $k = 1/\sqrt{K}$

and the placement of the structure factors on an absolute scale is

$$E_{hkl} = \frac{F'_{hkl}}{\varepsilon \sum_\theta}$$

where ε is a factor correcting for systematic absences and the structure factors have been corrected for thermal vibration:-

$$|F_{hkl}| = |F'_{hkl}| \exp(-B\sin^2\theta / \lambda^2)$$

These features are incorporated in the programme NORMAL. An additional feature of this programme is that calculation of $\sum_\theta$ can also be done on the basis of randomly oriented molecular fragments, which can help eliminate some of the irregularity resulting from the non-random nature of the structure.

SHELX has its own normalised structure factor calculation called using the MERG command. The method used is that of Karle et al (1958).

2.4 Direct Methods Programme

2.4.1 MULTAN

MULTAN is a computer programme for the largely automatic solution of crystal structures with up to about 150 atoms in the asymmetric unit; it is applicable to both centrosymmetric and non-centrosymmetric structures. As it now exists, it involves the following stages.

(i)   Input raw data and process this to give normalised structure factors (E's) (NORMAL).

(ii)   Statistically analyse the E's to give space group information or to reveal abnormal features of the structure.

(iii)   Find the set of largest structure factors from which the structure may be determined. Typically this is up to about 350 structure factors for a structure containing 50 non-hydrogen atoms in the asymmetric unit.

(iv)   Record triplets of the type $\underline{h}\ \underline{k}\ \underline{l}$, $\underline{h}'\ \underline{k}'\ \underline{l}'$ and $\underline{h}-\underline{h}'\ \underline{k}-\underline{k}'\ \underline{l}-\underline{l}'$ the expected sign of the reflection E(hkl) being given for a centrosymmetric structure by
$$S\left\{E(hkl)\right\} \approx S\left\{E(h'k'l')\right\} \cdot S\left\{E(h-h'k-k'l-l')\right\}$$
(where $S\left\{\ \right\}$ indicates the sign of the quantity in curly brackets) or in the case of general phases for non centrosymmetric structures:-
$$\phi_h \approx \phi_{h'} + \phi_{h-h'}$$
where the vector index h is used to represent hkl.
(SIGMA 2)

(v)   Find a good starting set for phase development from which other phases can be developed (CONVERGE)

(vi)   For each permuted starting point generate new phases and refine the complete set of phases with the tangent formula (FASTAN)

(vii)   Calculate figures of merit (FOMs) for each phase set developed. All figures of merit are combined into a single number - the combined figure of merit (CFOM)

(viii) For the set of phases with the highest CFOM (or any

other designated by the user) calculate an electron
density map with the Fast Fourier transform programme).

2.4.2 EEE's

An alternative centrosymmetric direct methods approach is that
of 'EEEs' used in the SHELX system. Here the approach rather than
recycling is to start with a very large number of permutations of
signs (of the order of 2^{10} to 2^{20}) and to eliminate early on during
sign expansion those sets which are giving poor agreement. If at any
stage, the agreement fails to reach a prescribed level (the level
itself increasing during expansion) then the expansion is
discontinued and the set rejected.

E-maps are computed and figures of merit calculated for the
surviving sets.

'The choice of which direct methods package to use was largely
predetermined by whether the structure was centrosymmetric or not.
Generally for centrosymmetric structures the relatively easy to use
EEEs programme in SHELX was used, whilst for non-centrosymmetric
structures MULTAN became the standard choice. Both have their
advantages. Namely in SHELX the initial large number of sign
permutations computed, whilst in MULTAN, the normalising procedure
and particularly the ability to use randomly oriented fragments in
calculating E's appears to be a distinct advantage.'

2.5 Least Squares Refinement

In all but one of the structuressolved in this work, full matrix
least squares refinement has proceeded using LS in SHELX.

The parameters usually refined are:-

    (i)   The $|Fo|$ scale factor

    (ii)   The atomic co-ordinates

    (iii)   Either the individual isotropic temperature factor or the six components of the individual atomic anisotropic vibration tensors U_{ij}.

The version of SHELX currently available on the CDC7600 at U.M.R.C.C. allows for the refinement of up to 306 parameters simultaneously with the option of blocking (BLOC instruction) where more parameters are to be refined. Small variations to these parameters are made on each cycle to produce a test agreement between observed and calculated structure factors. The least squares method proceeds by making the sum of the square of the errors in Fo a minimum namely minimising the function.

$$\sum w(\ |Fo| \ - \ |Fc| \ )^2$$

where w is a weight allocated to each structure factor. In SHELX the only permissible weighting scheme other than unit weights is

$$w = 1/[\sigma^2(Fo) + q \ |Fo| \ ^2]$$

where q is a refinable parameter or may be fixed at 0.

The procedure is a cyclic one and after each cycle, R, Rw and Rg, the various discrepancy indices, are printed, together with the shift/esd for each parameter. Refinement is terminated when there is no improvement in the R-factor, and the shift/esds are at an acceptable level for the various parameters, ie. at convergence..

The definitions of the various R-factors output in SHELX are as follows :-

$$R = \frac{\sum \Delta}{\sum Fobs}$$

$$Rw = \frac{\sum w^{1/2} \Delta}{\sum w^{1/2} Fobs}$$

$$Rg = \frac{\sum_w \Delta^2}{\sum_w |Fobs|^2}$$

where $\Delta = ||Fobs| - |Fcalc||$

During the refinement procedure, some groups of atoms may be refined as a rigid group, ie. a CH_3 or phenyl ring, (AFIX) or atomic positions refined to give a constrained bond distance with the operator specifying the standard deviation (DFIX). Where this has occurred in this work details are given under bond length data. (For further details see documentation on SHELX).

2.6 Least Squares Planes Calculations

Least squares planes calculations and torsion angles have all been computed using GEOM78 (Cambridge Crystallographic) on the CDC7600 at UMRCC. (For details see Cambridge Crystallographic documentation).

2.7 Molecular Plotting

PLUTO78 on the CDC7600 at UMRCC was used for all molecular plotting in this work. The programme allows plotting of either unit cell contents or individual molecules with options for projection, perspective or stereo representation. Either stick, ball and stick or space filling diagrams can be produced; a number of orientation options are allowed for, the default option being an approximately calculated minimum overlap orientation.

CHAPTER 3

Chapter 3

3.1.1   The Crystal Structures of Two Stable Salts of

Benzylpenicillin

Penicillin, the first of the antibiotics to come into general therapeutic use, is still in many ways the best.   It was discovered by Fleming, and its isolation and systematic study, by Florey, Chain and their co-workers spanned a period of 10 years.   It took many years of hard work to obtain penicillin in the pure state and the unit of activity by which it had first to be measured, now known to represent $0.6 \mu g$, has persisted to the present day although all later penicillins are prescribed by weight.

Benzylpenicillin (Figure 3.1) can be prepared in quantity only by the original process of cultivating a mould forming it (a high-yielding mutant of a strain of P. chrysogenum is now used) in a suitable liquid medium.   In the early stages of development it was found that four different penicillins were being formed known as F, G, X and K.   Of these, G or benzylpenicillin had the most desirable properties and its almost exclusive formation is ensured by adding the appropriate 'precursor' phenylacetic acid to the medium.

As formed in this process, penicillin is an unstable acid and in production is converted to either the potassium or sodium salts which are considerably more stable, exhibit a high degree of solubility in water and are characterised by their rapid absorption and excretion in the body.   Among the many incompatible chemicals affecting the stability of penicillin, the most important is acid, and indeed gastric acid accounts for most of the loss of a dose of benzylpenicillin if it is swallowed.   It is also destroyed by an enzyme penicillinase formed by various bacteria, including some staphylococci, various bacilli and some species of Proteus which in

many cases accounts for the abnormal resistance of many of these bacteria, particularly the staphylococci, to benzylpenicillin. Early work in producing different penicillins with enhanced properties was limited to adding different precursors to the cultivating medium, among which only derivatives of acetic acid were effective.

However, in 1959, Batchelor et al[6] noted enhanced quantities of the penicillin nucleus - 6 aminopenicillinic acid - in Penicillium chrysogenum fermentations in the absence of a precursor. A second discovery made at this time was that the penicillin nucleus could be prepared in quantity from benzyl or phenoxymethyl penicillin by the action of an enzyme derived from other micro-organisms: the effect of this amidase enzyme is to separate the side chain, Figure 3.1. p 29.

Thus, the availability of the nucleus itself, in reasonable quantities enabled side chains in a great variety to be attached by a semi-synthetic process. In Beecham Research Laboratories alone, where these discoveries were made, over 3000 new penicillins have been prepared by this process and those which have come into therapeutic use have one or more of the following advantages:

(1) resistance to acid, eg. Phenoxypenicillins

(2) resistance to penicillinase (a β-lactamase) eg. methacillin, oxacillin

(3) broad spectrum activity eg. Ampicillin

3.1.2 Mode of Action of the Penicillins

The outstanding difference between bacteria and mammalian cells is the tough thick cell wall external to the cell membrane which fixes the shape of bacterial cells and gives them their extraordinary resistance to osmotic damage. This structure is absent from mammalian cells and therefore any agent which interfered solely with its

construction would be entirely without effect on the mammalian host. Two such classes of compounds, the penicillins and cephalosporins, both β-lactam antibiotics, exert their effect in this way by interfering with the synthesis of the mucopeptide (peptidoglycan) component of the bacterial cell wall. The synthesis of the cell wall takes place in three distinct stages at three different sites in the cell[7]. It is the third stage which occurs outside the cell membrane which is specifically inhibited by penicillins and cephalosporins. At this stage, the linear peptidoglycan strands are cross-linked by a transpeptidation step in which a peptide bridge is formed between two adjacent strands with the elimination of D-alanine. It is postulated that the β-lactam might be a structural analogue of the D-alanyl-D-alanine end of the peptide chain. If the transpeptidase formed an acyl intermediate with the end of the pentapeptide eliminating D-alanine, it could also react with penicillin forming a penicilloyl intermediate and thus become inactivated[8]. The structural similarity between the penicillin nucleus and one of the conformations of the acyl-D-alanyl-D-alanine end of the nascent peptidoglycan as proposed by Tipper and Strominger[8] is shown in Figure 3.1-1. Obviously, this process of inhibiting cell wall formation can only take place whilst the cells are multiplying (ie. during cell division) and as such the β-lactam antibiotics are ineffective against dormant cells not producing new cell walls.

3.1.3 Pharmacokinetics

The salts of benzylpenicillin are very freely diffusible; after intramuscular injection, absorption occurs within a few minutes to produce a high concentation in the blood. Diffusion takes place into

the fetal circulation and into serous cavities, lower concentrations being found in glandular secretions and cerebro-spinal fluid. Concentrations two to five times that in the blood are found in bile, but excretion is mainly renal, accounting for about 60% of the dose. This excretion is mainly tubular and exceedingly rapid, thus making treatment with penicillin wasteful. Maintaining a sufficiently high blood level of the drug is usually carried out by giving very large doses - fortunately penicillin is virtually non-toxic. For example, the intramuscular dose required of a salt of benzylpenicillin, eg. potassium, to maintain a blood level of over 0.1μg/ml for 2 hours has been shown to be 50,000 units and thus if a continuous effect is expected from penicillin in this form, large and frequent doses are required. One way of prolonging the action of each dose of penicillin is to administer probenecid which interferes with tubular excretion.

3.1.4 Long-acting forms of benzylpenicillin

A second possible way of prolonging the effect of a dose is to delay the absorption. This was first achieved by suspending a calcium salt in a water-immiscible medium containing oil and beeswax. However, much more satisfactory results have since been obtained with penicillin compounds of a lesser solubility. A series of compounds which have this particular property are a group of penicillins in which the benzylpenicillin is complexed with another compound to form an ion-pair. The structures of two of these penicillins, namely, Procaine penicillin G monhydrate and Benethamine pencillin are reported in this work. A noticeable feature of these two compounds is their stability to hydrolysis and their relative insolubility in aqueous solvents, properties which are used in their

application as antibiotics. Both compounds are administered as an aqueous crystalline suspension and injected intramuscularly in order to provide a depot of the drug which may then slowly release Penicillin G, the active constituent of the complex, into the bloodstream. A dose of 300-900 mg of procaine benzylpenicillin may be expected to maintain an effective concentration in the blood for about 24 hours[9] so that injections need not be given more frequently than once per day. A similar dose of benethamine penicillin would provide an effective concentration in the blood for a period of four to five days. Clearly, even with these two compounds there is a marked difference in the stability of the two complexes, and initially the sole objectives of the study were to:-

a) identify the nature of the interactions between procaine, penicillin and water and benethamine and penicillin, ie. salt-like, hydrogen bonding or ring stacking

b) identify those aspects of the crystal structures and intermolecular interactions which might give rise to the stability of these compounds to aqueous environments thus rendering them useful as repository drugs.

However, shortly after we had solved the structure of Procaine Penicillin G, Dexter and van der Veen (1978)[10] published the structure, which although it had some minor errors involved, looked substantially the same as our results. Dexter and van der Veen had used a much smaller crystal, obtained by crystallising from ethanol, whilst in this work much larger crystals were produced from the significantly more polar methanol:water solvent. A further investigation into the pharmacokinetics of procaine penicillin showed that Buckwalter and Dickison (1958)[11] observed different blood levels from samples of a similar particle size but crystallised from

different   media.   Ballard   and   Nelson   (1975)[12]   suggested
polymorphising or solvate formation as an explanation, however, Macek
(1975)[13] reported that prisms and plates had the same crystal form,
by x-ray diffraction.

This therefore provided the additional objective of determining
whether crystals grown from methanol:water and those from ethanol are
the same material, and to suggest whether observed differences in
bioavailability are due to different internal structures or different
external habits.

Figure 3.1
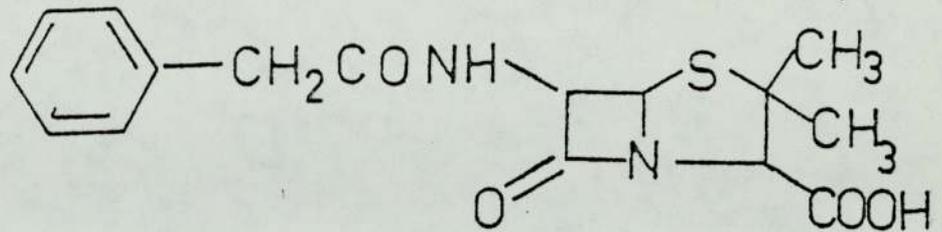
Benzylpenicillin (penicillin G)
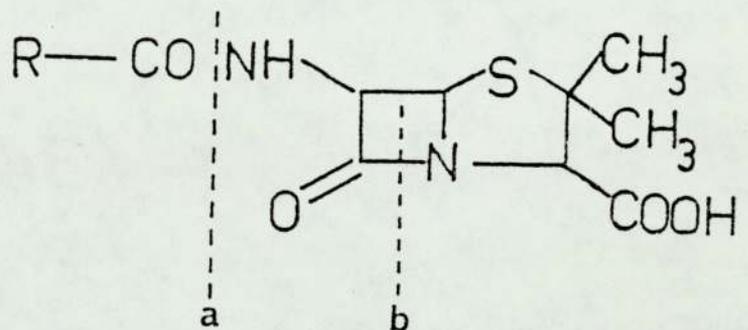
a) site of action by amidase to produce 6-amino penicillinic acid - 'penicillin nucleus'

b) site of action by β-lactamase (penicillinase) to inactivate penicillin

Figure 3.1-1
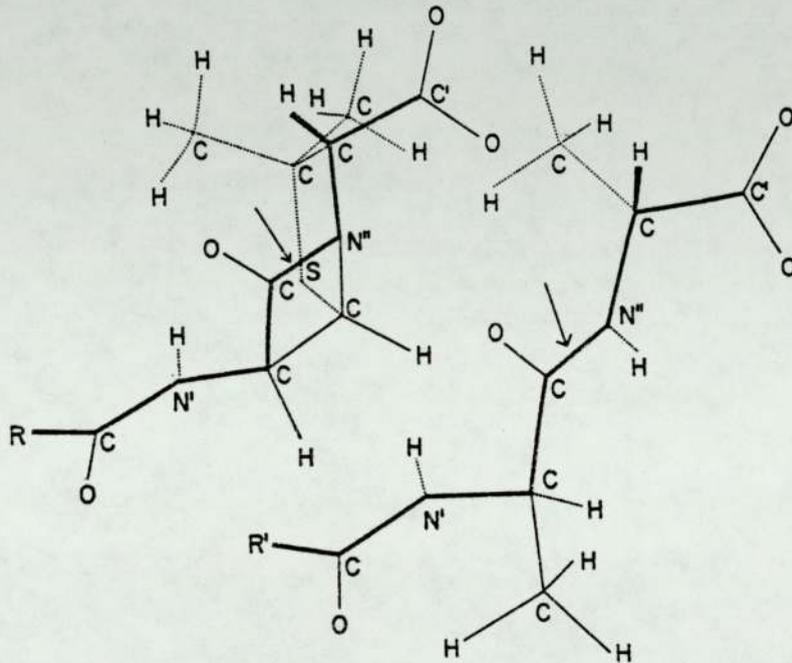
3.1-1   Drieding stereomodels of penicillin (*upper left*) and of the acyl-
D-alanyl-D-alanine end of the nascent peptidoglycan (*lower right*). Arrows
indicate the position of the CO—N bond in the β-lactam ring of penicillin and of
the CO—N peptide bond joining the two D-alanine residues. The portion of the
penicillin molecule which is believed to resemble the peptide backbone of acyl-D-
ala-D-ala is reproduced in heavy lines

Table 3.1.1   Solvents used in crystallisation procedures

Solvent	Polarity P'	Crystal Habit	Reference
H_2O	(9)	Small prisms	Buckwalter & Dickison, 1958
H_2O + colloid	(9)	Small plates	Sumner & Grenfell, 1955
MeOH/H_2O 3:1	7.2	Prisms 1 mm	This work
Acetone/H_2O	7.2	Prisms 1 mm	Buckwalter & Dickison, 1958
EtOH	5.2	Prisms 0.3mm	Dexter & van der Veen, 1978
PrOH	3.9	Small laths	Buckwalter & Dickison, 1958
BuOAc/H_2O 11:1	(3.7)		Bardolph, 1956

Buckwalter & Dickison (1958) observed different blood levels from samples of similar particle size but crystallized from different media. Ballard & Nelson (1975) suggest polymorphism or solvate formation as an explanation. However, Macek (1975) reports that prisms and plates have the same crystal form by X-ray diffraction.

3.2 THE CRYSTAL STRUCTURE OF PROCAINE PENICILLIN G MONOHYDRATE[14]

3.2.1 Experimental

A sample of 'Procaine Penicillin G Monohydrate' (APPG) in the form of a fine white microcrystalline powder was obtained from Glaxo Laboratories Ltd, Greenford, Middlesex.

A small quantity (approximately 100 mg) was dissolved in warm aqueous methanol (1:3) and allowed to cool. Slow evaporation (overnight) produced large, well formed, prismatic crystals ca 1 mm.

A single crystal (checked by polarizing microscope) of dimensions 0.6 x 0.7 x 1.0 mm was mounted along its longest dimension on a Stoe Weissenberg camera. An initial oscillation photograph showed the crystal to belong to the monoclinic class or higher. Subsequent zero and first level Weissenberg films confirmed the monoclinic class and the space group $P2_1$, with the crystal being mounted along its unique axis. Unit cell dimensions were calculated from the oscillation and Weissenberg films and a density determination (by flotation) confirmed the existence of two molecules per unit cell.

Data collection by film then ensued: seven levels of data were collected, ie. h0l to h6l inclusive using the non-integrating Weissenberg technique. Data for each level were collected in two film packs each of three films; one film pack had a long exposure time to enable the measurement of very weak reflections, the other a short exposure time to enable the measurement of strong reflections.

A final single film exposure of the h0l level was made as a check for any crystal deterioration during data collection. All film data were collected using Cu-Kα radiation ($\lambda = 1.5418$ Å).

In order to complete the data set and obtain higher order

reflections in 'k' a further data set was collected from the same crystal on Stoe 2-circle diffractometer at Birmingham University, courtesy of Dr T A Hamor. Five levels of data were collected (h7l to h11l inclusive) using graphite monochromated Mo-Kα radiation (λ = 0.71069 Å). After each level of data was collected, four check reflections were measured in order to assess any crystal deterioration or movement.

Whilst data reduction on the diffractometer data proceeded at Aston on the ICL 1904S, the film data were measured by the SRC Microdensitometer Service, Atlas Computer Division, Didcot, courtesy of Dr M Elder.

3.2.2   Crystal Data

Procaine Penicillin G Monohydrate ($C_{29}H_{38}N_4O_6S,H_2O$) crystallises in the monoclinic space group $P2_1$ with a=10.70(1), b=10.42(1), c=14.41(1)Å, β =104.74(8)°, Z=2 and V=1554(1)Å^3.

The molecular weight is 588.7 (F(ooo) = 616) and the density Dm (by flotation in carbon tetrachloride and n-propanol) = 1.26 g/ml with the calculated density Dc = 1.28 g/ml.

The absorption coefficient for Mo-Kα radiation (λ = 0.71069 Å) is μ = 1.55cm^{-1} and for Cu-Kα radiation (λ = 1.5418Å) is μ= 10.61 cm^{-1}.

3.2.3   Structure Determination and Refinement

The data reduction on both sets of data - film and diffractometer - involved the application of Lorentz and polarisation corrections and a linear intensity correction based on the measurements of the check reflections. No correction for absorption effects was made to either set. The scaling together of the

diffractometer and film data was performed at Aston using the programme FSC (see Chapter 2 on crystallographic software). To enable this, a number of reflections from each of the levels measured on film were also measured on the diffractometer; it was the intensities of these reflections which were used to provide inter level scale factors and place both Weissenberg and diffractometer data on a common scale. After data reduction and scaling, (this included the merging of Friedel pairs) the data set comprised 2962 unique observed reflections between $2° \leqslant \theta \leqslant 32°$ of which approximately 1500 were diffractometer data and the balance from film methods. The data were output in the form H, K, L, Fobs and σFobs where σFobs was based purely on counting statistics. Initial attempts at solving the structure using MULTAN[1] were unsuccessful in that the most promising fragment that appeared on an electron density map was the sulphur atom $(S(1))$ and its two adjoining carbon atoms $(C(2)$ and $C(5))$. The correct position of the sulphur atom in x and z was confirmed using a Patterson synthesis the result of which corresponded closely to the fragment determined using MULTAN. Attempts at phasing the remaining atoms by Fourier refinement failed presumably since we had insufficient of the total scattering power of the unit cell.

Headway was finally made by using atomic superposition of Patterson maps with the programme HASUP[15]. This yielded the atoms in the thiazolidine and β-lactam rings. Successive Fourier refinements finally located all of the non-hydrogen atoms.

Using the atomic positions determined from the final Fourier map, full matrix least squares refinement of positional parameters and isotropic temperature factors for all non-hydrogen atoms ensued using the XRAY 72 system[3] during which the origin along y in $P2_1$ was defined by fixing the y coordinate of the sulphur atom. During this

stage of the refinement, atoms were assigned their correct atomic scattering factors on the basis of the known structure of APPG and $R = \sum ||Fo| - |Fc|| \; / \; \sum |Fo|$ the unweighted discrepancy index decreased to 0.13. At this point, 26 of the 40 hydrogen atoms were located from a difference electron density synthesis and were assigned the isotropic temperature factor of their attached atom.

Successive full matrix least squares refinements of positional parameters and anisotropic thermal parameters for non-hydrogen atoms and positional parameters for hydrogen atoms resulted in all but one of the remaining hydrogen atoms being located from a difference electron density synthesis. Final anisotropic refinement during which all hydrogen atom positions and isotropic temperature factors were fixed, converged to R = 0.051 and Rw = 0.078.

The function minimised was $w \sum ||Fo| - (1/k)|Fc||^2$ and the choice of weighting scheme used during the final stages of refinement was w = 1.0/max (σ, 0.06Fo + 0.40) (University of Washington scheme)[16].

A final difference electron density synthesis showed no feature greater than 0.18 e$\overset{o}{A}^{-3}$.

Table 3.1   Positional parameters (fractional co-ordinates x 10^4)
with estimated standard deviations in parentheses.
(Non-hydrogen atoms).

ATOM	X/a	Y/b	Z/c
S(1)	6279(1)	3007(0)	10526(1)
C(2)	5821(4)	1745(3)	11274(3)
C(3)	5945(4)	446(3)	10760(3)
N(4)	6751(4)	684(3)	10102(2)
C(5)	7408(4)	1937(4)	10169(3)
C(6)	7276(4)	1838(4)	9068(3)
C(7)	6326(5)	739(4)	9113(3)
O(8)	5489(4)	189(3)	8550(3)
C(9)	4449(6)	1986(5)	11330(6)
C(10)	6735(7)	1820(6)	12292(4)
C(11)	3530(5)	4363(4)	8525(3)
O(12)	2388(4)	4009(5)	8442(3)
O(13)	4258(4)	3909(4)	8075(3)
N(14)	6785(4)	2875(4)	8428(3)
C(15)	7479(5)	3371(4)	7863(3)
O(16)	8624(4)	3093(5)	7942(3)
C(17)	6774(6)	4304(4)	7109(3)
C(18)	6227(5)	3646(4)	6149(3)
C(19)	4992(6)	3939(7)	5630(5)
C(20)	4493(7)	3397(11)	4743(6)
C(21)	5236(8)	2552(8)	4376(5)
C(22)	6450(7)	2257(7)	4884(6)
C(23)	6948(6)	2793(5)	5761(4)
N(24)	861(5)	1880(6)	951(3)
C(25)	835(5)	2375(5)	1824(4)

ATOM	X/a	Y/b	Z/c
C(26)	441(5)	1631(5)	2513(3)
C(27)	464(5)	2123(4)	3401(3)
C(28)	861(5)	3384(4)	3643(3)
C(29)	1250(6)	4121(5)	2958(4)
C(30)	1244(6)	3626(6)	2081(4)
C(31)	914(5)	3920(4)	4588(3)
O(32)	1073(5)	5038(4)	4809(3)
O(33)	795(4)	3020(3)	5224(2)
C(34)	788(6)	3429(4)	6182(3)
C(35)	816(5)	2255(4)	6779(3)
N(36)	2154(4)	1797(3)	7265(3)
C(37)	3019(6)	1564(6)	6615(4)
C(38)	2559(8)	473(7)	5909(5)
C(39)	2103(5)	668(5)	7897(4)
C(40)	1586(8)	981(6)	8756(4)
O(41)	529(4)	4194(4)	9481(3)

ATOM	X/a	Y/b	Z/c
C(26)	441(5)	1631(5)	2513(3)

Table 3.2   Anisotropic thermal parameters for non-hydrogen atoms

ATOM	U11	U22	U33	U23	U13	U12
S(1)	.0575	.0350	.0442	-.0054	.0168	.0007
C(2)	.0380	.0385	.0383	.0007	.0128	.0017
C(3)	.0458	.0321	.0321	.0006	.0099	.0027
N(4)	.0460	.0399	.0335	.0076	.0107	.0080
C(5)	.0333	.0519	.0487	-.0009	.0090	.0097
C(6)	.0392	.0567	.0435	.0036	.0142	.0060
C(7)	.0473	.0396	.0361	.0050	.0103	.0020
O(8)	.0823	.0512	.0449	-.0089	.0046	-.0032
C(9)	.0609	.0566	.1120	.0121	.0484	.0091
C(10)	.0897	.0634	.0429	.0004	.0007	-.0007
C(11)	.0504	.0357	.0366	-.0048	.0094	-.0061
O(12)	.0727	.0792	.0670	-.0407	.0286	-.0335
O(13)	.0677	.0580	.0505	.0159	.0076	-.0150
N(14)	.0445	.0493	.0429	.0093	.0133	.0122
C(15)	.0361	.0518	.0427	-.0071	.0138	.0014
O(16)	.0485	.0953	.0550	-.0056	.0180	.0047
C(17)	.0802	.0408	.0467	-.0099	.0249	.0065
C(18)	.0531	.0426	.0379	-.0060	.0138	.0066
C(19)	.0718	.0813	.0638	.0143	.0110	.0026
C(20)	.0570	.1477	.0653	-.0012	-.0065	-.0009
C(21)	.0852	.1060	.0580	-.0203	.0085	-.0143
C(22)	.0782	.0869	.0628	.0060	.0143	-.0088
C(23)	.0722	.0570	.0531	.0163	.0021	.0021
N(24)	.0735	.0872	.0443	-.0093	.0171	-.0095
C(25)	.0382	.0629	.0423	-.0026	.0052	.0048

ATOM	U11	U22	U33	U23	U13	U12
C(26)	.0536	.0510	.0424	-.0032	.0151	-.0003
C(27)	.0508	.0470	.0402	-.0025	.0071	.0047
C(28)	.0454	.0431	.0488	.0010	.0041	.0048
C(29)	.0650	.0466	.0538	-.0091	.0105	.0029
C(30)	.0503	.0618	.0507	-.0047	.0148	.0091
C(31)	.0457	.0438	.0461	.0051	-.0002	.0015
O(32)	.1072	.0432	.0587	-.0036	.0085	-.0031
O(33)	.0917	.0432	.0414	.0001	.0119	-.0018
C(34)	.0798	.0428	.0388	.0053	.0103	-.0097
C(35)	.0469	.0435	.0436	.0046	.0068	-.0064
N(36)	.0574	.0405	.0397	.0002	.0072	-.0091
C(37)	.0525	.0648	.0667	.0016	.0048	.0010
C(38)	.1014	.0756	.0771	.0144	.0428	-.0167
C(39)	.0601	.0493	.0373	.0113	-.0004	-.0018
C(40)	.1087	.0626	.0609	.0193	.0289	.0111
O(41)	.0712	.0697	.0599	-.0007	.0235	-.0053

Table 3.3 Positional parameters for hydrogen atoms (fractional co-
ordinates x 10^4) with isotropic temperature factors (U_{iso})

ATOM	X/a	Y/b	Z/c	U_{iso}
H(3)	5039	194	284	.0370
H(5)	8367	1902	10700	.0460
H(6)	8066	1486	8837	.0460
H(9A)	4420	2860	11740	.0720
H(9B)	4108	1245	11658	.0720
H(9C)	3838	2081	10669	.0720
H(10A)	6855	967	12790	.0650
H(10B)	7636	2009	12214	.0650
H(10C)	6462	2556	12639	.0650
H(14)	5992	3184	8418	.0368
H(17A)	7384	4988	7026	.0430
H(17B)	6042	4696	7331	.0430
H(19)	4453	4542	5902	.0646
H(20)	3594	3618	4358	.0706
H(21)	4862	2156	3729	.0691
H(22)	6973	1637	4606	.0623
H(23)	7834	2546	6122	.0603
H(24A)	972	2428	465	.0660
H(24B)	9701	6102	9302	.0660
H(26)	115	729	2351	.0420
H(27)	220	1557	3899	.0394
H(29)	1541	5020	3121	.0474
H(30)	1535	4190	1610	.0536
H(34A)	-38	3931	6162	.0431
H(34B)	1539	3976	6460	.0431
H(35A)	336	1547	6364	.0388

ATOM	X/a	Y/b	Z/c	U_{iso}
H(35B)	345	2443	7295	.0388
H(36)	2512	2507	7674	.0450
H(37A)	3913	1341	7033	.0514
H(37B)	3085	2357	6252	.0514
H(39A)	3020	331	8145	.0426
H(39B)	1571	-21	7501	.0426
H(38A)	1707	681	5506	.0657
H(38B)	2534	-335	6287	.0657
H(38C)	3172	303	5488	.0657
H(40A)	659	1301	8512	.0664
H(40B)	2107	1653	9156	.0664
H(40C)	1536	215	918	.0664
H(41)	1249	425	907	.0670

ATOM	X/a	Y/b	Z/c	U_{iso}
H(35B)	345	2443	7295	.0388

Table 3.4 Bond lengths [Å] for APPG with estimated standard deviations in parentheses

Bond distances

Benzylpenicillin:

S(1)-C(2)	1.843(4)	C(2)-C(9)	1.512(9)	C(17)-C(18)	1.522(6)
S(1)-C(5)	1.811(5)	C(2)-C(10)	1.543(6)	C(18)-C(23)	1.384(8)
C(2)-C(3)	1.565(6)	C(3)-C(11)	1.535(5)	C(18)-C(19)	1.377(5)
C(3)-N(4)	1.455(6)	C(11)-O(12)	1.252(7)	C(19)-C(20)	1.376(9)
N(4)-C(5)	1.474(6)	C(11)-O(13)	1.228(7)	C(20)-C(21)	1.379(13)
N(4)-C(7)	1.384(6)	C(6)-N(14)	1.431(6)	C(21)-C(22)	1.354(10)
C(5)-C(6)	1.561(7)	N(14)-C(15)	1.336(7)	C(22)-C(23)	1.360(9)
C(6)-C(7)	1.544(7)	C(15)-O(16)	1.238(6)		
C(7)-O(8)	1.191(6)	C(15)-C(17)	1.509(6)		

Procaine:

N(24)-C(25)	1.367(8)	C(28)-C(31)	1.460(7)	C(35)-N(36)	1.502(6)
C(25)-C(26)	1.406(8)	C(29)-C(30)	1.364(8)	N(36)-C(37)	1.494(8)
C(25)-C(30)	1.394(8)	C(31)-O(32)	1.210(6)	N(36)-C(39)	1.498(6)
C(26)-C(27)	1.374(7)	C(31)-O(33)	1.339(6)	C(37)-C(38)	1.522(9)
C(27)-C(28)	1.398(6)	O(33)-C(34)	1.447(6)	C(39)-C(40)	1.514(9)
C(28)-C(29)	1.394(8)	C(34)-C(35)	1.492(7)		

Bonds involving hydrogen atoms:

C(3)-H(3)	1.06	C(17)-H(17A)	1.00	C(27)-H(27)	1.01	C(39)-H(39A)	1.02
C(5)-H(5)	1.11	C(17)-H(17B)	1.01	C(29)-H(29)	1.00	C(39)-H(39B)	1.00
C(6)-H(6)	1.05	C(19)-H(19)	1.00	C(30)-H(30)	1.01	C(38)-H(38A)	.97
C(9)-H(9A)	1.09	C(20)-H(20)	1.01	C(34)-H(34A)	1.02	C(38)-H(38B)	1.01
C(9)-H(9B)	1.02	C(21)-H(21)	1.00	C(34)-H(34B)	.98	C(38)-H(38C)	1.02
C(9)-H(9C)	1.01	C(22)-H(22)	1.00	C(35)-H(35A)	1.00	C(40)-H(40A)	1.02

Bonds involving hydrogen atoms:

C(10)-H(10A)	1.12	C(23)-H(23)	1.00	C(35)-H(35B)	1.02	C(40)-H(40B)	.98
C(10)-H(10B)	1.02	N(24)-H(24A)	.93	N(36)-H(36)	.96	C(40)-H(40C)	1.02
C(10)-H(10C)	1.00	N(24)-H(24B)	1.02	C(37)-H(37A)	1.02	O(41)-H(41)	1.08
N(14)-H(14)	.90	C(26)-H(26)	1.01	C(37)-H(37B)	.99		

Table 3.5 Interatomic angles (°) with estimated standard
deviations in parentheses (for non-hydrogen atoms only)

Atoms	Interatomic Angle (°)	Atoms	Interatomic Angle (°)
C(2)-S(1)-C(5)	91.3(2)	C(17)-C(18)-C(19)	119.1(5)
S(1)-C(2)-C(3)	105.9(3)	C(17)-C(18)-C(23)	122.2(4)
S(1)-C(2)-C(9)	108.6(3)	C(18)-C(19)-C(20)	120.3(7)
S(1)-C(2)-C(10)	108.8(3)	C(19)-C(20)-C(21)	119.6(6)
C(9)-C(2)-C(10)	109.3(5)	C(20)-C(21)-C(22)	120.3(7)
C(2)-C(3)-N(4)	107.2(3)	C(21)-C(22)-C(23)	120.2(7)
C(3)-C(2)-C(10)	112.6(4)	C(22)-C(23)-C(18)	120.8(6)
C(3)-C(2)-C(9)	111.6(4)	C(23)-C(18)-C(19)	118.7(5)
C(2)-C(3)-C(11)	112.3(3)	N(24)-C(25)-C(30)	121.2(6)
N(4)-C(3)-C(11)	112.7(4)	N(24)-C(25)-C(26)	121.4(5)
C(3)-C(11)-O(12)	117.5(4)	C(25)-C(26)-C(27)	120.9(5)
C(3)-C(11)-O(13)	117.7(4)	C(26)-C(27)-C(28)	121.0(5)
O(12)-O(11)-O(13)	124.7(4)	C(27)-C(28)-C(29)	117.9(5)
C(3)-N(4)-C(5)	117.4(3)	C(27)-C(28)-C(31)	121.9(5)
C(3)-N(4)-C(7)	125.8(4)	C(29)-C(28)-C(31)	121.9(5)
N(4)-C(5)-S(1)	102.9(3)	C(29)-C(30)-C(25)	121.8(6)
N(4)-C(5)-C(6)	87.5(3)	C(30)-C(25)-C(26)	117.3(5)
C(5)-N(4)-C(7)	93.3(3)	C(28)-C(31)-O(33)	112.5(4)
S(1)-C(5)-C(6)	115.8(3)	C(28)-C(31)-O(32)	125.7(5)
N(4)-C(7)-O(8)	132.4(5)	O(32)-C(31)-O(33)	121.8(5)
N(4)-C(7)-C(6)	91.5(3)	C(31)-C(33)-C(34)	118.1(4)
O(8)-C(7)-C(6)	135.9(4)	O(33)-C(34)-C(35)	107.7(4)
C(5)-C(6)-C(7)	84.1(3)	C(34)-C(35)-N(36)	114.0(4)
C(7)-C(6)-N(14)	116.8(4)	C(35)-N(36)-C(39)	110.8(4)
C(5)-C(6)-N(14)	121.9(4)	C(35)-N(36)-C(37)	115.4(4)

Atoms	Interatomic Angle (°)	Atoms	Interatomic Angle (°)
C(6)-N(14)-C(15)	121.0(4)	C(37)-N(36)-C(39)	112.2(4)
N(14)-C(15)-O(16)	123.0(4)	N(36)-C(37)-C(38)	113.2(5)
O(16)-C(15)-C(17)	121.2(5)	N(36)-C(39)-C(40)	113.9(4)
C(15)-C(17)-C(18)	111.7(4)		

Table 3.6 Planarity of the thiazolidine and β-lactam rings

Deviations of non-hydrogen atoms from the least squares planes through (a) the thiazolidine ring and (b) the β-lactam ring.

(atoms used in the plane calculation are marked with an asterisk).

(a)

Atom	Deviation $\overset{o}{A}$
*S(1)	-.244
*C(2)	.207
*C(3)	-.072
*N(4)	-.144
*C(5)	.253

The equation of the plane is:-

0.790X - 0.068Y + 0.611Z + 13.374 = 0

where X, Y and Z are orthogonal coordinates in $\overset{o}{A}$ along a*, b and c

(b)

Atom	Deviation $\overset{o}{A}$
*N(4)	.100
*C(5)	-.089
*C(6)	.085
*C(7)	-.096
0(8)	-.381

The equation of the plane is:-

$0.806X - 0.574Y - 0.145Z + 3.274 = 0$

where X, Y and Z are orthogonal coordinates in $\overset{o}{A}$ along a*, b and c.

The angle between the normals to the two planes is

omega = 54.1°

Table 3.7 Planarity of the thiazolidine and β-lactam rings.

Deviations of non-hydrogen atoms from the least squares planes through:

a) the four most co-planar atoms in the thiazolidine ring

b) the most co-planar atoms in the β-lactam ring

c) the substituent atoms of N(4)

(atoms used in the plane calculation are marked with an asterisk)

(a)

Atom	Deviation $\overset{o}{A}$	Atom	Deviation $\overset{o}{A}$
S(1)	-.741	*N(4)	-.050
*C(2)	-.025	*C(5)	.028
*C(3)	.046		

The equation of the plane is:- $0.767X - 0.304Y + 0.566Z + 12.384 = 0$ where X, Y and Z are orthogonal coordinates in $\overset{o}{A}$ along a*, b and c

(b)

Atom	Deviation $\overset{o}{A}$	Atom	Deviation $\overset{o}{A}$
*N(4)	.006	*C(7)	-.022
C(5)	.415	*O(8)	.010
*C(6)	.006		

The equation of the plane is:- $0.676X + 0.683Y + 0.278Z - 0.705 = 0$ where X, Y and Z are orthogonal coordinates in $\overset{o}{A}$ along a*, b and c.

(c)

Atom	Deviation $\overset{\text{o}}{\text{A}}$
*C(3)	.000
N(4)	.401
*C(5)	.000
*C(7)	.000

The equation of the plane is:- $0.729X - 0.684Y + 0.032Z + 4.604 = 0$

where X, Y and Z are orthogonal coordinates in $\overset{\text{o}}{\text{A}}$ along a*, b and c

Table 3.8 Hydrogen bond contact distances

Hydrogen Bond	Donor-Acceptor Distance [$\overset{o}{A}$]
$N(36)_I-H(36) \ldots\ldots O(12)_{II}$	2.836
$O(41)_I-H(41) \ldots\ldots O(12)_{II}$	2.784
$O(41)_I \ldots\ldots\ldots\ldots O(16)_I$	2.612
$N(14)_I-H(14) \ldots\ldots O(13)_{II}$	2.722
$N(24)_I-H(24A) \ldots\ldots O(41)_{III}$	3.239
$N(24)_I-H(24B) \ldots\ldots (O(41)_{IV}$	3.103

The subscripts, I, II, III, IV refer to the equivalent positions:-

x,y,z; $-x,\frac{1}{2}y,-z$; $x,y,-1+z$; $-x,\frac{1}{2}y,1-z$

Figure 3.2 Molecular plot and numbering scheme for APPG
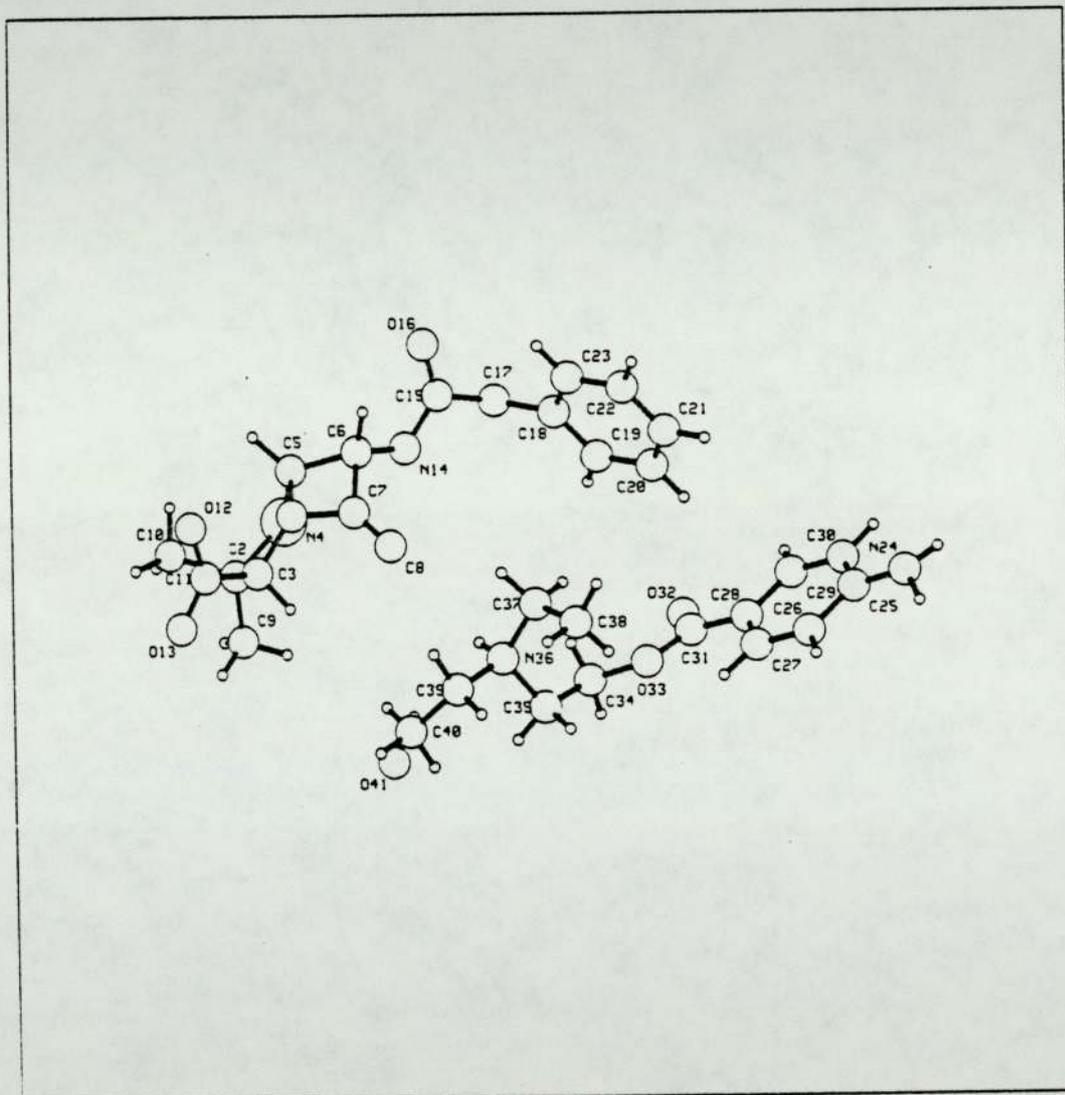
Figure 3.3a   Stereo diagram
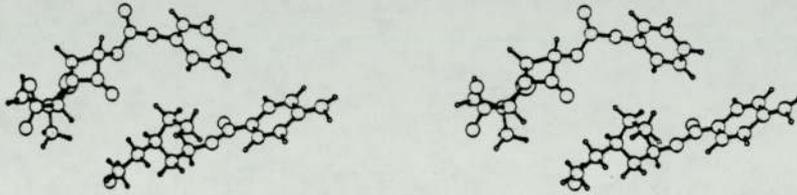
Figure 3.3b   Stereo packing diagram
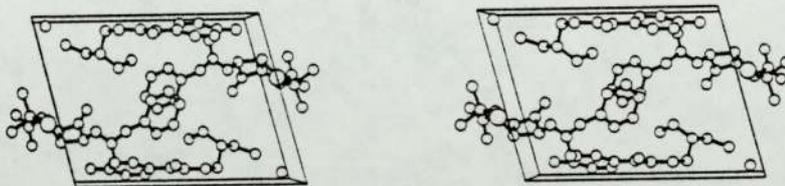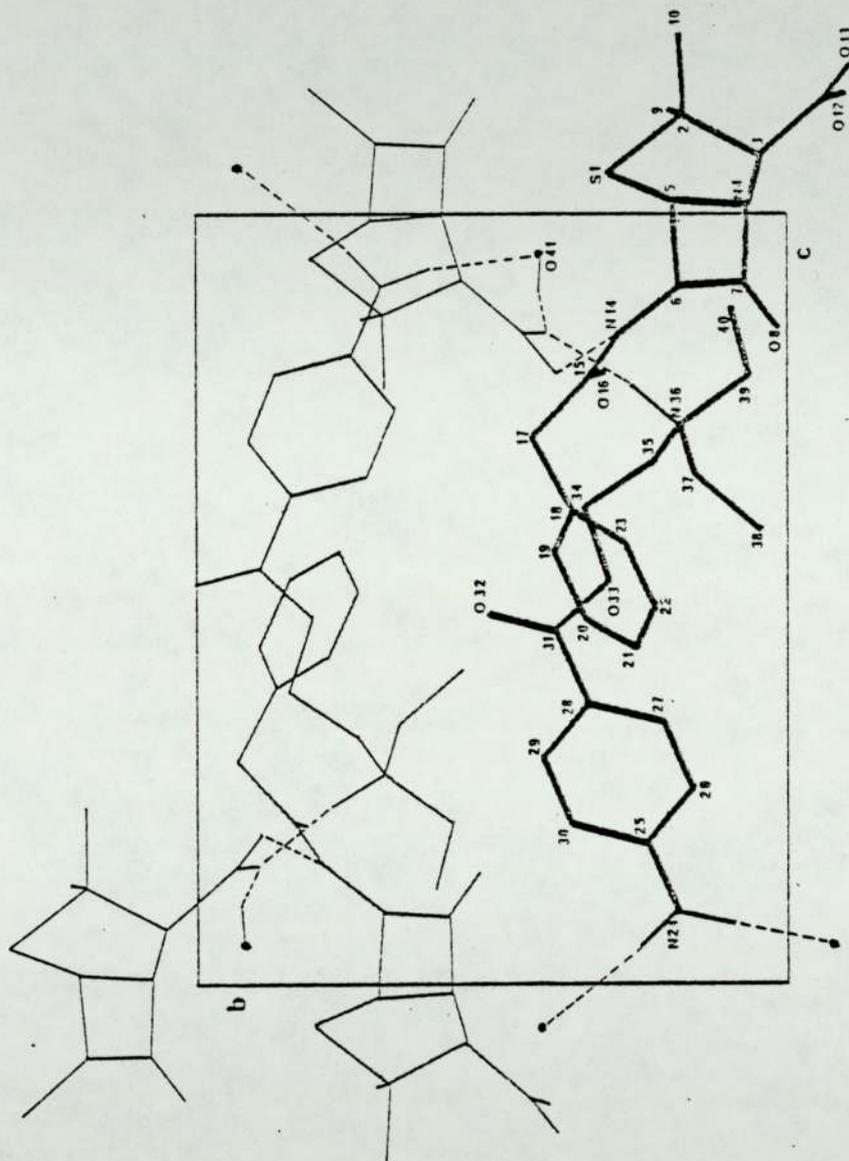
Figure 3.4 Packing diagram showing the unit cell contents projected onto the b-c plane (hydrogen bonds are shown in dotted lines)
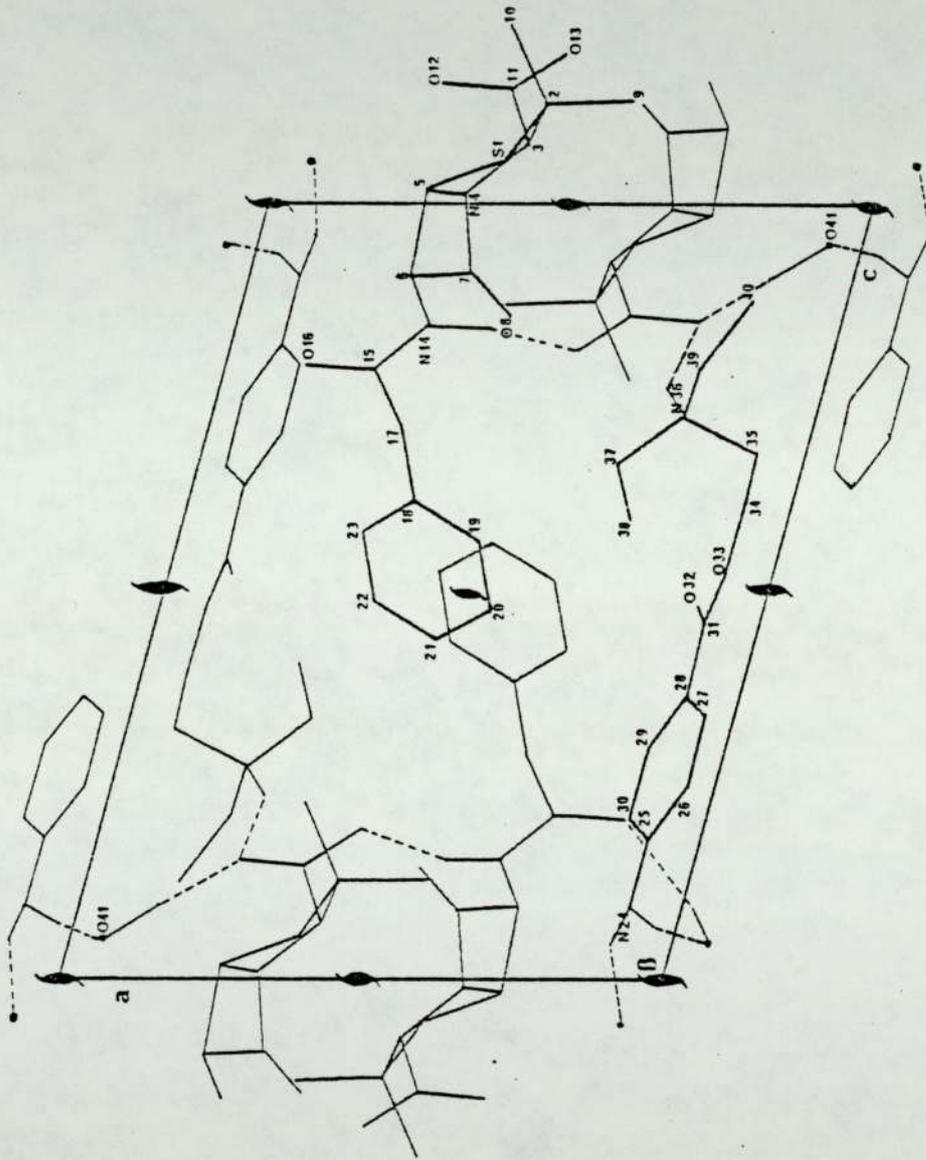
Figure 3.5 Packing diagram showing the unit cell contents projected onto the a-c plane (hydrogen bonds shown in dotted lines)

Figure 3.6   Space filling diagram for APPG
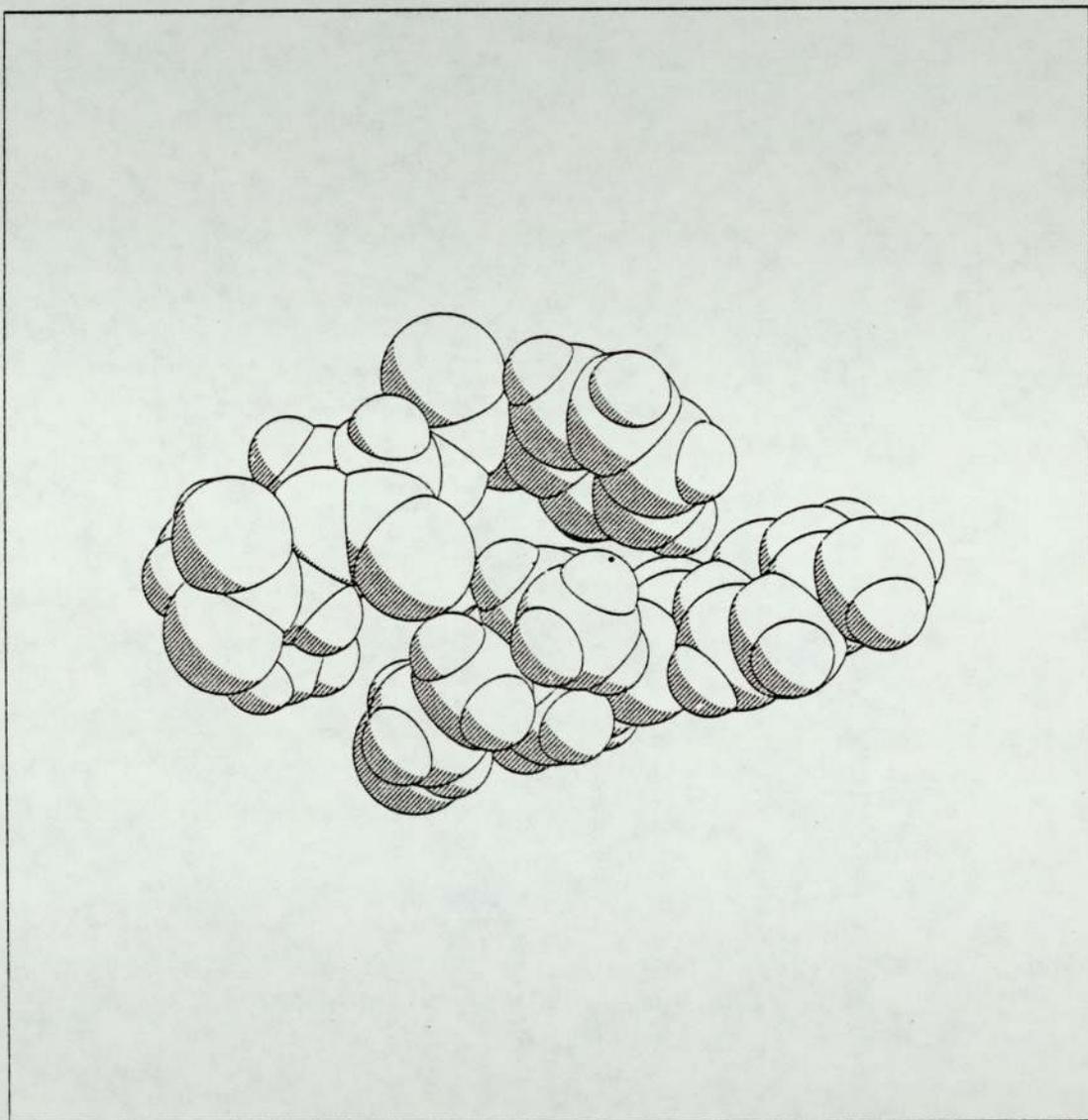
3.2.4   Results and Discussion

The structure of procaine benzylpenicillin monohydrate (APPG) crystallised from aqueous methanol and determined by single crystal x-ray diffractometry is depicted in Figure 3.2 with the numbering scheme used; positional parameters for all atoms with the exception of one of the hydrogens attached to the water molecule are given in Table 3.1 with anisotropic thermal parameters for non-hydrogen atoms and isotropic temperature factors for hydrogen atoms in Tables 3.2 and 3. Details of bond lengths ($\overset{o}{A}$) and interatomic angles (°) are in Tables 3.4 and 5 respectively.

Two independent structure determinations have been done on samples crystallised from two different solvents, the end products having different bioavailabilities. Dexter and Van der Veen (1978)[10] published the structure of APPG crystallised from ethanol finding a space group $P2_1$ with a unit cell volume of 1552.4 $\overset{o}{A}{}^3$ and refining their structure to R = 7.3% for 2718 data (structure DV). At the same time Lowe and Schwalbe (this work)[14] presented the structure of APPG also finding a space group of $P2_1$ with a cell volume of 1554(1)$\overset{o}{A}{}^3$ and refining to R = 5.1% for the 2962 collected data (structure LS).

Structures LS and DV agree well except for a discrepancy exceeding 30 σ in the y co-ordinate of O(12) which we attribute to a typographical error in the DV paper. If the crystals are identical, the xyz coordinates r_i found for each atom should be the same except for random errors. The quantity $\delta = | r_i{}^{LS} - r_i{}^{DV} |$ /(estimated error) should follow the positive half of a normal standard distribution with most discrepancies small, but a few large[17]. A plot of observed versus expected δ should give a straight line of unit slope. The plot for these two determinations is shown in Figure 3.7 the result

of which is a straight line of slope 1.7, intimating that the error estimates are too small but the materials are identical. This suggests that differences in bioavailability between crystals grown from different solvents are probably due to differences of habit not of internal strucure.

APPG is a complex salt in which ion pair formation is produced by protonation of the procaine tertiary nitrogen N(36) by the penicillin carboxylate group giving rise to a penicillin anion and procaine cation. Its structure determination is important from two aspects:-

(1) to see how closely the penicillin nucleus (the known active component) conforms in the solid state to those requirements postulated for therapeutic efficacy.

(2) to examine the structure as a whole in order to comment on its stability to aqueous environments.

Table 3.10a shows a comparison of the bond lengths involved in the thiazolidine and β-lactam rings for APPG and three other penicillin structures. It is immediately obvious that there is little variation in these dimensions and that the bond lengths reported are consistent with those determined in other penicillin structures. The lengthening of the C(2)-C(3) (1.565(6)$\overset{o}{A}$) and C(5)-C(6) (1.561(7)$\overset{o}{A}$) bonds over and above what is considered the norm (1.52-1.54 $\overset{o}{A}$) is probably due to the physical constraints placed upon them in terms of (a) ring strain, particularly in the case of the β-lactam ring and (b) a possible steric effect of the bulky substituents at C(2) and C(3). More interesting, however, is the geometry of the thiazolidine ring which does appear to vary from structure to structure.

Figure 3.7 Half normal probability plot of observed vs. expected δ
for H_2O:MeOH and EtOH

observed $\delta = \dfrac{|r_i^{MeOH} - r_i^{EtOH}|}{\text{est. error}}$

r_i are x,y,z coordinates for each atom

slope = 1.7
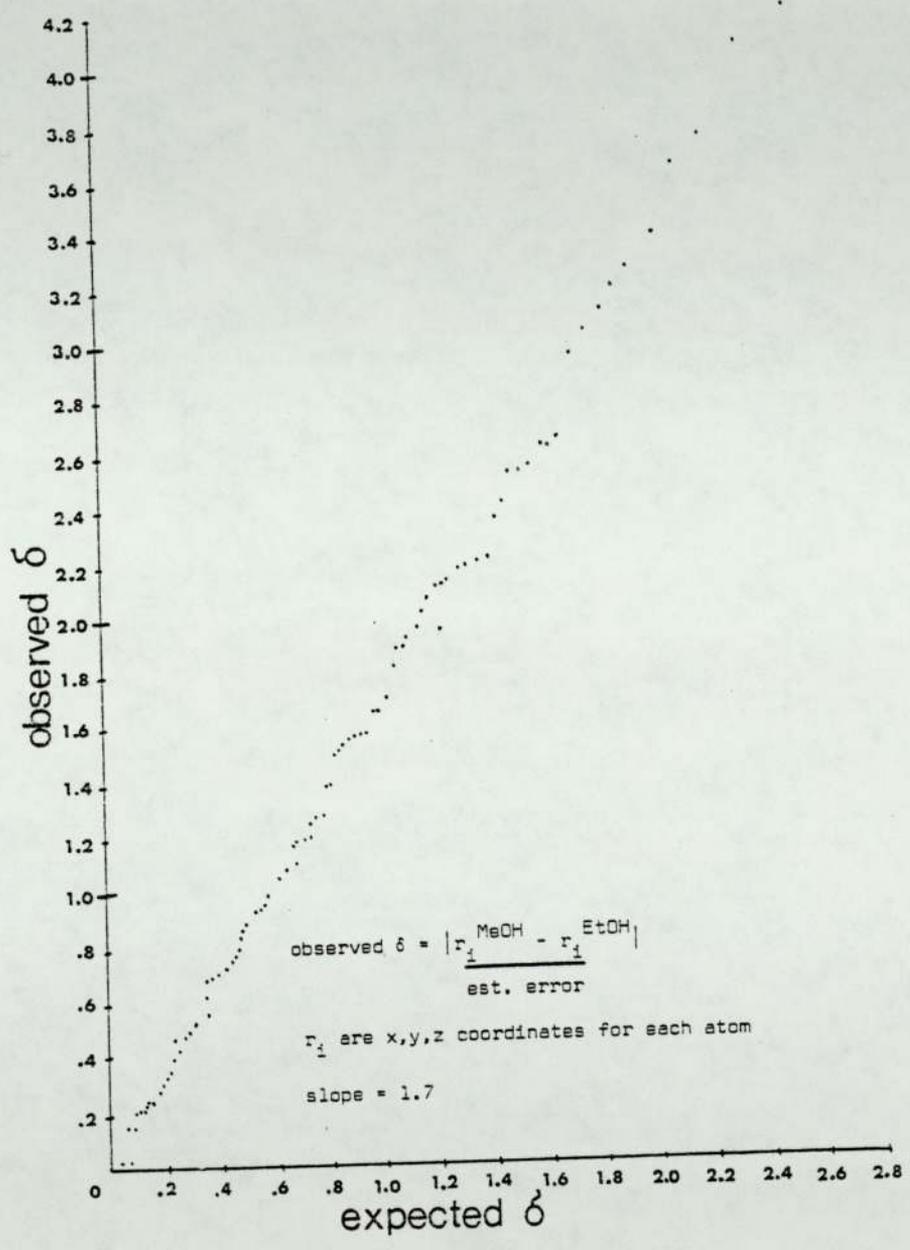
Table 3.10b shows a comparison of the internal bond angles in the thiazolidine and β-lactam rings for four penicillin compounds. The most striking variation in these dimensions is that of the interior angle at the sulphur atom S1 (91.3° for APPG). This angle is known to vary between 89.8° (ampicillin trihydrate) and 96° (6-amino-penicillanic acid and penicillin V)[10]. This variation in the geometry of the thiazolidine ring is further substantiated in a comparison of the S(1)-C(2)-C(3)-N(4) torsion angle for three penicillin structures Table 9; here a difference of 46.2° exists between potassium penicillin G and APPG. Table 3.6a shows the deviations in $\overset{o}{A}$ from the least squares plane through the thiazolidine ring and β-lactam ring. In the former, both S(1) and C(5) show large deviations from the plane (-.244$\overset{o}{A}$ and .253$\overset{o}{A}$ respectively).

Sweet[18] and Boles[19] suggest that four of the five atoms in the thiazolidine ring in penicillin structures tend to be co-planar. In APPG the best plane is formed by atoms C(2), C(3), N(4), C(5) the sulphur atom being displaced - 0.741 $\overset{o}{A}$ on the same side as the β-lactam group (Table 7a). Table 6b shows the β-lactam group to be reasonably planar with the carbonyl oxygen being displaced -0.381 $\overset{o}{A}$ from the least squares plane through the ring. However, atoms N(4), C(6), C(7), O(8) are more nearly co-planar than those forming the four-membered ring, with C(5) displaced 0.415$\overset{o}{A}$ on the same side of this plane as the sulphur atom (Table 3.7b). The nitrogen atom N(4) in the β-lactam is displaced 0.401$\overset{o}{A}$ from the plane of its substituents (C(3), C(5), C(7)) in a pyrimidal geometry (Table 3.7a) with the sum of the bond angles about it being 336.5(6)°. Clearly, N(4) is intermediate between the nitrogen atom in a tertiary amine and that in a normal amide where the atoms are co-planar. The displacement of N(4) serves to reduce electron delocalisation into

the N(4)-C(7) bond thus presumably weakening it. Such effects on the N(4)-C(7) and C(7)-O(8) bond lengths can be compared with the amide substituent at C(6) which has a normal amide geometry (N(4)-C(7) - 1.384(6)$\overset{o}{A}$; N(14)-C(15) - 1.336(7)$\overset{o}{A}$: C(7)-O(8) - 1.191(6)$\overset{o}{A}$; C(15)-O(16) - 1.238(6)$\overset{o}{A}$).

The bond lengths in the carboxylate group appear to reflect the degree to which O(12) and O(13) are hydrogen bonded. The C(11)-O(13) bond distance (1.228(7)$\overset{o}{A}$) is substantially shorter than C(11)-O(12) (1.252(7)$\overset{o}{A}$) the former participating in only one hydrogen bond, whilst the latter has two, one of which is to the procaine tertiary nitrogen. The angle between the least squares plane through the thiazolidine ring and the carboxylate group is 45.4°.

Differences in conformations of the side chain are of course to be expected from one penicillin to another. Examination of the torsion angles (Table 3.9) for the three penicillin structures show two major conformational differences between APPG and benethamine/potassium penicillin G in the C(5)-C(6)-N(14)-C(15) and N(14)-C(15)-C(17)-C(18) torsion angles. This results in an extended side chain in APPG changing through benethamine penicillin to a coiled side chain in potassium penicillin G and in the latter means a shielding of the amide proton by the phenyl group such that it cannot participate in hydrogen bonding.

Bond distances and angles in the procaine cation do not differ with any significance from those determined in other procaine cations, ie. the hydrochloride salt[20] and the dihydrogen orthophosphate salt[21]. Dexter and Van der Veen in their determination noted a diminution of the quinoid character of the p-aminobenzoate group compared to procaine hydrochloride; however, in this determination no diminution is apparent as a comparison of these

bond lengths in Table 3.11 shows.

A major conformational difference does, however, occur in the O(33)-C(34)-C(35)-N(36) torsion angle. Angles ranging from 60 to 72° have been found in other procaine salts, whilst in APPG this angle is 89.5°, this presumably being energically favourable with respect to (a) the hydrogen bonding constraints of N(36) and (b) the general packing of the structure.

Details of the hydrogen bonding are given in Table 3.8 and also in the packing diagrams Figures 3.4 and 5.

As mentioned earlier, the procaine tertiary nitrogen is protonated by the benzylpenicillin carboxyl group and produces an ion pair which is supplemented by an N(36)-H(36) .... O(12) hydrogen bond (2.836$\overset{o}{A}$) which links penicillin to procaine in the direction of the c axis. This is contrary to the description given by Dexter and Van der Veen who suggest an N(36)-H(36) .... O(13) interaction. Atom O(12) also participates in a further interaction in accepting a hydrogen bond from an adjacent water molecule(O(41)-H(41) .... O(13); 2.722$\overset{o}{A}$). This together with a head to tail interaction between adjacent penicillins (N(14)-H(14) ... O(13); 2.722$\overset{o}{A}$) and an O(41) .... O(16) interaction (2.612$\overset{o}{A}$) to a water molecule forms a continuous link along the a axis; the former producing helical chains along the screw axis. Atom O(41) also accepts two hydrogen bonds from the p-amino group of two different procaines (N(24)-H(24) .... O(41); 3.239 $\overset{o}{A}$ and N(24)-H(25) .... O(41); 3.103 $\overset{o}{A}$) to form an interconnected column along the b axis.

The net effect is that the structure is completely interlinked throughout the lattice, with the water molecules acting almost as a 'molecular glue' in holding four fragments together, namely two procaines and two penicillins; the interaction with water is mainly facilitated by the p-amino group of the procaine. As might be expected from the complex non-planar nature of the structure, stacking of the aromatic rings does not appear to be important.

Table 3.9 Important torsion angles in [°] for procaine benzylpenicillin, benethamine benzylpenicillin and potassium benzylpenicillin

Atoms	Procaine	Benethamine	Potassium[10]
S(1)-C(2)-C(3)-N(4)	18.3	-25.2	-27.9
S(1)-C(2)-C(3)-C(11)	140.0	95.9	94.2
S(1)-C(5)-C(6)-N(14)	28.2	21.4	15.0
S(1)-C(5)-N(4)-C(7)	101.1	108.5	114.3
S(1)-C(5)-C(6)-C(7)	-89.8	-95.8	-109.2
C(2)-S(1)-C(5)-C(6)	129.2	105.3	97.1
C(2)-C(3)-N(4)-C(5)	9.3	37.1	32.0
C(2)-C(3)-N(4)-C(7)	-107.6	-94.6	-86.0
C(2)-C(3)-C(11)-O(12)	-105.0	-73.9	-79.0
C(2)-C(3)-C(11)-O(13)	74.0	106.2	99.5
C(3)-N(4)-C(5)-S(1)	-32.5	-29.9	-19.5
C(3)-N(4)-C(5)-C(6)	148.4	-149.0	-140.0
C(3)-N(4)-C(7)-C(6)	142.5	141.5	136.9
N(4)-C(5)-S(1)-C(2)	35.7	9.8	0.1
N(4)-C(5)-C(6)-C(7)	13.3	9.4	5.4
N(4)-C(5)-C(6)-N(14)	131.3	126.6	121.2
N(4)-C(7)-C(6)-C(5)	-14.2	-10.1	-5.7
N(4)-C(7)-C(6)-N(14)	-137.1	-132.2	-123.5
N(4)-C(3)-C(11)-O(13)	165.7	-135.5	-141.8
N(4)-C(3)-C(11)-O(12)	15.7	44.2	40.0
C(5)-S(1)-C(2)-C(3)	-32.0	8.9	16.4
C(5)-N(4)-C(7)-C(6)	15.0	10.7	6.1
C(5)-C(6)-C(7)-N(4)	-14.2	10.1	-5.7
C(5)-C(6)-N(14)-C(15)	123.5	157.8	169.0

Atoms	Procaine	Benethamine	Potassium
C(6)-N(14)-C(15)-C(17)	170.2	177.3	176.0
C(7)-C(6)-N(14)-C(15)	-136.3	-157.0	-94.0
C(9)-C(2)-C(3)-C(11)	-100.0	-147.4	148.6
C(11)-C(3)-N(4)-C(5)	-115.0	-85.4	-91.1
O(11)-C(3)-N(4)-C(7)	-128.0	152.8	149.2
N(14)-C(15)-C(17)-C(18)	-95.1	-157.8	5.0
C(15)-C(17)-C(18)-C(23)	-46.0	-53.9	-91.0
C(15)-C(17)-C(18)-C(19)	136.5	128.0	84.0

Table 3.10a A comparison of the bond lengths [Å] in the thiazolidine and β-lactam rings for four penicillin structures.

length in [Å]

Atoms	BPG	APPG	KPG[10]	Ampicillin Anhyd.[19]
S(1)-C(2)	1.850(3)	1.834(4)	1.847(10)	1.85(1)
S(1)-C(5)	1.801(3)	1.811(5)	1.818(9)	1.79(1)
C(2)-C(3)	1.563(4)	1.565(6)	1.57(1)	1.55(1)
C(3)-N(4)	1.455(4)	1.455(6)	1.46(1)	1.46(1)
N(4)-C(5)	1.458(4)	1.474(6)	1.45(1)	1.45(1)
N(4)-C(7)	1.377(4)	1.384(6)	1.38(1)	1.38(1)
C(5)-C(6)	1.572(4)	1.561(7)	1.56(1)	1.53(1)
C(6)-C(7)	1.542(4)	1.544(7)	1.52(1)	1.52(1)
C(7)-O(8)	1.205(4)	1.191(6)	1.21(1)	1.18(1)
C(11)-O(12)	1.239(4)	1.252(7)	1.23(1)	1.25(1)
C(11)-O(13)	1.274(4)	1.228(7)	1.25(1)	1.19(1)

Table 3.10b A comparison of the internal bond angles [°] in the
thiazolidine and β-lactam rings for the four penicillin
structures

Atoms	BPG	APPG	KPG[10]	Ampicillin Anhyd.[19]
C(5)-S(1)-C(2)	95.5(1)	91.3(2)	95.2(4)	90.6(5)
S(1)-C(2)-C(3)	106.0(2)	105.9(3)	106.2(6)	102.5(7)
C(2)-C(3)-N(4)	105.7(2)	107.2(3)	105.2(7)	106.4(7)
C(3)-N(4)-C(5)	117.1(3)	117.4(3)	119.4(7)	117.8(9)
N(4)-C(5)-S(1)	104.8(2)	102.9(3)	105.1(6)	103.5(8)
C(5)-C(6)-C(7)	83.8(2)	84.1(3)	84.1(7)	85.5(8)
C(6)-C(7)-N(4)	91.2(2)	91.5(3)	93.0(10)	91.8(7)
C(5)-N(4)-C(7)	94.4(3)	93.3(3)	94.1(8)	93.6(7)
N(4)-C(5)-C(6)	87.9(2)	87.5(3)	88.3(7)	88.6(7)

BPG = Benethamine Penicillin G; APPG = Aqueous Procaine Penicillin G;
KPG = Potassium Penicillin G.

Table 3.11 A comparison of the bond lengths (Å) in the p-aminobenzoate group in APPG, procaine hydrochloride and procaine dihydrogen orthophosphate.

Bond	length (Å)		
	APPG	Procaine HCl[20]	Procaine dihydrogen orthophosphate[21]
N(24)-C(25)	1.367(8)	1.359(4)	1.357(3)
C(25)-C(26)	1.406(8)	1.405(4)	1.400(3)
C(25)-C(30)	1.394(8)	1.407(5)	1.395(4)
C(26)-C(27)	1.374(7)	1.367(4)	1.370(4)
C(27)-C(28)	1.398(6)	1.396(4)	1.386(4)
C(28)-C(29)	1.394(8)	1.409(4)	1.394(2)
C(29)-C(30)	1.364(8)	1.356(5)	1.364(3)
C(28)-C(31)	1.460(7)	1.457(4)	1.461(3)
C(31)-O(32)	1.210(6)	1.210(4)	1.206(3)
C(31)-O(33)	1.339(6)	1.356(4)	1.346(2)
O(33)-C(34)	1.447(6)	1.427(6)	1.440(2)

3.3   The Crystal Structure of Benethamine Penicillin G

3.3.1   Experimental

A 2 gram non-sterile sample of benethamine penicillin G was kindly donated by Glaxo Laboratories Ltd, Greenford, Middlesex, UB6 OHE.

A small quantity of the white microcrystalline powder (100-150 mg) was dissolved in warm aqueous methanol (1:3) and allowed to cool.   Slow evaporation overnight produced large well formed prismatic crystals of which the one chosen for data collection (1.0 x 0.44 x 0.16 mm) was a typical example.

The density of the crystals was determined by flotation in n-propanol and carbon tetrachloride and initial unit cell measurements and space-group determination were made by film methods using both oscillation and Weissenberg techniques.   The quality of the crystal was also checked from the films, prior to data collection on the four circle diffractometer.

It should be noted that before obtaining the four circle diffractometer, we had attempted to solve the crystal structure of benethamine penicillin from film data.   Unfortunately however, we were unsuccessful, possibly due to the lack of a sufficient number of good data (1500 reflections in total).   At this stage the diffractometer was installed at Aston and we did not pursue the film data.

3.3.2   Crystal Data

Benethamine Penicillin $(C_{31}H_{35}N_3O_4S)$ crystallises in the monoclinic space group $P2_1$ with a=10.161(6), b=11.020(1), c=13.770(3)$\overset{o}{A}$, β=108.85(5)$\overset{o}{}$, z=2 and v=1459(1)$\overset{o}{A}^3$.

The molecular weight is 545.7 (F(ooo)=580) and the density Dm
(by flotation in carbon tetrachloride and n-propanol) = 1.23 g/ml
with the calculated density Dc=1.24 g/ml. The absorption coefficient
for Mo-Kα radiation (λ = 0.71069Å) is μ = 1.13 cm^{-1}.

3.3.3 Structure Analysis

Data Collection

The data were collected from a crystal of dimensions 1.0 x 0.44
x 0.16 mm mounted along its long axis on an Enraf-Nonius CAD4
diffractometer using the ω-2Θ scan technique with graphite
monochromated Mo-kα radiation (λ = 0.71069Å). The ω scan width =
1.35 + 0.35° tanΘ and the ω scan speed = 1.18 - 3.35° min^{-1}. Two
reference reflections were monitored every 3600 sec of x-ray exposure
time and the orientation of the crystal checked after every 100
reflections. No loss of intensity, or crystal movement were detected
during the collection of the 2944 data between Θ = 0.5° and Θ = 25°
giving 2704 unique observed reflections for which the merged R =
0.098.

The unit cell parameters determined from 25 automatically
centred reflections are as given under 'Crystal Data'.

3.3.4 Structure Determination and Refinement

The data reduction was performed at Birmingham University on the
DEC 20 computer using a programme provided by Dr T A Hamor. Lorentz
and polarisation and intensity corrections were applied and the data
output in the form H K L, Fobs and σ Fobs, where σ Fobs is based on
counting statistics. No correction for absorption effects was made
to the data since for Mo-Kα radiation and the crystal size in
question, the maximum correction which would need to be applied would

be ± 6%.

Normalised structure factors were calculated by a modified Wilson plot method[22]. To aid the normalisation procedure certain geometrically known but randomly oriented fragments were incorporated, namely the thiazolidine and β-lactam rings and also the benzene rings. There are two each of the former and six of the latter in the unit cell. The curves from the Wilson and Debye calculation, and the least squares fit are shown in Figure 3.8 with the values calculated from the gradient and intercept being:-

Temperature factor (B) = 2.9113

Scale = 1.4281

The statistics on the E values followed closely the theoretical values for an acentric structure. Phase determination was undertaken using Multan[1] in which the 500 largest E values with E $\geqslant$ 1.354 produced 5000 unique $\sum_2$ relationships for use in phase refinement by the tangent formula. On the basis of a $\sum_1$ calculation no phases were determined with a probability greater than 95%. The 'Converge' procedure produced the following origin determining reflections:-
11 2 8 (360°); 8 5 $\bar{3}$ (± 45°); 6 4 $\bar{5}$ (45°) with the latter being used to fix the enantiomorph.

Two other reflections used in the starting set were the 2 1 1 and the 6 10 6 the former being given values of ± 45° or ± 135°, the latter being varied in the permutations by 60° commencing at 30°.

The set of phases with the highest CFOM gave the best solution, in which all but seven of the non-hydrogen atoms were located in an E-map. The set of starting phases which produced this had the following initial values:

2	1	1	(135)
11	2	8	(360)
6	10	6	(210)
8	5	$\bar{3}$	(45)
6	4	$\bar{5}$	(45)

Using the atomic positions determined from the E-map, two successive Fourier refinements located the positions of the remaining seven non-hydrogen atoms. Full matrix least squares refinement (SHELX)[2] of positional parameters and isotropic temperature factors for all non-hydrogen atoms ensued during which the origin along y in $P2_1$ was defined by fixing the y coordinate of the sulphur atom. During this stage of refinement atoms were assigned their correct atomic scattering factors on the basis of the known structure of benethamine penicillin; and R, the unweighted discrepancy index, decreased to 0.176. Further attempts to refine the structure all failed with a minimum R factor of 0.169 being obtained and certain bond distances deviating considerably from what would be expected.

