AN X-RAY DIFFRACTION STUDY OF THE MOLECULAR MORPHOLOGY AND STRUCTURAL TRANSITIONS OF DNA

by

Victor Trevor Forsyth, B.Sc.(Keele), M.Inst.P.

A thesis submitted to the University of Keele
for the degree of Doctor of Philosophy

Molecular Biophysics Group,
Department of Physics,
University of Keele,
Staffordshire, U.K.

1984

IMAGING SERVICES NORTH
Boston Spa, Wetherby
West Yorkshire, LS23 7BQ
www.bl.uk

Figure 1.2 page 5

Plate 1.1 page 23

Figure 3.10 page 74

Figure 3.11 page 75

Figure 5.1 page 132

Figure 6.6 page 159

Excluded at the request of the university

TO

MY PARENTS

AND

PAUL AND IAN

ACKNOWLEDGEMENTS

I wish firstly to acknowledge the respective contributions of Dr. A. Mahendrasingam, Dr. N.J. Rhodes and Professor W. Fuller to the data that is described and analysed in the following chapters.

I would like to thank my supervisor, Professor Watson Fuller, for his supervision, encouragement and provision of research facilities throught the course of this project. I have also benefited enormously from discussions with an advice from my colleagues Drs. R.J. Greenall, A. Mahendrasingam, W.J. Pigram and N.J. Rhodes over the last few years. I would like also to acknowledge the help of Dr. Colin Nave, SERC Daresbury Laboratory, who has provided invaluable assistance during our experimental work at the SRS. I am also very grateful to Dr. Colm Ward who has on numerous occasions patiently advised and assisted on the implementation of various computer programs.

Mr. F. Rowerth and the technical staff of this department have provided an excellent service. I would like to make special mention of Messrs. G. Dudley and T. Greasley who were responsible fo r the construction of the x-ray camera used at Daresbury. Work undertaken using the Hilger and Elliott generators has been greatly facilitated by Mr. P. Clarke.

I am totally indebted to the stunningly rapid and proficient secretarial assistance I have obtained in the first instance from my mother and more latterly from Miss Helen Martin.

I would like to take this opportunity to thank my parents and family for their unfailing interest and ecnouragement during the course of all my studies.

This work was financially supported by an SERC Research Scholarship.

Finally, I would like to leave a space for Humphrey whose vocal contribution to this group will, no doubt, not pass unremembered.

ABSTRACT

This thesis is introduced with a brief review of the structure and function of nucleic acids in biological systems. DNA and its synthetic analogues have been shown to be capable of adopting a wide range of molecular conformations which differ in pitch, turn angle and handedness. The nature of this 'polymorphism' is found to depend critically on base sequence, hydration, and ionic strength, and as such, may be of great help in visualising the structural basis for processes such as transcription, replication and chromatin condensation.

The technique of x-ray scattering from oriented fibres of DNA/RNA has been instrumental in the determination of the structures of different nucleic acids and in characterising the transitions that occur between their allomorphs. Chapter Two describes the experimental methods that have been used to obtain diffraction data. Chapter Three outlines the physical theory underlying the interpretation of this data.

In Chapter Four the C' form, a close relative of C-DNA, is analysed. The preferred model is a right-handed nine fold double helix having furanose rings that are puckered C3-exo.

The B' conformation, recently heralded as a 'heteronomous' molecule[21], has been observed in poly d(A-T).poly d(A-T) and is analysed in Chapter Five.

In Chapter Six, right and left-handed models for the β-D conformation of poly d(A-T).poly d(A-T) are compared.

The effect of proflavine, an acridine drug, on the transitions which normally occur in a variety of polynucleotides is examined in Chapter Seven.

Chapter Eight describes a number of preliminary experiments in which the intense x-ray beam provided by the SERC Daresbury Synchrotron Radiation Source (SRS) has been used to undertake the first real-time studies of transitions within DNA fibres.

Chapter Nine concludes with a short summary of the state of nucleic acid crystallography to date and of the methods which are likely to be most productive in the near future.

CONTENTS

CHAPTER ONE

INTRODUCTION

1.1 THE BIOLOGICAL STRUCTURE AND FUNCTION OF NUCLEIC ACIDS

1.1.1 Nucleic Acids

Nucleic acids are nucleotide polymers. The repeating nucleotide unit consists of a base, either a purine or a pyrimidine, a pentose sugar and a phosphate. Nucleic acids fall broadly into two groups: deoxyribonucleic acids (DNA) in which the sugar residue is deoxyribose, and ribosynucleic acids (RNA) in which the sugar is ribose.

1.1.2 The DNA Molecule

The double helical DNA molecule is a stable 'helical-ladder' like structure consisting of two polynucleotide chains held together by hydrogen bonds between paired bases. The DNA molecule is seen as having either a right or a left-handed twist, usually forming a double-helix which is characterised by a pitch of $\sim$ 28Å - 43Å. However, it can be single, double, triple or four-stranded (Chandrasekaran and Leslie, 1976b; Arnott and Bond, 1973a,b; Arnott et al., 1974; Arnott et al., 1976a). The diameter of the molecule is typically in the region of 20Å. Each turn of the helix has anything from eight to twelve base-pairs strutting the spirals (see Figure 1.1).

DNA contains, in the main, four different types of base, two of which are classed as pyrimidines and two as purines. The two purines are adenine (A) and guanine (G) and the two pyrimidines are cytosine (C) and thymine (T). In terms of the double helical model for DNA, the symmetry of

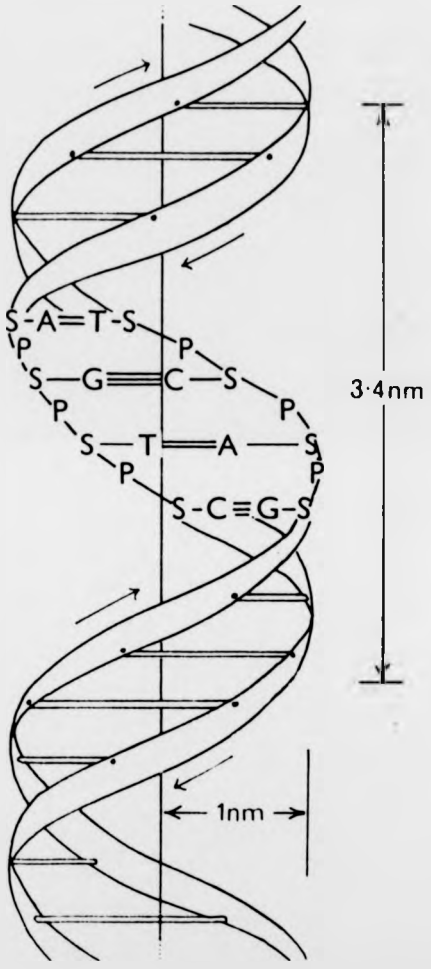


FIGURE 1.1 Diagramatic representation of the double-helical structure of B-DNA

the molecule and the nature of the hydrogen bonding between the bases means that each base pair consists of a purine and a pyrimidine.

In the polynucleotide chain itself, adjacent sugars are linked through their 3' and 5' carbons by a phosphodiester linkage. If the sugars in one strand are linked 3'-5', the complementary strand is related by diad axes perpendicular to the helix axis so that the chains of double helical DNA can be termed 'anti-parallel'.

1.1.3 Three Different Types of RNA

RNA differs from DNA in a number of important aspects: whereas in DNA there are two hydrogens attached to the 2' carbon of the sugar residue, in RNA the carbon carries a hydrogen and hydroxyl group. A second important difference is that throughout the RNA molecule the base uracil (U) is found in place of the thymine associated with DNA.

In most cells, RNA is generally believed to be less stable and not as 'long-lived' as DNA. The relative stability of DNA, and in particular of the base sequences, is perhaps to be expected given that accurate mitosis and protein production occur throughout the lifetime of most organisms.

The conformations of RNA show greater variability than those of DNA, and a number of widely different models have been suggested - double helices, a single helix folded back and in regions complementary to itself. The three principal forms of RNA found in cells are (a) messenger RNA (m-RNA), (b) ribosomal RNA (r-RNA), and (c) transfer RNA (t-RNA):

(a) Messenger RNA is mostly single-stranded and serves to convey coded information from the DNA in the nucleus to the cytoplasm where proteins are then synthesised. It is known that m-RNA is complementary to one strand of a section of a DNA molecule.

(b) Ribosomal RNA forms the bulk of cellular RNA, especially in cells that are undergoing rapid protein synthesis. In contrast to m-RNA, r-RNA seems

to have a relatively non-specific base sequence and in the case of E Coli 5S RNA is ∿ 60% double-helical (Erdmann, 1979). It is also more stable than m-RNA. Complexed with protein, r-RNA forms the cellular ribosomes, which are the sites of protein synthesis in the cytoplasm. It is thought that the intranuclear nucleolus (which is rich in r-RNA) is the site of r-RNA synthesis.

(c) Transfer-RNA molecules are relatively small, consisting of about 70-90 nucleotides and having molecular weights of around 25000. Their function in the cell is to collect free amino acids in the cytoplasm and selectively attach them to the end of a growing polypeptide. Their ultimate role is thus to accurately translate the base sequence carried by m-RNA, via a 'coding system' into a corresponding sequence of amino acids, hence eventually constructing a cellular protein. Transfer RNA contains bases not found in other nucleic acids (e.g. methylated pyrimidines) and it is possible that the sequences are deliberately (excuse the anthropomorphism) made non-complementary in places. The clover leaf model of Holley (1968) was based on sequence studies of different yeast t-RNA's and is shown in Figure 1.2. Single crystals of yeast t-RNA have been grown and x-ray diffraction data obtained therefrom have confirmed the clover leaf model and furthermore elucidated the tertiary structure of yeast phenylalanine t-RNA (Kim et al., 1974; Robertus et al., 1974; Ladner et al., 1975).

As stated previously, t-RNA functions to collect amino acids from the cytoplasm and 'tack' them together in a sequence determined by the message encoded in the template m-RNA. To this end there is a region, the 'anti-codon', of the t-RNA molecule which is capable of complementary bonding to a suitable part of the m-RNA (known as the 'codon'). A non-specific terminal triplet (cytosine-cytosine-adenine-OH) forms a point of attachment of the amino acid (Rich and Raj Bhandary, 1976).

1.1.4 Nucleic Acid Synthesis and Replication

Nucleic acids are synthesised by an energy consuming cellular

FIGURE 1.2 The complete nucleotide sequence of alanine t-RNA showing the unusual bases and codon/anticodon position (from Watson, 1976)

process involving the polymerisation of nucleotides. Adenosine triphosphate (ATP) causes phosphorylation of the nucleotides and this provides enough energy for the formation of the polymerising phosphodiester linkage.

After fertilization, one cell (the zygote) is ultimately responsible for the subsequent morphogenesis of the whole organism. The zygote undergoes successive mitotic divisions until all the tissues and organs pertaining to a given individual are formed. Each cell in the individual (with the exception of meiotically divided sex cells) has an identical genetic complement, and it is thus clear that DNA molecules must duplicate themselves with a very high degree of accuracy. The proposal by Watson and Crick (1953) that DNA was double-helical immediately suggested (barring a few topological problems) a mechanism for replication : if the molecule were to unwind, breaking at the interbase hydrogen bonds, each strand could then serve as a template for the construction of an identical DNA duplex (see Figure 1.3). Further evidence for this hypothesis was provided by Meselson and Stahl (1958) who performed a biochemical assay using a heavy nitrogen isotope to show that such replication did indeed occur.

1.1.5 The Cistron, the Genetic Code, Transcription, Translation and Polypeptides

In adherence with the 'one gene, one enzyme' hypothesis, one gene, or perhaps more definitively, one 'cistron', codes for the production of one polypeptide chain in the cytoplasm. A cistron is therefore a section of a DNA molecule having a specific base sequence, and which serves as a template for the production of m-RNA, which, in turn, then determines a sequence of amino acids. m-RNA thus transcribed leaves the nucleus through 'nuclear pores', and becomes attached to ribosomes (which mostly lie on the internal aspect of the endoplasmic reticulum).

The base sequence determines the amino acid sequence of a

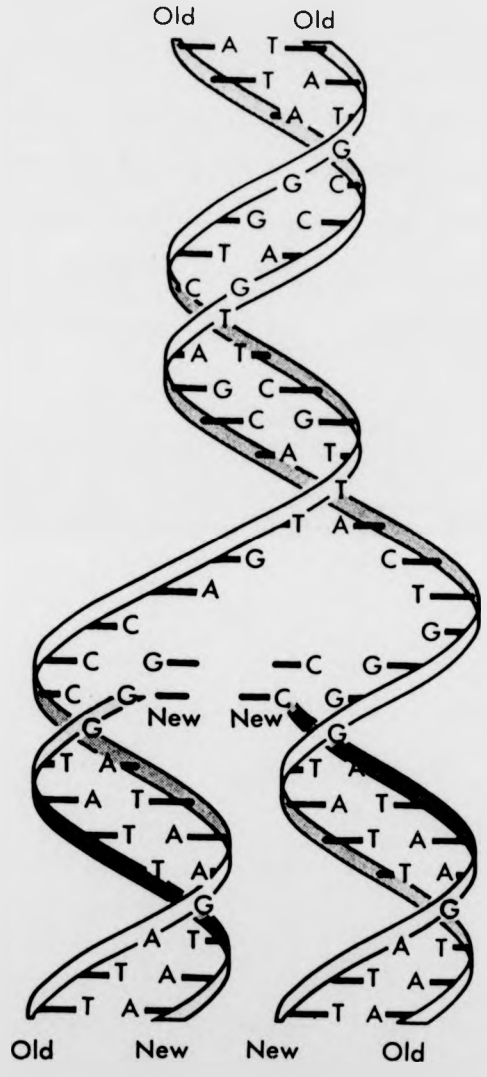


FIGURE 1.3 The Watson-Crick scheme for DNA replication

polypeptide by means of a 'triplet code' in which three consecutive bases code for a particular amino acid (e.g. UUU codes for phenylalanine). The basis of this genetic code was partially predicted on theoretical grounds, before the 'alphabet' had been determined. The number of amino acids (twenty) compared with the number of nucleotides (four) implied that a triplet code was necessary: only a combination of three nucleotides could yield more than twenty different amino acid codewords. By this hypothesis, the total possible number of different trinucleotides is sixty-four (see Table 1.1).

SECOND LETTER

FIRST LETTER	U	C	A	G	THIRD LETTER
U	UUU, UUC — Phe UUA, UUG — Leu	UCU, UCC, UCA, UCG — Ser	UAU, UAC — Tyr UAA, UAG — Ter	UGU, UGC — Cys UGA — Ter UGG — Trp	U C A G
C	CUU, CUC, CUA, CUG — Leu	CCU, CCC, CCA, CCG — Pro	CAU, CAC — His CAA, CAG — Gln	CGU, CGC, CGA, CGG — Arg	U C A G
A	AUU, AUC — Ile AUA — Met AUG	ACU, ACC, ACA, ACG — Thr	AAU, AAC — Asn AAA, AAG — Lys	AGU, AGC — Ser AGA, AGG — Arg	U C A G
G	GUU, GUC, GUA, GUG — Val	GCU, GCC, GCA, GCG — Ala	GAU, GAC — Asp GAA, GAG — Glu	GGU, GGC, GGA, GGG — Gly	U C A G

TABLE 1.1 The Genetic Code

The fact that the figure of sixty-four is way in excess of the necessary twenty, indicates that either (a) the code is degenerate and that several codons are used to produce the same amino acid, or (b) there exist a number of 'nonsense' codons - codewords which do not correspond to any

amino acid and are not used in the message. The two different possibilities (a) and (b), are significant in terms of the structure of m-RNA: a degenerate code would allow more flexibility since different primary structures could still carry the same information. The determination of the genetic code, and the analysis of nucleotide sequences has now clearly established the degeneracy of the code, and the expected flexibility in the m-RNA structure has also become apparent. There are three 'nonsense' codons, and it is believed that all of them serve as termination signals in protein synthesis.

Another feature of the genetic code is that the nucleotide sequence can be interpreted in any of three different reading frames, each giving totally different amino acid sequences. The genetic code is not 'punctuated'. It has been established experimentally (Barrell et al. (1976); Smith et al. (1977); Shaw et al (1978)) in a study of the bacteriophage ØX 174 that the same nucleic acid sequence can comprise two different messages, and can direct the synthesis of two different polypeptide chains which are encoded in the same sequence but in different reading frames. See Figure 1.4.

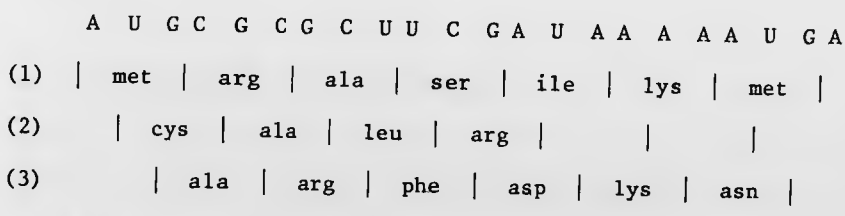


FIGURE 1.4 Decoding of a message in three different reading frames

It is thus apparent that for faithful decoding it is vital that the correct reading frame is recognised. It is thought possible that two codons can define initiation sites: AUG and GUG (which is less frequent). However, it is noteworthy that both these triplets can code for amino acids at internal positions (AUG for methionine, GUG for valine) and, as such,

comprise the only ambiguities in the genetic code.

An m-RNA molecules, once adsorbed into a ribosome, is considered to be an active site of protein synthesis. One m-RNA molecule does not, however, act as a protein template indefinitely, and may only function in the construction of about ten to twenty proteins before stopping.

The mechanism by which t-RNA collects amino acids from the cytoplasm and adds them, in sequence, to the polypeptide chain, is aided by GTP which is required to form the peptide bond between the 'enzyme activiated' amino acid and the polypeptide chain. The t-RNA/amino acid complex is then slotted into position at the ribosome when the codon of the m-RNA strand is complementary to the anticodon of t-RNA. A peptide bond is formed between the amino acid and the polypeptide chain. Upon completion the chain folds up into a globular form having a characteristic secondary and tertiary configuration. To some extent the tertiary structure of a protein is inevitably give certain primary and secondary constraints, but it is possible that certain 'folding enzymes' may be present to ensure correct tertiary structure.

A very simplified summary of nucleic acid pathways in protein synthesis is shown in Figure 1.5.

1. 1.6 Nucleoprotein and Chromatin

As previously mentioned, the nature of nucleic acid structure makes an enormous amount of conformational variability. It is possible that the extensive polymorphism of DNA is involved in the dramatic behaviour of chromatin during the life of the mitotic cell. The DNA of eukaryote cells *in vivo* exists in complex with various histone and non-histone proteins. The nucleo protein so formed, is, in the living cell, capable of great morphological variability throughout the cycle of the cell: this is clearly realised when one considers the extent of condensation that occurs within the eukaryote nucleus. The precise structure and functioning of chromatin (particularly in terms of replication, transcription and condensation) has yet to be fully understood, but work undertaken during the 1960's and 1970's

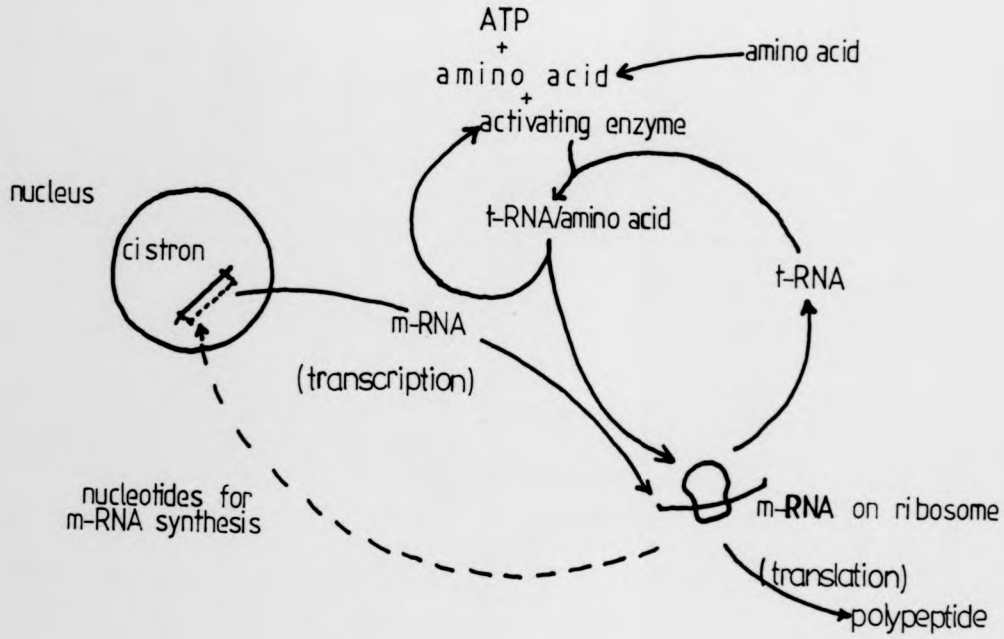


FIGURE 1.5 Summary of Nucleic Acid Pathways in Protein Synthesis

has established that the first level of organisation in the 'contraction' of DNA is the 'nucleosome structure'. The structure consists of a core particle of one hundred and forty base pairs of DNA wound around an octomer core of histones comprising two each of the histones called H2A, H2B, H3 and H4. These core particles are joined by lengths of 'linker' DNA. Another protein, H1, is thought to be associated with the linker region (Butler et al., 1969; Kornberg and Thomas, 1974; Olins and Olins, 1974). See Figure 1.6.

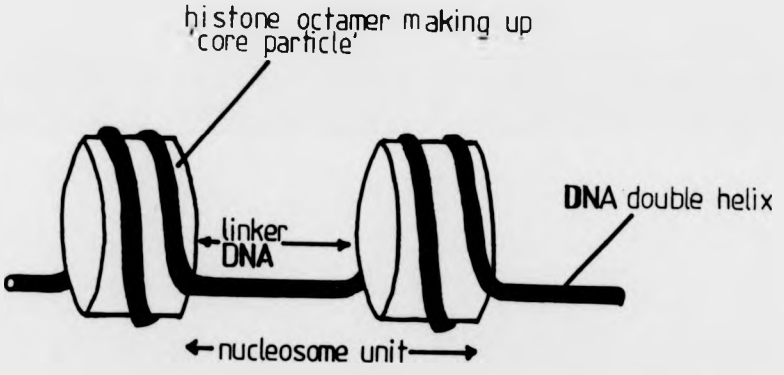


Fig 1.6: Schematic diagram illustrating the nucleosome structure of chromatin

Although the chemistry of chromatin contraction is not fully understood, various models have been put forward (Finch and Klug, 1976; Bak et al., 1977) suggesting that the nucleosome threads coil into a helix having a pitch of $\sim$ 110Å and an outer diameter of 300Å. This helix then further supercoils into a chromatid which is a hollow cylinder of $\sim$ 4000Å diameter. There is, as yet, no further detail at the molecular level for the construction of higher ordered structures.

It is interesting that the chromosomes themselves are capable of reversible condensation and that different stages of condensation can

be identified by the presence of a number of different liquid crystal phases ranging from isotropic liquid to mesophase. Chambers (1924) has noted that when the nuclear membrane of the living grasshopper spermotocyte is pricked during prophase, chromosomal filaments appear which gradually thicken until they look like those of a normal metaphase stage. This observation implies that certain 'concentrations' in the nucleus must be modified by injury to the nuclear membrane, and probably by cytoplasmic contents rushing into the nucleus. It is remarkable that the natural condensation of chromosomes is preceeded by the disappearance of the nuclear envelope.

The implications of the above observations are, as yet, difficult to estimate. Bearing in mind that chromosomes are condensed phases of DNA - protein complexes and other chemical species, it is likely that certain aspects of polymer crystallography (at both high and low resolution) ought to yield information on the 'modus operandi' of the DNA molecule itself and the methods by which it manages to pack into different regular liquid crystal phases. The effect of ions and other groups on the structure of chromatin is of great interest here, as is their effect on the conformation of 'naked' DNA's.

1.2 THE STUDY OF NUCLEOTIDE POLYMERS

A great number of different methods have been employed in attempts to clarify the structure of nucleic acids, various nucleoprotein complexes, and chromatin. This thesis is primarily devoted to the study of various types of polynucleotide and polynucleotide-drug complex. Whilst it might seem that the most realistic method of studying DNA is in the examination of the liquid state, a lot more information is yielded by the study of polymer fibres. However, the extent to which the semi-crystalline or crystalline state may produce artefactual information is obviously very important and fibre studies must therefore be undertaken with some caution.

Single crystal X-ray studies allow structure determination to atomic resolution, but must be treated with even greater reservation than fibre studies, since the single crystal examinations so far have been confined to short nucleotide repeats and 'end-effects' are likely to be large.

1.2.1 Solution Studies

Studies of different types of DNA solutions have provided a wealth of information which has added to and corroborated evidence obtained elsewhere. Optical rotary dispersion (ORD) was used to show the existence of structural transitions (as a function of salt concentration) in DNA solutions (Cheng, 1965; Tunis and Hearst, 1968) and later investigations on the effects of various neutral electrolytes on their circular dichroic (CD) spectral properties have supported this conclusion. Tunis-Schneider and Maestre (1970) have identified the CD spectra appropriate for the A, B and C forms of DNA, and in 1972 Erfurth et al. published results which showed that the different forms could be distinguished by an examination of several Raman bands that arise from the vibration of the sugar phosphate backbone of nucleic acid polymers. Wide angle X-ray scattering has also been applied to investigate molecular structure in fairly concentrated solutions. Bram (1976a)has studied the secondary structure of DNA in solution and in nucleohistone by this method and concluded that DNA in such conditions adopts a 'B' type structure.

Later work has revealed the very interesting behaviour of poly (dG-dC).poly (dG-dC) as a function of salt concentration (Pohl and Jovin, 1972). This behaviour has since been correlated with the structure of the oligomer d(CpGpCpGpCpG) as described by Wang et al. (1979) after Laser-Raman spectroscopic studies indicated that the high salt form of the aqueous polymer has the same structure as left-handed Z-DNA in the crystalline oligomer (Thamann et al., 1981).

It is well recognised in this field that although correspondence between solution data and solid state data is of great importance, conformations are likely to differ structurally in the solution, fibre, and single crystal states (Bram, 1976b; Hanlon et al., 1975). The details of such differences are of great interest.

1.2.2 Fibre Diffraction Studies

As a result of the presence of various types of disorder in polymer fibres (these will be discussed in section 3.5), the data obtained from fibre diffraction provides a resolution ($\sim$ 3Å) which does not enable the unambiguous determination of atomic positions. Whilst resolution is low, certain features of the helical conformation can readily be ascertained (e.g. pitch of helix, number of residues per turn) if the specimen is reasonable ordered. It is a combination of these parameters, the observed diffraction and a relatively sound knowledge of the overall stereochemistry of the molecule that enables a structural refinement to give an acceptable 'best fit' with the observed data. Structure determination by this method is still, however, by no means totally definitive and for a given diffraction pattern, models having both helical senses can often be generated to give equally acceptable agreement with the observed diffraction. This is rather a harrowing predicament, since in recent years, the handedness of the DNA helix has emerged as a rather crucial problem, especially in the light of certain dynamic transitions that are found to occur in fibres and in solution as a function of ionic strength and hydration.

1.2.3 Single Crystal X-ray Studies

The highest level of solid state order is a single crystal. The structure of such crystals can usually be solved to atomic resolution but the main difficulty with respect to nucleic acid polymer research is the fact that the oligomers used have to be relatively short to enable the acquisition of high quality crystalline data. This restraint is a significant

one since, as mentioned in Section 1.2.1, the shorter the polymer, the greater the risk of 'end-effects' distorting the true conformational nature of the helix and producing misleading information that cannot be sensibly correlated with fibre and solution studies.