Examination of a listing of Fobs and Fcalc showed serious discrepancies in the first 600 reflections collected; these had been collected by the installation engineer using a very poorly refined orientation matrix. Removal of those reflections from the data set and subsequent recollection, followed by isotropic refinement with separate scale factors for the two sets of data, unit weights and the three benzene rings refined as rigid groups, reduced R to 0.102.

Further refinement proceeded with the following constraints applied:

a) all benzene rings with their attached hydrogen atoms refined as rigid groups (AFIX 66 and AFIX 85 respectively).

b) all methyl and methylene hydrogens placed in calculated positions

c) all other non-hydrogen atoms refined anisotropically

d) reflections weighted according to $w = k/\sigma^2(Fo)$ where k is as described below.

This reduced R to 0.078 with four of the five remaining hydrogen atoms appearing on a difference electron density map.

A final sequence of refinements of the complete structure of positional parameters and anisotropic thermal parameters for non-hydrogen atoms and positional parameters and isotropic temperature factors for hydrogens with the remaining outstanding hydrogen atom placed in a calculated position reduced the final agreement parameters to $R = 0.040$ and $Rg = 0.053$. During this refinement the benethamine and the benzylpenicillin were blocked, each block being refined on alternate cycles and the reflections were weighted according to $w = k/[\sigma^2(Fo) + q(Fo)^2]$ where q is a refinable parameter which coverged at 0.0027 and k which should approach unity if there are no serious systematic errors equalled 0.848. The various constraints placed upon the hydrogen atoms during refinement are given in Table 3.14. This stage of the refinement involved 492 parameters and 2457 unique data for which $Fo > 3\sigma Fo$. When the refinement was finally terminated, a difference electron density map showed no feature greater than 0.19 eA^{-3} and the maximum shift associated with a non-hydrogen positional parameter of 0.10σ (0.50σ for the maximum hydrogen positional parameter).

Figure 3.8   Wilson Plot
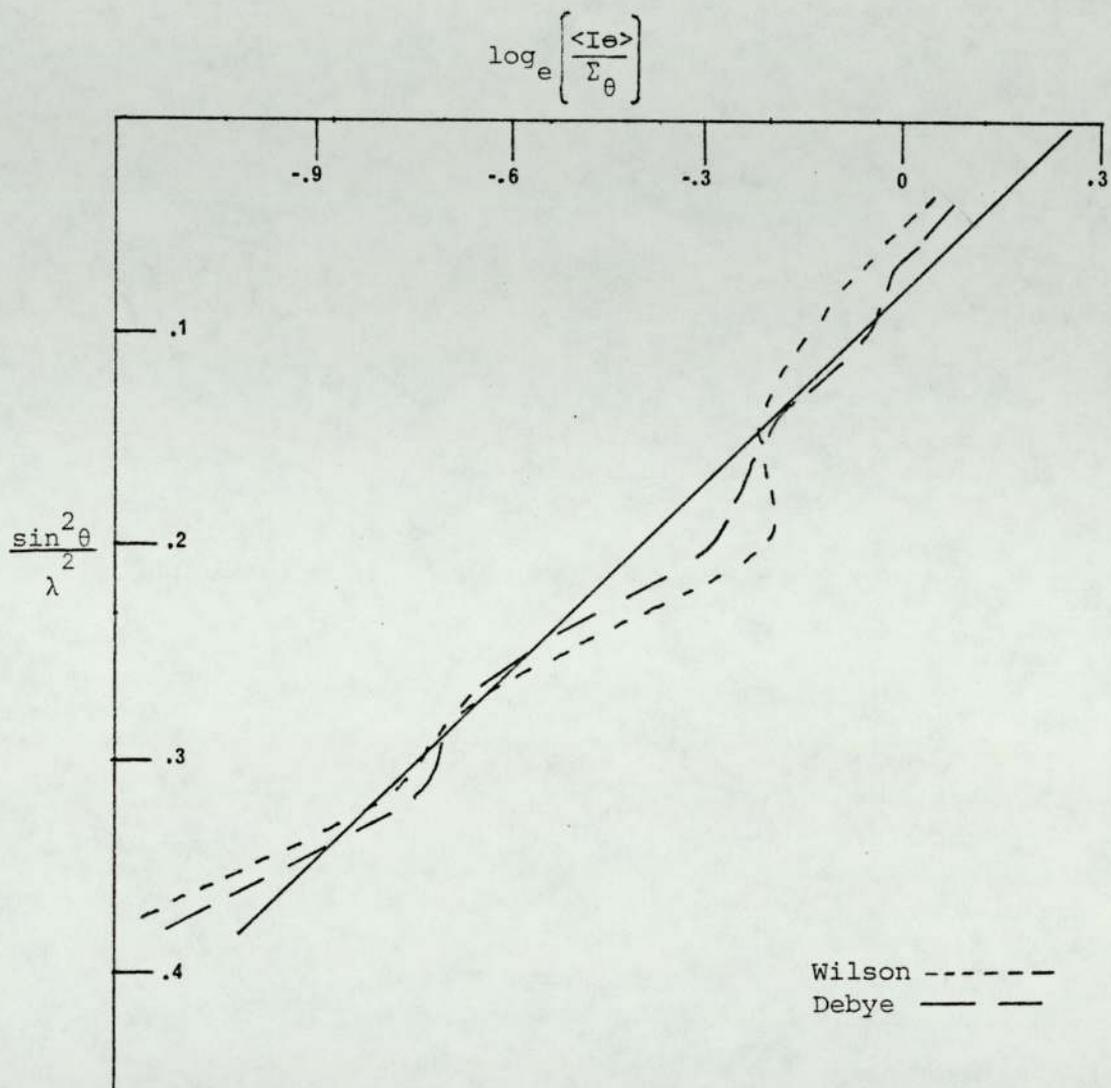
Table 3.12   Positional parameters (fractional co-ordinates x 10^4) with estimated standard deviations in parentheses.

Atom	X/a	Y/b	Z/c
S(1)	2971(1)	2495	5456(1)
C(2)	4089(3)	1648(3)	6048(2)
C(3)	4099(3)	0305(3)	5680(2)
N(4)	3804(2)	0365(2)	4715(2)
C(5)	2803(3)	1275(3)	4634(2)
C(6)	3637(3)	1359(3)	3455(2)
C(7)	4708(3)	0568(3)	3741(2)
O(8)	5887(2)	0238(3)	3323(2)
C(9)	5568(4)	2167(3)	5637(4)
C(10)	3506(5)	1807(3)	7198(3)
C(11)	3504(3)	-0500(3)	6447(2)
O(12)	1782(2)	-0354(2)	6520(2)
O(13)	3486(2)	-1239(2)	6951(2)
N(14)	4126(3)	2492(3)	2975(2)
C(15)	3418(3)	3067(3)	2403(2)
O(16)	2434(2)	2675(3)	2256(2)
C(17)	4116(4)	4281(4)	1984(3)
C(18)	3848(3)	4684(7)	1018(4)
C(19)	4188(5)	3980(5)	0144(4)
C(20)	3965(5)	4437(3)	-0745(3)
C(21)	3422(5)	5572(4)	-0746(3)
C(22)	3080(5)	6261(5)	0097(4)
C(23)	3308(5)	5829(6)	-0986(4)
C(24)	0974(4)	6103(5)	3971(3)
C(25)	0741(4)	7263(4)	4293(3)
C(26)	0757(3)	7465(4)	5283(3)

Atom	X/a	Y/b	Z/c
C(27)	0984(3)	6504(3)	5968(3)
C(28)	1218(4)	5363(4)	5539(3)
C(29)	1221(4)	5162(4)	4544(4)
C(30)	1103(4)	6654(3)	7018(3)
N(31)	0415(2)	7753(2)	7573(2)
C(32)	0591(3)	7842(3)	8608(2)
C(33)	0020(4)	9004(4)	9154(3)
C(34)	0467(4)	9123(4)	10091(3)
C(35)	1632(5)	9809(5)	10065(4)
C(36)	2079(5)	9833(6)	10915(4)
C(37)	1402(5)	9196(6)	11785(4)
C(38)	0251(6)	8521(6)	11813(3)
C(39)	-0219(5)	8500(5)	10975(3)
H(3)	4837(48)	-0029(45)	5469(31)
H(5)	1859(44)	0991(42)	4823(31)
H(6)	3157(40)	0921(37)	2991(28)
H(9A)	5616(50)	3067(27)	5880(34)
H(9B)	5964(40)	2022(37)	4811(19)
H(9C)	6215(40)	1681(40)	5998(31)
H(10A)	2461(27)	1499(34)	7456(25)
H(10B)	3581(52)	2752(25)	7213(37)
H(10C)	4189(34)	1367(36)	7539(26)
H(14)	4902(37)	2908(42)	3114(33)
H(17A)	3485(90)	4821(87)	2626(54)
H(17B)	5226(28)	4152(45)	1760(32)
H(19)	4546(43)	3090(46)	0234(31)
H(20)	4129(55)	4238(51)	-1382(44)
H(21)	3730(86)	5696(85)	-1595(63)

Atom	X/a	Y/b	Z/c
H(22)	2340(110)	7397(112)	0275(71)
H(23)	3122(51)	6381(50)	1611(38)
H(24)	1398(69)	-4053(73)	3145(24)
H(25)	0684(57)	-1976(37)	3942(36)
H(26)	0532(42)	-1677(27)	5720(29)
H(28)	1418(46)	-5416(34)	5935(32)
H(29)	1049(64)	-5700(34)	4257(44)
H(30A)	0603(47)	-4125(33)	7409(33)
H(30B)	2104(29)	-3233(42)	7053(30)
H(31A)	-0664	-1647	7632
H(31B)	0726(27)	8386(25)	7266(19)
H(32A)	0360(57)	7042(34)	9021(37)
H(32B)	1622(24)	7813(29)	8529(22)
H(33A)	-1093(29)	8843(48)	9408(34)
H(33B)	0497(41)	9692(32)	8638(28)
H(35)	1954(53)	10349(57)	9425(40)
H(36)	2682(78)	10308(81)	10956(56)
H(37)	2297(84)	8982(80)	12371(59)
H(38)	-0304(59)	7964(60)	12490(43)
H(39)	-1018(54)	7984(51)	11000(36)

TABLE 3.13   Anisotropic temperature factors (non-hydrogen atoms) and
              isotropic temperature factors (hydrogen atoms) with standard
              deviations in parentheses.

ATOM	U11	U22	U33	U23	U13	U12
S(1)	.0434(4)	.0395(4)	.0517(4)	-.0033(4)	-.0201(3)	.0160(4)
C(2)	.0224(14)	.0320(15)	.0483(16)	.0006(13)	-.0158(12)	-.0018(12)
C(3)	.0188(12)	.0304(14)	.0390(14)	.0000(12)	-.0118(11)	.0051(11)
N(4)	.0289(12)	.0326(13)	.0364(12)	.0016(10)	-.0125(10)	.0014(10)
C(5)	.0264(15)	.0429(17)	.0403(16)	.0063(14)	-.0139(12)	-.0026(13)
C(6)	.0343(16)	.0427(17)	.0363(15)	.0019(13)	-.0170(12)	-.0058(13)
C(7)	.0378(18)	.0355(16)	.0398(15)	-.0002(13)	-.0179(13)	-.0035(13)
O(8)	.0371(12)	.0595(15)	.0444(12)	.0000(11)	-.0049(10)	.0084(12)
C(9)	.0346(17)	.0378(19)	.0846(27)	.0019(17)	-.0221(18)	-.0072(13)
C(10)	.0638(24)	.0425(19)	.0481(19)	-.0066(16)	-.0260(18)	.0045(18)
C(11)	.0290(14)	.0285(14)	.0394(15)	.0006(12)	-.0164(12)	-.0003(12)
O(12)	.0215(10)	.0463(12)	.0574(13)	.0154(11)	-.0172(9)	-.0035(9)
O(13)	.0317(11)	.0508(13)	.0527(13)	.0186(11)	-.0195(10)	.0004(10)
N(14)	.0342(13)	.0459(14)	.0424(13)	.0083(13)	-.0238(11)	-.0100(13)
C(15)	.0288(15)	.0454(16)	.0440(16)	.0043(14)	-.0189(13)	-.0014(14)
O(16)	.0370(12)	.0687(18)	.0883(17)	.0224(15)	-.0428(12)	-.0138(13)
C(17)	.0572(23)	.0564(22)	.0734(25)	.0195(20)	-.0448(20)	-.0145(19)
C(18)	.0378(17)	.0514(19)	.0458(17)	.0093(15)	-.0231(14)	-.0079(15)
C(19)	.0538(24)	.0751(30)	.0748(27)	-.0048(23)	-.0267(21)	.0033(22)
C(20)	.0693(29)	.1292(52)	.0431(22)	-.0142(27)	-.0155(21)	-.0158(32)
C(21)	.0061(28)	.1017(40)	.0638(28)	.0315(28)	-.0272(22)	-.0084(27)
C(22)	.0780(31)	.0772(30)	.0662(27)	.0246(25)	-.0306(23)	.0020(25)
C(23)	.0670(26)	.0621(26)	.0541(22)	.0124(20)	-.0245(19)	-.0027(21)

ATOM	U11	U22	U33	U23	U13	U12
C(24)	.0361(19)	.0993(24)	.0497(21)	-.0198(23)	-.0140(23)	-.0103(21)
C(25)	.0400(18)	.0767(28)	.0541(20)	-.0070(19)	-.0239(16)	.0077(18)
C(26)	.0379(16)	.0555(20)	.0518(17)	-.0067(18)	-.0223(14)	.0032(17)
C(27)	.0266(15)	.0395(17)	.0525(17)	.0009(14)	-.0170(13)	-.0042(12)
C(27)	.0266(15)	.0395(17)	.0525(17)	.0009(14)	-.0170(13)	-.0042(12)
C(28)	.0455(20)	.0434(20)	.0655(23)	-.0039(18)	-.0064(17)	-.0047(16)
C(29)	.0462(21)	.0626(26)	.0748(27)	-.0216(23)	-.0078(19)	-.0106(19)
C(30)	.0406(19)	.0424(19)	.0597(20)	.0049(16)	-.0243(16)	.0051(15)
N(31)	.0250(11)	.0421(16)	.0412(12)	.0094(11)	-.0142(10)	-.0037(10)
C(32)	.0363(16)	.0588(23)	.0379(16)	.0087(14)	-.0164(13)	-.0028(15)
C(33)	.0482(20)	.0738(26)	.0433(18)	-.0028(18)	-.0172(16)	.0155(20)
C(34)	.0433(19)	.0657(23)	.0437(18)	.0107(17)	-.0133(15)	-.0111(18)
C(35)	.0541(24)	.0798(30)	.0647(26)	.0068(23)	-.0147(20)	.0031(23)
C(36)	.0553(25)	.0976(39)	.0906(36)	.0301(31)	-.0342(25)	.0018(26)
C(37)	.0695(29)	.1128(42)	.0593(25)	.0236(28)	-.0290(23)	-.0115(30)
C(38)	.0829(33)	.1007(37)	.0425(20)	.0076(23)	-.0233(21)	-.0007(29)
C(39)	.0528(23)	.0892(31)	.0448(20)	-.0010(20)	-.0115(18)	.0060(23)

	Uiso
H(3)	.069(13)
H(5)	.059(12)
H(6)	.052(10)
H(9A)	.071(13)
H(9B)	.054(10)
H(9C)	.069(12)
H(10A)	.043(9)
H(10B)	.077(14)

Atom	Uiso
H(10C)	.051(10)
H(14)	.072(13)
H(17A)	.200(31)
H(17B)	.069(12)
H(19)	.059(11)
H(20)	.079(15)
H(21)	.156(28)
H(22)	.195(33)
H(23)	.073(13)
H(24)	.138(27)
H(25)	.086(15)
H(26)	.053(10)
H(28)	.068(13)
H(29)	.104(19)
H(30A)	.079(15)
H(30B)	.061(11)
H(31A)	.050
H(31B)	.015(6)
H(32A)	.087(16)
H(32B)	.030(7)
H(33A)	.081(14)
H(33B)	.058(11)
H(35)	.084(15)
H(36)	.124(24)
H(37)	.153(29)
H(38)	.100(18)
H(39)	.079(15)

Table 3.14 Bond distances (Å) with estimated standard deviations in parentheses

Benzylpenicillin:

Bond	Interatomic Distance (Å)
S(1)-C(2)	1.850(3)
S(1)-C(5)	1.801(3)
C(2)-C(3)	1.563(4)
C(3)-N(4)	1.455(4)
N(4)-C(5)	1.458(4)
N(4)-C(7)	1.377(4)
C(5)-C(6)	1.572(4)
C(6)-C(7)	1.542(5)
C(7)-O(8)	1.205(4)
C(2)-C(9)	1.535(5)
C(2)-C(10)	1.510(5)
C(3)-C(11)	1.518(4)
C(11)-O(12)	1.274(4)
C(11)-O(13)	1.239(4)
C(6)-N(14)	1.425(4)
N(14)-C(15)	1.336(5)
C(15)-O(16)	1.231(5)
C(15)-C(17)	1.517(5)
C(17)-C(18)	1.507(6)
C(18)-C(23)	1.383(6)
C(18)-C(19)	1.380(6)
C(19)-C(20)	1.409(8)
C(20)-C(21)	1.366(9)
C(21)-C(22)	1.336(8)
C(22)-C(23)	1.401(8)

Benethamine:

Bond	Interatomic Distance (Å)
C(24)-C(25)	1.392(7)
C(25)-C(26)	1.386(6)
C(26)-C(27)	1.396(6)
C(27)-C(28)	1.377(5)
C(28)-C(29)	1.387(7)
C(29)-C(24)	1.375(7)
C(27)-C(30)	1.498(6)
C(30)-N(31)	1.481(4)
N(31)-C(32)	1.496(5)
C(32)-C(33)	1.503(6)
C(33)-C(34)	1.507(6)
C(34)-C(35)	1.394(6)
C(35)-C(36)	1.387(9)
C(36)-C(37)	1.359(8)
C(37)-C(38)	1.370(8)
C(38)-C(39)	1.395(7)
C(39)-C(34)	1.375(5)

Bonds involving hydrogen atoms:

Bond	Interatomic Distance (Å)
C(3)-H(3)	0.80(5)
C(5)-H(5)	0.84(5)
C(6)-H(6)	1.04(5)
C(9)-H(9A)	[a]1.05(3)
C(9)-H(9B)	[a]1.09(3)
C(9)-H(9C)	[a]1.09(5)
C(10)-H(10A)	[a]1.06(3)

Bond	Interatomic Distance ($\overset{o}{A}$)
C(10)-H(10B)	[a]1.05(3)
C(10)-H(10C)	[a]1.07(4)
N(14)-H(14)	[b]0.98(4)
C(17)-H(17A)	1.09(8)
C(17)-H(17B)	1.08(3)
C(19)-H(19)	1.07(5)
C(20)-H(20)	0.87(6)
C(21)-H(21)	1.11(9)
C(22)-H(22)	1.09(6)
C(23)-H(23)	1.02(5)
C(24)-H(24)	[a]1.05(3)
C(25)-H(25)	[a]1.06(5)
C(26)-H(26)	[a]1.05(3)
C(28)-H(28)	[a]1.07(4)
C(29)-H(29)	[a]1.06(5)
C(30)-H(30A)	[a]1.05(4)
C(30)-H(30B)	[a]1.04(3)
N(31)-H(31A)	[c]1.08
N(31)-H(31B)	0.81(3)
C(32)-H(32A)	[a]1.03(4)
C(32)-H(32B)	[a]1.02(3)
C(33)-H(33A)	[a]1.08(4)
C(33)-H(33B)	[a]1.05(3)
C(35)-H(35)	1.02(6)
C(36)-H(36)	0.82(8)
C(37)-H(37)	1.10(9)
C(38)-H(38)	1.10(6)
C(39)-H(39)	0.98(6)

a   Hydrogen positions refined with a constrained bond distance of 1.07$\overset{o}{A}$ with a standard deviation of 0.03.

b   Hydrogen positions refined with a constrained bond distance of 1.01$\overset{o}{A}$ with a standard deviation of 0.03.

c   Hydrogen position calculated using AFIX command (SHELX) and fixed during least squares refinement.

Table 3.15   Interatomic angles (°) (non-hydrogen atoms) with
estimated standard deviations in parentheses.

Atoms	angle°	Atoms	angle°
C(2)-S(1)-C(5)	95.5(1)	C(17)-C(18)-C(19)	122.6(5)
S(1)-C(2)-C(3)	106.0(2)	C(17)-C(18)-C(23)	118.9(5)
S(1)-C(2)-C(9)	108.5(2)	C(19)-C(18)-C(23)	118.4(5)
S(1)-C(2)-C(10)	108.3(2)	C(18)-C(19)-C(20)	119.5(5)
C(3)-C(2)-C(9)	109.3(2)	C(19)-C(20)-C(21)	120.4(5)
C(3)-C(2)-C(10)	114.1(3)	C(20)-C(21)-C(22)	120.7(5)
C(9)-C(2)-C(10)	110.4(3)	C(21)-C(22)-C(23)	119.8(5)
C(2)-C(3)-N(4)	105.7(2)	C(18)-C(23)-C(22)	121.2(4)
C(2)-C(3)-C(11)	114.2(2)	C(25)-C(24)-C(29)	124.9(5)
N(4)-C(3)-C(11)	124.5(2)	C(24)-C(25)-C(26)	114.8(4)
C(3)-N(4)-C(5)	117.1(3)	C(25)-C(26)-C(27)	124.3(4)
C(3)-N(4)-C(7)	137.1(3)	C(26)-C(27)-C(28)	114.8(4)
C(5)-N(4)-C(7)	94.4(3)	C(26)-C(27)-C(30)	127.7(3)
S(1)-C(5)-N(4)	104.8(2)	C(28)-C(27)-C(30)	117.3(3)
S(1)-C(5)-C(6)	118.4(2)	C(27)-C(28)-C(29)	121.0(4)
N(4)-C(5)-C(6)	87.9(2)	C(24)-C(29)-C(28)	120.3(4)
C(5)-C(6)-C(7)	83.8(2)	C(27)-C(30)-N(31)	114.6(3)
C(5)-C(6)-N(11)	121.4(3)	C(30)-N(31)-C(32)	111.3(3)
C(7)-C(6)-N(14)	113.4(3)	N(31)-C(32)-C(33)	112.8(3)
N(4)-C(7)-C(6)	91.2(2)	C(32)-C(33)-C(34)	109.2(3)
N(4)-C(7)-O(8)	108.6(3)	C(33)-C(34)-C(35)	120.9(3)
C(6)-C(7)-O(8)	157.1(3)	C(33)-C(34)-C(39)	120.5(3)
C(3)-C(11)-O(12)	118.6(3)	C(35)-C(34)-C(39)	118.5(4)
C(3)-C(11)-O(13)	116.4(3)	C(34)-C(35)-C(36)	119.7(4)
O(12)-C(11)-O(13)	125.0(3)	C(35)-C(36)-C(37)	121.4(5)

C(6)-N(14)-C(15)	121.7(3)	C(34)-C(39)-C(38)	121.1(5)
N(14)-C(15)-O(16)	114.4(3)	C(37)-C(37)-C(38)	119.0(5)
O(16)-C(15)-C(17)	123.0(3)		
C(15)-C(17)-C(18)	114.8(4)		

Table 3.16   Planarity of the thiazolidine and β-lactam rings

Deviations of non-hydrogen atoms from the least squares planes through (a) the thiazolidine ring and (b) the β-lactam ring.

(atoms used in the plane calculation are marked with an asterisk)

(a)

Atom	Deviation $\overset{o}{A}$
*S(1)	-.007
*C(2)	.100
*C(3)	-.183
*N(4)	.194
*C(5)	-.100

The equation of the plane is:- $0.845X - 0.335Y + 0.417Z + 0.195 = 0$
where X, Y and Z are orthogonal coordinates in $\overset{o}{A}$ along a*, b and c

(b)

Atom	Deviation Å
*N(4)	-.025
*C(5)	.027
*C(6)	-.028
*C(7)	.026
O(8)	.435

The equation of the plane is:- $0.533X - 0.742Y + 0.408Z + 5.428 = 0$

where X, Y and Z are orthogonal coordinates in Å along a*, b and c

The angle between the normals to the two planes is Omega = 58.1°

Table 3.17   Planarity of the thiazolidine and β-lactam rings

Deviations of non-hydrogen atoms from the least squares planes through:-

a)   the four most co-planar atoms in the thiazolidine ring

b)   the most co-planar atoms in the β-lactam system

c)   the substituent atoms of N(4)

(atoms used in the plane calculation are marked with an asterisk)

(a)

Atom	Deviation $\overset{o}{A}$	Atom	Deviation $\overset{o}{A}$
*S(1)	-.057	N(4)	.403
*C(2)	.065	*C(5)	.041
*C(3)	-.049		

The equation of the plane is:- $0.846X + 0.244Y - 0.474Z - 0.878 = 0$
where X, Y and Z are orthogonal coordinates in $\overset{o}{A}$ along a*, b and c

(b)

Atom	Deviation $\overset{o}{A}$	Atom	Deviation $\overset{o}{A}$
*N(4)	.013	*C(7)	-.127
C(5)	.293	*O(8)	.063
*C(6)	.051		

The equation of the plane is:- $0.358X + 0.819Y + 0.447Z + 5.093 = 0$

where X, Y and Z are orthogonal coordinates in $\overset{o}{A}$ along a*, b and c

(c)

Atom	Deviation $\overset{o}{A}$
*C(3)	.000
N(4)	-.338
*C(5)	.000
*C(7)	.000

The equation of the plane is:- $0.466X + 0.865Y + 0.184Z + 3.818 = 0$

where X, Y and Z are orthogonal coordinates in $\overset{o}{A}$ along a*, b and c

Table 3.18   Hydrogen bond contact distances and angles

Hydrogen Bond	Donor-Acceptor Distance [$\overset{o}{A}$]
$N(14)_I-H(14) \ldots\ldots O(13)_{II}$	2.831
$N(31)_I-H(31A) \ldots\ldots O(16)_{III}$	2.830
$N(31)_I-H(31B) \ldots\ldots O(12)_{IV}$	2.662

The subscripts I, II, III, IV refer to the equivalent positions:-
x,y,z; $1-x, \frac{1}{2}y, 1-z$; $-x, \frac{1}{2}y, 1-z$; $x, 1+y, z$

Figure 3.9   Molecular plot and numbering scheme for BPG

Figure 3.10a Stereo diagram
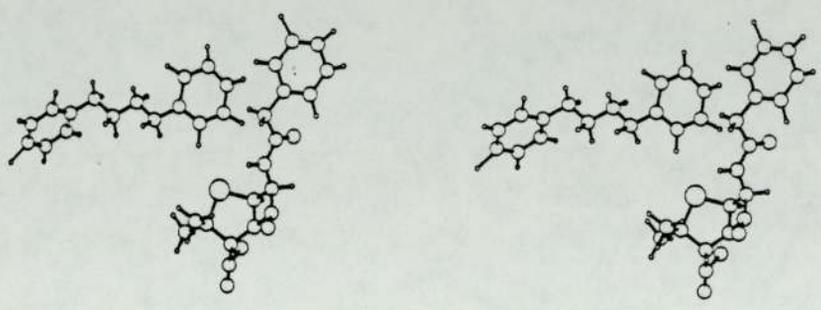
Figure 3.10b Stereo packing diagram
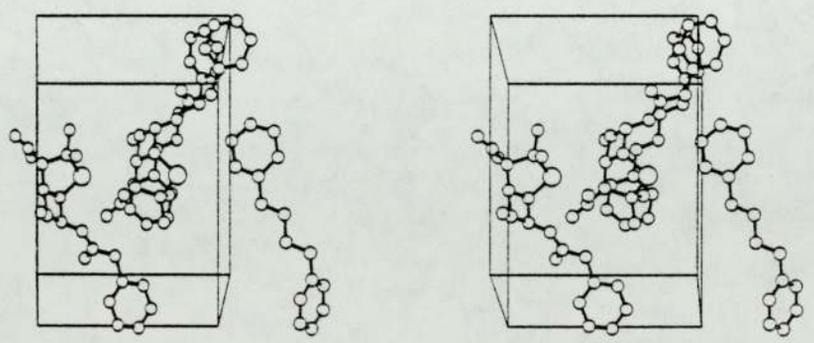
Figure 3.11 Packing diagram showing hydrogen bonding in dotted lines (b-c plane)
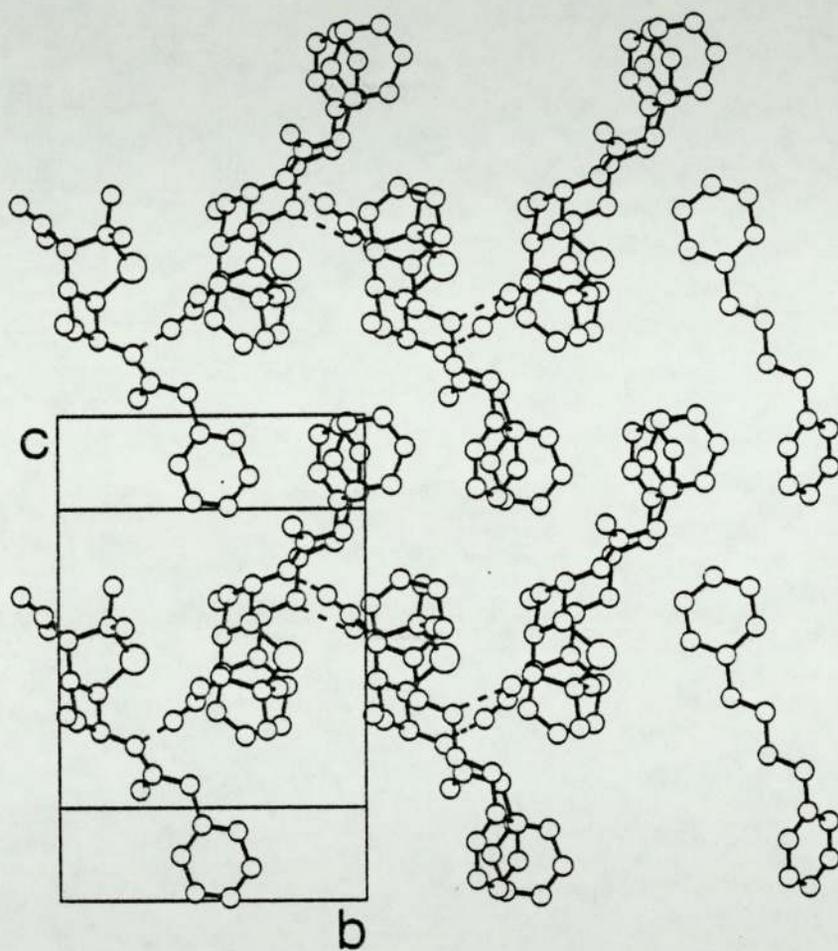
Figure 3.12 Packing diagram showing hydrogen bonding in dotted lines (a-c plane)
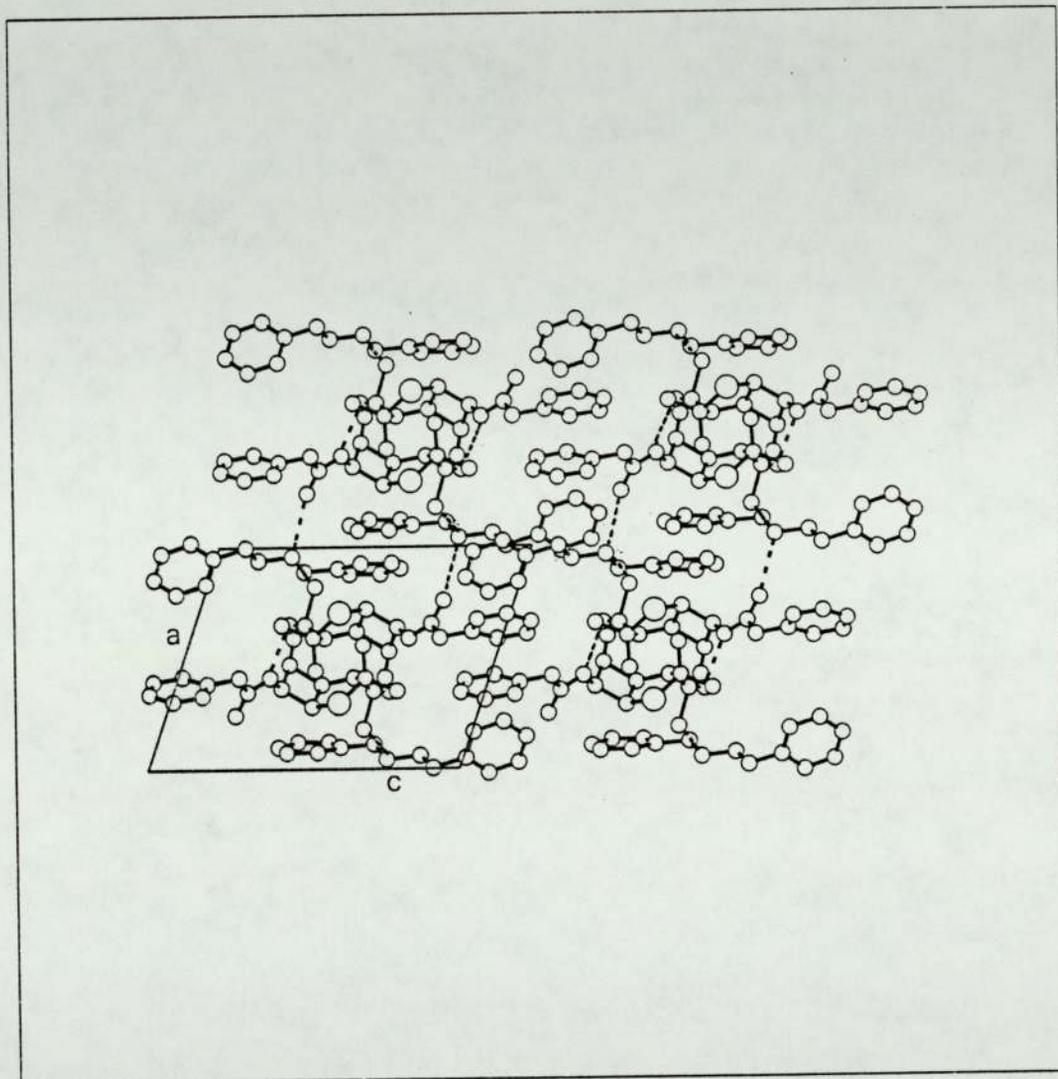
Figure 3.13   Space filling diagram for BPG
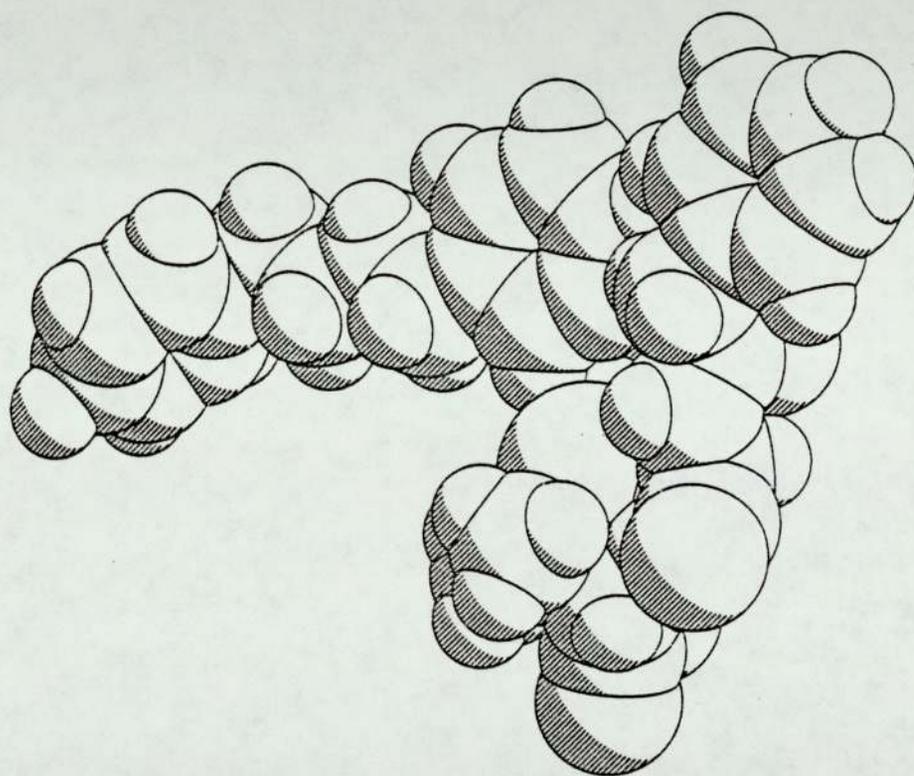
3.3.5   Results and Discussion

The structure of benethamine benzylpenicillin (BPG) as determined by single crystal x-ray diffractometry is depicted in Figure 3.9 together with the numbering scheme used; positional parameters for all atoms are given in Table 3.12 with anisotropic thermal parameters for non-hydrogen atoms and isotropic temperature factors for hydrogen atoms in Table 3.13.

Details of bond lengths ($\overset{o}{A}$) and interatomic angles (°) are in Tables 3.14 and 15 respectively.

As with APPG, BPG is a complexed salt in which ion-pair formation is produced by protonation of the secondary nitrogen N(31) by the penicillin carboxylate group giving rise to a penicillin anion and a benethamine cation.

The bond lengths in the thiazolidine and β-lactam rings are again fairly typical of those found in penicillin structures as the comparison in Table 3.10a shows. The slight differences in the S(1)-C(2), S(1)-C(5) and N(4)-C(5) bond lengths ($1.850(3)\overset{o}{A}$ to $1.834(4)\overset{o}{A}$; $1.801(3)\overset{o}{A}$ to $1.811(5)\overset{o}{A}$; $1.458(4)\overset{o}{A}$ to $1.474(6)\overset{o}{A}$) for BPG and APPG are probably due to the differing geometry of the two thiazolidine rings as mentioned previously in the discussion of APPG.

The comparison of the internal angles in the two ring systems again shows the significant variation in the interior angle at the sulphur atom S(1) and in BPG (95.5(1)°) this is at the upper limit of the expected range, similar to that of potassium penicillin G[10], penicillin V[23] and 6-aminopenicillanic acid[24]. Deviations in $\overset{o}{A}$ from the least squares plane through the five atoms in the thiazolidine ring (Table 3.16a) show C(3) and N(4) to be the most non-co-planar with deviations of $-0.183\overset{o}{A}$ and $0.194\overset{o}{A}$ respectively. Table 3.17a however, shows the best plane is formed by atoms S(1), C(2), C(3) and

C(5) with N(4) being displaced 0.403Å. This differs from APPG where
the sulphur atom S(1) is the least co-planar and is almost identical
to 6-aminopenicillanic acid which also has N(4) displaced 0.4Å from
the plane through the other four atoms in the thiazolidine ring.
This difference as might be expected is reflected in the torsion
angle S(1)-C(2)-C(3)-N(4) for the two compounds (Table 3.9) in which
the variation between the two is 43.5°.

Table 3.17b shows the four-membered lactam ring to be very
nearly planar with the carbonyl oxygen atom O(8) displaced 0.435Å
from the plane on the same side as the thiazolidine ring.   Like
APPG, in this case, atoms N(4), C(6), C(7) and O(8) are still
reasonably co-planar, with atom C(5) deviating by 0.293Å from the
least squares plane through the four atoms.   As seems to be
consistent throughout the penicillins, the nitrogen atom N(4) is
displaced 0.338Å from the plane of its substituents (C(3), C(5) and
C(7)) in a pyrimidal geometry (Table 3.17c) with the sum of the bond
angles about it being 343.9(5)°.   As with APPG this reduces the
electron delocalisation into the N(4)-C(7) bond compared with that of
a planar amide. The fact that N(4) in BPG is less pyrimidal than in
APPG is reflected in the N(4)-C(7) and C(7)-O(8) bond lengths for the
two compounds (BPG:- N(4)-C(7) - 1.377(4)Å, C(7)-O(8) - 1.205(4)Å:
APPG:- N(4)-C(7) - 1.384(6)Å C(7)-O(8) - 1.191(6)Å) in that in BPG
N(4)-C(7) is shortened slightly whilst C(7)-O(8) is lengthened.

The N(14)-C(15) and C(15)-O(16) bonds, 1.336(5)Å and 1.231(5)Å
respectively, are almost identical to those in APPG with N(14)-C(15)
showing considerable double bond character as might be expected in a
normal planar amide.

Again the carboxylate group shows two significantly varying C-O
bond lengths in C(11)-O(12) (1.239(4)Å) and C(11)-O(13) (1.274(4)Å).

Both oxygens participate in a single hydrogen bond and therefore at first sight it appears difficult to rationalise the difference in bond lengths in terms of the degree of hydrogen bonding. However, O(12) forms a very short hydrogen bond with the charged secondary nitrogen N(31) (2.662Å) and it seems likely that there is some polarisation such that $C \overset{\overline{O}}{\underset{O}{\lessgtr}} \cdots\cdots H\overset{+}{N}$ is the dominant resonance form.

The conformation of the side chain substituent at C(6) in BPG as discussed in the previous section is intermediate betwen the extended form in APPG and the coiled conformation in potassium penicillin G, as suggested by the C(5)-C(6)-N(14)-C(15) and N(14)-C(15)-C(17)-C(18) torsion angles (Table 3.9). Unlike potassium penicillin G however, there is no shielding of the N(14) hydrogen atom by the phenyl group, and this hydrogen participates effectively in the hydrogen bonding of the structure.

The bond lengths and angles in the benethamine cation show little variation from what might be expected in a residue which has neither electron withdrawing nor donating groups attached to the phenyl rings. It is fully extended in the direction of c with a classical aliphatic backbone with the two phenyl rings rotated approximately 64.4° with respect to each other. The details of the hydrogen bonding are given in Table 3.18 with Figures 3.11 and 12 showing the molecular packing and hydrogen bonds in dotted lines. Like APPG, BPG forms an ion pair in which the secondary nitrogen in the benethamine residue (N(31)) is protonated by the carboxyl group of the benzylpenicillin, this being supplemented by an N(31)-H(31B) ... O(12) hydrogen bond which this time is shorter and stronger than in APPG at 2.662Å with an NHO angle of 160.2°. Adjacent ion pairs are linked in the direction of b by an N(14)-H(14) ... O(13) interaction (2.831Å) and in the direction of a by an N(31)-H(31A) ...

O(16) hydrogen bond. The net result is that molecules are hydrogen bonded in the lattice in two dimensions <u>a</u> and <u>b</u>, but not in <u>c</u>, which has only the hydrophobic phenyl rings in the direction of its axis.

3.4 Conclusions

The most striking structural feature of the penicillins and the cephalosporins is the large pyrimidal character of the β-lactam nitrogen atom N(4).   Sweet et al (1970)[25] report the structures of two cephalosporins with a pyrimidal nitrogen in the β-lactam ring, both of which are biologically active and one in which the β-lactam nitrogen is nearly planar, which is inactive. Both BPG and APPG exhibit the pyrimidal nitrogen at the N(4) position. The ease of base hydrolysis of the lactam amide bond in these antibiotics seems to correlate with biological activity[25].   Since the pyrimidal nitrogen (N(4)) is likely to weaken the amide C-N bond in the β-lactam ring compared to a normal planar amide where the unshared electron pair of the nitrogen atom can be involved in π-bonding with the adjacent carbonyl carbon atom, it is hypothesised that the nature of the C-N amide bond is related to biological activity. Certainly the results of this work show the C(7)-N(4) bond in both APPG and BPG to have considerably more single bond character than an equivalent bond in a planar amide.

Tipper and Strominger[8] suggest that the penicillins imitate a possible conformation of D-alanyl-D-alanine, the substrate of peptidoglycan transpeptidase; the enzyme mistakes the penicillin for its proper substrate, cleaves the amide bond in the penicillin's β-lactam ring, is thus acylated and is blocked irreversibly from further activity.   The assertion that valid structure-activity relationships can be derived from solid state geometries is based on the premise that the fused ring conformations found in the crystalline state for the penicillins and cephalosporin derivatives remain unchanged upon their dissolution. Prime evidence in support of this basic assumption comes from the structural investigation of

penicillin V sulphoxide.[26] X-ray diffraction and nuclear Overhauser effect studies showed that the fused ring system in this compound adopts the same detailed conformation both in the solid state and in solution.

3.4.1 The stability of APPG and BPG to aqueous environments

Both APPG and BPG are complexed salts in which the carboxyl group of the penicillin protonates N(36) (APPG) or N(31)(BPG) in the counter-ion. In both cases, the ion-pair is supplemented by a hydrogen bond between the nitrogen atom and O(12) of the carboxylate anion.

The C(11)-O(12) and C(11)-O(13) (1.252(7)$\overset{o}{A}$, 1.228(7)$\overset{o}{A}$; APPG and 1.274(4)$\overset{o}{A}$, 1.239(4)$\overset{o}{A}$; BPG) bond lengths in both compounds are markedly different; in the case of APPG the difference in length is 0.025(9)$\overset{o}{A}$ whilst for BPG it is 0.035(5)$\overset{o}{A}$. With APPG it might be possible to explain this difference due to O(12) being doubly hydrogen bonded whilst O(13) only participates in a single hydrogen bond. This, however, is not the case for BPG where both O(12) and O(13) participate in a single hydrogen bond. It is the case, however, that with both compounds it is O(12) which is hydrogen bonded to the counter-ion and which exhibits the most single bond character. Indeed in BPG this is an extremely short contact distance of only 2.662$\overset{o}{A}$. It seems likely therefore that in both APPG and BPG there is some polarisation of charge in the carboxylate anion due to the proximity of the positively charged nitrogen atom such that the dominant   form is:

$$-C(11)\begin{array}{c} \diagup \overset{-}{O}(12) \cdots\cdots H-\overset{+}{N} \\ \diagdown\!\!= O(13) \end{array}$$
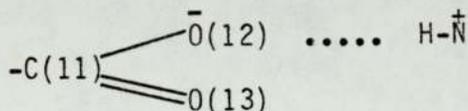
As stated earlier, APPG and BPG are both quite resistant to decomposition in aqueous environments, BPG being the more stable of the two. APPG is a monohydrate salt in which the water of crystallisation is extensively hydrogen bonded, supplemented by the strong hydrogen bond between the ion pair and the $N(14)-H(14)...O(13)$ interaction between adjacent penicillins. It is this extensive three dimensional hydrogen bonding in which the water molecules act almost as a 'molecular glue' which imparts this stability to aqueous environments.

BPG, unlike APPG, does not contain water of crystallisation and yet is even more stable. Like APPG it too forms a hydrogen bond between the two charged groups but this time substantially shorter and stronger. Further hydrogen bonding between adjacent ion-pairs holds the structure together in the direction of the a and b axes. The remainder of the complex, largely in the direction of the c axis consists mainly of hydrophobic areas, namely the three terminal phenyl groups which are inhospitable to solvation by water thus resulting in the compound's stability to aqueous environments.

In both cases it is the stability of the two drugs to aqueous decomposition which renders them useful as repositories for the slow release of the antibacterial agent benzylpenicillin.

Various solvents have been used in the preparation of APPG with a range of polarities. Buckwalter and Dickinson (1958)[11] observed different blood levels from samples of similar particle size but crystallised from different media. Table 3.1-1 shows the range of solvents with their differing polarities together with the crystal size and habit. Macek (1975)[13] reports from x-ray diffraction studies that prisms and plates have the same crystal form, whilst this work shows that prisms of different sizes crystallised from two

solvents of differing polarity have the same detailed internal structure. It is almost certain, therefore, that the differing bioavailabilities are an effect of differing crystal habit; however, an additional detailed crystallographic study on either the laths obtained from propanol, or the plates from water and colloid would finally verify this.

CHAPTER 4

CHAPTER 4

4.1.1   The importance of Folic Acid

Folic acid - vitamin Bc - (Figure 4.1) is an essential factor to mammals and micro-organisms alike. Its importance is embodied in the further conversion of its reduced form FAH_4 (5,6,7,8 tetrahydrofolic acid) to a number of one-carbon unit carrying co-factors which are then involved, for example, in amino acid interconversions (Figure 4.3) and purine and pyrimidine synthesis (Figure 4.2), the purine and pyrimidine bases being essential precursors for DNA biosynthesis.

In mammals, it is essential for folic acid to be ingested in the diet (approximately 50 μg/day for an adult human) whereas in certain micro-organisms, eg. bacterial cells, folic acid is synthesised from its basic subunits namely a glutamoylpteridine and p-aminobenzoic acid (PABA). Herein lies an important difference, in that mammalian cells must ingest folic acid whilst bacterial cells synthesise it; indeed, certain plasmodia and pathogenic bacteria synthesise folate containing cofactors de novo[27] from the folic acid constituents. Since folic acid is so important in the growth of normal cells and the proliferation of abnormal cells, it is an obvious important target for chemotherapy.

4.1.2   The metabolism of Folic Acid

Initially, folic acid is reduced in the cell enzymatically to 5,6,7,8 tetrahydrofolic acid (FAH_4) via 7,8 dihydrofolic acid by the enzyme dihydrofolate reductase (DHFR)[28] (Figure 4.1), the reaction being driven by the conversion of NADPH to $NADP^+$[29]. FAH_4 then picks up a one-carbon group which may be a formyl, methyl or hydroxymethyl and is converted into the one-carbon unit carrying cofactors mentioned earlier. The one-carbon unit can then be transferred to other compounds which are then utilised in various essential pathways, eg. a one-carbon

Figure 4.1   The Reduction of Folic Acid
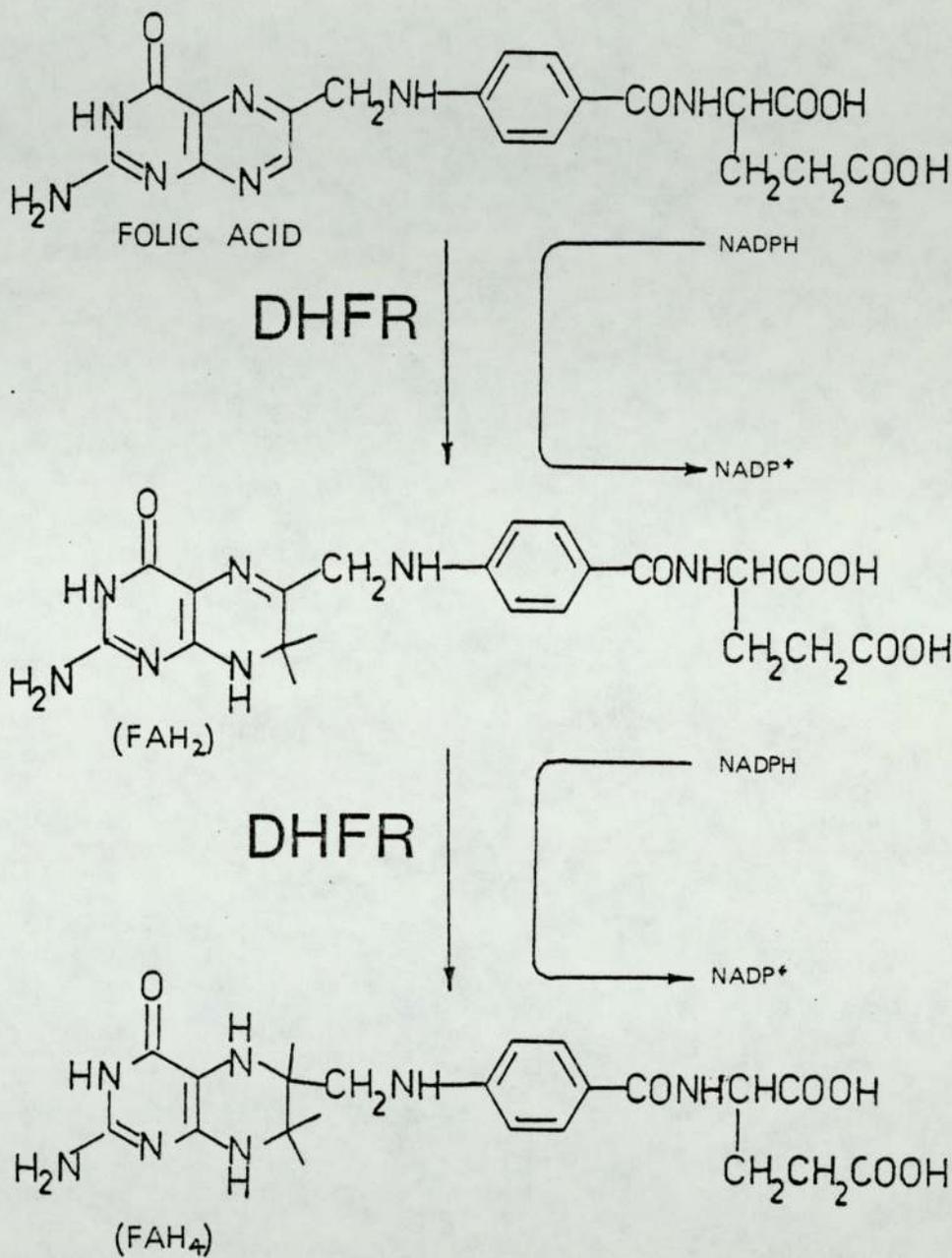
unit is transferred into C(2) and C(8) positions of the purine bases, the C(5) of deoxyuridylic acid (Fig 4.2) and the methyl group of thyamine.

Figure 4.2   The conversion of dUMP to dTMP
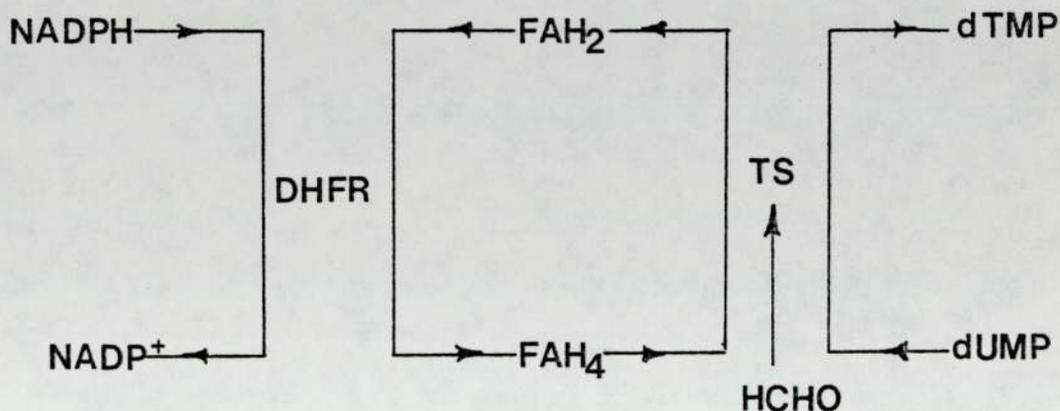
In the cycle shown in Figure 4.2, deoxyuridylic acid (dUMP) is converted to deoxythymidylic acid (dTMP) by the enzyme thymidylate synthetase (TS). During this cycle the continuous regeneration of FAH_4 by DHFR is required, and 'thymine less' death of cells has been reported where dihydrofolate reductase has been irreversibly inhibited.

Similarly, the interconversion of glycine and serine is folate dependent.[30]

Figure 4.3   The conversion of glycine to L-serine.

$$5,10\text{-methylene-}FAH_4 + Gly + H_2O \rightleftharpoons L\text{-serine} + FAH_4$$
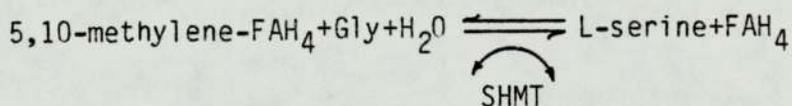
$$\overset{\frown}{SHMT}$$

SHMT   =   serine — hydroxymethyltransferase

(Blakely,   1954)[31]

Clearly, the importance of the folate cofactors in cell metabolism makes them an important target for chemotherapy where control of rapidly dividing cells is important, indeed, the indirect inhibition of the enzyme thymidilate synthetase (ie. by preventing the production of FAH_4-coenzymes) is a primary mechanism by which antifolate drugs antagonise DNA synthesis (Jukes & Broquist)[32].

4.1.3 The sulphonamides (inhibition of folic acid synthesis in bacterial cells)

As stated earlier, in bacterial cells folic acid is synthesised from its basic subunits, namely a glutamoylpteridine and PABA, this being achieved by a condensing enzyme. The use of the sulphonamides as antibacterial agents is related to their close structural similarity to PABA (p-aminobenzoic acid) these combining competitively with the condensing enzyme and inhibiting the normal entry of PABA into folic acid. The ultimate consequence is a folic acid deficiency manifested as a reversible inhibition of bacterial growth[33]. Whilst it is considered in part that the action of the sulponamides is in blocking the condensing enzyme, it is also accepted that they may participate as substrates with the consequent formation of false folic acids within the bacterium[34]. An action of this kind might be more effective than straight competitive antagonism in blocking bacterial metabolism, since it would be less readily reversible.

The simplest member of the sulphonamide series is sulphanilamide (Figure 4.32) in which both hydrogen atoms on the amide are unsubstituted; the other sulphonamides carry substituents. Just as a free p-amino group is essential for substrate activity in PABA so in a sulphonamide drug, a free p-amino group is essential for antibacterial action. In contrast, however, there are almost no

limitations on the possible substituents on the amide nitrogen compatible with antibacterial efficacy. It may be deduced that whereas the p-amino group combines with the enzyme active site, the amide nitrogen with its substituent probably does not.

A plausible explanation for the role of the substituent is suggested by the relationship between the $-SO_2-$ group in sulphonamides and the $-CO_2-$ group of PABA. The latter is completely ionised at neutral pH, and there is evidence that a negative charge here promotes combination with the enzyme. Thus it may be inferred that sulphonamides should become more potent with increasing electron density in the $-SO_2-$ region; this property should be influenced by electron withdrawing substituents at the amide nitrogen, an effect which can be measured by the pKa of each compound.

Bell and Roblin[35] tested a great many sulphonamides for antibacterial potency versus acid strength and confirmed that at pH7 maximal potency was found with those compounds with a pKa of 6 which related to about 91% ionisation. The 'tail off' in potency with greater acidities than pKa 6 can probably be explained as follows: when complete ionisation has already been attained, a still more strongly electrophilic substituent will begin to draw electrons away from the sulphone group. Table 4.1 shows a range of therapeutic sulphonamides with different acid strengths.

Antagonists of Dihydrofolate Reductase

The structure of folic acid is shown in Figure 4.1.

(I)   (II)

$$H_2N - \underset{4}{C} ═══ \underset{3}{N} ════ \underset{2}{C}(NH_2) ════ \underset{1}{N} \qquad\qquad O ═══ \underset{4}{C} - \underset{3}{N}(H) ════ \underset{2}{C}(NH_2) ════ \underset{1}{N}$$
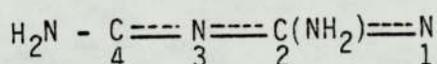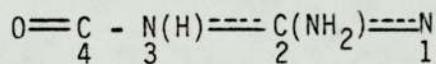
Simple substitution at the 4-position of folate (II) by an amino group to that of aminopterin (1) enhances the binding to DHFR by several orders of magnitude[36] preventing the reduction of FAH_2 to

FAH_4 and thus for example blocking the incorporation of one carbon units into thymine, a requirement for DNA synthesis.

Methylation of N(10) yields amethopterin (methotrexate, BP - - MTX); these compounds are potent inhibitors of DHFR whatever the source and attack all rapidly dividing cells. Methotrexate (Fig 4.4a) has been used in the treatment of leukemia (Li, et al, 1958)[37] and in combination therapies with other agents has been found useful in the treatment of other carcinomas (Goth, 1972)[38]. Unfortunately, the non-specificity of this and related compounds to malignant cells has caused considerable problems with respect to undesirable side effects. Attention of late, however, has moved away from the classical inhibitors of DHFR (minor modifications of folic acid) to compounds which still contain the 2-amino-4-oxo (substrate) or 2,4-diamino moeity (antagonist) but with a modified ring system eg. derivatives of triazines, quinazolines, or more importantly to this work, pyrimidines.

In 1948, it was found that many 2,4-diaminopyrimidines were powerful antagonists of folic acid in cultures of Lactobacillus Casei[39]. A formal analogy betwen some of these structures and proguanil (Fig 4.4b) and the discovery that proguanil was also an antagonist of folic acid suggested that the pyrimidine compounds might have antimalarial properties[40]; this was in fact shown to be the case[41]. One such compound which belongs to this class of folate antagonists is pyrimethamine (Fig 4.4b). Although this compound is not particularly active against the reductase from bacteria, it has an exceptional affinity for the enzyme from Plasmodium Vinckei which may well explain its specific antimalarial action, assuming similar activity with other plasmodia.

Table 4.1 Some sulphonamides with different acid strengths. (Bell and Roblin[35])

Compound	*R	pKa
Sulphanilamide	-H	10.43
Sulphapyridine		8.43
Sulphathiazole		7.12
Sulphadiazine		6.48
Sulphacetamide	$-\overset{\overset{\displaystyle O}{\|\|}}{C}-CH_3$	5.38
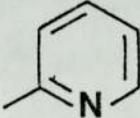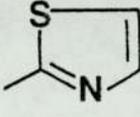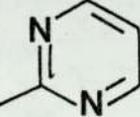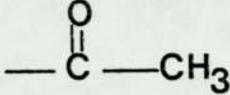
*
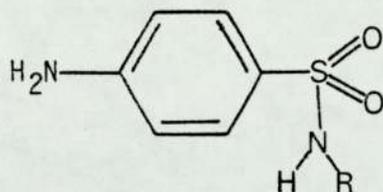
Figure 4.4a   The structure of two classical 2,4 diamino analogues of folic acid
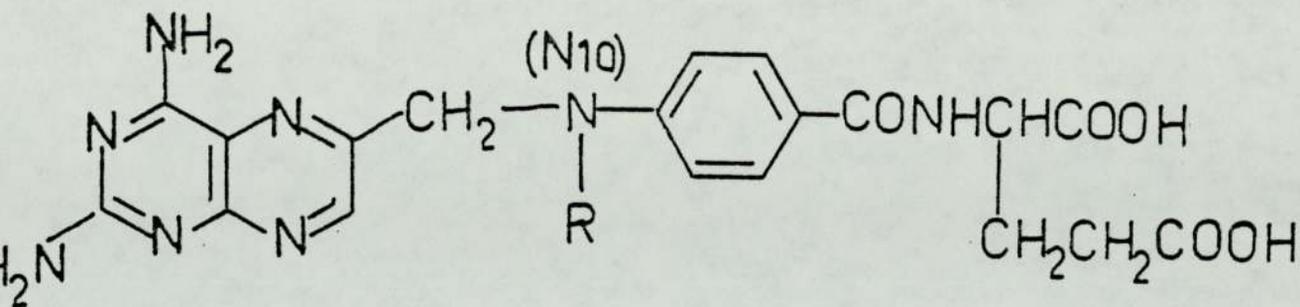
Aminopterin   R = -H

Methotrexate   R = -CH$_3$

Figure 4.4b   The structure of two non-classical inhibitors of DHFR

Proguanil

Cycloguanil

Pyrimethamine
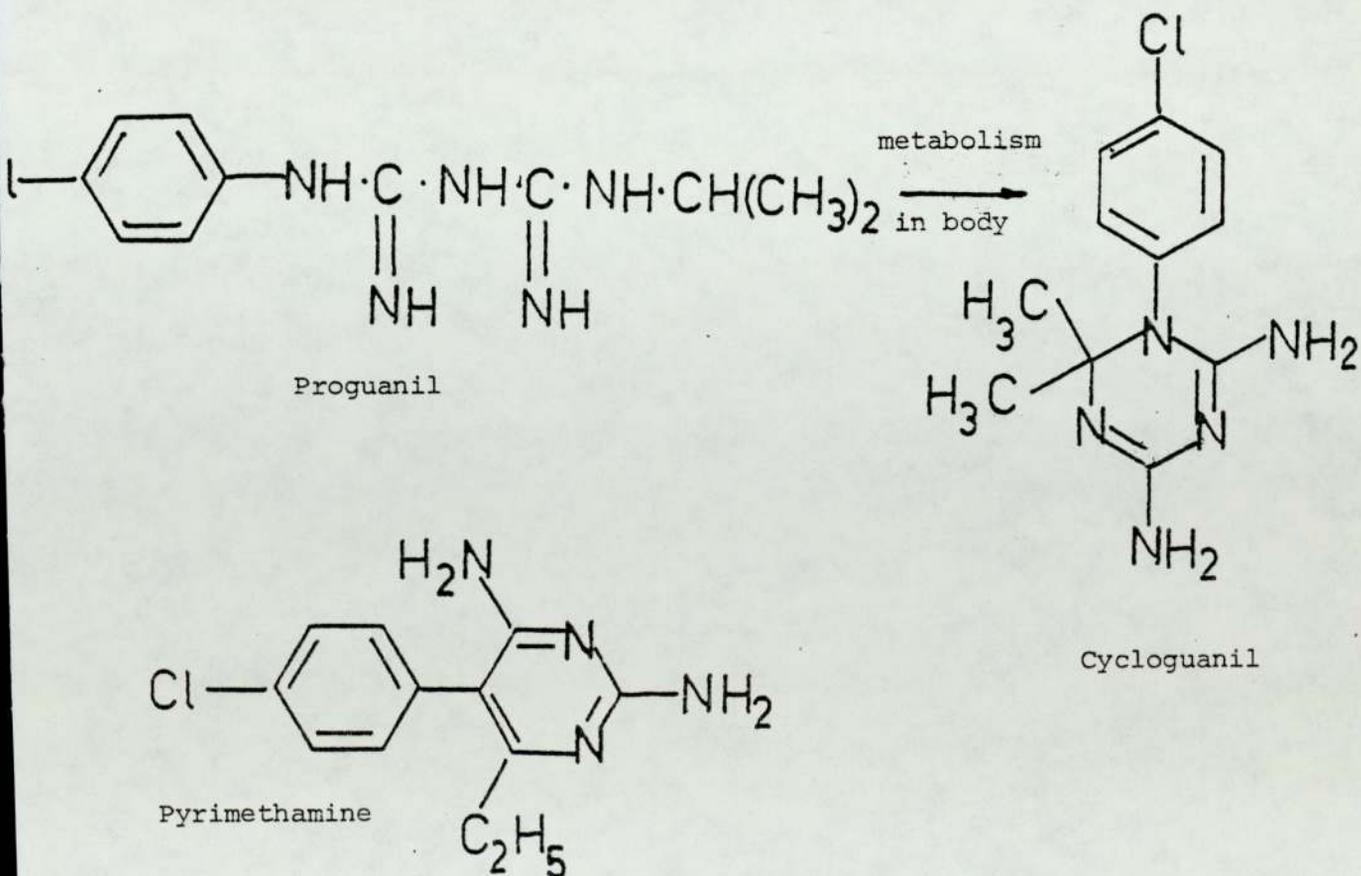
The observation that the antimalarials are not necessarily good antibacterial agents goes some way to showing the specificity of DHFR to the particular organism, such specificity being an obvious target for drug design. In an attempt to obtain a better understanding of the mode of action of the antifolates, particularly with respect to the binding of the drug to the enzyme, the structure of a number of these compounds has been determined crystallographically[42,43]. The two compounds reported in this work both bear structural similarities to that of pyrimethamine and indeed much of the work performed by Schwalbe et al (Aston) and Cody et al (Buffalo, New York) highlights the properties and common features of the diaminopyrimidine moeity.

4.1.5   Dihydrofolate Reductase Binding

The natural substrates of DHFR contain a 2-amino-4-oxo-pteridine fragment whereas the most effective inhibitors are the 2,4-diamino derivatives of pyrimidines, triazines, pteridines and quinazolines. Explanations for the enhanced affinity of the folate inhibitors have focussed on the modified pattern of hydrogen bond donors and acceptors[44] or the change in electron density in the heterocycle leading to increased basicity[45]. In the antifolates, both amino groups can act as hydrogen bond donors whilst $N(1)$ and $N(3)$ can act as hydrogen bond acceptors or be protonated. These patterns are in contrast to the natural substrates in which only the 2-amino group can be a proton donor, where $N(3)$ has a proton and where the enzymatic protonation site is $N(5)$, although $N(1)$ and $N(8)$ can be protonated. In all of the crystallographic studies of DHFR inhibitors, it has been shown that when the molecule is protonated the hydrogen is attached at $N(1)$ or its equivalent $N(3)$ position in S-triazines.

Information on the importance of hydrogen bonding power and basicity to DHFR inhibitor binding is available from crystal structure determinations by Matthews et al of DHFR - methotrexate complexes[46,47] one with DHFR from E. coli, the other from Lactobacillus casei which was also complexed with NADPH. Both show considerable hydrogen bonding between enzyme and inhibitor. The enzymes supply a carboxylate anion near the protonated N(1) (hence the importance of protonation at N(1) and thus basicity), a proton acceptor for the 2-amino group and another proton acceptor near the 4-substituent. In addition the E. coli but not the L. casei appears to furnish a proton donor in the vicinity of the 4-substituent. If the 4-amino group is capable of rehybridisation it is possible for it to act as both proton donor and acceptor thus forming two hydrogen bonds to the E. coli enzyme compared with just one for the carbonyl oxygen of the substrate[48]. Acting purely as a proton donor, it can form one hydrogen bond to the L.casei DHFR as compared with none for the carbonyl oxygen of the substrate. It has been suggested that the pteridine ring of bound folates may be turned over from its orientation in bound methotrexate[47]; if so a ring nitrogen capable of accepting but not donating a proton could approach the carbonyl position, and the above reasoning would still hold[48].

4.1.6 Hydrogen bonding of crystalline antifolate compounds

Hydrogen bonding studies of the pyrimidines and related compounds show that base-paired dimerization about a centre or pseudo centre of symmetry in the crystal lattice using NH_2 as a proton donor, and a ring nitrogen as an acceptor is an almost universal phenomenon. Unsolvated neutral molecules[49] form two such base pairs each utilizing only one of the amine hydrogens. However, when oxygen functions are available from solvents[50] N-H ... O bonds are formed

usually at the expense of one base pair.

In general hydrogen bonds formed by the 2 and 4 amino groups show that in the free-base both tend to make the same type of interaction (base pair dimerization); however, when N(1) is protonated, the 2-amino groups tends to use both protons to hydrogen bond to solvent oxygens or counter ions whilst the 4-amino group maintains a preference to form the base pair with one proton and form a hydrogen bond with a solvent oxygen or counter ion with the other[42]. This effect is well exemplified by the 2,4 diaminopyrimidine structure determined in this work which has both solvent and counter ion, the details of which are discussed later.

4.1.7   The Rationale

The two antifolate structures determined in this work, one containing the 2,4-diamino-pyrimidine moeity, the other the 2-amino-4-oxo-pyrimidine were part of a study on modified pyrimethamine structures as inhibitors of DHFR with particular reference to the enzyme from protozoal sources[51]. It was recognised that whilst the diaminopyrimidine moeity is a requirement for good binding to DHFR some modification of the chlorophenyl unit is permissable. The modified systems investigated (Figure 4.7) incorporated an electron-donating group (v) an electron withdrawing group (vi) and both in an altered pattern with increased bulk (iv). The 2 and 4-oxo compounds with an unmodified chlorophenyl unit were also examined. This, together with the structure of the sulphonamide covers the following areas with respect to antifolates:-

1) an inhibitor of DHFR with the classical 2,4-diaminopyrimidine
    moiety

2) an inhibitor of DHFR but with the 2-amino-4-oxo pyrimidine
    moiety - a requirement for the substrate of DHFR

3) a sulphonamide compound - sulphonamides being known inhibitors
    of the producton of folic acid in bacteria.

4.2 The crystal structure of 2,4-diamino-6-ethyl-5-(4-piperidino-3-nitrophenyl)-pyrimidine hydrochloride monohydrate (Nitropiperidino-pyrimethamine HCl; NPPHCl)

4.2.1 <u>Experimental</u> Preparation of NPPHCl [52]

a) Preparation of Nitropyrimethamine

pyrimethamine   nitropyrimethamine
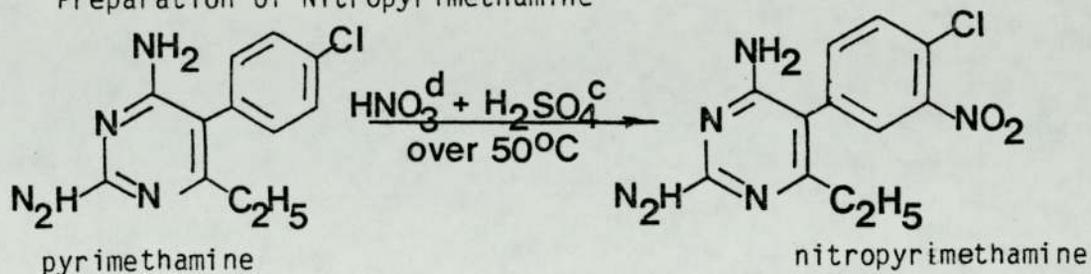
b) The addition of Piperidine

Nitropyrimethamine is boiled with excess piperidine for a period of 1 hour. On cooling, after a period of about 5 minutes, red crystals of NPP free base are deposited. The yield is about 95%. The free base of NPP can be crystallised from ethanol.

c) The preparation of NPP hydrochloride

The free base of NPP is warmed in an ethanol/conc. hydrochloric acid mixture to give a straw coloured solution. From this solution a cream crystalline product is deposited; this is probably the dihydrochloride salt of NPP. On recrystallising this product from water, thick red needles of nitropiperidinopyrimethamine monohydrochloride are produced.
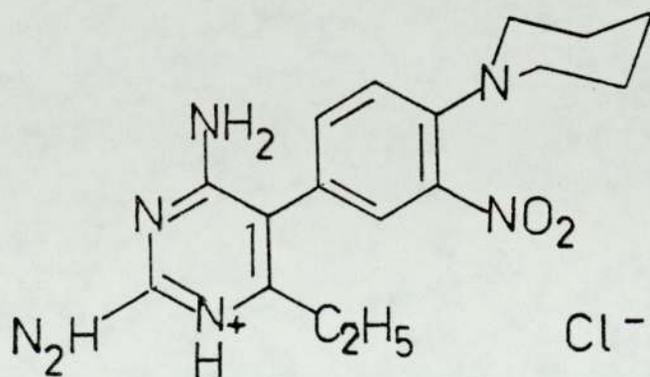
NPP hydrochloride

4.2.2   Crystal Data

2,4 -diamino-6-ethyl-5-(4-piperidino-3-nitrophenyl)-pyrimidine hydrochloride monohydrate ($C_{17}H_{23}N_6O_2ClH_2O$) crystallises in the monoclinic space group C2/c with a=18.213(7), b=12.385(3), c=19.179(9)$\overset{o}{A}$, β=116.41(4)°, z=8 and v=3874(1)$\overset{o}{A}{}^3$.

The molecular weight is 396.9 (F(ooo)=1680) and the density Dm (by flotation in butan-1-ol and carbon tetrachloride) = 1.36 gcm^{-1} with the calculated density Dc=1.35 gcm^{-3}. The absorption coefficient for Mo-Kα radiation (λ = 0.710690$\overset{o}{A}$) is μ=1.85 cm^{-1}.

4.2.3   Structure Analysis

Data Collection

The data were collected from a crystal of dimensions 1.25 x 0.4 x 0.26 mm mounted along its needle axis on an Enraf-Nonius CAD4 diffractometer using the ω -2Θ scan technique with graphite monochromated Mo-Kα radiation (λ =0.71069$\overset{o}{A}$). The ω scan width = 1.4 + 0.35 tan Θ and ω scan speed 1.11 - 6.7° min^{-1}. Two reference reflections were monitored every 5400 sec of x-ray exposure time and the orientation of the crystal checked after every 150 reflections. A loss of intensity of about 3% was observed over the whole data collection on the reference reflections, and this was corrected for during data reduction. No movement of the crystal was observed during the collection of the 6636 data between Θ =2 and Θ =24° giving 3068 unique observed reflections for which the merged R=0.05.

The unit cell parameters determined from 25 automatically centred reflections are as given above under 'Crystal Data'.

Structure Determination and Refinement

The data reduction was performed at Birmingham University on the DEC20 computer using a programme provided by Dr T A Hamor. Lorentz and polarisation and intensity corrections were applied, and the data

output in the form HKL, Fobs and σ Fobs, where σ Fobs is based on counting statistics and includes a correction for instrument instability. No correction for absorption was made to the data, since for Mo-Kα radiation and the crystal size in question the maximum correction which would need to be applied to intensity would be $\pm$ 8%.

Normalised structure factors were calculated using a modified Wilson plot method[22]. The curve from the Wilson calculation, and the least squares fit are shown in Figure 4.5 with the values calculated from the gradient and the intercept being:-

Temperature factor B = 2.7248

Scale (K) = 3.3382

The distribution of E's for the complete data set and the $\left| E^2 - 1 \right|$ statistic closely follow that of a centric structure. Phase determination was undertaken by Multan[1] in which the 208 largest E's all with $E \geqslant 2.22$ produced 1335 unique $\sum_2$ relationships for use in phase refinement. On the basis of a $\sum_1$ calculation, the following invariant phases were determined with a probability greater than 95%:-

    0 10 0   (-)

    2 2 10 (+)

The 'Converge' procedure produced the following origin determining reflections:- 17 3 2 (+); 5 5 11 (+). Eight starting sets were produced by permuting the signs of the 17 3 1, 16 4 -3 and 19 1 -6 reflections; the set of signs (+) (+) (-) produced an E-map in which 24 of the 27 non-hydrogen atoms could be located. This solution had the highest combined figure of merit.

Figure 4.5   Modified Wilson Plot
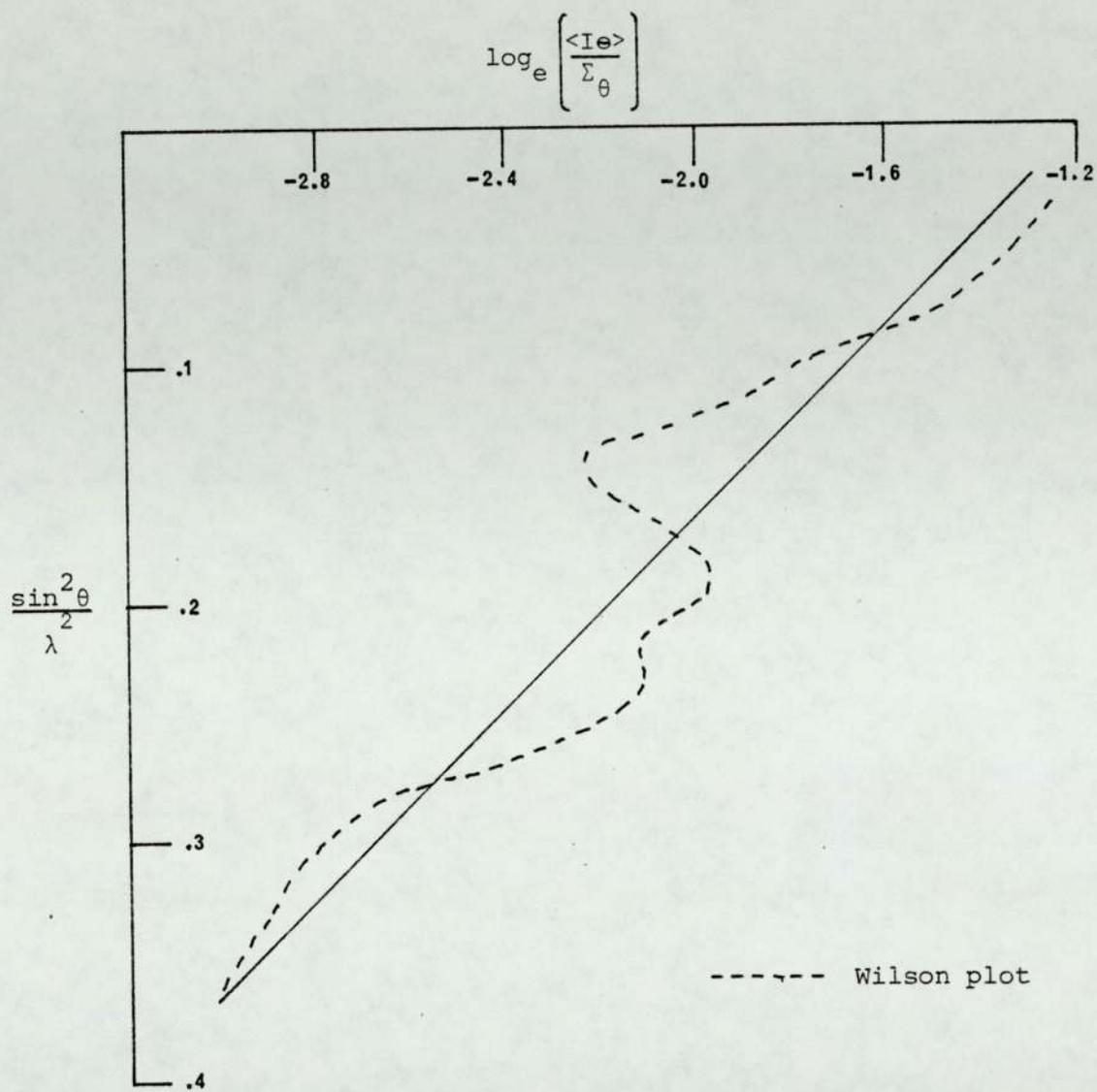
A subsequent Fourier refinement and Fourier map located the positions of the 3 remaining non-hydrogen atoms.

Successive full matrix least squares refinements of positional parameters and isotropic temperature factors for all non-hydrogen atoms using unit weights, during which atoms were assigned their correct scattering factors on the basis of conventional analytical techniques, reduced R, the unweighted discrepancy index to 0.12. At this point, all hydrogen atoms were located on a difference electron density map.

Final anisotropic refinement of the complete structure with isotropic temperature factors for hydrogen atoms, during which all hydrogen atoms were refined freely without bond length constraints, reduced the final agreement parameters to R=0.058 and Rg=0.082. During this stage of the refinement, the hydrogen atoms were 'blocked', each block being refined on alternate cycles. The reflections were weighted according to $W=k/[\sigma^2(Fo)+qFo^2]$ where k and q (both described earlier) converged at 1.0000 and 0.0153 respectively. This stage of the refinement involved 263 parameters and 2524 unique data for which $Fo > 3\sigma(Fo)$. When the refinement was finally terminated, a difference electron density map showed no feature greater than $0.40e\overset{o}{A}^{-3}$, this being a peak associated with the chlorine atom, and the maximum shift associated with a positional parameter was 0.47σ for the hydrogen atoms, and 0.1σ for non-hydrogen atoms. The final temperature factors and their values throughout the refinement confirmed the correct assignment of atomic scattering factors during the early stages of refinement.

Table 4.2 Positional parameters (fractional co-ordinates x 10^4) with
estimated standard deviations in parentheses.

Atom	X/a	Y/b	Z/c
Cl(1)	7298(1)	4412(1)	4662(1)
O(1)	6382(1)	-290(2)	5747(1)
N(1)	5544(1)	1670(2)	5569(1)
C(2)	5714(1)	2567(2)	5272(1)
N(3)	5311(1)	3479(1)	5208(1)
C(4)	4712(1)	3487(2)	5440(1)
C(5)	4527(1)	2572(2)	5782(1)
C(6)	4971(1)	1660(2)	5844(1)
N(7)	6290(1)	2515(2)	5030(1)
N(8)	4285(1)	4392(2)	5319(1)
C(9)	4956(1)	652(2)	6259(1)
C(10)	5669(2)	643(2)	7076(2)
C(1')	3904(1)	2627(2)	6084(1)
C(2')	3977(1)	3379(2)	6643(1)
C(3')	8401(1)	3435(2)	6920(1)
C(4')	2731(1)	2718(2)	6685(1)
C(5')	2672(1)	1972(2)	6123(1)
C(6')	3227(1)	1936(2)	5815(1)
N(7')	3490(1)	4334(2)	7440(1)
O(8')	4186(1)	4614(2)	7901(1)
O(9')	2873(1)	4805(2)	7381(1)
N(1")	2169(1)	2767(2)	6998(1)
C(2")	2509(1)	2646(2)	7851(1)
C(3")	1914(2)	3075(3)	8126(1)
C(4")	1101(2)	2497(3)	7723(1)
C(5")	776(2)	2548(3)	6849(2)

Table **4.2** continued

Atom	X/a	Y/b	Z/c
C(6")	1403(1)	2161(2)	6585(2)
H(1)	5801(19)	1100(27)	5591(18)
H(2)	6628(22)	1933(29)	5088(20)
H(3)	6469(16)	3069(23)	4888(16)
H(4)	4458(15)	4972(23)	5175(15)
H(5)	3916(20)	4476(23)	5454(19)
H(6)	4466(21)	607(24)	6264(19)
H(7)	5044(15)	31(24)	5986(15)
H(8)	5638(23)	1222(33)	7304(23)
H(9)	6215(21)	668(25)	7081(19)
H(10)	6562(18)	4(26)	7312(18)
H(11)	4426(15)	3918(20)	6826(14)
H(12)	2292(18)	1475(22)	5986(17)
H(13)	3161(14)	1396(21)	5395(15)
H(14)	3054(19)	3084(25)	8079(18)
H(15)	2593(16)	1939(25)	7970(17)
H(16)	1877(16)	3841(25)	8062(15)
H(17)	2160(24)	2897(33)	8685(25)
H(18)	771(20)	2930(29)	7990(20)
H(19)	1168(22)	1779(31)	7913(23)
H(20)	668(16)	3331(23)	6679(16)
H(21)	283(19)	2120(26)	6580(18)
H(22)	1441(19)	1444(28)	6676(19)
H(23)	1213(17)	2388(21)	6045(17)
H(24)	6731(28)	-133(36)	5662(26)
H(25)	6605(25)	-580(32)	6148(26)

Table **4.3**   Anisotropic temperature factors (non-hydrogen atoms).

Isotropic temperature factors (hydrogen atoms) .

With standard deviations in parentheses.

ATOM	U11	U22	U33	U23	U13	U12
Cl(1)	.0459(5)	.0453(5)	.0775(6)	.0112(3)	.0379(4)	.0080(3)
O(1)	.0363(11)	.0512(12)	.0515(14)	.0092(10)	.0200(10)	.0108(9)
N(1)	.0250(10)	.0317(10)	.0268(10)	.0025(8)	.0094(8)	.0047(8)
C(2)	.0194(11)	.0367(13)	.0203(12)	.0022(9)	.0029(9)	.0024(9)
N(3)	.0227(9)	.0322(10)	.0286(10)	.0031(8)	.0092(8)	.0016(8)
C(4)	.0192(10)	.0340(12)	.0190(11)	.0006(9)	.0033(9)	.0000(9)
C(5)	.0216(11)	.0352(12)	.0200(11)	.0007(9)	.0051(9)	-.0004(9)
C(6)	.0203(10)	.0320(12)	.0215(11)	-.0028(9)	.0043(9)	-.0014(9)
N(7)	.0317(11)	.0415(12)	.0426(13)	.0100(9)	.0209(10)	.0066(9)
N(8)	.0306(11)	.0313(11)	.0456(13)	.0099(9)	.0217(10)	.0076(8)
C(9)	.0310(13)	.0307(12)	.0365(13)	.0049(10)	.0140(11)	.0022(9)
C(10)	.0470(16)	.0509(16)	.0342(14)	.0117(12)	.0148(13)	.0061(12)
C(1')	.0208(11)	.0318(12)	.0228(12)	.0073(9)	.0047(9)	.0041(9)
C(2')	.0212(11)	.0339(13)	.0257(12)	.0029(10)	.0053(9)	-.0001(9)
C(3')	.0232(11)	.0322(12)	.0212(11)	-.0011(9)	.0047(9)	.0028(9)
C(4')	.0179(10)	.0334(12)	.0249(12)	.0050(9)	.0049(9)	.0040(9)
C(5')	.0208(11)	.0338(12)	.0280(13)	-.0019(9)	.0057(10)	-.0039(9)
C(6')	.0238(11)	.0336(12)	.0247(12)	-.0028(9)	.0060(10)	.0000(9)
N(7')	.0294(11)	.0373(11)	.0234(11)	-.0055(9)	.0090(9)	-.0010(9)
O(8')	.0295(10)	.0734(14)	.0569(14)	-.0342(11)	.0024(10)	-.0031(9)
O(9')	.0343(10)	.0473(11)	.0633(13)	-.0162(9)	.0226(9)	.0026(8)
N(1")	.0187(9)	.0455(11)	.0205(10)	.0001(8)	.0054(8)	-.0002(8)
C(2")	.0256(12)	.0487(17)	.0240(13)	.0045(11)	.0053(10)	.0034(11)

Table 4.3 continued

ATOM	U11	U22	U33	U23	U13	U12
C(3")	.0341(14)	.0709(21)	.0298(15)	.0041(13)	.0121(12)	.0078(13)
C(4")	.0340(15)	.0898(26)	.0424(17)	.0071(17)	.0208(13)	.0047(15)
C(5")	.0203(12)	.0653(20)	.0409(16)	-.0019(14)	.0088(11)	-.0017(!4)
C(6")	.0216(12)	.0441(15)	.0351(15)	.0011(12)	.0075(11)	-.0024(11)

	Uiso
H(1)	.0475(84)
H(2)	.0599(94)
H(3)	.0297(68)
H(4)	.0267(61)
H(5)	.0449(84)
H(6)	.0504(90)
H(7)	.0309(66)
H(8)	.0731(116)
H(9)	.0528(86)
H(10)	.0461(74)
H(11)	.0301(62)
H(12)	.0427(74)
H(13)	.0306(62)
H(14)	.0520(82)
H(15)	.0356(70)
H(16)	.0393(71)
H(17)	.0763(110)
H(18)	.0632(94)
H(19)	.0647(106)
H(20)	.0376(71)

Table 4.3 continued

H(21)   .0486(81)

H(22)   .0533(89)

H(23)   .0351(68)

H(24)   .0756(145)

H(25)   .0643(130)

Table 4.3 continued

Table 4.4 Bond distances in $\overset{o}{A}$ with estimated standard deviations in parentheses.

Bond	Interatomic distance $\overset{o}{A}$	Bond	Interatomic distance $\overset{o}{A}$
N(1)-C(2)	1.345(3)	N(1)-H(1)	.84(4)
C(2)-N(7)	1.324(4)	N(7)-H(2)	.92(4)
C(2)-N(3)	1.322(3)	N(7)-H(3)	.86(3)
N(3)-C(4)	1.349(4)	N(8)-H(4)	.88(3)
C(4)-N(8)	1.325(3)	N(8)-H(5)	.83(4)
C(4)-C(5)	1.423(3)	C(9)-H(6)	.94(4)
C(5)-C(1')	1.488(4)	C(9)-H(7)	.98(3)
C(5)-C(6)	1.362(3)	C(10)-H(8)	.86(4)
C(6)-N(1)	1.364(4)	C(10)-H(9)	.99(4)
C(6)-C(9)	1.487(3)	C(10)-H(10)	.97(4)
C(9)-C(10)	1.527(3)	C(2')-H(11)	.99(3)
C(1')-C(2')	1.380(3)	C(5')-H(12)	.87(3)
C(2)'-C(3')	1.375(4)	C(6')-H(13)	1.01(3)
C(3')-N(7')	1.459(3)	C(2")-H(14)	1.04(3)
C(3')-C(4')	1.409(3)	C(2")-H(15)	.90(3)
C(4')-C(5')	1.388(4)	C(3")-H(16)	.96(3)
C(5')-C(6')	1.379(4)	C(3")-H(17)	.99(4)
C(6')-C(1')	1.398(3)	C(4")-H(18)	1.09(4)
N(7')-O(8')	1.229(2)	C(4")-H(19)	.95(4)
N(7')-O(9')	1.223(3)	C(5")-H(20)	1.01(3)
C(4')-N(1")	1.399(4)	C(5")-H(21)	.97(3)
N(1")-C(2")	1.478(3)	C(6")-H(22)	.90(4)
C(2")-C(3")	1.497(5)	C(6")-H(23)	.98(3)
C(3")-C(4")	1.512(4)	O(1)-H(24)	.75(6)
C(4")-C(5")	1.510(4)	O(1)-H(25)	.78(4)
C(5")-C(6")	1.518(5)		
C(6")-N(1")	1.468(3)		

Table 4.5   Interatomic angles [°] with estimted standard deviations in
             parentheses (for non-hydrogen atoms and those hydrogens
             attached to primary amine groups).

Atoms	Bond Angle [°]	Atoms	Bond Angle [°]
N(1)-C(2)-N(3)	121.5(3)	C(5')-C(6')-C(1')	121.1(2)
N(1)-C(2)-N(7)	118.4(2)	C(6')-C(1')-C(2')	117.7(3)
N(7)-C(2)-N(3)	120.1(2)	C(4')-N(1")-C(2")	116.4(2)
C(2)-N(3)-C(4)	118.2(2)	C(4')-N(1")-C(6")	116.7(2)
N(3)-C(4)-C(5)	122.5(2)	N(1")-C(2")-C(3")	110.0(2)
N(3)-C(4)-N(8)	116.6(2)	C(2")-C(3")-C(4")	110.4(3)
N(8)-C(4)-C(5)	120.9(2)	C(3")-C(4")-C(5")	111.0(3)
C(4)-C(5)-C(6)	116.6(2)	C(4")-C(5")-C(6")	112.0(2)
C(4)-C(5)-C(1')	121.4(2)	C(5")-C(6")-N(1")	109.5(2)
C(1')-C(5)-C(6)	122.0(2)	C(6")-N(1")-C(2")	112.0(2)
C(5)-C(6)-N(1)	119.0(2)	C(2)-N(7)-H(2)	126(3)
C(5)-C(6)-C(9)	114.5(2)	C(2)-N(7)-H(3)	123(2)
C(9)-C(6)-N(1)	126.2(3)	H(2)-N(7)-H(3)	110(3)
C(6)-C(9)-C(10)	110.3(2)	C(4)-N(8)-H(4)	119(2)
C(6)-N(1)-C(2)	122.1(2)	C(4)-N(8)-H(5)	123(2)
C(5)-C(1')-C(2')	120.4(2)	H(4)-N(8)-H(5)	117(3)
C(5)-C(1')-C(6')	121.9(2)		
C(1')-C(2')-C(3')	120.5(2)		
C(2')-C(3')-C(4')	123.1(2)		
C(2')-C(3')-C(4')	115.5(2)		
N(7')-C(3')-C(4')	121.3(2)		
O(8')-N(7')-O(9')	123.2(2)		
C(3')-N(7')-O(8')	118.1(2)		
C(3')-N(7')-O(9')	118.7(2)		
C(3')-C(4')-C(5')	115.1(3)		

Table 4.5 continued

C(3')-C(4')-N(1")   122.0(2)

N(1")-C(4')-C(5')   122.9(2)

C(4')-C(5')-C(6')   122.4(2)

Table 4.6 Deviations of non-hydrogen (in $\overset{o}{A}$) atoms from the least squares planes through (a) the pyrimidine ring, and (b) the phenyl ring and (c) the piperidine ring[+] (atoms used in the plane calculation are marked with an asterisk).

a)

ATOM	DEVIATION [$\overset{o}{A}$]	ATOM	DEVIATION [$\overset{o}{A}$]
*N(1)	-.020	*C(6)	.015
*C(2)	.004	N(7)	-.004
*N(3)	.017	N(8)	-.101
*C(4)	-.021	C(9)	.189
*C(5)	.005	C(10)	1.667

[+] The equation of the plane is:-

$0.680X + 0.282Y + 0.676Z + 12.002 = 0$ where X, Y and Z are orthogonal coordinates in $\overset{o}{A}$ along a*, b and c.

b)

ATOM	DEVIATION [$\overset{o}{A}$]	ATOM	DEVIATION [$\overset{o}{A}$]
*C(1')	-.012	*C(6')	.021
*C(2')	-.010	N(7')	.188
*C(3')	.023	O(8')	-.380
*C(4')	-.013	O(9')	.897
*C(5')	-.009	N(1")	-.024

[+] The equation of the plane is:-

$.0544X + 0.651Y - 0.529Z - 12.168 = 0$ where X, Y and Z are orthogonal coordinates in $\overset{o}{A}$ along a*, b and c

Table 4.6 continued

c)

ATOM	DEVIATION [$\overset{o}{A}$]	ATOM	DEVIATION [$\overset{o}{A}$]
*N(1")	.253	*C(4")	-.218
*C(2")	-.255	*C(5")	.213
*C(3")	.239	*C(6")	-.233

[+] The equation of the plane is:-

$-0.169X + 0.967Y - 0.190Z + 6.353 = 0$ where X, Y and Z are orthogonal coordinates in $\overset{o}{A}$ along a*, b and c

Table 4.7   Important Torsion Angles [°]

Atoms	Angle [°]
C(6)-N(1)-C(2)-N(7)	178.6
C(6)-N(1)-C(2)-N(3)	-2.4
N(1)-C(2)-N(3)-C(4)	-1.3
C(2)-N(3)-C(4)-N(8)	175.4
C(2)-N(3)-C(4)-C(5)	3.8
N(7)-C(2)-N(3)-C(4)	177.8
N(3)-C(4)-C(5)-C(6)	-2.7
N(3)-C(4)-C(5)-C(1')	175.4
N(8)-C(4)-C(5)-C(6)	176.5
N(8)-C(4)-C(5)-C(1')	-5.4
C(4)-C(5)-C(6)-N(1)	-.9
C(4)-C(5)-C(1')-C(2')	-57.3
C(4)-C(5)-C(1')-C(6')	122.3
C(4)-C(5)-C(6)-C(9)	173.0
C(5)-C(6)-N(1)-C(2)	3.4
C(5)-C(6)-C(9)-C(10)	-96.9
C(2')-C(3')-N(7')-O(8')	-37.2
C(2')-C(3')-N(7')-O(9')	141.0
C(4')-C(3')-N(7')-O(8')	146.6
C(4')-C(3')-N(7')-O(9')	-35.2

The torsion angle A-B-C-D is the projected angle between A-B and C-D when viewed down the B-C bond; the clockwise rotation of the C-D bond with reference to bond A-B is considered positive.

Table 4.8   Hydrogen bond contact distances and angles

Hydrogen Bond	Angle at H[°]	Donor-Acceptor Distance $\overset{o}{A}$
$N(1)_I - H(1) \ldots\ldots O(1)_I$	171.8	2.806
$N(7)_I - H(3) \ldots\ldots Cl(1)_I$	164.2	3.247
$N(7)_I - H(2) \ldots\ldots Cl(1)_{IV}$	172.3	3.364
$N(8)_I - H(4) \ldots\ldots N(3)_{II}$	169.3	3.030
$N(8)_I - H(5) \ldots\ldots Cl(1)_{II}$	146.7	3.255
$O(1)_{IV} - H(24) \ldots. Cl(1)_I$	169.9	3.042
$O(1)_{III} - H(25) \ldots O(9')_I$	134.3	3.104
$O(1)_{III} - H(25) \ldots N(1'')_I$	152.9	3.247

The subscripts I, II, III, IV refer to the equivalent positions:-
x, y, z; $\frac{1}{2}$x, $\frac{1}{2}$y, 1-z; -1+x, 1+y, z; 2-x, -y, 1-z

4.2.4   Results and Discussion

The structure of 2,4-diamino-6-ethyl-5-(4-piperdino-3-nitrophenyl) pyrimidine hydrochloride monohydrate (NPP hydrochloride) as determined crystallographically is depicted in Figure 4.8 with the numbering scheme used; positional parameters for all atoms are given in Table 4.2 with anisotropic thermal parameters for non-hydrogen atoms and isotropic temperature factors for hydrogen atoms in Table 4.3.

Details of bond distances ($\overset{O}{A}$) and interatomic angles (°) are in Tables 4.4 and 5 respectively.

Protonation in producing the hydrochloride salt occurs at N(1), this being confirmed by location of the attached hydrogen atom from a difference Fourier synthesis. Further evidence for protonation at N(1) as opposed to N(3) can be seen from the two nitrogen valence angles C(6)-N(1)-C(2) (122.1(2)° and C(2)-N(3)-C(4) (118.2 (2)°), which comply with Chatar Singh's Rule - for protonated and unprotonated nitrogen atoms respectively, or more fundamentally, the Valence Shell Electron Pair Repulsion Theory[53] (VSEPR).

Inspection of the C(2)-N(3) bond distance (1.322(3)$\overset{O}{A}$) shows it to have considerable double bond character, more so than any of the other structures depicted in Figure 4.7 and in particular is significantly shorter than the average value for this bond determined from a number of protonated 2,4 – diaminopyrimidine structures (Schwalbe & Cody)[42]. This suggests a considerable contribution to the structure from resonance form II in Figure 4.6. The N(1)-C(2) and N(3)-C(4) bonds (1.345(3)$\overset{O}{A}$ and 1.349(4)$\overset{O}{A}$) are almost identical in length both showing partial double bond character; this seems to be consistent throughout the 2,4 diamino pyrimidine structures reported in this section and compatible with resonance form I, Ia and II in

Figure 4.6. The C(4)-C(5) bond distance $(1.423(3)\overset{o}{A})$ and C(5)-C(6) $(1.362(3)\overset{o}{A})$ are variable throughout the range of compounds.

In the protonated compounds, C(5)-C(6) is shorter than in the free bases, whilst the reverse is the case of C(4)-C(5). The difference in these bond distances can again be rationalised by the canonical forms in Figure4.6, where, in all but Ia, the C(5)-C(6) bond has more double bond character than C(4)-C(5). In general, all of these compounds show delocalisation of electrons within the pyrimidine ring; however, in those protonated at N(1) a contribution from resonance form II seems to be important. The C(2)-N(7) $(1.324(4)\overset{o}{A})$ and C(4)-N(8) $(1.325(3)\overset{o}{A})$ exocyclic bonds are significantly shorter than those suggested by Schwalbe and Cody for protonated 2,4 diamino-pyrimidines from their values determined from a number of structures. The exocyclic $C-NH_2$ bond distances are considerably shorter than the ring C-N bonds. This is consistent with electron donation from N(7) and N(8), particularly contributions from canonical forms (II), (III), (IV) & (V). The sum of the angles about the two nitrogens N(7) and N(8), Table 4.5, are both 359(6)° suggesting Sp^2 hybridisation. Schwalbe et al[51] suggested that despite the large dihedral angle between the pyrimidine and phenyl rings (57°) withdrawal of electrons at the 3' position (eg. by the nitro group) appeared to elicit electron donation from the 2 and 4 amino groups; however, as demonstrated by NPPHCl (Fig 4.7) DDMP (VII) and pyrimethamine hydrochloride (II), by far the most substantial effect on the exocyclic C-N bond distances is that due to protonation at N(1). Such protonation at N(1) can therefore be seen to enhance the immonium character of N(7) and N(8) thus increasing the hydrogen bond donor potential. The bridging bond C(5)-C(1') $(1.488(4)\overset{o}{A})$ is slightly shorter than that normally associated with a $C(Sp^2)-C(Sp^2)$

bond[54], however, with such a large dihedral angle, very little orbital overlap would be expected; this distance is similar to that commonly found in this class of compounds. The moiety shows as expected[55] internal ring angles of $< 120°$ at the electron donating substituent (angle at C(4) = 115.1°) and $> 120°$ at the electron withdrawing substituent (angle at C(3') = 123.1°) with, however, there being little evidence of conjugative interaction between the substituents. The plane of the nitro group intersects the plane of the phenyl ring at 38.1°; such a dihedral angle would suggest little interaction between the two as is confirmed by the C(3')-N(7') distance (1.459(3)$\overset{o}{A}$). The C(4')-N(1") bond distance (1.399(4)$\overset{o}{A}$) suggests some electron donation from the piperidine nitrogen and the sum of the bond angles about this nitrogen (344(6)°) suggest a hybridisation intermediate between Sp^2 and Sp^3.

The bond distances in the phenyl ring all approach 1.38$\overset{o}{A}$ with the exception of the C(3')-C(4') (1.409(3)$\overset{o}{A}$). This bond has the nitro group attached at one end (C(3')) and the piperidine at the other (C(4')). This longer bond may reflect some electronic interaction of the groups with the ring, but more likely is purely physical, reflecting the steric effect of two bulky groups in a cis configuration. Deviations in $\overset{o}{A}$ from the least squares planes through the three ring systems and important torsion angles are given in Tables 4.6 and 7.

The pyrimidine ring is essentially coplanar with its 2,4 substituents, the N(8) substituent at C(4) showing the larger deviation possibly due to the hydrogen bonding constraints placed upon it. The phenyl ring, as expected, is also planar with the angle of intersection between it and the pyrimidine ring being 57°. Similarly, the intersecting angle between the plane through the

piperidine ring which is in the chair conformation and the phenyl ring is 34.7°, with the direction of rotation being opposite to that of the pyrimidine ring.

The molecular packing, as might be expected from such a 'bulky' molecule in a centred space group, is quite complicated as can be seen from the packing diagram in Figure 4.13. For clarity, Figures 4.10, 11 and 12 show the unit cell but with the centred molecules omitted. Because of lack of coplanarity of the three ring systems, stacking forces as such do not appear to be important. The hydrogen bonding scheme is depicted in Figure 4.15 with the contact distances, angles and relevant symmetry operations given in Table 4.8. The proposed scheme is similar to that found for other antifolate structures with a counter ion and solvent oxygens available as described by Schwalbe and Cody[42]. The main interaction is a base pair dimerization across a centre of symmetry between N(8) and N(3) of the one molecule and the corresponding pair on the centrosymmetrically related one (N(8)-H(4) ... N(3) - 3.030Å). The chloride counter ion is fully hydrogen bonded with each chloride ion interacting with the two centrosymmetrically related molecules via N(7)-H(3) ... Cl(1) (3.247Å) and N(8)-H(5) ... Cl(1) (3.255Å) contacts. Further hydrogen bonding with the counter ion occurs with the remaining hydrogen atom on N(7) (N(7)-H(2) ... Cl(1) - 3.364Å) and with the solvent oxygen O(1) (O(1)-H(24) ... Cl(1) - 3.042Å). The remaining interactions involve the solvent oxygen acting as a hydrogen bond receptor via an N(1)-H(1) ... O(1) contact (2.806), thus mimicking the N(1)-carboxylate anion hydrogen bond found in the enzyme-inhibitor complex, and as a proton donor in what appears to be a bifurcated hydrogen bond with the nitro oxygen O(9') (O(1)-H(25) ... O(9') - 3.104Å) and the piperidine nitrogen N(1") (O(1)-H(25) ... N(1") - 3.247Å).

Thus with the exception of O(8'), the remaining nitro oxygen, all possible sites of hydrogen bond interactions, either donor or acceptor, participate maximally in producing eight intermolecular interactions. Analysis of the types of hydrogen bonds formed by N(7) and N(8) in this structure are consistent with the results found by Schwalbe and Cody[42]; whereas, in the free bases both N(7) and N(8) tend to be involved in base-pair dimerization, when N(1) is protonated N(7) tends to use both protons to hydrogen bond to solvent oxygens or counter ions, whilst N(8) maintains a preference to form a base-pair dimer with the one proton and a solvent oxygen/counter ion interaction with the other.

Figure 4.6   Possible canonical forms for NPP monohydrochloride monohydrate

(R = 4-piperidino-3-nitrophenyl)
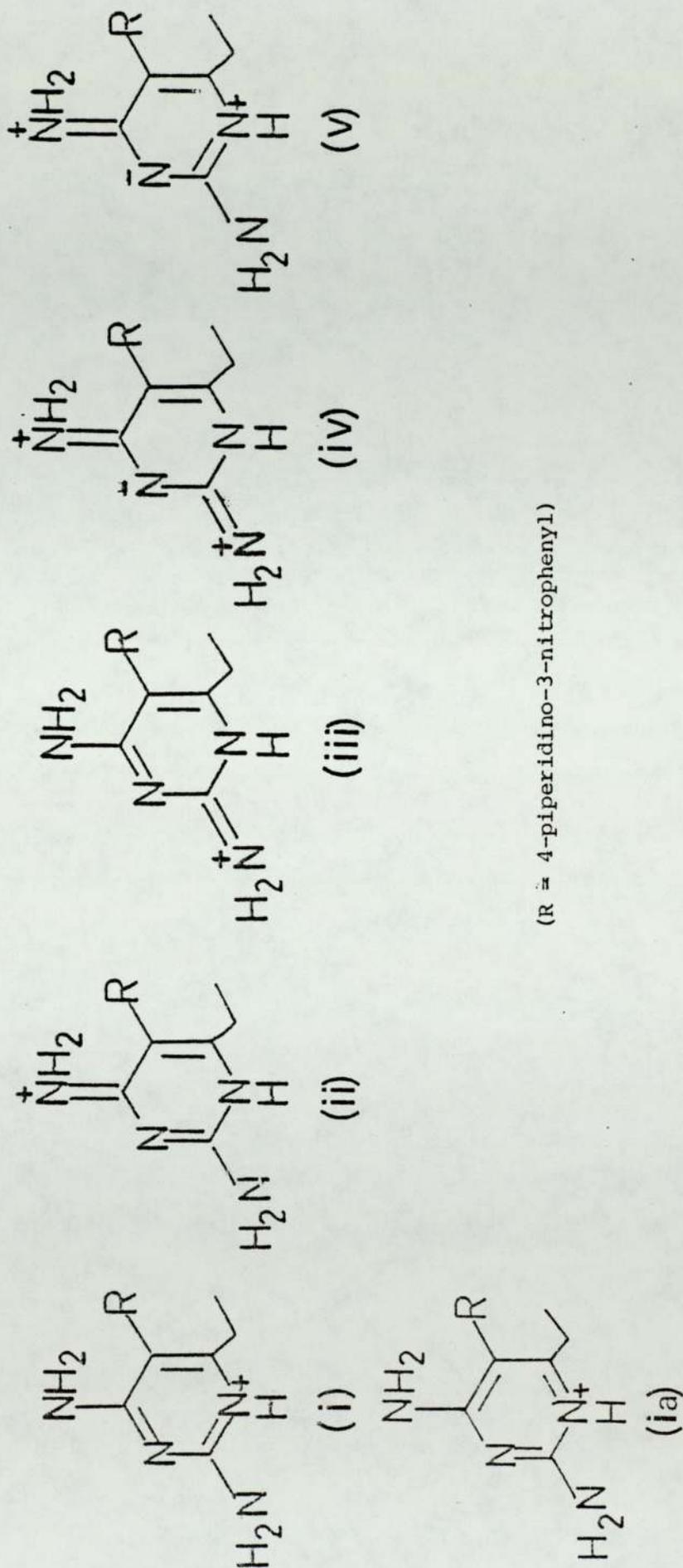
Figure 4.7 Important bond distances [Å] for a number of
derivatives of the antimalarial drug pyrimethamine
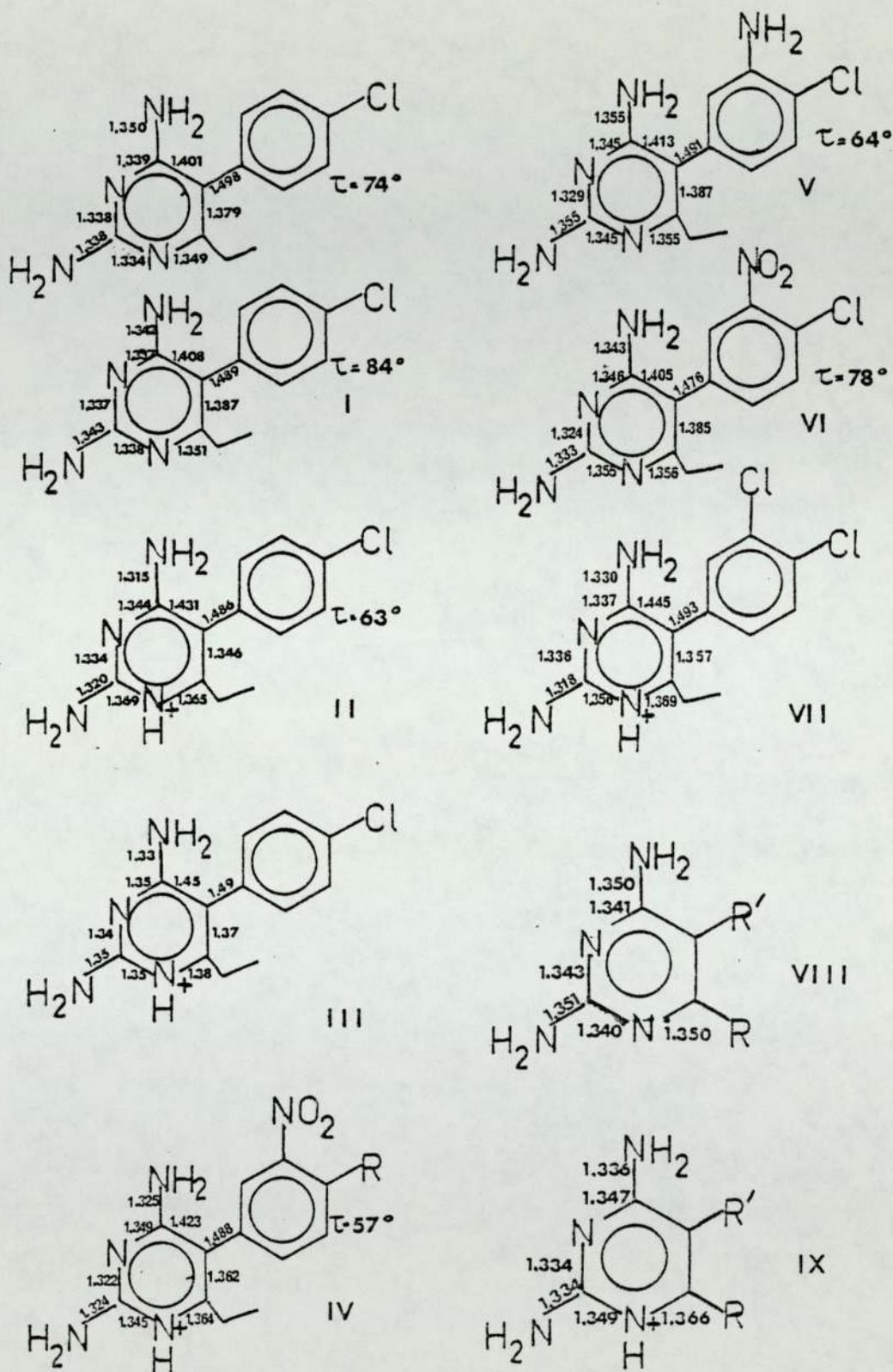
Key to Figure 4.7

(I)   Pyrimethamine[69] (2 independent molecules per
            asymmetric unit)

(II)   Pyrimethamine hydrochloride[69]

(III)   Pyrimethamine hydrobromide[70]

(IV)   NPP Hydrochloride (This work)

(V)   m-aminopyrimethamine[51]

(VI)   m-nitropyrimethamine[51]

(VII)   DDMP Ethanesulphonate salt[71]

(VIII)   Average values for the pyrimidine ring unprotonated
            (Schwalbe and Cody)

(IX)   Average values for the pyrimidine ring protonated at
            N(1) (Schwalbe and Cody)

Figure 4.8   Molecular plot and numbering scheme
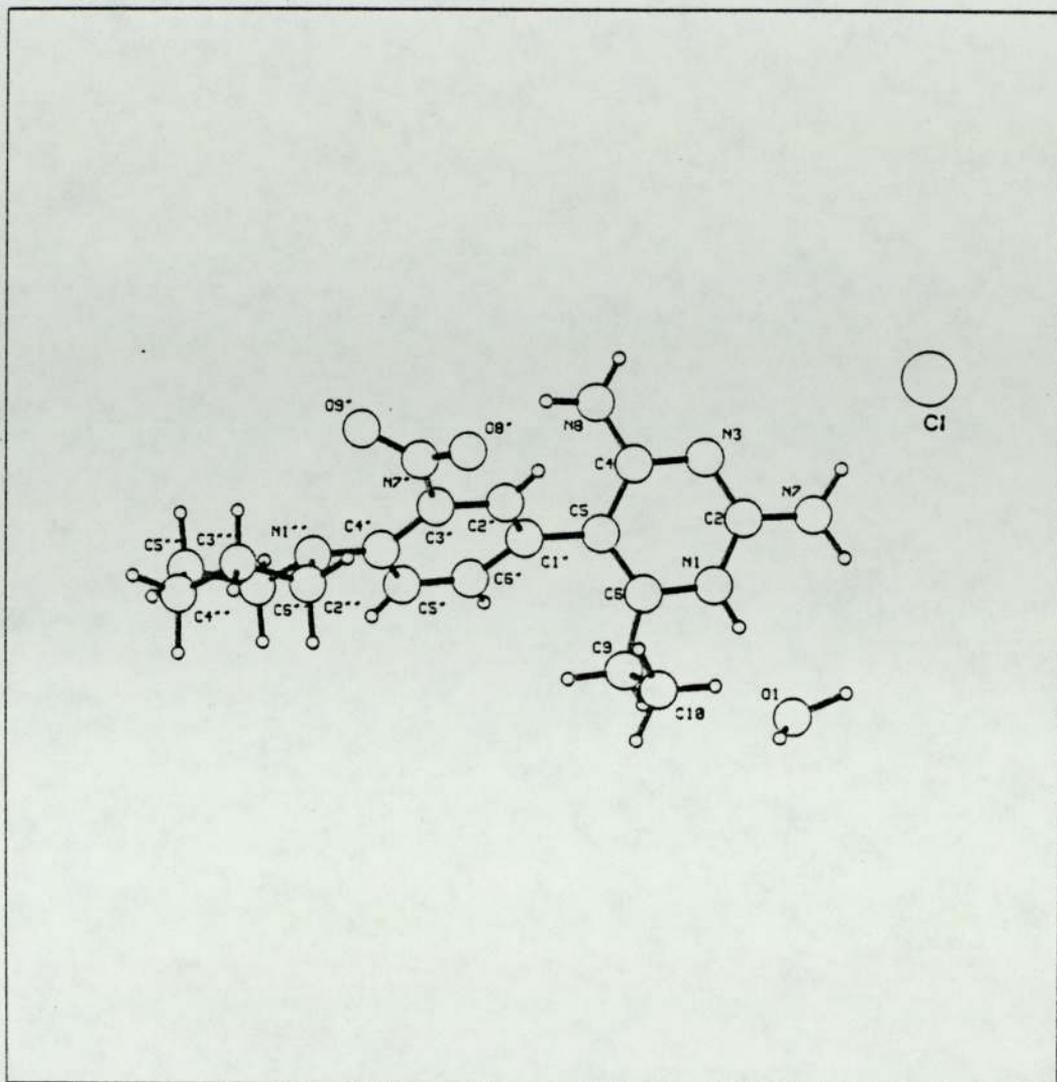
Figure 4.9   Stereo diagram
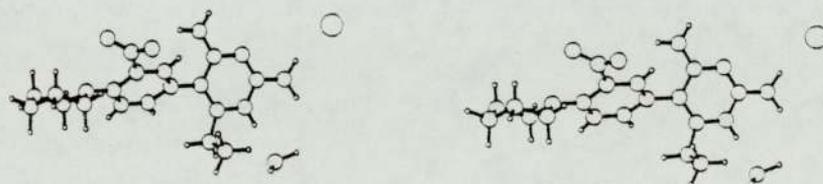
Figure 4.10 Packing diagram viewed along the b axis with c
centering omitted for clarity
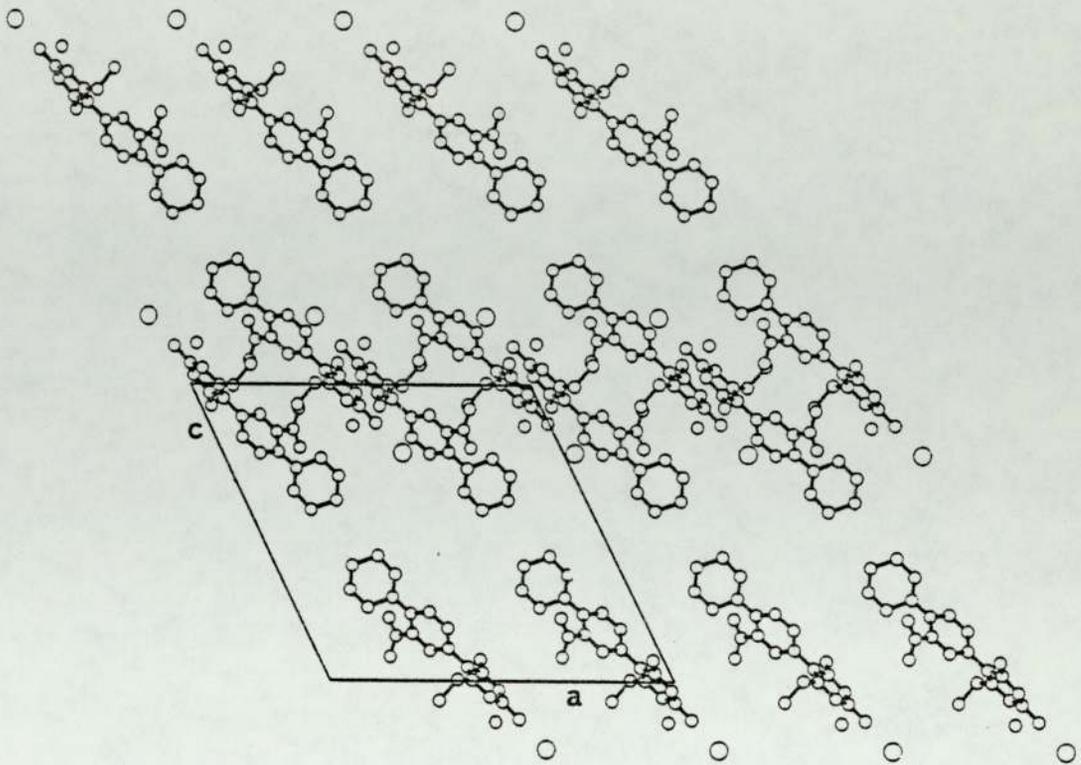
Figure 4.11 Stereo diagram
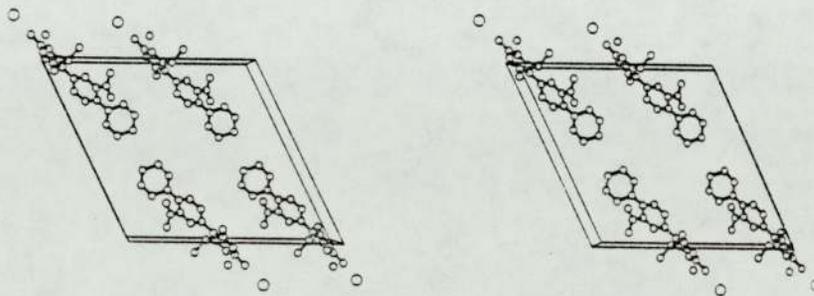
Figure 4.12 Packing diagram viewed along the c axis with the c
centering omitted for clarity
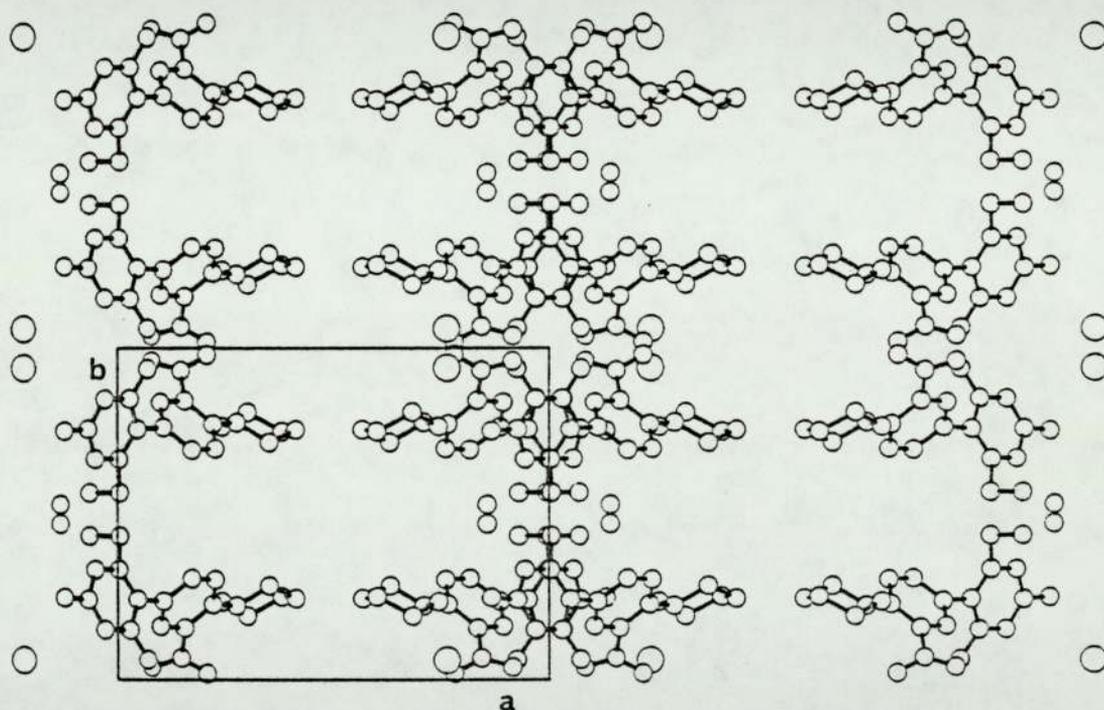
Figure 4.13   Packing diagram viewed along the b axis incorporating
the c centering
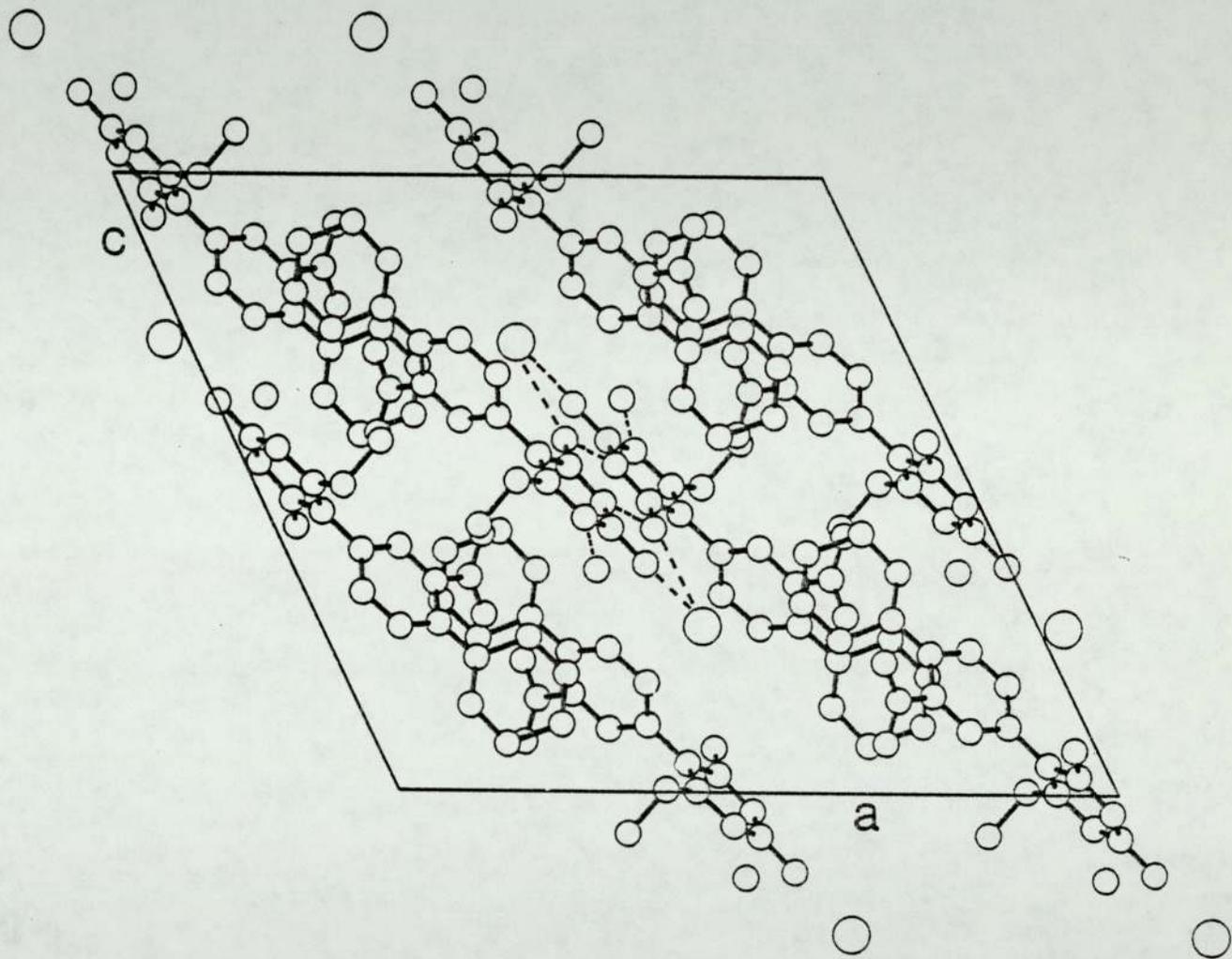
Figure 4.14   Space filling diagram
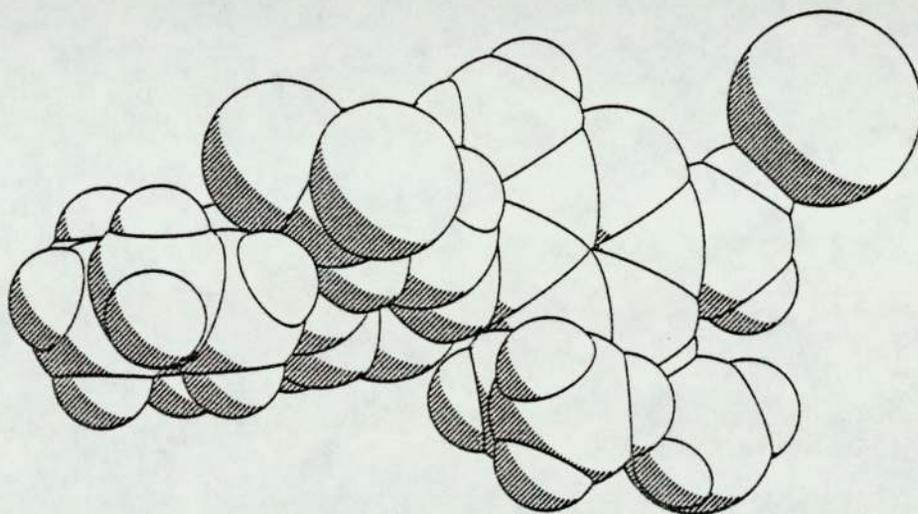
Figure 4.15   Hydrogen Bonding
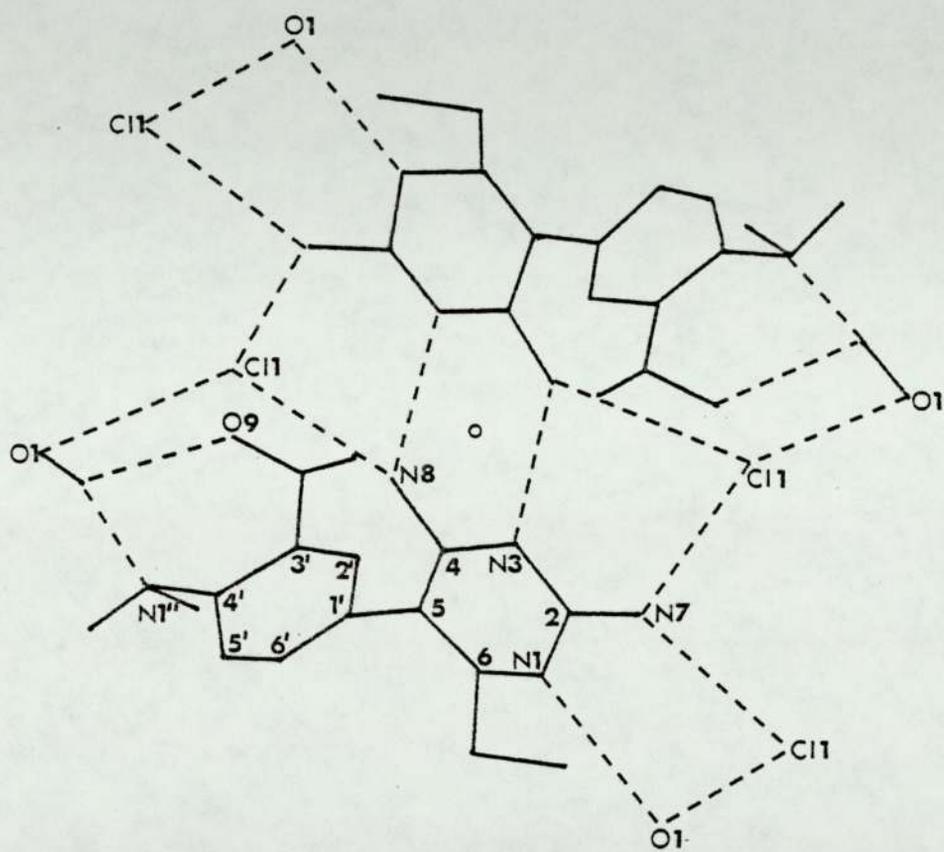
4.3   The crystal structure of 2-Amino-5-(4-chlorophenyl)-6-ethylpyrimidin-4-(3H)-one-monohydrochloride

4.3.1   Experimental

The synthesis of 2-Amino-5-(4-chlorophenyl)-6-ethylpyrimidin-4(3H)-one-monohydrochloride[56].

Pyrimethamine (10 g) was hydrolysed for 24 h in boiling 6N-hydrochloric acid (400 ml). The mixture was stored in a refrigerator for 4 days and colourless crystals of a mixture of 4-amino-5-(4-chlorophenyl)-6-ethylpyrimidin-2(1H)-one   (b)   and   2-amino-5-(4-chlorophenyl)-6-ethylpyrimidin-4(3H)-one   (a)   hydrochlorides   were collected (total 9.4 g).

The mixture was crystallised from 2N-hydrochloric acid (325 ml). After 16 h colourless flakes of the '4-one' hydrochloride (a) deposited (3.0 g) which were further recrystallised from 2N-hydrochloric acid, m.p. 290-305 °C (decomp).

The mother liquors left after removal of the '4-one' hydrochloride were kept at 4 °C for 4 days. Colourless prisms of the '2-one' hydrochloride were collected (4.75 g) and were further crystallised from 2N-hydrochloric acid, m.p. > 310 °C (with decomposition).

Figure 4.16

pyrimethamine-4-one
hydrochloride

(a)

pyrimethamine-2-one
hydrochloride
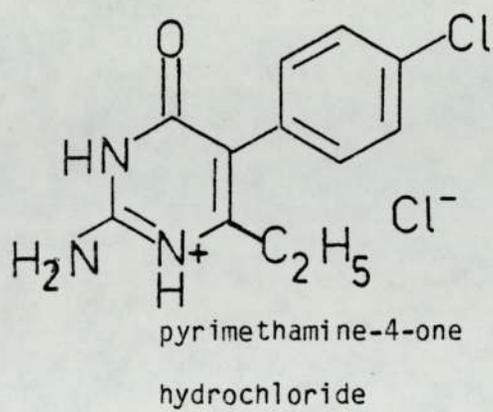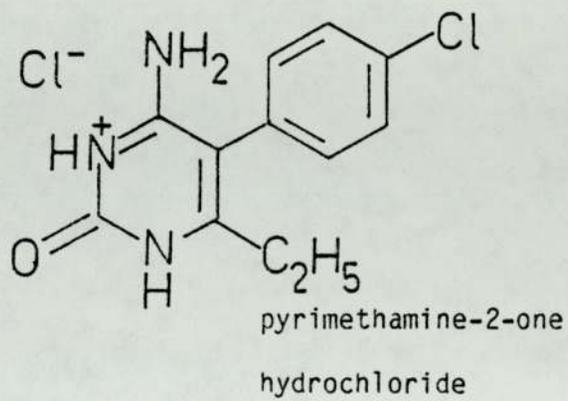
(b)
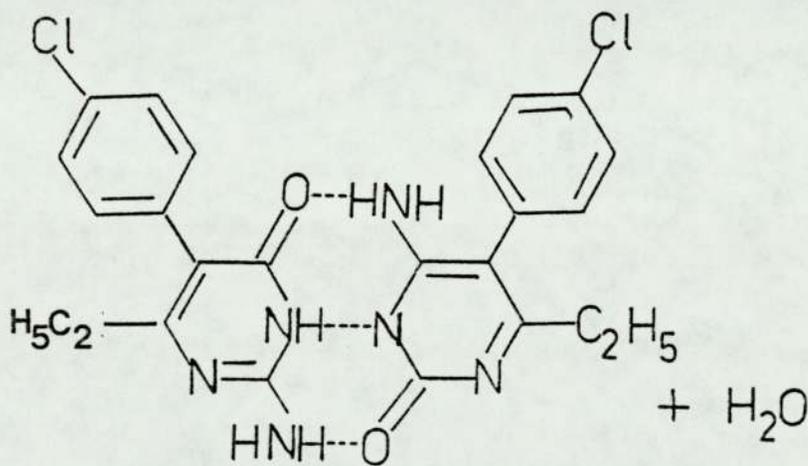
pyrimethamine-4-one/2-one complex

(c)

4.3.2   Crystal Data

2-amino-5-(4-chlorophenyl)-6-ethyl-pyrimidin-4(3H)one   mono-hydrochloride ($C_{12}H_{12}N_3OCl.HCl$) crystallises in the monoclinic space group $P2_1/c$ with a=17.853(5), b=8.581(8), c=9.071(6)$\overset{o}{A}$, β= 93.44(4)°, z=4 and v=1387(1)$\overset{o}{A}{}^3$.

The molecular weight is 286.16 (F(ooo)=592) and the calculated density Dc=1.36 gcm^{-3}.   The absorption coefficient is μ=2.23 cm^{-1} for Mo-Kα radiation (λ = 0.71069$\overset{o}{A}$).

4.3.3   Structure Analysis

Data Collection

The data were collected from a flat, plate-like crystal of dimensions 1.25 x 1.0 x 0.28 mm mounted along its largest dimension on an Enraf-Nonius CAD4 diffractometer using an ω -2θ scan with graphite monochromated Mo-Kα radiation (λ =0.71069$\overset{o}{A}$).   Because of the nature of the crystal - a thin plate - the 'flat' option was used in the data collection parameters[57].   The ω scan speed ranged fom 1.11-5.0 deg min^{-1} dependent upon intensity and the ω scan width = 1.08+0.35 tanθ .   Two reference reflections were measured every 3600 sec of x-ray exposure time and the orientation checked every 150 reflections.   There was no significant loss of intensity between commencement and termination of the data collection and no signs of crystal movement.   2976 reflections were collected between θ =2 and θ =25° ($\pm$h, k, 1) giving 2438 unique observed reflections for which the merged R=0.126.

The unit cell parameters determined from 22 automatically centred reflections are as given above under 'crystal data'.

Structure Determination and Refinement

The data reduction was performed at Birmingham University on the DEC 20 computer using a programme provided by Dr T A Hamor. Lorentz and polarization and intensity corrections were applied and the data output in the form HKL, Fobs and σ (Fobs), where σ (Fobs) is based on counting statistics. No correction for absorption effects were made to the data, since, for Mo-Kα radiation and the crystal size in question, the maximum correction to intensity which would need to be applied would be $\pm10\%$.

Normalised structure factors were calculated using SHELX[2] and the overall scale and estimated temperature factors determined as 0.6586 and 0.047 respectively.

Phases were determined using the SHELX automatic centrosymmetric direct methods package, during which Emin, the minimum value for a normalised structure factor used in the determination, was set to 1.1, and NP the number of phase permutations was assigned the value of 14 (ie. 2^{14} permutations). From this, a trial structure was produced in which all of the non-hydrogen atoms appeared on an E-map. The solution which gave the correct structure was the one with the best figures of merit and the programme assigned the following origin determining reflections:- 7 0 6 (+), $\bar{1}$ 2 1 (+), 8 3 6 (+).

Since it was important to determine the point of acid hydrolysis (ie. 2-one or 4-one) full matrix least squares refinement of positional parameters and isotropic temperature factors ensued during which all C, N and O atoms were assigned the scattering factors of carbon. This ensured there was no "operator" bias to the structural formula. From isotropic temperature factors, and bond distances, the atoms were then assigned their correct scattering factors. A sequence of further refinements of co-ordinates and anisotropic

temperature factors for the chlorine atoms and co-ordinates and isotropic temperature factors for the remaining non-hydrogen atoms using a $1/\sigma^2$(Fobs) weighting scheme led to the determination of all hydrogen atom positions from difference electron density maps and reduction of R to 0.094. Final full matrix least squares refinement of co-ordinates and anisotropic thermal parameters for non-hydrogens and co-ordinates and isotropic temperature factors for hydrogens, using SHELX[2] resulted in R and Rg decreasing to 0.050 and 0.087 respectively. During this stage of refinement, reflections were weighted according to $w=k/[\sigma^2(Fobs)+qFobs^2]$ where k and q (both mentioned earlier) refined to 1.0000 and 0.0013 respectively and eight of the thirteen hydrogen atoms were refined with constrained bond distances. This involved 1944 unique observed reflections for which Fobs $> 3\sigma$(Fobs) for the 215 refinable parameters. A total of 18 reflections had been removed from the data set, some as showing signs of extinction; a secondary extinction correction was refined for the data set. The remainder of these reflections removed were those for which the crystal had been visibly obscured from the x-ray beam by the goniometer head, this apparently being due to the 'flat option' used during the data collection; this was checked by looking at the profiles of the peaks. The refinement was terminated when no positional parameter shifted by more than 0.359σ (y co-ordinate of a hydrogen atom) at which point a final difference electron density map showed no feature greater than $0.33e\overset{o}{A}^{-3}$.

Table 4.9 Positional parameters (fractional co-ordinates x 10^4) with estimated standard deviations in parentheses.

	X/a	Y/b	Z/c
Cl(1)	4273(0)	3047(1)	-5200(1)
Cl(2)	10290(1)	4056(2)	2917(1)
N(1)	6343(1)	5211(3)	-2011(2)
C(2)	5988(2)	4402(3)	-1512(3)
N(3)	5910(2)	3607(3)	-236(3)
C(4)	6586(2)	3586(3)	593(3)
C(5)	7202(2)	4414(3)	-60(3)
C(6)	7051(2)	5225(3)	-1325(3)
N(7)	5119(2)	4391(4)	-2210(3)
O(8)	6629(2)	2897(3)	1753(3)
C(9)	7584(2)	6234(4)	-2105(4)
C(10)	7760(3)	5642(7)	-3617(4)
C(1')	7962(2)	4305(3)	699(3)
C(2')	8089(2)	4701(4)	2175(3)
C(3')	8802(2)	4638(5)	2855(4)
C(4')	9381(2)	4147(5)	2065(4)
C(5')	9281(2)	3703(5)	621(4)
C(6')	8566(2)	3794(4)	-65(4)
H(1)	6194(21)	5863(43)	-2926(42)
H(2)	5005(22)	4966(42)	-3049(32)
H(3)	4771(18)	3762(38)	-1982(39)
H(4)	5423(22)	3067(56)	120(54)
H(5)	7305(21)	7181(46)	-2275(36)
H(6)	8073(23)	6354(45)	-1584(40)
H(7)	7245(18)	5564(51)	-4166(44)
H(8)	8167(22)	6427(46)	-3845(52)

Table 4.9 continued

	X/a	Y/b	Z/c
H(9)	8009(23)	4576(36)	-3695(46)
H(10)	7618(15)	5023(38)	2672(31)
H(11)	8930(25)	5066(50)	3947(31)
H(12)	9615(26)	3342(51)	27(47)
H(13)	8457(20)	3571(41)	-1090(42)

Table 4.10 Anisotropic temperature factors (non-hydrogen atoms)

Isotropic temperature factors (hydrogen atoms) with estimated

standard deviations in parentheses

ATOM	U11	U22	U33	U23	U13	U12
Cl(1)	.0605(6)	.0592(6)	.0438(5)	.0113(3)	-.0028(3)	.0053(3)
Cl(2)	.0591(7)	.1718(14)	.0752(7)	.0241(7)	-.0151(5)	-.0112(7)
N(1)	.0569(15)	.0407(13)	.0317(11)	.0030(9)	.0039(10)	.0057(11)
C(2)	.0553(18)	.0409(15)	.0336(13)	-.0041(11)	.0038(12)	.0083(13)
N(3)	.0550(16)	.0478(14)	.0393(13)	.0047(10)	.0031(11)	-.0005(12)
C(4)	.0613(18)	.0411(15)	.0420(13)	.0082(11)	.0016(12)	.0008(14)
C(5)	.0458(17)	.0369(14)	.0383(13)	-.0005(11)	.0010(11)	.0016(12)
C(6)	.0542(17)	.0347(13)	.0368(13)	-.0001(11)	.0048(12)	.0022(12)
N(7)	.0567(18)	.0676(18)	.0445(15)	.0011(13)	-.0054(12)	.0015(14)
O(8)	.0723(17)	.0811(17)	.0573(14)	.0340(13)	-.0044(12)	-.0177(13)
C(9)	.0676(23)	.0522(20)	.0524(19)	.0132(15)	.0073(16)	-.0062(18)
C(10)	.0620(25)	.1240(38)	.0453(18)	.0145(21)	.0146(16)	-.0074(25)
C(1')	.0543(18)	.0364(14)	.0412(15)	.0043(11)	.0015(13)	-.0007(12)
C(2')	.0560(20)	.0532(18)	.0432(16)	-.0004(13)	.0006(13)	.0034(14)
C(3')	.0706(25)	.0692(22)	.0451(17)	.0033(15)	-.0052(16)	-.0041(17)
C(4')	.0512(20)	.0806(25)	.0545(19)	.0131(17)	-.0031(15)	-.0054(17)
C(5')	.0556(22)	.0862(26)	.0599(22)	.0028(19)	.0053(17)	.0058(18)
C(6')	.0625(21)	.0670(21)	.0429(17)	-.0033(15)	.0040(15)	.0000(17)
H(1')	.0668(108)					
H(2)	.0699(107)					
H(3')	.0576(99)					
H(4)	.1230(181)					
H(5)	.0593(98)					

Table **4.10** continued

H(6)	.0651(106)
H(7)	.1002(150)
H(8)	.1046(155)
H(9)	.0858(138)
H(10)	.0509(83)
H(11)	.0987(142)
H(12)	.0833(128)
H(13)	.0599(94)

Table 4.11   Bond distances in $\overset{\text{o}}{\text{A}}$ with estimated standard deviations in parentheses.

BOND	INTERATOMIC DISTANCE $\overset{\text{o}}{\text{A}}$
N(1)-C(2)	1.312(4)
C(2)-N(3)	1.350(4)
C(2)-N(7)	1.317(4)
N(3)-C(4)	1.383(4)
C(4)-C(5)	1.465(4)
C(4)-O(8)	1.205(3)
C(5)-C(6)	1.356(4)
C(5)-C(1')	1.487(4)
C(6)-N(1)	1.375(4)
C(6)-C(9)	1.496(4)
C(9)-C(10)	1.513(5)
C(1)-C(2')	1.382(5)
C(3')-C(4')	1.359(5)
C(4')-Cl(2)	1.756(4)
C(4')-C(5')	1.366(5)
C(5')-C(6')	1.387(5)
C(6')-C(1')	1.387(4)
N(1)-H(1)	1.02(4)
N(7)-H(2)	[a] .92(2)
N(7)-H(3)	[a] .91(2)
N(3)-H(4)	[a] 1.05(3)
C(9)-H(5)	.96(4)
C(9)-H(6)	.97(4)
C(10)-H(7)	[b] 1.02(3)
C(10)-H(8)	[b] 1.02(3)
C(10)-H(9)	[b] 1.02(3)

Table **4.11** continued

C(2')-H(10)	b	1.02(2)
C(3')-H(11)	b	1.07(3)
C(5')-H(12)		.92(4)
C(6')-H(13)		.96(4)

a) Hydrogen positions refined with a constrained bond distance of 1.01 $\overset{\text{o}}{\text{A}}$ and a standard deviation of 0.03

b) Hydrogen positions refined with a constrained bond distance of 1.05 $\overset{\text{o}}{\text{A}}$ and a standard deviation of 0.03

Table 4.12 Interatomic angles [°] with estimated standard
deviations in parentheses (for non-hydrogen atoms, and
those hydrogens attached to a primary amine group)

Atoms	Bond Angle [°]	Atoms	Bond Angle
N(1)-C(2)-N(3)	118.5(3)	C(2)-N(7)-H(2)	123(3)
N(1)-C(2)-N(7)	121.2(3)	C(2)-N(7)-H(3)	121(2)
N(7)-C(2)-N(3)	120.3(3)	H(2)-N(7)-H(3)	116(4)
C(2)-N(3)-C(4)	124.3(3)		
N(3)-C(4)-C(5)	115.2(2)		
N(3)-C(4)-O(8)	119.3(3)		
O(8)-C(4)-C(5)	125.5(3)		
C(4)-C(5)-C(6)	118.3(3)		
C(4)-C(5)-C(1')	118.0(2)		
C(5')-C(5)-C(6)	123.6(3)		
C(5)-C(6)-N(1)	120.6(3)		
C(5)-C(6)-C(9)	126.7(3)		
C(9)-C(6)-N(1)	112.7(3)		
C(6)-N(1)-C(2)	122.8(2)		
C(6)-C(9)-C(10)	114.0(3)		
C(5)-C(1')-C(2')	121.5(3)		
C(5)-C(1')-C(6')	120.3(3)		
C(6')-C(1')-C(2')	118.2(3)		
C(1')-C(2')-C(3')	120.9(3)		
C(2')-C(3')-C(4')	119.1(3)		
C(3')-C(4')-C(5')	122.2(4)		
C(3')-C(4')-Cl(2)	119.6(3)		
Cl(2)-C(4')-C(5')	118.3(3)		
C(4')-C(5')-C(6')	118.6(3)		
C(5')-C(6')-C(1')	121.0(3)		

Table 4.13   Deviations of non-hydrogen atoms from the least squares planes through (a) the pyrimidine ring, and (b) the phenyl ring[+] (atoms used in the plane calculation are marked with an asterisk).

a)

ATOM	DEVIATION [Å]	ATOM	DEVIATION [Å]
*N(1)	.021	O(8)	.092
*C(2)	-.015	C(1')	-.1218
*N(3)	-.011	N(7)	-.019
*C(4)	.030	C(9)	.052
*C(5)	-.026	C(10)	-1.179
*C(6)	.001		

[+] The equation of the plane is:-

$-.254X + .807Y + .533Z - .616 = 0$

where X, Y and Z are orthogonal coordinates in Å along a*, b and c

b)

ATOM	DEVIATION [Å]	ATOM	DEVIATION [Å]
*C(1')	.009	*C(5')	-.009
*C(2')	-.012	*C(6')	.001
*C(3')	.003	Cl(2)	.021
*C(4)'	.008		

+The equation of the plane is:-

$.173X + .942Y - .287Z + 5.991 = 0$

where X, Y and Z are orthogonal coordinates in Å along a*, b and c

The twist angle between the two planes 'Omega' = 55.7°

Table 4.14   Important torsion angles [°].

Atoms	Angle [°]
C(6)-N(1)-C(2)-N(7)	-178.8
C(6)-N(1)-C(2)-N(3)	3.0
N(1)-C(2)-N(3)-C(4)	.1
C(2)-N(3)-C(4)-C(8)	176.6
C(2)-N(3)-C(4)-C(5)	-4.2
N(7)-C(2)-N(3)-C(4)	-178.1
N(3)-C(4)-C(5)-C(6)	5.4
N(3)-C(4)-C(5)-C(1')	-174.0
O(8)-C(4)-C(5)-C(6)	-175.5
O(8)-C(4)-C(5)-C(1')	5.1
C(4)-C(5)-C(6)-N(1)	-2.9
C(4)-C(5)-C(1)-C(2')	-54.8
C(4)-C(5)-C(1')-C(6')	125.0
C(4)-C(5)-C(6)-C(9)	175.3
C(5)-C(6)-N(1)-C(2)	-1.6
C(5)-C(6)-C(9)-C(10)	115.0

The torsion angle A-B-C-D is the projected angle between A-B and C-D
when viewed down the B-C bond; the clockwise rotation of the C-D bond
with reference to bond A-B is considered positive

Table 4.15   Hydrogen bond contact distances and angles

Hydrogen bond	Angle at H [°]	Donor-Acceptor Distance Å
N(3)$_I$-H(4) .... Cl(1)$_{II}$	155.3	3.250
N(1)$_I$-H(1) .... Cl(1)$_{III}$	171.3	3.084
N(7)$_I$-H(3) .... Cl(1)$_{II}$	148.8	3.211
N(7)$_I$-H(2) .... Cl(1)$_{III}$	137.2	3.440

The subscripts I, II, III refer to the equivalent positions:-
x, y, z; x, $\frac{1}{2}$y, $\frac{1}{2}$z; 1-x, 1-y, -1-z

Figure 4.17 Molecular structure and numbering scheme
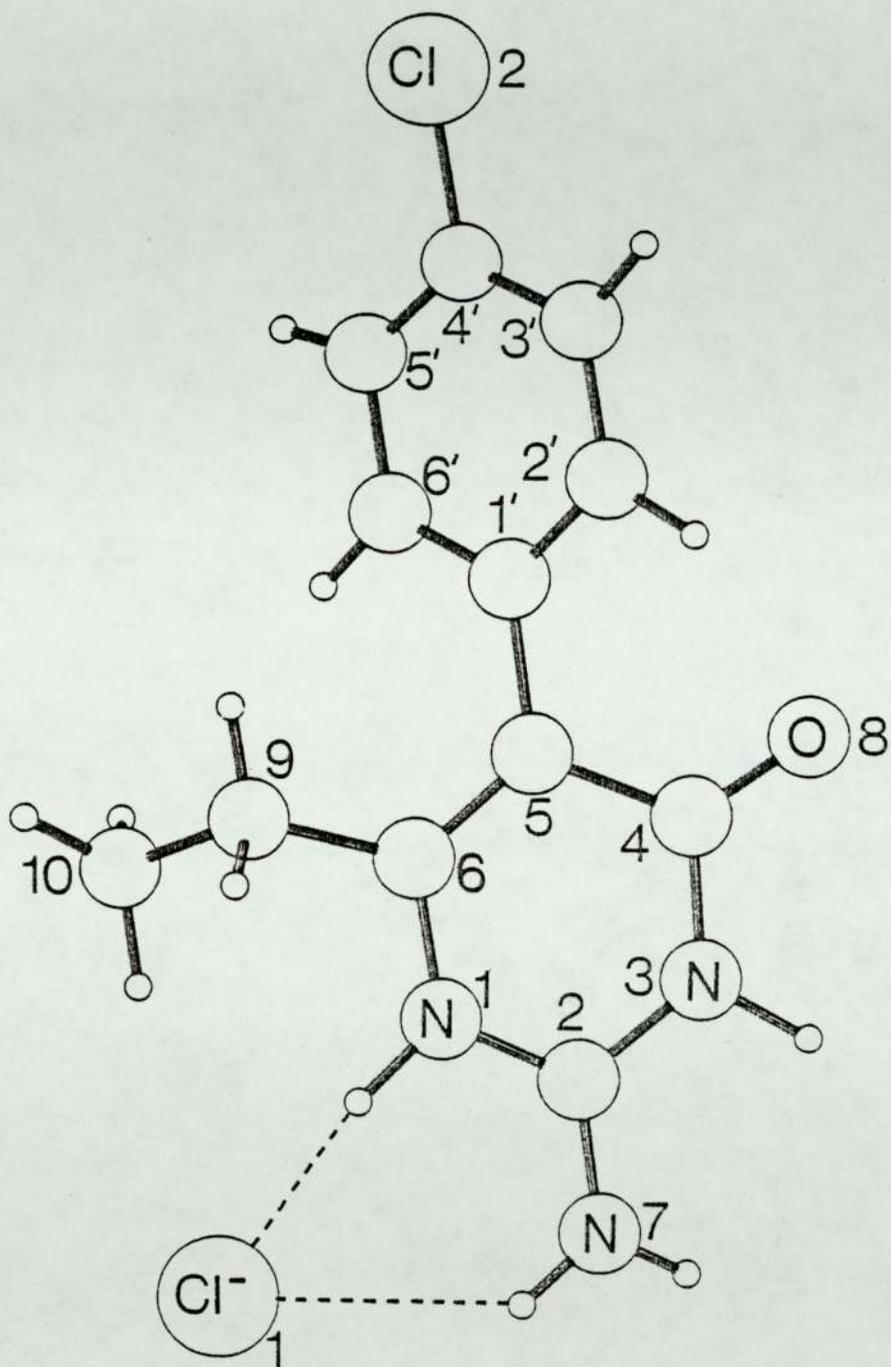
Figure 4.18   Stereo diagram
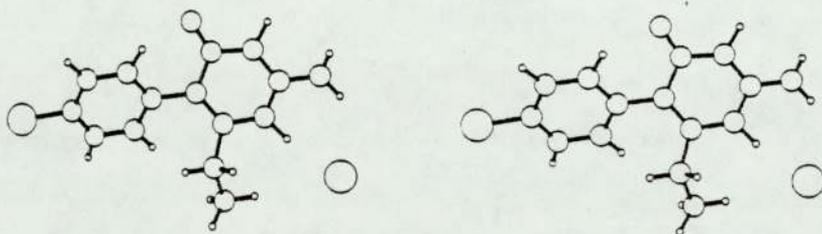
Figure 4.19   Stereo diagram (packing)
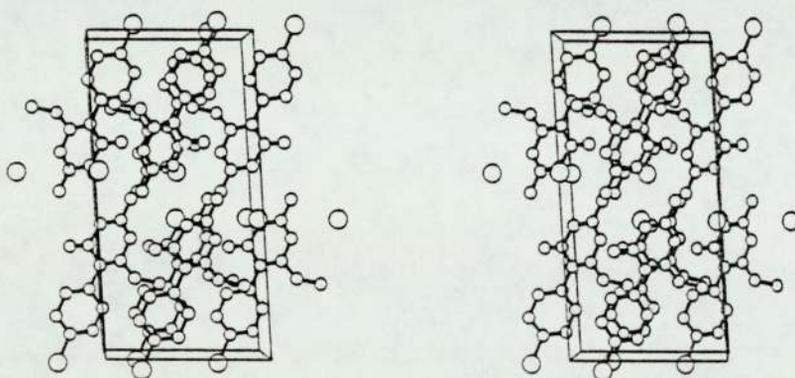
Figure 4.20   Packing diagram (hydrogen bonds in dotted lines)
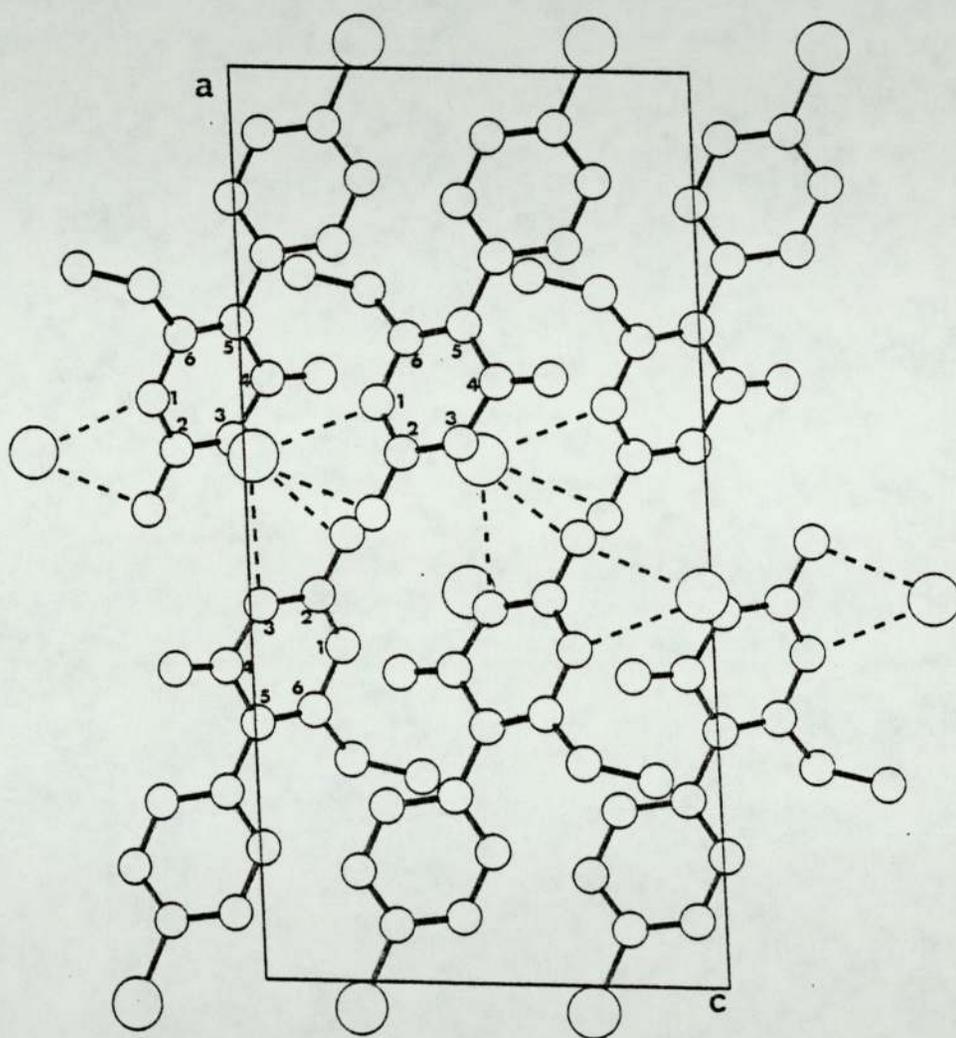
Figure 4.21   Space filling diagram
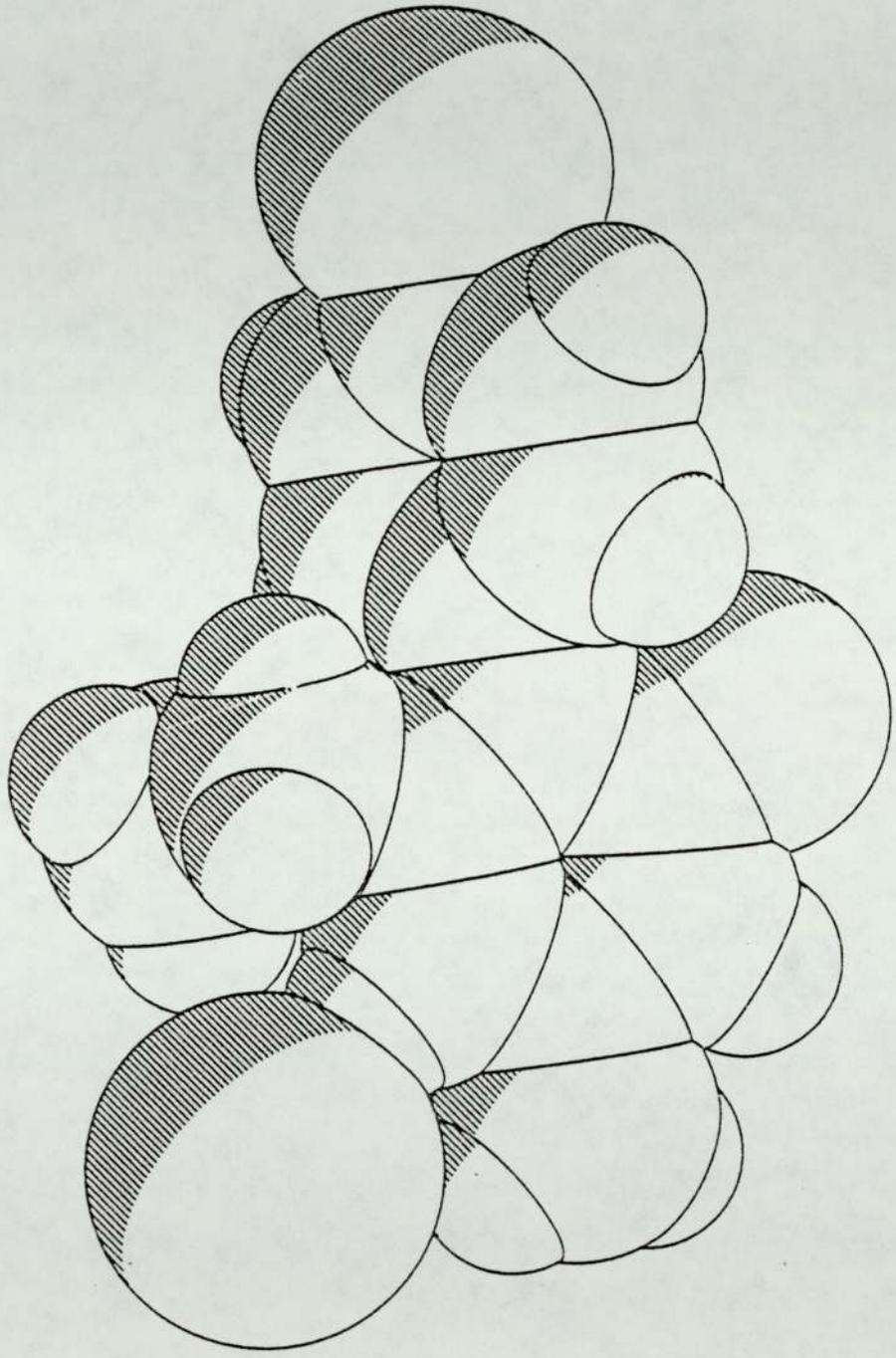
Figure 4.22   Space filling diagram of the -2-one hydrochloride
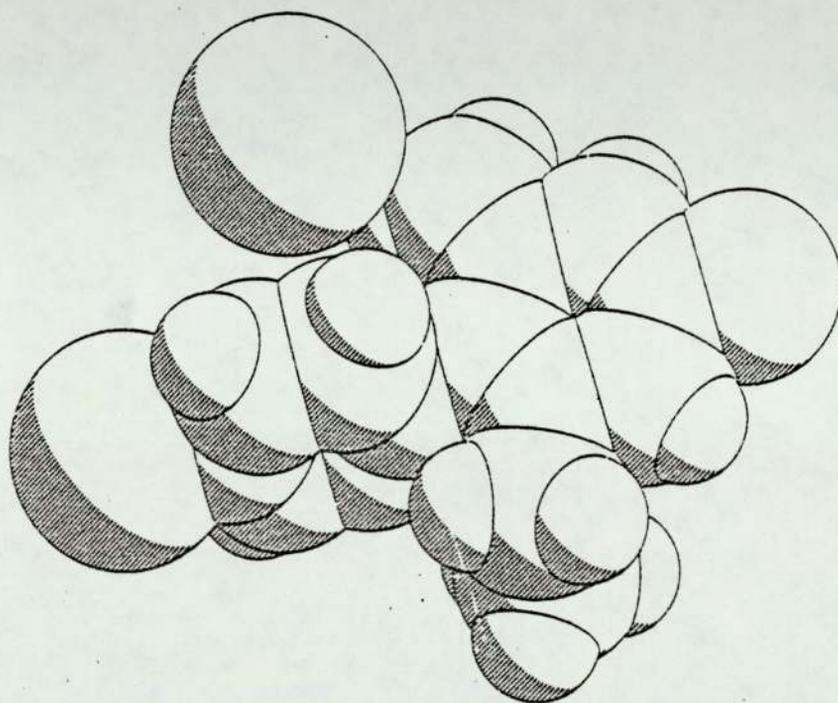
4.3.4   Results and Discussion

The structure of 2-amino-5-(4-chlorophenyl)-6-ethyl-pyrimidin-4(3H)-one monohydrochloride (trivial nomenclature pyrimethamine-4-one monohydrochloride), as determined crystallographically, is depicted in Figure 4.17 with the numbering scheme used; positional parameters for all atoms are given in Table 4.9 with anisotropic thermal parameters for non-hydrogen atoms and isotropic temperature factors for hydrogen atoms in Table 4.10.