The most interesting oligomers that have been studied to date are those containing G-C base pairs. The oligomer d(CpGpCpGpCpG) (or $d(CG)_3$) has been found to crystallise into a novel left-handed 'Z' form (Wang et al., 1979) having Watson-Crick base-pairs and alternating C2-endo and C3-endo sugar puckers. Further work on this oligomer by Wang et al. (1981) has outlined the presence of two Z helices, nominated Z_I and Z_{II}. Another structure that has been solved from a single crystal examination of $d(CG)_2$ has been called Z'-DNA (Drew, Dickerson, Itakura, 1978). The Z'form is very similar to the Z helices described by Wang et al. (1979) and Wang et al. (1981). Crawford et al. (1980) have also studied crystals of $d(CG)_2$ and have published similar Z structures. Crawford et al. (1980) maintain that the small differences that exist between the Z molecules thus far solved can be attributed to the different ions present in each of the crystals. It therefore seems likely, as suggested by Neidle and Berman (1983), that the Z_I, Z_{II} and Z' forms are just three of a whole range of Z structures that can be obtained, and that there is what could be called 'restricted' polymorphism in single crystals. Fibre studies have since (Arnott et al., 1980) elucidated the structure of high salt poly d(G-C).poly d(G-C) and have interestingly enough shown that a model having the overall features of Z-DNA accounts well for the diffraction obtained from such fibres. The idealised and continuous 6_5 helix has been called S-DNA. Other studies have been undertaken which verify the presence of a B type helix in the duodecamer d(CGCGAATTCGCG) (Wang et al. 1980; Dickerson and Drew, 1981a). The structure has 10.1 residues per turn, Watson-Crick base pairs, and an average axial rise per base pair of 3.4Å. As might be expected, all the conformational parameters vary slightly from one residue to the next.

The value of single crystal work is clear: it has been correlated with previous fibre and solution studies and has opened up new areas of immense interest. The importance of left-handed DNA is apparent, especially in relation to transitions and change of helix sense. Single crystal work has also emphasised the significance of sequence dependent local molecular variation.

1.3 THE STEREOCHEMICAL DESCRIPTION OF POLYNUCLEOTIDE HELICES

A schematic outline of the double helical chain is shown in Figure 1.7 and it is immediately evident that there are six torsion angles capable of variation (5 along the backbone and one about the glycosidic bond that links base and sugar). The resolution enabled by fibre diffraction justifies assuming standard and constant covalent bond lengths and angles.

The torsion angles shown in Figure 1.7 are broadly classified (Arnott and Hukins, 1972a, b) so that the 'cis' position is defined as zero, and a positive torsion angle occurs when, looking along the bond length, the nearest atom is reached by an anticlockwise rotation. The backbone torsion angles are said to be 'gauche$^+$' (g+) if $0 < \tau < 120^o$, 'gauche$^-$' (g^-) if $-120^o < \tau < 0^o$, and 'trans' (t) if $120^o < \tau < 240^o$ (see Figure 1.8). The glycosidic torsion angle, χ, which characterises the link between sugar and base is defined by the atoms C2', C1', N and C4 if the base is a purine, and by C2', C1', N and C2 if the base is a pyrimidine. Two classes of this conformation angle are commonly observed - anti when $\chi \sim 90^o$, and syn and $\chi \sim 300^o$.

Another feature relating to the flexibility of the chain is the fact that the sugar residues are not necessarily flat. The furanose rings of nucleic acids are either β-D-ribose or β-D-erythropentose (deoxyribose). Although early molecular model building studies (Crick and Watson, 1954) assumed a planar sugar ring, Spencer (1959) asserted that in all likelihood

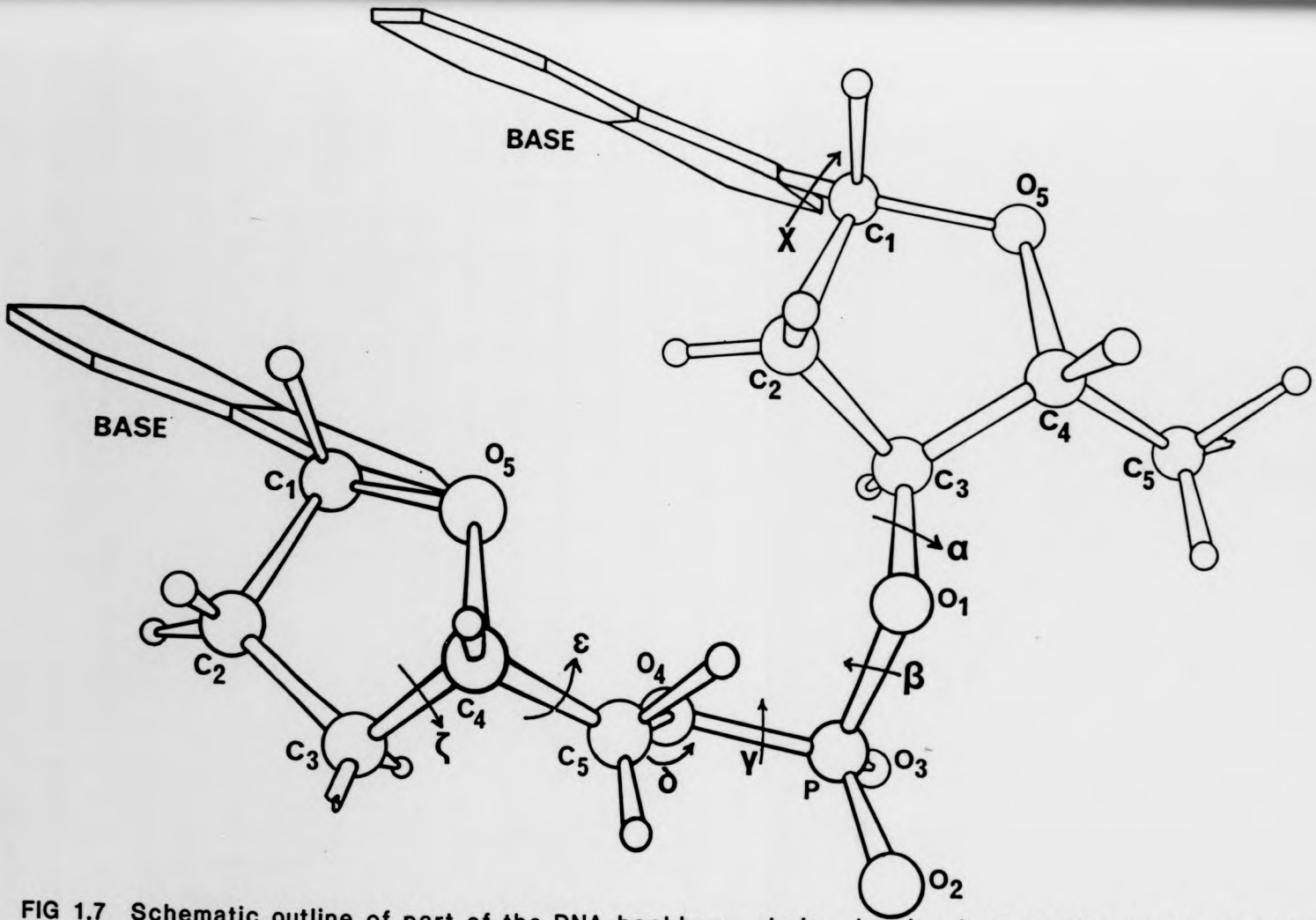


FIG 1.7 Schematic outline of part of the DNA backbone chain, showing how torsion angles are used to describe a mononucleotide repeat.

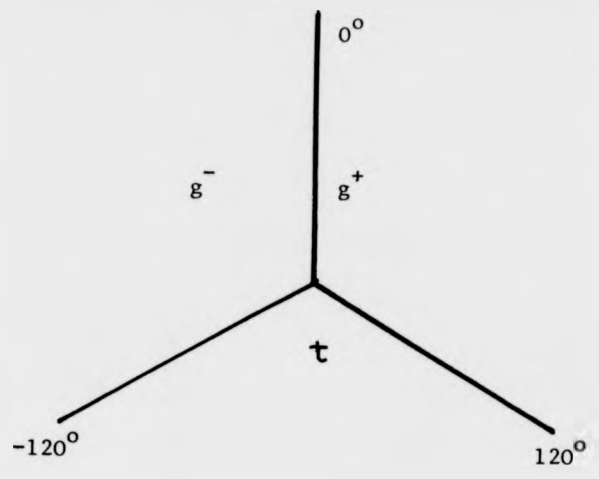


FIGURE 1.8 The convention used in evaluating and describing torsion angles

steric interference between hydrogens of neighbouring carbons would cause a puckering of the rings. A variety of furanose puckering arrangements have been described by Sundralingam and Jensen (1965), and are defined by the offsets of the C2 and C3 atoms from a plane formed by the C1, O5 and C4 atoms. Puckering conformations are then classified in terms of which of the two atoms, C2 and C3, is most heavily offset from the defined plane, and on which side of the plane this offset occurs. If it is on the same side as the C5 atom, the pucker is denoted 'endo', and if it is on the opposite side the pucker is correspondingly called 'exo'. Figure 1.9 shows a projection (edge on the plane C1-O5-C4) of four different puckering arrangements.

Although the overall structure of a nucleotide is unambiguously described by coordinates, it is often found convenient and instructive to depict a molecule in terms of its various torsion angles, the sugar pucker(s), and base position. A summary of these parameters for each structure of relevance to this work is given in the appendix. The base atoms are assumed to lie in one plane, and base position is established by expressing the magnitude of the distance D(the base displacement from the helix axis), a twist angle, and a tilt angle. The two types of base pair commonly

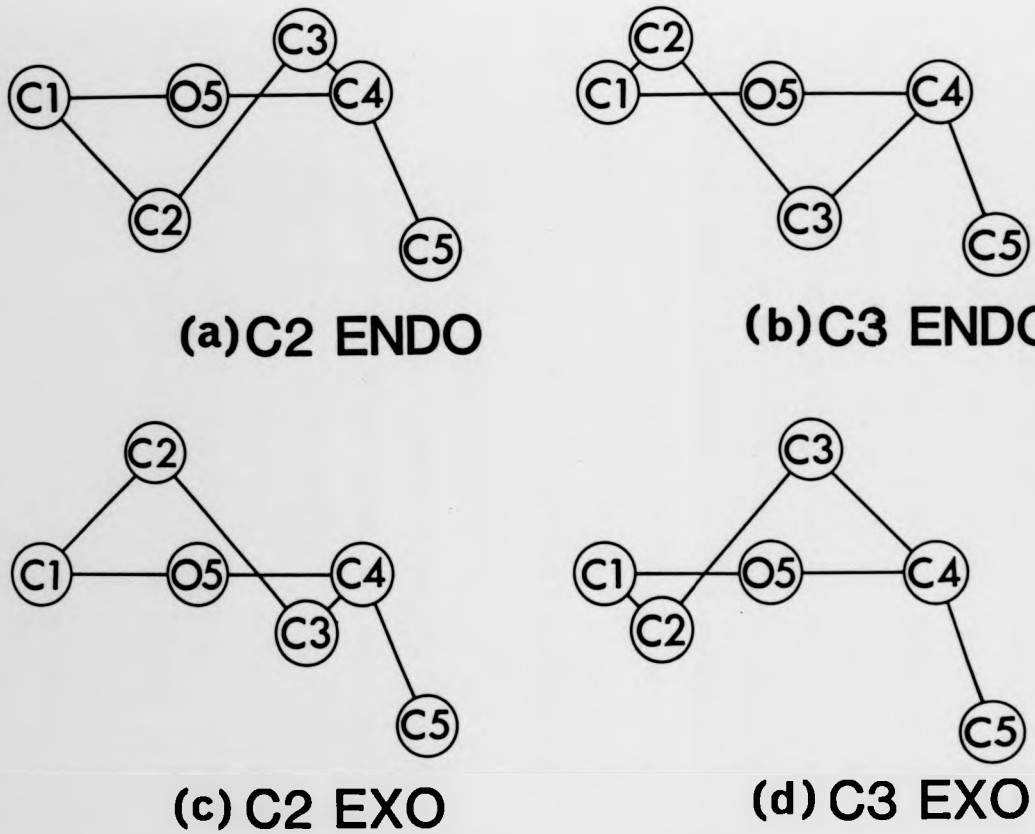


FIG 1.9 Projections(edge on the plane C1–O5–C4) of four different ring puckering arrangements.

found in DNA are shown in Figure 1.9(a) and (b) and Figure 1.9(c) shows the way in which base position is described.

1.4 A SUMMARY OF THE DIFFERENT MORPHOLOGICAL STRAINS OF DNA AND RNA AS OBSERVED IN X-RAY DIFFRACTION

Early diffraction experiments on fibres made from a variety of natural DNA's established the existence of three distinct, though related, structures (the A, B and C conformations). Since then, a large variety of synthetic polynucleotides (having known base sequence) have become available and the observed diffraction from such materials has forced the classification of nucleic acid helices to be broadened and extended. Leslie et al. (1980) have surveyed the polymorphism that has been observed in a wide range of polynucleotides. Minor variations in structure from one type of polynucleotide to the next have necessitated the classification of nucleic acid polymer conformations into families.

1.4.1 The 'A' Set of Structures

The A family is of particular interest because it is found to be the only conformation adopted by RNA, and, as suggested by Arnott (1968) may be the structure used by DNA during the process of transcription.

The A-form of DNA gives a crystalline diffraction pattern (see Plate 1.1) and was first obtained from the sodium salt at low relative humidities and low excess salt (Franklin and Gosling, 1953). A-DNA patterns have been obtained from salts of many different ions, and with the exception of lithium, there appears to be no salt in which DNA is prevented from occuring in the A-form.