Bond distances [$\overset{o}{A}$] and interatomic angles [°] are given in Tables 4.11 and 12. Inspection of the N(1)-C(2) bond distance (1.312(4)$\overset{o}{}$) shows it to have a strong double bond character, suggesting that protonation of the free base to produce the monohydrochloride occurs at N(1) and that the free base itself is protonated at N(3). This is further confirmed by the structure of the complex of pyrimethamine-4-one with the 2-one[58] which shows protonation of the 4-one at N(3). Both of the hydrogen atoms (at N(3) and N(1)) were found from a difference electron density map and further evidence of protonation at these positions can be seen from the two nitrogen valence angles C(6)-N(1)-C(2) (122.8(2)°) and C(2)-N(3)-C(4) (124.3(3)°) both of which comply with Chatar Singh's rule[59]. The C(5)-C(6) bond (1.356(4)$\overset{o}{A}$) also shows considerable double bond character compared with the C(4)-C(5) bond (1.465(4)$\overset{o}{}$). The possible resonance forms for the 4-one hydrochloride are given in Figure 4.23 and the bonding mentioned so far, is consistent with a large contribution from resonance form (1). The C(2)-N(3) bond (1.350(4)$\overset{o}{A}$) also shows some significant shortening compared to the N(1)-C(6) (1.375(4)$\overset{o}{A}$) and N(3)-C(4) (1.383(4)$\overset{o}{A}$) bonds suggesting some electron donation by the N(3) lone pair to the pyrimidine ring, this being consistent with a contribution from resonance form (II) in

Figure 4.23.

The ring carbonyl C(4)-O(8) bond (1.205(3)$\overset{o}{A}$) is substantially shorter and stronger than in other compounds determined in the series (pyrimethamine-2-one hydrochloride and pyrimethamine-2-one-4-one complex)[51] this being in part due to the fact that it does not participate in any intermolecular hydrogen bonding, but also suggesting minimal contribution from resonance forms (III) and (V) in Figure 4.23. The C(2)-N(7) attachment to the extra-annular amino group (1.317(4)$\overset{o}{A}$) suggests very considerable double bond character, the sp^2 nature of the nitrogen N(7) being confirmed by the sum of the C(2)-N(7)-H(2), C(2)-N(7)-H(3) and H(2)-N(7)-H(3) bond angles (360(6)°). This electron donating effect of the amine nitrogen N(7) is consistent with a substantial contribution from resonance form (IV) in Figure 23.

The C(5)-C(1') bridging bond (1.487(4)$\overset{o}{A}$) is slightly shorter than the norm for a c(sp^2) single bond, but is nevertheless consistent with it; it is difficult to envisage any real double bond character with the dihedral angle between the two rings being 55.7°. The bond lengths in the benzene ring, albeit consistent with partial double bonding, show a systematic shortening in the 2-one and 4-one hydrochlorides of the C(3')-C(4') and C(4')-C(5') bonds compared with other bonds in the ring. This shortening in both cases is of the order of 0.015-0.025$\overset{o}{A}$ and is probably due to the inductive effect of the more electronegative chlorine atom. The C(4')-Cl(2) bond distance (1.756(4)$\overset{o}{A}$) (1.740(2)$\overset{o}{A}$ in compound (b) Figure 4.16) suggest some overlap of the chlorine p orbitals with those of the attached carbon as compared with the C-Cl single bond in mitozolomide [1.785(3)$\overset{o}{A}$] even allowing for an effect due to the different hybridisation of the carbon atoms. The bond angles in the benzene ring (Table 4.12) are all as expected, close to 120°.

Figure 4.23 Possible canonical forms for pyrimethamine-4-one monohydrochloride
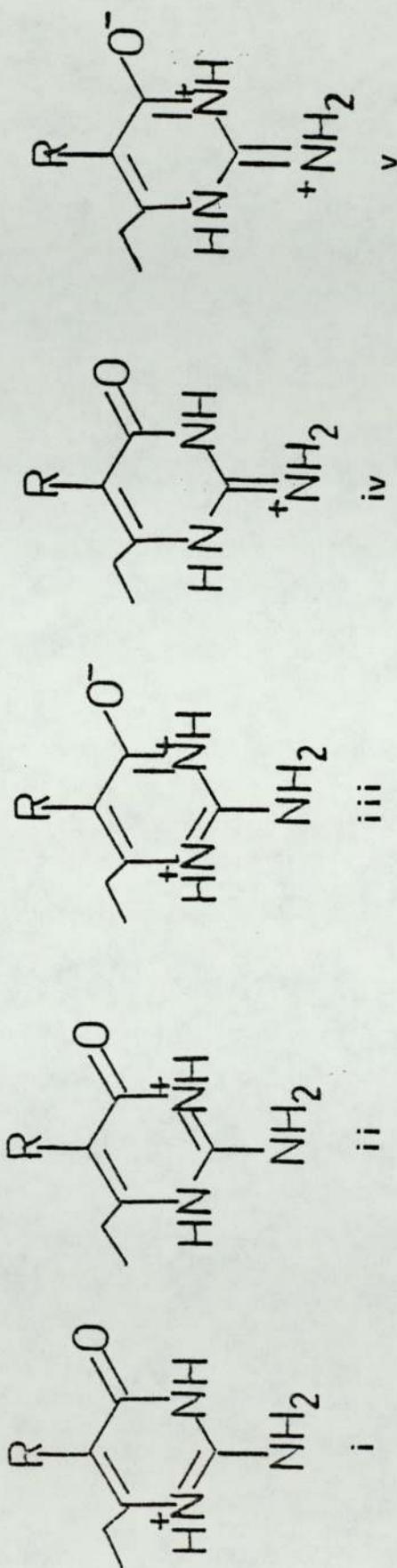
(R=p-Chlorophenyl)

Deviations in $\overset{o}{A}$ from the least squares planes through the pyrimidine ring and the benzene ring are given in Table 4.13. Together with the torsion angles (Table 4.14) these show the pyrimidine ring and the substituents at the 2,4 and C(9) at the 6 position to be essentially coplanar. The C(1') atom of the benzene ring is slightly out of the plane of the pyrimdine ring with a deviation of $-0.122\overset{o}{A}$. The dihedral angle between the planes through the pyrimidine and benzene rings is 55.7° compared with 80° in compound (b). The torsion angle N(1)-C(6)-C(9)-C(10) in (a) (-66.6°) is considerably less than in (b) (-83.4) thus bringing the methyl group in the latter into much closer proximity with the benzene ring. The remaining torsion angles involved with the pyrimidine ring are all close to either 0° or 180° confirming the planar nature of the system.

The molecular packing arrangement is shown in Figures 4.19 and 20 and the hydrogen bond contact distances and angles given in Table 4.15. Unlike compound (b) the carbonyl oxygen in (a) does not participate in intermolecular hydrogen bonding. Protonation in producing the hydrochloride salt occurs at N(1) which then in turn hydrogen bonds to Cl(1) via an N(1)-H(1) .... Cl(1) interaction [$3.084\overset{o}{A}$]. The same chloride ion also hydrogen bonds to the adjacent exocyclic N(7) atom via an N(7)-H(2) .... Cl(1) contact [$3.440\overset{o}{A}$]; this contact is substantially longer and hence weaker. Further hydrogen bonding occurs between N(7) and Cl(1) and N(3) and Cl(1) via N(3)-H(4) .... Cl(1) [$3.250\overset{o}{A}$] and N(7)-H(3) .... Cl(1) [$3.211\overset{o}{A}$] interactions, this chloride being related to the previous one by the action of the screw axis. The net result of this arrangement is that each chloride ion participates in four hydrogen bonds linking pairs of molecules to form infinite chains in the direction of the b

axis.   Consideration of the stereo pair diagram (Figure 4.19) suggests that stacking forces are not important.

So far in this series of compounds, the following structures have been determined at Aston:-

pyrimethamine-4-one monohydrochloride (this work) (a)

pyrimethamine-2-one monohydrochloride[58] (b) (Figure 16)

pyrimethamine-2-one/4-one complex[58] (c) (Figure 16)

A comparison of (a) with the '4-one' free base of (c) shows only one significant difference in terms of bond lengths.   Namely, that the carbonyl bond in the latter $(1.263(10)\overset{o}{A})$ is significantly longer than in the monohydrochloride; this is partly due to its involvement in hydrogen bonding in the complex.   The dihedral angle in the free base (67°) is also larger than in the monohydrochloride, but both are still significantly less than their '2-one' counterparts.

More interesting and significant differences occur between the two hydrochlorides (a) and (b), and these may be summarized as follows:-

a)   The site of protonation to produce the salt; N(1) in (a) and N(3) in (b).

b)   The large difference in dihedral angle; 55.7° for (a) and 80.0° for (b).

c)   The greater delocalisation of electrons in the pyrimidine ring in (b) compared to (a)

d)   The intermolecular hydrogen bonding; compound (a) does not involve the carbonyl oxygen in hydrogen bonds whilst compound (b) does.

The large difference in the dihedral angle can possibly be explained in terms of steric hindrance.   There are three factors involved here:-

(i)   In the '2-one' compound (b), the N(1)-C(6)-C(9)-C(10) torsion angle is 16.8° larger than in the '4-one' (a) bringing the C(10) methyl group into closer proximity with the benzene ring.

(ii)   The slightly bulkier exocyclic amino group at C(4) in (b) as opposed to the carbonyl oxygen in (a).

(iii)   The siting of the chloride ion in compound (b) compared with that of (a) (see Figures 4.21 and 4.22).

The difference in delocalisation of electrons in the pyrimidine ring can be seen from a comparison of the bond lengths involved (Table 4.16).

Table 4.16   Pyrimidine ring bond lengths [$\overset{o}{A}$] for pyrimethamine-4-one and -2-one hydrochlorides

Bond	Distance [$\overset{o}{A}$] (a)	Distance [$\overset{o}{A}$] (b)
N(1)-C(2)	1.312(4)	1.363(2)
C(2)-N(3)	1.350(4)	1.375(2)
N(3)-C(4)	1.383(4)	1.358(2)
C(4)-C(5)	1.465(4)	1.423(2)
C(5)-C(6)	1.356(4)	1.368(2)
C(6)-N(1)	1.375(4)	1.371(2)

From the proposed structures, the distances underlined should be associated with double bonds, however, it can be seen that this is more the case in (a) than in structure (b). This difference in delocalisation in the pyrimidine ring is also true for the two free bases in the '2-one' - '4-one' complex (c).

Although the two compounds (a and b) differ as mentioned in their hydrogen bonding interactions, due mainly to the siting of the

chloride ion and the packing arrangement, they are similar in their extensive use of the chloride counter ion interactions and their inability to form cyclic dimers linked by paired N-H ... N bonds as is very common in crystalline antifolate drugs.

A comparison of the pyrimidines discussed in this section, with the 2,4 diaminopyrimidines shows fundamental differences in terms of hydrogen bonding and bond distances both in the pyrimidine ring and to associated amino substituents. In the case of 2,4 diaminopyrimidine free bases, base-pair dimerization is maximised with each molecule forming two base-pair dimers (double inverse base pair). Protonation of these compounds at N(1) to form the salts naturally eliminates one of the base-pairs with both protons on the C(2) substituent now participating in counter ion interactions. With the pyrimidine compounds, the free bases in the complex of pyrimethamine-2-one and 4-one again form a base pair relationship of the type found in isocytosine[60] - namely a 'Watson and Crick' base-pair. The effect of protonation of the "-2-one" and "-4-one" compounds is to eliminate base pairing completely, the interaction with the counter ion now being of fundamental importance. Protonation in this class of compounds, therefore, appears quite naturally to reduce/eliminate base pair formation, and elicit a strong tendency to hydrogen bonding with the counter ion.

The differences in the bond distances in the pyrimidine ring and the ring to exocyclic nitrogen(s) attachment can be summarized as follows: Taking pyrimethamine-4-one as the reference, NPP hydrochloride shows a lengthening of the N(1)-C(2) bond (1.312$\overset{o}{A}$ to 1.345$\overset{o}{A}$), a shortening of the C(2)-N(3) bond (1.350$\overset{o}{A}$ to 1.322$\overset{o}{A}$), a shortening of the N(3)-C(4) bond (1.383$\overset{o}{A}$ to 1.349$\overset{o}{A}$) and a lengthening of the C(4)-C(5) bond (1.423$\overset{o}{A}$ to 1.465$\overset{o}{A}$). The C-NH$_2$ bond distance in

the protonated "-2-one" and "-4-one" compounds is generally significantly shorter than in the protonated 2,4-diamino compounds. Presumably this difference is due to the removal of one electron-donating amino group and its replacement with an electron-withdrawing carbonyl group which can elicit additional electron donation from the remaining exocyclic nitrogen, thus increasing the immonium character and enhancing the proton donor effect in hydrogen bonding.

4.4　The crystal structure of 5-sulphamoyl-3H-azepin-2(1H)-one[61]

4.4.1　Experimental

Preparation of 5-sulphamoyl-3H-azepin-2(1H)-one

4-Azidobenzenesulphonamide (2.9 g) was dissolved in redistilled tetrahydrofuran (370 ml) and water (630 ml). The mixture was photolysed for 20 hrs through a quartz filter with a 100W medium pressure lamp in an Hanovia Photochemical Reactor and the solution was vacuum evaporated to give a brown solid (2.4 g). A methanol solution of the solid was chromatographically fractionated on a neutral alumina column, and a pale yellow band was eluted with toluene-methanol (10:3). Evaporation of solvent afforded the crude azepinone (1.9 g) which crystallised from water as cream prisms, m.p. 176-178 (decomp)[61].

Initial structure determination was undertaken by film methods using a Weissenberg camera. A total of 700 reflections were collected on film and measured by the SERC micro-densitometer service, Daresbury. The structure was solved from these data using Multan[1] and refined to an unweighted discrepancy index R=0.102 using SHELX[2]. At this point the refinement of the structure from the film data had converged. The advent of the CAD4 diffractometer at Aston inspired the collection of a new set of data for this compound with the coordinates from the film data taken as the starting point for refinement to the diffractometer data.

Whilst the interpretation of the structure in the remainder of this section is based on the parameters determined from diffractometer data, the unit cell parameters and bond distances determined from film methods are reported.

4.4.2   Crystal Data   (diffractometer)

5-sulphamoyl-3H-azapin-2(1H)-one ($C_6H_8N_2O_3S$) crystallises in the monoclinic space group $P2_1/a$ with a=11.243(6), b=8.124(3), c=8.795(6)Å, β=100.78(5)°, z=4 and V=789.1(8)Å^3.

The molecular weight is 188.2 (F(ooo)=392) and the density Dm (by flotation) = 1.59(1) gcm^{-3} with the calculated density Dx=1.58 gcm^{-3}. The absorption coefficient for Mo-Kα radiation (λ =0.71069Å) is μ=3.2 cm^{-1}.

Unit Cell Dimensions (film data)

a=11.04, b=8.08, c=8.65Å, β=100.8°

Structure Analysis (diffractometer)

The data were collected from a crystal of dimensions 1.08 x 0.80 x 0.32 mm on an Enraf-Nonius CAD4 diffractometer using the ω-2θ scan technique with graphite monochromated Mo-Kα radiation (λ =0.71069Å). The scan width=1.35+0.35 tanθ and the ω scan speed of 1.1-9.0 deg min^{-1} depending upon intensity. Two reference reflections were monitored every 5400 sec of x-ray exposure time and the orientation of the crystal checked after every 150 reflections. No appreciable decomposition or movement of the crystal was detected during the collection of the 1489 reflections between θ=2 and θ=25° ($\pm$h, k, l) which merged to give 1389 unique reflections at R=0.036.

The unit cell parameters determined from 25 automatically centred reflections are as given above under "crystal data".

Figure 4.24   Wilson Plot
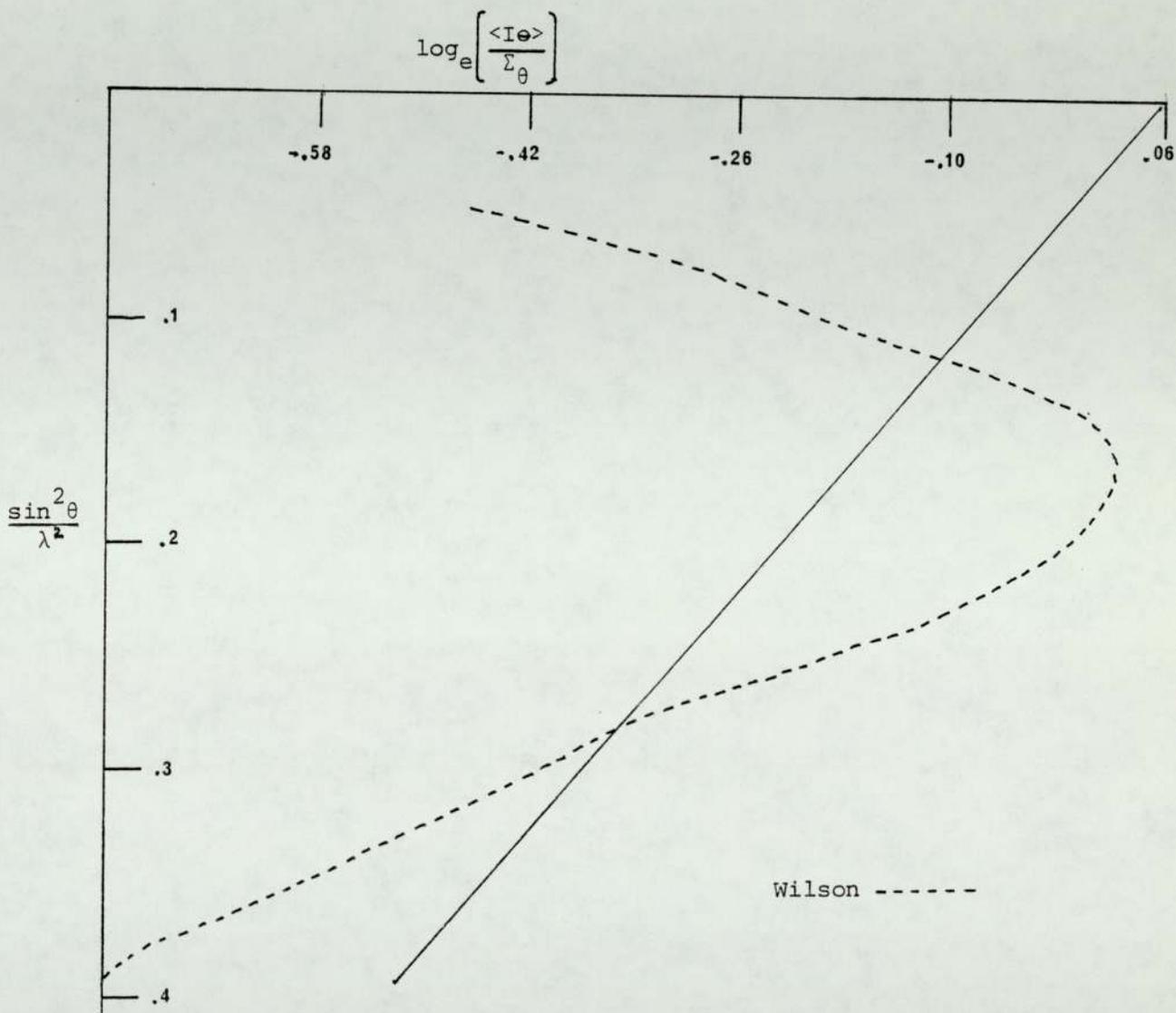
4.4.5   Structure Determination and Refinement

The reduction of the film data was performed at Aston on the ICL 1904S computer, this routine applying Lorentz and polarisation corrections and scaling together the various independently collected levels of data by the Hamilton Rollett and Sparks algorithm.   As stated earlier, it was from this set of reduced data that the structure was solved.

The data reduction of the diffractometer data was performed at Birmingham University using the DEC 20 computer and a programme provided by Dr T A Hamor. Lorentz and polarisation corrections and a simple linear function relating intensity and time were applied and the data output in the form H K L, Fobs and σ Fobs, where σ Fobs is based on counting statistics.   No corrections to the data were made for absorption effects since, for Mo-Kα radiation and the crystal size in question, transmission factors for I for straight paths through the maximum and minimum thickness of the crystal differ by 18% - this assumes the maximum and minimum dimensions of the crystal. Again as stated earlier it was from this data set that the final refined version of the structure was produced.   Normalised structure factors for the 700 film data were calculated using a modified Wilson plot method[22]. The curve from the Wilson calculation and the least squares fit are shown in Figure 4.24 with the values calculated from the gradient and the intercept being:-

<div align="center">

Temperature factor (B) = .7482

Scale = .9353

</div>

Although systematic absences and successful refinement show the structure to be centrosymmetric, the distribution of Es for the complete data set, and the $|E^2-1|$ statistic closely followed that of an acentric structure.

Phase determination was undertaken using Multan[1] in which the 148 largest E values with $E \geqslant 1.212$ produced 967 unique $\sum_2$ $\sum_1$ relationships for use in phase refinement. On the basis of a calculation, no phases were determined with a probability greater than 95%. The 'Converge' procedure produced the following origin determining reflections:-

$3\ 7\ \bar{1}\ (+)$; $5\ 5\ \bar{4}\ (+)$; $6\ 1\ \bar{7}\ (+)$ and four starting sets were produced by permuting the phases the 6 1 1 and 3 7 5 reflections. The set of signs (+) (-) produced an Emap from which all non-hydrogen atoms were located; this solution had by far the highest combined figure of merit.

Using the atomic positions determined from the film data solution, isotropic refinement of non-hydrogen atoms ensued, with atom scattering factors assigned on the basis of the previously refined film structure and knowledge gained from conventional analytical procedures. A difference Fourier synthesis located four of the hydrogen atoms, the remaining hydrogens being entered in calculated positions since they were all bonded to ring carbon atoms of well defined geometry. Each hydrogen atom was assigned the isotropic temperature factor of its attached atom.

Further refinement of coordinates and anisotropic thermal parameters for non-hydrogen atoms and co-ordinates and isotropic temperature factors for hydrogens was carried out with SHELX[2] with those hydrogens in calculated positions constrained to 'ride' on the corresponding carbon atom (Table 4.19). This reduced the unweighted discrepancy index to R=0.035 for the 1287 reflections deemed observed with Fobs $\rangle$ 3 σ (Fobs). In the latter stages of refinement reflections were weighted according to $w=1/\sigma^2$(Fobs) and there were 129 refinable parameters. Refinement was terminated when no

positional parameter shifted by more than 0.13σ . At convergence, the weighted discrepancy index Rg=0.050, and a final difference electron density map showed no feature greater than $0.33e\overset{\circ}{A}^{-3}$.

Table 4.17 Positional parameters (fractional co-ordinates x 10^4) with estimated standard deviations in parentheses.

	X/a	Y/b	Z/c
N(1)	7302(1)	-3152(2)	2789(2)
C(2)	8252(2)	-2611(2)	2180(2)
C(3)	8716(2)	-902(2)	2619(3)
C(4)	7770(2)	397(2)	2178(2)
C(5)	6788(2)	470(2)	2836(2)
C(6)	6449(2)	-756(2)	3878(2)
C(7)	6692(2)	-2352(2)	3817(3)
O(8)	8725(1)	-3525(2)	1348(2)
S(9)	5817.4(4)	2212.0(6)	2442.9(6)
N(10)	4750(1)	1823(2)	993(2)
O(11)	5265(1)	2442(2)	3771(2)
O(12)	6504(1)	3528(2)	1956(2)
H(13)	5994	-348	4789
H(14)	6390(18)	-3040(25)	4490(25)
H(15)	7140(24)	-4270(27)	2580(32)
H(16)	9039	-856	3854
H(17)	9453	-633	2027
H(18)	7863	1272	1283
H(19)	4330(20)	940(27)	1310(29)
H(20)	5000(25)	1760(29)	30(26)

Table 4.18   Anisotropic temperature factors (non-hydrogen atoms) and
              isotropic temperature factors (hydrogen atoms) with standard
              deviations in parentheses

ATOM	U11	U22	U33	U23	U13	U12
N(1)	.0344(9)	.0201(8)	.0415(10)	-.0018(7)	.0114(8)	-.0031(7)
C(2)	.0299(10)	.0252(9)	.0266(10)	.0032(8)	.0044(8)	.0051(8)
C(3)	.0258(9)	.0253(9)	.0463(12)	.0000(9)	.0103(9)	-.0012(8)
C(4)	.0317(9)	.0207(9)	.0347(11)	-.0016(8)	.0088(8)	-.0027(8)
C(5)	.0271(9)	.0206(9)	.0311(10)	-.0044(8)	.0063(8)	-.0006(7)
C(6)	.0314(9)	.0337(10)	.0280(10)	.0016(8)	.0095(8)	.0020(8)
C(7)	.0308(10)	.0312(10)	.0342(11)	.0066(9)	.0115(9)	-.0011(8)
O(8)	.0495(9)	.0340(8)	.0429(9)	-.0037(7)	.0200(7)	.0078(7)
S(9)	.0340(3)	.0198(3)	.0365(3)	-.0052(2)	.0094(2)	.0026(2)
N(10)	.0334(9)	.0284(8)	.0372(10)	.0008(8)	.0078(8)	.0010(7)
O(11)	.0569(10)	.0460(9)	.0409(10)	-.0096(8)	.0187(8)	.0177(8)
O(12)	.0460(8)	.0200(7)	.0715(12)	-.0026(7)	.0102(8)	-.0056(6)
H(13)	.0600					
H(14)	.0319(42)					
H(15)	.0813(54)					
H(16)	.0500					
H(17)	.0583					
H(18)	.0480					
H(19)	.0571(49)					
H(20)	.0531(47)					

Table 4.19   Bond distances in $\overset{o}{A}$ with estimated standard deviations
              in parentheses

Bond	Interatomic Distances $\overset{o}{A}$ (Diffractometer data)	[+]Interatomic Distances $\overset{o}{A}$ Film Data
N(1)-C(2)	1.355(2)	1.331(12)
C(2)-C(3)	1.508(3)	1.495(11)
C(3)-C(4)	1.497(3)	1.469(13)
C(4)-C(5)	1.340(3)	1.300(14)
C(5)-C(6)	1.452(3)	1.434(12)
C(6)-C(7)	1.328(3)	1.321(12)
C(7)-N(1)	1.394(3)	1.377(12)
C(2)-O(8)	1.231(2)	1.223(10)
S(9)-C(5)	1.781(2)	1.757(9)
S(9)-N(10)	1.610(2)	1.580(8)
S(9)-O(11)	1.434(2)	1.418(7)
S(9)-O(12)	1.431(2)	1.413(7)
N(1)-H(15)	.94(2)	.97(4)
C(3)-H(16)	*1.08	*1.08
C(3)-H(17)	*1.08	*1.08
C(4)-H(18)	*1.08	*1.08
C(6)-H(13)	*1.08	*1.08
C(7)-H(14)	.92(2)	1.02(4)
N(10)-H(19)	.93(2)	.97(7)
N(10)-H(20)	.95(2)	1.00(4)

* Hydrogen atoms entered in calculated positions and constrained to
  'ride' on their corresponding carbon atom.

[+] Included for comparison purposes only

Table 4.20   Interatomic angles [°] for C, N, O and S atoms with standard deviations in parentheses.

Atoms	Bond angle [°]
N(1)-C(2)-C(3)	117.3(2)
N(1)-C(2)-O(8)	120.0(2)
O(8)-C(2)-C(3)	122.7(2)
C(2)-C(3)-C(4)	112.8(2)
C(3)-C(4)-C(5)	121.5(2)
C(4)-C(5)-C(6)	125.2(2)
C(4)-C(5)-S(9)	118.4(2)
S(9)-C(5)-C(6)	116.3(1)
C(5)-C(6)-C(7)	124.3(2)
C(6)-C(7)-N(1)	127.7(2)
C(7)-N(1)-C(2)	128.7(2)
O(11)-S(9)-O(12)	119.1(1)
O(11)-S(9)-C(5)	106.8(1)
O(12)-S(9)-C(5)	107.8(1)
N(10)-S(9)-O(11)	107.4(1)
N(10)-S(9)-O(12)	105.9(1)
N(10)-S(9)-C(5)	109.7(1)

Table 4.21 Deviations of non-hydrogen [in $\overset{o}{A}$] atoms from the least squares planes through the seven membered ring.[+] (atoms used in the plane calculation are marked with an asterisk)

ATOM	DEVIATION [$\overset{o}{A}$]	ATOM	DEVIATION [$\overset{o}{A}$]
*N(1)	-.271	O(8)	-.558
*C(2)	-.164	S(9)	-.897
*C(3)	.497	N(10)	-2.473
*C(4)	-.202	O(11)	-.250
*C(5)	-.249	O(12)	-.787
*C(6)	.209		
*C(7)	.180		

[+] The equation of the plane is:-

0.615X + 0.086Y + 0.784Z + 5.730 = 0

where X, Y and Z are orthogonal co-ordinates in $\overset{o}{A}$ along a*, b and c.

Table 4.22   Torsion angles (°) for bonds involving non-hydrogen
atoms

Atoms	Angle (°)
*N(1)-C(2)-C(3)-C(4)	-59.1
*C(2)-C(3)-C(4)-C(5)	65.6
*C(4)-C(5)-C(6)-C(7)	-31.0
C(4)-C(5)-S(9)-N(10)	92.8
C(4)-C(5)-S(9)-O(11)	-151.1
C(4)-C(5)-S(9)-O(12)	-22.0
*C(5)-C(6)-C(7)-N(1)	0.3
*C(6)-C(7)-N(1)-C(2)	36.0
C(6)-C(5)-S(9)-N(10)	-88.0
C(6)-C(5)-S(9)-O(11)	28.0
C(6)-C(5)-S(9)-O(12)	157.1
*C(6)-C(5)-C(4)-C(3)	-7.5
*C(7)-N(1)-C(2)-C(3)	-1.6
C(7)-N(1)-C(2)-O(8)	176.0
O(8)-C(2)-C(3)-C(4)	123.4
S(9)-C(5)-C(4)-C(3)	171.5
S(9)-C(5)-C(6)-C(7)	150.0

The torsion angle A-B-C-D is the projected angle between AB and CD
when viewed down the B-C bond; the clockwise rotation of the C-D bond
with reference to bond A-B is considered positive.

* Torsion angles involving only ring atoms

Table 4.23   Hydrogen bond contact distances and angles

Hydrogen bond	Angle at H[°]	N ... O Distance [Å]
N(1)$_{II}$-H(15) .... O(12)$_I$	170	2.893
N(10)$_{III}$-H(19) .. O(8)$_I$	156	2.956
N(10)$_I$-H(20) .... O(8)$_{IV}$	153	2.929

The subscripts I, II, III, IV refer to the equivalent positions:-
x, y, z;   x, 1+y, z; $\frac{1}{2}$x, $\frac{1}{2}$y, z;   1$\frac{1}{2}$x, $\frac{1}{2}$y, -z respectively.

Figure 4.25   Molecular structure and numbering scheme
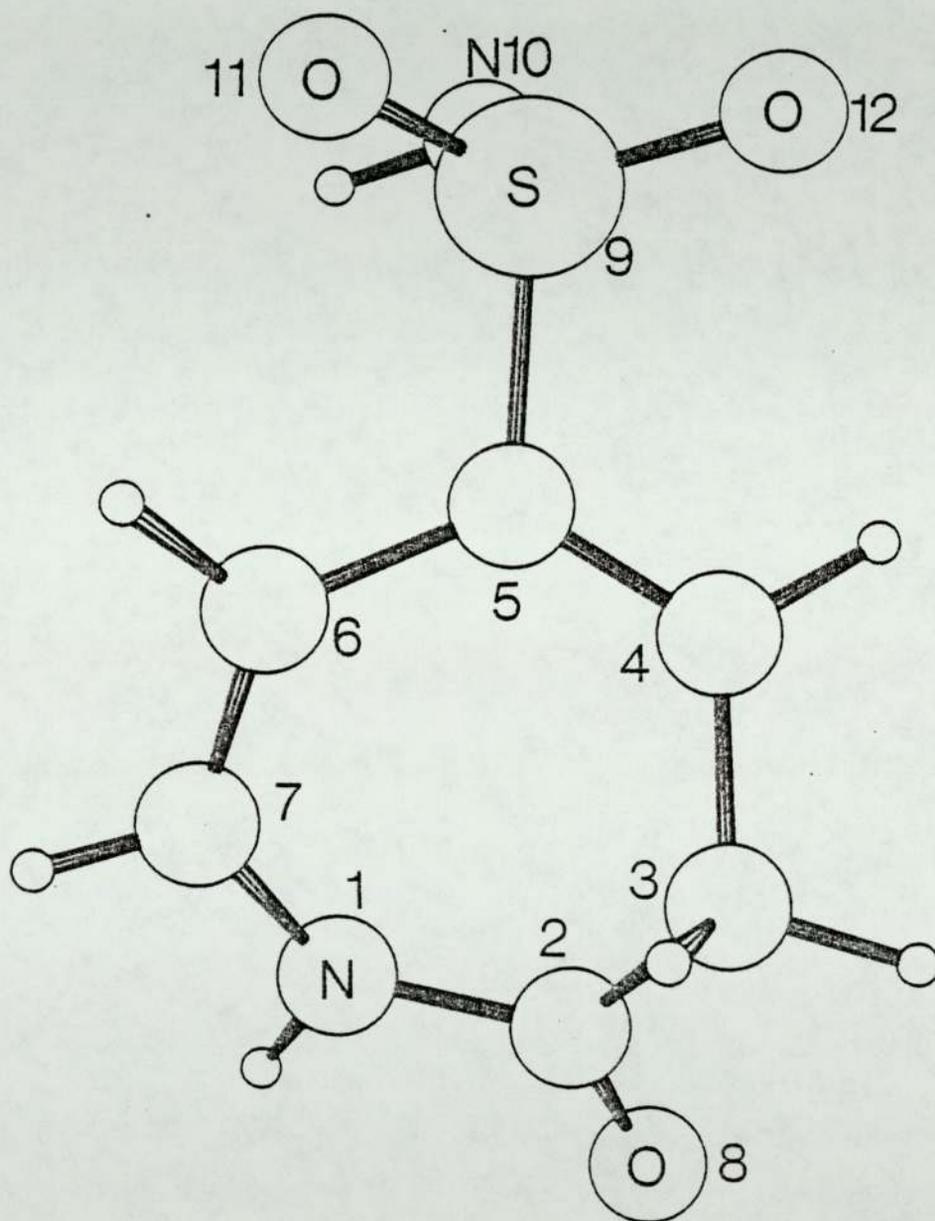
Figure 4.26   Stereo diagram
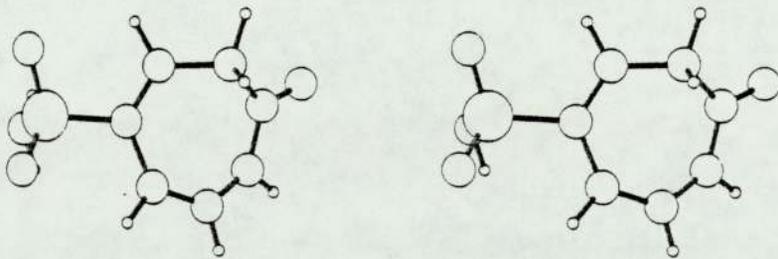
Figure 4.27   Stereo packing diagram
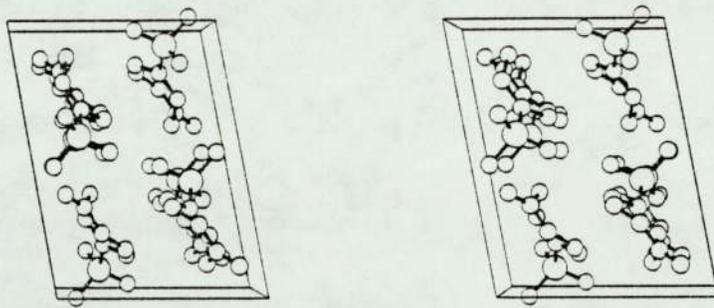
Figure 4.28   Packing diagram
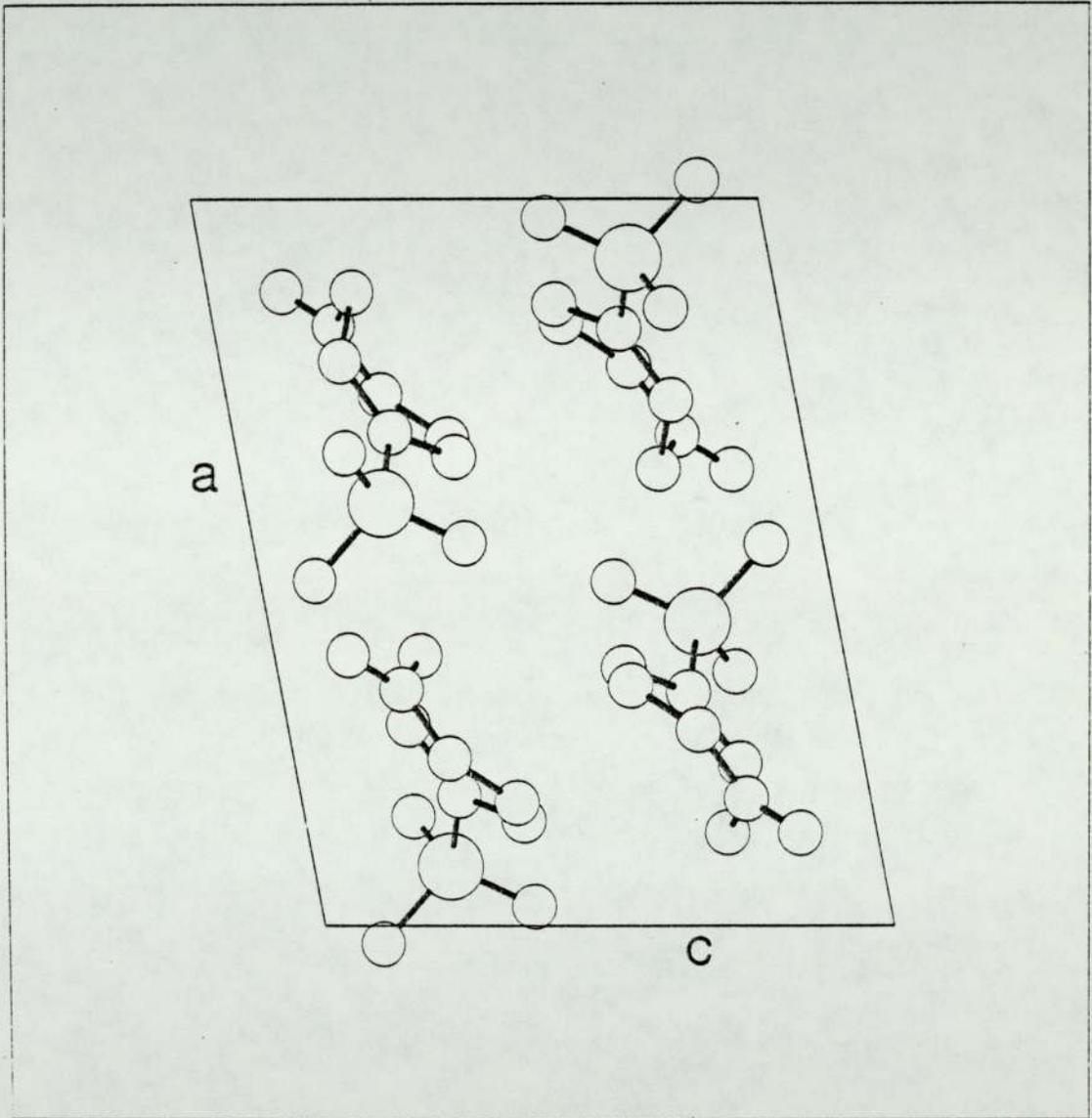
Figure 4.29 Extended packing diagram showing hydrogen bonding in
             dotted lines.   The N(1)-H(15) ... O(12) hydrogen bond
             is not shown.
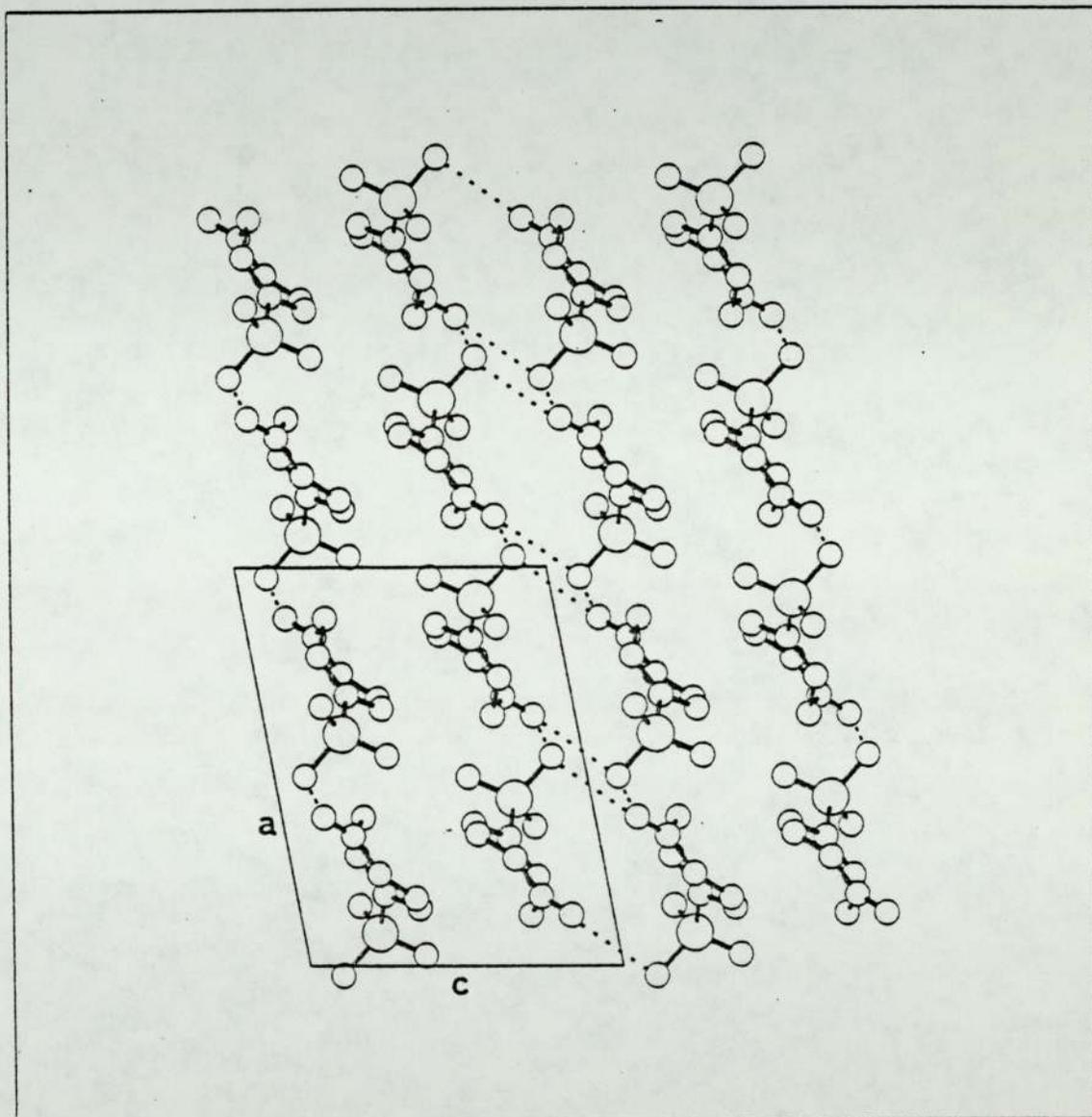
Figure 4.30   Extended packing diagram viewed along the c axis
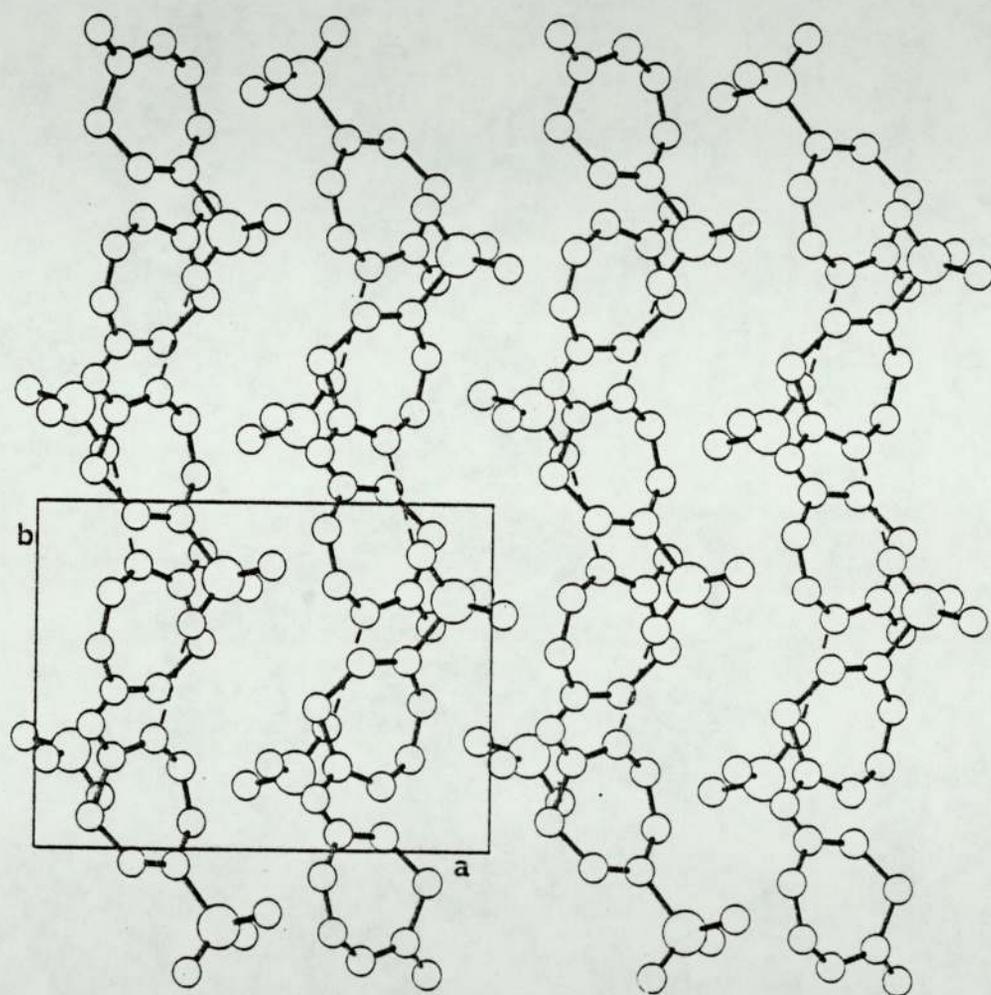
Figure 4.31   Space filling diagram
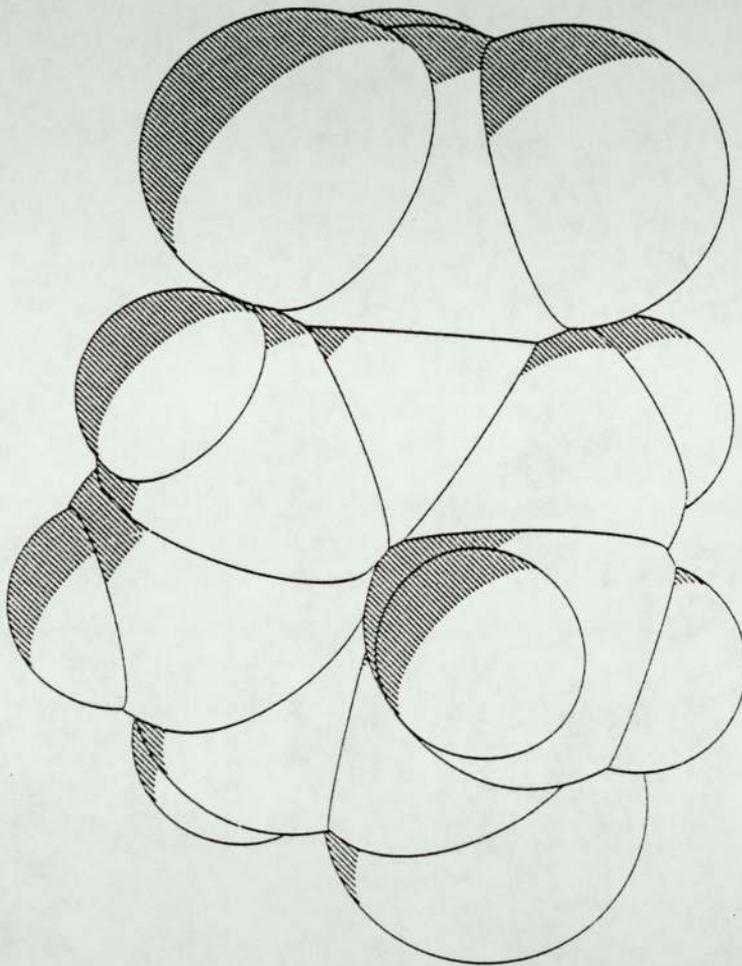
4.4.5   Results and Discussion

The structure of 5-sulphamoyl-3H-azepin-2(1H)-one as determined crystallographically is depicted in Figure 4.25 together with the numbering scheme used; positional parameters for all atoms are given in Table 4.17 with anisotropic thermal parameters for non-hydrogens and isotropic temperature factors for hydrogens in Table 4.18.

The seven-membered ring and the lactam configuration are confirmed, the latter having a C=O bond length of 1.231(2)$\overset{o}{A}$ (Table 4.19) and the hydrogen to N(1) having been located in a difference Fourier synthesis.

As with sulphanilamide[62] the sulphamoyl group is approximately tetrahedral (Table 4.20) with the largest difference from the ideal tetrahedron occurring in the O(11)-S(9)-O(12) bond angle [119.1(2)°] compared with sulphanilamide [118.2(1)°]. The two S-O bonds in the azepinone are almost identical in length 1.434(2)$\overset{o}{A}$ and 1.431(2)$\overset{o}{A}$, albeit that only O(12) is hydrogen bonded, suggesting bond orders of ca. 0.66[63]. The S(9)-N(10) and S(9)-C(5) bond lengths of 1.610(2)$\overset{o}{A}$ and 1.781(2)$\overset{o}{A}$ respectively show some double bond character compared to observed single bond distances[64, 65]; however, the π bond order of the S-C bond appears to be somewhat less than in sulphanilamide where the S-C distance is 1.750(18)$\overset{o}{A}$. The amide portion of the azepinone system shows considerable delocalisation of electrons, but other double bonds are largely isolated. The torsion angles (Table 4.22; see also Figures 4.27 and 28) show the molecule to be substantially non-planar with the seven membered ring puckering into a boat form. The sulphamoyl group aligns with the two oxygens O(11) and O(12) closest to the least squares plane through the seven-membered ring (Table 4.21), the deviations being -0.250$\overset{o}{A}$ and -0.787$\overset{o}{A}$ respectively with the nitrogen pointing out of the plane with a

deviation of $-2.473\overset{o}{A}$. The sum of the two SNH bond angles and the HNH angle at N(10) is 338(6)° (Table 4.20) suggesting a hybridisation that is intermediate between sp^2 and sp^3 for this nitrogen atom. The molecular packing occurs as infinite chains of molecules linked by N(10)-H(19) .... O(8) hydrogen bonds $[2.956(3)\overset{o}{A}]$ running parallel to the _a_ axis, pairs of chains being firmly linked in the direction of _c_ by N(10)-H(20) ... O(8) interactions $[2.929(3)\overset{o}{A}]$. Further hydrogen bonding of the type N(1)-H(15) ... O(12) $[2.893(3)\overset{o}{A}]$ occurs between molecules parallel to the _b_ axis, the whole structure thus comprising pairs of layers linked by intermolecular hydrogen bonds (Table 4.23 and Figures 4.29 and 30).

A comparison of the bond distances determined from diffractometer and film data shows there to be a systematic shortening of all values calculated from film data, this being observed as a second place error. This effect can be rationalised in terms of an equivalent systematic shortening this time in the first decimal place, in each of the unit cell dimensions determined by film. A much closer agreement between the two sets of bond length data would be observed if the least squares refined unit cell dimensions determined from the diffractometer data were used in the refinement of the film data.

A range of bacteriological tests on the azepinone using sulphanilamide and gentamicin Neo-sensitabs as references against Pseudomonas aeruginosa, Staphylococcus aureus, Streptococcus faecalis, Salmonella typhimurium and others (Abdullah, University of Aston)[66] showed that the azepinone did not inhibit bacterial growth and thus did not have antibacterial activity. One of the features required for the molecule to have anti-PABA activity is a primary aromatic amine group at position 4. This N(4) group must not be

substituted in any way unless by a group which will readily break down in the body and liberate the primary amine group[67]; this is clearly not the case with the azepinone.

Further evidence crystallographically which would suggest preclusion of this compound as an antibacterial sulphonamide are its molecular dimensions. Certain important dimensions for sulphanilamide[62], PABA[68] and the azepinone are shown in Figure 4.32.

The distances from the amine nitrogen to the carboxylate oxygens for PABA, a substrate of dihydrofolate synthetase, can be seen to be very similar to those of the amine nitrogen to the sulphamoyl oxygens of sulphanilamide; thus sulphanilamide can and does competitively inhibit folic acid production. These critical distances are obviously important in binding to the enzyme. If one examines similar distances between the ring nitrogen and the sulphamoyl oxygens in the azepinone these can be seen to be substantially different from those in PABA and sulphanilamide. This together with the very non-planar nature of the seven-membered ring, compared with the benzene ring in both sulphanilamide and PABA, make it not too difficult to understand the lack of antibacterial activity in the azepinone.

Figure 4.32   Critical N ... O distances for PABA, sulphanilamide and
5-sulphamoyl-3H-azepin-2(1H)-one

p-aminobenzoic acid (anion)

R=H (sulphanilamide)
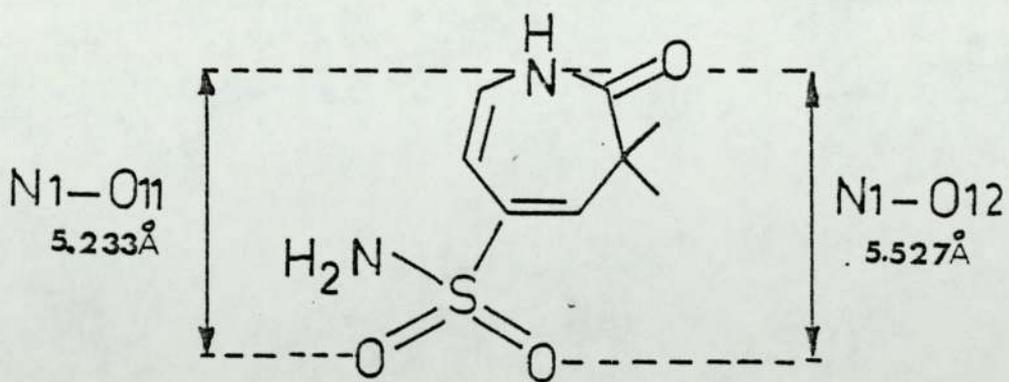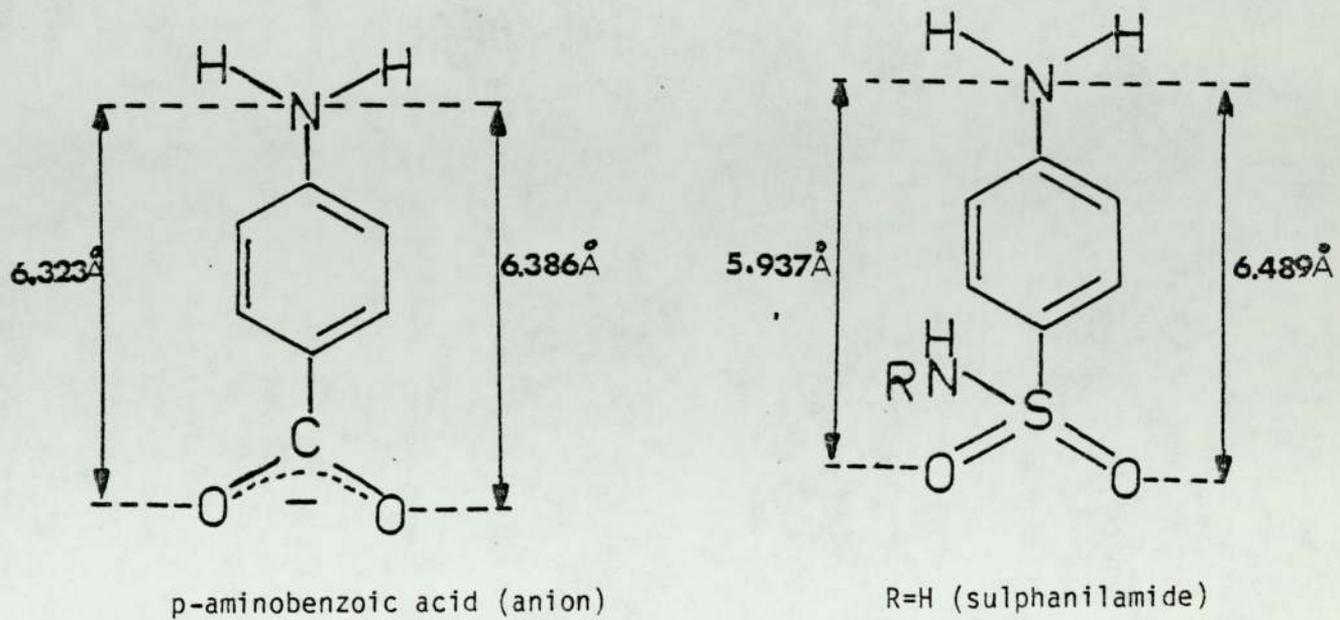
CHAPTER 5

CHAPTER 5

5.1   THE CRYSTAL STRUCTURES OF 6-PHENYL AND 6-BENZYL 1,3,5 TRIAZIN-
       2,4 [1H,3H]-DIONES, AND 6-METHYLISOCYTOSINE

The nucleic acids DNA and RNA are large molecules which store
and transfer genetic information in cells. They are composed of a
polymeric structure of nucleotides which are in turn composed of a
nitrogenous base, a pentose sugar and phosphoric acid.

The bases, purines and pyrimidines are attached to a pentose
sugar at the C(1) position; in DNA the pentose sugar is 2-deoxy-D-
ribose, whilst in RNA it is D-ribose. The attachment of phosphate to
the C(5) position of the pentose sugar yields a nucleoside
monophosphate or nucleotide. The bases in DNA are adenine, guanine,
cytosine and thymine whilst in RNA, thymine is replaced by uracil
(Fig 5.1). The polymerisation of the doxyribonucleotides into a
polynucleotide strand of DNA is shown in Figure 5.2[72]. Two strands
then combine to form a double helical DNA molecule, the strands being
held together by hydrogen bonding interactions of the type adenine to
thymine and guanine to cytosine (Fig 5.3)[72]. The carbonyl groups in
these compounds have an important function in producing these base
pairs as described by Watson and Crick (1953)[73] for the DNA double
helix.

Prior to 1950, each of the known classes of nucleic acids were
thought to consist only of four basic monomer nucleosides; this
concept proved to be quite naive and since then numerous additional
components have been found to be playing important roles in DNA and
RNA in differing organisms[74]. The compounds studied in this section,
although probably not incorporated directly into nucleic acids, have
characteristics similar to the pyrimidine bases in Figure 5.1 which
form the nucleoside monomers in such nucleic acids.

Figure 5.1   STRUCTURES OF COMPOUNDS REFERRED TO IN THIS SECTION
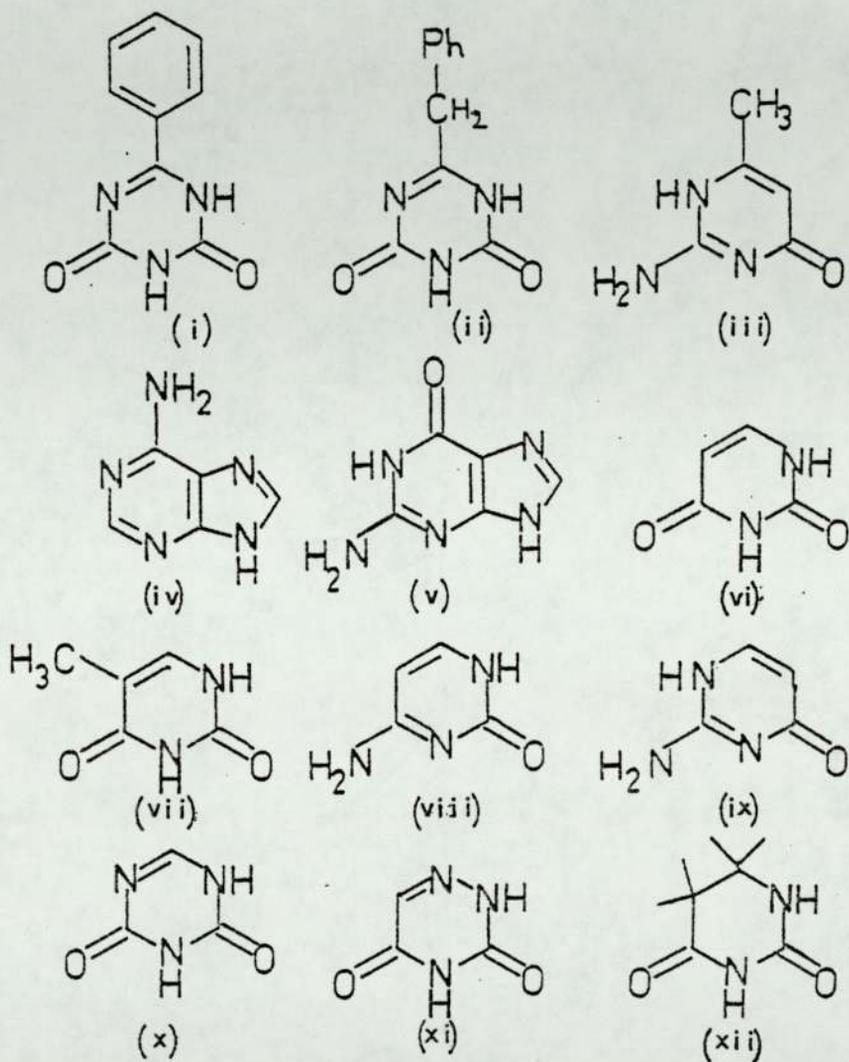
(i)	6-phenyl-5-azauracil	(viii)	cytosine
(ii)	6-benzyl-5-azauracil	(ix)	isocytosine
(iii)	6-methylisocytosine	(x)	5-azauracil
(iv)	adenine	(xi)	6-azauracil
(v)	guanine	(xii)	dihydrouracil
(vi)	uracil		
(vii)	thymine		

Figure 5.2   POLYMERIZATION OF DEOXYRIBONUCLEOTIDES INTO A
                 POLYNUCLEOTIDE STRAND OF DNA
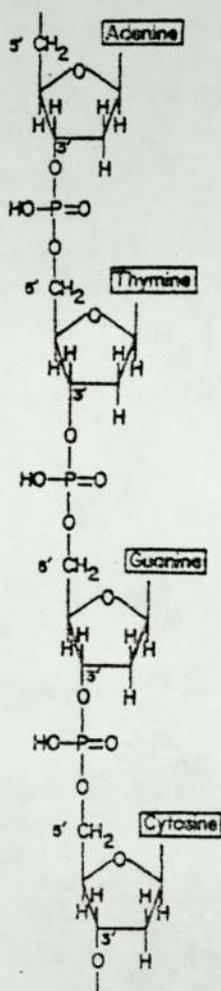
Figure 5.3a A DNA DOUBLE HELIX SHOWING THE PROCESS OF REPLICATION
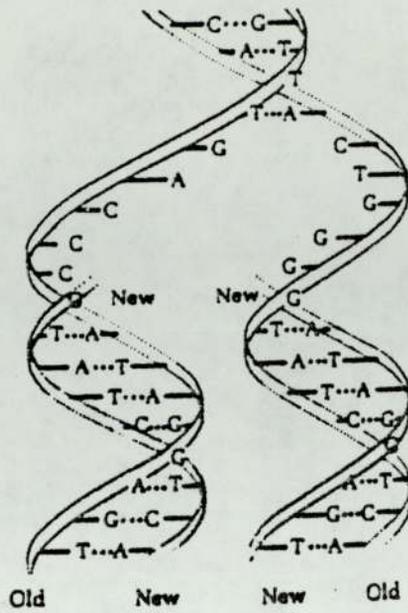
Figure 5.3b Hydrogen bonding between ADENINE and THYMINE and GUANINE and CYTOSINE
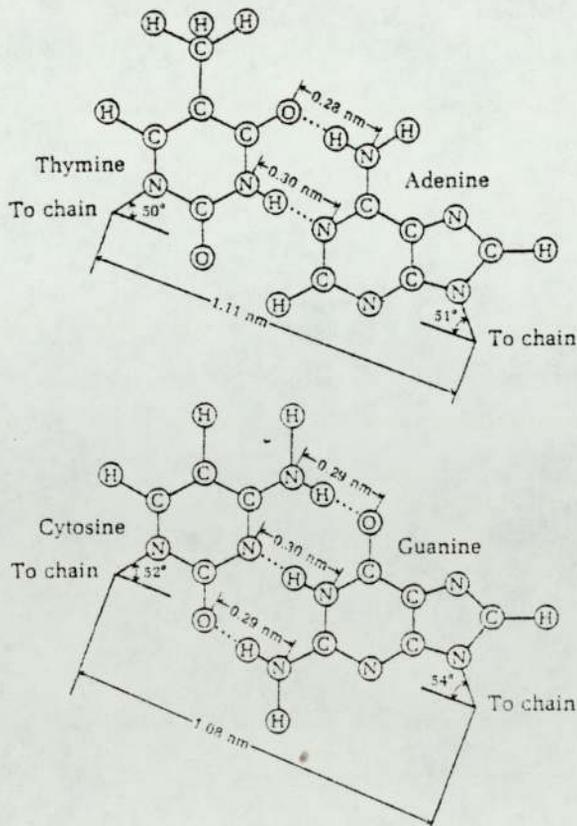
Two of the compounds, 6-phenyl-1,3,5,triazine-2,4[1H,3H]-dione and the 6-benzyl substituted analogue (Figures 5.1 (I) and (II)) are triazines and with the exception of a nitrogen atom at the N(5) position have the uracil configuration, and hence the trivial nomenclature of 6-phenyl and 6-benzyl-5-azauracils. The third compound, 6-methylisocytosine (Figure 5.1 (III)) has a pyrimidine ring which is an isomeric form of cytosine.

Although (I) and (II) have shown little in the way of biological activity, some azauracils have shown considerable promise in cancer chemotherapy. The parent compound 5-azauracil, for example, has reported activity as an antimetabolite[75] inhibiting the enzyme orotidylic pyrophosphorylase and thus blocking the production of orotic acid from orotidylic acid. This therefore inhibits pyrimidine biosynthesis (Figure 5.4)[76] and thus the production of nucleic acids; 6-azauracil has also been shown to have antineoplastic activity in animal tumours[77].

Since the carbonyl groups in such compounds may be implicated in base pairing and are likely to be important in binding to enzymes, their structures were determined to confirm the tautomeric forms suggested by spectroscopic analysis, and to compare the intermolecular interactions with those of similar compounds for which the structures have already been determined.

Figure 5.5 shows the possible tautomeric forms for 6-substituted-5-azauracils, and early workers favoured the mono-enol form since the metal salts of the compound contained only a single equivalent of a metal ion[78]. Later, however, a group of Czech workers published much IR and UV data on 1,3,5-triazine-diones and related compounds[79,80,81] which suggested that these compounds did in fact exist in the diketo form. A further noticeable feature of the

IR spectra was the occurrence of two peaks in the C=O stretching region, an effect which had been explained by Randall and co-workers[82] as being due to non-equivalence of the two carbonyl groups; and that the bands were, in fact, the symmetrical and antisymmetrical stretching modes[83], the antisymmetrical being the higher of the two frequencies in this case, was later proposed by Czech workers.

The choice between the theories of non-equivalence and vibrational coupling as a reason for the occurrence of two bands in the C=O stretching region requires the consideration of several points. To invoke a hypothesis of non-equivalence one would need to show a significant difference in the environment of the two carbonyl groups as would be manifested in their bond lengths, bond orders, lengths of adjacent bonds or masses of adjoining atoms or groups of atoms. Large differences in any of these, but especially in the carbonyl bond lengths may be sufficient to produce separate bands in the spectrum. Similarly, there are certain requirements for efficient vibrational coupling to be observed[83]. The two most important of these are that the bond lengths of the two carbonyls should be close in magnitude, (the closer they are, the more efficient the coupling) and that the two groups should be essentially co-planar; the greater the deviation from co-planarity, the weaker being the coupling effect.

In comparing the infrared spectra of 5-azauracil and 6-substituted-5-azauracils with analogous compounds, it was noticeable that the 5-aza compounds had significantly higher carbonyl stretching frequencies, suggesting a higher bond order in their case. With 6-phenyl-5-azauracil there was also some evidence of conjugation between the triazin-dione system and the phenyl ring, a feature which made it of additional interest. The 6-benzyl compound although

incapable of any conjugation between the two rings is also of interest because of its overall non-planarity. In these and related compounds, planarity appears to be an important feature where stacking and base-pairing are concerned and it was of considerable interest to see in such a non-planar molecule, what the effect would be.

6-methylisocytosine attracted interest from two points of view. Its possible tautomeric forms are shown in Figure 5.6 and since for isocytosine[84], tautomers of types (I) and (II) in a 1:1 ratio were found in the structure determination, it was important to establish what effect the 6-substituent had on this compound both in terms of its tautomeric form and intermolecular interactions. Second, if the molecule existed in the keto form as spectroscopic data suggested, it appeared to be an ideal control compound having just a single carbonyl group, to compare its carbonyl stretching frequency with those of the diketo compounds.

On the basis of this, the structures of the three compounds are reported.

Figure 5.4   BIOSYNTHESIS OF PYRIMIDINE NUCLEOTIDES
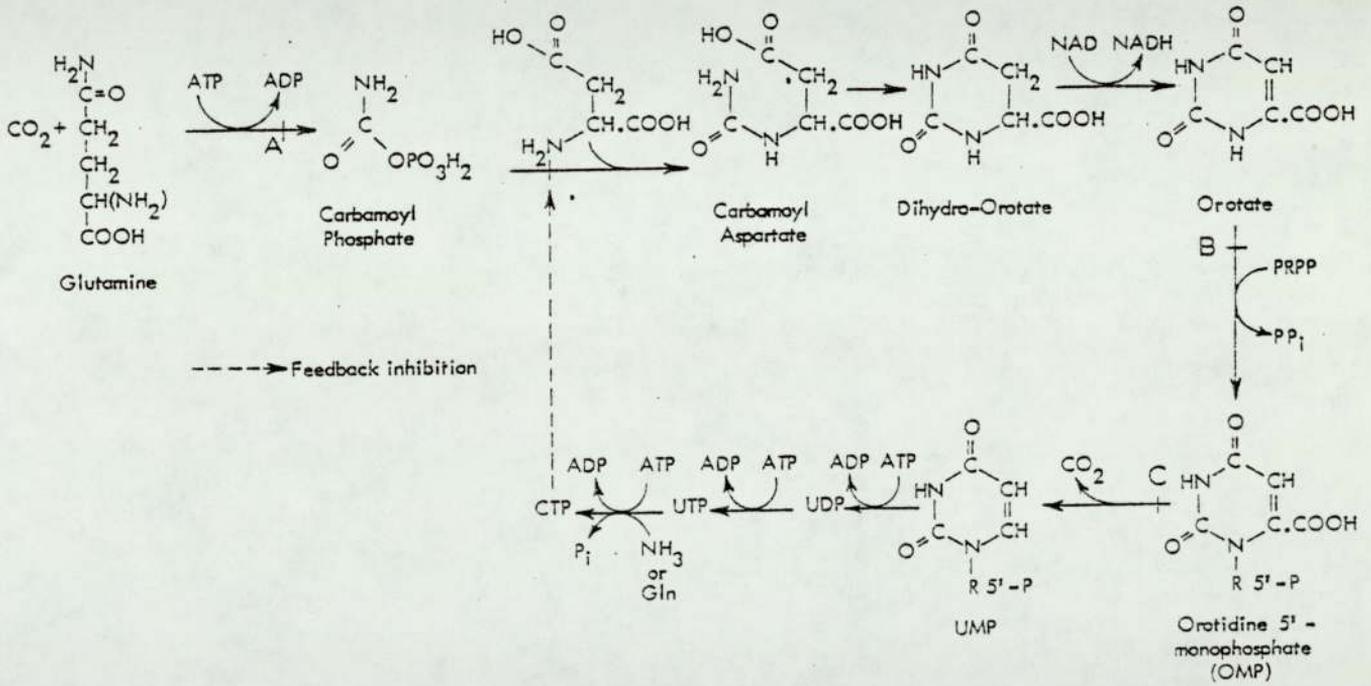
Figure 5.5 POSSIBLE TAUTOMERIC FORMS FOR 6-SUBSTITUTED-5-AZAURACILS
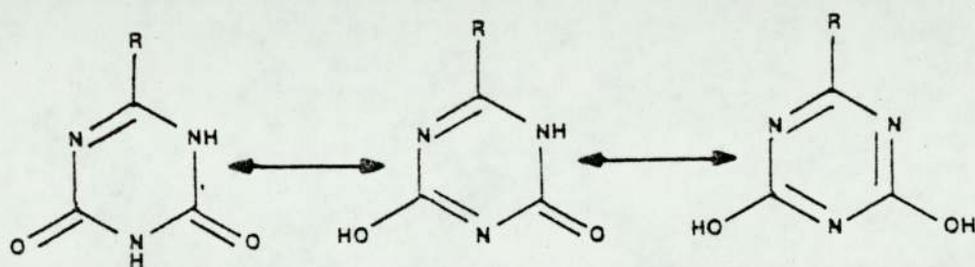
Figure 5.6   POSSIBLE TAUTOMERIC FORMS OF ISOCYTOSINE AND ITS 6-
SUBSTITUTED ANALOGUES
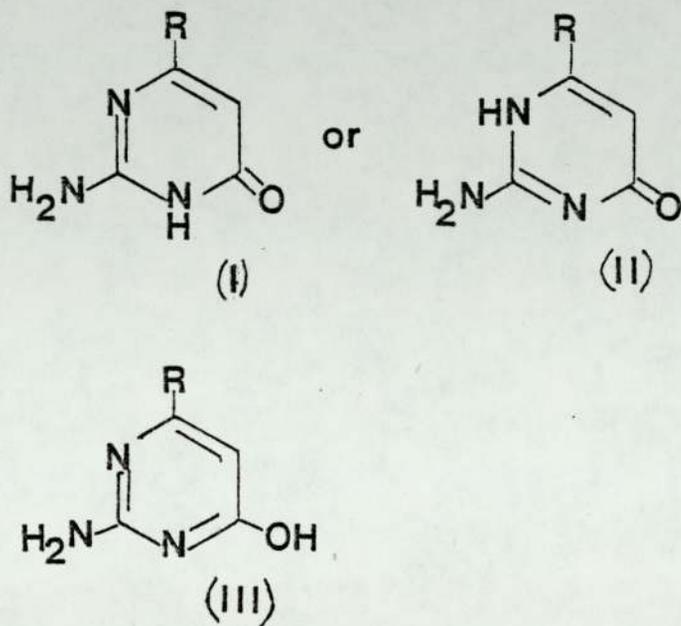
5.2   6-phenyl-1,3,5 triazin-2,4 [1H,3H]-dione

5.2.1   EXPERIMENTAL

The compound 6-phenyl-5-azauracil (trivial nomenclature based on the uracil configuration) was prepared by the treatment of the corresponding acyl biguanide with aqueous base[85]. Crystallisation from water produced colourless needles elongated along the b axis with a melting point of 293-294°C.

Prior to the department obtaining a four circle diffractometer, structure determination on this compound was undertaken on a 'Stoe Reciprocal Lattice Explorer' in the precession mode and also a Weissenberg camera. For the precession data, the camera was operated in the integrating mode with Zr filtered MoKα radiation (λ=0.71069Å) and films exposed singly for short, medium and long exposure times for each level. With the crystal mounted along the b axis, five levels of data were collected whilst precessing about the a axis (0kl to 4kl) and three levels whilst precessing about $\underline{c}$(hk0 to hk2). Integrated intensities were determined using a Joyce Loebl densitometer, and indices assigned to each reflection. With the Weissenberg films, five levels were taken in film packs of three (h0l to h4l), with Ni filtered CuKα radiation (λ =1.5418Å) and the films measured and indexed on the SRC scanning microdensitometer at the Atlas Computer Laboratory. Data reduction ensued with the two sets of data being treated separately for Lorentz and polarisation corrections; the data were then scaled together by means of the Hamilton Rollett and Sparkes algorithm[5]. The structure was solved with MULTAN[1] and in the final stages of refinement 804 unique reflections were used.

However, because of the rather poor R factor (0.105) presumably casting some doubt on the quality of the data, redetermination of the

structure was carried out on the advent of the CAD4 in the department. Whilst, from this point on, discussion of the structure will be confined to those parameters determined on the diffractometer, bond distances and angles and unit cell parameters determined from the film work will be included for comparison purposes. Infrared spectra were obtained on those compounds either made, or readily available in the laboratory, including 6-methyl-5-azauracil and uracil. These spectra were run as nujol mulls using a Pye Unicam SP200 scanning spectrophotometer. Data on other compounds documented, including solution spectra in dioxan, were obtained from the literature.

5.2.2 Crystal Data (Diffractometer determination)

6-phenyl-1,3,5-triazin-2,4(1H,3H)-dione ($C_9H_7N_3O_2$) crystallises in the monoclinic space group $P2_1/c$ with a=9.299(1)$\mathring{A}$, b=7.279(3)$\mathring{A}$, c=12.404(1)$\mathring{A}$, β=90.05(1)°, z=4 and v=839(11)A^3.

The molecular weight is 189.24 ($F_{(ooo)}$=392) and the density Dm (by flotation in butan-1-ol and carbon tetrachloride) =1.50 gm cm^{-3} with the calculated density Dc=1.490 g cm^{-3}. The absorption coefficient for MoKα radiation (λ =071069$\mathring{A}$) is μ=0.69 cm^{-1}.

Crystal Data (Film determination)

The early film work gave the following cell parameters:
a=9.27(3)$\mathring{A}$, b=7.28(3)$\mathring{A}$, c=12.35(3)$\mathring{A}$, β=90.4(2)°, z=4 and v=833A^3.

Systematic absences from Weissenberg films:- h0l when l is odd; 0k0 when k is odd.

5.2.3   Structure Analysis

Data Collection

The data were collected from a crystal of dimensions 0.7 x 0.3 x 0.15 mm, mounted along its needle axis, on an Enraf-Nonius CAD4 diffractometer, using the ω-2θ scan technique with monochromated Mo-Kα radiation (λ =0.71069Å). Two reference reflections were measured every 3600 secs of x-ray exposure time and the orientation of the crystal checked after every 100 reflections. No decomposition or movement of the crystal was detected during the collection of the 2079 unique reflections between θ=2 and θ=25°.

After completion of the data collection, those reflections used for the unit cell determination were examined. These showed some relatively high order reflections in k and l but only low order reflections in h. Accordingly, some higher order reflections in h were chosen from the data collection and substituted into this set. The cell dimensions were then refined using the new set of reflections, and are as given above under 'crystal data'.

Structure Determination and Refinement

The data reduction for the diffractometer data was performed at Birmingham University on the DEC 20 using a programme provided by Dr T A Hamor. Lorentz and polarisation and intensity corrections were applied and the data output in the form Hkl, Fobs and σFobs where σ Fobs is based on counting statistics. No correction to the data was made for absorption effects since for MoKα radiation, the maximum change in intensity along a straight path over the length of the crystal is only 4%. The change in intensity between the minimum and maximum paths through the crystal is of the order of 3%.

Normalised structure factors were calculated using SHELX[2] and the overall scale factor and estimated temperature factor were 0.7341 and 0.061 respectively. Phases were then determined using the SHELX automatic centrosymmetric direct methods package in which E min, the minimum value for a normalised structure factor used in the determination, was set to 1.1. NP the number of phase permutations was assigned the value of 20 (ie. 2^{20} permutations) and from this, a trial structure was produced in which all of the non-hydrogen atoms, and all but one of the hydrogens appeared on an Emap. The solution which gave the correct structure was that with the highest figures of merit and the programme assigned the following origin determining reflections:- $\bar{5}$ 5 9 (+); $\bar{3}$ 1 10 (+); $\bar{9}$ 2 1 (+).

Full matrix least squares refinement of all positional parameters and isotropic temperature factors by SHELX using unit weights, ensued, during which the programme assigned default values to all temperature factors and then refined them. After six cycles, $R = \sum [\ ||Fo| -S|Fc||\ ] / \sum |Fo|$ where S is a scale factor, decreased to 0.11. At this stage, the remaining hydrogen atom, as predicted in the structure, appeared on a difference Fourier synthesis.

A further five cycles of least squares refinement of coordinators and anisotropic thermal parameters for non-hydrogen atoms and coordinates and isotropic temperature factors for hydrogens, reduced R to 0.062. During this stage of refinement, reflections were weighted according to $w=1/\sigma^2 [Fo]$.

At this point, two reflections showing signs of serious extinction (2 0 0 and 1 2 1) were excluded from the data set and a final sequence of refinements for the complete structure using SHELX resulted in an unweighted discrepancy index R=0.039 for the 1421 observed data with Fo > 3σ. During the latter stage of refinement,

reflections were weighted according to $w=k/[\sigma^2(Fo) + qFo^2]$ where q, a refinable parameter, converged at 0.0062 and k which should approach unity if there are no serious systematic errors equalled 1.00.

The refinement was terminated when no positional parameter shifted by more than 0.22σ and the weighted discrepancy index Rg=0.068.

A final difference electron density map showed no feature greater than $0.15e\text{Å}^{-3}$.

During refinement, atoms were assigned their correct atomic scattering factors on the basis of the structure proposed from IR, NMR, MS and CHNO analysis. The isotropic temperature factors during refinement and the final discrepancy index and difference electron density map appear to confirm the correctness of these assignments.

Table 5.1   Positional parameters (fractional co-ordinates x 10^4) with estimated standard deviations in parentheses.

	X/a	Y/b	Z/c
N(1)	3190(1)	1865(2)	5752(1)
C(2)	1892(1)	1124(2)	5435(1)
N(3)	1769(1)	868(2)	4349(1)
C(4)	2814(2)	1349(2)	3605(1)
N(5)	4121(1)	1956(2)	3996(1)
C(6)	4263(1)	2212(1)	5025(1)
O(2)	955(1)	751(2)	6080(1)
O(4)	2564(1)	1202(2)	2645(1)
C(1)'	5660(1)	2877(2)	5452(1)
C(2)'	6707(2)	3432(2)	4719(1)
C(3)'	8036(2)	4010(3)	5086(2)
C(4)'	8330(2)	4072(3)	6177(2)
C(5)'	7284(2)	3554(3)	6907(1)
C(6)'	5956(2)	2945(2)	6551(1)
H(1)	3178(21)	2284(26)	6481(16)
H(3)	925(26)	356(28)	4141(17)
H(2)'	6484(20)	3388(27)	3962(16)
H(3)'	8722(28)	4424(31)	4554(20)
H(4)'	9301(29)	4539(31)	6390(21)
H(5)'	7481(24)	3737(35)	7716(20)
H(6)'	5254(23)	2479(31)	7110(16)

TABLE 5.2   Anisotropic temperature factors (non-hydrogen atoms)

Isotropic temperature factors (hydrogen atoms) with standard

deviations in parentheses

ATOM	U11	U22	U33	U23	U13	U12
N(1)	.0243(6)	.0390(2)	.0215(6)	-.0012(5)	.0021(4)	-.0044(5)
C(2)	.0249(7)	.0328(7)	.0258(7)	.0025(5)	-.0001(5)	-.0018(5)
N(3)	.0277(7)	.0413(7)	.0234(6)	.0007(5)	-.0002(5)	-.0073(5)
C(4)	.0322(8)	.0319(7)	.0230(6)	.0022(5)	.0019(5)	-.0013(6)
N(5)	.0292(6)	.0365(7)	.0247(6)	-.0002(5)	.0042(5)	-.0047(5)
C(6)	.0260(7)	.0247(7)	.0256(6)	.0013(5)	.0043(5)	.0002(5)
O(2)	.0271(6)	.0627(8)	.0275(5)	.0003(5)	.0034(4)	-.0115(5)
O(4)	.0436(7)	.0568(8)	.0210(5)	.0024(5)	-.0019(4)	-.0087(5)
C(1)'	.0241(7)	.0257(7)	.0322(8)	.0000(5)	.0032(5)	-.0013(5)
C(2)'	.0322(8)	.0400(8)	.0331(8)	.0008(6)	.0070(6)	-.0041(6)
C(3)'	.0301(8)	.0532(10)	.0515(10)	.0049(8)	.0104(7)	-.0097(7)
C(4)'	.0292(8)	.0473(10)	.0580(11)	-.0008(8)	-.0041(7)	-.0110(7)
C(5)'	.0361(9)	.0516(10)	.0378(9)	-.0019(7)	-.0063(7)	-.0088(7)
C(6)'	.0292(7)	.0446(9)	.0324(8)	-.0006(7)	.0025(6)	-.0062(6)

	Uiso
H(1)	.0493(53)
H(3)	.0485(54)
H(2)'	.0412(50)
H(3)'	.0589(63)
H(4)'	.0636(65)
H(5)'	.0632(69)
H(6)'	.0482(54)

Figure 5.7 Interatomic angles [°] as determined by film methods (mean estimated standard deviations 0.5°)

Table 5.3   Bond distances [Å] as determined by film methods

Bond	Dist Å	Bond	Dist Å
N(1)-C(2)	1.361(7)	C(6)-C(1)'	1.488(8)
C(2)-N(3)	1.360(7)	C(1)'-C(2)'	1.396(8)
N(3)-C(4)	1.390(7)	C(2)'-C(3)'	1.361(9)
C(4)-N(5)	1.378(8)	C(3)'-C(4)'	1.371(9)
N(5)-C(6)	1.286(7)	C(4)'-C(5)'	1.375(9)
C(6)-N(1)	1.362(7)	C(5)'-C(6)'	1.375(9)
C(2)-0(2)	1.202(7)	C(6)'-C(1)'	1.393(8)
C(4)-0(4)	1.203(6)	N(1)-H(1)	0.86
		N(3)-H(3)	0.92

Table 5.4 Bond distances in Å with estimated standard deviations in parentheses.

Bond	Interatomic distance Å
N(1)-C(2)	1.379(2)
C(2)-N(3)	1.364(2)
N(3)-C(4)	1.387(2)
C(4)-N(5)	1.381(2)
N(5)-C(6)	1.296(2)
C(6)-N(1)	1.369(2)
C(2)-O(2)	1.214(2)
C(4)-O(4)	1.218(2)
C(6)-C(1)'	1.483(2)
C(1)'-C(2)'	1.392(2)
C(2)'-C(3)'	1.383(2)
C(3)'-C(4)'	1.381(3)
C(4)'-C(5)'	1.382(3)
C(5)'-C(6)'	1.384(2)
C(6)'-C(1)'	1.392(2)
N(1)-H(1)	0.96(2)
N(3)-H(3)	0.91(2)
C(2)'-H(2)'	0.96(2)
C(3)'-H(3)'	0.97(3)
C(4)'-H(4)'	1.00(3)
C(5)'-H(5)'	1.03(2)
C(6)'-H(6)'	1.01(2)

Table 5.5 Interatomic angles [°] with estimated standard deviations in parentheses (for non-hydrogen atoms, and those hydrogens attached to nitrogen atoms)

Atoms	Bond Angle [°]
N(1)-C(2)-N(3)	114.0(1)
N(1)-C(2)-O(2)	121.9(1)
O(2)-C(2)-N(3)	124.1(1)
C(2)-N(3)-C(4)	124.4(1)
N(3)-C(4)-N(5)	117.6(1)
N(3)-C(4)-O(4)	119.7(1)
O(4)-C(4)-N(5)	122.7(1)
C(4)-N(5)-C(6)	118.8(1)
N(5)-C(6)-N(1)	123.2(1)
C(6)-N(1)-C(2)	121.6(1)
N(5)-C(6)-C(1)'	119.1(1)
N(1)-C(6)-C(1)'	117.7(1)
C(6)-C(1)'-C(2)'	118.3(1)
C(6)-C(1)'-C(6)'	122.2(1)
C(1)'-C(2)'-C(3)'	119.9(2)
C(2)'-C(3)'-C(4)'	120.6(2)
C(3)'-C(4)'-C(5)'	119.7(2)
C(4)'-C(5)'-C(6)'	120.4(2)
C(5)'-C(6)'-C(1)'	120.0(1)
C(6)'-C(1)'-C(2)'	119.5(1)
H(1)-N(1)-C(2)	113(1)
H(1)-N(1)-C(6)	125(1)
C(2)-N(3)-H(3)	114(1)
H(3)-N(3)-C(4)	122(1)

Table 5.6 Deviations of non-hydrogen atoms in $\overset{o}{A}$ from the least squares planes through (a) the triazine ring and (b) the phenyl ring.[+] (Atoms used in the plane calculation are marked with an asterisk).

(a)

ATOM	DEVIATION [$\overset{o}{A}$]	ATOM	DEVIATION [$\overset{o}{A}$]
*N(1)	.028	O(4)	.129
*C(2)	-.011	C(1)'	-.068
*N(3)	-.023	C(2)'	.043
*C(4)	.040	C(3)'	-.044
*N(5)	-.024	C(4)'	-.220
*C(6)	-.011	C(5)'	-.306
O(2)	-.025	C(6)'	-.241

[+] The equation of the plane is:-

$-.355X + .931Y - .089Z - .451 = 0$ where X, Y and Z are orthogonal coordinates in $\overset{o}{A}$ along a*, b and c.

Table 5.6 (b)

ATOM	DEVIATION [Å]	ATOM	DEVIATION [Å]
N(1)	.103	O(4)	.293
C(2)	-.005	*C(1)'	-.005
N(3)	-.191	*C(2)'	.008
C(4)	-.228	*C(3)'	-.003
N(5)	-.211	*C(4)'	-.005
C(6)	-.035	*C(5)'	-.007
O(2)	.067	*C(6)'	-.003

+ The equation of the plane is:-

-.342X + .940Y + .037Z + .4193 = 0   where X, Y and Z are

orthogonal coordinates in Å along a*, b and c.

Table 5.7 Important torsion angles [°]

ATOMS	ANGLE [°]
C(6)-N(1)-C(2)-N(3)	-3.0
N(1)-C(2)-N(3)-C(4)	-2.0
C(2)-N(3)-C(4)-N(5)	6.8
N(3)-C(4)-N(5)-C(6)	-6.4
C(4)-N(5)-C(6)-N(1)	1.7
N(5)-C(6)-N(1)-C(2)	3.3
C(6)-N(1)-C(2)-O(2)	177.3
O(2)-C(2)-N(3)-C(4)	177.7
C(2)-N(3)-C(4)-O(4)	174.2
O(4)-C(4)-N(5)-C(6)	174.6
N(1)-C(6)-C(1)'-C(6)'	7.7
N(1)-C(6)-C(1)'-C(2)'	-172.9
N(5)-C(6)-C(1)'-C(2)'	8.3
N(5)-C(6)-C(1)'-C(6)'	171.1

The torsion angle A-B-C-D is the projected angle between AB and CD when viewed down the B-C bond; the clockwise rotation of the C-D bond with reference to bond A-B is considered positive.

Table 5.8 Hydrogen bond contact distances and angles

HYDROGEN BOND	ANGLE AT H[°]	N·········O DISTANCE $\overset{o}{A}$
$N(1)_{II}$-H(1) ..... $O(4)_I$	155	2.799
$N(3)_{III}$-H(3) .... $O(2)_I$	169	2.843

I, II, III refer the the symmetry operations x, y, z; x, $\frac{1}{2}$-y, $\frac{1}{2}$-z;

-x, -y, 1-z

Table 5.9

Important intermolecular contacts

$$C(2)'_I - C(6)_{III} \quad - \quad 3.31\overset{o}{A}$$
$$C(2)'_I - C(2)_{II} \quad - \quad 3.58\overset{o}{A}$$
$$C(4)'_I - C(4)_{III} \quad - \quad 3.49\overset{o}{A}$$
$$C(6)_I - C(6)_{II} \quad - \quad 3.51\overset{o}{A}$$

The subscripts refer to the following equivalent positions:-

I   x, y, z

II 1-x, -y, 1-z

III 1-x, 1-y, 1-z

Figure 5.7a Molecular structure and numbering scheme for 6-phenyl-
1,3,5 triazin-2,4[1H,3H] dione
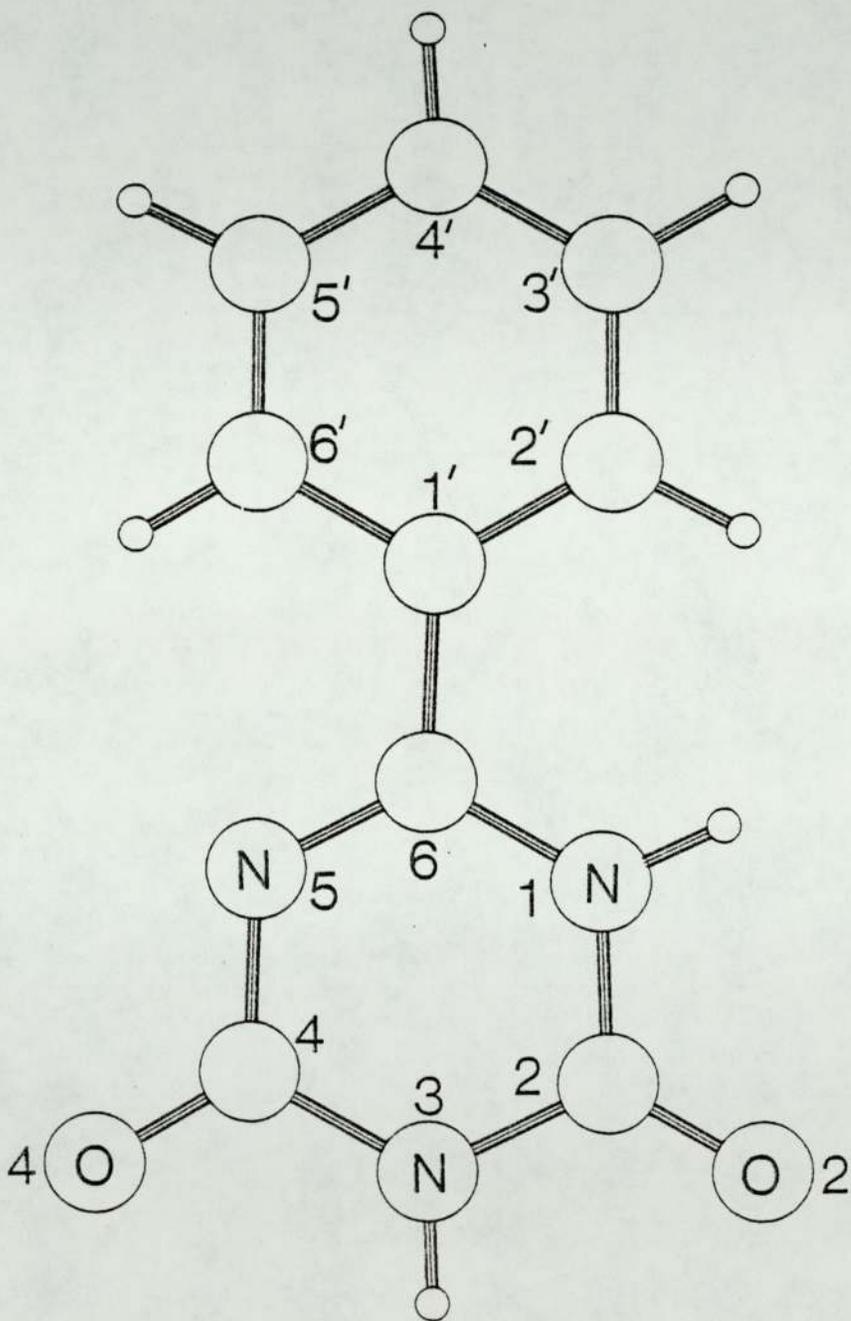
Figure 5.8a Stereo diagram of 6-phenyl-1,3,5,—triazin-2,4[1H,3H]-dione
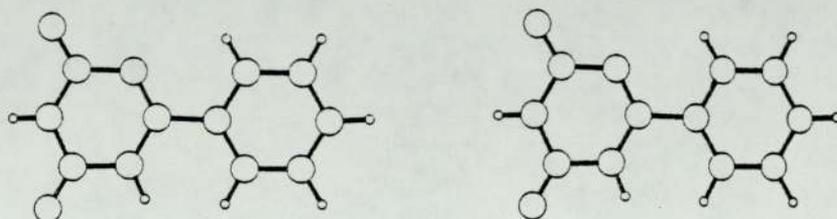
Figure 5.8b Stereo packing diagram
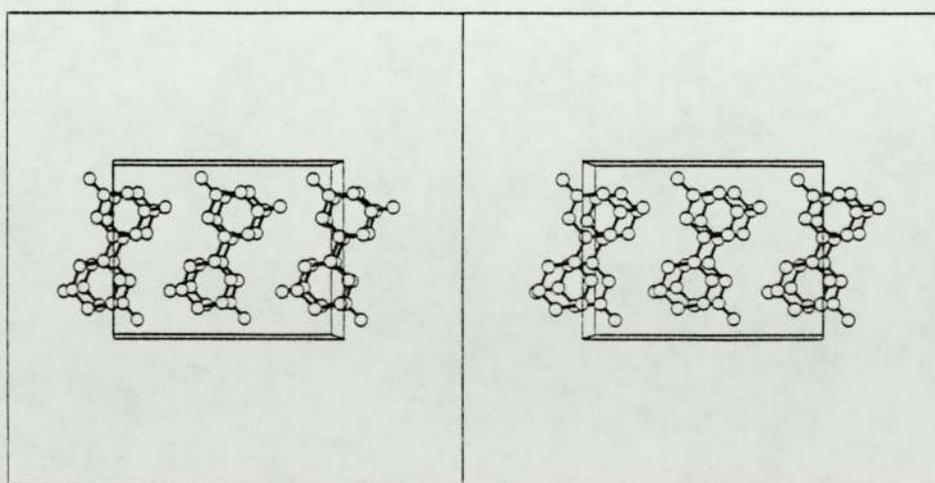
Figure 5.9   Packing diagram of 6-phenyl-1,3,5—triazin-2,4[1H,3H]-
dione
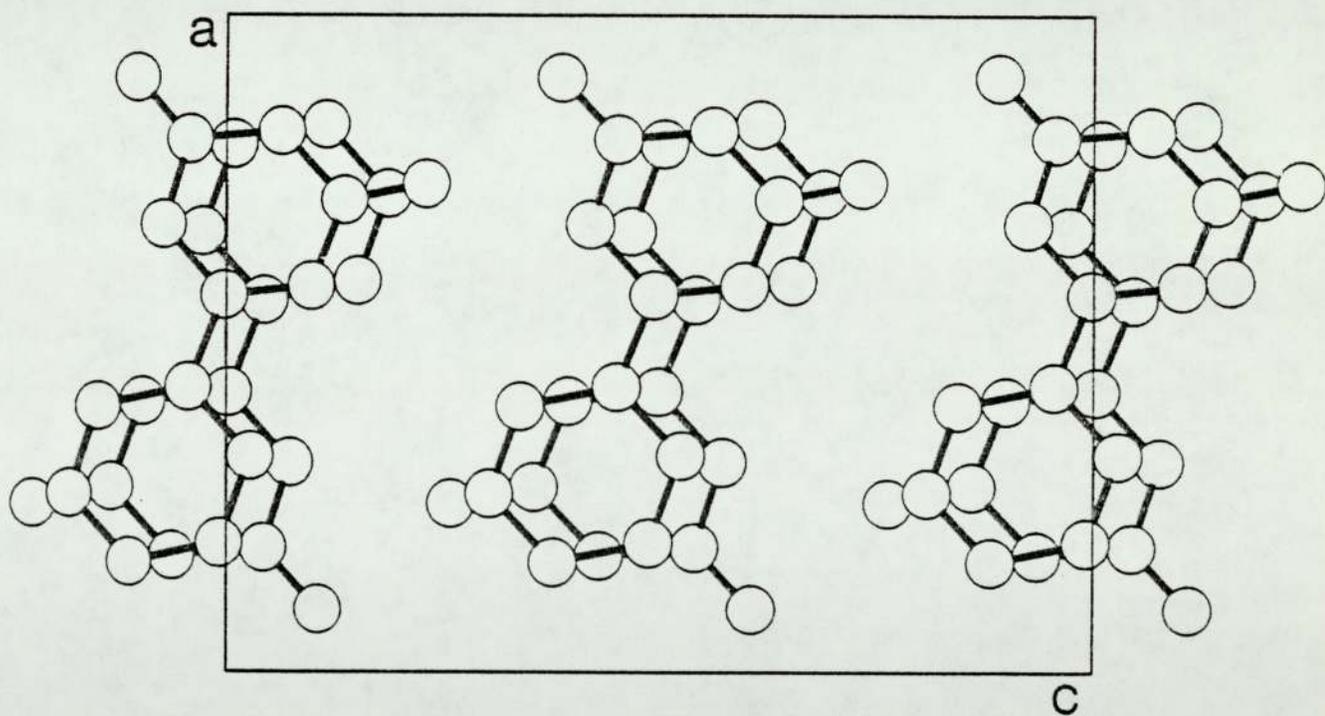
Figure 5.10 Packing diagram showing hydrogen bonding in the a-c plane. (Hydrogen bonds are shown in dotted lines).
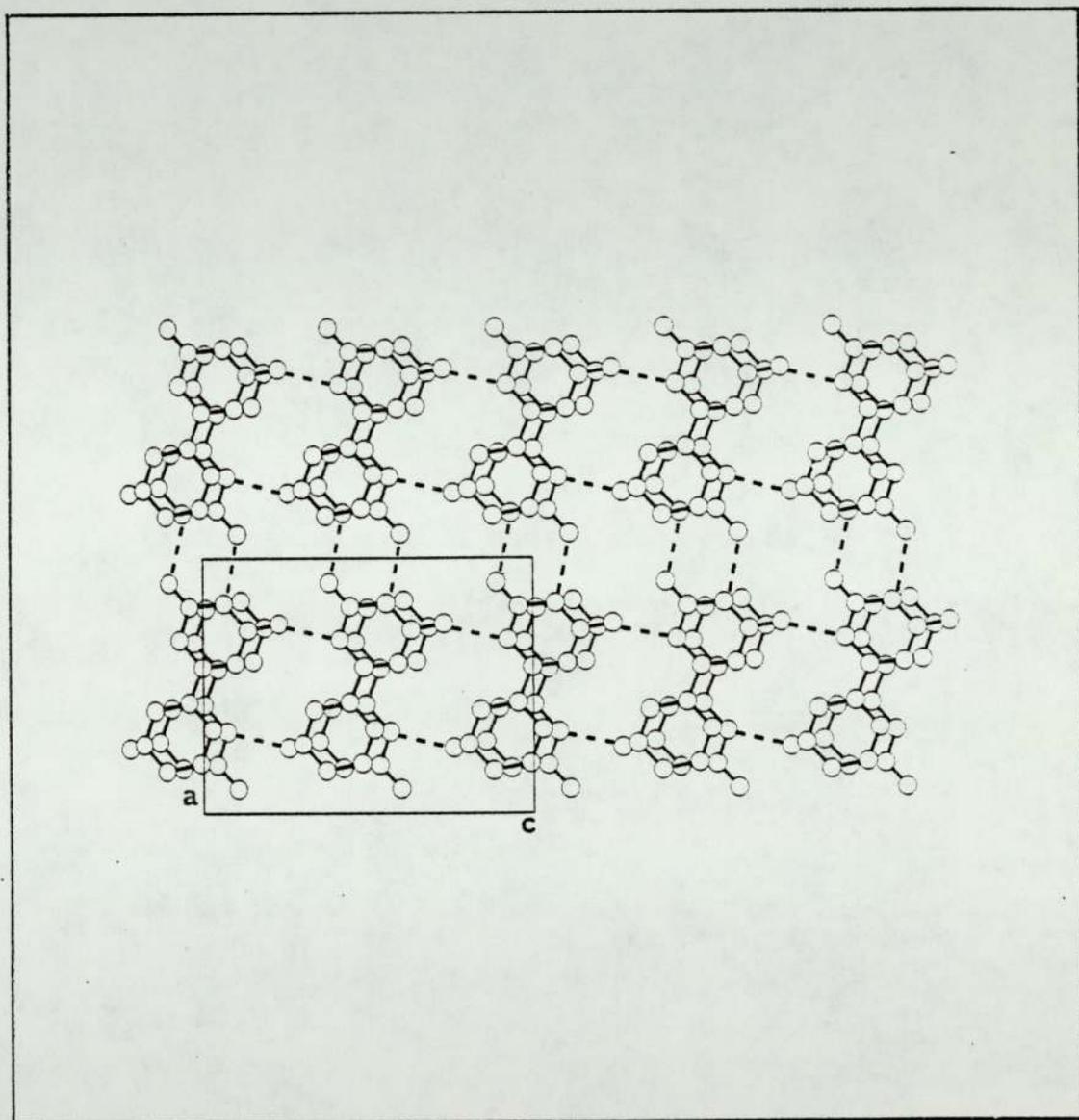
Figure 5.11 Packing diagram showing hydrogen bonding in the b-c plane. (Hydrogen bonds are shown in dotted lines).
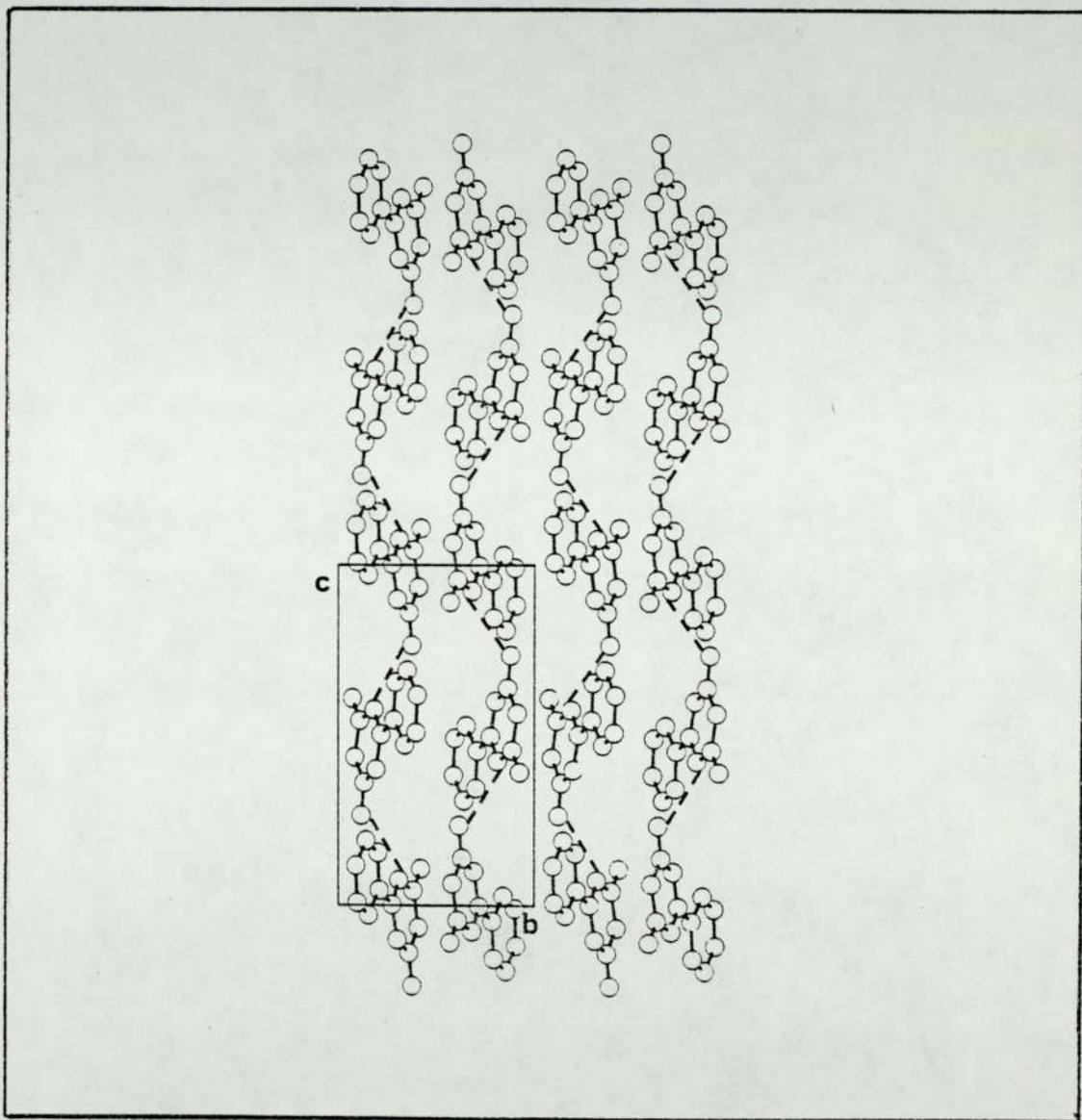
Figure 5.12 Space filling diagram of 6-phenyl-1,3,5—triazin-2,4[1H,3H]-dione
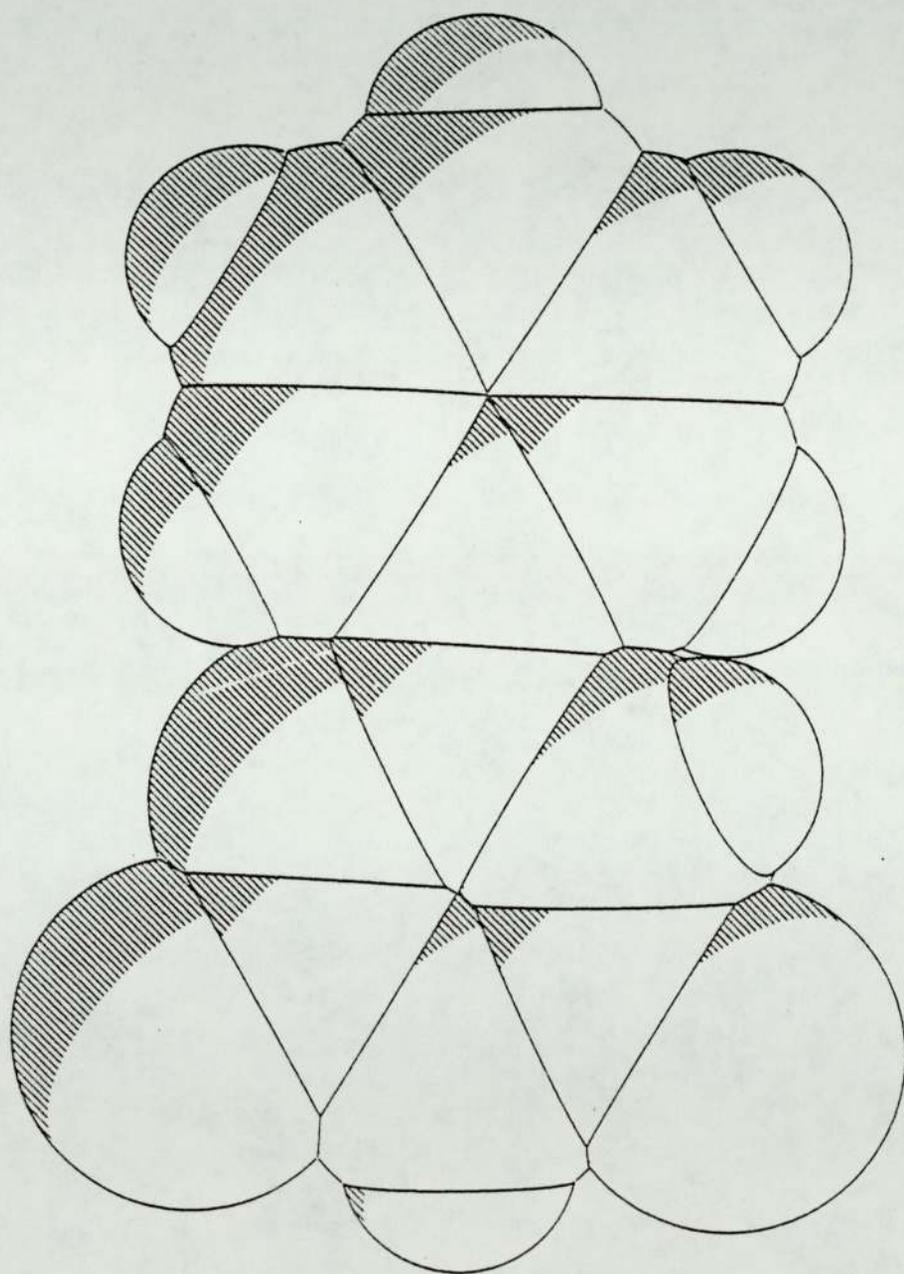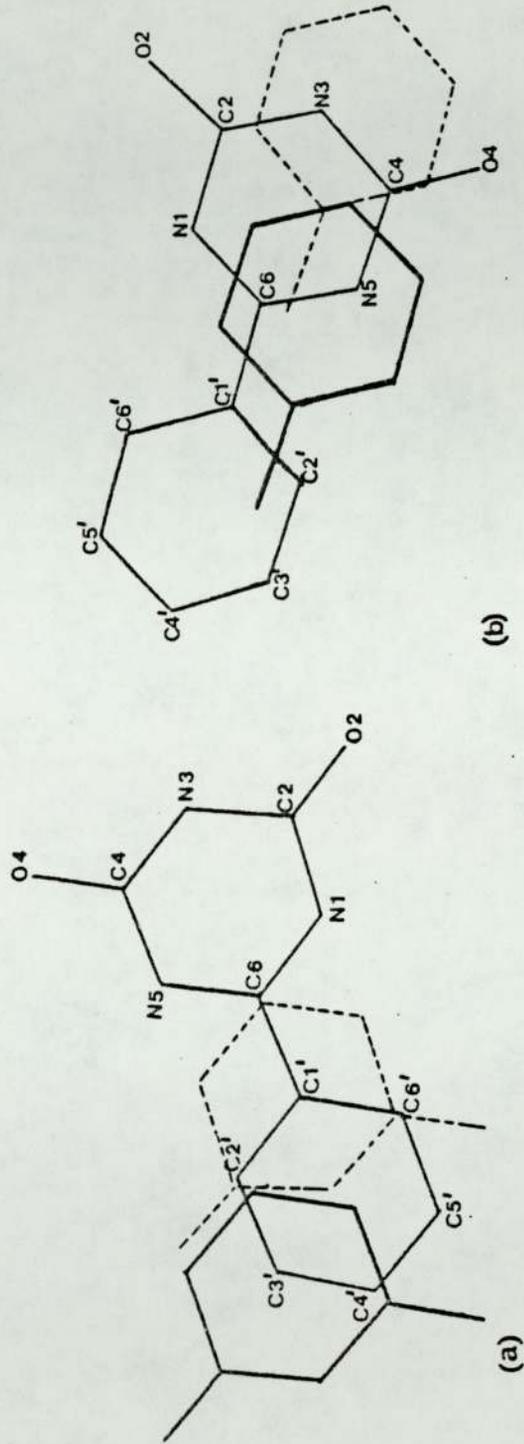
Figure 5.13 Stacking arrangement of 6-phenyl-1,3,5 triazin-2,4[1H,3H]-dione

(a) as projected onto the phenyl plane

(b) as projected onto the triazine plane

Resonance structures for 6-substituted-5-azauracils

a) For any R (eg. 6-phenyl-5-azauracil and 6-benzyl-5-azauracil)
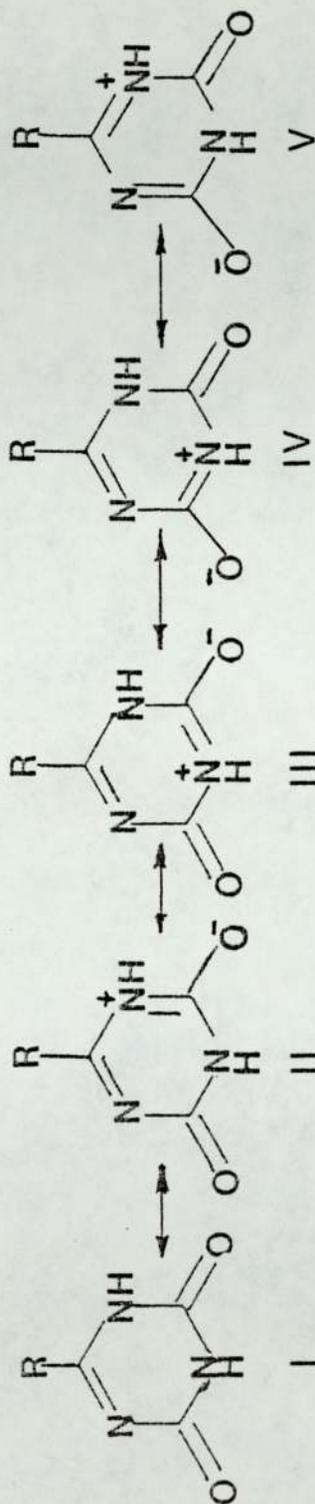
b) Additional resonance structures if R can conjugate (ie. 6-phenyl-5-azauracil)
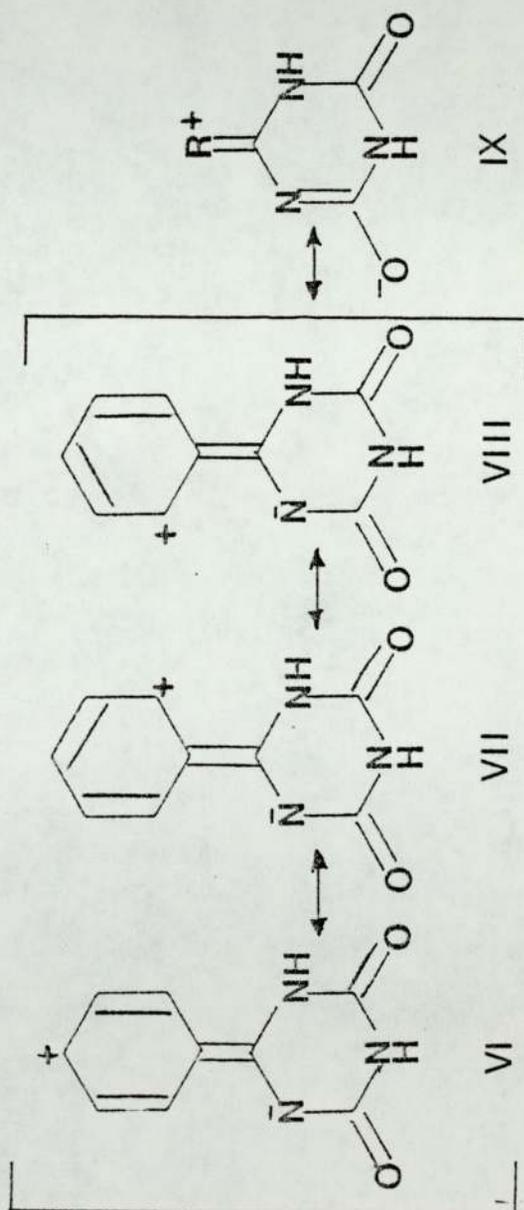
Figure 5.14

5.2.4   RESULTS AND DISCUSSION

The   structure   of   6-phenyl-5-azauracil   as   determined crystallographically is shown in Figure 5.7 together with the numbering scheme used; positional parameters for all atoms are given in Table 5.1, with anisotropic thermal parameters for non hydrogen atoms, and isotropic temperature factors for hydrogens in Table 5.2.

The bond distances and interatomic angles are given in Tables 5.4 and 5 respectively. Inspection of the two carbon to oxygen bonds C(2)-O(2) and C(4)-O(4) shows them to be almost equal in length (1.214(2)$\overset{o}{A}$ and 1.218(2)$\overset{o}{A}$) and on average somewhat shorter than found in the analogous compounds 6-azauracil[90], 6-benzyl-5-azauracil and uracil[87]. The bond lengths calculated for these two bonds suggest they are of double bond character and therefore, that in the crystalline state the molecule exists in the diketo form as opposed to the mono-enol form suggested by earlier workers[78]. The N(1)-C(2) and C(4)-N(5) bond distances (1.379(2)$\overset{o}{A}$ and 1.381(2)$\overset{o}{A}$) are similar in magnitude, the length of the latter suggesting little conjugation between C(6), N(5), C(4) and O(4). The N(3)-C(4) and C(2)-N(3) bonds (1.387(2)$\overset{o}{A}$ and 1.364(2)$\overset{o}{A}$) albeit similar, do show a significant variation. The exact reason for the reduced bond distance of the latter is at present uncertain. One possibility considered was the conjugation of the N(3) lone pair electrons and the carbonyl group C(2)-O(2). Equally likely however, would be a similar contribution from the N(1) lone pair, and since N(5) is unable to participate in a similar fashion, one might have expected N(1) lone pair to conjugate with C(2)-O(2) and N(3) lone pair with C(4)-O(4) thus contradicting the argument. Further evidence which refutes this initial hypothesis are the carbonyl bond distances. Since both are hydrogen bonded, if the N(3) non bonded electrons were conjugated with the C(2)-O(2)

system one might have expected this to be elongated with respect to C(4)-O(4). This is not, however, so and indeed the reverse is the case. The equivalence of the residual charges on the two carbonyl oxygens as determined by molecular orbital calculations with GAUSSIAN 70 in the STO-3G approximation further substantiates this; (O(2) -0.46, O(4) -0.47).

A second and possible hypothesis involves the nature of the intermolecular hydrogen bonding. The N(3) and O(2) atoms form a base pair arrangement with corresponding atoms on a molecule related by a centre of symmetry. It is possible therefore that such an arrangement could place a physical constraint on the enclosed C(2)-N(3) bond. This is not the case for the O(4) oxygen which participates in a single hydrogen bond with a neighbouring molecule. A similar shortening of the equivalent bond distance occurs in 6-azauracil[90] which also exhibits the same form of hydrogen bonding arrangement. The main difference in bonding in the triazine ring exists between the N(5)-C(6) and N(1)-C(6) bond distances (1.296(2)$\overset{o}{A}$ and 1.369(2)$\overset{o}{A}$), the former showing true double bond character and thus suggesting the correct location of the hydrogen atom at N(1).

The bond lengths in the phenyl ring and the length of the C(6)-C(1)' bridging bond show some evidence of conjugation between the two rings albeit quite small. The average length of the C(1)'-C(2)' and C(1)'-C(6)' bonds compared with the averages of the other pairs of symmetrically opposite bonds, is of the order of 0.01$\overset{o}{A}$ more, thus adding weight to the resonance forms (VI) (VII) and (VIII) for this structure as depicted in Figure 5.14b. This suggestion is further enhanced by examining the C(6)-C(1)' bridging bond (1.483(2)$\overset{o}{A}$), which is somewhat shorter than the bridging bond in biphenyl (1.496$\overset{o}{A}$) which

may be considered representative of a true $C(Sp^2)-C(Sp^2)$ single bond[54]. The degree to which electrons are attracted by the triazine ring from the benzene ring, however, seems limited, as can be seen from the bond distances in the triazine ring and the major resonance form consistant with this is that of (1) in Figure 5.14a with there being some small contribution from (VI), (VII) and (VIII) in Figure 5.14b. The bond angles in the molecule are shown in Table 5.5. The internal bond angles in the phenyl ring are all, as expected, close to 120°. The important angles are, however, the nitrogen valence angles in the triazine ring, since these provide additional evidence for the correct location of the hydrogen atoms on the triazine nitrogens. From bond length data it is suggested that there is an extra-annular attachment at N(1) and N(3) nitrogens, but not at N(5); this being further substantiated by the location of these hydrogens by a difference Fourier synthesis. The nitrogen valence angles in the triazine ring are 121.6(1)°, 124.4(1)° and 118.8(1)° for N(1), N(3) and N(5) respectively. They can be interpreted according to Chatar Singh's Rule[59] which is basically an empirical form of the Valence Shell Electron Pair Repulsion Theory and states that 'Nitrogens in a six membered heterocyclic ring with an extra-annular hydrogen atom have nitrogen valence angles within the range 125±3°, whilst those without extra-annular hydrogens have angles within the range 116±3°'. This suggests that the location of the hydrogens at N(1) and N(3) but not at N(5) is correct.

In the crystalline state, 6-phenyl-5-azauracil is essentially planar as can be seen from the stereo diagrams (Figures 5.8a & b). The angle between the least squares planes through the phenyl and triazine rings is only 6.9°. Deviations in Å from the least squares plane through the two rings (Table 5.6) show all of the carbon atoms

in the phenyl ring to lie within $\pm 0.01\overset{\text{o}}{\text{A}}$ of the phenyl plane, and the triazine ring atoms to lie within $\pm 0.04\overset{\text{o}}{\text{A}}$ of the triazine plane. The deviation of the O(2) and O(4) carbonyl oxygens from the triazine plane are $-0.025\overset{\text{o}}{\text{A}}$ and $0.129\overset{\text{o}}{\text{A}}$ respectively. The planarity of the triazine ring and the coplanarity of the carbonyl groups with it, is an important factor where vibrational coupling between the two carbonyls is concerned and the small twist angle between the two planes would be conducive to a degree of conjugation between the two rings.

The packing arrangement of the molecules is shown in Figures 5.9, 10 and 11. The stacking takes the form of infinite columns of molecules, phenyl and triazine rings alternating, along the b axis; this is shown projected a) onto the phenyl plane, and b) onto the triazine plane in Figure 5.13, and from this important intermolecular contacts can be seen. These are shown in Table 5.9 and with the exception of the slightly short C(2)'-C(6) $(3.312\overset{\text{o}}{\text{A}})$ are indicative of distances corresponding to normal Van der Waals electrostatic interaction.

The suggested hydrogen bonds (shown as dotted lines in Figures 10 and 11) are an important feature of the packing, and the distances and angles that characterise these bonds together with the symmetry operations are given in Table 5.8. Both carbonyl oxygens participate in strong hydrogen bonds. The N(1)-H(1) .... O(4) interaction $(2.799\overset{\text{o}}{\text{A}})$ links chains of molecules related by the glide plane, along the c axis, whilst the N(3)-H(3) .... O(2) interaction $(2.843\overset{\text{o}}{\text{A}})$ produces a dimer arrangement across a centre of symmetry along the a axis, thus forming a base pair. The net result is that the triazine ring is firmly anchored in four positions, a feature possibly reflected in its high melting point of 293-294 °C.

The parameters as determined by film methods, Figure 5.7 and Table 5.3, albeit having higher standard deviations, compare well with those found from diffractometer data. The only barely significant differences in bond lengths occur in the N(1)-C(2) bond (1.361(7)$\overset{o}{A}$ film 1.379(2)$\overset{o}{A}$ diffractometer) the two carbonyl bond distances which appear slightly shorter from the film data, and the C(2)'-C(3)' bond (1.361(9)$\overset{o}{A}$ film and 1.383(2)$\overset{o}{A}$ diffractometer). It should be added that although there are differences, the conclusions drawn from the two structure determinations would at the most be only marginally different.

5.3   6-BENZYL-1,3,5-TRIAZIN-2,4(1H,3H)-DIONE

5.3.1   EXPERIMENTAL

6-benzyl-5-azauracil (trivial nomenclature based on the uracil configuration) was prepared by Traynor[85] by the treatment of the corresponding acyl biguanide with aqueous base.   On crystallising from aqueous ethanol, this produced colourless needles with a melting point of 254-255 °C.

The infrared spectrum of the compound was recorded as a nujol mull on a Pye Unicam SP200 spectrophotometer.

5.3.2   Crystal Data

6-benzyl-1,3,5-triazin-2,4($1\underline{H}$,$3\underline{H}$)-dione ($C_{10}H_9N_3O_2$) crystallises in the monoclinic space group $P2_1/c$ with a = 9.712(2)$\overset{o}{A}$, b = 8.310(1)$\overset{o}{A}$, c = 12.661(5)$\overset{o}{A}$, β = 109.18(3)° z = 4 and v = 965(1)$\overset{o}{A}^3$.

The molecular weight is 203.27 (F(ooo) = 424) and the density Dc (calculated) = 1.390 gcm^{-3}.   The absorption coefficient for MoKα radiation (λ = 0.71069$\overset{o}{A}$) is μ = 0.62 cm^{-1}.

5.3.3   Structure Analysis

The data were collected from a crystal of dimensions 0.80 x 0.28 x 0.24 mm mounted along its longest dimension on an Enraf-Nonius CAD4 diffractometer using the ω-2θ scan technique with monochromated MoKα radiation (λ = 0.71069$\overset{o}{A}$).   Two reference reflections were measured every 7200 sec of x-ray exposure time and the orientation of the crystal was checked after every 150 reflections.   No decomposition or movement of the crystal was detected during the collection of the 1775 observed reflections between θ = 2 and θ = 25°.

Structure Determination and Refinement

The data reduction for the diffractometer data was performed at Birmingham University on the DEC 20 using a programme provided by Dr T A Hamor. Lorentz and polarisation and intensity corrections were applied, and the data output in the form of HKL Fobs and σFobs where σ Fobs is based on counting statistics.

No correction to the data was made for absorption effects, since, for MoKα radiation, the maximum change in intensity along a straight path over the length of the crystal (ie. $\frac{Io - I}{Io}$ x 100) is only 4.8%. The change in intensity between the minimum and maximum paths through the crystal is of the order of 3%.

Normalised structure factors were calculated using SHELX[2] and the overall scale factor and estimated temperature factor determined as 1.1300 and 0.064 respectively.

Phases were determined using the SHELX automatic centrosymmetric direct methods package during which Emin, the minimum value for a normalised structure factor used in the determination, was set to 1.1, and NP, the number of phase permutations, was assigned the value of 14 (ie. 2^{14} permutations). From this, a trial structure was produced in which all of the non-hydrogen atoms appeared on an Emap. The solution which gave the correct structure was that with the second highest figures of merit and the programme assigned the following origin determining reflections:- $\bar{5}$ 1 10 (+); 3 3 1 (+); 1 2 1 (+).

Using SHELX[2], full matrix least squares refinement of positional parameters and isotropic temperature factors ensued. The programme assigned default values for the isotropic temperature factors, and the atomic scattering factors for each atom were assigned on the basis of previous analytical information available (ie. NMR, IR, MS

and CHNO analysis). During this stage of refinement unit weights were used, and after six cycles, $R = \sum [\; ||F_o| \; -s \; |F_c|| \quad]/ \sum |F_o|$ decreased to 0.143. At this point, all of the predicted hydrogen atoms in the structure appeared on a difference Fourier synthesis.

A further seven cycles of least squares refinement of coordinates and anisotropic thermal parameters for non-hydrogen atoms and coordinates and isotropic temperature factors for hydrogens reduced R to 0.064. During this stage of refinement, reflections were weighted according to $w = 1/\sigma^2[F_o]$, and isotropic temperature factors for hydrogens were initiated at that of their attached atom. A final sequence of refinements for the complete structure during which two hydrogen atoms were refined with constrained bond distances, resulted in an unweighted discrepancy index R = 0.045 for the 1341 unique observed data with $F_o > 3\sigma$. During this latter stage of refinement, reflections were weighted according to $w = k/[\sigma^2(F_o) + qF_o^2]$ where q, a refineable parameter, converged at 0.0011 and k which should approach unity if there are no serious systematic errors, equalled 1.6309.

The refinement was finally terminated when no positional parameter shifted by more than 0.31 and the weighted discrepancy index $R_g = 0.066$.

A final difference electron density map showed no feature greater than $0.22 e\mathring{A}^{-3}$.

Table 5.10 Positional parameters (fractional co-ordinates x 10^4) with estimated standard deviations in parentheses.

	X/a	Y/b	Z/c
N(1)	8272(2)	2842(2)	1937(1)
C(2)	9325(2)	3916(3)	2514(2)
N(3)	9373(2)	4108(2)	3599(1)
C(4)	8487(2)	3322(2)	4076(1)
N(5)	7491(2)	2247(2)	3455(1)
C(6)	7442(2)	2024(2)	2422(2)
O(2)	10114(2)	4617(2)	2107(1)
O(4)	8624(2)	3607(2)	5065(1)
C(7)	6352(2)	879(3)	1687(2)
C(1)'	4918(2)	1711(3)	1165(2)
C(2)'	4621(3)	2546(3)	171(2)
C(3)'	3326(4)	3356(4)	-290(3)
C(4)'	2319(4)	3379(4)	236(4)
C(5)'	2575(3)	2544(4)	1229(4)
C(6)'	3871(3)	1707(4)	1693(3)
H(1)	8222(29)	2529(34)	1269(25)
H(3)	10108(33)	4809(34)	3963(25)
H(71)	6781(28)	440(30)	1085(23)
H(72)	6292(27)	4(28)	2107(21)
H(2)'	5341(33)	2542(40)	-184(27)
H(3)'	3077(46)	4104(44)	-950(24)
H(4)'	1407(44)	3979(58)	-311(37)
H(5)'	1801(37)	2560(57)	1604(34)
H(6)'	4057(34)	1064(43)	2366(31)

TABLE 5.11   Anisotropic temperature factors (non-hydrogen atoms)

Isotropic temperature factors (hydrogen atoms)

with estimated standard deviations in parentheses

ATOM	U11	U22	U33	U23	U13	U12
N(1)	.0434(10)	.0479(10)	.0230(8)	-.0048(7)	.0135(7)	-.0014(8)
C(2)	.0407(11)	.0492(12)	.0245(10)	-.0015(8)	.0126(8)	-.0011(9)
N(3)	.0429(10)	.0513(11)	.0246(8)	-.0046(7)	.0130(7)	-.0096(9)
C(4)	.0421(11)	.0412(11)	.0235(9)	.0024(8)	.0097(8)	.0000(9)
N(5)	.0465(10)	.0435(10)	.0299(9)	-.0009(7)	.0146(7)	-.0052(8)
C(6)	.0419(11)	.0367(10)	.0305(10)	-.0014(8)	.0113(8)	.0023(9)
O(2)	.0621(10)	.0843(13)	.0362(8)	-.0073(8)	.0272(8)	-.0248(9)
O(4)	.0617(10)	.0620(10)	.0214(7)	-.0017(6)	.0164(7)	-.0119(8)
C(7)	.0521(13)	.0397(12)	.0420(12)	-.0088(10)	.0135(10)	-.0047(10)
C(1)'	.0481(12)	.0414(12)	.0423(12)	-.0146(9)	.0110(10)	-.0099(10)
C(2)	.0675(18)	.0710(18)	.0488(14)	-.0025(13)	.0132(13)	.0021(14)
C(3)'	.0718(20)	.0894(25)	.0733(20)	.0044(18)	.0037(17)	.0135(18)
C(4)'	.0673(20)	.0643(20)	.1031(27)	-.0181(18)	-.0077(20)	.0048(16)
C(5)'	.0512(17)	.0796(21)	.1272(31)	-.0406(21)	.0381(20)	-.0144(16)
C(6)'	.0610(17)	.0666(17)	.0720(18)	-.0126(15)	.0298(14)	-.0136(14)

	Uiso
H(1)	.0611(71)
H(3)	.0724(82)
H(71)	.0608(67)
H(72)	.0544(67)
H(2)'	.0874(106)
H(3)'	.1311(145)
H(4)'	.1363(145)
H(5)'	.1539(161)
H(6)'	.0856(99)

Table 5.12 Bond distances in $\overset{o}{A}$ with estimated standard deviations in parentheses

Bond	Interatomic distance $\overset{o}{A}$
N(1)-C(2)	1.373(3)
C(2)-N(3)	1.368(2)
N(3)-C(4)	1.369(3)
C(4)-N(5)	1.360(2)
N(5)-C(6)	1.307(2)
C(6)-N(1)	1.346(3)
C(2)-O(2)	1.205(2)
C(4)-O(4)	1.238(2)
C(6)-C(7)	1.499(3)
C(7)-C(1)'	1.500(3)
C(1)'-C(2)'	1.382(4)
C(2)'-C(3)'	1.376(4)
C(3)'-C(4)'	1.351(5)
C(4)'-C(5)'	1.386(6)
C(5)'-C(6)'	1.388(5)
C(6)'-C(1)'	1.388(3)
N(1)-H(1)	0.87(3)
N(3)-H(3)	0.92(3)
C(7)-H(71)	1.05(3)
C(7)-H(72)	0.91(2)
C(2)'-H(2)'	0.95(3)
C(3)'-H(3)'	1.01(1)
C(4)'-H(4)'	1.05(5)
C(5)'-H(5)'	1.01(1)
C(6)'-H(6)'	.97(4)

<u>Note</u>: Hydrogen atom positions H(3)' and H(5)' were refined with constrained bond distances of 1.01$\overset{o}{A}$ with a standard deviation of 0.01 using the DFIX option in SHELX.

Table 5.13 Interatomic angles with estimated standard deviations in parentheses (for non-hydrogen atoms, and those hydrogens attached to nitrogen atoms).

Atoms	Bond Angle [°]
N(1)-C(2)-N(3)	112.5(2)
N(1)-C(2)-C(2)	123.7(2)
O(2)-C(2)-N(3)	123.8(2)
C(2)-N(3)-C(4)	124.6(2)
N(3)-C(4)-N(5)	119.3(2)
N(3)-C(4)-O(4)	118.7(2)
O(4)-C(4)-N(5)	122.0(2)
C(4)-N(5)-C(6)	117.1(2)
N(5)-C(6)-N(1)	123.7(2)
N(5)-C(6)-C(7)	119.8(2)
C(6)-N(1)-C(2)	122.6(2)
N(1)-C(6)-C(7)	116.4(2)
C(6)-C(7)-C(1)'	110.2(2)
C(7)-C(1)'-C(2)'	121.2(2)
C(7)-C(1)'-C(6)'	120.5(2)
C(1)'-C(2)'-C(3)'	121.3(3)
C(2)'-C(3)'-C(4)'	120.4(3)
C(3)'-C(4)'-C(5)'	119.8(3)
C(4)'-C(5)'-C(6)'	120.2(3)
C(5)'-C(6)'-C(1)'	120.1(3)
C(6)-C(1)'-C(2)'	118.2(3)
C(6)-N(1)-H(1)	116(2)
C(2)-N(1)-H(1)	120(2)
C(2)-N(3)-H(3)	110(2)
C(4)-N(3)-H(3)	125(2)

Table 5.14 Deviations of non-hydrogen atoms in $\overset{o}{A}$ from the least squares planes through (a) the triazine ring, and (b) the phenyl ring[+] (atoms used in the plane calculation are marked with an asterisk)

a)

ATOM	DEVIATION $[\overset{o}{A}]$	ATOM	DEVIATION $[\overset{o}{A}]$
*N(1)	.020	C(7)	.000
*C(2)	-.001	C(1)'	1.411
*N(3)	-.007	C(2)'	2.179
*C(4)	.009	C(3)'	3.488
*N(5)	.004	C(4)'	4.065
*C(6)	-.018	C(5)'	3.322
O(2)	-.022	C(6)'	1.995
O(4)	.019		

+ The equation of the plane is:-

$-.678X + .732Y - .064Z - 3.422 = 0$ where X, Y and Z are orthogonal co-ordinates in $\overset{o}{A}$ along a*, b and c

Table 5.14 b)

ATOM	DEVIATION [Å]	ATOM	DEVIATION [Å]
N(1)	2.024	C(7)	.068
C(2)	3.314	*C(1)'	.005
N(3)	3.996	*C(2)'	.003
C(4)	3.469	*C(3)'	-.009
N(5)	2.172	*C(4)'	.008
C(6)	1.494	*C(5)'	-.001
O(2).	3.792	*C(6)'	-.006
O(4)	4.190		

The equation of the plane is:-

$.412X + .819Y + .393Z + 3.008 = 0$ where X, Y and Z are orthogonal coordinates in Å along a*, b and c.

Table 5.15 Important torsion angles

ATOMS	ANGLE [°]
C(6)-N(1)-C(2)-N(3)	-3.1
N(1)-C(2)-N(3)-C(4)	.4
C(2)-N(3)-C(4)-N(5)	1.1
N(3)-C(4)-N(5)-C(6)	.1
C(4)-N(5)-C(6)-N(1)	-2.9
N(5)-C(6)-N(1)-C(2)	4.6
C(6)-N(1)-C(2)-O(2)	177.3
O(2)-C(2)-N(3)-C(4)	-180.0
C(2)-N(3)-C(4)-O(4)	-179.3
O(4)-C(4)-N(5)-C(6)	-179.5
N(5)-C(6)-C(7)-C(1)'	84.9
N(1)-C(6)-C(7)-C(1)'	-91.5
C(6)-C(7)-C(1)'-C(2)'	87.2
C(6)-C(7)-C(1)'-C(6)'	-90.5

The torsion angle A-B-C-D is the projected angle between AB and CD
when viewed down the B-C bond; the clockwise rotation of the C-D bond
with reference to bond A-B is considered positive.

Table 5.16 Hydrogen bond contact distances and angles

HYDROGEN BOND	ANGLE AT H[°]	N ... O DISTANCE Å
$N(1)_{II}-H(1)\ldots\ldots O(4)_I$	162	2.778
$N(3)_{III}-H(3)\ldots\ldots O(4)_I$	166	2.845

I, II, III refer to the symmetry operations x, y, z; $x, \frac{1}{2}-y, \frac{1}{2}+z$;

$2-x, 1-y, 1-z$

Figure 5.15 Molecular structure and numbering scheme for 6-benzyl-1,3,5 triazin-2,4[1H,3H]-dione
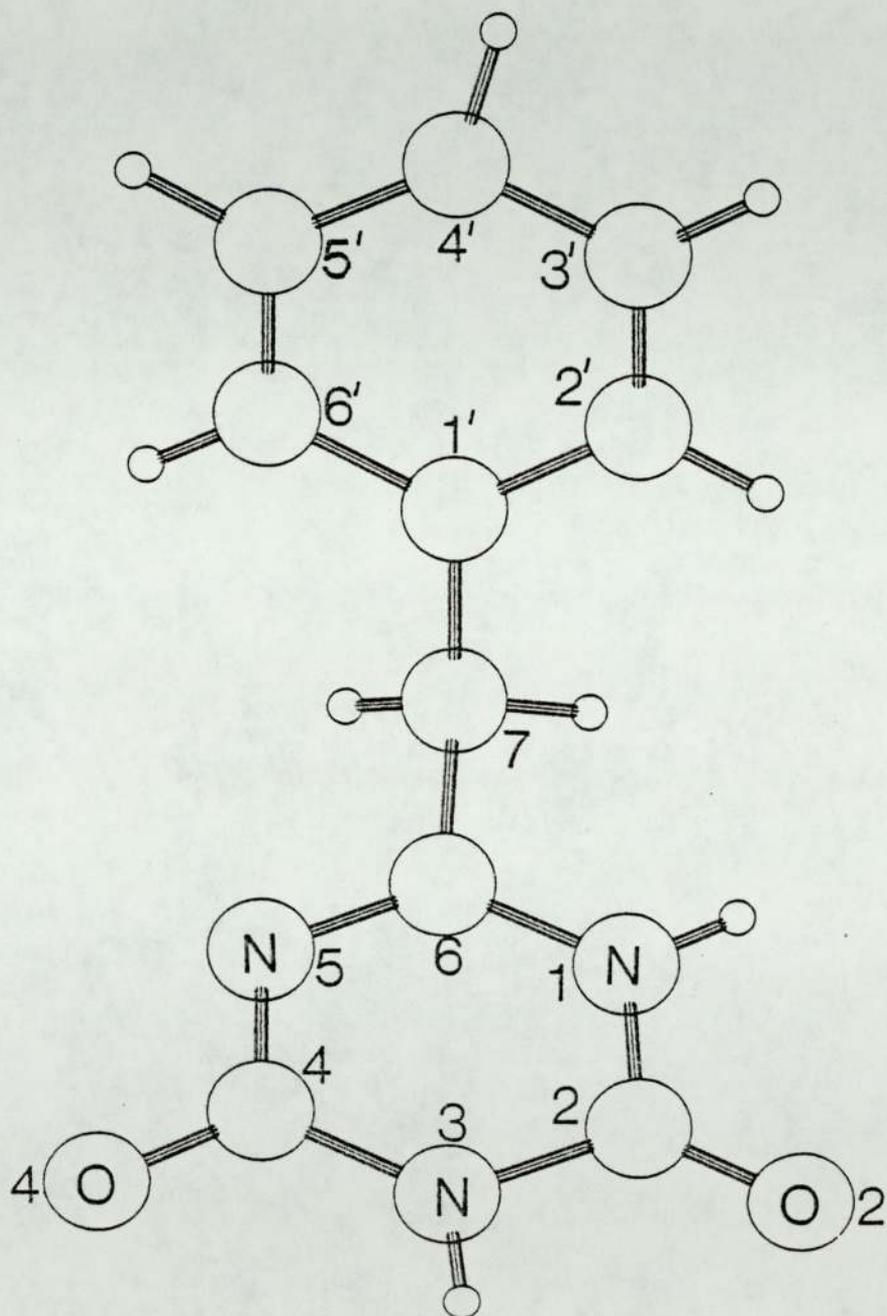
Figure 5.15a Stereo diagram of 6-benzyl-1,3,5 triazin-2,4[1H,3H]-dione
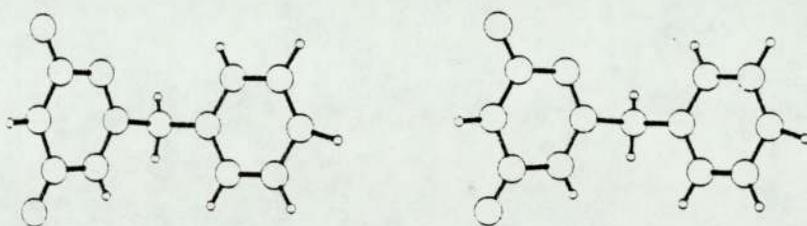
Figure 5.15b Stereo packing diagram viewed along b
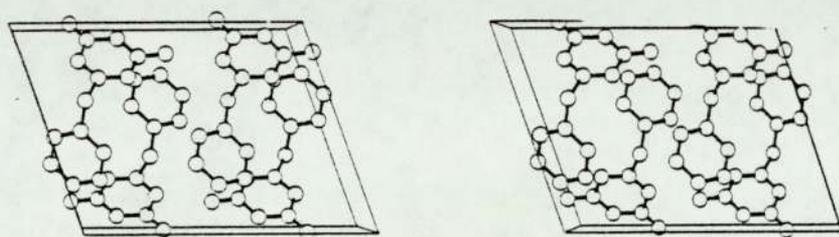
Figure 5.16 Packing diagram of 6-benzyl-1,3,5 triazin-2,4[1H,3H]-
dione
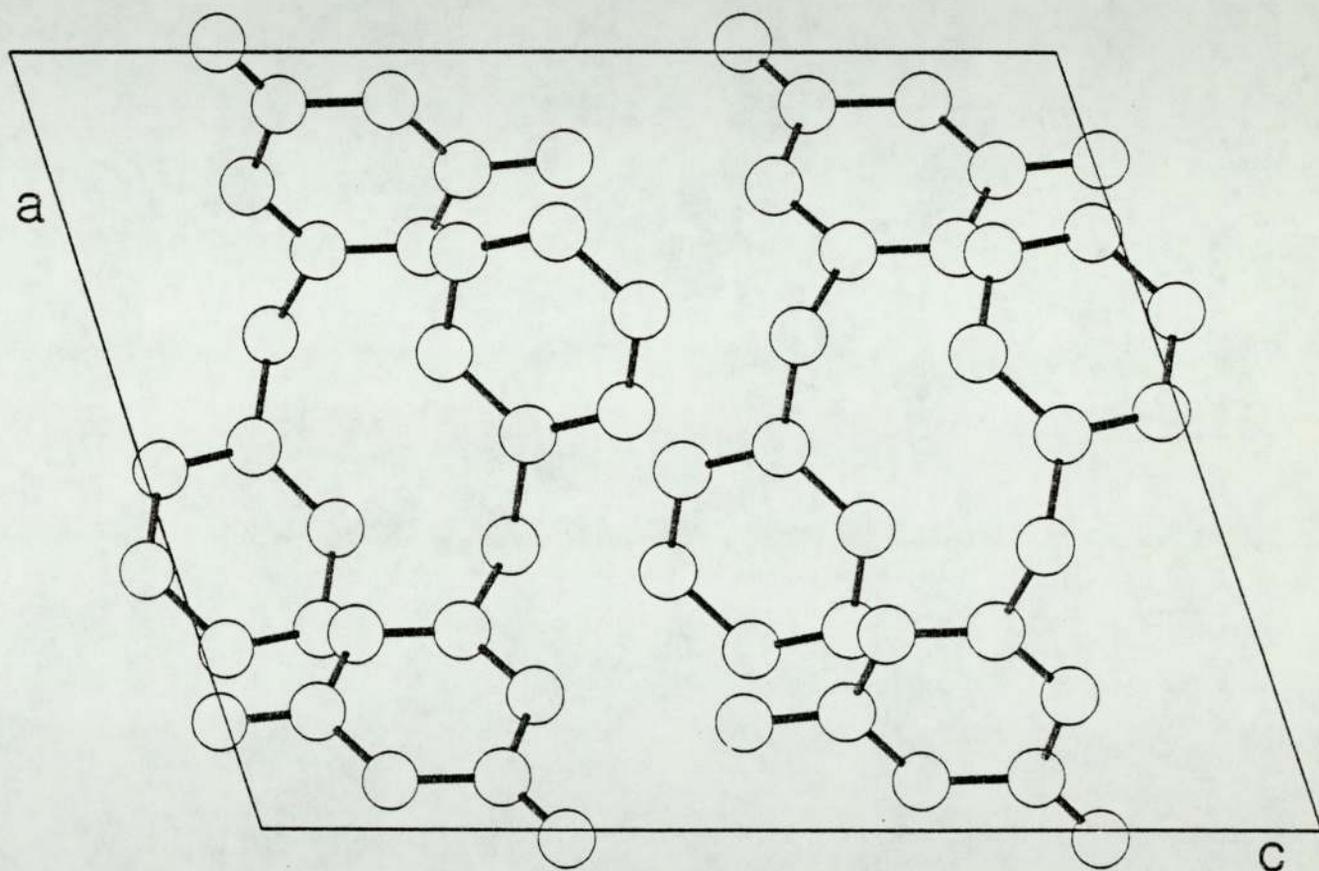
Figure 5.17 Packing diagram showing hydrogen bonding in the a-c plane (hydrogen bonds are shown in dotted lines)
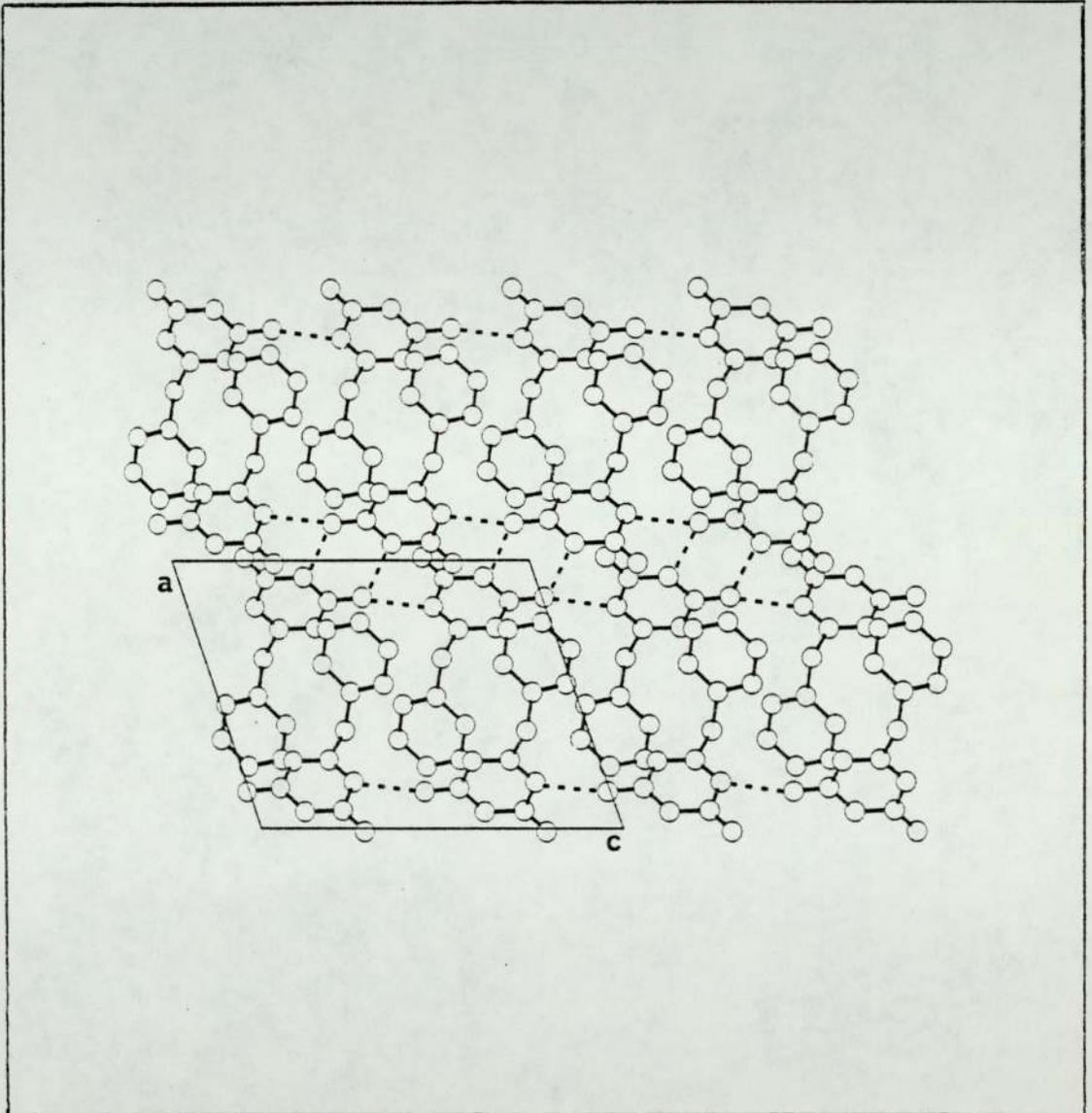
Figure 5.18 Packing diagram viewed along c showing the a-b plane
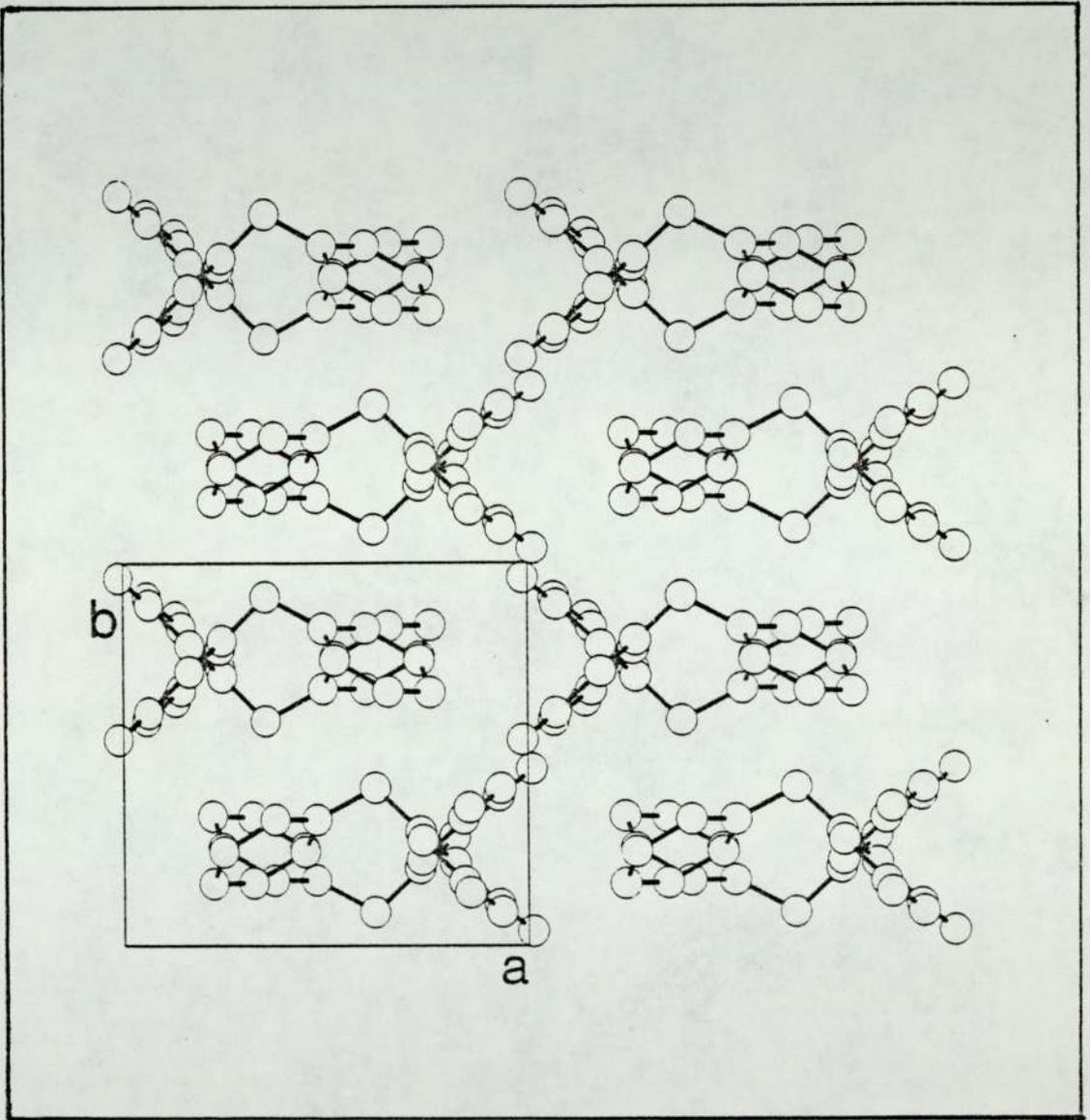
Figure 5.19 Space filling diagram of 6-benzyl-1,3,5 triazin-
2,4[1H,3H] dione
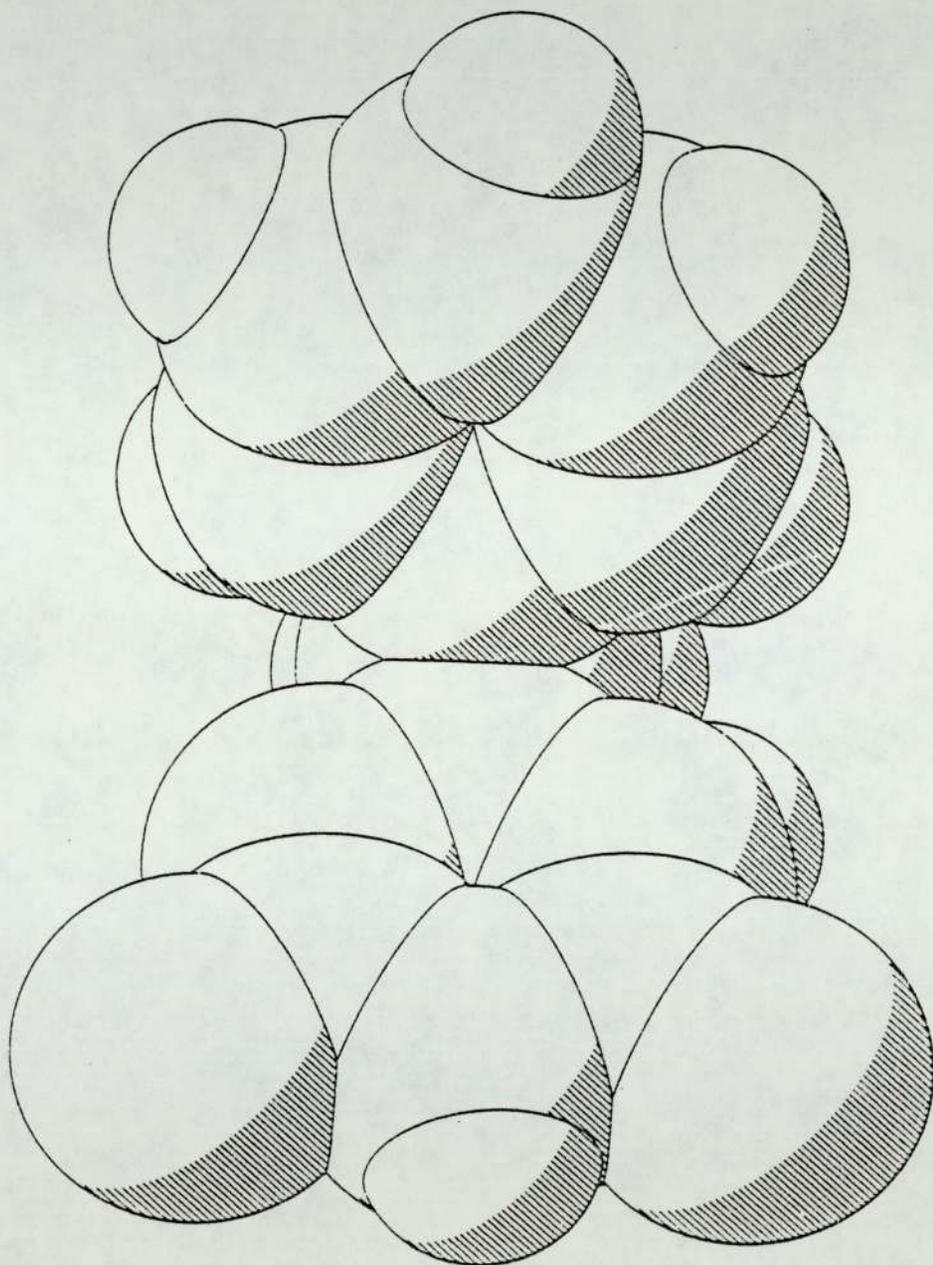
5.3.4   Results and Discussion

The structure as determined crystallographically is shown in Figure 5.15 together with the numbering scheme used; positional parameters for all atoms are given in Table 5.10 with anisotropic thermal parameters for non-hydrogens and isotropic temperature factors for hydrogens in Table 5.11. The bond distances and interatomic angles are given in Tables 5.12 and 13 respectively.