The first detailed analysis of A-DNA was undertaken by Fuller et al. (1965). In this work the helix was asserted to have a right-handed helical sense with 11_1 symmetry and a pitch of 28.2Å. The sugar rings have C3-*endo* puckering, and the bases have a tilt of 20°, a twist of -8°

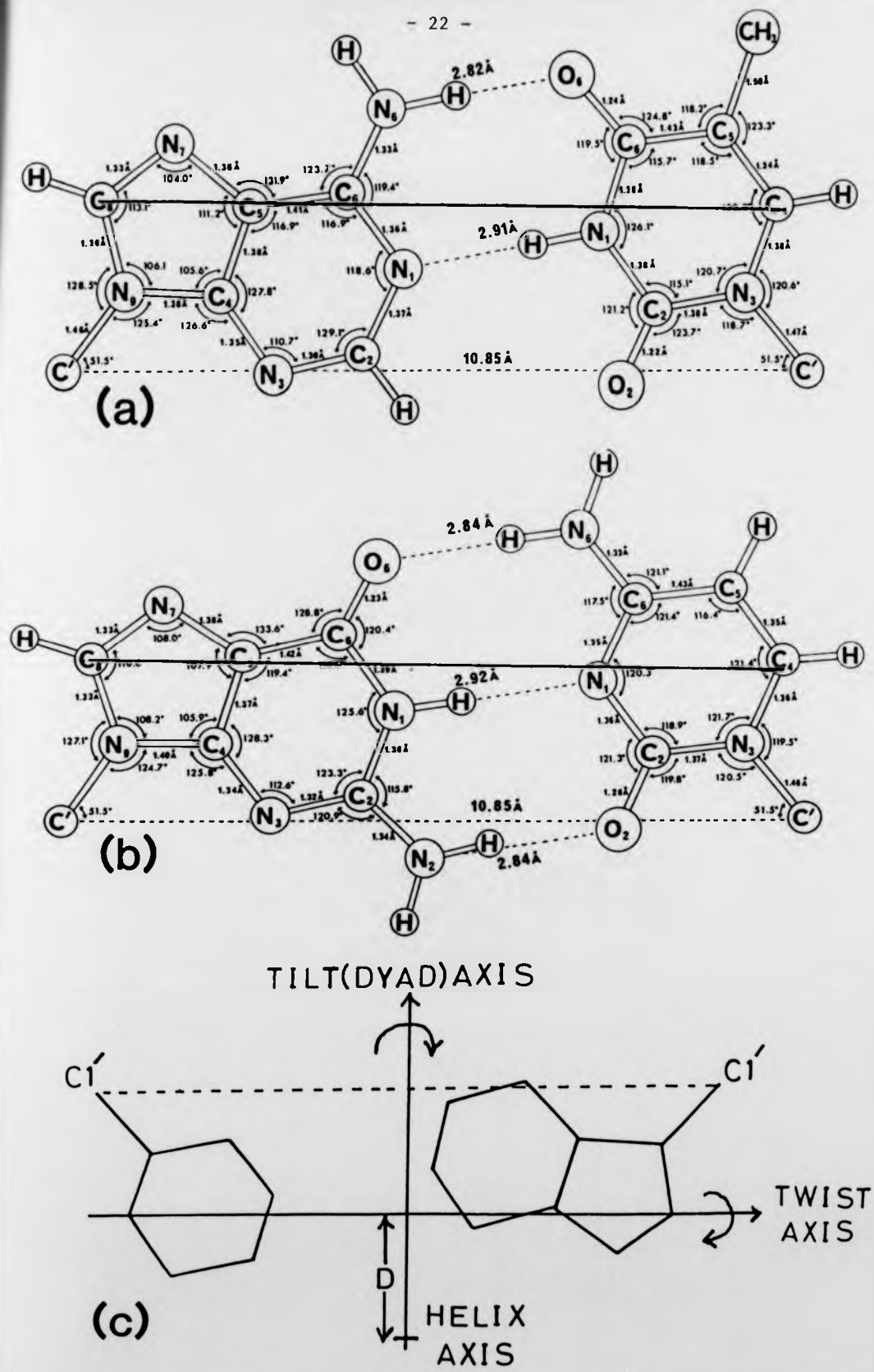


FIG 1.10 Scale drawings of atomic positions in (a) an A–T base pair (b) a G–C base pair and (c) sketch showing the way in which base position is described.

PLATE 1.1 X-ray Diffraction Photograph of A-DNA from Calf Thymus (from Fuller et al., 1965)

PLATE 1.1 X-ray Diffraction Photograph of A-DNA from Calf Thymus (from Fuller et al., 1965)

and a base displacement distance, D, of 4.25Å.

The Fuller model has since been modified twice to incorporate standard stereochemistry (Arnott and Hukins, 1972b; Arnott et al., 1980), and the overall features of the refined models are shown in the Appendix.

It has been suggested that the A conformation of DNA could be a structural peculiarity associated with the crystallisation of DNA into a monoclinic lattice. This idea has been rejected by Leslie et al., (1980) who maintain that the A structure is preserved even in fibres which crystallise with different types of lattice.

A-RNA is essentially similar to A-DNA (see Appendix), having a pitch of 30.9Å, a positive base displacement of 4Å from the axis, a tilt of 16° and a twist of -7°. The analyses of the diffraction patterns obtained from a number of different types of RNA shows that the molecules can pack into two different hexagonal lattices, termed α and β (Langridge and Gomatos (1963), Tomita and Rich (1964), Arnott et al. (1966)).

Another form of RNA is A'-RNA (diffraction pattern is shown in Plate 1.2b as compared with A-RNA in Plate 1.2a). The diffraction from this conformation was observed by Arnott et al. (1968) in fibres of poly (I).poly (C) and poly (A).poly (U). The helical structure shows 12_1 symmetry and the pitch of 36.2Å is slightly higher than that of A-RNA, but otherwise the conformation angles of A'-RNA are similar to those of A-RNA and A-DNA (see Appendix). A third RNA structure has been nominated A"-RNA and was observed by Arnott et al . (1968) in low salt fibres made from poly (A-U), poly (G-C) and poly (I-C).

The fact that m-RNA is associated with the process of transcription from the 'master code' of DNA makes the examination of DNA-RNA hybrids of special interest. Milman et al. (1967) undertook such a study and concluded (without detailed elucidation of the structure) that the hybrid was in basis similar to A-DNA. O'Brien and MacEwan (1970) discovered that

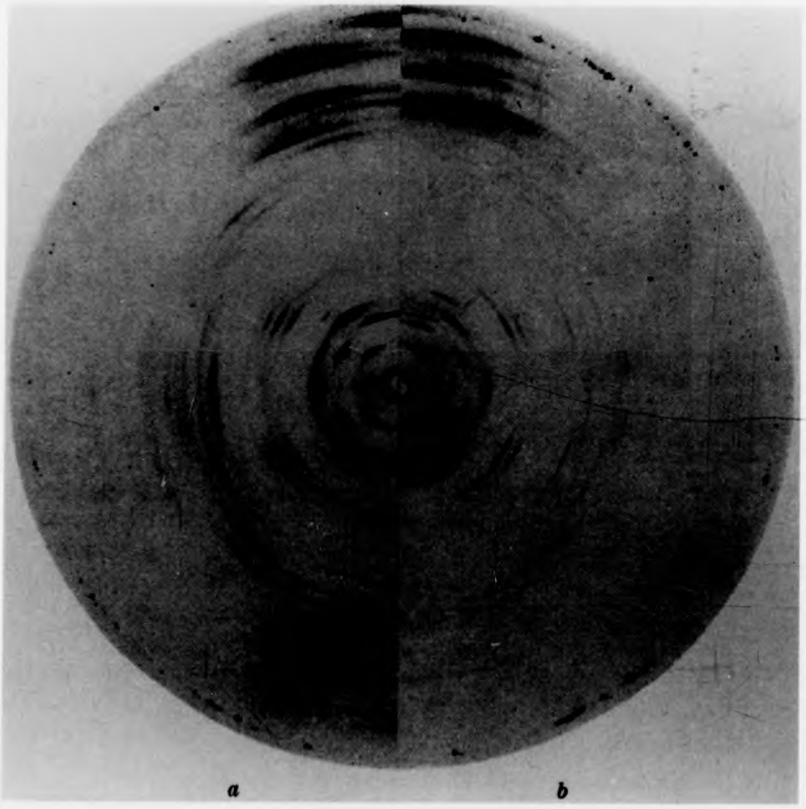


PLATE 1.2 X-ray diffraction photographs comparing the A-form of poly (I+C) in 1.2a with the twelve-fold A' conformation in 1.2b (taken from Arnott et al., 1968).

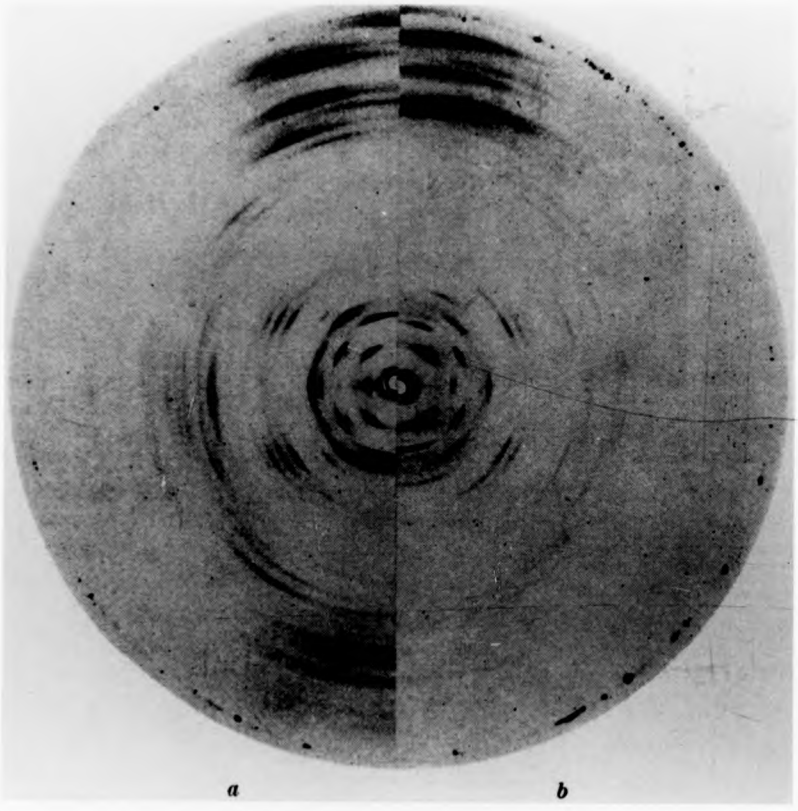


PLATE 1.2 X-ray diffraction photographs comparing the A-form of poly (I+C) in 1.2a with the twelve-fold A' conformation in 1.2b (taken from Arnott et al., 1968).

the hybrid poly rI.poly dC is similar to A'-RNA.

Another form of RNA (A*-RNA) has been presented by Chandrasekaran et al. (1980) and Arnott (1980) from fibres of poly dI. poly rC. The A*-RNA structure is novel in that the torsion angles fall into a different class (see Appendix). The conformation has 10_1 symmetry and the sugar rings have C3'-endo puckering. In other respects, such as base displacement and orientation the structure is similar to that of A-DNA.

1.4.2 The 'B' Set of Structures

The B-conformation of DNA is also a widely occuring structure, and is found to prevail when relative humidity is high or ionic concentration is high. It is of special interest because the structure has been found to exist in cells and in nucleosome structures (Wilkins and Randall (1953); Finch et al. (1981); Bently, Finch, Lewitt-Bently (1981)). The B-forms first obtained were those from the sodium and lithium salts. Sodium DNA was found to give a semi-crystalline diffraction pattern (Plate 1.3) whereas the lithium salt gave a highly crystalline pattern (Plate 1.4). Whilst the semi-crystalline pattern does not yield as much information as the crystalline one, it is generally accepted that the molecular conformation is the same in both cases.

The initial analysis of DNA was based on a semi-crystalline pattern similar to Plate 1.3, and led Watson and Crick (1953) to their now well established double-helical model for DNA. The Watson-Crick model for B-DNA was, however, far from an exact solution and was refined by Langridge et al. (1960a, b) using the more extensive data obtained from crystalline patterns (see Plate 1.4).

A further refinement of B-DNA was published by Arnott and Hukins (1972b, 1973) in which the Langridge model was adjusted to incorporate standard covalent stereochemistry.