Unlike the 6-phenyl compound the carbonyl bond distances C(2)-O(2) and C(4)-O(4) are not identical $(1.205(2)\overset{o}{A}$ and $1.238(2)\overset{o}{A}$ respectively). This is attributed later mainly to O(4) being doubly hydrogen bonded, whilst O(2) does not participate in any hydrogen bonds. Such a feature is also found in the structures of uracil[87] and dihydrouracil[88] which hydrogen bond in a similar fashion. However, the bond lengths determined for these two bonds do, as with 6-phenyl-5-azauracil suggest that in the crystalline state the molecule exists as the diketo tautomer.

The bond distances in the triazine ring albeit similar to those in the 6-phenyl compound, do show some significant differences. The N(1)-C(2) and C(2)-N(3) bond distances in both compounds are almost identical; however, the C(6)-N(1) in the 6-benzyl compound is shorter by $0.023\overset{o}{A}$ $(1.346(3)\overset{o}{A})$ suggesting considerably more double bond character. This is accompanied by a lengthening of the N(5)-C(6) bond $(1.307(2)\overset{o}{A})$ by $0.01\overset{o}{A}$ implying more single bond character. The C(4)-N(5) bond $(1.360(2)\overset{o}{A})$ is 0.021 $\overset{o}{A}$ shorter than in the 6-phenyl compound, and such changes in these bond distances would be consistent with resonance form (V) in Fig 5.14a being the more important for the 6-benzyl compound. It is difficult to assess the effect of conjugation on the carbonyl bond distance which such resonance would imply, since by far the larger effect on any

lengthening of the C(4)-O(4) with respect to the C(2)-O(2) bonds may be due to their differences in hydrogen bonding. A further contribution from resonance form (IV) might explain the decrease in the N(3)-C(4) bond distance in the 6-benzyl compound of $0.018\overset{\text{o}}{\text{A}}$.

The C(6)-C(7) and C(7)-C(1') bond lengths ($1.499(3)\overset{\text{o}}{\text{A}}$ and $1.500(3)\overset{\text{o}}{\text{A}}$) are almost identical and are consistent with typical $C(sp^2)$-$C(sp^3)$ bond lengths[89]. In the phenyl ring, with the exception of the C(3)'-C(4)' bond which is inexplicably short ($1.351(5)\overset{\text{o}}{\text{A}}$) these are as might be expected for a benzene ring, and exemplify the difference found in the 6-phenyl compound for which conjugation with the triazine ring is a possibility.

The bond angles in the molecule are shown in Table 5.13. The internal angles in the phenyl ring are all as expected close to 120°. The C(6)-C(7)-C(1)' angle (110.2(2)°) which bridges the phenyl and triazine rings is close to the tetrahedral angle. The important angles again, are the nitrogen valence angles in the triazine ring, since these provide additional information about the correct location of the hydrogen atoms on the triazine nitrogens. From the bond length data N(5)-C(6) is assigned as a double bond with there being extra-annular attachments (hydrogen atoms) at N(1) and N(3) nitrogens but not at N(5); this being further substantiated by the location of these hydrogens by a difference Fourier synthesis. The nitrogen valence angles in the triazine ring are 122.6(2)°, 124.6(2)° and 117.1(2)° for N(1), N(3) and N(5) respectively. As with the structure of the 6-phenyl compound, these values conform to Chatar Singh's rule[59] and confirm the correct location of the triazine hydrogens at N(1) and N(3).

The deviations in Angstroms from the least squares planes through the triazine and phenyl rings are given in Tables 5.14a and

b.   All of the atoms in the triazine ring lie within $\pm 0.02\overset{o}{A}$ of the plane, with the deviations of the two carbonyl oxygens, O(2) and O(4), being $-0.022\overset{o}{A}$ and 0.019 respectively; the atoms in the phenyl ring all lie within $\pm 0.009\overset{o}{A}$ of the least squares plane through the six atoms. The planarity of the triazine ring, and the co-planarity of the carbonyl oxygens is an important feature where vibrational coupling of the two carbonyl stretching frequencies is concerned. The approximate twist angle between the two rings obtained from the torsion angles (Table 5.15) is $2.2°$; the angle between the least squares planes is 72.8°. This is consistent with the low torsion angle since 180°-72.8° (107.2°) is very close to the tetrahedral geometry.   The packing arrangement of the molecules is shown in Figures 5.16, 17 and 18 with the hydrogen bonding scheme in dotted lines.   As might be expected from the very non-planar nature of the structure, stacking forces appear to be unimportant. The fundamental difference between the 6-benzyl and 6-phenyl compounds is in the intermolecular interactions.   Unlike the 6-phenyl compound and 6-azauracil[90], the 6-benzyl compound behaves like uracil[87] and dihydrouracil[88], in that only one of the carbonyl groups is hydrogen bonded, namely, C(4) and this hydrogen bonds to both N(1) and N(3). Pairs of molecules related by a centre of symmetry, form a base pair interaction via an N(3)-H(3) ... O4 interaction $(2.845\overset{o}{A})$ in the direction of the a axis, whilst adjacent molecules related by the c glide plane interact via an N(1)-H(1) ... O(4) $(2.778\overset{o}{A})$. Details of distances, angles and symmetry operations involved in the intermolecular hydrogen bonds are given in Table 5.16.

5.4   6-METHYLISOCYTOSINE

5.4.1   Experimental

A sample of 6-methylisocytosine was purchased from the Sigma Chemical Company, St Louis, MO 63178, USA. The compound as purchased was, in the crystalline state, in the form of 'tiny' prisms. Without any further chemical manipulations, a single crystal of dimensions 0.3 x 0.14 x 0.13 mm was selected from the sample for diffractometry.

The infrared spectrum of the compound was recorded both as a nujol mull and KBr disc on a Pye Unicam SP200 scanning spectrophotometer.

5.4.2   Crystal Data

6-methylisocytosine ($C_5H_7N_3O$) crystallises in the monoclinic space group $P2_1/n$ with a=7.653(1)$\overset{o}{A}$, b=6.567(1)$\overset{o}{A}$, c=11.815(3)$\overset{o}{A}$, β=98.12(2)° with z=4 and v=588(1)$\overset{o}{A}$. The molecular weight is 125.14 (F(ooo)=264) and the calculated density Dc=1.404 gcm^{-3}. The absorption coefficient for CuKα radiation (λ =1.5418$\overset{o}{A}$) is μ=8.81 cm^{-1}.

5.4.3   Structure Analysis

The data were collected from a prismatic crystal of dimensions 0.3 x 0.14 x 0.13 mm mounted parallel to the 00$\overline{1}$ axis on an Enraf-Nonius CAD4 diffractometer; the ω -2Θ scan technique was used with monochromated CuKα radiation (λ =1.5418$\overset{o}{A}$). Two reference reflections were measured every 3600 sec of exposure time, and the orientation checked after every 100 reflections. No decomposition or movement of the crystal was detected during the collection of the 2796 unique reflections between Θ=1 and Θ=78°. The unit cell parameters determined from the data collection procedure are given above under 'Crystal Data'.

The assistance of Dr G B Williams of Brookhaven National Laboratory in data collection is gratefully acknowledged.

Structure Determination and Refinement

The data reduction for the diffractometer data was performed at Birmingham University on the DEC 20 using a programme provided by Dr T A Hamor. Lorentz and polarization and intensity corrections were applied and the data output in the form HKL, Fobs and σFobs where σFobs is based on counting statistics. No correction to the data was made for absorption effects since for CuKα radiation, the maximum change in intensity along a straight path over the length of the crystal (ie. $\frac{Io-I}{Io} \times 100$) is of the order of 20%. The change in intensity between the minimum and maximum paths through the crystal is of the order of 10%.

Normalised structure factors were calculated using a modified Wilson plot method[22]. The curve from the Wilson calculation and a least squares fit are shown in Figure 5.20, with the values calculated from the gradient and intercept being:

Temperature factor (B) = 2.3281

Scale (k)   = .5489

The distribution of E's for the complete data set and and the $|E^2-1|$ statistic are consistent with that of a centrosymmetric structure.

Phase determination was undertaken by MULTAN[1] in which 400 reflections all with E $\geqslant$ 1.425 provided 3900 unique $\sum_2$ relationships for use in phase refinement. On the basis of a $\sum_1$ calculation, the following invarient phases were determined, all with probabilities greater than 95%:-

        2 0 6   (+)
        4 0 12   (+)
        0 0 4   (-)
        0 6 0   (-)
        0 2 0   (-)

The converge procedure produced the following origin determining reflections:- 1 0 3 (+), 3 5 0 (+), 5 1 $\bar{7}$ (+). Sixteen starting sets were produced by permuting the signs of the 2 4 $\bar{6}$, 3 5 $\bar{4}$, 1 6 0 and 6 1 $\bar{6}$ reflections; the set of signs (+), (-), (+), (+) respectively, produced an E-map from which all of the non-hydrogen atoms could be located.

Figure 5.20
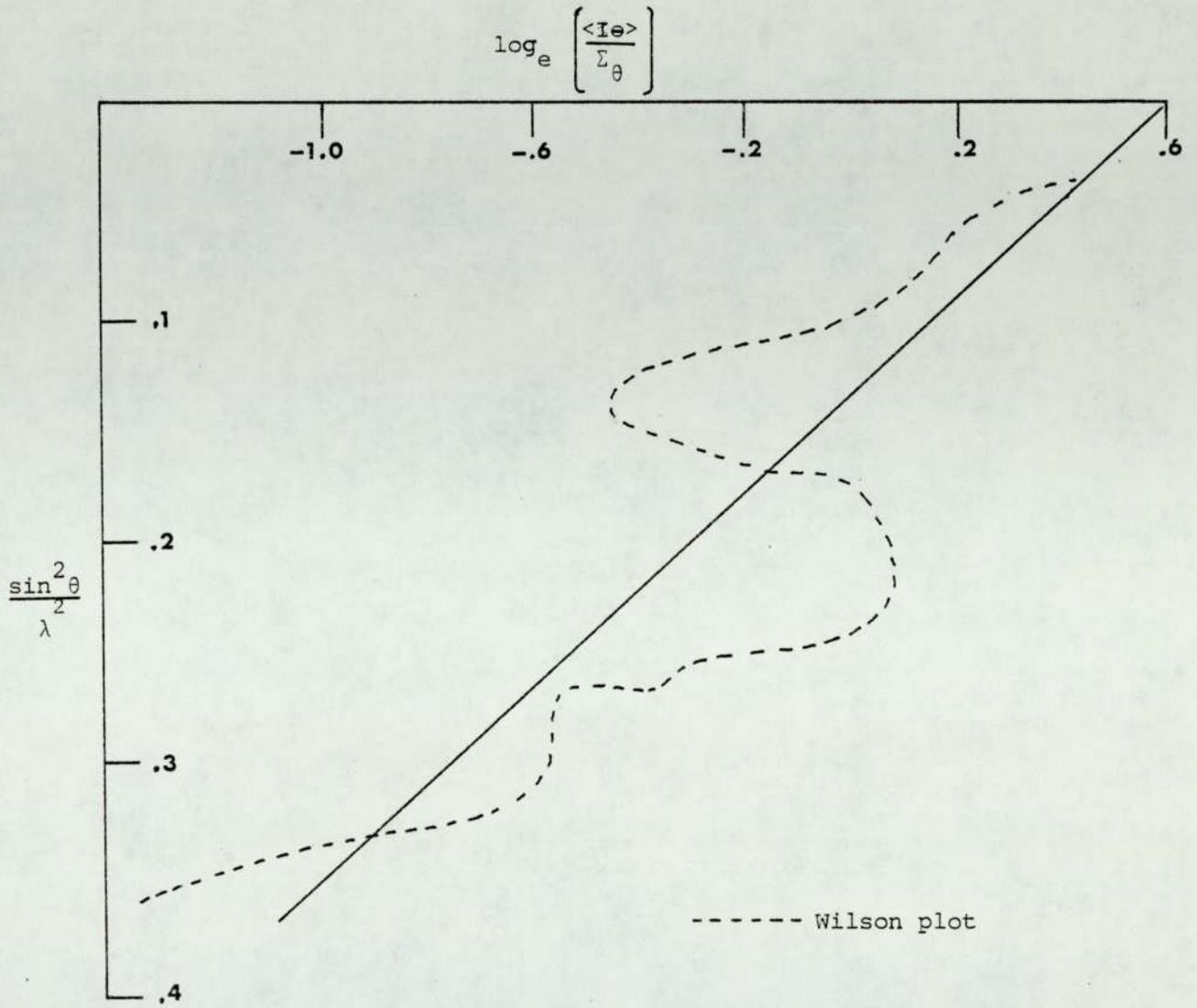
$\log_e \left[\dfrac{<I\theta>}{\Sigma_\theta} \right]$

$\dfrac{\sin^2\theta}{\lambda^2}$

------ Wilson plot

Full matrix least squares refinement of positional parameters and isotropic temperature factors using SHELX[2], during which all atoms were assigned the scattering factor for carbon, reduced the unweighted discrepancy index R to 0.19. At this stage, on the basis of isotropic temperature factors and the correct location of the carbonyl oxygen in view of the C=O bond distance, atoms were assigned their correct scattering factors and a further five cycles of isotropic refinement ensued after which, all but one of the hydrogens (one attached to the exocyclic amine group) appeared in a difference Fourier synthesis. A further sequence of least squares refinement of coordinates and anisotropic temperature factors for hydrogens during which the missing hydrogen was placed in a calculated position and allowed to refine, reduced R to 0.071. During this stage of refinement, reflections were weighted according to $w=1/\sigma^2(Fo)$.

Final anisotropic refinement of the complete structure with isotropic temperature factors for hydrogen atoms reduced the final agreement parameters to $R=0.060$ and $Rg = 0.068$ for the observed reflections. During this stage of the refinement procedure, the methyl and amine hydrogens were refined with constrained bond distances (Table 5.19) and the reflections weighted according to $w=k/[\sigma^2(Fo) + qFo^2]$ where k and q (both described earlier in this section) converged at 4.6126 and 0.0001 respectively. This stage of the refinement involved 123 parameters and 1094 unique observed data with $Fo > 3\sigma$.

The refinement was terminated when none of the positional parameters shifted by more than 0.011σ. A final difference electron density map showed no feature greater than $0.32e\text{Å}^{-3}$.

Table 5.17 Positional parameters (fractional co-ordinates $\times 10^4$) with estimated standard deviations in parentheses.

	X/a	Y/b	Z/c
N(1)	6480(3)	-2130(3)	1208(2)
C(2)	7873(3)	-1315(4)	766(2)
N(3)	8122(3)	682(3)	681(2)
C(4)	6910(3)	1965(3)	1030(2)
C(5)	5437(3)	1150(4)	1518(2)
C(6)	5252(3)	-888(4)	1599(2)
N(7)	9008(3)	-2602(3)	401(2)
O(8)	7122(3)	3848(3)	924(2)
C(9)	3793(4)	-1944(5)	2072(3)
H(1)	6330(36)	-3645(33)	1227(23)
H(2)	10007(35)	-2143(48)	25(24)
H(3)	8651(44)	-3975(37)	444(29)
H(4)	4436(44)	2109(54)	1740(27)
H(5)	3147(50)	-2941(54)	1519(30)
H(6)	2863(55)	-1175(71)	2410(43)
H(7)	4184(48)	-2867(50)	2750(24)

TABLE 5.18   Anisotropic temperature factors (non-hydrogen atoms)

Isotropic temperature factors (hydrogen atoms) with standard

deviations in parentheses

ATOM	U11	U22	U33	U23	U13	U12
N(1)	.0321(10)	.0213(8)	.0415(11)	.0028(8)	.0103(9)	-.0023(8)
C(2)	.0312(12)	.0242(10)	.0390(12)	.0029(9)	.0102(10)	-.0013(9)
N(3)	.0328(11)	.0221(9)	.0445(12)	.0017(8)	.0120(9)	-.0027(8)
C(4)	.0348(12)	.0251(11)	.0406(13)	-.0002(9)	.0070(10)	-.0007(10)
C(5)	.0365(13)	.0261(12)	.0442(14)	-.0018(10)	.0138(11)	.0023(10)
C(6)	.0302(12)	.0288(11)	.0405(13)	.0010(10)	.0121(10)	-.0036(9)
N(7)	.0422(12)	.0227(10)	.0709(16)	-.0018(10)	.0286(12)	-.0008(9)
O(8)	.0527(12)	.0174(8)	.0720(14)	-.0006(8)	.0187(10)	-.0034(8)
C(9)	.0386(14)	.0433(15)	.0544(17)	.0084(13)	.0192(13)	-.0058(12)
	Uiso					
H(1)	.0424(76)					
H(2)	.0568(89)					
H(3)	.0670(107)					
H(4)	.0654(101)					
H(5)	.0973(141)					
H(6)	.1474(205)					
H(7)	.0784(116)					

Table 5.19 Bond distances in $\overset{o}{A}$ with estimated standard deviations in parentheses

Bond		Interatomic distance $\overset{o}{A}$
N(1)-C(2)		1.361(3)
C(2)-N(3)		1.331(3)
C(2)-N(7)		1.327(3)
N(3)-C(4)		1.360(3)
C(4)-O(8)		1.256(3)
C(4)-C(5)		1.440(3)
C(5)-C(6)		1.351(3)
C(6)-C(9)		1.489(3)
C(6)-N(1)		1.373(3)
N(1)-H(1)	a	1.00(2)
N(7)-H(2)	a	0.98(2)
N(7)-H(3)	a	0.95(2)
C(5)-H(4)		1.05(3)
C(9)-H(5)	b	1.00(2)
C(9)-H(6)	b	1.00(3)
C(9)-H(7)	b	1.02(3)

a) Hydrogen positions refined with a constrained bond distance of 1.01$\overset{o}{A}$ with a standard deviation of 0.03

b) Hydrogen positions refined with a constrained bond distance of 1.08$\overset{o}{A}$ with a standard deviation of 0.03

Table 5.20 Interatomic angles [°] with estimated standard deviations in parentheses (for non-hydrogen atoms, and those hydrogens attached to nitrogen atoms)

Atoms	Bond Angle [°]
N(1)-C(2)-N(3)	123.0(2)
N(1)-C(2)-N(7)	117.3(2)
N(7)-C(2)-N(3)	119.7(2)
C(2)-N(3)-C(4)	118.5(2)
N(3)-C(4)-C(5)	119.8(2)
N(3)-C(4)-O(8)	118.5(2)
C(5)-C(4)-O(8)	121.7(2)
C(4)-C(5)-C(6)	119.5(2)
C(5)-C(6)-N(1)	118.8(2)
C(5)-C(6)-C(9)	125.4(2)
N(1)-C(6)-C(9)	115.8(2)
C(6)-N(1)-C(2)	120.4(2)
C(2)-N(1)-H(1)	120 (2)
C(6)-N(1)-H(1)	120 (2)
C(2)-N(7)-H(2)	123 (2)
C(2)-N(7)-H(3)	113 (2)
H(2)-N(7)-H(3)	124(3)

Table 5.21   Deviations in [Å] of non-hydrogen atoms from the least squares plane through the pyrimidine ring[+]. (Atoms used in the plane calculation are marked with an asterisk)

ATOM	DEVIATION [Å]
*N(1)	-.008
*C(2)	-.001
*N(3)	.012
*C(4)	-.012
*C(5)	.003
*C(6)	.007
N(7)	-.020
O(8)	-.033
C(9)	-.008

The equation of the plane is:-

.537X + .014Y + .844Z + 3.236 = 0 where X, Y and Z are orthogonal coordinates in Å along a*, b and c

Table 5.22 Important Torsion Angles [°]

ATOMS	ANGLE [°]
N(1)-C(2)-N(3)-C(4)	-1.4
C(2)-N(3)-C(4)-C(5)	2.3
C(2)-N(3)-C(4)-O(8)	-178.3
N(3)-C(4)-C(5)-C(6)	-1.6
C(4)-C(5)-C(6)-N(1)	-.2
C(4)-C(5)-C(6)-C(9)	-179.4
O(8)-C(4)-C(5)-C(6)	179.1
C(5)-C(6)-N(1)-C(2)	1.1
C(6)-N(1)-C(2)-N(3)	-.4
C(6)-N(1)-C(2)-N(7)	-179.8
C(9)-C(6)-N(1)-C(2)	-179.6
N(7)-C(2)-N(3)-C(4)	178.1

The torsion angle A-B-C-D is the projected angle between AB and CD when viewed down the B-C bond; the clockwise rotation of the C-D bond with reference to the A-B bond is considered positive.

Table 5.23 Hydrogen bond contact distances and angles

Hydrogen bond	Angle at H [°]	N ..... O Distance [Å] or N ..... N
N(1)$_{II}$-H(1) ... O(8)$_I$	150	2.716
N(7)$_I$-H(2) .... N(3)$_{III}$	171	2.973
N(7)$_{II}$-H(3) ... O(8)$_I$	152	2.855

I, II and III refer to the symmetry operations:- x, y, z;

x, 1+y, z; 2-x, -y, -z

Table 24 Important intermolecular contacts involved in stacking

ATOMS	DISTANCE [Å]
C(4)$_I$-N(1)$_{II}$	3.44
C(5)$_I$-N(1)$_{II}$	3.41
C(2)$_I$-C(5)$_{II}$	3.43
C(6)$_I$-C(4)$_{II}$	3.38

I and II refer to the symmetry operations:- x, y, z; 1-x, -y, -z

Figure 5.21 Molecular structure and numbering scheme for 6-methylisocytosine
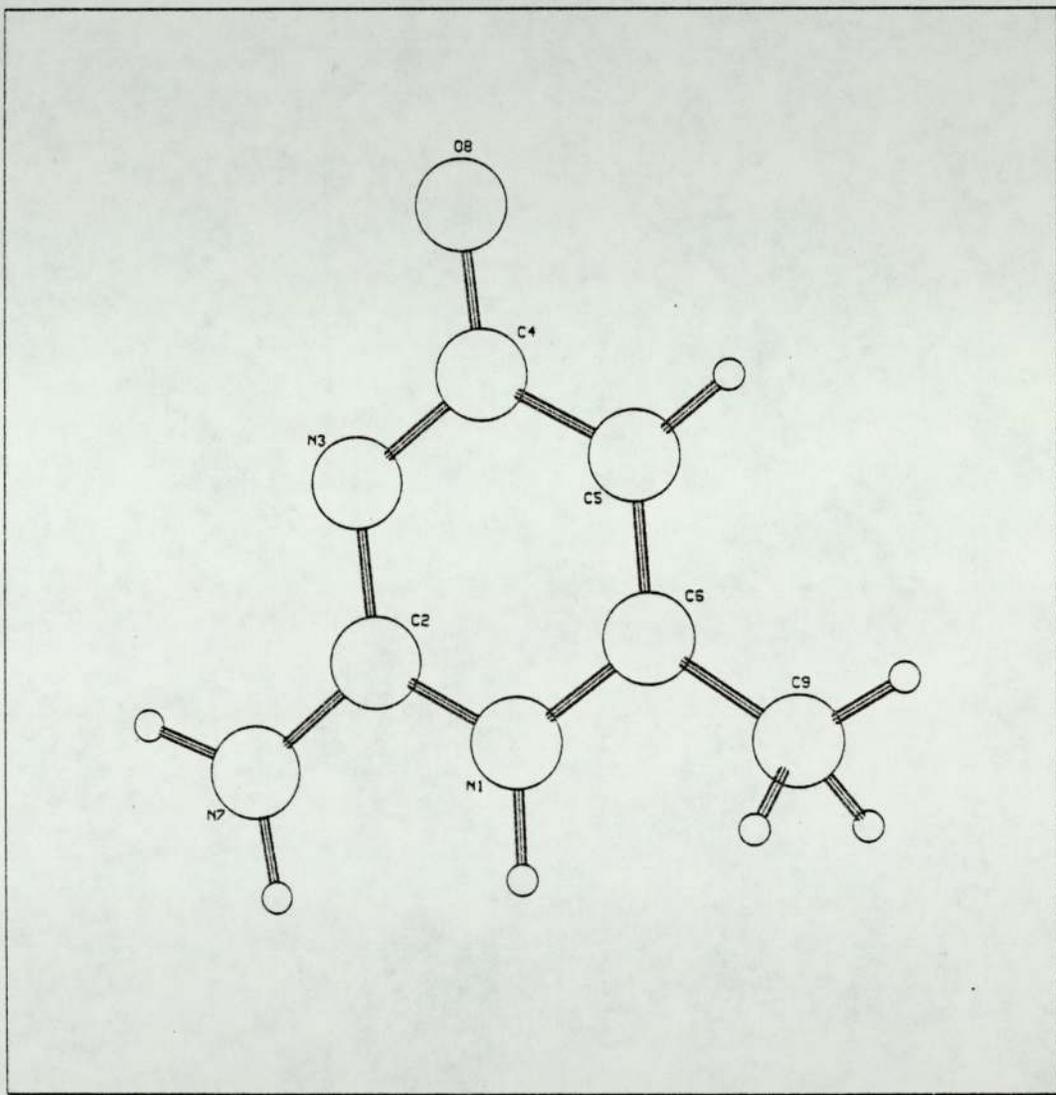
Figure 5.22a Stereo diagram of 6-methylisocytosine
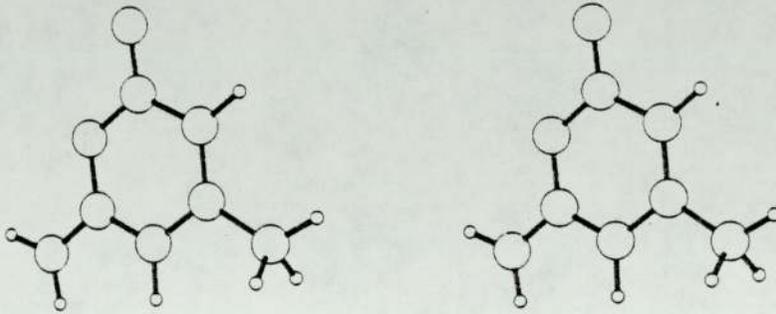
Figure 5.22b Stereo packing diagram viewed along <u>b</u>
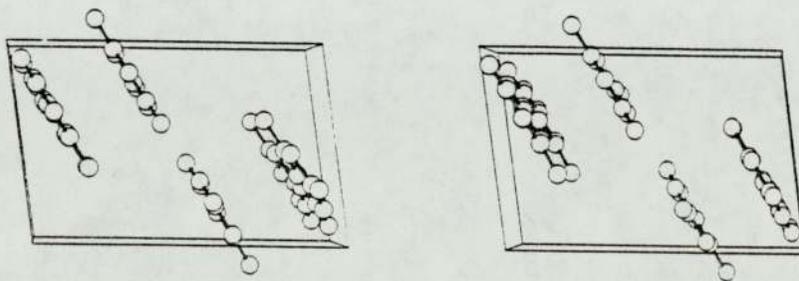
Figure 5.23 Packing diagram for 6-methylisocytosine showing close coincidence of the plane of the molecule with the 103 crystallographic planes
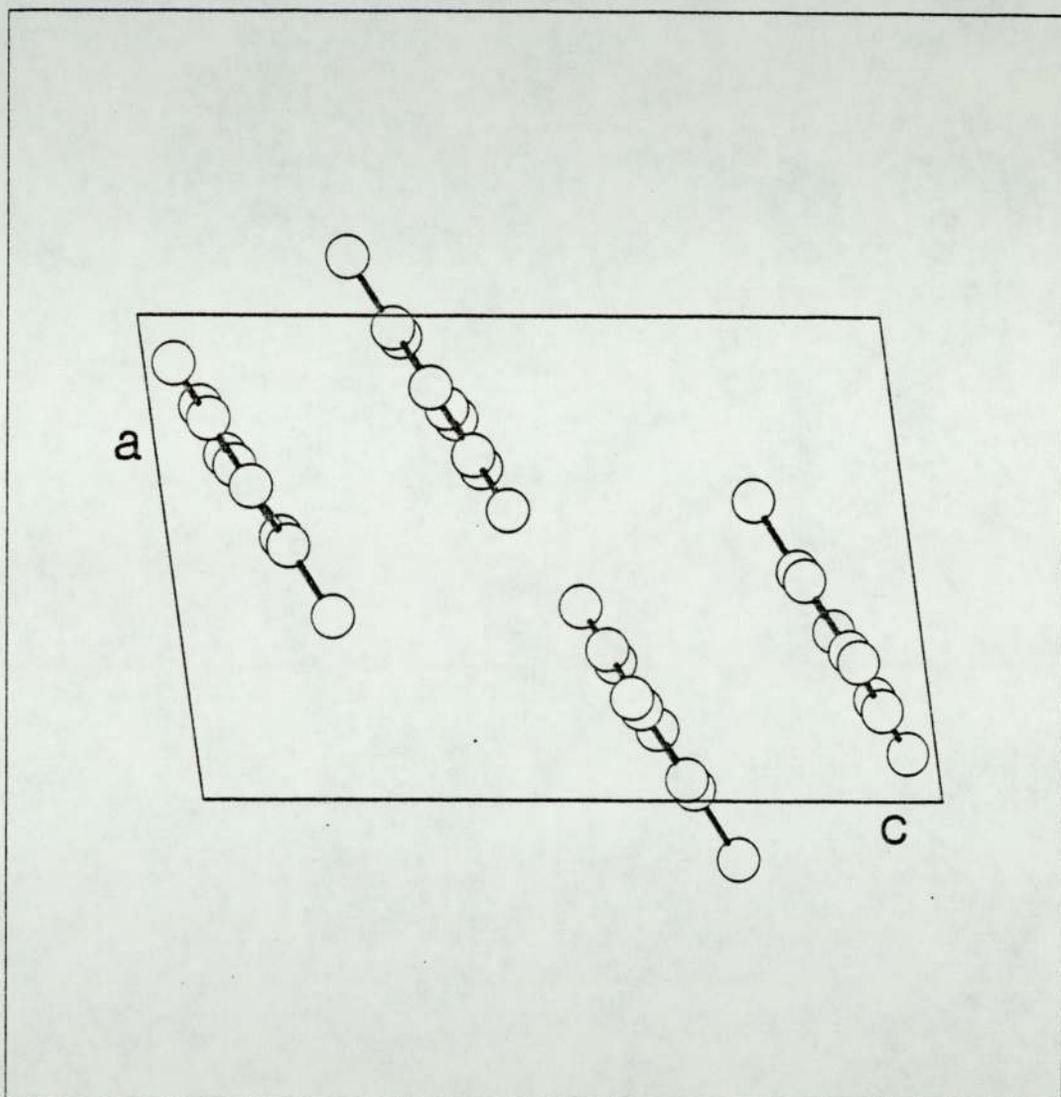
Figure 5.24 Packing diagram viewed along c showing the b-c plane
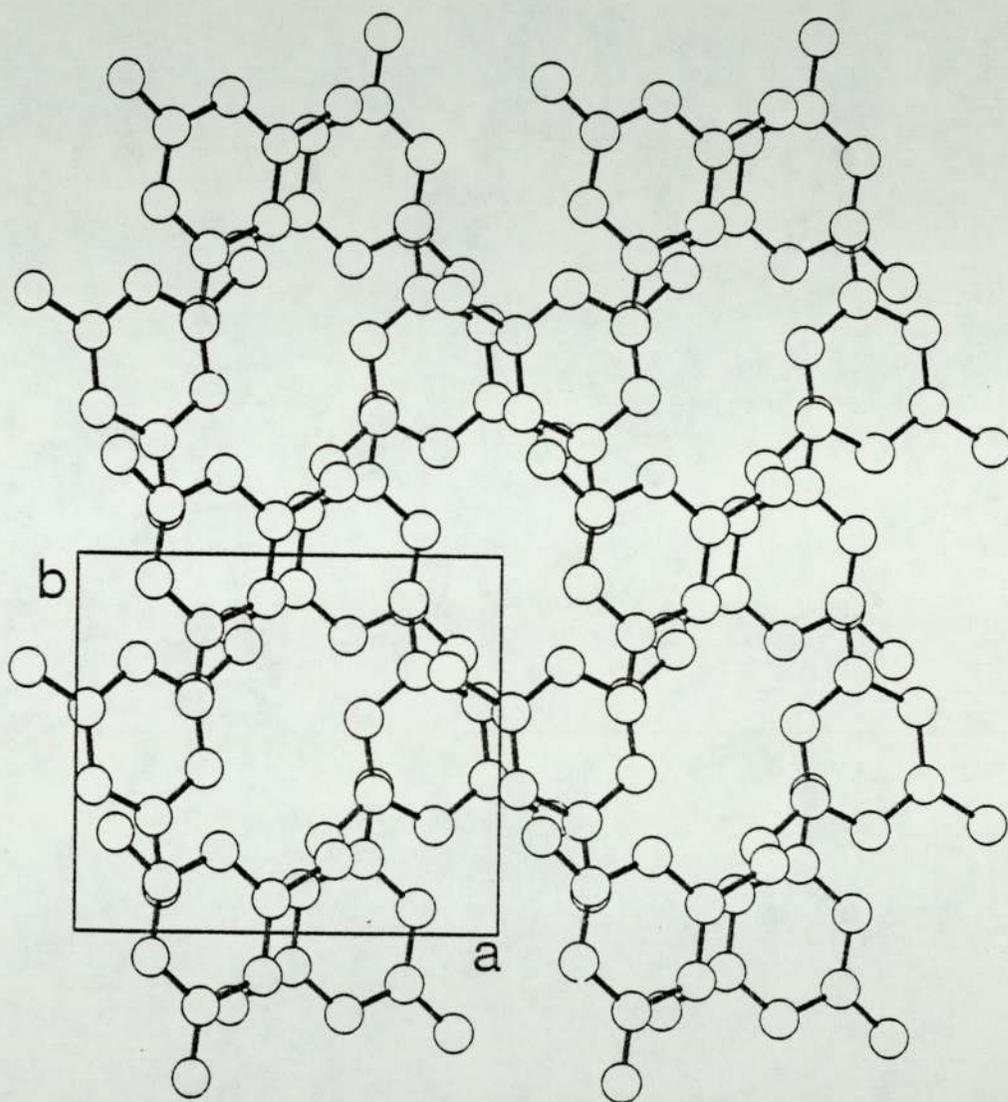
Figure 5.25   The same view as Figure 24 but with the screw axis and
               glide plane related molecules removed to add clarity to
               the hydrogen bonding.   (Hydrogen bonds are in dotted
               lines)
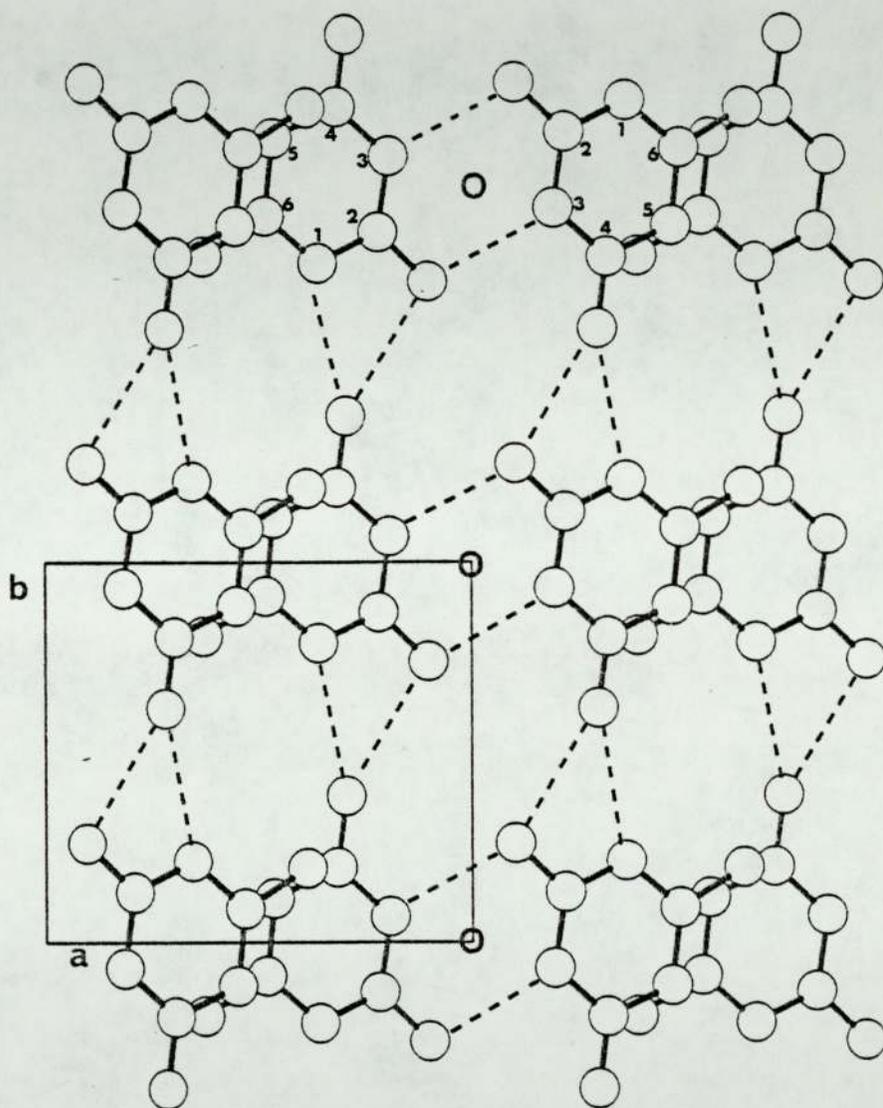
Figure 5.26   Space filling diagram of 6-methylisocytosine
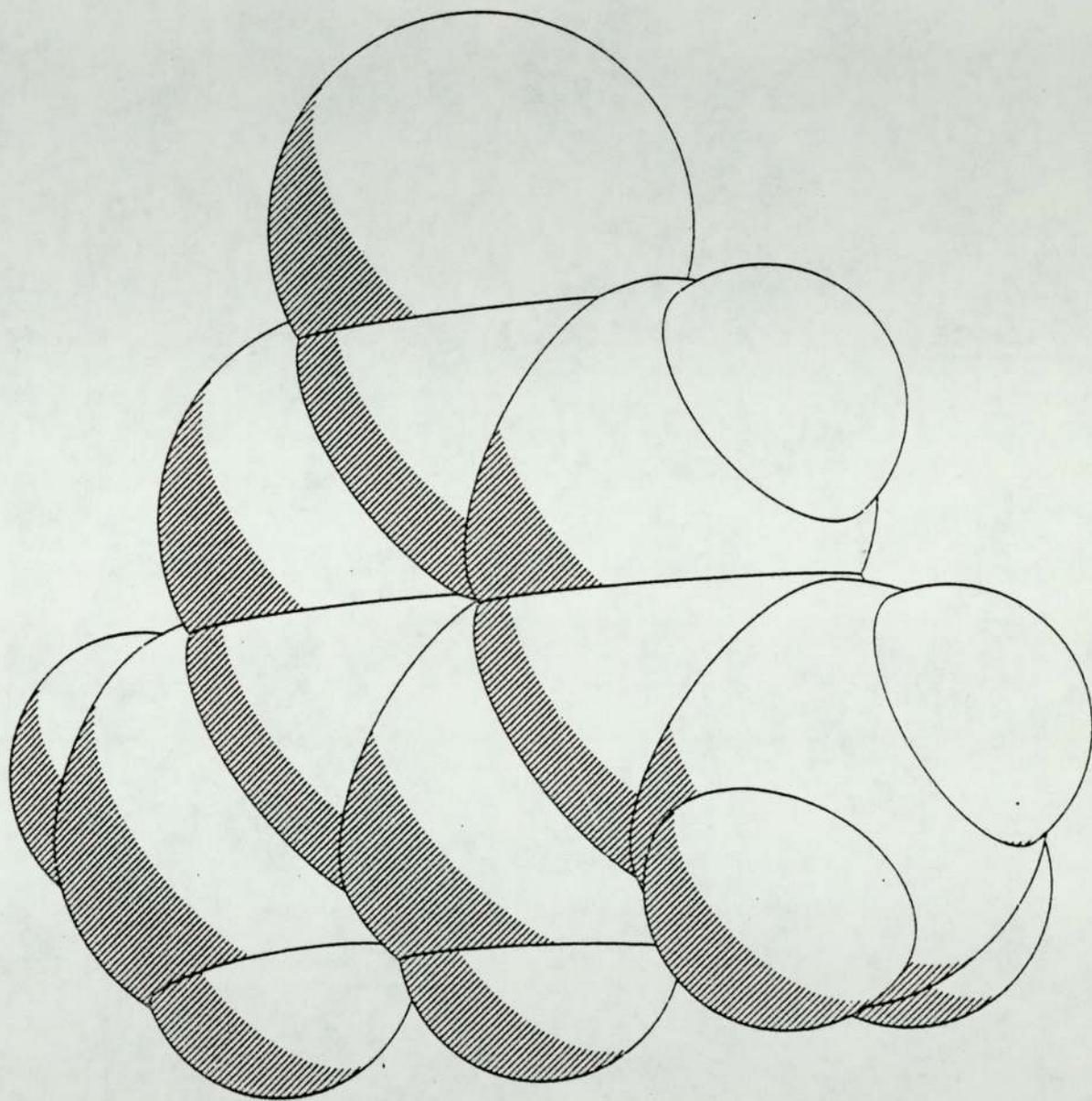
Figure 5.27   Resonance forms for 6-methylisocytosine

i

ii

iii

iv

v

vi
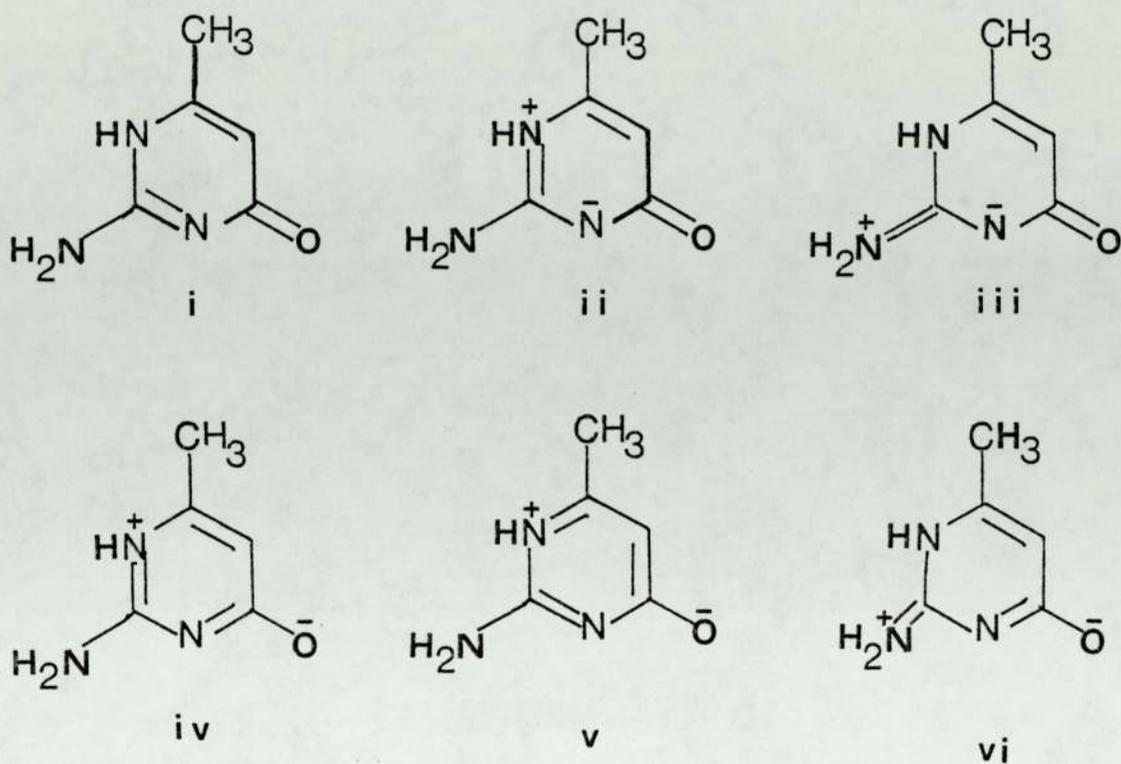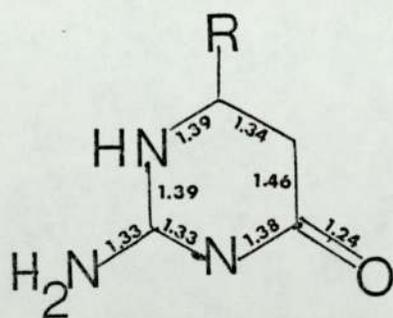
Figure 5.28

Table 5.25

I   - 35%

II   - 10%

III - 15%

IV   - 10%

V   - 10%

VI   - 20%

5.4.4   Results and Discussion

The structure of 6-methylisocytosine as determined crystallographically is shown in Figure 5.21, together with the numbering scheme used; positional parameters for all atoms are given in Table 5.17, with anisotropic thermal parameters for non-hydrogen atoms and isotropic temperature factors for hydrogens given in Table 5.18.

The bond distances and interatomic angles are given in Tables 5.19 and 20 respectively. Inspection of the C(4)-O(8) carbonyl bond distance $(1.256(3)\overset{o}{A})$ confirms the keto tautomeric form, and location of H(10) bonded to N(1) from a difference Fourier synthesis confirms this to be the tautomer shown in Figure 5.6 II. The correct location of this proton, and hence identification of the tautomeric form is further substantiated by the double bond nature of C(2)-N(3) $(1.331(3)\overset{o}{A})$ as opposed to N(1)-C(2) $(1.361(3)\overset{o}{A})$. This constitutes a major difference between the structures of 6-methylisocytosine and isocytosine[84]. With isocytosine, both tautomers (ie. protonated at N(1) and N(3)) are present in a 1:1 relationship and are hydrogen bonded to each other in a manner analogous to that proposed for guanine and cytosine pairing in DNA Fig 3b; with only the one tautomer present in the 6-methyl compound, such hydrogen bonding is prohibited.

The possible resonance forms of this tautomer of 6-methyl isocytosine are shown in Fig 5.27. The observed bond distances Table 5.19 can be explained by considering relative contributions from each of these canonical structures. The C(2)-N(3) bond $(1.331(3)\overset{o}{A})$ shows considerably more double bond character than either the N(1)-C(2) $(1.361(3)\overset{o}{A})$ or the N(3)-C(4) $(1.359(3)\overset{o}{A})$. The C(5)-C(6) bond $(1.350(3)\overset{o}{A})$ in a similar fashion has a strong double bond character,

suggesting that a major contribution to this structure comes from resonance form I in Fig 5.27. The C(2)-N(7) bond distance (1.326(3)Å) is characteristic of this type of bond in this group of compounds, eg. isocytosine[84] (1.323Å and 1.324Å), guanine monohydrate[91] (1.333Å). This can be explained by considering the overall structure to have some contribution from resonance forms III and VI in Figure 5.27. The sp^2 hybridized nature of N(7) is further confirmed by the sum of the bond angles about it, which is 360(4)°, Table 5.20. (This value should, however, be treated with some caution bearing in mind the constraints placed upon the N(7)-H(2) and N(7)-H(3) bonds during refinement). The C(4)-O(8) bond (1.255Å) is similar to that found in both tautomers of isocytosine[84] (1.246Å and 1.248Å) albeit slightly longer, and as with isocytosine O(8) participates in two hydrogen bonds. With cytosine, however, this bond distance is somewhat shorter (1.234Å) although its carbonyl oxygen also participates in two hydrogen bonds. This suggests that conjugation in the pyrimidine ring which contributes to the lengthening of the carbonyl bond in ie. tautomeric forms IV, V and VI may well feature somewhat greater in 6-methylisocytosine than in isocytosine and cytosine.

Sharma et al[84] have calculated the bond distances for isocytosine Fig 28 which result from the contributions shown in Table 5.25 of the canonical structures in Fig 5.27; these compare quite favourably with the distances determined experimentally for 6-methylisocytosine.

The bond angles (Table 5.19) show the nitrogen valence angles at N(1) and N(3) to be 120.4(2)° and 118.5(2)° respectively. Whilst this does not comply exactly with Chatar Singh's Rule, application of the Valence Shell Electron Pair Repulsion Theory still implies that

the location of the hydrogen attached to N(1) rather than N(3) is correct. Other internal angles in the pyrimidine ring are, as expected, close to 120°. The C(5)-C(6)-C(9) and N(1)-C(6)-C(9) bond angles (125.4(2)° and 115.8(2)°) suggest some lateral displacement of the methyl group.

The deviations in $\overset{o}{A}$ from the least squares plane through the pyrimidine ring atoms are given in Table 5.21. As can be seen, the whole molecule is essentially planar, with small deviations occurring in N(3) (0.012$\overset{o}{A}$), C(4) (-0.012$\overset{o}{A}$), N(7) (-0.020) and O(8) (-0.033$\overset{o}{A}$). A simple interpretation of Figure 5.23 shows the molecules to be planar and lying on or closely parallel to the 1 0 3 planes. This observation is confirmed by the fact that 1 0 3 reflection has the largest observed structure amplitude (approximately three times as large as the next highest value) and consequently the largest E value. The interplanar spacing between the 1 0 3 planes is 3.28$\overset{o}{A}$, and agrees well with the distance between the least squares planes through adjacent stacking molecules of 3.26$\overset{o}{A}$. Equivalent stacking distances in N-methylcytosine[93], cytosine[92] and isocytosine[84] are 3.40$\overset{o}{A}$, 3.36$\overset{o}{A}$ and 3.36$\overset{o}{A}$ respectively. Important intermolecular contacts involved in stacking are given in Table 5.24 and these are indicative of distances corresponding to normal Van der Waals electrostatic interaction. The packing arrangement of the molecules is shown in Figures 5.23, 24 and 25 with the hydrogen bonding scheme shown in dotted lines. Details of distances, angles and symmetry operations involved in the intermolecular hydrogen bonds are given in Table 5.23. Molecules related by a centre of symmetry, form a planar base pair interaction in the direction of a via N(7)-H(2) ... N(3) hydrogen bonds (2.973$\overset{o}{A}$); this same interaction is found in the structure of N-methylcytosine[93]. In the direction of b, molecules

related by a unit translation in <u>b</u> form a unique type of hydrogen bond. The interaction involves $O(8)$ (x, y, z) forming hydrogen bonds with $N(1)$ and $N(7)$ in the same molecule (x, 1+y, z) via the following contacts:-

$N(1)-H(1)$ ... $O(8)$ (2.716Å) and $N(7)-H(3)$ ... $O(8)$ (2.855Å).
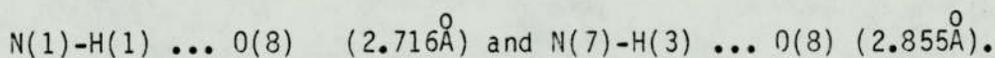
The result of this is to produce two opposing chains of molecules in the direction of b, linked by the base pair interaction in the direction of <u>a</u>.

As stated earlier, 6-methylisocytosine is quite different in structure to isocytosine and it is not immediately obvious why either the lack of, or addition of, a methyl group should effect such changes. By far a greater resemblance in intermolecular interactions is noted between 6-methylisocytosine and N-methylcytosine which forms the same base pair interaction. However, all of the bases mentioned in this section tend to form pairwise arrangements via hydrogen bonding and maximise their use of both donors and receptors. The diketo compounds achieve this by $N(3)$ to carbonyl oxygen pairs this being the case for all of the diketo structures mentioned in this section. With the exception of 6-benzyl-5-azauracil which has a substantially non-coplanar and 'bulky' benzyl group, the other more planar molecules tend to stack with an interplanar distance of between 3.25-3.4Å (Table 5.25). MacIntyre et al[94] interpret the close spacing in 8-azaguanine (3.25Å) and hence in 6-methylisocytosine (3.26Å) as indicating charge transfer interactions; however, Slatten et al[95] point out this is not unusually close for nucleic acid bases.

Table 5.25   Interplanar (stacking) distances

Compound	Stacking Distance $\overset{o}{A}$
6-phenyl-5-azauracil	3.31
6-methylisocytosine	3.26
cytosine	3.36
N-methylcytosine	3.40
isocytosine	3.36
guanine monohydrate	3.30
dihydrouracil	3.34
8-azaguanine	3.25

5.5   Spectroscopic Data

The spectroscopic data relating to the carbonyl stretching frequencies of uracil, 5-azauracil, 6-azauracil and related compounds are shown in Table 5.26a. It is noticeable that with the exception of the dihydrocompounds (and 6-methylisocytosine which has only one C=O group) all show two peaks in this region with a separation ranging from 30-70 cm^{-1}; the dihydrocompounds showing only a single peak.

Examination of the O-H stretching region of compounds (I), (V), (VI), (VIII) and (X) shows no evidence for the presence of -OH in them. This is in agreement with the crystal structures for the three compounds in this section in that both suggest the absence of any enol form in the solid state. On considering the mean of the two ν(C=O) frequencies for each compound (Table 5.26b) as being representative of the respective force constants involved, it can be seen that they follow the order 5-azauracil and derivatives> 6-azauracil and derivatives > uracil. Direct comparison of mean ν(C=O) with crystallographically determined bond lengths is made difficult however in that the two carbonyl bond distances within the molecules of 6-phenyl-5-azauracil and 6-azauracil are almost identical, whilst the 6-benzyl compound and uracil have differing carbonyl bond lengths. This indeed poses the question as to why there are two carbonyl stretching frequencies observed in the infrared spectrum. An immediate response might be that of non-equivalence of the two C-O bonds, yet as just stated, in two of the compounds these bond distances are almost identical, whilst in dihydrouracil with marginally differing C-O bond distances only a single C-O absorption is seen.

Table 5.26a Carbonyl stretching frequencies of uracil, 5-azauracil and related compounds.

COMPOUNDS	$(C=O)cm^{-1}$ (nujol)
1. uracil	1730-1715 , 1670
2. 5-methyl-6-azauracil	1720 , 1675
3. 5-azauracil	1740 , 1710
4. 6-methyl-5-azauracil	1760 , 1705
5. 6-phenyl-5-azauracil	1740 , 1685
6. 6-benzyl-5-azauracil	1750 , 1680
7. 5,6 dihydro-6-methyl-5-azauracil	1727 (dioxan)
8. 6-azauracil	1730 , 1695
9. dihydrouracil	1709
10. 6-methylisocytosine	1660

Table 5.26b Mean values of the two $(C=O)$ stretching absorptions where applicable

COMPOUND	MEAN VALUE OF THE TWO $(C=O)$ ABSORPTIONS cm^{-1}
uracil	1696
5-methyl-6-azauracil	1698
5-azauracil	1725
6-methyl-5-azauracil	1732
6-phenyl-5-azauracil	1713
6-benzyl-5-azauracil	1715
6-azauracil	1712
6-methylisocytosine	1660

Table 5.27   Torsion angles for dihydrouracil, uracil, 6-azauracil, 6-phenyl-5-azauracil and 6-benzyl-5-azauracil

URACIL NUMBERING SCHEME	DIHYDROURACIL[88]	URACIL[87]	6-AZAURACIL[90]	6-PHENYL-5-AZAURACIL	6-BENZYL-5-AZAURACIL
C(6)-N(1)-C(2)-N(3)	13.0°	0.0°	-3.1	-3.0°	-3.1°
N(1)-C(2)-N(3)-C(4)	11.1°	-0.5°	-3.0°	-2.0°	.4°
C(2)-N(3)-C(4)-C(5)	-3.0°	0.8°	2.2°	6.8°	1.1°
N(3)-C(4)-C(5)-C(6)	-26.2°	-0.5°	-0.8°	-6.4°	.1°
C(4)-C(5)-C(6)-N(1)	-45.4°	0.0°	0.3°	1.7°	-2.9°
C(5)-C(6)-N(1)-C(2)	-41.0°	0.2°	-1.2°	3.3°	4.6°
C(6)-N(1)-C(2)-O(2)	-166.9°	179.7°	-179.0°	177.3°	177.3°
O(2)-C(2)-N(3)-C(4)	168.9°	179.7°	178.3°	177.7°	-180.0°
C(2)-N(3)-C(4)-O(4)	176.4°	-179.0°	-178.8°	174.2°	-179.3°
O(4)-C(4)-C(5)-C(6)	154.3°	179.3°	-179.7°	174.6°	-179.5°

NOTE   The numbering scheme refers to uracil.   For 6-azauracil replace C(6) with N(6), and for the 6-substituted-5-azauracil replace C(5) with N(5).

The torsion angle A-B-C-D is the projected angle between AB and CD when viewed down the B-C bond; the clockwise rotation of the C-D bond with reference to bond A-B is considered positive.

The main criteria for a non-equivalence argument would be a significant difference in the carbonyl bond lengths or a significantly different environment for the two groups, be it in terms of a mass effect or force constants of adjacent bonds. In the case of 6-phenyl and 6-benzyl-5-azauracil, any mass effect would be marginal and limited to the difference of a single hydrogen atom located at $N(1)$ but not $N(5)$, whilst bond length data show that in both compounds the carbonyl groups have a similar environment in terms of nearest neighbour bonds. As for the $C=O$ bonds in particular, in 6-phenyl-5-azauracil these are insignificantly different [Δ d=0.004(3)$\overset{o}{A}$], and in 6-azauracil are determined crystallographically, to be equal [Δ d=0.000(4)$\overset{o}{A}$]. Since the separation of the two carbonyl frequencies in most cases is greater than 35 cm^{-1}, a non-equivalence argument appears inadequate to explain the occurrance of the two bands. It can also be seen that in the case of 6-methylisocytosine which has only a single carbonyl group, with an equivalent environment, the single stretching frequency occurs at 1660 cm^{-1}; this is of the order of 80-100 cm^{-1} lower than the upper bands of the 6-substituted-5-azauracils.

A more adequate explanation of the observed separation of these bands is that of frequency interactions between the two carbonyl groups. The extent to which frequency coupling between such groups occurs depends upon a number of factors[83]. These include (a) the proximity of the two groups, (b) the projected angle between the two groups, (c) the degree to which individual frequencies would approach each other, and (d) the degree of planarity which the system imposes on the two groups. The first of these is of less importance since interaction is known to persist in 1,4 systems where there is sufficient coplanarity. The degree to which the angle between the

groups affects the coupling is related to the fact that carbonyl bonds at right angles to each other are unable to interact, whereas, those in CO_2 for example, may do so to a maximum extent, this being a simple mechanical phenomenon.

In considering the 6-phenyl and 6-benzyl-5-azauracils, and analogous compounds, eg. 6-azauracil and uracil, we find that these all satisfy requirements (a) and (b) by being 2,4 substituted and all having angles between the carbonyl groups projected onto the plane of the ring, of 117-118°. Each of these compounds, and especially 6-phenyl-5-azauracil and 6-azauracil have their carbonyl bond distances of similar magnitudes and one might therefore suggest that their carbonyl frequencies are sufficiently similar to interact.

The last requirement is that of planarity and can be shown to be of prime importance. Deviations in $\overset{o}{A}$ from the least squares planes through the triazine rings in the 6-phenyl and 6-benzyl compounds show the carbonyl groups in both cases to be essentially planar with the triazine rings. This can be futher demonstrated by examining the torsion angles involved (Table 5.27) in the triazindione system in these and related compounds.

Uracil, 6-azauracil and the 6-phenyl and 6-benzyl-5-azauracils all have torsion angles in the ring approaching 0°, and those involving the carbonyl groups approaching 180°, suggesting a good degree of planarity in each case. As previously stated, each of these compounds shows two distinct peaks in the carbonyl stretching region. However, the torsion angles in the dihydrocompound (dihydrouracil) and more important those involving the carbonyl groups, demonstrate a significant deviation from planarity, and in this case only a single absorption is noticable in the carbonyl stretching region. Since the carbonyl bond lengths in dihydrouracil

(1.21 and 1.22 $\overset{O}{A}$)[88] are again almost equal, it would appear to be the lack of planarity due to the absence of the double bond in the ring which seriously reduces the coupling effect in this case. With the four other compounds in Table 5.27, planarity is maintained by the presence of a double bond in the (5)-(6) position in the ring.

In general, the results of this work have shown by a crystallographic study and a comparison with other model compounds that:-

a)   The structures studied in this section exist solely in the keto tautomeric form in the solid state.

b)   The 6-phenyl and 6-benzyl-5-azauracils satisfy the general requirements for good vibrational coupling between the carbonyl groups and that in this case, a non-equivalence argument is insufficient to explain the occurrance of the two bands in the   (C=O) stretching region.

CHAPTER 6

CHAPTER 6   The crystal structure of 8-carbamoyl-3-(2-chloroethyl)
imidazo [5,1-d]-1,2,3,5-tetrazine-4(3H)-one.   A Novel
Bicyclic Antitumour Agent, and Two of its Analogues.

6.1   As mentioned earlier, the approach to controlling cancer is a
multifaceted attack, of which chemotherapy is just one of the
components used.   The range of drugs presently available is of the
same order as the number of different cancers needing treatment.   The
cytotoxic drugs used in the treatment of cancer like ionising
radiation, do not kill tumour cells directly but affect the cell
division process, and thereby cell proliferation.   Exactly at which
point the cell division process is affected depends upon the drug
used, e.g. antipyrimidines and antipurines inhibit the synthesis of
DNA, whilst certain alkylating agents such as BCNU or CCNU may cause
cross linking or breakage of DNA thus inhibiting the process of
replication.

A particular class of small molecules bearing NNN linkages
either as 1,2,3 cyclic triazines or acyclic triazenes possess
versatile chemical reactivity upon which attempts have been made to
achieve selective antitumour effects[96-98].   One such compound 5-(3,3-
dimethyltriazenyl) imidazole-4-carboxamide (DTIC) is currently in
clinical use as a single agent for the treatment of malignant
melanoma [99] and in combination regimes against soft tissue
sarcomas[100] and Hodgkin's disease[101].   The consensus of views on the
mode   of   action   of   DTIC   and   the   family   of   1-aryl-3,3-
dialkyltriazenes[102] favours the metabolic activation hypothesis and
implicates the monomethyl-triazene (8,9; MTIC) as the bioactive
species[103] possibly after enzymic hydroxylation as shown in Figure
6.2[104].   Shealy and co-workers[105,106] and Skibba et al[104] suggest

final decomposition of DTIC and DCTIC (4) to diazeno derivatives finally liberating alkylating reagents.

Two related imidazoles are also worthy of note: the bis(chloroethyl)triazene (4; DCTIC) which has potent antitumour activity against L1210 leukemia in mice[107] but proved disappointing in the clinic[108]; and the monochloroethyltriazene (5,6; MCTIC) which is too unstable to be a realistic clinical candidate although it demonstrates good antitumour activity in animal systems[109]. Shealy and co-workers suggested again that DCTIC decomposed via a diazeno derivative to liberate for instance $NH(C_2H_4Cl)_2$[110] nornitrogen mustard, a known alkylating agent.

Results of activity-structure studies on drugs related to aryltriazenes indicated that the terminal substituents on the triazene group were important in determining the potency[111]. It appeared that at least one of the substituents must be a methyl group, or the compound is inactive. The exception to this rule is DCTIC which shows a different specificity to various cancers than DTIC and possibly acts by a different mechanism, the former giving rise to a bifunctional alkylating agent, the latter a monofunctional one. Connors and co-workers[112] came to similar conclusions in working with the arylalkyltriazenes and found that large alkyl groups suppress the antitumour activity substantially and for example compounds with tertiary butyl groups are totally inactive even when the other group is a methyl. These results were interpreted by Connors as follows; only those compounds which can be metabolised to the arylmonomethyl (or monochloroethyl) triazenes have antitumour activity (at least against TLX5 lymphoma). This principle would appear to confirm the importance of the monoalkyltriazene which has no OC-CH group amenable to metabolic oxidation and thus cannot be metabolised to the monomethyltriazene.

Figure 6.1
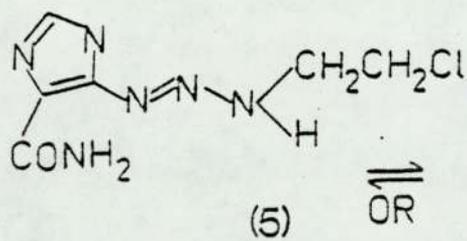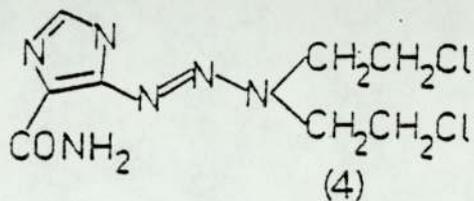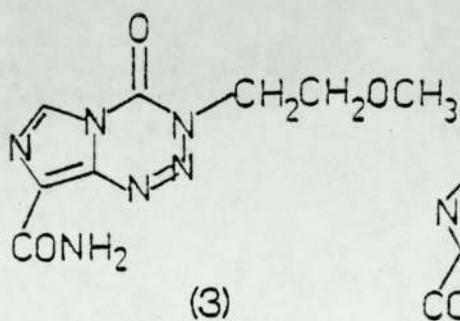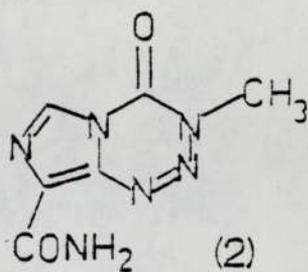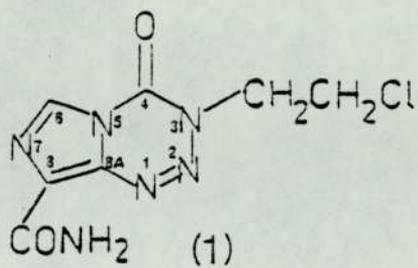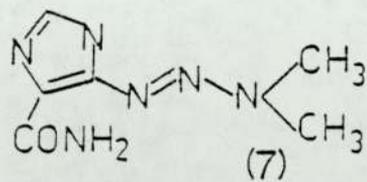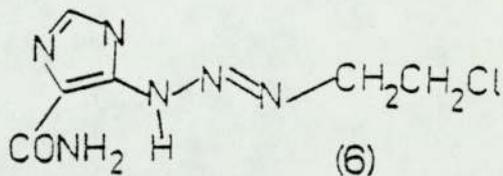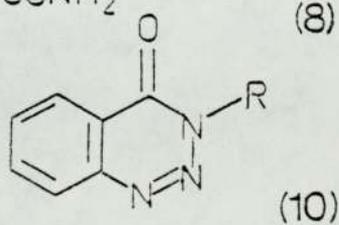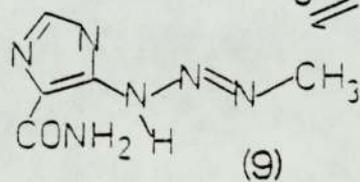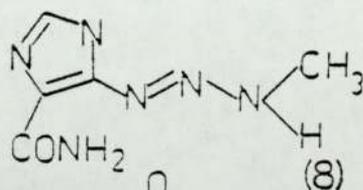
Figure 6.2
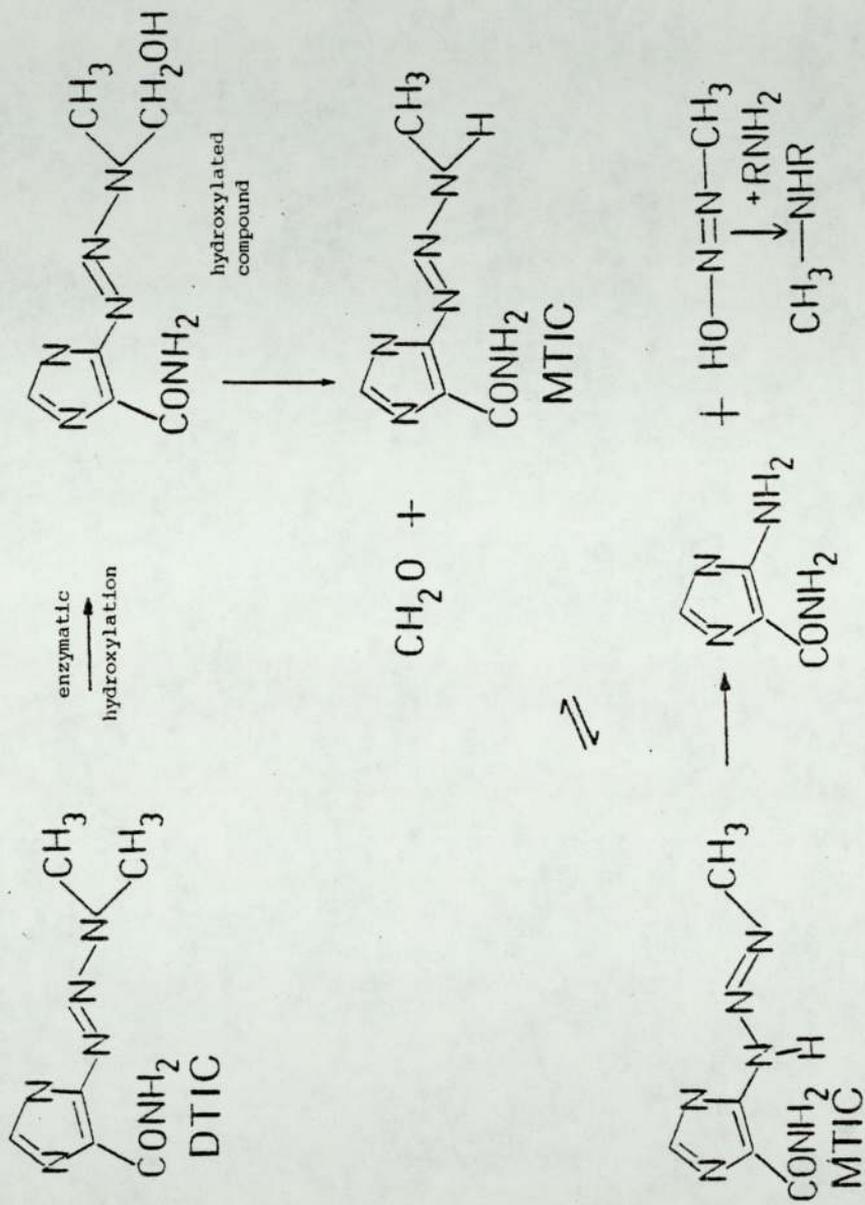
Stevens et al[113,114] having worked on similar cyclic triazines and acyclic triazenes were interested in the effect which may be elicited by producing a bicyclic triazene with a bridgehead nitrogen atom but nevertheless, containing some of the important features of the known antitumour triazenes, e.g. the imidazocarboxamide moiety and either a chloroethyl or methyl exocyclic substituent at N(3). On the basis of their previous work involving cyclic variants of 3-(substituted) 1,2,3-benzotriazin-4[3H]-ones[113] [e.g. 10] they found that depending upon conditions, these underwent fission at the 1,8a-, 2,3- or 3,4 bonds whilst ring opening of the 3-(substituted)-imidazo[5,1-c]-1,2,4-triazin-4(3H)-ones in the presence of hydrazines proceeded by cleavage of the 4,5 bond only[115]. An attractive proposition was the possibility of 3-(substituted)-imidazo[5,1-d]-1,2,3,5-tetrazin-4(3H)ones [e.g. 1,2 and 3] cleaving at four different bonds (1,8a; 2,3; 3,4; and 4,5) based on the precedents of the aforementioned compounds, and thus producing a cascade of reactive moieties some of which are known to possess potent antitumour activity.

Compound (1) the chloroethyl derivative (mitozolomide, CCRG 81010) was the first of the series to be synthesised from the new bicyclic ring system[116] and entered Phase 1 clinical trial in 1983. The potent broad-spectrum antitumour activity of (1) in experimental systems[116,117] was found to be markedly superior to that of its structurally related triazene DTIC (7) and in the Day 1 and Days 1-9 schedules in tests against L1210 leukemia, closely parallelled that reported for MCTIC (5,6)[109]. Compound 1 also showed pronounced activity against P388 leukemia with 100% cures being produced at more than one dose level in Day 1 and Day 1-5 schedules. The drug has activity rated (++) on the NCI activity scale against B16 melanoma,

Colon 38 tumour, Lewis lung carcinoma and the LX-1 lung tumour xenograft[118]. The compound also has curative activity against TLX-5 lymphoma, M5076 reticulum cell sarcoma, and ADJ/PC6A plasmacytoma in mice[119] but is inactive against CD8F mammary tumour, the CX-1 colon tumour xenograft and the MX-1 mammary tumour xenograft[118].

Because of the obvious importance of this compound from _in vivo_ and _in vitro_ tests on its antitumour activity, and the commercial and clinical interests in it, unequivocal proof of its structure was required. At this stage, however, new analogues of the compound were already being prepared, both at Aston and by a drug company. Two such analogues were compounds 2 and 3, these being the 3-methyl, and 3-methoxyethyl substituted ones respectively. These compounds were of structural interest in that they maintained the essential features of compound 1 with the only variant being the N(3) substituent; their activities were, however, substantially reduced, compound 2 being still active but compound 3 having no antitumour activity at all. The crystal structures of each of these compounds has been studied in this work, not only in order to provide unequivocal proof of structure but also in the hope of providing additional information which may assist in determining possible modes of action.

6.2.1   Preparation of 8-carbamoyl-3-(2-chloroethyl) imidazo [5,1-d]-
        1,2,3,5-tetrazine-4(3H)-one[116] (compound 1)

A suspension of 5-diazoimidazole-4-carboxamide (0.3 g) in
anhydrous dichloromethane (10 ml) was stirred with 2-
chloroethylisocyanate (1.0 ml) for 20 days in the dark at 25 °C.
Addition of ether (30 ml) to the mixture gave a roseate precipitate
(90%) of the title compound (1) which was collected and washed with
ether.

Well formed plate-like crystals of the compound were grown by
slow evaporation from 50% aqueous acetone. The product has a melting
point of 158 °C.
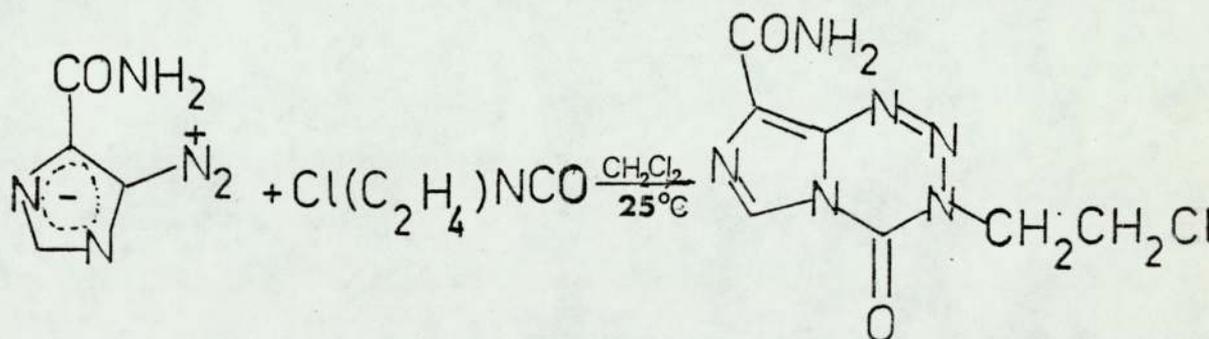
Preparative work by courtesy of Mrs G U Baig, Department of Pharmacy,
University of Aston.
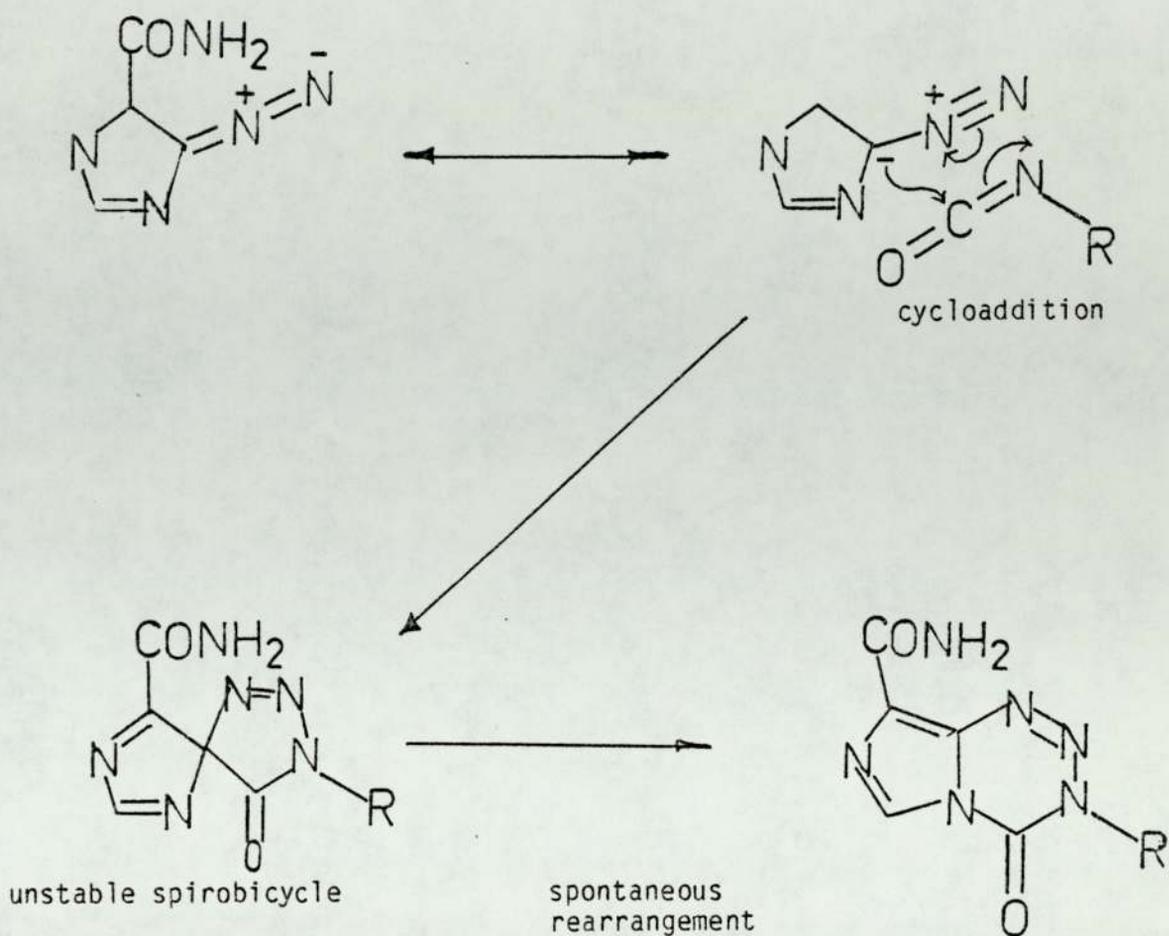
Figure 6.3

Probable mechanism, suggested by Stevens et al[116], for the formation of CCRG 81010 and its two analogues from 5-diazoimidazole-4-carboxamide and an appropriately substituted isocyanate.

1: R = $-CH_2CH_2Cl$ (CCRG 81010)

2: R = $-CH_3$

3: R = $-CH_2CH_2OCH_3$

6.2.2   Crystal Data

8-carbamoyl-3-(2-chloroethyl)   imidazo[5,1-d]-1,2,3,5-tetrazin-4(3H)-one ($C_7H_7N_6O_2Cl$) crystallises in the space group $P\bar{1}$ with a = 7.003(4)Å, b = 8.680(4)Å, c = 16.041(9)Å, α = 93.76(5)°, β = 93.99(5)°, γ = 92.08(7)°, Z = 4, V = 969(1)Å^3. The molecular weight is 242.5 (F(ooo) = 536), and the density Dm (by flotation in CCl_4 and C_2H_5I) = 1.670 gcm^{-3} with the calculated density Dc = 1.660 gcm^{-3}. The absorption coefficient for MoKα radiation (0.71069Å) is μ= 3.44 cm^{-1}.

6.2.3   Structure Analysis

Data Collection

The data were collected from a crystal of dimensions 1.2 x 0.8 x 0.35 mm on an Enraf-Nonius CAD4 diffractometer using the ω-2θ scan technique with monochromated Mo-Kα radiation λ = 0.71069Å.   Two reference reflections were measured every 3600 sec of X-ray exposure time and the orientation of the crystal checked after every 150 reflections.   No decomposition or movement of the crystal was detected during the collection of the 3792 unique reflections between θ = 2° and θ = 25°.

Structure determination and refinement

The data reduction was performed at Birmingham University using a programme provided by Dr T A Hamor. Lorentz and polarisation and a simple linear intensity correction were applied and the data output in the form HKL, Fobs and σ Fobs where the latter is based on counting statistics together with an allowance for instrument instability.   No correction to the data was made for absorption effects.   Normalised structure factors were calculated using SHELX[2] and the overall scale factor and estimated temperature factor determined as 0.5623 and 0.061 respectively.

Phase determination was undertaken using SHELX automatic centrosymmetric direct methods. E min was set to 1.1 and the number of phase permutations NP to 14 (i.e. 2^{14} permutations). From this a trial structure was produced in which all but one of the non-hydrogen atoms appeared on an E-map. Subsequent Fourier refinement determined the position of the remaining atom. The solution which gave the correct structure was the one with the second highest figures of merit and the programme assigned the following origin determining reflections; $3\bar{8}11$ (+), $2\bar{4}5$ (+), and $\bar{3}\bar{7}2$ (+).

Full matrix least squares refinement of positional parameters and isotropic temperature factors with only unit weights, ensued after which $R = \sum [||Fo| -S |Fc| |]/ \sum|Fo|$ where S is a scale factor decreased to 0.14. At this point nine of the fourteen hydrogen atoms predicted in the structure appeared on a difference Fourier synthesis. Further refinement of co-ordinates and anisotropic thermal parameters for non-hydrogen atoms and co-ordinates and isotropic temperature factors for hydrogens reduced R to 0.062; then the remaining five hydrogen atoms were located in an (Fo-Fc) synthesis.

A final sequence of refinements by SHELX for the complete structure using the constraints described in the immediately previous cycles resulted in an unweighted discrepancy index R = 0.052 for the 3244 observed data with Fobs>3σ. In the latter stages of refinement, reflections were weighted according to $W = K/[\sigma^2(Fo) + qFo^2]$ where q is a refinable parameter which converged at 0.0013 and K which should approach unity if there are no serious systematic errors equalled 4.710. When refinement was terminated, no positional parameter shifted by more than 0.22σ and the weighted discrepancy index

$Rg = [\sum w(|Fo| - S |Fc|)^2 / \sum w |Fo|^2]^{1/2} = 0.074$ where S and w are the scale and weighting factors previously described. A final difference electron density map showed no feature greater than 0.35 eÅ^{-3}.

During the refinement procedure, atoms were assigned their correct atomic scattering factors on the basis of the structure proposed from IR, NMR, MS and CHNO analysis. The isotropic temperature factors during refinement and the final discrepancy index and difference electron density map appeared to confirm the correctness of these assignments.

Table 6.1 Positional parameters (fractional co-ordinates x 10^4) with estimated standard deviations in parentheses.

	X/a	Y/b	Z/c
N(1)	1206(3)	4413(2)	7329(1)
N(2)	1978(3)	3844(2)	6694(1)
N(3)	1020(3)	2687(2)	6187(1)
C(4)	-768(3)	2037(2)	6287(1)
N(5)	-1549(2)	2715(2)	6993(1)
C(6)	-3258(3)	2431(2)	7337(1)
N(7)	-3394(3)	3325(2)	8016(1)
C(8)	-1743(3)	4223(2)	8133(1)
C(8A)	-576(3)	3870(2)	7503(1)
C(31)	2052(3)	2119(2)	5469(1)
C(32)	1398(4)	2789(3)	4662(1)
Cl	2336(1)	4721(1)	4623(1)
O(4)	-1567(3)	1028(2)	5832(1)
C(81)	-1453(3)	5368(2)	8869(1)
N(82)	267(3)	6068(3)	8996(1)
O(82)	-2768(3)	5644(3)	9306(1)
N(1)'	671(2)	1003(2)	2365(1)
N(2)'	2022(2)	1941(2)	2629(1)
N(3)'	3666(2)	2061(2)	2200(1)
C(4)'	3995(3)	1279(2)	1452(1)
N(5)'	2457(2)	262(2)	1181(1)
C(6)'	2194(3)	-782(2)	505(1)
N(7)'	570(3)	-1559(2)	504(1)
C(8)'	-310(3)	-1031(2)	1206(1)
C(8A)'	849(3)	109(2)	1637(1)
C(31)'	5143(3)	3125(2)	2622(1)

	X/a	Y/b	Z/c
C(32)'	6694(3)	2288(3)	3079(2)
CL'	5770(1)	1299(1)	3195(1)
O(4)'	5420(2)	1437(2)	1081(1)
C(81)'	-2181(3)	-1630(2)	1464(1)
N(82)'	-3159(3)	-2673(2)	940(1)
O(82)'	-2717(2)	-1180(2)	2149(1)
H(1)	1802(29)	997(27)	5402(17)
H(2)	3384(26)	2348(23)	5596(13)
H(3)	1967(29)	2220(27)	4123(18)
H(4)	38(29)	2746(25)	4593(15)
H(5)	-4200(27)	1729(25)	7150(16)
H(6)	545(35)	6681(33)	9467(24)
H(7)	1118(35)	5789(32)	8671(23)
H(1)'	4441(26)	3786(24)	3023(15)
H(2)'	5716(28)	3728(26)	2146(16)
H(3)'	7573(31)	3000(29)	3340(19)
H(4)'	7260(28)	1488(25)	2688(16)
H(5)'	3044(27)	-901(25)	124(15)
H(6)'	-2766(30)	-2954(27)	458(18)
H(7)'	-4377(30)	-3084(26)	1106(17)

	X/a	Y/b	Z/c
C(32)'			
CL'			

TABLE 6.2 Anisotropic temperature factors (non-hydrogen atoms)

ATOM	U11	U22	U33	U23	U13	U12
N(1)	.0401(9)	.0295(8)	.0274(9)	.0012(7)	.0074(7)	-.0088(7)
N(2)	.0418(9)	.0341(9)	.0300(9)	.0022(7)	.0066(7)	-.0071(8)
N(3)	.0408(9)	.0272(8)	.0264(8)	.0019(6)	.0051(7)	.0000(7)
C(4)	.0451(11)	.0257(9)	.0232(9)	.0038(7)	.0032(8)	-.0009(8)
N(5)	.0366(8)	.0233(8)	.0237(8)	.0028(6)	.0030(6)	-.0040(7)
C(6)	.0350(10)	.0299(10)	.0332(10)	.0049(8)	.0032(8)	-.0065(8)
N(7)	.0351(9)	.0359(9)	.0343(9)	.0059(7)	.0064(7)	.0009(7)
C(8)	.0372(10)	.0256(9)	.0269(10)	.0060(7)	.0035(8)	.0018(8)
C(8A)	.0373(10)	.0214(9)	.0245(9)	.0036(7)	.0023(7)	-.0028(7)
C(31)	.0474(12)	.0342(11)	.0299(10)	.0027(8)	.0100(9)	.0103(9)
C(32)	.0485(12)	.0492(13)	.0286(10)	.0041(9)	.0102(9)	.0071(10)
Cl	.0705(4)	.0507(4)	.0710(5)	.0275(3)	.0284(4)	.0159(3)
O(4)	.0631(10)	.0398(9)	.0369(9)	-.0114(7)	.0066(7)	-.0144(8)
C(81)	.0443(11)	.0319(10)	.0249(9)	.0032(8)	.0039(8)	.0070(9)
N(82)	.0596(12)	.0492(12)	.0418(12)	-.0170(9)	.0105(10)	-.0096(10)
O(82)	.0550(10)	.0749(13)	.0427(10)	-.0180(9)	.0123(8)	.0095(9)
N(1)'	.0300(8)	.0367(9)	.0280(8)	-.0033(7)	.0076(7)	-.0006(7)
N(2)'	.0318(8)	.0410(10)	.0271(8)	-.0020(7)	.0084(7)	-.0002(7)
N(3)'	.0288(8)	.0336(9)	.0288(8)	.0018(7)	.0080(6)	.0005(7)
C(4)'	.0308(9)	.0314(10)	.0274(9)	.0076(7)	.0069(7)	.0030(8)
N(5)'	.0297(8)	.0309(8)	.0232(8)	.0040(6)	.0077(6)	.0027(7)
C(6)'	.0369(10)	.0349(10)	.0245(9)	.0009(8)	.0095(8)	.0035(8)
N(7)'	.0383(9)	.0323(9)	.0252(8)	.0003(7)	.0060(7)	.0031(7)
C(8)'	.0327(9)	.0268(9)	.0224(9)	.0040(7)	.0033(7)	.0043(8)
C(8A)'	.0268(9)	.0303(9)	.0239(9)	.0053(7)	.0074(7)	.0035(7)

ATOM	U11	U22	U33	U23	U13	U12
C(31)'	.0354(10)	.0288(10)	.0366(11)	.0038(8)	.0062(8)	-.0024(8)
C(32)'	.0304(10)	.0370(11)	.0496(13)	.0037(10)	.0031(9)	-.0021(9)
CL'	.0626(4)	.0522(4)	.0558(4)	.0233(3)	.0041(3)	.0137(3)
O(4)'	.0359(8)	.0483(9)	.0364(8)	.0028(7)	.0161(6)	-.0011(7)
C(81)'	.0323(9)	.0264(9)	.0315(10)	.0051(8)	.0055(8)	.0030(8)
N(82)'	.0358(9)	.0450(11)	.0358(10)	-.0035(8)	.0079(8)	-.0068(8)
O(82)'	.0433(8)	.0378(8)	.0378(9)	-.0033(7)	.0152(6)	-.0073(7)

ATOM	U11	U22	U33	U23	U13	U12
C(31)'	.0354(10)	.0288(10)	.0366(11)	.0038(8)	.0062(8)	-.0024(8)
C(32)'	.0304(10)	.0370(11)	.0496(13)	.0037(10)	.0031(9)	-.0021(9)

TABLE **6.2b** Isotropic temperature factors for hydrogen atoms

ATOM	U iso
H(1)	.046(4)
H(2)	.024(3)
H(3)	.056(4)
H(4)	.042(4)
H(5)	.047(4)
H(6)	.102(4)
H(7)	.091(4)
H(1)'	.030(3)
H(2)'	.043(4)
H(3)'	.061(4)
H(4)'	.042(4)
H(5)'	.035(4)
H(6)'	.052(4)
H(7)'	.048(4)

Table 6.3 Bond distances (in Å) of the two molecules in the asymmetric unit (unprimed and primed) with estimated standard deviations in parentheses.

Bond	unprimed Å	primed Å
N(1)-N(2)	1.266(3)	1.257(3)
N(2)-N(3)	1.374(3)	1.386(3)
N(3)-C(31)	1.469(3)	1.460(3)
C(31)-C(32)	1.503(3)	1.504(3)
C(32)-C(1)	1.785(3)	1.785(3)
N(3)-C(4)	1.378(3)	1.377(3)
C(4)-O(4)	1.198(3)	1.205(3)
C(4)-N(5)	1.396(3)	1.394(3)
N(5)-C(6)	1.373(3)	1.364(3)
C(6)-N(7)	1.306(3)	1.300(3)
N(7)-C(8)	1.365(3)	1.381(3)
C(8)-C(8A)	1.369(3)	1.374(3)
C(8A)-N(5)	1.379(3)	1.391(3)
C(8)-C(81)	1.490(3)	1.487(3)
C(81)-O(82)	1.216(3)	1.229(3)
C(81)-N(82)	1.326(4)	1.330(3)
C(8A)-N(1)	1.371(3)	1.374(3)
C(31)-H(1)	0.98(2)	1.00(2)
C(31)-H(2)	0.95(2)	1.05(2)
C(32)-H(3)	1.08(3)	0.92(2)
C(32)-H(4)	0.95(2)	1.02(2)
C(6)-H(5)	0.90(2)	0.89(2)
N(82)-H(6)	0.90(3)	0.86(3)
N(82)-H(7)	0.85(3)	0.97(2)

TABLE 6.4   Important interatomic angles (^{0}) for the two molecules in the asymmetric unit (unprimed and primed) with estimated standard deviations in parentheses.

Atoms	Bond angle [^{0}] (unprimed)	Bond angle [^{0}] primed
C(8A)-N(1)-N(2)	119.7(2)	118.6(2)
N(1)-N(2)-N(3)	119.6(2)	120.7(2)
N(2)-N(3)-C(4)	126.8(2)	126.4(2)
N(2)-N(3)-C(31)	114.8(2)	113.6(2)
C(32)-N(3)-C(4)	118.4(2)	120.0(2)
N(3)-C(4)-N(5)	111.0(2)	110.8(2)
N(3)-C(4)-O(4)	125.2(2)	125.0(2)
O(4)-C(4)-N(5)	123.8(2)	124.2(2)
C(4)-N(5)-C(8A)	122.0(2)	122.1(2)
C(4)-N(5)-C(6)	131.4(2)	131.2(2)
C(6)-N(5)-C(8A)	106.6(2)	106.7(2)
N(5)-C(6)-N(7)	110.9(2)	111.4(2)
C(6)-N(7)-C(8)	106.9(2)	107.2(2)
N(7)-C(8)-C(8A)	109.8(2)	108.8(2)
N(7)-C(8)-C(81)	119.9(2)	125.5(2)
C(8A)-C(8)-C(81)	130.4(2)	125.7(2)
C(8)-C(81)-O(82)	120.2(2)	119.2(2)
C(8)-C(81)-N(82)	116.4(2)	117.2(2)
O(82)-C(81)-N(82)	123.4(2)	123.6(2)
C(8)-C(8A)-N(5)	105.8(2)	105.9(2)
N(1)-C(8A)-N(5)	120.9(2)	121.3(2)
N(1)-C(8A)-C(8)	133.2(2)	132.8(2)
C(81)-N(82)-H(6)	119.(2)	122.(2)
C(81)-N(82)-H(7)	118.(2)	118.(2)
H(6)-N(82)-H(7)	122.(3)	120.(3)

TABLE 6.5 Deviations (Å) of non-hydrogen atoms from the least squares plane through the imidazo-tetrazine ring system[+]. (The two molecules in the asymmetric unit are treated separately. Atoms used in the plane calculation are marked with an asterisk).

ATOM	DEVIATION [Å] unprimed	DEVIATION [Å] primed
*N(1)	-.004(3)	-.017(3)
*N(2)	.002(3)	-.002(3)
*N(3)	.007(3)	.035(3)
*C(4)	.003(3)	-.004(3)
*N(5)	-.010(3)	-.029(3)
*C(6)	-.003(3)	-.009(3)
*N(7)	.005(3)	.017(3)
*C(8)	.008(3)	.019(3)
*C(8A)	-.007(3)	-.010(3)
C(31)	.011(3)	.142(3)
C(32)	-1.346(3)	1.547(3)
Cl	-2.355(2)	2.646(2)
O(4)	.006(3)	-.006(3)
C(81)	.030(3)	.080(3)
O(82)	-.126(3)	.220(3)
N(82)	.188(3)	.019(3)

+ The equations of the plane are:-

unprimed - $0.422X - 0.705Y + 0.571Z + 4.221 = 0$

primed - $0.433X - 0.721Y + 0.541Z + 1.611 = 0$

where X, Y and Z are orthogonal co-ordinates in Å along a*, cxa* and c.

TABLE 6.6   Important torsion angles [^{o}]

Atoms	angle [^{o}] unprimed	angle [^{o}] primed
N(7)-C(8)-C(81)-N(82)	172.4	-5.8
N(7)-C(8)-C(81)-O(82)	-9.2	172.2
C(8A)-C(8)-C(81)-O(82)	170.9	-6.0
C(8A)-C(8)-C(81)-N(82)	-7.5	176.0
N(3)-C(31)-C(32)-Cl	76.8	-64.9
N(2)-N(3)-C(31)-C(32)	-99.6	101.0
C(4)-N(5)-C(8A)-N(1)	.7	2.5
C(4)-N(5)-C(8A)-C(8)	-178.6	-177.3
C(6)-N(5)-C(8A)-N(1)	179.6	179.7
C(6)-N(5)-C(8A)-C(8)	.3	0.0

TABLE 6.7   Hydrogen bond contact distances and angles

Hydrogen Bond	Angle at H [$^{\circ}$]	N....0 distance ($\overset{\text{o}}{\text{A}}$)
N(82)$_I$-H(7) .... N(1)$_I$	137	3.075
N(82)$_I$-H(6) .... N(7)'$_{II}$	166	3.063
N(82)'$_{II}$-H(6)' ... O(82)$_I$	156	2.951
N(82)'$_{III}$-H(7)'... N(7)$_I$	153	3.089

I, II, III refer to the symmetry operations x, y, z; x, 1+y, 1+z;

-1-x, -y, 1-z

6.2.4   Results and Discussion

The structure of 1 as proposed by Stevens et al[116] is confirmed by this crystallographic study and the numbering scheme and general conformation are shown in Figure 6.4; positional parameters of the atoms are given in Table 6.1.

The asymmetric unit consists of two conformationally distinctive molecules; these conformers are rotamers about the C(8)-C(81) ring to carboxamide bond, the one amide being intramolecularly hydrogen bonded to N(1) of the tetrazine ring, the other not.   The molecular dimensions, bond distances ($\overset{o}{A}$) and angles (°) are given in Tables 6.3 and 4.

A comparison of the dimensions shows the two independent molecules to be structurally similar, with however, significant differences betwen the lengths of N(7)-C(8) (4σ) and also C(8A)-N(5) and N(2)-N(3) (both 3σ) bonds in the primed and unprimed molecules. The N(7)-C(8) bond shows significantly more double bond character than N(7)'-C(8)' this possibly being influenced by the intramolecular interaction N(82)-H(7)......N(1) 3.075 $\overset{o}{A}$ (Table 6.7).

The orientation of the carboxamide groups is the fundamental difference between the two molecules, the one being rotated approximately 180° to the other with respect to the imidazole ring in each molecule.   In both cases, however, the carboxamide groups are approximately coplanar with the imidazole rings (Table 6.5 and 6) with the twist angles between being approximately 5.0° and 8.0° respectively.   This planarity affords the possibility of some conjugation between the carboxamide groups and their respective imidazole rings and indeed the C(8)-C(81) and C(8)'-C(81)' distances of 1.470(3)$\overset{o}{A}$ and 1.487(3)$\overset{o}{A}$ suggest some shortening of these bonds relative to the $C(sp^2)-C(sp^2)$ bridging bond in biphenyl

$(1.496\overset{\text{o}}{\text{A}})$[54] . The accommodation of the intramolecular hydrogen bond does, however, appear to impose some lateral displacement on the C(8)-C(81) bond as demonstrated by the difference in bond angles N(7)-C(8)-C(81) and C(8A)'-C(8)'-C(81)', Table 6.4.

The C(81)-N(82) bond distance in both the unprimed and primed molecules of $1.326(4)\overset{\text{o}}{\text{A}}$ and $1.330(3)\overset{\text{o}}{\text{A}}$, respectively suggests considerable double bond character consistent with sp^2 hybridisation of the amide nitrogen in each case. This is further demonstrated by the sum of the bond angles around N(82) and N(82)', these being 359(6)° and 360(6)°.

The N(1)-N(2) bond distance [$1.266(3)\overset{\text{o}}{\text{A}}$ unprimed, $1.257(3)\overset{\text{o}}{\text{A}}$ primed] and the C(6)-N(7) [$1.306(3)\overset{\text{o}}{\text{A}}$ unprimed, $1.300(3)\overset{\text{o}}{\text{A}}$ primed] show true double bond character , whilst the C(8)-C(8A) bonds [$1.369(3)\overset{\text{o}}{\text{A}}$ unprimed, $1.374(3)\overset{\text{o}}{\text{A}}$ primed] are compatible with a partial . double bond suggesting some delocalisation. Deviations in $\overset{\text{o}}{\text{A}}$ from the least squares plane through the ring atoms of the bicyclic systems in the two molecules (Table 6.5) show the imidazoletetrazine moeities to be substantially planar, albeit that N(1)', N(3)', N(5)' and N(7)' show significantly greater deviations than in the unprimed molecule. Consideration of the deviations of N(5) and C(8A), which form the bridging bond, and also the torsion angles about this bond (Table 6.6) suggest that in both the unprimed and the primed molecules the imidazole and tetrazine rings make an angle to each other of no more than 1.5° from coplanarity. The chloroethyl groups are both of similar geometry, but as can been seen from the least squares planes calculations and torsion angles, are approximate opposites with respect to their tetrazine rings.

The molecular packing and hydrogen bonding are shown in Figure 6.5 and the contact distances and symmetry operations given in Table

6.7. Groups of four molecules hydrogen bond around a centre of symmetry facilitated by a 'base pairing' interaction of the type N(82)-H(7) .....N(7)' (3.063Å) and N(82)'-H(6)' ....O(82) (2.95Å) . The intramolecular hydrogen bond N(82)-H(7) ....N(1) (3.075Å) maintains the planarity of the carboxamide group with the ring system of the unprimed molecule. This feature is also found in the crystal structure of DTIC[120] (Figure 6.7) and its hydrochloride (hydrate) salt[121] and ensures a reasonably planar 'base pair'. The interaction between the two sets of bases is then completed by an N(82)'-H(7)' ....N(7) contact (3.089Å) to produce a group of four.

The original specimen of 1 synthesised by the interaction of 5-diazoimidazole-4-carboxamide and 2-chloroethyisocyanate in dichloromethane at 25° had IR absorptions (KBr) at 2500 cm^{-1} and 3240 cm^{-1} (broad) and carbonyl bands at 1740 and 1680 cm^{-1}. Samples synthesised employing ethyl acetate as solvent or obtained by crystallising the dichloromethane products from aqueous acetone on the other hand, differed in that the NH absorptions appeared at 3450, 3350 and 3220 (broad) and 3120 cm^{-1} with strong broad carbonyl frequencies at 1748 and 1673 cm^{-1}. The solution IR and ^{1}HNMR spectra of different samples were identical. The discrepancies in the solid phase IR spectra of samples obtained from either dichloromethane or ethyl acetate can be explained possibly if there is some variation in hydrogen bonding or presence of rotamers between the two.
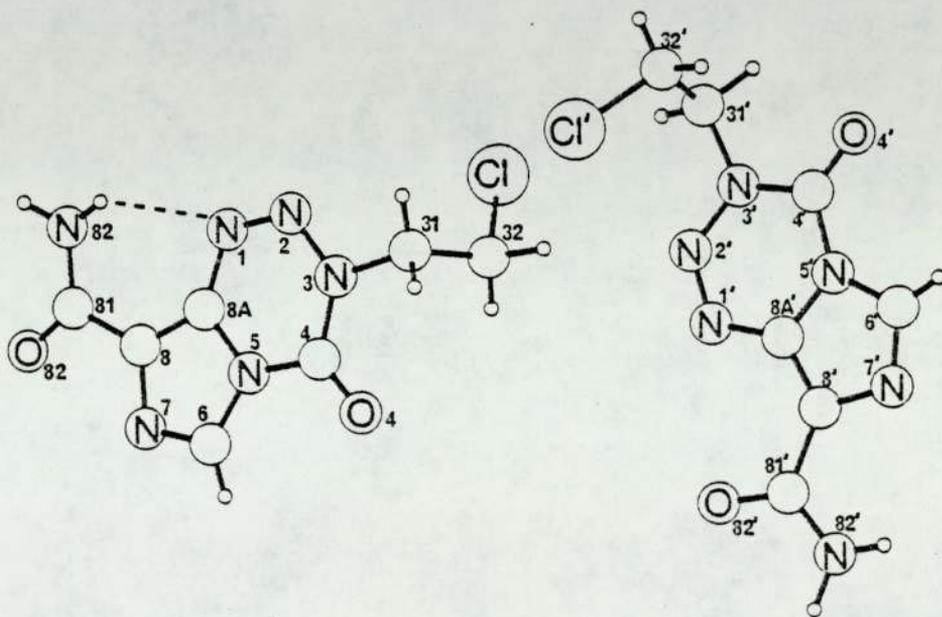
Fig.6.4 Molecular structure and numbering scheme for 8-carbamoyl-3-(2-chloroethyl) imidazo-[5,1-d]-1,2,3,5-tetrazine-4(3H)-one.
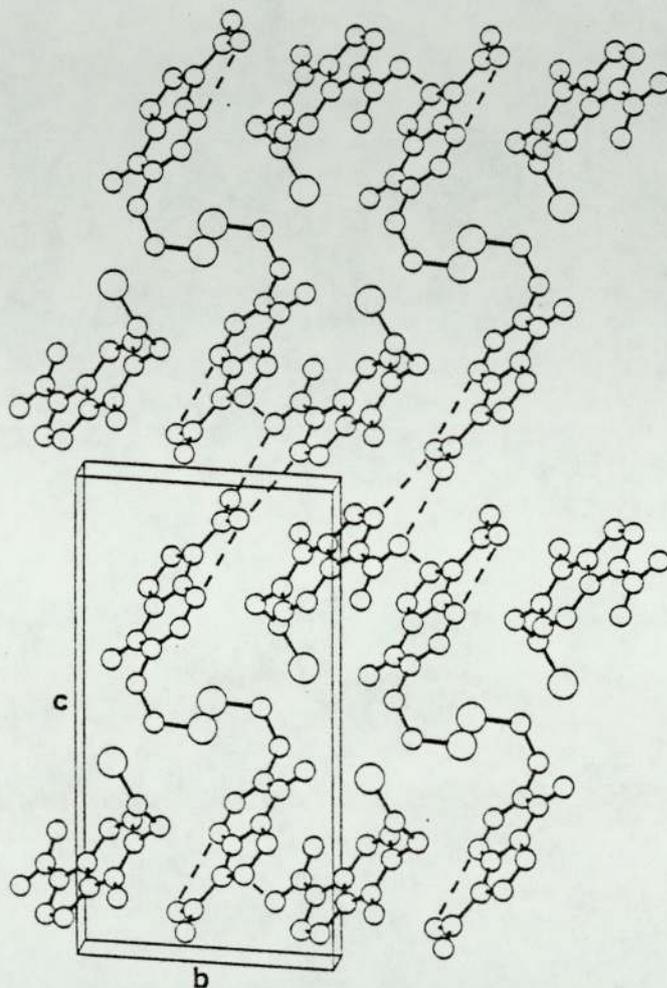
Fig.6.5 Molecular packing and hydrogen bonding for 8-carbamoyl-3-(2-chloroethyl) imidazo- [5,1-d] -1,2,3,5-tetrazine-4-(3H)-one.

Figure 6.5a   Stereo diagram
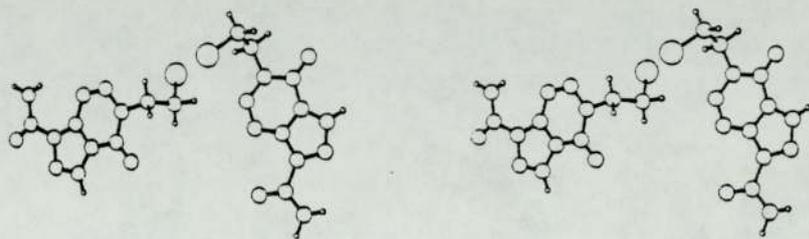
Figure 6.5b   Stereo packing diagram
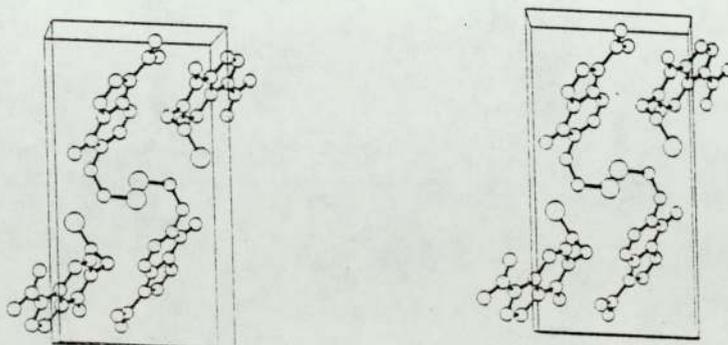
Figure 6.6   Space filling diagram
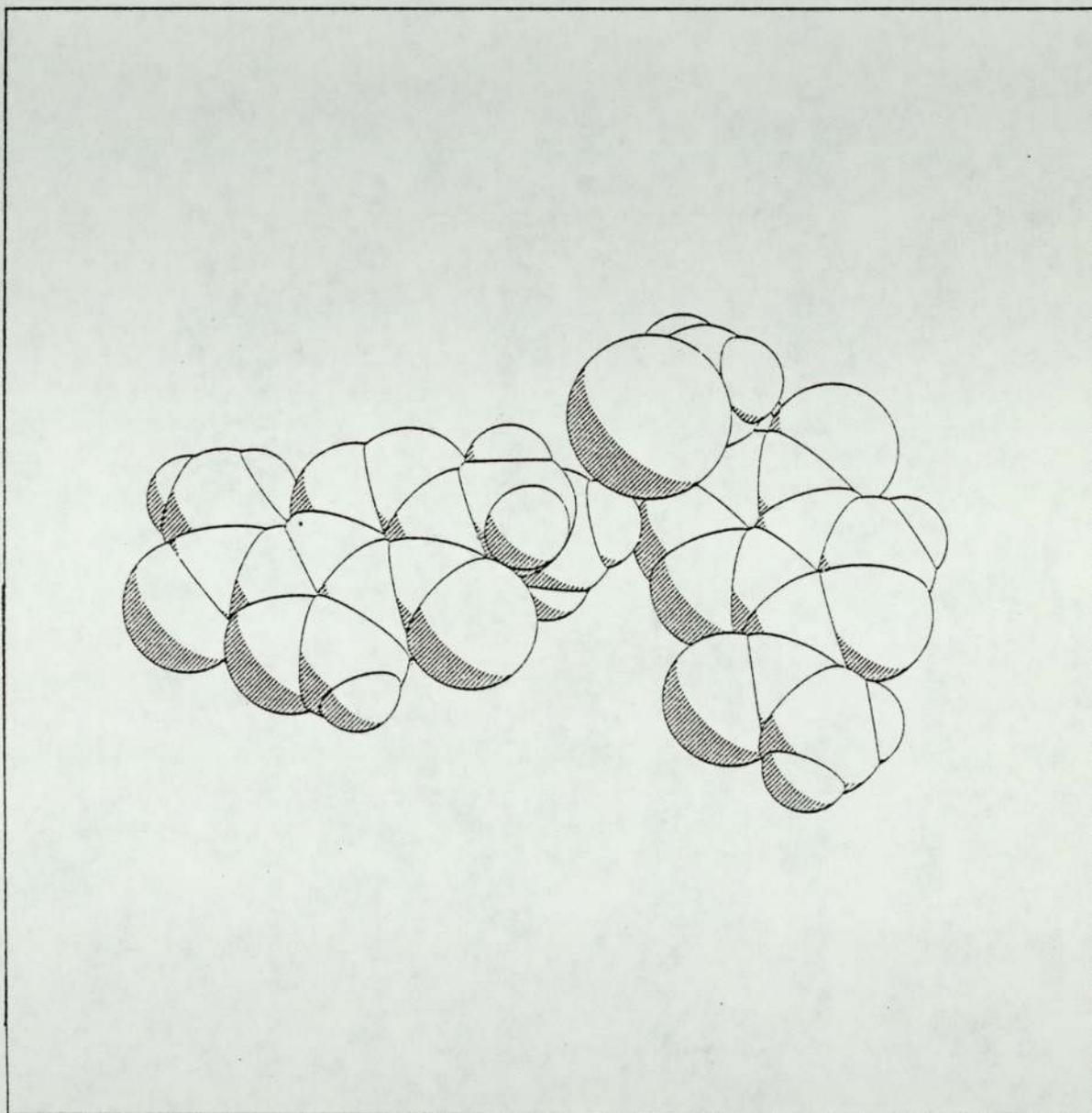
Figure 6.7 The structure of DTIC
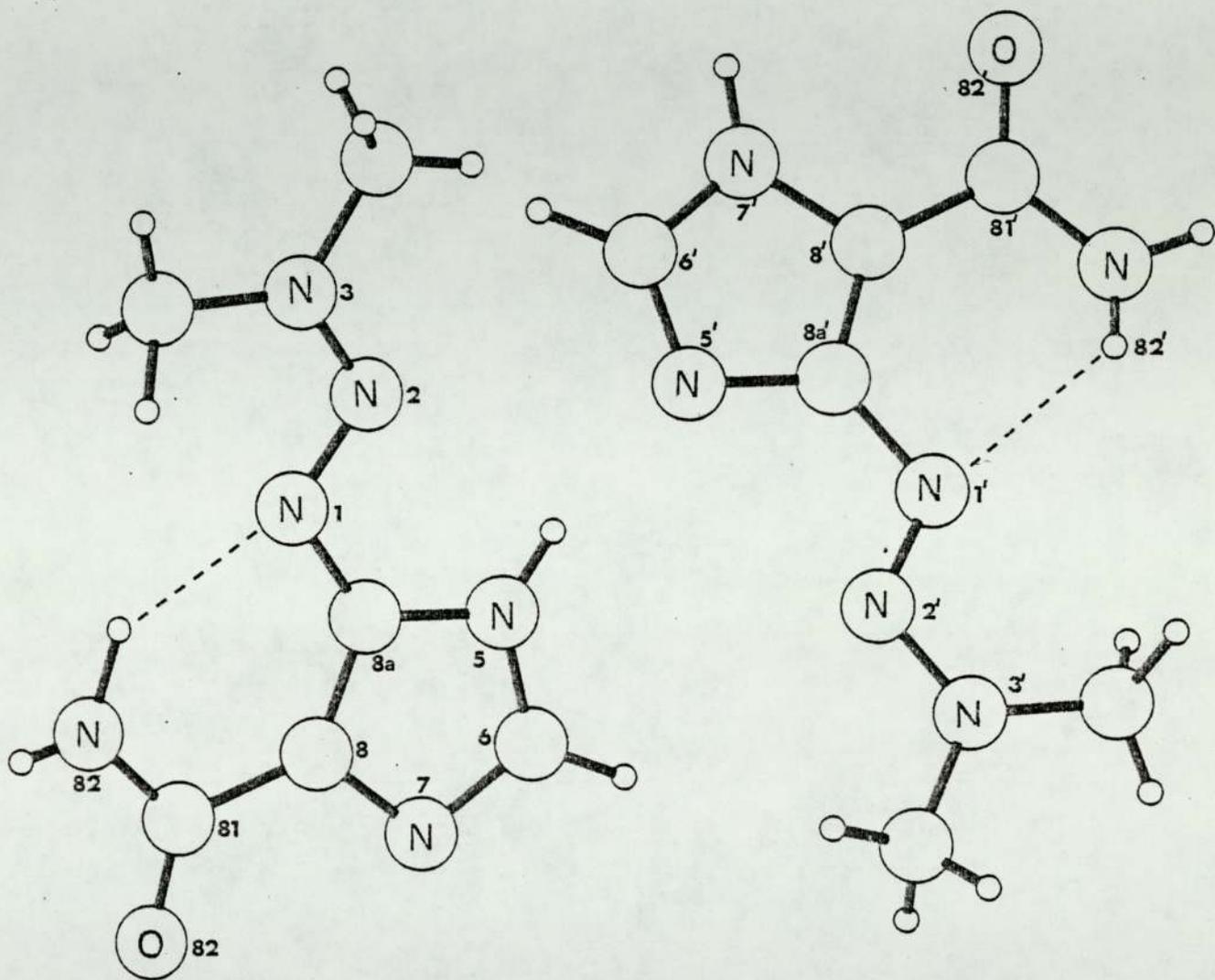
6.3.1   Preparation of 8-carbamoyl-3-methylimidazo [5,1-d] 1,2,3,5

tetrazin 4(3H)-one[22] (Compound 2)

5-diazoimidazole-4-carboxamide (1.37 g) was stirred in ethyl acetate with methyl isocyanate (7.0 g) in the dark at room temperature in a tightly closed container for three weeks. The cream coloured solid produced was filtered, washed with ether and dried under suction.   The calculated yield was 99% and the solid had a melting point of 212 °C.

Recrystallisation from three different solvent systems produced the following products each showing differences in the carbonyl stretching region of the infrared spectra of the solid compounds.

2a)   Acetone-water 3:1; colourless laths/plates

2b)   Acetone-water 1:3; a white microcrystalline solid

2c)   Hot water; a granular yellow solid *

Conventional analytical procedures confirmed that products (a) and (b) had the same molecular structure, and it was at this stage postulated that the differences in the infrared spectra of the solid compounds might be due to variations in the crystal structures.
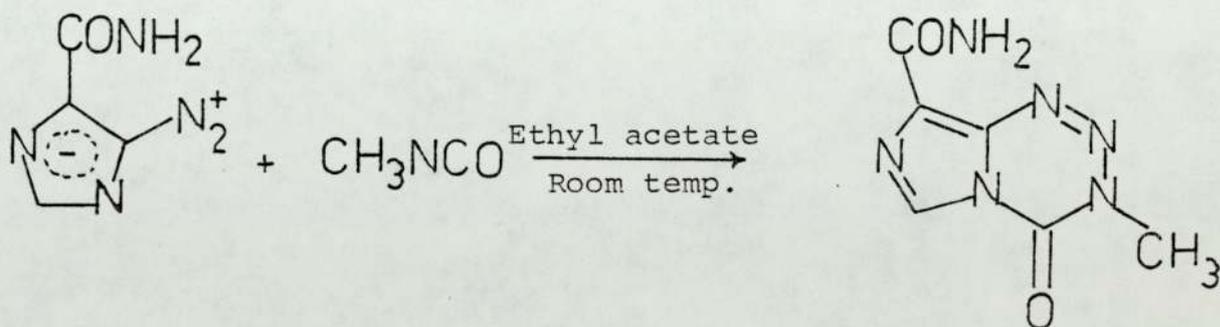
* Preparative work by courtesy of Mrs G U Baig, Department of Pharmacy, University of Aston.

6.3.2   Crystal Data

8-Carbamoyl-3-methylimidazo [5,1-d] 1,2,3,5 tetrazin-4(3H)-one ($C_6H_6N_6O_2$) crystallises in the space group $P2_1/c$ with a = 17.332 (3) Å , b = 7.351(2)Å , c = 13.247(1)Å , β = 109.56(1)°, z = 8, v = 1590(1) Å^3 . The molecular weight is 194.16 [F(ooo) = 760) and the calculated density Dc = 1.612 gcm^{-3}. The absorption coefficient for MoKα (λ= 0.71069 Å) radiation is μ = 0.78 cm^{-1}.

6.3.3   Structure Analysis

Data Collection

A well formed crystal of product 2(a) of dimensions 1.0 x 0.48 x 0.16 mm grown by slow evaporation from 1:3 aqueous acetone was mounted on an Enraf-Nonius CAD4 diffractometer. After obtaining an orientation matrix and determining a suitable unit cell a total of 2492 unique reflections were collected using the ω -2θ scan technique between θ = 2° and 24° with monochromated Mo-Kα radiation (λ = 0.71069 Å). Two reference reflections were measured every 3600 secs of x-ray exposure and the orientation of the crystal was checked after every 100 reflections. No decomposition or movement of the crystal was detected during the data collection. A final sequence of refinements on the unit cell dimensions were performed using selected reflections from the data set; these are as shown under crystal data.

Structure determination and refinement

The data reduction was performed at Birmingham University using a programme provided by Dr T A Hamor. Lorentz and polarisation and intensity corrections were applied, and the data output in the form HKL, Fobs and σ Fobs where σ Fobs is based on counting statistics and an allowance for instrument instability. Normalised structure factors were calculated using SHELX[2] and the overall scale factor and

estimated temperature factor determined as 2.5836 and 0.067 respectively.

Phase determination was undertaken using SHELX[2] automatic centrosymmetric direct methods. E min was set to 1.1 and NP, the number of phase permutations, to 14 (ie. 2^{14} permutations).

The first attempt at solving the structure did not produce a refinable solution, albeit that certain fragments of the two molecules in the asymmetric unit did appear on several of the E-maps. However, on removing the $\bar{5}41$ reflection from the starting set, this having a dominantly high E value of 7.305, an E-map was obtained in which all of the non-hydrogen atoms in the asymmetric unit appeared. This solution was the one with the highest figures of merit and the programme assigned the following origin determining reflections; $\bar{1}$ 1 10 (+), $\bar{1}6$ 1 6 (+) and $\bar{3}$ 2 1 (+). This determination was then followed by six cycles of full matrix least squares refinement of positional parameters and isotropic temperature factors using only unit weights, after which $R = \sum_i [\, ||F_o| - S\,|F_c|| \,]/ \sum |F_o|$ decreased to 0.097 and 's', the overall scale factor, refined to a value of 1.3805. From the difference Fourier synthesis produced by SHELX, all of the hydrogen atoms in the structure were located. Further refinement of co-ordinates and anisotropic thermal parameters for the non-hydrogen atoms proceeded. At this stage the hydrogen atoms were assigned the isotropic temperature factor of their attached atom, but no hydrogen atom positional parameters or temperature factors were refined. The unweighted R factor after seven cycles of refinement reduced to 0.051. A final sequence of refinements by SHELX for the complete structure of co-ordinates and anisotropic thermal parameters for non-hydrogen atoms and co-ordinates and isotropic temperature factors for hydrogens resulted in

an unweighted discrepancy index R = 0.043 for the 1749 observed data with Fo > 3σ.

In the final sequence of refinements, reflections were weighted according to W = k/ $[\sigma^2(Fo) + qFo^2]$ where q is a refinable parameter which converged at 0.0014, and k, which should approach unity if there are no serious systematic errors, equalled 1.5642. After the last cycle of refinement, no positional parameter had shifted by more than 0.25σ and the weighted discrepancy index Rg = 0.062 with the overall scale factor s = 1.4211. The final difference electron density map showed no feature greater than 0.20 e $\overset{o-3}{A}$. During the refinement procedure, atoms were assigned their correct atomic scattering factors on the basis of the structure proposed from IR, NMR, MS and CHNO analysis. The isotropic temperature factors during refinement and the final discrepancy index and difference electron density map appeared to confirm the correctness of these assignments.

TABLE 6.8   Positional parameters (fractional co-ordinates x 10^4) with estimated standard deviations in parentheses.

Atom	x/a	y/b	z/c
N(1)	9419(2)	2377(4)	3123(2)
N(2)	10123(2)	3105(5)	3362(2)
N(3)	10594(2)	3368(4)	4423(2)
C(31)	11411(3)	4091(8)	4580(4)
C(4)	10384(2)	2896(5)	5298(3)
O(4)	10801(2)	3254(4)	6210(2)
N(5)	9614(1)	2098(4)	4990(2)
C(6)	9184(2)	1394(5)	5597(3)
N(7)	8483(2)	737(4)	4992(2)
C(8)	8439(2)	1014(4)	3955(2)
C(81)	7725(2)	0402(5)	3034(3)
N(82)	7122(2)	0354(5)	3270(3)
O(82)	7735(2)	0628(4)	2121(2)
C(8A)	9140(2)	1843(5)	3928(2)
N(1)'	4152(1)	8571(4)	4906(2)
N(2)'	3462(2)	9344(4)	4534(2)
N(3)'	3106(1)	9601(4)	3445(2)
C(31)'	2313(2)	10502(6)	3114(3)
C(4)'	3424(2)	9094(4)	2669(2)
O(4)'	3082(1)	9288(3)	1720(2)
N(5)'	4186(1)	8281(4)	3114(2)
C(6)'	4715(2)	7599(5)	1645(3)
N(7)'	5368(2)	6916(4)	3364(2)
C(8)'	5270(2)	7145(4)	4346(2)
C(81)'	5901(2)	6522(4)	5344(2)
N(82)'	6524(2)	5661(4)	5198(2)

Atom	x/a	y/b	z/c
O(82)'	5838(1)	6819(3)	6230(2)
C(8A)'	4538(2)	8014(4)	4210(2)
H(1)	11797(26)	3281(66)	4826(35)
H(2)	11428(32)	4455(74)	3823(43)
H(3)	11538(29)	4984(71)	5106(40)
H(4)	9425(21)	1385(53)	6383(32)
H(5)	7122(25)	- 511(57)	3928(36)
H(6)	6596(25)	- 787(59)	2687(33)
H(1)'	1874(38)	9590(90)	2776(49)
H(2)'	2326(29)	11462(81)	2738(43)
H(3)'	2166(29)	11076(67)	3716(44)
H(4)'	4597(18)	7688(46)	1900(27)
H(5)'	6911(25)	5303(59)	5775(35)
H(6)'	6545(24)	5499(56)	4493(35)

Atom	x/a	y/b	z/c
O(82)'	5838(1)	6819(3)	6230(2)
C(8A)'	4538(2)	8014(4)	4210(2)
H(1)'	1874(38)	9590(90)	2776(49)

TABLE 6.9   Anisotropic temperature factors

ATOM	U11	U22	U33	U23	U13	U12
N(1)	.0497(17)	.0502(20)	.0326(15)	.0020(14)	.0132(13)	-.0004(15)
N(2)	.0510(18)	.0663(24)	.0428(18)	-.0002(16)	.0216(15)	-.0073(17)
N(3)	.0394(16)	.0490(20)	.0468(17)	-.0038(14)	.0159(14)	-.0044(14)
C(31)	.0408(23)	.0624(32)	.0781(32)	-.0084(27)	.0276(22)	-.0083(22)
C(4)	.0393(18)	.0406(22)	.0411(20)	-.0042(16)	.0101(16)	-.0012(16)
O(4)	.0515(15)	.0850(23)	.0405(15)	-.0125(14)	.0030(12)	-.0199(15)
N(5)	.0352(14)	.0399(18)	.0277(14)	-.0021(12)	.0071(11)	-.0052(13)
C(6)	.0523(21)	.0584(24)	.0258(17)	-.0008(16)	.0086(16)	-.0130(19)
N(7)	.0448(17)	.0550(20)	.0305(15)	-.0020(13)	.0118(13)	-.0116(14)
C(8)	.0348(17)	.0356(20)	.0291(16)	-.0032(14)	.0062(13)	-.0010(15)
C(81)	.0401(19)	.0340(20)	.0291(16)	-.0078(15)	.0082(15)	-.0011(15)
N(82)	.0410(19)	.0546(22)	.0448(19)	-.0057(16)	.0080(16)	-.0081(15)
O(82)	.0545(16)	.0684(19)	.0301(13)	-.0093(12)	.0057(11)	-.0137(13)
C(8A)	.0388(17)	.0352(19)	.0231(15)	-.0043(13)	.0078(13)	.0008(15)
N(1)'	.0333(14)	.0379(16)	.0277(13)	-.0005(12)	.0097(11)	-.0012(13)
N(2)'	.0380(15)	.0348(16)	.0290(14)	-.0006(12)	.0087(12)	-.0009(13)
N(3)'	.0031(14)	.0300(15)	.0304(14)	.0029(12)	.0094(11)	.0004(12)
C(31)'	.0382(20)	.0517(26)	.0413(21)	.0078(20)	.0120(17)	.0133(18)
C(4)'	.0291(16)	.0280(18)	.0299(17)	.0033(13)	.0057(14)	-.0047(13)
O(4)'	.0393(13)	.0561(17)	.0280(13)	.0086(11)	.0060(10)	.0015(11)
N(5)'	.0312(13)	.0355(16)	.0229(13)	.0039(11)	.0067(10)	.0003(12)
C(6)'	.0377(18)	.0597(24)	.0237(17)	.0036(16)	.0105(14)	.0063(18)
N(7)'	.0381(15)	.0466(18)	.0291(14)	.0038(13)	.0108(12)	.0050(13)
C(8)'	.0307(16)	.0322(18)	.0249(15)	-.0003(13)	.0070(13)	-.0028(14)
C(81)'	.0306(16)	.0335(19)	.0277(17)	.0004(14)	.0088(13)	-.0088(15)

ATOM	U11	U22	U33	U23	U13	.U12
N(82)'	.0369(16)	.0497(20)	.0273(15)	.0040(13)	.0079(13)	.0102(14)
O(82)'	.0436(13)	.0555(16)	.0250(12)	.0034(11)	.0100(10)	.0072(11)
C(8A)'	.0323(16)	.0269(18)	.0233(15)	.0003(13)	.0058(13)	-.0042(14)

TABLE 6.10   Isotropic temperature factors for hydrogen atoms

ATOM	Uiso
H(1)	.074(16)
H(2)	.108(19)
H(3)	.083(16)
H(4)	.064(11)
H(5)	.064(13)
H(6)	.071(13)
H(1)'	.143(24)
H(2)'	.097(19)
H(3)'	.095(17)
H(4)'	.041(8)
H(5)'	.061(13)
H(6)'	.066(12)

TABLE 6.11 Bond distances in Å of the two molecules in the asymmetric unit (unprimed and primed) with estimated standard deviations in parentheses.

Bond	Unprimed [Å]	Primed [Å]
N(1)-N(2)	1.272(4)	1.265(4)
N(2)-N(3)	1.382(4)	1.379(3)
N(3)-C(31)	1.461(6)	1.455(5)
N(3)-C(4)	1.371(5)	1.371(5)
C(4)-O(4)	1.196(4)	1.205(3)
C(4)-N(5)	1.389(4)	1.388(4)
N(5)-C(6)	1.368(5)	1.364(5)
C(6)-N(7)	1.303(4)	1.311(4)
N(7)-C(8)	1.368(4)	1.378(4)
C(8)-C(8A)	1.370(5)	1.378(4)
C(8A)-N(5)	1.384(3)	1.388(3)
C(8)-C(81)	1.486(4)	1.478(4)
C(81)-O(82)	1.227(4)	1.234(4)
C(81)-N(82)	1.326(5)	1.322(5)
C(8A)-N(1)	1.368(5)	1.370(4)
C(31)-H(1)	0.88(4)	1.03(6)
C(31)-H(2)	1.05(6)	0.87(6)
C(31)-H(3)	0.93(5)	0.96(6)
C(6)-H(4)	0.98(4)	0.94(4)
N(82)-H(5)	0.87(5)	0.87(4)
N(82)-H(6)	1.02(4)	0.96(5)

TABLE 6.12   Interatomic angles (°) for the two molecules in the asymmetric unit (unprimed and primed) with estimated standard deviations in parentheses.

Atoms	Bond angle [°] (unprimed)	Bond angle [°] (primed)
N(1)- N(2)-N(3)	120.2(3)	120.2(3)
N(2)-N(3)-C(4)	126.3(3)	126.5(2)
N(2)-N(3)-C(31)	114.2(3)	115.1(3)
C(31)-N(3)-C(4)	119.3(3)	118.3(3)
N(3)-C(4)-O(4)	124.9(3)	124.6(3)
N(3)-C(4)-N(5)	111.0(3)	111.3(2)
O(4)-C(4)-N(5)	124.1(4)	124.1(3)
C(4)-N(5)-C(6)	130.2(3)	130.8(3)
C(4)-N(5)-C(8A)	122.8(3)	121.8(3)
C(6)-N(5)-C(8A)	106.9(3)	107.3(2)
N(5)-C(6)-N(7)	111.1(3)	111.0(3)
C(6)-N(7)-C(8)	106.8(3)	107.0(3)
N(7)-C(8)-C(8A)	109.9(2)	109.5(2)
N(7)-C(8)-C(81)	122.1(3)	121.3(3)
C(81)-C(8)-C(8A)	128.0(3)	129.3(3)
C(8)-C(81)-O(82)	119.0(3)	121.5(3)
C(8)-C(81)-N(82)	116.5(3)	114.4(3)
O(82)-C(81)-N(82)	124.5(3)	124.1(3)
C(8)-C(8A)-N(5)	105.3(3)	105.4(3)
C(8)-C(8A)-N(1)	134.1(3)	133.3(2)
N(1)-C(8A)-N(5)	120.6(3)	121.3(2)
C(8A)-N(1)-N(2)	119.1(3)	118.8(2)
N(3)-C(31)-H(1)	113(3)	109(4)
N(3)-C(31)-H(2)	107(3)	110(4)

N(3)-C(31)-H(3)	110(3)	112(3)
H(1)-C(31)-H(2)	106(4)	111(5)
H(1)-C(31)-H(3)	104(4)	108(5)
H(2)-C(31)-H(3)	117(4)	108(5)
N(5)-C(6)-H(4)	120(2)	121(2)
H(4)-C(6)-N(7)	127(2)	128(2)
C(81)-N(82)-H(5)	123(3)	116(3)
C(81)-N(82)-H(6)	121(3)	120(2)
H(5)-N(82)-H(6)	116(4)	123(4)

TABLE 6.13 Deviations [Å] of non-hydrogen atoms from the least
squares plane through the imidazotetrazine ring system[+]
(The two molecules in the asymmetric unit are treated
separately. Atoms used in the plane calculation are
marked with an asterisk).

ATOM	DEVIATION [Å] unprimed	DEVIATION [Å] primed
*N(1)	.001	.079
*N(2)	-.015	-.012
*N(3)	-.005	-.013
*C(4)	.007	-.019
*N(5)	.015	-.003
*C(6)	-.007	.009
*N(7)	-.018	-.002
*C(8)	.009	-.027
*C(8A)	.013	-.013
C(31)	-.125	-.005
O(4)	.027	-.063
C(81)	.018	-.043
O(82)	-.003	-.108
N(82)	-.047	.014

[+] The equations of the planes are:-

unprimed $-.420X + .897Y + .140Z - 5.077 = 0$

primed $.449X + .890Y - .075Z + 8.366 = 0$

where X, Y and Z are orthogonal co-ordinates in Å along a*, b and c

TABLE 6.14   Important Torsion Angles [°]

ATOMS	Angle [°] unprimed	Angle [°] primed
N(7)-C(8)-C(81)-N(82)	3.2	-3.6
N(7)-C(8)-C(81)-O(82)	-177.1	175.7
C(8A)-C(8)-C(81)-N(82)	-170.3	177.4
C(8A)-C(8)-C(81)-O(82)	.3	-3.3
C(31)-N(3)-C(4)-O(4)	5.9	-2.6
N(1)-N(2)-N(3)-C(31)	175.7	-176.5
C(4)-N(5)-C(8A)-N(1)	.1	-9.3
C(4)-N(5)-C(8A)-C(8)	179.4	178.5
C(6)-N(5)-C(8A)-C(8)	.8	-.8
C(6)-N(5)-C(8A)-N(1)	-178.5	171.3

TABLE 6.15   Hydrogen bond contact distances and angles

Hydrogen Bond	Angle at H[°]	N....O distance ($\mathring{A}$)
$N(82)_I$-H(6).....$O(82)^!_{II}$	159	3.057
$N(82)^!_I$ - H(5)'.....$O(82)_{III}$	179	2.865
$N(82)^!_I$-H(6)'....$O(4)^!_{IV}$	155	3.018

I, II, III and IV refer to the symmetry operations*-

x, y, z; x,1/2-y, -1/2+z; x, 1/2-y, 1/2+z; 1-x, -1/2+y, 1/2-z

6.3.4 Results and Discussion

The structure of (2) as proposed by Stevens et al[116] is confirmed by this crystallographic study and the numbering scheme and general conformation are shown in Figure 6.8; the positional parameters of the atoms are given in Table 6.8.

Again, as with the structure of (1), the asymmetric unit consists of two distinctive molecules, however, unlike mitozolomide (1) they are not rotamers about the C(8)-C(81) bond, albeit that it is a hydrogen bonding constraint that confers the unique nature of the two molecules.

Comparison of the dimensions of the two independent molecules, Tables 6.11 and 12, shows that they are structurally similar with the only appreciable differences in bond lengths occurring between the N(7)-C(8) and C(4)-O(4) (2σ) and the C(8)-C(81) and C(6)-N(7) ($\frac{1}{2}\sigma$) bonds in the unprimed and primed molecules. The length of the C(4)-O(4) bond is probably influenced by hydrogen bonding, in that the longer of the two, in the primed molecule, is hydrogen bonded, whilst the other is not. The other bonds mentioned are in the immediate vicinity of the carboxamide group in each molecule, and may well be affected by differences in hydrogen bonding or the degree of coplanarity of this moiety with the ring system.

Unlike 1 (mitozolomide) the orientation of the carboxamide group with respect to the ring system is the same in the two independent molecules (unprimed and primed) and is such, so as not to facilitate intramolecular hydrogen bonding. The carboxamide groups are, however, still approximately coplanar with their respective imidazole rings, the twist angles between the two moeities about the C(8)-C(81) and C(8)'-C(81)' being 1.7° and 4.2°. The larger twist angle associated with the primed molecule may well be related to the

additional hydrogen bonding constraint placed upon the carboxamide nitrogen.

As with 1 this planarity again affords the possibility of some conjugation between the carboxamide groups and their respective imidazole rings, and indeed the C(8)-C(81) and C(8)'-C(81)' are of the same order or slightly shorter than those found in 1 (1.486(4) $\overset{o}{A}$ unprimed, 1.478(4) $\overset{o}{A}$ primed).

The C(81)-N(82) bond distance in the primed and unprimed molecules of 1.326(5)$\overset{o}{A}$ and 1.322(5)$\overset{o}{A}$ again suggests considerable double bond character with the sum of the bond angles around N(82) and N(82)' of 360(6)° and 359(6)° being consistent with sp^2 hybridisation of the nitrogen.

The true double bond nature of the N(1)-N(2) bonds (1.272(4)$\overset{o}{A}$ unprimed, 1.265(4) $\overset{o}{A}$ primed) and C(6)-N(7) (1.303(4)$\overset{o}{A}$ unprimed, 1.311(4) $\overset{o}{A}$ primed) is confirmed whilst the C(8)-C(8A) bonds (1.370(5) $\overset{o}{A}$ unprimed, 1.378(4) $\overset{o}{A}$ primed) show partial double bond character similar to those found in aromatic ring systems[122] suggesting some delocalisation.

The bond angles of the two molecules in the asymmetric unit (Table 6.12) do not show any major differences. Deviations in $\overset{o}{A}$ from the least squares plane through the ring atoms of the two imidazotetrazine rings (Table 6.13) show the whole molecule to be essentially planar. Small, but significant deviations from planarity occur with N(1)' in the tetrazine ring (0.079$\overset{o}{A}$) and C(31) (-.125$\overset{o}{A}$) in the unprimed molecule, both of these being confirmed by torsion angles (Table 6.14); the approximate planarity of the carboxamide groups with the ring systems is also confirmed. Consideration of the torsion angles about the N(5)-C(8A) bridging bond shows the imidazole and tetrazine rings to make an angle of approximately 1.5° to each

other in both of the molecules. The molecular packing and hydrogen
bonding are shown in Figures 6.9 and 10 and the contact distances and
symmetry operations are given in Table 6.15. The hydrogen bonding is
the feature which confers the difference between the two molecules in
the asymmetric unit. The unprimed and primed molecules form a
carboxamide pair with each other via N(82)-H(6) ... O(82)' (3.057Å)
and N(82)'-H(5)' ... O(82) (2.865Å) contacts, a similar pairing being
found in DTIC[120]. The primed molecule, however, forms an additional
hydrogen bond between O(4)' and an adjacent N(82)' (3.018Å). This
results in two adjacent primed molecules being linked by two N(82)'-
H(6)' ... O(4)' contacts 'head to tail' each one then forming a
carboxamide pair with an unprimed molecule to produce a group of four
molecules firmly hydrogen bonded. The infrared spectrum of this
compound shows three peaks in the carbonyl stretching region at 1755
cm^{-1}, 1725 cm^{-1} and 1675 cm^{-1}. The two higher wavenumber peaks are
likely to refer to the two ring carbonyl groups these being
substantially shorter and stronger than the carboxamide C=O bonds.
The difference in stretching frequency between the two ring carbonyls
is probably due to one being hydrogen bonded, and the other not. The
third and much lower stretching frequency (also more intense) is
probably due to the unprimed and primed carboxamide C=O absorptions
and since they are both hydrogen bonded in a similar environment, give
rise to a single peak.

The two carbonyl compound mentioned earlier (2b) gives rise to
two broad peaks centred at 1745 cm^{-1} and 1670 cm^{-1} both of which
appear to show signs of splitting. It is likely that this compound,
albeit molecularly the same, varies in the way it hydrogen bonds as a
solid. This infrared spectrum is similar to that of 1 (mitozolomide)
crystallised from aqueous acetone which shows two similar broadened
peaks with pronounced shoulders.

Figure 6.8   Molecular plot for the 3-methyl substituted

compound (2)
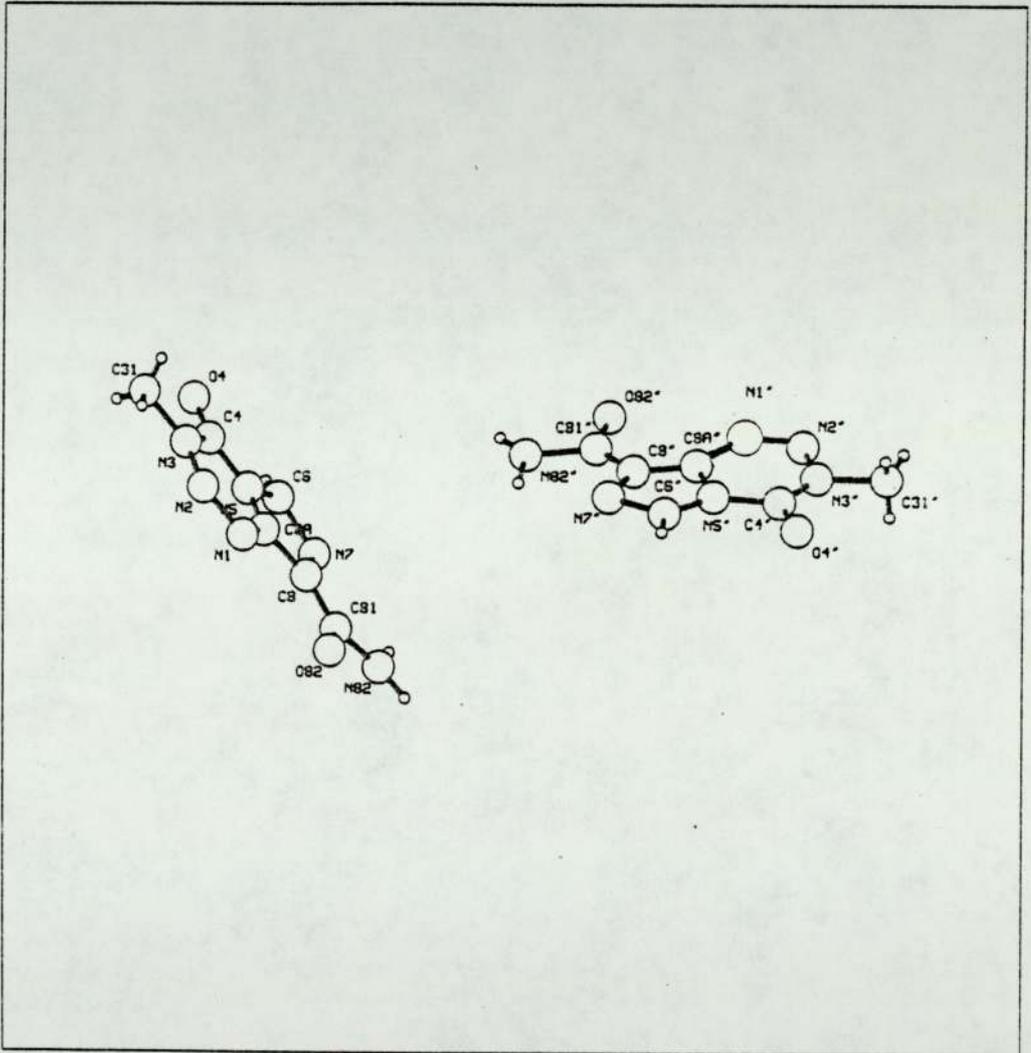
Figure 6.8a   Stereo plot

Figure 6.8b   Stereo packing diagram
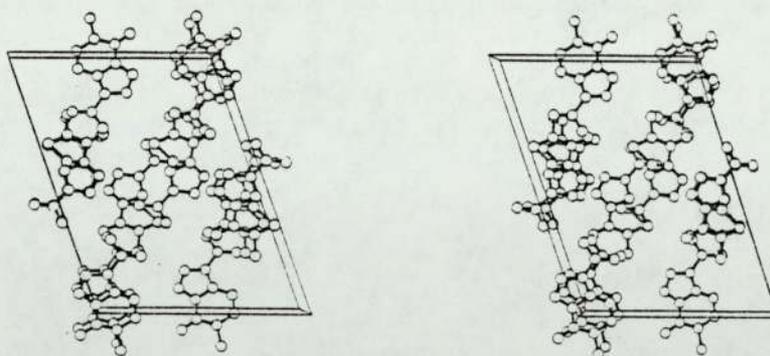
Figure 6.9   Packing diagram for compound 2 - a-c plane

(hydrogen bonds in dotted lines)
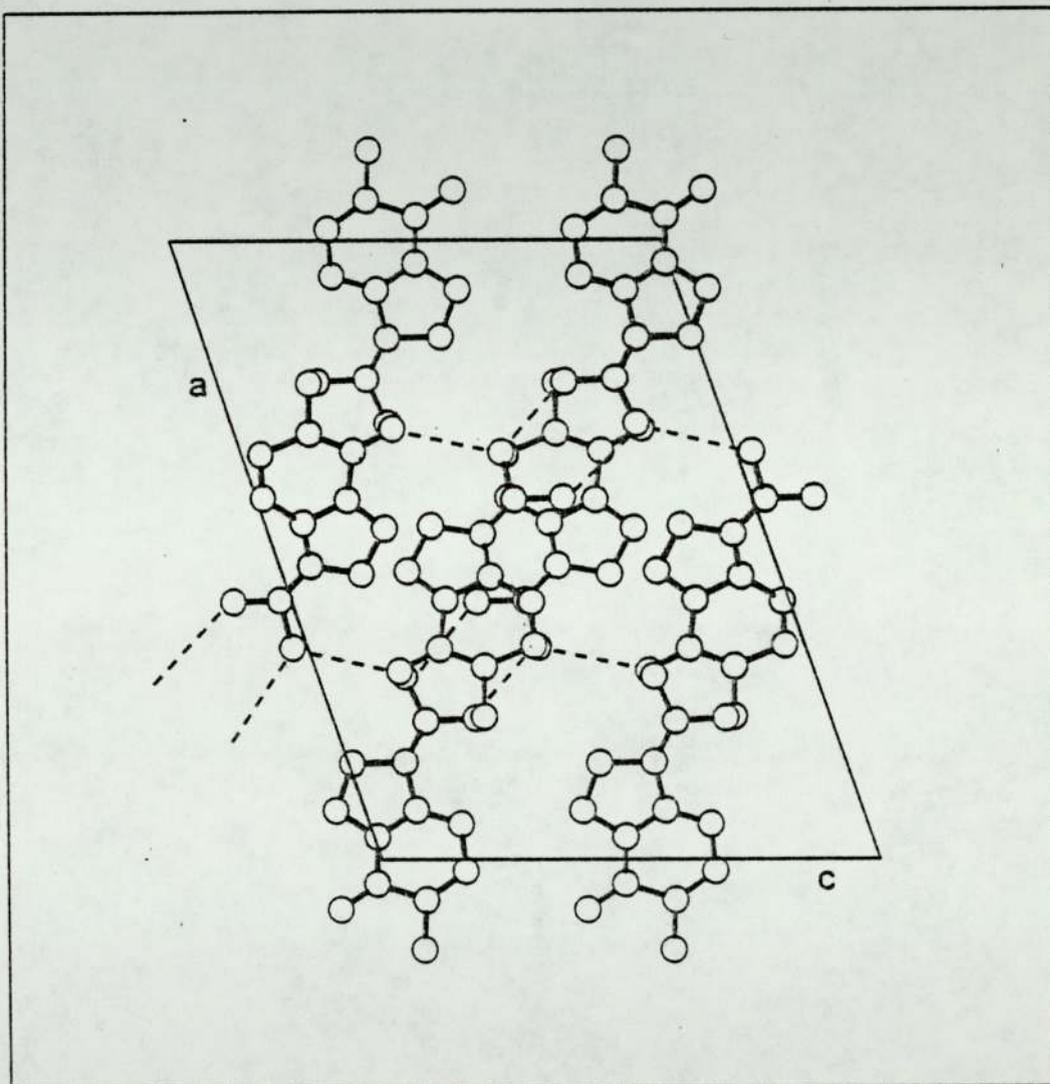
Figure 6.10   Packing diagram for compound 2 - b-a plane
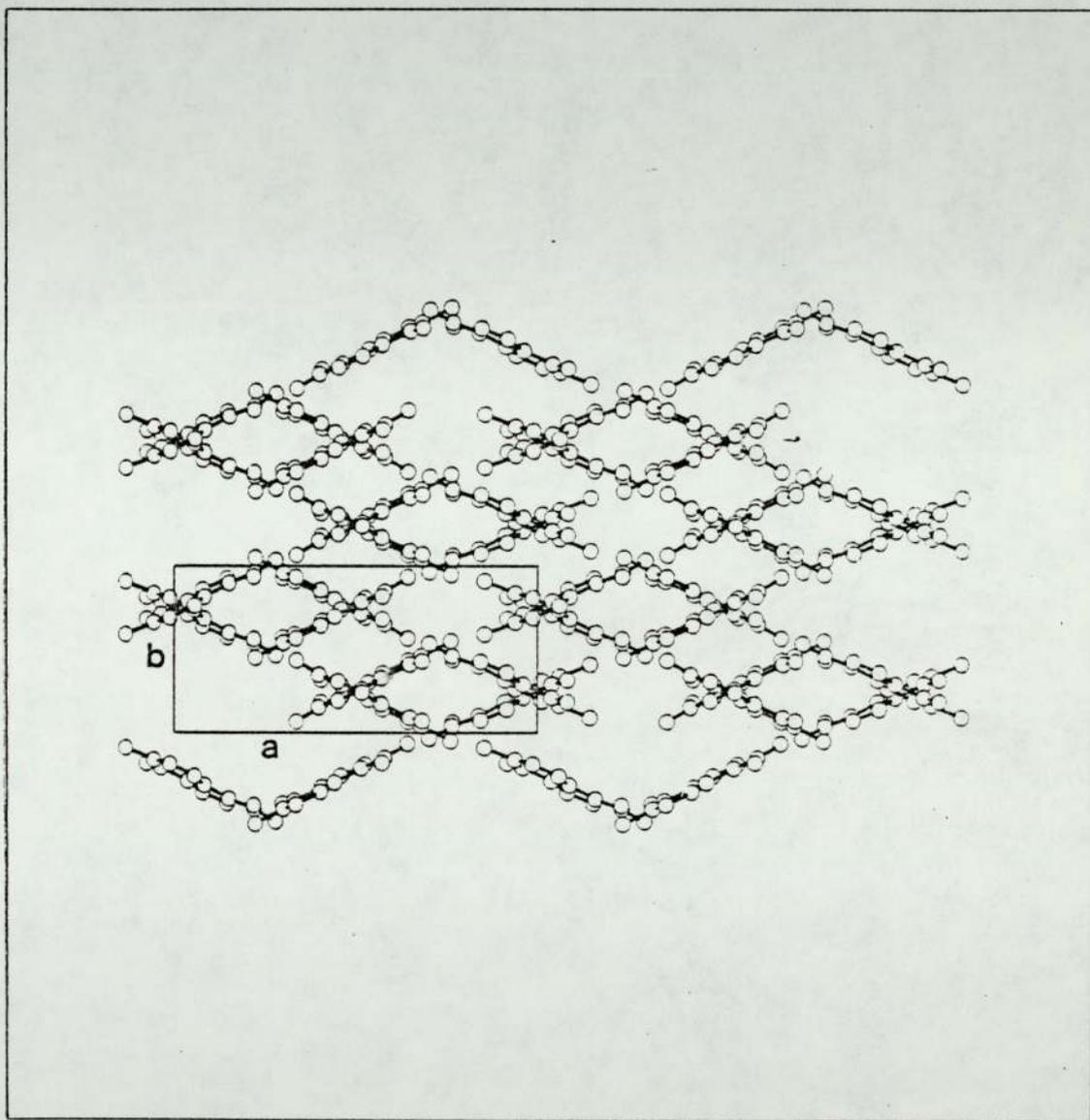
Figure 6.10a   Space filling diagram - compound 2
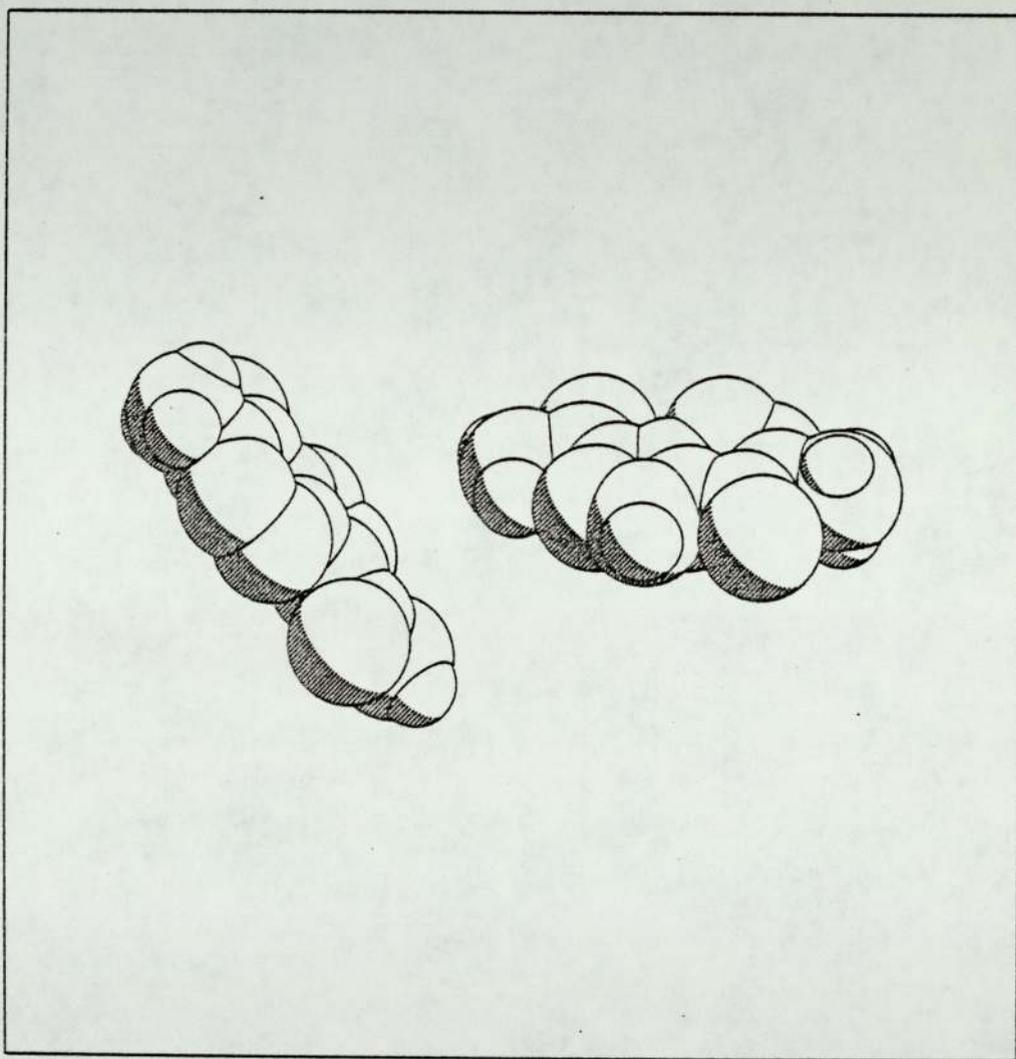
6.4.1   Preparation of 8-carbamoyl-3-(2-methoxyethyl) imidazo [5,1-d]
1,2,3,5-tetrazin-4(3H)-one[116]

5-Diazoimidazole-4-carboxamide (0.4 g) was suspended in
acetonitrile (10 ml) and 2-methoxyethyl isocyanate (0.4 g) added.
The mixture was heated to 45 °C and stirred in the dark for a period
of 8 hours. The reaction mixture was cooled and then filtered, and
the product thoroughly washed with ether. On evaporating to near
dryness, further product was obtained giving an eventual yield of
85%.