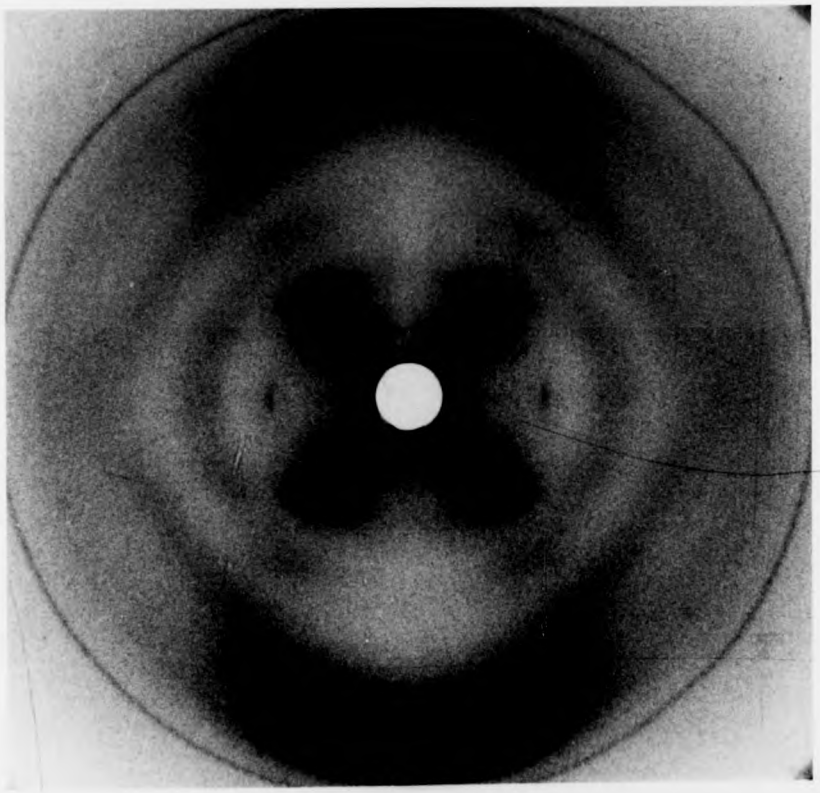


PLATE 1.3 The Semicrystalline B-form obtained from Na DNA

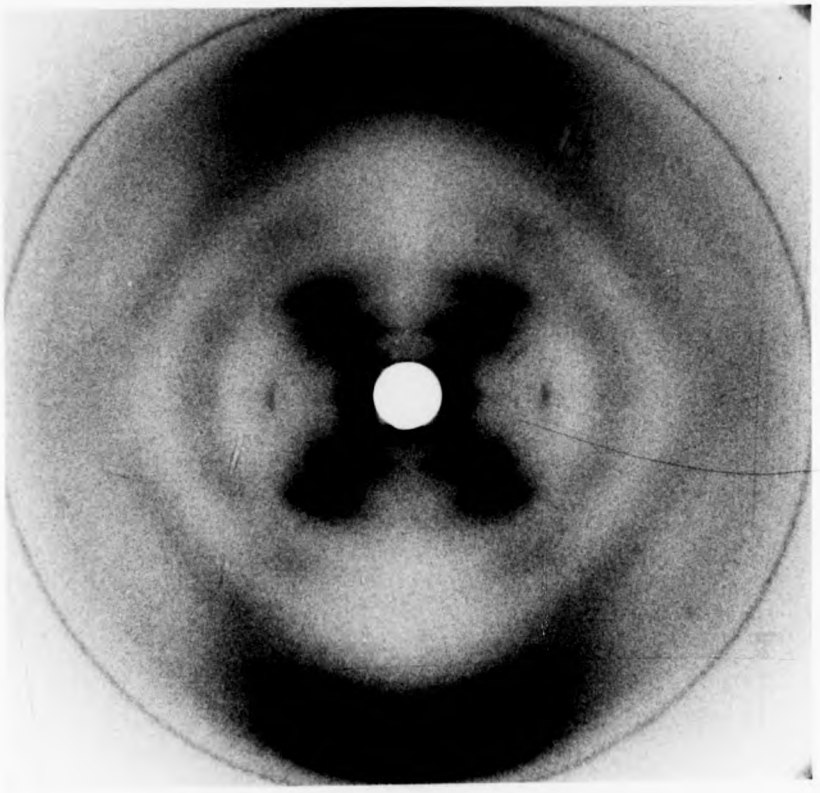


PLATE 1.3 The Semicrystalline B-form obtained from Na DNA

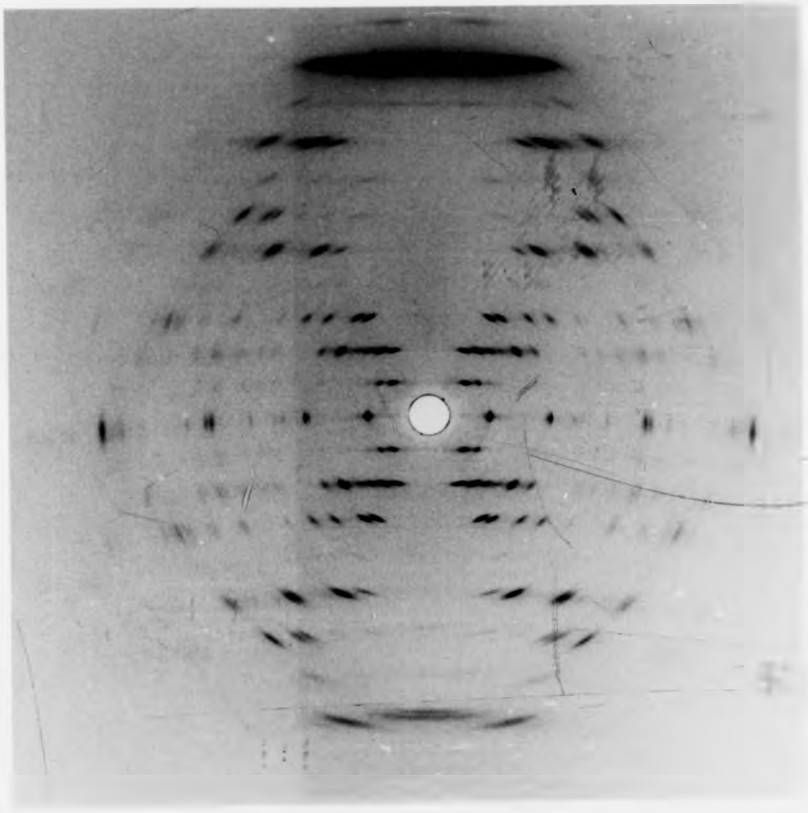


PLATE 1.4 The crystalline B-form obtained from Li DNA (Courtesy of Prof. M.H.F. Wilkins, FRS).

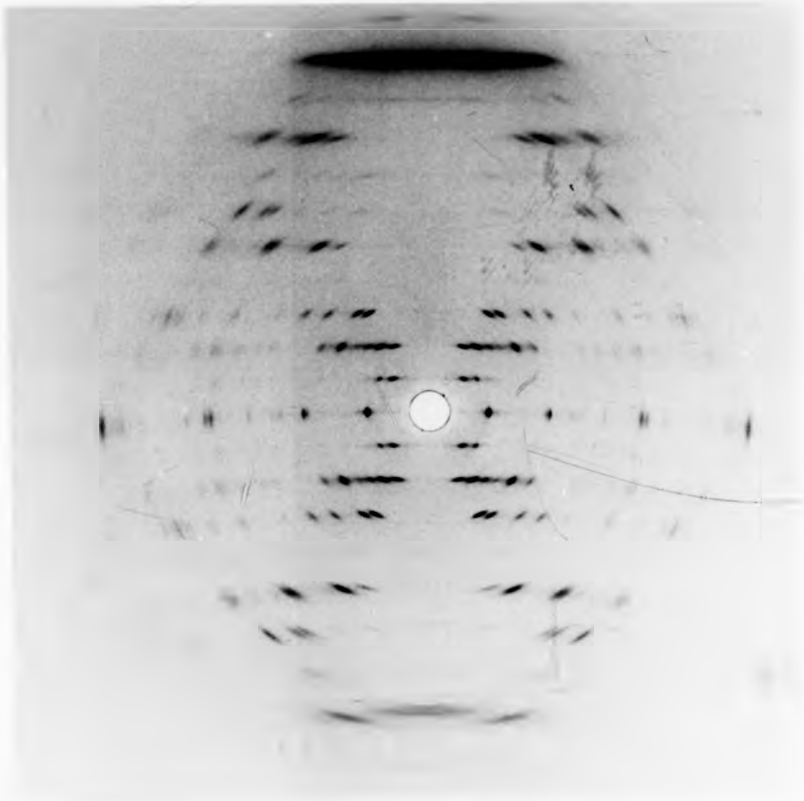


PLATE 1.4 The crystalline B-form obtained from Li DNA (Courtesy of Prof. M.H.F. Wilkins, FRS).

The overall features of the molecular conformation are a helix pitch of $\sim 34\text{Å}$, a 10_1 mononucleotide helical repeat, and C3-*exo* puckering of the sugar rings. The bases are displaced so that D = -0.16Å, having a twist of -2.1° and a tilt of -6°.

There have been a number of different models in which the details of the above conformation have been altered. Arnott and Hukins (1973) have tried using a C2'-*endo* puckering instead of C3'-*exo*, but found it hard to judge which gave the better fit on the basis of diffraction data alone. With the acquisition of better data, however, Arnott and Chandrasekaran (unpublished) came to the conclusion that the C2'-*endo* model was preferable.

From the crystalline Li B-DNA patterns, the molecules are found to pack into an orthorhombic lattice having a space group $P2_12_12_1$ where the molecular diad of the base points along the *b* direction. There are two molecules per unit cell, and the one in the centre of the cell is displaced by 0.33 c along the Z axis. It is noteworthy (Dover, 1977) that the ratio of a/b is equal to tan 36° and that this gives equivalent inter-helical contacts throughout the lattice.

The semi-crystalline Na B-DNA patterns indicate hexagonal packing and although the molecular conformation is thought to be the same as that for the orthorhombic patterns, lack of clarity in the sodium patterns make this difficult to assert. Work on poly d(G-C).poly d(G-C) and on poly d(A-C).poly d(G-T) by Leslie et al. (1980) does, however, tend to support this idea.

Another member of the B family is observed in fibres made from Na poly(dA).poly(dT) and is called B'-DNA (Arnott and Selsing, 1974). B'-DNA observed in poly(dA).poly(dT) was assumed (Arnott and Selsing, 1974) to be capable of packing into two different lattices: one hexagonal (α-B'-DNA) and one orthorhombic (β-B'-DNA). The α-B' form has also been obtained from poly d(A-T).poly d(A-T) and the two diffraction patterns (one

from d(A-T)) are compared in Plate 5.2.

A structure that has been designated B" has been discovered by Mahendrasingam (1984) in fibres made from the potassium salt of poly d(G-C).poly d(G-C). This molecule has a pitch of $\sim 34\text{Å}$, ten residues per helix turn and possibly a dinucleotide asymmetric unit. The diffraction pattern obtained from B"-DNA is shown in Plate 1.5.

A summary of the lattice and conformational parameters associated with each of the various forms discussed above is given in the Appendix.

1.4.3 The 'C' Set of Structures

The C conformation was first described by Marvin et al. (1958, 1961) for fibres of Li DNA at relative humidities of $\sim$ 66% or below. The molecule was found to pack either into a hexagonal or an orthorhombic lattice (depending on the amount of chloride present as LiCℓ) and produced a fibre diffraction pattern similar to the one shown in Plate 1.6. Marvin analysed the C patterns in terms of a clockwise helix having 28_3 symmetry and a pitch of $\sim 31\text{Å}$.

The importance of the C conformation has recently become apparent with the realisation that it is not just a structural distortion of the B conformation associated with the lithium ion, but a 'low salt' structure that can be routinely obtained in the presence of a variety of counterions (Zimmerman and Pheiffer, 1980; Rhodes et al., 1982). In the study of native and synthetic DNAs it has become evident that the conformation of C-DNA shows great variability, having a pitch which can be anything from 28Å to 32Å and containing between 8 and 9.6 residues per helix turn. (Zimmerman and Pheiffer, 1980). The features associated with a variety of different C patterns (and a variety of proposed models) are given in the Appendix.

Arnott and Selsing (1975) have revised the Marvin C model so that it more closely resembles the structure of B-DNA, although in the light of the agreement between their calculated and Marvin's observed data, the

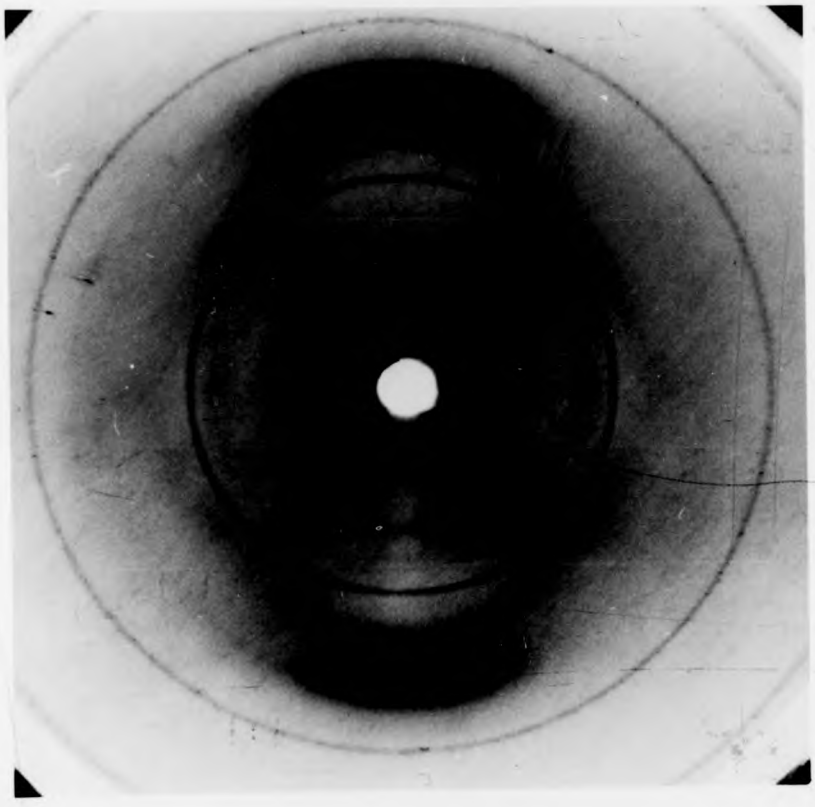


PLATE 1.5 The B" conformation of K Poly d(G-C).poly d(G-C) (from Mahendrasingam, 1984)

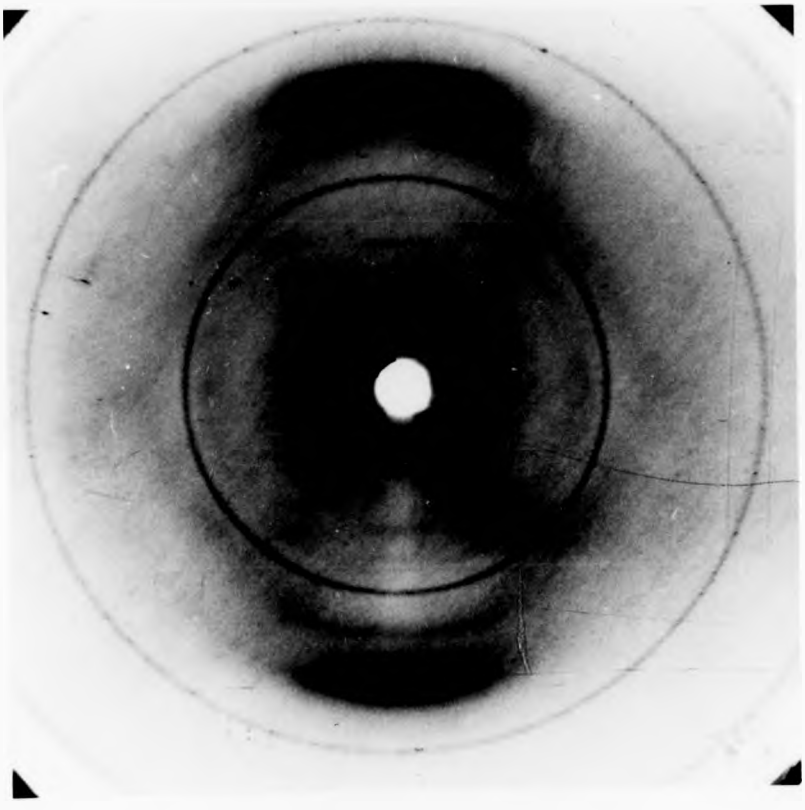


PLATE 1.5 The B" conformation of K Poly d(G-C).poly d(G-C) (from Mahendrasingam, 1984)