The compound was recrystallised using dimethylsulphoxide (DMSO)
which produced needle-like crystals with a melting point of 145   -
148 °C.*
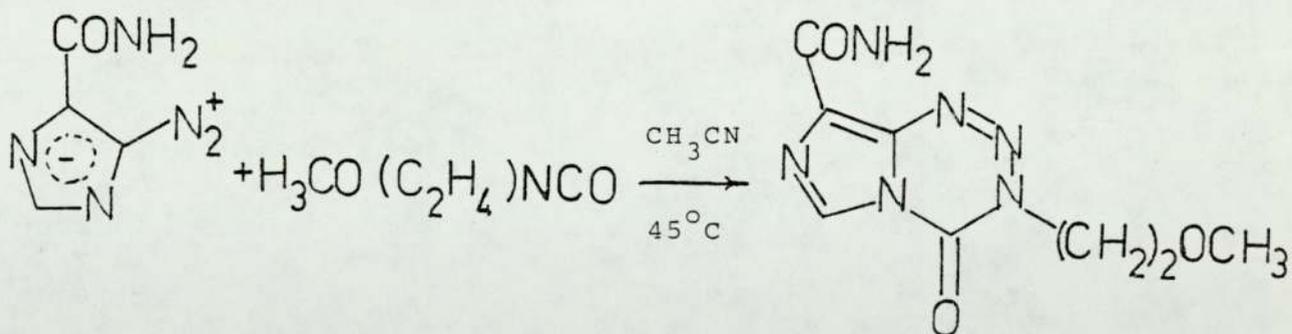
* Preparative work by courtesy of Mrs G U Baig, Department of
Pharmacy, University of Aston.

6.4.2 Crystal Data

8-carbamoyl-3-(2-methoxyethyl) imidazo[5,1-d] 1,2,3,5-tetrazin-4(3H)-one ($C_8H_{10}N_6O_3$) crystallises in the space group $P2_1/a$, with a = 6.905(3) $\overset{o}{A}$, b = 14.380(2) $\overset{o}{A}$, c = 10.649(2) $\overset{o}{A}$, β = 104.94(3)°, z = 4, v = 1022(1) $\overset{o}{A}{}^3$. The molecular weight is 238.22 (F(ooo) = 496) and the calculated density Dc = 1.538 g cm^{-3}. The absorption coefficient for MoKα radiation (λ = 0.71069 $\overset{o}{A}$) is μ= 0.80 cm^{-1}.

6.4.3 Structure Analysis

Data Collection

A crystal of dimensions 0.6 x 0.2 x 0.1 mm was mounted along its needle axis on an Enraf-Nenius CAD4 diffractometer. After an orientation matrix was obtained and refined and the appropriate monoclinic unit cell was determined, a total of 1771 unique reflections were collected by the ω-2θ scan technique between θ = 2° and 25° using monochromated Mo-Kα radiation (λ = 0.71069 $\overset{o}{A}$). Two reference reflections were measured every 7200 secs of x-ray exposure and the orientation of the crystal was checked after every 150 reflections. During the data collection, no decomposition or movement of the crystal was detected. A final sequence of refinements on the unit cell dimensions were performed using two high order reflections in 'h', since the initial search and index procedure only found and used low order 'h' reflections in determining the cell. The final unit cell dimensions are as shown under Crystal Data.

Structure Determination and refinement

The data reduction, as with the previous two structures in this section, was performed at Birmingham University using Dr T A Hamor's programme. Lorentz and polarisation and intensity corrections were applied and the data output in form HKL, Fobs and σFobs where the latter is based on counting statistics. No correction to the data was made for absorption effects, this decision being based on the nature of the crystal, lack of any heavy atoms and low 'μ' value.

Initially, phase determination was undertaken using SHELX[2] automatic centrosymmetric direct methods. Emin was set to 1.2 and NP, the number of phase permutations to 14 (ie. 2^{14} permutations). Although this produced one Emap with a recognisable fragment, neither Fourier nor least squares refinement would enhance the structure. It was therefore decided to try the MULTAN[1] direct methods package.

Normalised structure factors were calculated using a modified Wilson Plot method[22]. The curve from the Wilson calculation and a least squares fit are shown in Figure 6.11 with the values calculated from the gradient and intercept being:-

Temperature factor (B) = 2.5962

Scale K = 4.6836.

Phase determination was undertaken by MULTAN in which 168 reflections all with E$\geqslant$1.76 provided 1213 unique $\sum_2$ relationships for use in phase refinement.

The convergence mapping produced the following three origin fixing reflections; 4 0 $\bar{3}$ (+), 2 13 $\bar{4}$ (+), 3 8 4 (+). Sixteen starting sets were then obtained by permuting the signs of the 2 0 $\bar{2}$, 2 1 $\bar{2}$, 4 2 7 and 2 11 $\bar{3}$ reflections. The set with these reflections as -, +, +, +, gave an Emap in which all of the non-hydrogen atoms in the structure appeared. Using SHELX, seven cycles of full matrix

least squares refinement of positional parameters and isotropic temperature factors ensued and with unit weights, R the unweighted discrepancy index reduced to 0.088. At this point, a difference Fourier synthesis revealed nine of the ten proposed hydrogen atoms in the structure. Further refinement of co-ordinates and anisotropic thermal parameters for non-hydrogen atoms, and co-ordinates and isotropic temperature factors for hydrogens reduced R to 0.043; the difference Fourier synthesis, however, did not reveal the missing $-NH_2$ hydrogen. It was decided at this stage to insert the 'missing' hydrogen in a calculated position and refine the positions of both of the NH_2 hydrogens under a bond length constraint (DFIX parameter in SHELX).

After a final sequence of refinements using the latter mentioned constraints, and reflections weighted according to $w = k/[\sigma^2(Fo) + qFo^2]$ where q the refinable parameter was fixed at 0.0, R the unweighted discrepancy index reduced to 0.038 for the 1145 observed data with $Fo > 3\sigma$. The overall scale factor s, and k used in the weighting scheme, converged at 1.3720 and 0.7422 respectively.

After the last cycle of refinement no positional parameter had shifted by more than 0.15σ , and the weighted discrepancy index Rg = 0.036. A final difference electron density synthesis showed no feature greater than 0.16 e $\overset{o}{A}^{-3}$.

Figure 6.11

$$\log_e\left[\frac{\langle I\theta \rangle}{\Sigma_\theta}\right]$$

$$\frac{\sin^2\theta}{\lambda^2}$$

W= Wilson plot
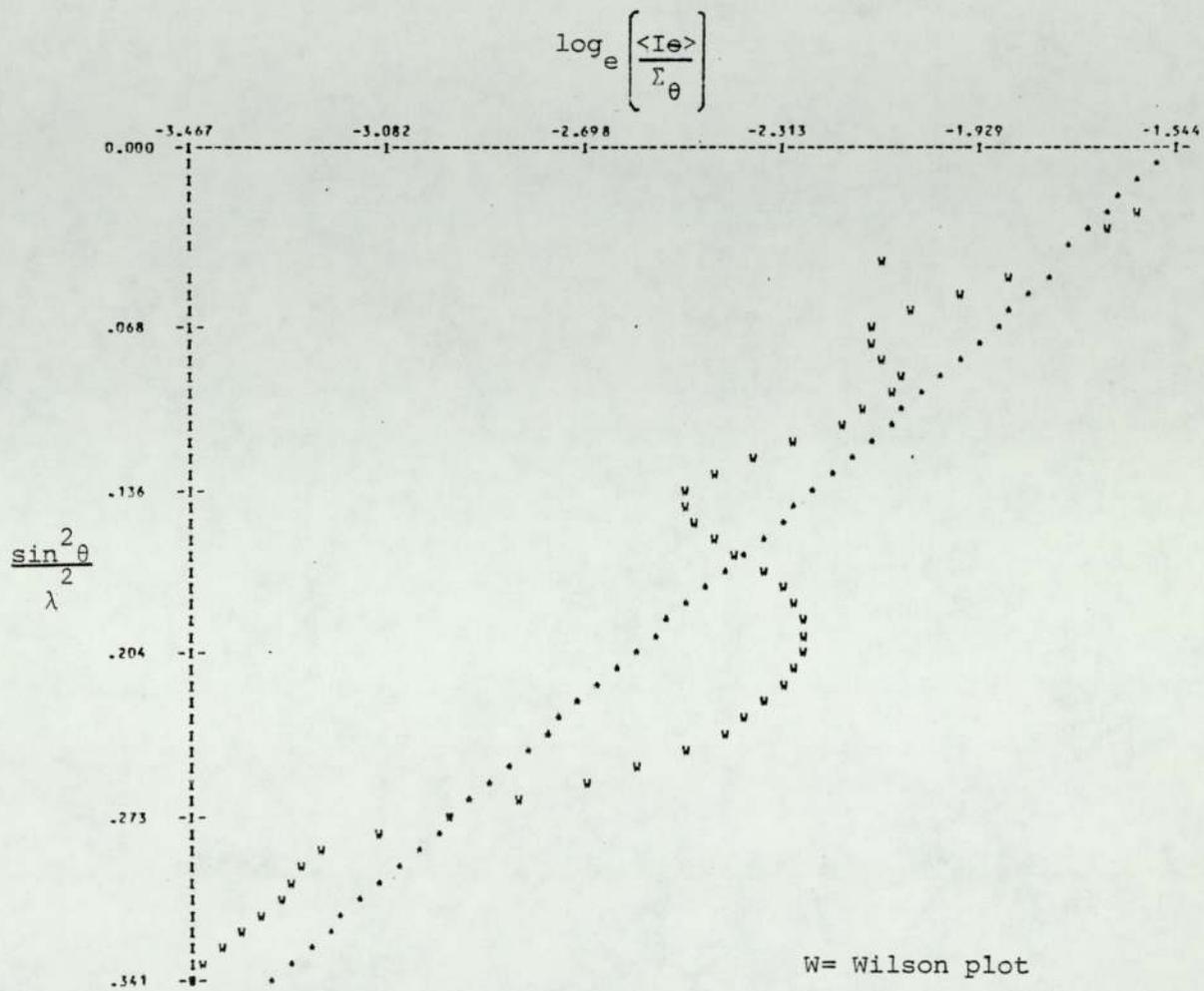
TABLE 6.16   Positional parameters (fractional co-ordinates x 10^4) with estimated standard deviations in parentheses.

Atom	x/a	y/b	z/c
N(1)	1494(4)	1810(2)	8654(2)
N(2)	1715(4)	966(2)	9007(2)
N(3)	954(3)	260(1)	8134(2)
C(4)	-156(4)	363(2)	6874(2)
O(4)	-918(3)	-254(1)	6155(2)
N(5)	-288(3)	1300(1)	6520(2)
C(6)	-1219(4)	1717(2)	5368(3)
N(7)	-1098(3)	2619(2)	5458(2)
C(8)	-35(4)	2814(2)	6714(2)
C(81)	357(4)	3788(2)	7152(2)
N(82)	-695(3)	4420(2)	6334(2)
O(82)	1538(3)	3991(1)	8192(2)
C(8A)	484(4)	2010(2)	7389(2)
C(31)	1320(5)	-675(2)	8720(3)
C(32)	3403(5)	-1026(2)	8841(3)
O(33)	3659(3)	-1255(1)	7596(2)
C(34)	5539(8)	-1696(4)	7703(5)
H(1)	1008(47)	-635(19)	8105(28)
H(2)	252(42)	-1083(18)	9619(25)
H(3)	4358(43)	-585(18)	9226(24)
H(4)	3658(44)	-1604(19)	9437(25)
H(5)	5745(62)	-1830(23)	6845(38)
H(6)	5786(81)	-2213(31)	8334(41)
H(7)	6609(80)	-1609(32)	8088(47)
H(8)	-1846(39)	1372(16)	4588(23)

Atom	x/a	y/b	z/c
H(9)	-1454(37)	4272(19)	5540(14)
H(10)	-683(46)	5023(8)	6562(26)

TABLE 6.17   Anisotropic temperature factors (non-hydrogen atoms)

isotropic temperature factors (hydrogen atoms)

ATOM	U11	U22	U33	U23	U13	U12
N(1)	.0375(15)	.0405(15)	.0273(12)	.0004(10)	.0025(11)	.0008(12)
N(2)	.0369(15)	.0450(16)	.0314(13)	.0003(12)	.0027(12)	.0006(13)
N(3)	.0310(14)	.0358(14)	.0299(12)	.0014(11)	.0064(11)	.0041(12)
C(31)	.0415(19)	.0407(19)	.0377(17)	.0062(15)	.0120(15)	.0040(16)
C(32)	.0401(19)	.0369(19)	.0278(16)	-.0017(14)	.0018(14)	.0080(16)
O(33)	.0426(13)	.0415(12)	.0305(10)	-.0052(9)	.0037(9)	.0092(11)
C(34)	.0676(29)	.0818(33)	.0617(27)	-.0010(24)	.0192(24)	.0382(25)
C(4)	.0259(16)	.0376(18)	.0321(15)	-.0024(13)	.0082(13)	.0049(14)
O(4)	.0049(14)	.0379(12)	.0405(11)	-.0078(10)	.0038(11)	-.0020(11)
N(5)	.0278(13)	.0371(13)	.0252(12)	-.0029(10)	.0043(10)	.0025(11)
C(6)	.0348(18)	.0416(18)	.0245(15)	-.0036(14)	.0024(14)	.0048(15)
N(7)	.0320(14)	.0352(15)	.0270(12)	-.0011(10)	.0060(10)	.0015(12)
C(8)	.0216(15)	.0378(16)	.0265(13)	-.0004(12)	.0096(12)	.0030(13)
C(81)	.0302(16)	.0402(18)	.0320(16)	-.0003(14)	.0138(14)	-.0034(15)
N(82)	.0479(18)	.0322(16)	.0445(16)	-.0007(14)	.0004(14)	.0038(14)
O(82)	.0494(13)	.0469(13)	.0340(11)	-.0066(9)	-.0008(11)	-.0070(11)
C(8A)	.0240(14)	.0345(16)	.0256(13)	-.0054(12)	.0063(12)	.0006(14)

	Uiso
H(1)	.0606(91)
H(2)	.0443(82)
H(3)	.0463(82)
H(4)	.0455(87)
H(5)	.1003(111)

H(6)	.1439(203)
H(7)	.1447(232)
H(8)	.0360(73)
H(9)	.0526(95)
H(10)	.0469(93)

TABLE 6.18   Bond distances in $\overset{\text{O}}{\text{A}}$ with estimated standard deviations in parentheses.

Bond	Interatomic distance $\overset{\text{O}}{\text{A}}$
N(1)-N(2)	1.268(3)
N(2)-N(3)	1.385(3)
N(3)-C(31)	1.475(3)
C(31)-C(32)	1.498(4)
C(32)-O(33)	1.421(3)
O(33)-C(34)	1.422(4)
N(3)-C(4)	1.372(3)
C(4)-O(4)	1.201(3)
C(4)-N(5)	1.395(3)
N(5)-C(6)	1.368(3)
N(5)-C(8A)	1.388(3)
C(6)-N(7)	1.302(3)
N(7)-C(8)	1.379(3)
C(8)-C(81)	1.479(4)
C(81)-O(82)	1.229(3)
C(81)-N(82)	1.336(3)
C(8)-C(8A)	1.360(3)
C(8A)-N(1)	1.378(3)
C(31)-H(1)	1.03 (3)
C(31)-H(2)	1.04 (3)
C(32)-H(3)	1.03 (3)
C(32)-H(4)	.94 (2)
C(34)-H(5)	.98 (3)
C(34)-H(6)	.99 (3)
C(34)-H(7)	.95 (3)

C(6)-H)8)	.97(2)
N(82)-H(9)	.90 (2)
N(82)-H(10)	.90 (2)

1) H(9) initially located geometrically

2) H(9) and H(10) allowed to move under refinement, but with a
    constrained bond distance.

TABLE 6.19 Interatomic angles [°] with estimated standard deviations in parentheses. (Only important angles involving hydrogen atoms are included).

Atoms	Bond Angle [°]
N(1)-N(2)-N(3)	120.4(2)
N(2)-N(3)-C(4)	126.7(2)
N(2)-N(3)-C(31)	112.9(2)
N(3)-C(31)-C(32)	113.5(3)
C(31)-C(32)-O(33)	110.2(2)
C(32)-O(33)-C(34)	111.0(3)
C(31)-N(3)-C(4)	120.2(2)
N(3)-C(4)-O(4)	125.9(3)
N(3)-C(4)-N(5)	110.5(2)
O(4)-C(4)-N(5)	123.6(2)
C(4)-N(5)-C(8A)	122.9(2)
C(4)-N(5)-C(6)	130.4(2)
N(5)-C(6)-N(7)	111.4(2)
C(6)-N(5)-C(8A)	106.6(2)
C(6)-N(7)-C(8)	106.4(2)
N(7)-C(8)-C(8A)	110.0(2)
N(7)-C(8)-C(81)	120.5(2)
C(8)-C(81)-N(82)	114.4(2)
C(8)-C(81)-O(82)	122.3(2)
N(82)-C(81)-O(82)	123.3(3)
C(8)-C(8A)-N(5)	105.6(2)
C(8)-C(8A)-N(1)	133.8(2)
C(8A)-C(8)-C(81)	129.5(2)
C(8A)-N(1)-N(2)	118.8(2)

N(1)-C(8A)-N(5)	120.6(2)
C(81)-N(82)-H(9)	123(2)
C(81)-N(82)-H(10)	121(2)
H(9)-N(82)-H(10)	116(3)

N(1)-C(8A)-N(5)   120.6(2)

C(81)-N(82)-H(9)   123(2)

C(81)-N(82)-H(10)   121(2)

H(9)-N(82)-H(10)   116(3)

TABLE 6.20   Deviations [$\overset{o}{A}$] of non-hydrogen atoms from the least squares plane through the imidazotetrazine ring system[+]. (Atoms used in the plane calculation are marked with an asterisk).

ATOM	DEVIATION [$\overset{o}{A}$]
*N(1)	-.005
*N(2)	-.003
*N(3)	.015
*C(4)	-.032
*N(5)	.020
*C(6)	.013
*N(7)	-.010
*C(8)	-.009
*C(8A)	.011
C(31)	-.010
C(32)	1.324
O(33)	2.193
C(34)	3.405
O(4)	-.105
C(81)	-.048
O(82)	.150
N(82)	-.331

[+]   The equation of the plane is:-

$$.855x - .041y - .5172 - 3.877 = 0$$

where x, y and z are orthogonal co-ordinates in [$\overset{o}{A}$] along a*, b and c.

TABLE 6.21   Important torsion angles [°]

ATOMS	ANGLE [°]
N(7)-C(8)-C(81)-N(82)	11.8
N(7)-C(8)-C(81)-O(82)	-169.0
C(8A)-C(8)-C(81)-N(82)	-167.0
C(8A)-C(8)-C(81)-O(82)	12.1
N(1)-N(2)-N(3)-C(31)	-179.0
N(2)-N(3)-C(31)-C(32)	-79.4
N(3)-C(31)-C(32)-O(33)	-71.5
C(31)-C(32)-O(33)-C(34)	-172.8
O(4)-C(4)-N(3)-C(31)	.5
C(4)-N(3)-C(31)-C(32)	104.7
C(4)-N(5)-C(8A)-N(1)	2.3
C(4)-N(5)-C(8A)-C(8)	-176.5
C(6)-N(5)-C(8A)-C(8)	.3
C(6)-N(5)-C(8A)-N(1)	179.0

TABLE 6.22   Hydrogen bond contact distances and angles

Hydrogen Bond	Angle at H [°]	N .... O distance [°]
$N(82)_I$-H(9) .... $O(4)_{II}$	145	3.085
$N(82)_I$-H(10) ... $O(33)_{III}$	158	3.045

I, II, III refer to the symmetry operations:-

x, y, z; -1/2-x, 1/2+y, 1-z; -1/2+x, 1/2-y, z

6.4.4   Results and Discussion

The structure of 3 as proposed by Stevens et al[116] is confirmed by this crystallographic study and the numbering scheme and general conformation shown in Fig 6.12; positional parameters of the atoms are given in Table 6.16.

Unlike the chloroethyl (1) and methyl (2) analogues, this compound contains only a single molecule in the asymmetric unit. It also shows a fundamental difference in that the carboxamide oxygen does not participate in any form of intermolecular hydrogen bonding.

The orientation of the carboxamide group is the same as that found in the methyl analogue (2) and is not conducive to the intramolecular hydrogen bonding encountered in mitazolamide (1) and DTIC. Reasonable coplanarity with the imidazole ring, however, is maintained, and the twist angle about the $C(8)-C(81)$ bond between the two moieties is approximately 12.0° which is substantially larger than the same angles determined in 1 and 2. The $C(8)-C(81)$ bond distance $(1.479(4)\overset{o}{A})$ still exhibits some shortening of this bond relative to the bridging bond found in biphenyl[54] (pure Csp^2-Csp^2) again suggesting the possibility of some conjugation between the carboxamide group and the ring system. The $C(81)-N(82)$ bond $(1.336(3)\overset{o}{A})$ still shows considerable double bond character, with the sum of the bond angles about $N(82)$ being 360(6)° consistent with sp^2 hybridisation of the amide nitrogen, albeit that one should be cautious with these angles since the hydrogens attached to $N(82)$ were refined to constrained bond lengths.

The molecular dimensions, bond lengths and interatomic angles are given in Tables 6.18 and 19.

The true double bond character of $N(1)-N(2)$ $(1.268(3)\overset{o}{A})$ and $C(6)-N(7)$ $(1.302(3)\overset{o}{A})$ is confirmed, with $C(8)-C(8A)$ $(1.360(3)\overset{o}{A})$ still showing as a partial double bond, but being significantly shorter

(3σ) than the equivalent bonds in the other two analogues (1 and 2). A further interesting increase in bond length with respect to the other two analogues occurs in the C(8A)-N(1) bond (1.378(3)$\overset{o}{A}$) which may suggest something about the conjugation of this system (see later discussion). The C(4)-O(4) bond (1.201(3)$\overset{o}{A}$) is slightly shorter than may be expected in comparing it with an equivalently hydrogen bonded carbonyl group in the methyl analogue (1.205(3)$\overset{o}{A}$) but nonetheless falls in the midpoint of the C(4)-O(4) and C(4)'-O(4)' bond lengths of mitazolamide (1) neither of which participate in intermolecular contacts.

Interatomic angles determined for this structure show excellent agreement with those found in 1 and 2. Deviations in $\overset{o}{A}$ from the least squares plane through the ring atoms of the imidazotetrazine ring (Table 6.20) show that with the exception of the methoxyethyl group, the molecule is essentially planar; however, as also demonstrated by the torsion angles (Table 6.21) the carboxamide group in this compound is not as coplanar as in the other two analogues with the deviations from the least squares plane being .150$\overset{o}{A}$ and -.331$\overset{o}{A}$ for O(82) and N(82) respectively. The carbonyl oxygen O(4) also shows a greater deviation than in the other two analogues, (-.105$\overset{o}{A}$) and it seems likely that this results from hydrogen bonding since by far the closest to it is that of O(4)' in the methyl analogue (-.063$\overset{o}{A}$) which is also hydrogen bonded. The small but positive deviations of N(5) (.020$\overset{o}{A}$) and C(8A) (.011$\overset{o}{A}$) and the torsion angles about the bridging bond N(5)-C(8A) suggest an angle between the two rings of about 1.7°.

As far as the methoxyethyl group is concerned, C(31) is close to being coplanar with the ring system (-.010$\overset{o}{A}$) whilst C(32), O(33) and C(34) move systematically out of the plane with C(34) having a

deviation of 3.405Å, and the O(33)-C(34) bond being almost perpendicular to the plane.

The molecular packing and hydrogen bonding are shown in Figures 6.14 and 6.15 and the contact distances given in Table 6.22 together with the symmetry operations involved. A fundamental difference between the methoxyethyl compound and the others is that it does not associate into pairs, either of bases or carboxamide groups and as such the carboxamide oxygen O(82) and the imidazo nitrogen N(7) play no part in the hydrogen bonding scheme. Instead, N(82) donates two hydrogen bonds, and O(4) and O(33) each receive one. The N(82)-H(10) ... O(33) hydrogen bond (3.045Å) links molecules in a head to tail relationship by the operation of the screw axis and centre of symmetry whilst the N(82)-H(9) ... O(4) hydrogen bond (3.085Å) links adjacent molecules by the operation of the screw axis. The overall effect of this is to produce a weak three dimensional hydrogen bonded lattice.

The infrared spectrum shows two single peaks, one centred at 1665 cm^{-1} (carboxamide C=O) and the other at 1720 cm^{-1} (tetrazine ring C=O) consistent with the solid state structure of the compound.

Figure 6.12 Molecular structure and numbering scheme for 8-
carbamoyl-3-(2-methoxyethyl) imidazo [5,1-d] 1,2,3,5-
tetrazin-4(3H)-one(3)
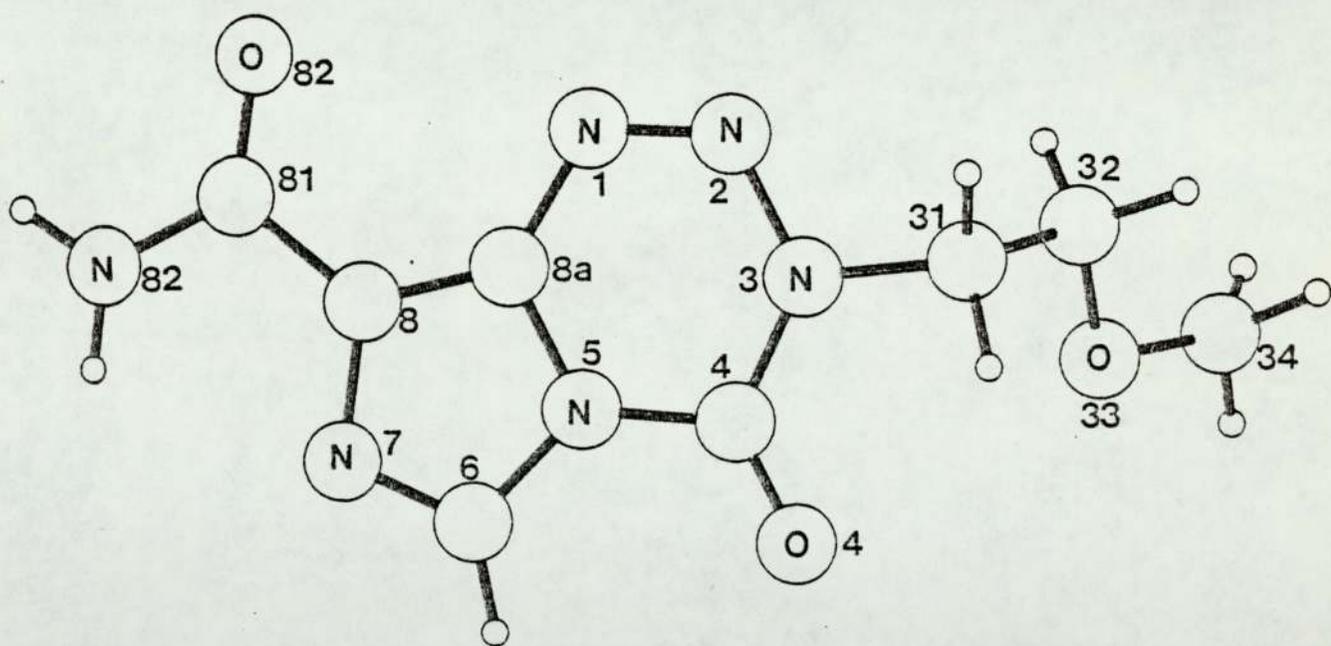
Figure 6.13   Stereo pair (3)
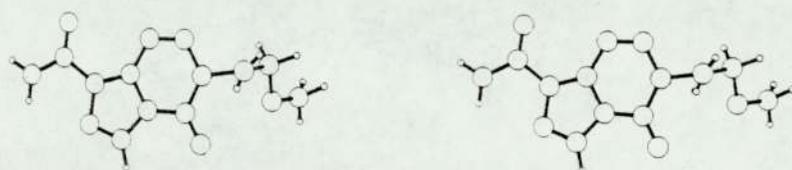
Figure 6.14   Molecular packing and hydrogen bonding (3)
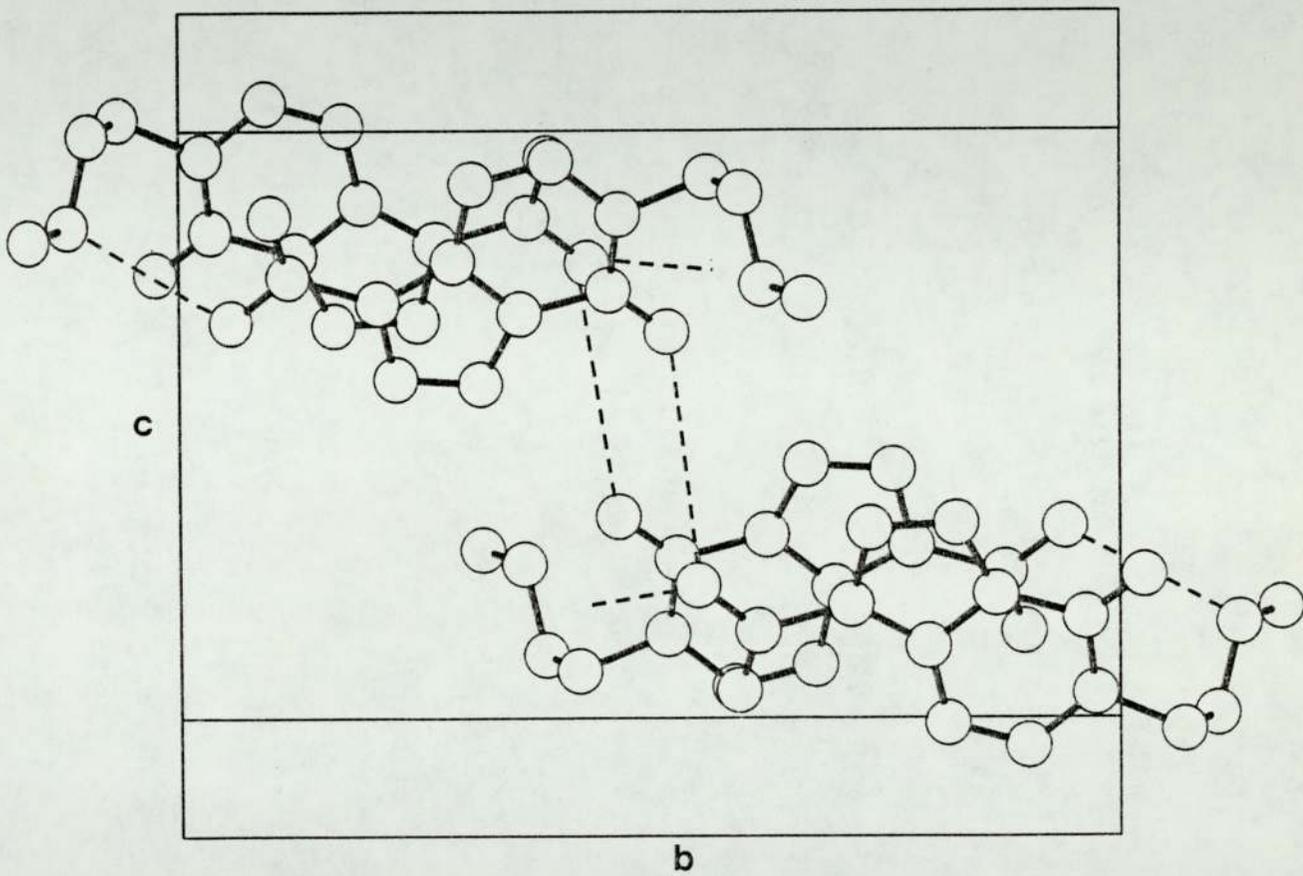
Figure 6.15   Extended packing diagram for (3)
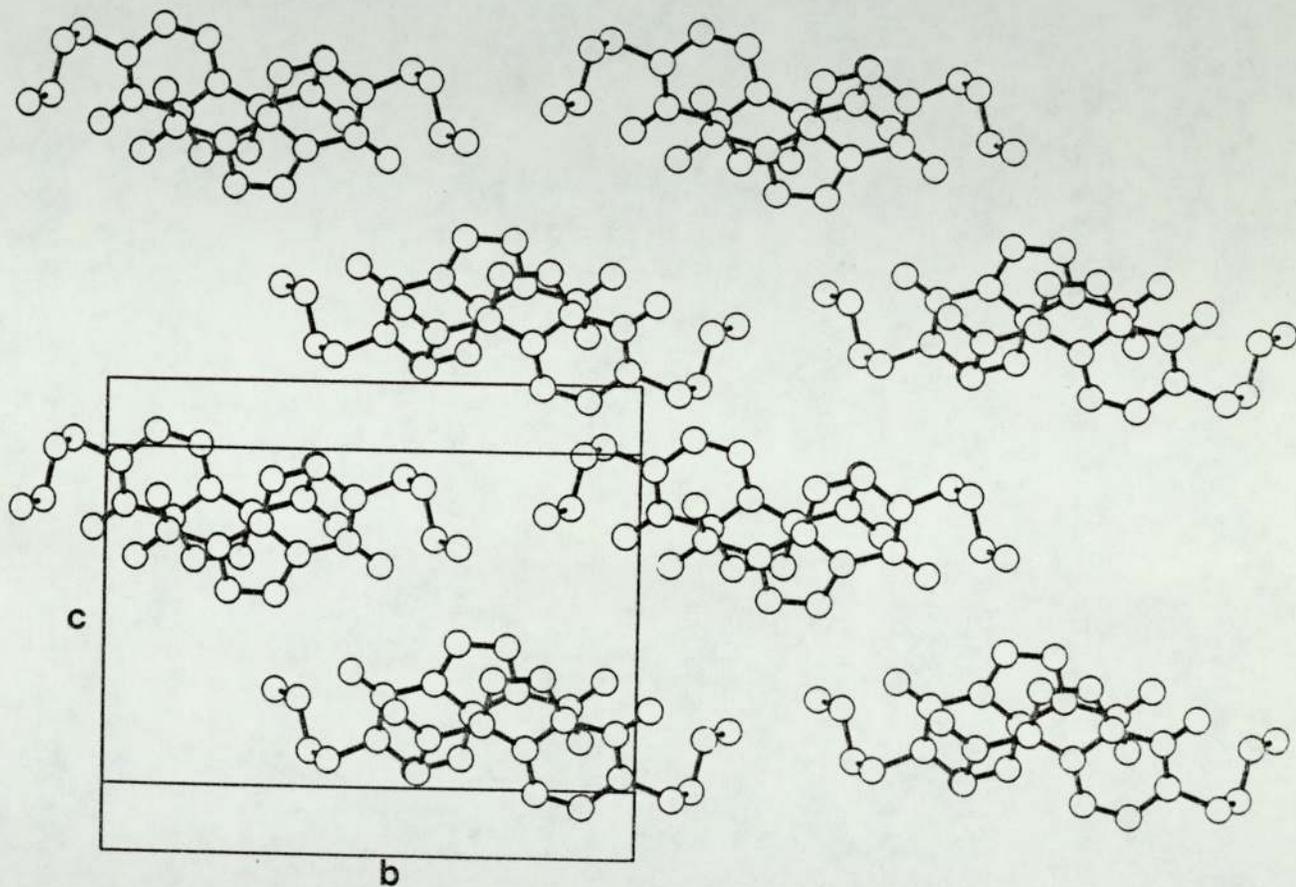
6.5   Discussion on Mitozolamide (CCRG 81010),

             its analogues and related compounds

Although the structures of the three compounds have been unequivocally confirmed and individually discussed, certain questions still remain unanswered:-

Why do the three drugs differ so much in their effectiveness as anti-cancer agents?

Does the imidazocarboxamide moiety play an active role in the effectiveness of these compounds as antitumour agents?

Is there any evidence to support the proposed mechanism of degradation to the monoalkytriazene?

Can any obvious comparisons be made between these compounds and known similar antitumour agents?

The exact mode of action of the active drugs is as yet undetermined, albeit that current hypotheses are under examination. In the case of the title compound (1) a considerable amount of evidence suggests that the active species involved is not the intact molecule, but the monoalkyltriazene (monochloroethyl triazene MCTIC for 1) after some form of activation process[116]. Depending upon reaction conditions, the title compound (1) may decompose in two ways[116]. At elevated temperatures, in organic solvents, retro-cycloaddition in the dominant reaction (the reverse of the synthesis) in which the $C(4)-N(5)$ and $N(2)-N(3)$ bonds are expected to cleave. Application of the formula used by Hjortas[123] to obtain π-bond orders from bond distances in 1,2,3-benzotriazin-4(3H)-one yields orders of 0.35 and 0.41 for these bonds averaged over the two independent molecules of compound 1. These are the weakest and presumably most labile ring bonds in the entire molecule, but in view of their significant multiple bond character, it is understandable that input

of heat is required to rupture them. The alternative decomposition takes place in aqueous media and involves attack by water at C(4) followed by ring opening. The C(4)-O(4) bond is short and strong, but the π-bond orders from C(4) to N(3) and N(5) (0.45 and 0.35) are weaker than corresponding values in urea[124] (0.64 each). Presumably, less electron density is being transferred from adjacent atoms in the C(4)-O(4) system here than in urea, and nucleophilic attack by water may be facilitated.

Stevens et al[116] on the basis of such a reaction in 5% aqueous sodium carbonate have proposed the possible degradation shown in Figure 6.16 of which are one of the final products is AIC (azidoimidazocarboxamide). Using [6-C[14]] labelled CCRG 81010, Tisdale and Horgan[125] found at 80% conversion to AIC occurs in tissue culture media at 37° over a 24 hour period.

It seems, therefore, quite likely that the title compound (1) is a stable prodrug form of MCTIC particularly since mitozolomide and MCTIC have similar activities _in vivo_ and _in vitro_. It is however, not possible to explain the reduced activity of 2 and 3 on the basis of less tendency to hydrolysis and ring opening inasmuch as the important tetrazinone bond lengths in compounds 1, 2 and 3 are identical within 0.010Å (Table 6.23) and would therefore be likely to undergo a similar hydrolysis. The one possible exception to this conclusion might be if the hydrolysis were enzymically controlled (Figure 6.2) in which case they may differ in binding power to such an enzyme. This would be quite likely in view of the differences in size and polarity of the side chains, particularly in the case of the methoxyethyl compound (3) which displays an ability to hydrogen bond to O(33).

Comparisons in activity have up to now been made between mitazolomide (CCRG 81010) and DTIC (Dacarbazine) basically since Dacarbazine is a clinically available drug. If the preceding theory is correct, and the monoalkyltriazene is the active species, the comparisons should be between mitazolomide (1) and DCTIC (4) and the 3-methyl substituted analogues (2) and DTIC (7).

DCTIC is a known antitumour agent[107] with a potency somewhat greater than DTIC[105] but unfortunately undergoes a spontaneous transformation at room temperature in the solid state to form an inactive isomer[126,127]. Abrahams and co-workers[127] suggested a structural similarity between DCTIC and guanine (although because of its instability they were unable to carry out structure determination) and suggested that the major activity of the compound was due not to the N-di(chloroethyl) [or N-bis(chloroethyl)] part of the molecule, but the remainder, which may well bind and subsequently bond covalently by means of alkylating groups to, for example, the active site of enzymes for which guanine is a substrate. A similar proposal for compound 1 was forwarded by Stevens, albeit as purely a target for criticism, and is shown in figures 6.17 and 6.18. A comparison of the structures from this work and DTIC is given in terms of bond distances in Table 6.23.

A direct comparison is made somewhat difficult by the fact that DTIC exists in the solid state as two tautomers, one of each per asymmetric unit, with in the one case N(7) being protonated whilst in the other N(5) (Fig 6.7). Assuming that such specificity was conferred by the imidazo–carboxamide moeity as Abrahams et al suggest, such tautomerism could seriously affect the potency if it persists when the molecule is solvated and if it imposes a specific pattern of hydrogen bonding. DTIC resembles the unprimed molecule

of compound 1, in that it contains an equivalent internal hydrogen
bond, in this case in both molecules in the asymmetric unit. The
carboxamide groups again are approximately coplanar with the
imidazole ring at 2.5° and 1.0° twist angles about the C-C bond. Its
intermolecular hydrogen bonding shows a closer relationship to
compound 2 in that it forms a dicarboxamide link (Figure 6.19). What
of the importance of the carboxamide group attached to C(8)? In each
of the cases mentioned so far, compounds 1, 2 and 3 and DTIC, the
carboxamide group is essentially coplanar with the imidazole ring,
with the greatest deviation from coplanarity occurring in the
methoxyethyl compound with a twist angle of 12.0°. The orientation
of the carboxamide group with respect to the imidazole ring has been
seen to be vital for internal hydrogen bonding since only one of the
two coplanar orientations will allow this. In the case of
mitazolomide 1, NMR studies suggest that this internal hydrogen bond
is maintained in solution at least in the solvent in which the NMR
was performed (Figure 6.20). In the case of the open chain monoalkyl
triazenes, e.g. MCTIC, both orientations of the carboxamide group can
give rise to internal hydrogen bonding, in one of the two possible
tautomeric forms, and Stevens et al[116] suggest that intra-molecular
hydrogen bonding may stabilise the aminoimidazole tautomer thus
activating the electrophilic ∝-methylene group of MCTIC to attack by
nucleophiles (Figure 6.16). In compound 1 it can be seen that the
coplanarity of the carboxamide groups with the imidazole rings
facilitates a planar base pairing interaction involving N(7) of the
imidazole ring: such base pairing would be unlikely for a
significantly out of plane carboxamide group. Similarly, in compound
2 and DTIC planar carboxamide pairs are facilitated. It is possible
that such maximisation of hydrogen bonding facilities might be
important if these compounds bind to a target site.

The existence of the carboxamide group and its coplanarity with the ring system may be significant in one additional respect: conjugation between the two moieties. An important question in considering the possible activation of these compounds to the monoalkytriazenes concerns which bonds are most likely to cleave under given reaction conditions. Schwalbe et al[128] have studied the bond distances in 1,2,3-benzotriazinones (10). In the two structures they examined (R=OH, R=H) the N(1) to ring bond distances were 1.398Å and 1.402Å respectively and indeed this class of compounds do undergo bond fission at this point. In compounds 1 to 3, this bond C(8A)-N(1) is substantially shorter and stronger, ranging from 1.368(5)Å to 1.378(3)Å in the three compounds, the methoxyethyl substituted (3) which has largest twist angle about the C(8)-C(81) bond (12.0°) also has the longest and presumably weakest C(8A)-N(1) bond. In terms, therefore, of available bonds to cleave in the tetrazinone ring, in all cases in compounds 1 to 3, the C(4)-N(5) bond rather than C(8A)-N(1) is the longest and presumably the weakest (Table 6.23) (C(4)-N(5) bonds range from 1.388Å to 1.396Å). Accordingly, the mechanism proposed by Stevens et al (Figure 6.16) of hydrolysis and subsequent cleavage at this point, is a reasonable one. Whether the significant shortening of the C(8A)-N(1) bond in compounds 1 to 3 compared to the benzotriazinones is due to the imidazole ring, or the conjugation of the carboxamide group with the ring system, or indeed both, will need to be tested by examining the crystal structure of a compound with the same ring system but no carboxamide attachment.

The preparation of alternative active analogues to mitazolamide is continuing at Aston and the associated drug company with substituents other than a carboxamide moeity. Their crystal structures are currently being studied, and correlation will be made

between these and existing structures.

In conclusion it appears likely that compound 1 acts as a stable prodrug form of MCTIC, an analogous function possibly applying to compound 2. Their differences in activity on this basis would be as expected for DCTIC and DTIC and indeed determined experimentally. It is possible that whilst 1 degrades to produce a bifunctional alkylating agent, 2 produces a monofunctional one and bifunctional alkylating agents are known to be more potent. However, it is also possible that both act via different routes, as it as yet unexplained as to why, when the carboxamide groups in 1 and 2 are exchanged for a cyano group, the 3-methyl compound remains active whilst the chloroethyl compound is rendered inactive. The exact reason for the inactivity of 3 has only been hypothesised in this section; however it has been found experimentally that a fundamental requirement for activity of these compounds is a methyl or chloroethyl substituent, all others with the exception of the ethyl substituted compound, which shows very slight activity at high doses, are inactive.

Figure 6.16 Possible degradation of mitozolomide (CCRG 81010) Stevens et al[116]
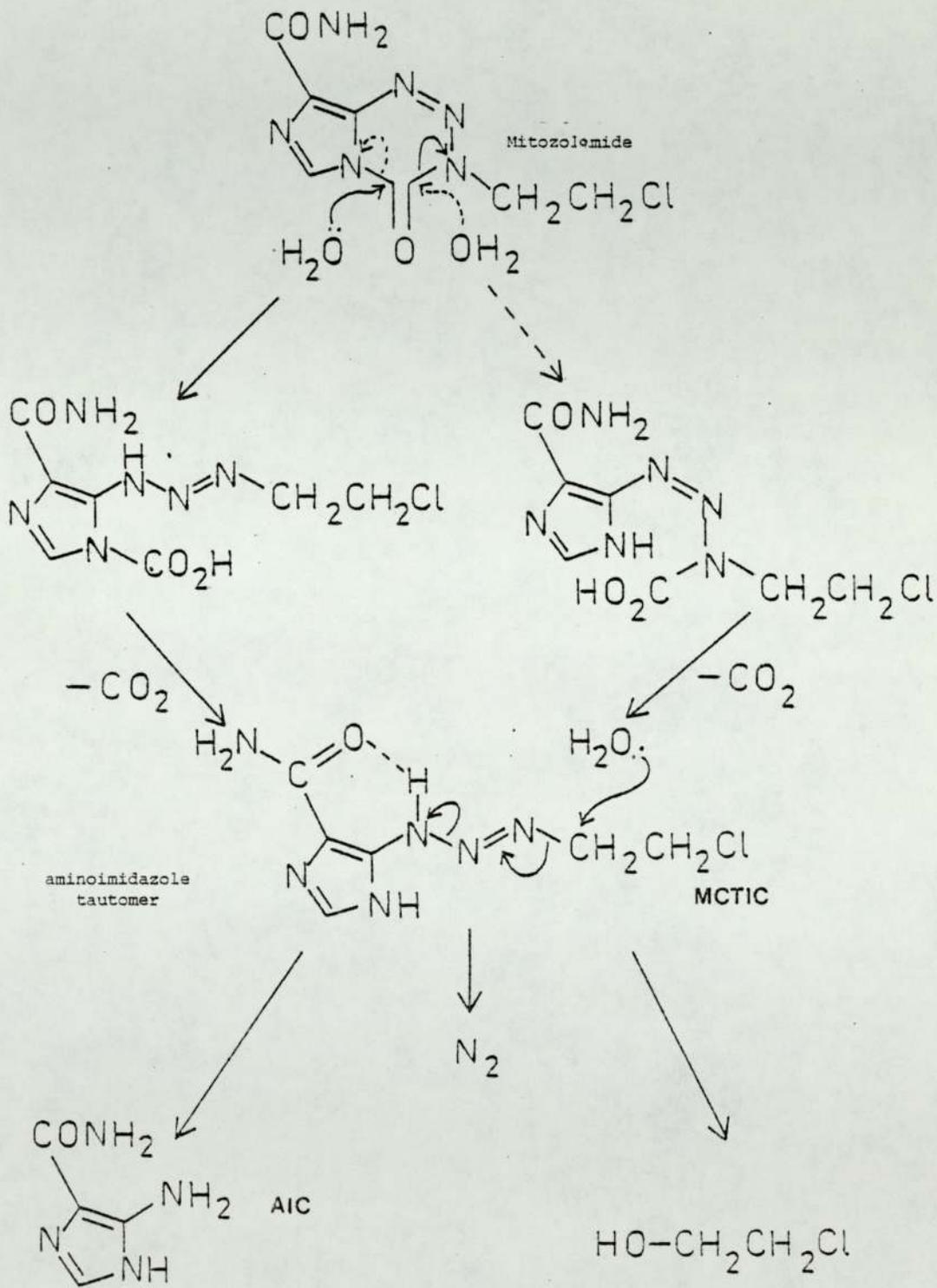
Figure 6.17 MITOZOLOMIDE: A SITE DIRECTED IRREVERSIBLE INHIBITOR?
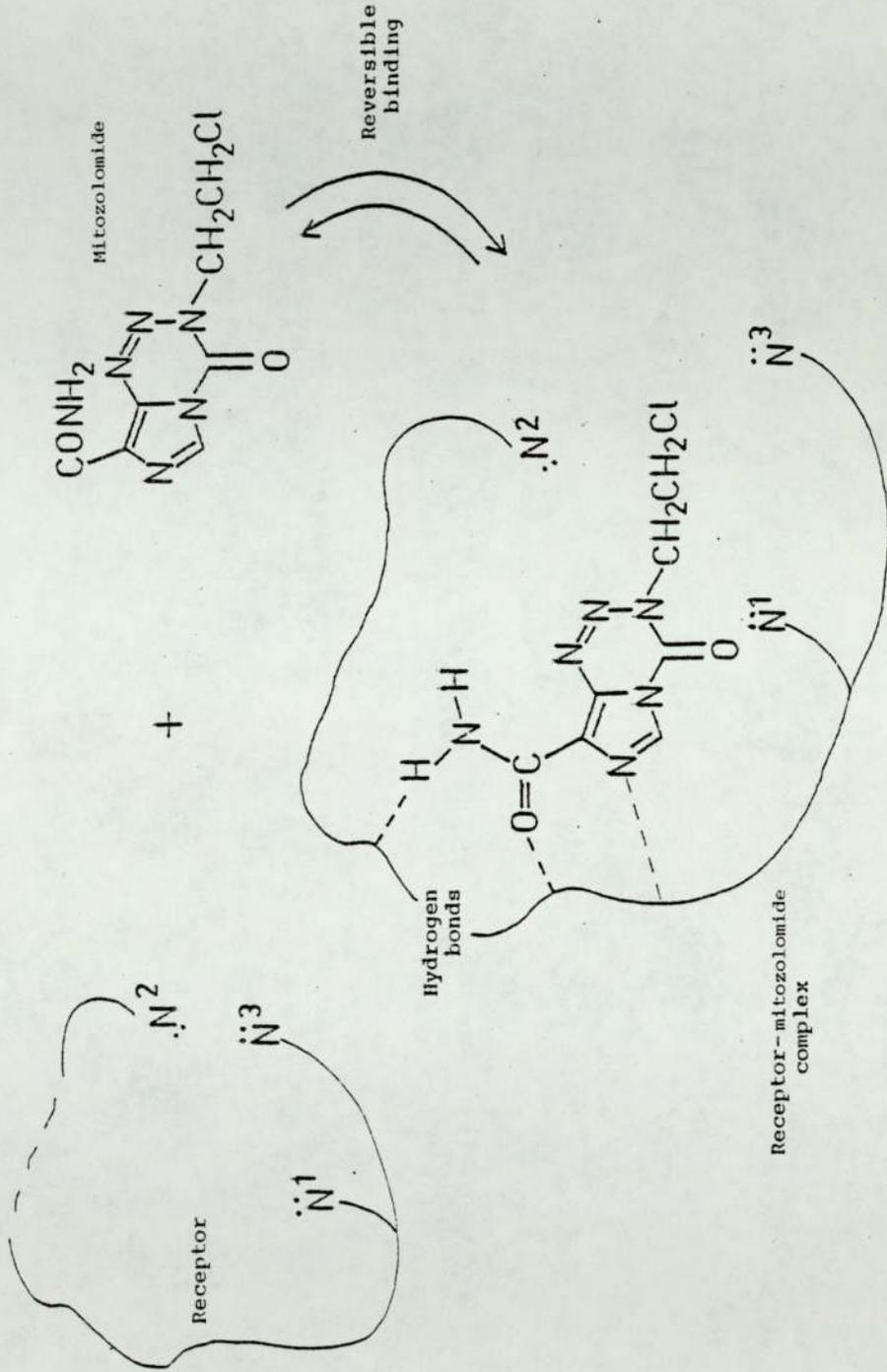
Figure 6.18 THE CARBAMDUBALK (CARBAMOYLATION-DOUBLE ALKYLATION) REACTION
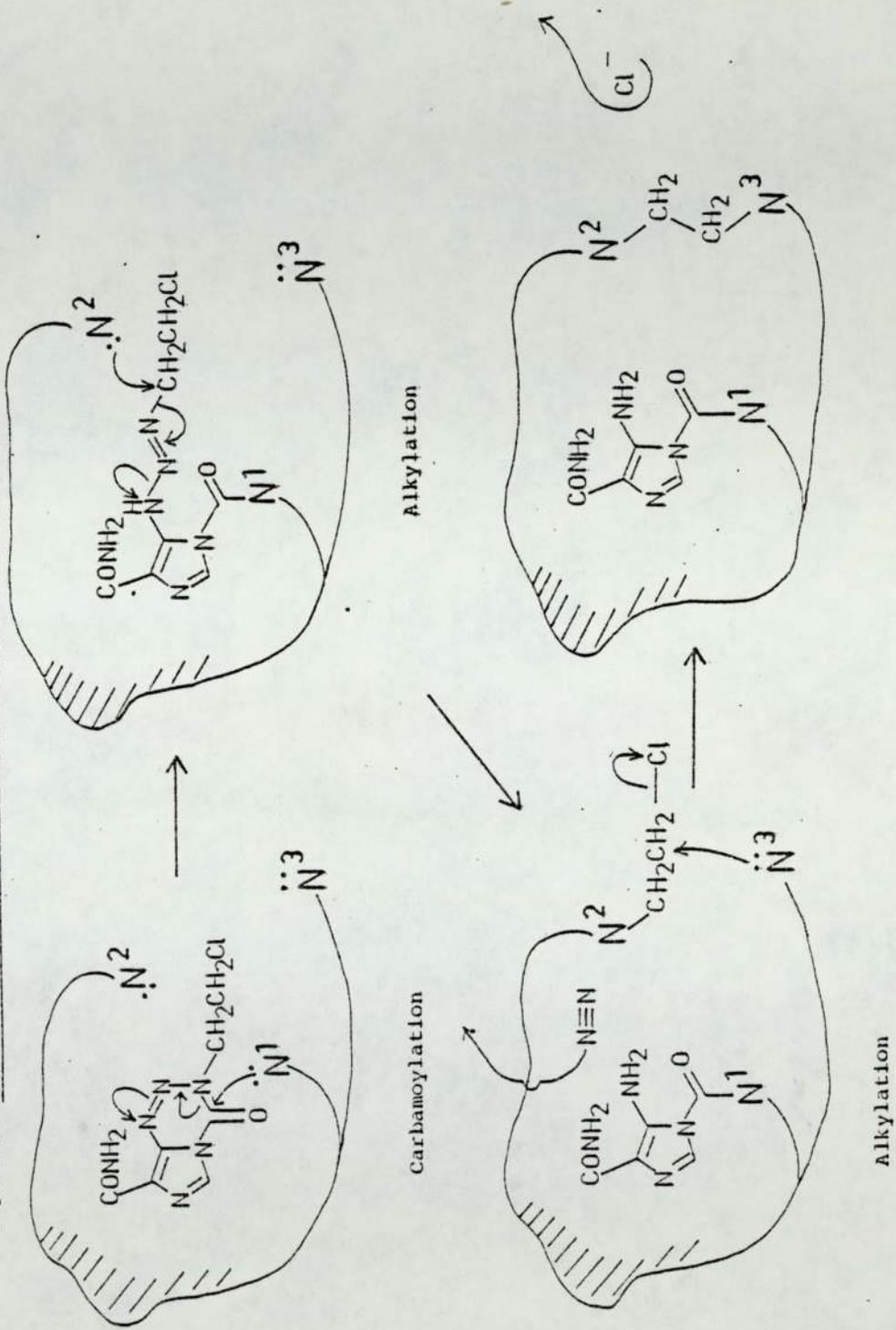
Figure 6.19   Hydrogen bonding

R=CH₂CH₂Cl
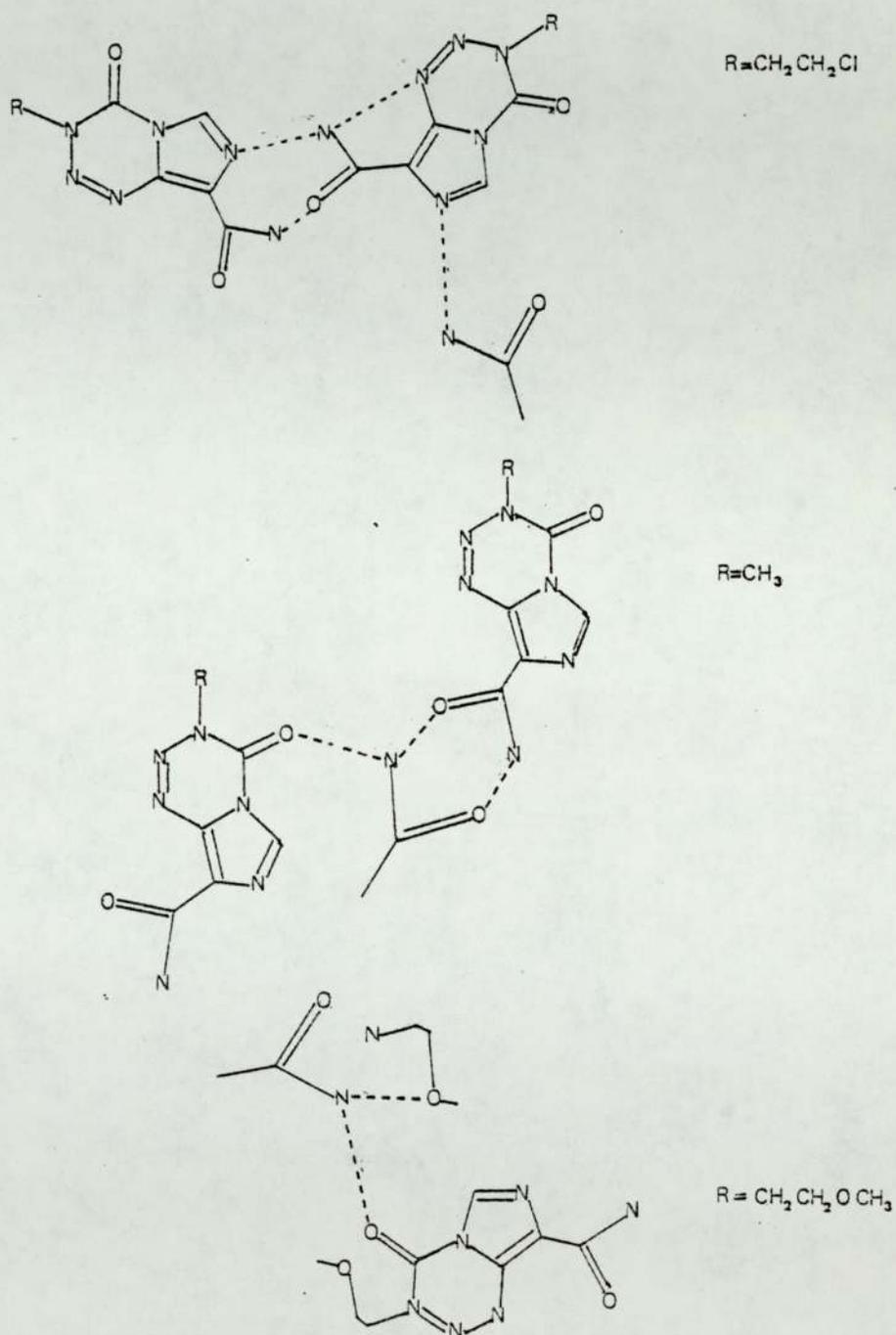
R=CH₃

R = CH₂CH₂OCH₃

Table 6.23

BOND DISTANCES IN Å WITH STANDARD DEVIATIONS IN PARENTHESES FOR MITOZOLOMIDE, ITS TWO ANALOGUES AND DTIC

BOND	1 $R=CH_2CH_2Cl$ Unprimed	Primed	2 $R=CH_3$ Unprimed	Primed	3 $R=-CH_2CH_2OCH_3$	*DTIC equivalent bonds Unprimed	Primed
N(1)-N(2)	1.266(3)	1.257(3)	1.272(4)	1.265(4)	1.268(3)	1.285(4)	1.285(4)
N(2)-N(3)	1.374(3)	1.386(3)	1.382(4)	1.379(3)	1.385(3)	1.304(4)	1.309(4)
N(3)-R	1.469(3)	1.460(3)	1.461(6)	1.455(5)	1.475(3)		
N(3)-C(4)	1.378(3)	1.377(3)	1.371(5)	1.371(5)	1.372(3)		
C(4)-O(4)	1.198(3)	1.205(3)	1.196(4)	1.205(3)	1.201(3)		
C(4)-N(5)	1.396(3)	1.394(3)	1.389(4)	1.388(4)	1.395(3)		
N(5)-C(6)	1.373(3)	1.364(3)	1.368(5)	1.364(5)	1.368(3)	1.352(4)	1.328(4)
C(6)-N(7)	1.306(3)	1.300(3)	1.303(4)	1.311(4)	1.302(3)	1.332(4)	1.338(4)
N(7)-C(8)	1.365(3)	1.381(3)	1.368(4)	1.378(4)	1.379(3)	1.388(4)	1.380(4)
C(8)-C(8A)	1.369(3)	1.374(3)	1.370(5)	1.378(4)	1.360(3)	1.379(4)	1.370(4)
C(8A)-N(5)	1.379(3)	1.391(3)	1.384(3)	1.388(3)	1.388(3)	1.375(4)	1.376(4)
C(8)-C(81)	1.490(3)	1.487(3)	1.486(4)	1.478(4)	1.479(4)	1.470(4)	1.477(4)
C(81)-O(82)	1.216(3)	1.229(3)	1.227(4)	1.234(4)	1.229(3)	1.229(4)	1.235(4)
C(81)-N(82)	1.326(4)	1.330(3)	1.326(5)	1.322(5)	1.336(3)	1.338(4)	1.323(4)
C(8A)-N(1)	1.371(3)	1.374(3)	1.368(5)	1.370(4)	1.378(3)	1.384(4)	1.393(4)

Figure 6.20 NMR MITOZOLOMIDE
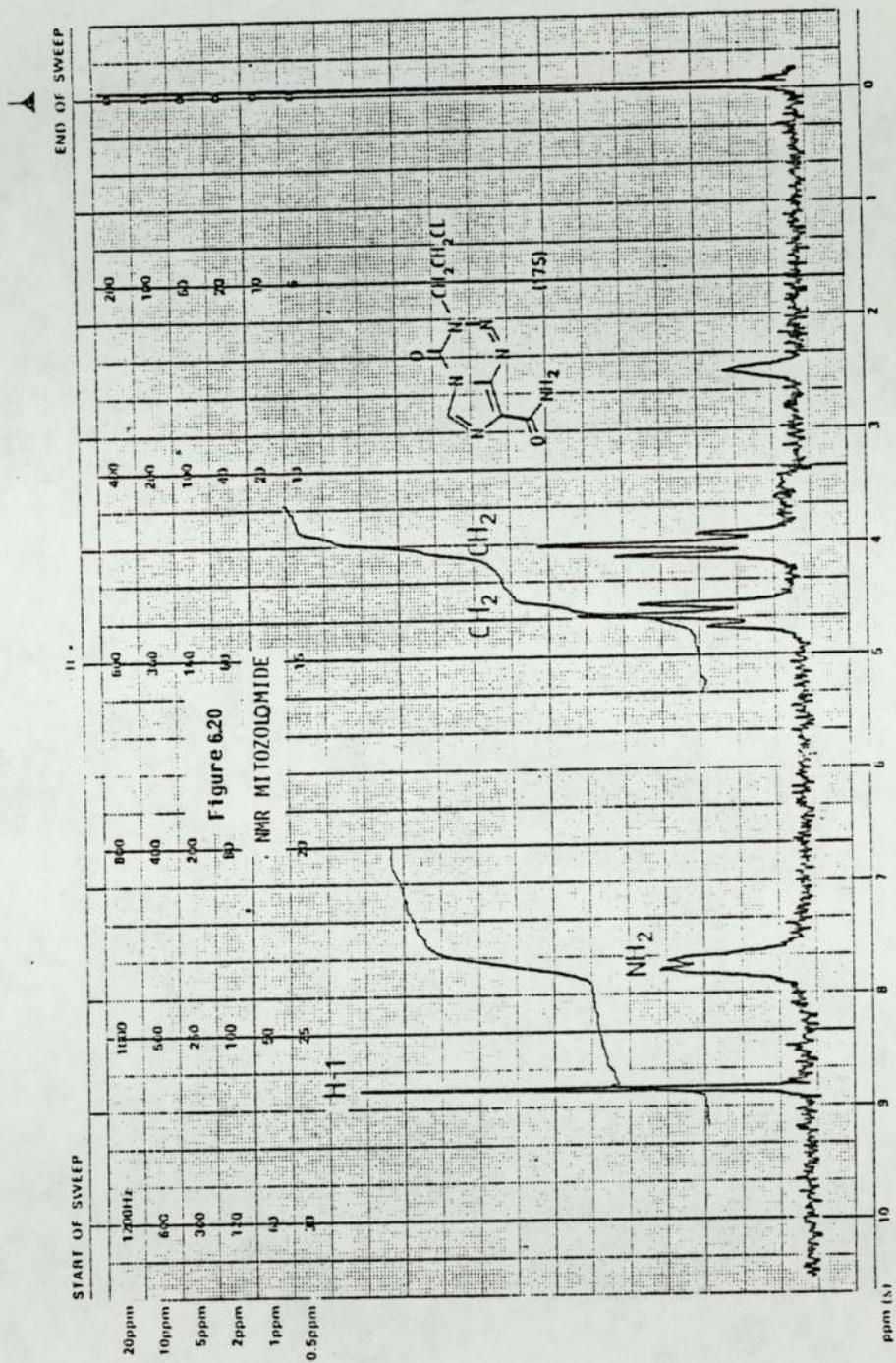
APPENDIX 1

References

1. Main, P., Woolfson, M. M. and Germain, G. (1971). MULTAN: A Computer Programme for the Automatic Solution of Crystal Structures. University of York.

2. Sheldrick, G. (1976). SHELX: Programme for Crystal Structure Determination. Cambridge University.

3. Stewart, J. M., Kruger, G. J., Ammon, H. L., Dickinson, C. and Hall, S. R. (1972). The XRAY System - version of June 1972. Tech. Rep. TR-192, Computer Science Centre, University of Maryland.

4. Motherwell, W. D. S. (1972). Pluto - A Programme for Plotting Molecular and Crystal Structures. University of Cambridge, England.

5. Hamilton, W. C., Rollett, J. S. and Sparkes, R. A. (1965). Acta Cryst., $\underline{18}$, 129.

6. Batchelor, F. R., Doyle, F. P., Naylor, J.H.C. and Rollinson, G.N. (1959). Nature (London) $\underline{183}$, 257.

7. Strominger, J. L., Izaki, K., Matsuhashi, M. and Tipper, D. J. (1967). Proc. Fed. Am. Soc. Exp. Biol., $\underline{26}$, 9.

8. Tipper, D. J. and Strominger, J. L. (1965). Proc. Nat. Acad. Sci. USA, $\underline{54}$, 1133.

9. The British Pharmaceutical Codex, Eleventh Edition. The Pharmaceutical Press.

10. Dexter, D.D. and van der Veen, J.M. (1978). J. Chem. Soc. Perkin I, 185-190.

11. Buckwalter, F.H. and Dickinson, H.L. (1958) J. Amer. Pharm. Ass. (Sci. Edn.) 47, 661-666.

12. Ballard, B. E. and Nelson, E. (1975). In Remington's Pharmaceutical Sciences, 15th Edn., p. 1642, Easton, Pa.: Mack.

13. Macek, T.J. (1975). In Remington's Pharmaceutical Sciences, 15th Edn., p. 1359, Easton, Pa.: Mack.

14. Lowe, P.R. and Schwalbe, C.H. (1978). J. Pharm. Pharmac. p.92.

15. Simpson, P.G, Dobrott, R.D. and Lipscomb, W.N. (1965). Acta Cryst., 18, 169.

16. Hughes, E.W. (1941). J. Amer. Chem. Soc., 63, 1737.

17. Abrahams, S.C. and Keve, E.T. (1971). Acta Cryst., A27, 157-165.

18. Sweet, R.M. in 'Cephalosporins and Penicillins', Academic Press, New York, 1972, 280-309.

19. Boles, M.O. and Girven, R.J. (1976). Acta Cryst. B32, 2279-2284.

20. Dexter, D.D. (1972). Acta Cryst., B28, 77-82.

21. Freeman, G.R and Bugg, C.E. (1975). Acta Cryst., B31, 96-101.

22. Wilson, A.J.C. (1942). Nature, 150, 152.

23. Abrahamsson, S, Crowfoot Hodgkin, D. and Maslen, E.N. (1963). Biochem. J., 86, 514-535.

24. Diamond, R.D. (1963). D.Phil Thesis: University of Oxford.

25. Sweet, R.M. and Dahl, L.F. (1970). J. Amer. Chem. Soc., 92, 5489-5507.

26. Hodgkin, D.C. and Maslen, E.N. (1961). Biochem. J., 79, 383.

27. Ferone, R. and Hitchings, G.H. (1966). J. Protozool., 13, 504.

28. Zakrzewski, S.F. and Nichol, C.A. (1960). J. Biol. Chem., 235, 2984.

29. Futterman, S. (1957). J. Biol. Chem., 228, 1301.

30. Rader, J.I. and Huennekens, F.M. (1973). Folate Coenzymediated Transfer of One-carbon Groups, in "The Enzymes, Vol IX", Academic Press, New York.

31. Blakely, R.L. (1954). Biochem. J., 58, 448.

32. Jukes, T.H. and Broquist, H.P. (1963). Metabolic Inhibitors - eds. Hochester, R.M and Quatzel, J.H. Academic Press, New York, 481-534.

33. Goldstein, A. "Principles of Drug Action". (1974). 2nd Edition. John Wiley and Sons.

34. Franklin, T.J. and Snow, G.A. (1975). "Biochemistry of Antimicrobial Action", Chapman and Hall, London, 2nd edition, 139-147.

35. Bell, P.H. and Roblin, R.O. (1942). J. Amer. Chem. Soc., 64, 2905.

36. Werkheiser, W.C. (1961). J. Biol. Chem., 236, 888-893.

37. Li, M C, Hertz, R. and Bergenstal, D.M. (1958). New England J. Med., 259, 66.

38. Goth, A. (1972). "Medicinal Pharmacology", Chapter 54, Mosby, St Louis.

39. Hitchings, G.H, Elion, G.B, van der Werff, H. and Falco, E.A. (1948). J. Biol. Chem., 174, 765.

40. Falco, E.A, Goodwin, L.G, Hitchings, G.H, Rollo, I.M. and Russel, P.B. (1951). Brit. J. Pharmacol., 6, 185.

41. Goodwin, L.G. (1949). Nature, 164, 1133.

42. Schwalbe, C.H. and Cody, V. "Structural Studies of Antifolate Drugs", in "Chemistry and Biology of Pteridines". (1983). Editor J.A.Blair, Walter de Gruyter and Co, Berlin, New York.

43. Cody, V. and Zakrzewski, S.F. (1982). J. Med. Chem., 25, 427-430.

44. Zakrzewski, S.F. (1963). J. Biol. Chem., 238, 1485-1490.

45. Baker, B.R. (1959). Cancer Chemother. Rep., 4, 1-10.

46. Matthews, D.A, Alden, R.A, Bolin, J.T., Freer, S.T., Hamlin, R., Xuong, N., Kraut, J., Poe, M., Williams, M. and Hoogsteen, K. (1977). Science, 197, 452-455.

47. Matthews, D.A, Alden, R.A, Bolin, J.T., Filman, D.T., Freer, S.T., Hamlin, R., Hol, W.G.T., Kisluik, R.L, Pastore, E.J., Plante, L.T., Xuong, N. and Kraut J. (1978). J. Biol. Chem., 253, 6946-6954.

48. Hunt, W.E, Schwalbe, C.H, Bird, K. and Mallinson, P.D. (1980). J. Biochem., 187, 533-536.

49. Koetzle, T.F. and Williams, G.B. (1976). J. Amer. Chem. Soc., 98, 2074-2078.

50. Cody, V. and Zakrzewski, S.F. (1982). J. Med. Chem., 25, 427-430.

51. Schwalbe, C.H., Wong, K.P., Lowe, P.R., Randall, A.J. and Williams, G. J. B. (1982). British Crystallographic Association Conference, Durham, UK.

52. Stevens, M.F.G. Private Communication.

53. Gillespie, R.J. and Nyholm, R.S. (1957). Q. Rev. Chem. Soc. 11, 339-380.

54. Robertson, G.B. (1961). Nature, 191, 593.

55. Domenicano, A., Vaciago, A. and Coulson, C.A. (1975). Acta Cryst., B31, 221-234.

56. Stevens, M.F.G. Private Communication.

57. CAD4 Operating Manual, Enraf-Nonius, Delft, Holland.

58. Wong, K.P. (1982). BSc Project, University of Aston.

59. Singh, C. (1965). Acta Cryst., 19, 861-864.

60. Sharma, B.D. and McConnell, J.F. (1965). Acta Cryst., 19, 797.

61. Brown, T.B., Lowe, P.R., Schwalbe, C.H. and Stevens, M.F.G. (1983). J. Chem. Soc. Perkin Trans. I, 2485-2490.

62. O'Connell, A.M. and Maslen, E.N. (1967). Acta Cryst., 22, 134.

63. Cruikshank, D.W.J. (1961). J. Chem. Soc., 5486.

64. Sass, R.L. (1960). Acta Cryst., 13, 320.

65. 'International Tables for X-Ray Crystallography', (1968). 2nd edn., Kynoch Press, Birmingham, vol. 3, 276.

66. Abdullah, N.A. (1983). BSc Project, University of Aston.

67. Goldstein, A. "Principles of Drug Action". (1974). 2nd edn., John Wiley and Sons.

68. Lai, T.F. and Marsh, R.E. (1967). Acta Cryst., 22, 884.

69. Jenks, R.J. (1983). BSc Project, Wolverhampton Polytechnic.

70. Haltiwanger, R.C. (1971). MSc Thesis, University of Virginia, Charlottesville, USA.

71. Cody, V. (1983). Cancer Biochem. Biophys., 6, 173-177.

72. Crooke, S.T. and Prestayko, A.N. "Cancer and Chemotherapy", (1981). Vol. 1, Academic Press.

73. Watson, J.D. and Crick, F.H.C. (1953). Nature (London), 171, 737-738.

74. Hall, R.H. "The Modified Nucleosides in Nucleic Acids", (1971). Columbia University Press.

75. Suhadolnik, R. J.   "Nucleoside Antibiotics", (1970).   Wiley Interscience.

76. Adams, R.L.P. and Davidson, J.N. (1981). 'The Biochemistry of Nucleic Acids', 9th edn., Chapter 7, p186-187.

77. Swindler, R. and Welch, H.D. (1957). Science, 125, 548.

78. Ostrogovich, A. and Cadariv, J. (1940). Gazz. Chim. Ital., 71, 505-524.

79. Jonas, J., Horak, M., Piskala, A. and Gut, J. (1962). Coll. Czech. Chem. Comm., 27, 2754.

80. Horak, M. and Gut, J. (1961). Coll. Czech. Chem. Comm., 26, 1681.

81. Piskala, A. (1963). Coll. Czech Chem. Comm., 28, 2365.

82. Randall, A.M., Fowler, R.G., Fuson, N. and Dangl, J.R. (1949). "Infrared Determination of Organic Structures", D. van Nostrand, New York.

83. Bellamy, L. J.   "Advances in Infrared Group Frequencies", (1975).   Vol. II, "The Infrared Structure of Complex Molecules", Chapman and Hall, London.

84. Sharma, B.D. and McConnell, J.F. (1965). Acta Cryst., 19, 797.

85. Traynor, J.R. (1973). PhD Thesis, University of Aston.

86. Ostrogovich, A. and Tanislav, I. Gazz. Chim. Ital. (1934). 64, 824.

87. Steward, R.F. and Jensen, L.H. (1967). Acta Cryst., 23, 1102.

88. Rohrer, D.C. and Sundaralingam, M. (1970). Acta Cryst., B 26, 546.

89. 'International Tables for X-Ray Crystallography', (1968). 2nd edn., Kynoch Press, Birmingham, Vol. 3, 276.

90. Singh, P. and Hodgeson, D.J. (1973). J.C.S. Chem. Comm., 439-440.

91. Thewalt, U., Bugg, C.E. and Marsh, R.E. (1971). Acta Cryst., B 27, 2358.

92. Barker, D.L. and Marsh R.E. (1964). Acta Cryst., 17, 1581.

93. Matthews, F.S. and Rich, A. (1964). Nature, 201, 179.

94. MacIntyre, W.M, Singh, P. and Werkema, M. (1965). Biophys. J., 5, 697.

95. Sletten, J., Sletten, E. and Jensen, L.H. (1968). Acta Cryst., B 24, 1692.

96. Hickman, J.A. (1978). Biochemie, 60, 997.

97. Stevens, M.F. G, Gescher, A. and Turnbull, C.P.   (1977).
     Biochem. Pharmacol., 28, 769.

98. Gescher, A, Hickman, J.A, Simmonds, R.J., Stevens, M.F.G.and
     Vaughan, K. (1981). Biochem. Pharmacol., 30, 89.

99. Conius, R.L. (1976). Cancer Treat. Rep., 60, 165.

100. Blum, R.H, Corson, J.M, Wilson, R.E, Greenberger, J.S,
     Canellor, G.P. and Frei, E. (1980). Cancer, 46, 1722.

101. Krikorian, J.G, Portlock, C.S. and Rosenberg, S.A.   (1978).
     Cancer, 41, 2107.

102. Stevens, M.F.G. (1983). "Structure Activity Relationships of
     Antitumour Agents", Editors D.N. Reinhoudt, T.A.Connors, H.M.
     Pinedo and K.W.van de Poll, Martinus Nijhoff Publishers, The
     Hague, pp183-218.

103. Mizuno, N.S, Decker, R.W.and Zakis, B.   (1975).   Biochem.
     Pharmacol., 24, 615.

104. Skibba, J.L, Beal, D.D, Ramirez, G.and Bryan, G.T.   (1970).
     Cancer Res., 30, 147-150.

105. Shealy, Y.F., Montgomery, J.A. and Laster, W.R.   (1962).
     Biochem. Pharmacol., 11, 674.

106. Shealy Y. F., Struck, R. P., Holum, L. B. and Montgomery, J. A. (1961). 26, 2396.

107. Shealy, Y. F. and Kranth, C. A. (1966). Nature, 210, 208.

108. Falkson, G., van der Merwe, A. M. and Falkson, H. C. (1972). Cancer Chemother. Rep., 56, 671.

109. Shealy, Y. F., O'Dell, C. A. and Kranth, C. A. (1975). J. Pharm. Sci., 64, 177.

110. Vogel, C. L., Denham, C., Waalkes, T. P. and De Vita, T. (1970). Cancer Res., 30, 1651.

111. Hansch, C., Smith, N., Engle, R. and Wood, H. (1972). Cancer Chemother. Rep., 56, 443.

112. Connors, T. A., Goddard, P. M., Merai, K., Ross, W. C. J., Wilman, D. (1976). Prog. Med. Chem., 25, 241-246.

113. Stevens, M. F. G. (1976). Prog. Med. Chem., 13, 305.

114. Vaughan, K. and Stevens, M. F. G. (1978). Chem. Soc. Rev., 7, 377.

115. Baig, G. U., Stevens, M. F. G., Stone, R. and Lunt, E. (1982). J. Chem. Soc. Perkin I., 1811.

116. Stevens, M.F.G., Hickman, J.A., Stone, R, Gibson, N.W., Baig, G U., Newton, C.G. and Lunt, E. (1983). J. Medicin. Chem. In press.

117. Stevens, M.F.G., Hickman, J.A., Atassi, G., Gibson, N.W., Stone, R., Lunt, E., Newton, C.G., Langdon, S.P., Lavelle, F. and Fizames, C. (1983). To be published.

118. Unpublished data from the National Cancer Institute.

119. Stevens, M.F.G., Hickman, J.A., Gibson, N.W., Langdon, S.P. and Chubb, D.R. (1983). In preparation.

120. Freeman, H.C. and Hutchinson, N.D. (1979). Acta Cryst., B35, 2051-2054.

121. Edwards, S.L., Sherfinski, J.S. and Marsh, R.E. (1974). J. Amer. Chem. Soc., 96, 2593-2597.

122. Schwalbe, C.H. and Lowe, P. R. (1979). American Crystallographic Association Summer Meeting, Boston, Massachusetts, Abstracts, p. 44.

123. Hjortas, J. (1973). Acta Cryst., B29, 1916.

124. Mullen, D. and Hellner, E. (1978). Acta Cryst., B34, 1624.

125. Tisdale, M. and Horgan, C. (1983). Private communication.

126. Shealy, Y. F., Kranth, C. A., Holum, L. B. and Fitzgibbon, W. E.
     (1968). J. Pharm. Sci., _57_, 83.

127. Abraham, D., Rutherford, J. and Rosenstein, R. (1969). J. Med.
     Chem., _12_, 189.

128. Schwalbe, C. H., Randall, A., Lai, K. K., Simmonds, R. J., Whiston,
     C. J. and Vaughan K. (1983). Joint British Crystallographic
     Association and Royal Society of Chemistry Autumn Meeting,
     Swansea.