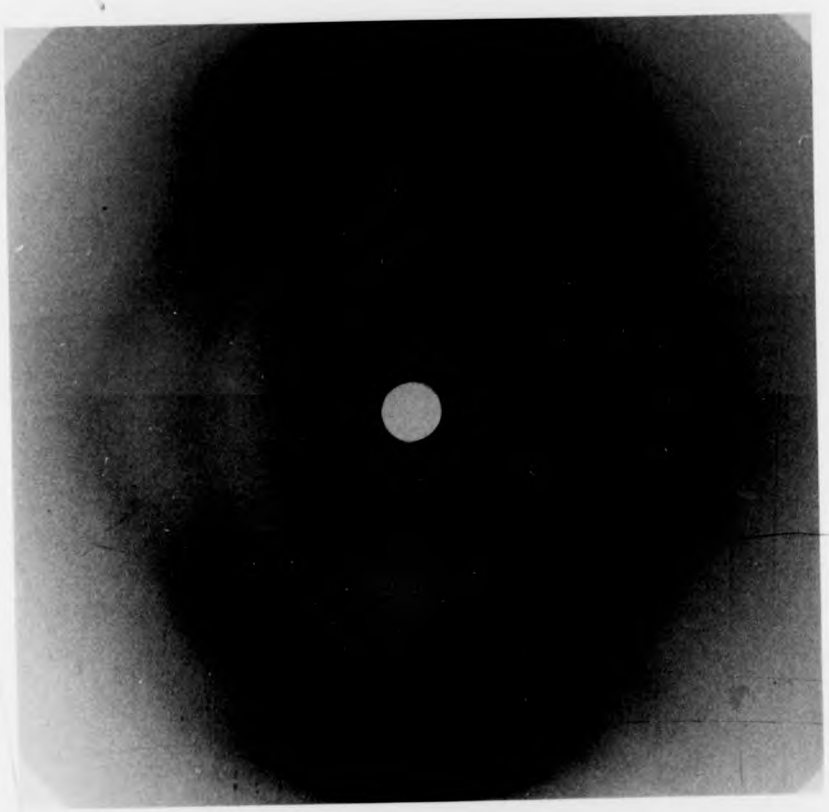


PLATE 1.6 A Semicrystalline C-DNA pattern obtained from Na DNA at 33% RH.

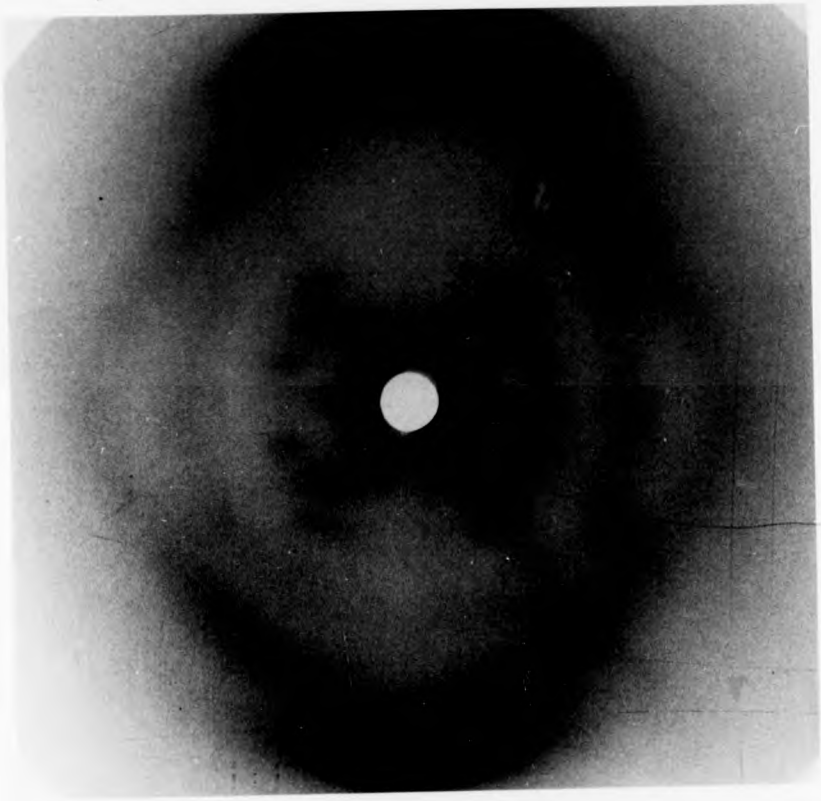


PLATE 1.6 A Semicrystalline C-DNA pattern obtained from Na DNA at 33% RH.

reasons for doing this are not clear. The data obtained from the Marvin patterns is not very extensive and bearing in mind the complex packing arrangements found in C-DNA lattices, it is thought wise not to assume too much about the relationship between B-DNA and C-DNA.

C'-DNA has been observed for the lithium salt of poly d(G-G-T).poly d(A-C-C) (Leslie et al., 1980) and has been found to be an integral nine-fold helix having a pitch of $\sim 29\text{Å}$. A conformation fairly similar to this has been observed by Rhodes (1982) and by Mahendrasingam (1984) and is analysed in Chapter Four (see Plates 4.2, 4.3, 4.4).

Leslie et al. have also observed a paracrystalline conformation from Na poly d(A-G).poly d(C-T) which they call C". This has 9_2 symmetry, a pitch of 29.1Å and crystallises into a hexagonal lattice.

1.4.4 D-DNA

In 1963, Davies and Baldwin (1963) reported a unique fibre diagram from samples of poly d(A-T).poly d(A-T). The conformation had a pitch of $\sim 24.5\text{Å}$. A pattern of this type is shown in Plate 1.7. Mitsui et al. (1970) reported a similar diagram from fibres of poly d(I-C).poly d(I-C) and such a pattern is shown in Plate 1.8. Mitsui et al. proposed that the structure was left-handed, having 8_7 symmetry, O5-endo puckering and bases displaced 2Å behind the helix axis.

Arnott et al. (1974), however, disagreed with the Mitsui analysis and basing their structure determination on data obtained from poly d(A-T).poly d(A-T), suggested a molecular conformation having 8_1 symmetry. The D form from poly d(A-T).poly d(A-T) packs into a tetragonal lattice and is named α-D-DNA, whereas the D form from poly d(A-T-T).poly d(A-A-T) adopts a hexagonal lattice (see Plate 6.5) and is called β-D-DNA (Selsing, Arnott and Ratliff, 1975). Although the overall intensity distribution is similar in both α and β patterns, there are some features which can only be attributed to differences in molecular conformation.

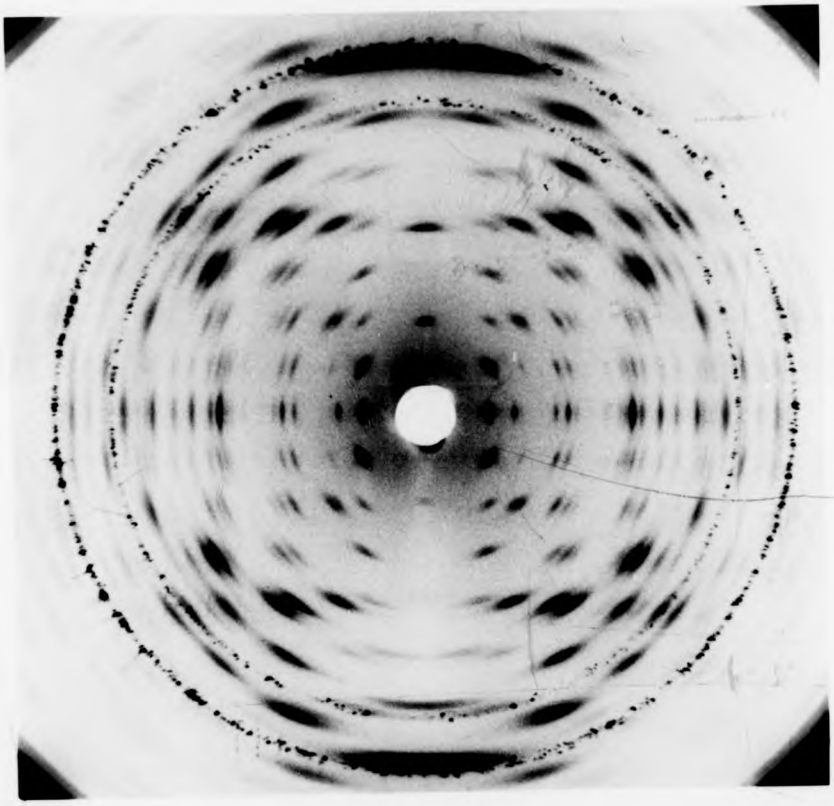


PLATE 1.7 The α-D conformation seen in a fibre of Rb poly d(A-T). poly d(A-T) (from Mahendrasingam, 1984).

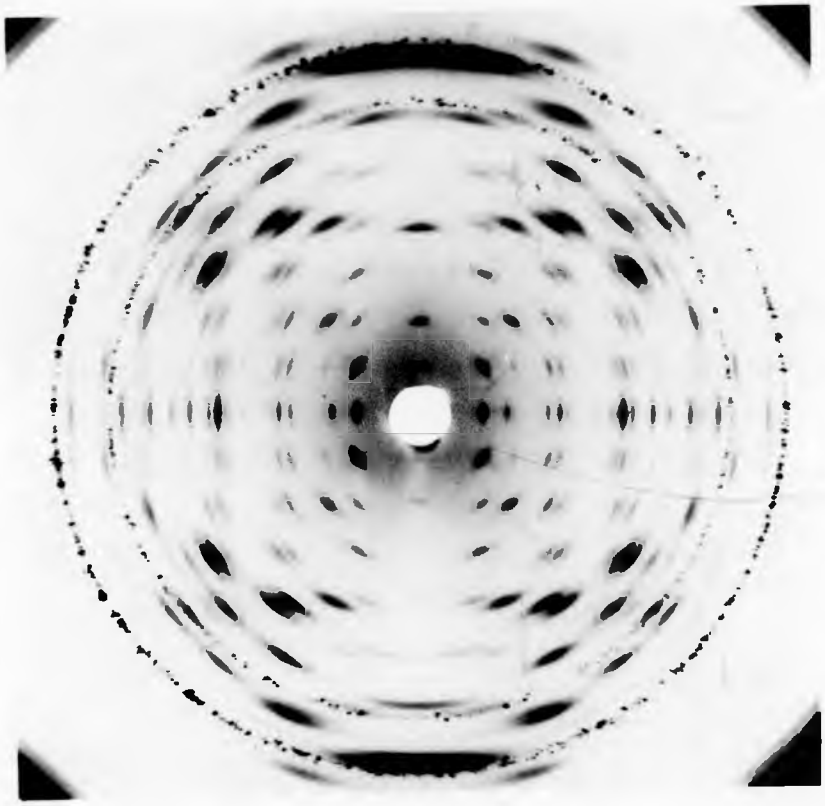


PLATE 1.7 The α-D conformation seen in a fibre of Rb poly d(A-T). poly d(A-T) (from Mahendrasingam, 1984).

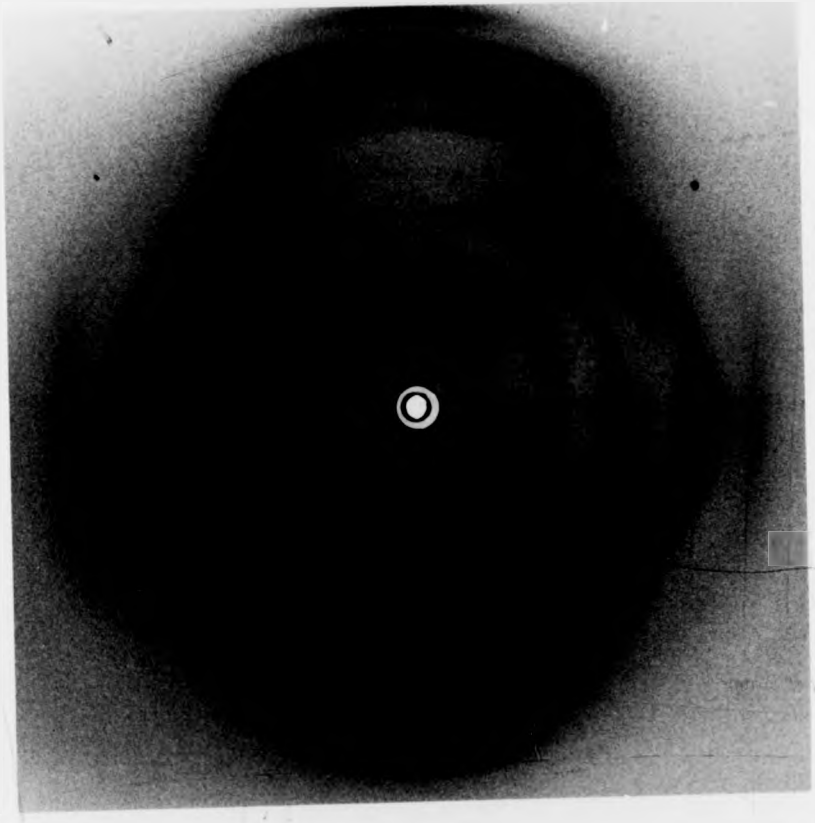


PLATE 1.8 The α-D conformation seen in a fibre of poly d(I-C). poly d(I-C) (from Mahendrasingam, 1984).

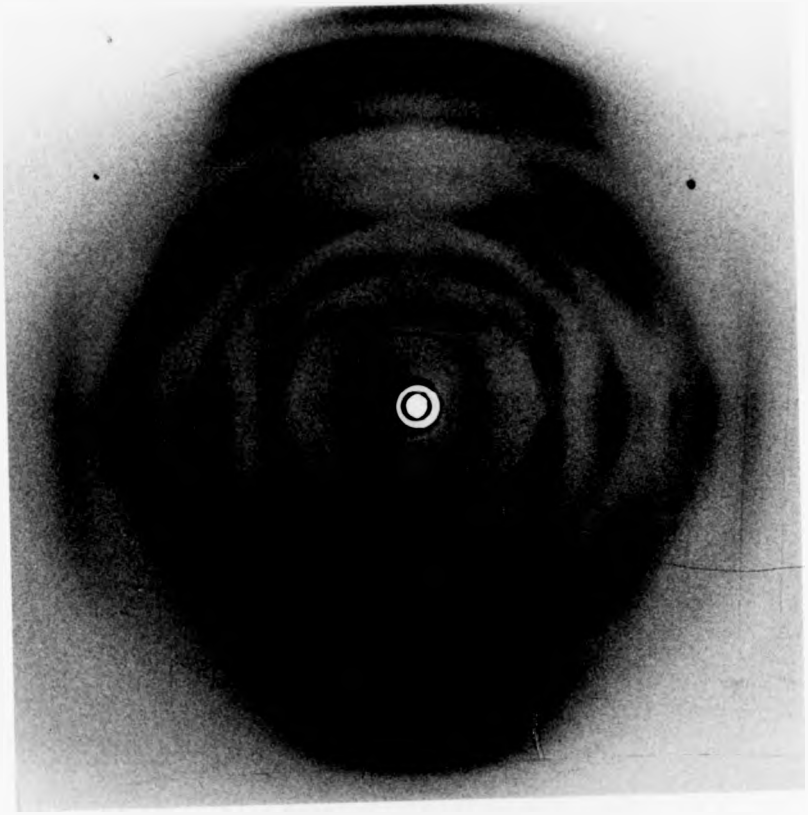


PLATE 1.8 The α-D conformation seen in a fibre of poly d(I-C). poly d(I-C) (from Mahendrasingam, 1984).

Since the right-handed models of Arnott et al. (1974, 1975), a variety of other structures have been proposed for D-DNA: a left-handed model having 7_6 symmetry and Hoogsteen base pairs (Donohue and Trueblood, 1960) was proposed by Drew and Dickerson (1982) but has since been rejected. A report by Ramaswamy et al. (1982) claims that equally satisfactory models can be generated for both helix senses, and that the data does not permit discrimination between the two.

Work undertaken in this laboratory (Mahendrasingam, unpublished; Forsyth, Chapter Six of this work) tends to favour the idea that the D helix is left-handed. The details of these studies are given in Chapter Six.

Two other variants of the D-type have recently been obtained from poly d(A-T).poly d(A-T). The F conformation is shown in Plate 1.9 and has been obtained from the lithium salt of this polymer (Mahendrasingam, 1984; Arnott, personal communication) and from the sodium salt (Mahendrasingam, unpublished). Another D conformation (referred to here as γ-D) has been obtained from a fibre of Li poly d(A-T).poly d(A-T) made by Dr. A. Mahendrasingam when the salt concentration is significantly lower than the level required to produce the F form. A γ-D pattern recorded at the Daresbury SRS by the author and his colleagues is shown in Plate 1.10. The intensities diffracted from these fibres indicate that the molecular structure is the same as the α-D form but the lattice, which is still approximately tetragonal is distorted so that either $a^* \neq b^*$ or $\gamma^* \neq 90^\circ$. This results in a splitting of reflections throughout the pattern which would normally overlap and consequently in the acquisition of a great deal more data relating to molecular transform of the D helix.

1.4.5 E-DNA

At low relative humidities the synthetic polynucleotide poly d(I-I-T).poly d(A-C-C) takes up an unusual conformation denoted E-DNA

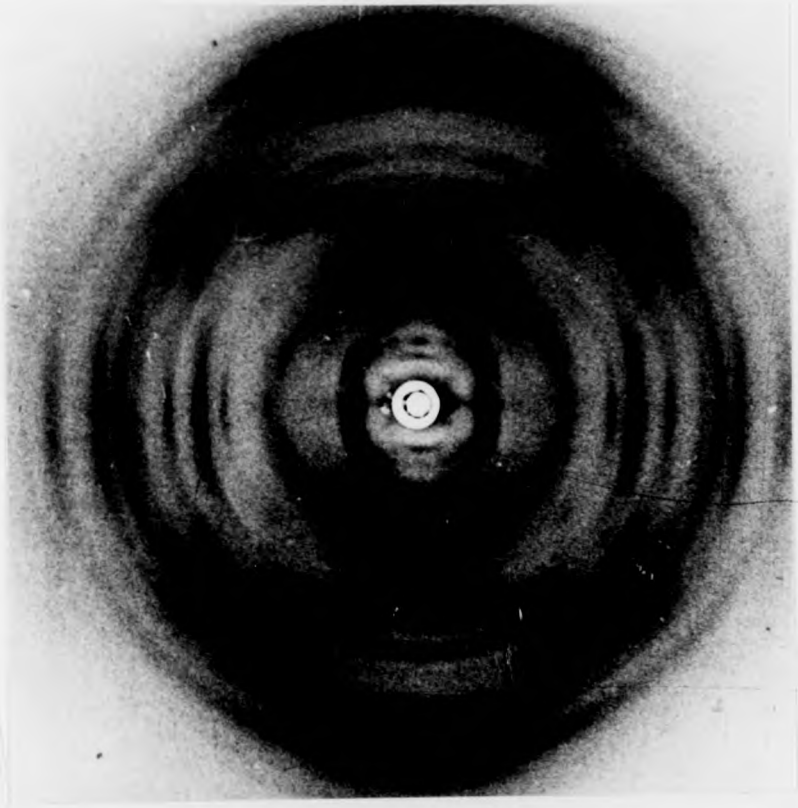


PLATE 1.9 The 'F' conformation seen in a fibre of Li poly d(A-T). poly d(A-T) (from Mahendrasingam, 1984).

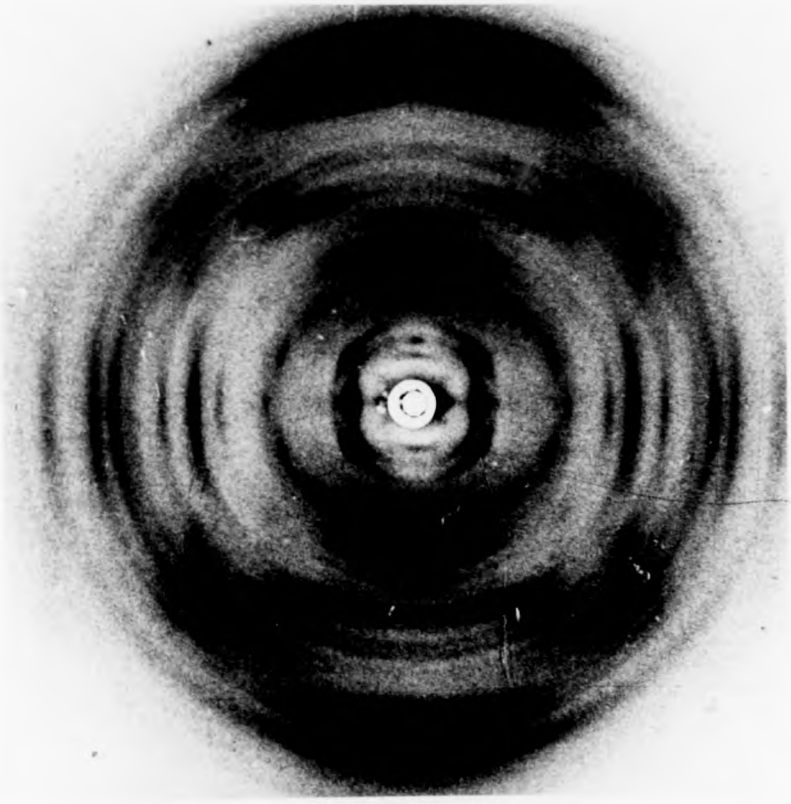


PLATE 1.9 The 'F' conformation seen in a fibre of Li poly d(A-T). poly d(A-T) (from Mahendrasingam, 1984).

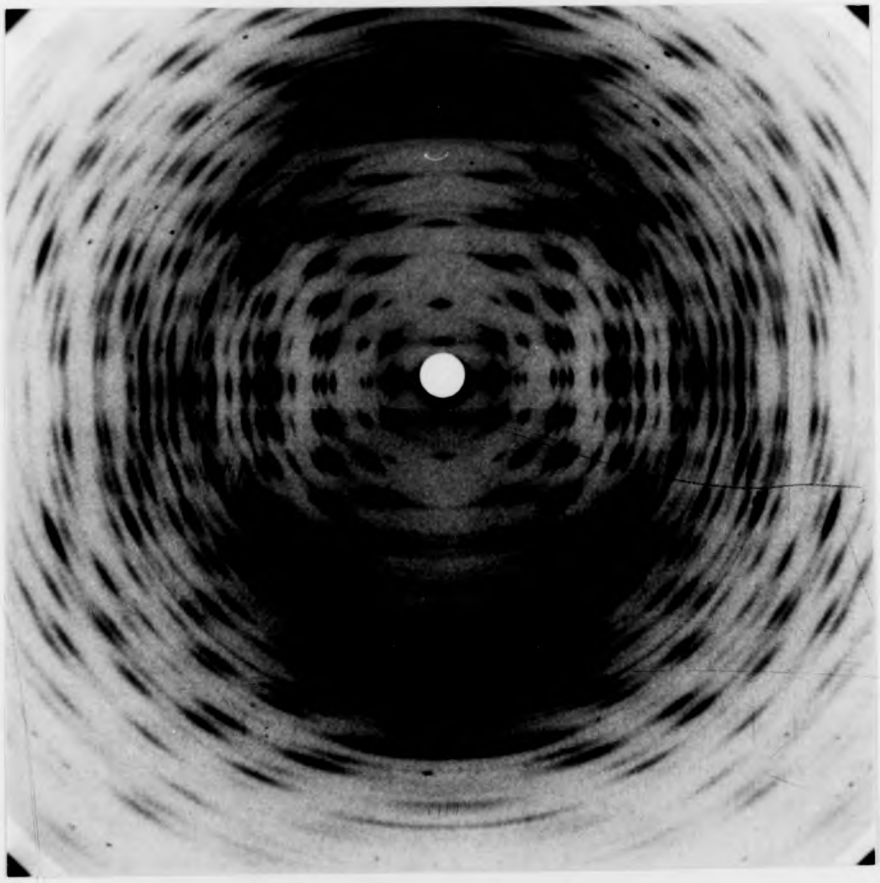


PLATE 1.10 The 'γ-D' pattern obtained from the lithium salt of poly d(A-T).poly d(A-T). The fibre was prepared by Dr. A. Mahendrasingam and the photograph recorded at the Daresbury SRS.

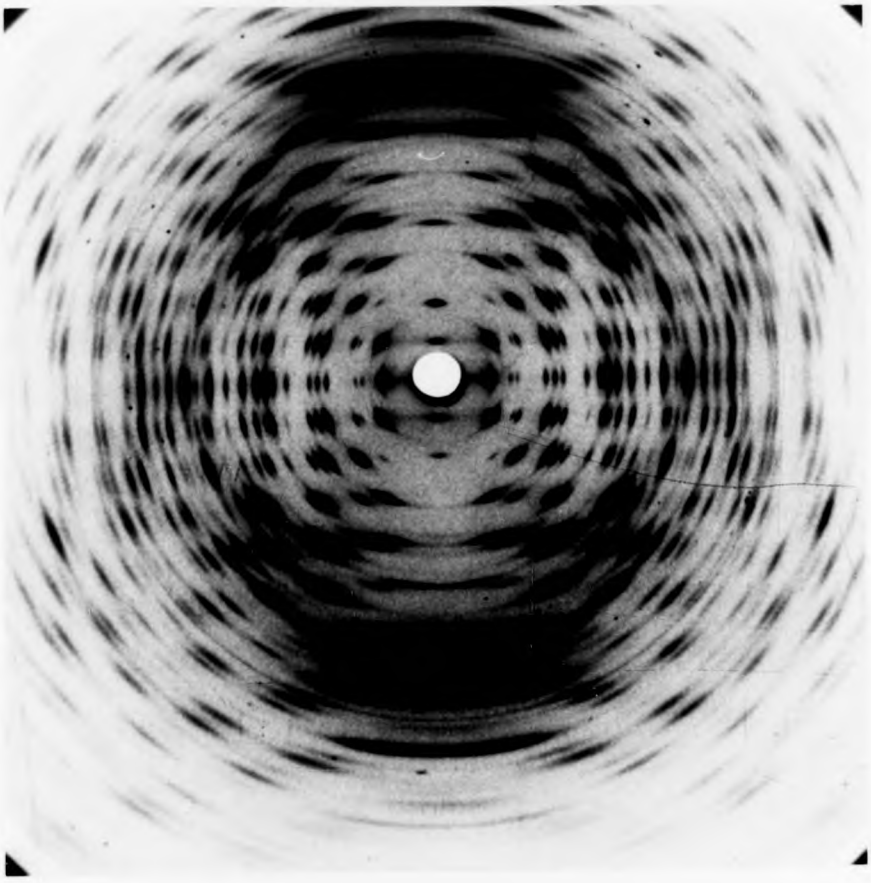


PLATE 1.10 The 'γ-D' pattern obtained from the lithium salt of poly d(A-T).poly d(A-T). The fibre was prepared by Dr. A. Mahendrasingam and the photograph recorded at the Daresbury SRS.

(Leslie et al. 1980). This molecule has a pitch of 48.7Å and has 3_2 symmetry where the asymmetric unit has five nucleotides (see Plate 1.11). A starting model put forward by Chandrasekaran et al. (1980) suggests that E-DNA has all its conformation angles in the trans range.

1.4.6 Z-DNA and S-DNA

The analysis by Wang et al. (1979) of single crystal diffraction data from the oligomer d(CpGpCpGpCpG) has sparked off intense interest in the possible polymer conformations of poly d(G-C).poly d(G-C). Wang et al. solved the structure of a $d(CG)_3$ oligomer to a resolution of 0.9Å and showed that the molecule consisted of a left-handed double helix (pitch of 44.6Å) having Watson-Crick base pairs. Because of the zig-zag nature of the backbone chain, the structure has been called 'Z-DNA', and is novel in a number of respects other than the sense of the chain: firstly the cytosine containing nucleotides have an anti base-sugar orientation and a C2'-endo sugar pucker, whereas the guanine containing nucleotides have a syn orientation and the sugars have a C3'-endo pucker; secondly, the base stacking is unusual in that bases are 'sheared' with respect to each other in a fashion again distinctive of a dinucleotide repeat. As a result of this, the diad axis lies in between the base planes. Because of the fact that the bases are pulled away from the helix axis by approximately 3Å, there is only one deep groove (corresponding to the minor groove of B-DNA) along the double helical chain.

Since the work of Wang et al. (1979) several other single crystal studies of G-C oligomers have been undertaken. All of them indicate a similar type of conformation (there are variations which possibly occur as a result of the presence of different ions - these will be discussed in Chapter Six).

Fibre diffraction data from the polymer poly d(G-C).poly d(G-C) indicates the occurence of the A and B conformations (depending on the ionic

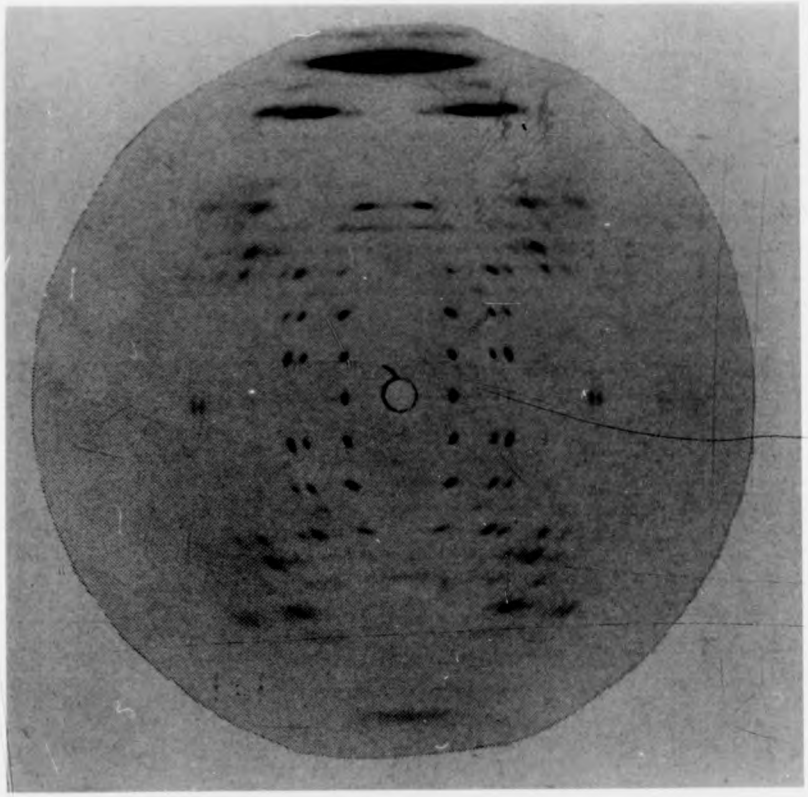


PLATE 1.11 The 'E' conformation of poly d(I-I-T).poly d(A-C-C) (from Chandrasekaran et al., 1980)

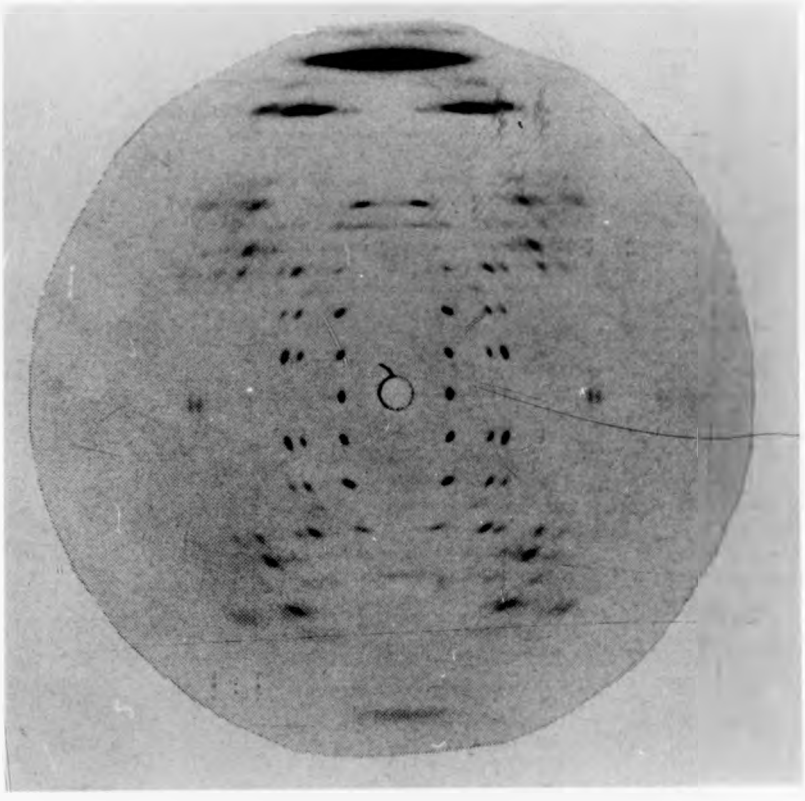


PLATE 1.11 The 'E' conformation of poly d(I-I-T).poly d(A-C-C) (from Chandrasekaran et al., 1980)

and water content) and also a unique conformation called S-DNA (Leslie et al., 1980; Arnott et al., 1980). The fibre diagram associated with S-DNA (see Plate 1.12) was first obtained by Arnott et al. (1980) from the sodium salt and showed a high degree of disorder. The overall features of this data are well explained by a model having the basic features of Z-DNA. The stereochemical complexity of transitions which occur within fibres of poly d(G-C).poly d(G-C) is immediately evident.

The S-conformation has also been reported to occur in fibres of poly d(A-C).polyd(G-T) (Arnott et al., 1980), but it has not (as yet) been observed from fibres of poly d(A-T).poly d(A-T). It is thus conceivable that the S-form may only occur in the presence of G-C base pairs.

1.5 THE PROJECTS UNDERTAKEN IN THIS WORK

Recent experimental work now suggests not only the presence of left-handed DNA in intact cells, but also the possibility of dynamic transitions involving a complete change of helical sense as a function of ionic and hydration conditions. In this study emphasis has been placed on the significance of left-handed structures and also on the importance of new experimental facilities which may enable the various mechanisms of transition to be detailed.

In the light of recently acquired high quality x-ray data for C'-DNA this work starts off with a definitive structural analysis of the C'-conformation, in particular focussing on the assumption that the C family as a whole is a minor variant of the B-form.

In Chapter Five the α-B' conformation of poly d(A-T).poly d(A-T) is analysed and compared with the 'heteronomous' model of Arnott et al. (1983) which has been proposed for both the α and β forms of B'-DNA.

The structure of β-D-DNA is examined in Chapter Six, and special attention is given to the correspondence that appears to exist between the

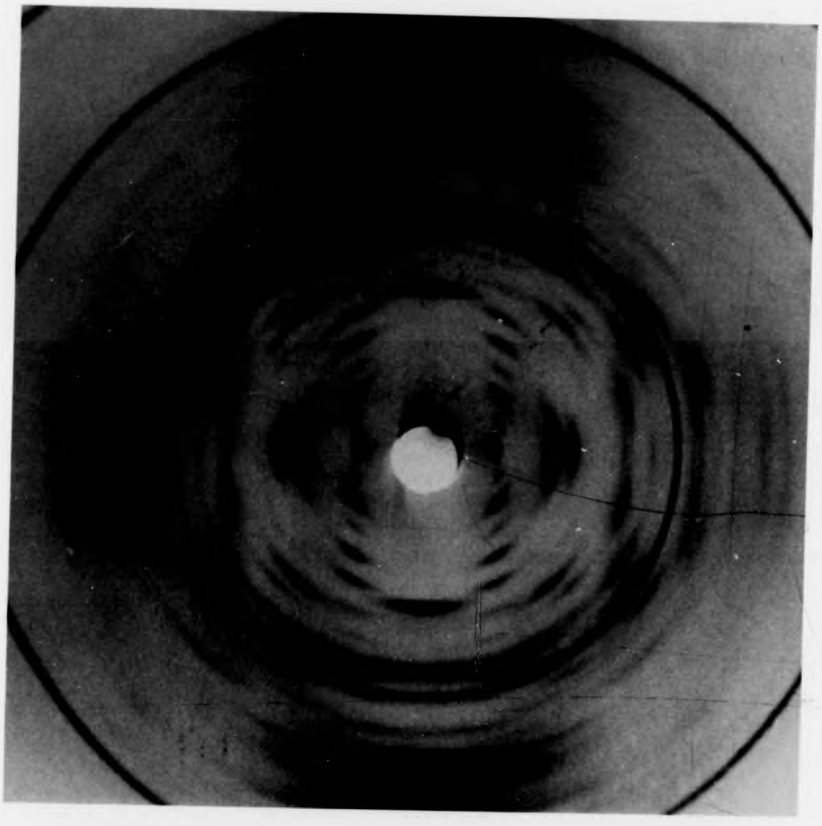


PLATE 1.12 The 'S' conformation as obtained from fibres of KF poly d(G-C).poly d(G-C) (photograph recorded on the SAS station at Dareshury).

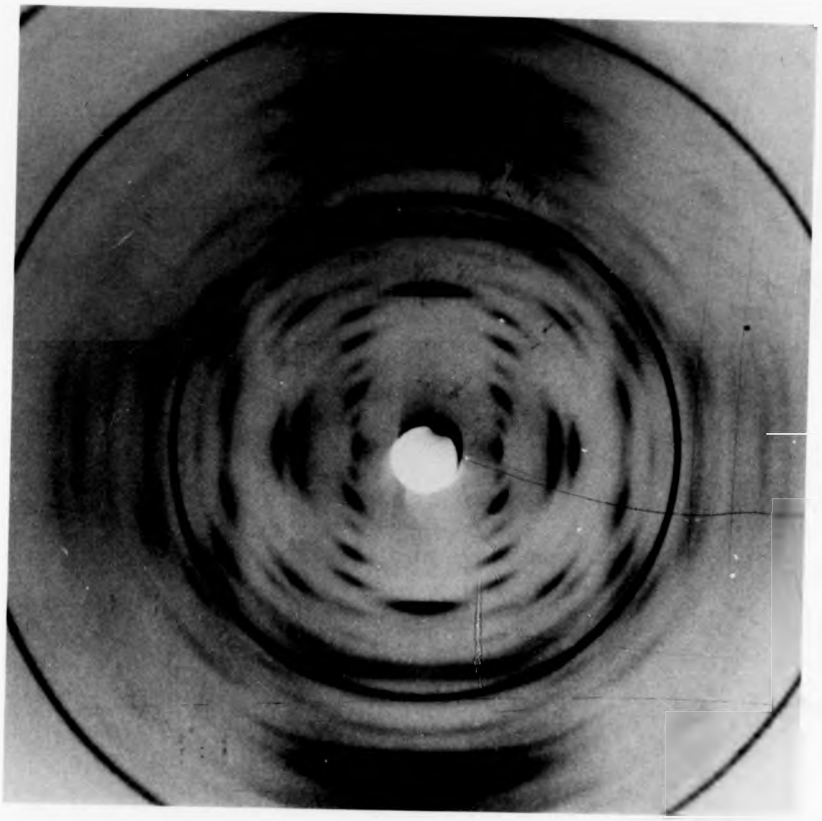


PLATE 1.12 The 'S' conformation as obtained from fibres of KF poly d(G-C).poly d(G-C) (photograph recorded on the SAS station at Daresbury).

transitional sequences found in poly d(A-T).poly d(A-T) and those of poly d(G-C).poly d(G-C).

Chapter Seven qualitatively examines the effect of a bacteriostatic and mutagenic acridine drug, proflavine, on the sequence of transitions normally observed in fibres of poly d(A-T).poly d(A-T), fibres of poly d(G-C).poly d(G-C), and also fibres made from Micrococcus Lysodeikticus DNA. Whilst proflavine has been extensively studied, the emphasis has largely been structural whereas here the effect of the drug on the nature of fibre transitions is examined.

A number of 'time resolved' experiments are described in Chapter Eight. These were undertaken at the SERC Daresbury synchrotron radiation source (SRS) in England and were novel experiments using recently developed equipment. The ultimate aim of such projects is to detail the stereochemistry of fibre transitions and are of particular significance in the study of changes between conformatinons which are thought to have different helical twists.

Chapter Nine concludes the thesis with a brief overview of the results obtained and a few suggestions for the most profitable directions in which to indulge further effort.

CHAPTER TWO

MATERIALS AND EXPERIMENTAL METHODS

2.1 THE MATERIAL

All DNA material was obtained from the Sigma Chemical Company, Boehringer, or from J. Brahms at the University of Paris. The material was classified as 'highly polymerised' and previous experiments had indicated that it was on the whole of adequate purity for the purposes of x-ray fibre diffraction.

2.2 PREPARATION OF FIBRES

The DNA obtained from the sources above was known to be stabilised either by sodium or potassium cations. In order to perform quantitative studies relating to differing concentrations of excess ion, it was necessary to first eliminate as much as possible of the original ion and replace it with a certain concentration of the ion under investigation. A known amount of the supplied polymer ($\sim$ 2 mg) was dissolved and dialysed against a concentrated solution ($\sim$ 1 M) containing the new ion. Subsequent dialysis against rather less concentrated solutions ($\sim$ 2-10 mM) was used to attain the desired level. DNA was precipitated from the salt solution by cold ethanol, wound onto a glass rod, dried, and stored for use. Another method of varying ionic strength within DNA samples rather more controllably was by centrifugation from solutions of known concentration. This method was frequently used in conjunction with the dialysis technique described above. In general, centrifugation was performed on an MSE pegasus ultracentrifuge for a period of 12 hours at 60,000 rpm, and the supernatant so obtained was kept

for later studies on exact ionic concentration (flame emission spectroscopy). The sedimented gel was either stored or used directly in fibre preparation.

The precipitation of DNA by ethanol from a solution is a very difficult process to control and for a good yield of the original DNA, high salt strengths are required. This makes the acquisition of low salt samples, difficult at the best of times, rather hazardous when the material in question is expensive and available only in small quantities. On the whole, the centrifugation method was used in such situations.

Fibres were drawn from concentrated gels by making use of a fibre pulling system as shown in Figure 2.1. Care here was required to obtain highly oriented fibres. The glass rods were made to an approximate diameter of 100 μ (larger diameter rods tended to draw too much of the gel onto the rod itself, and so waste some of the material). Particularly important for the acquisition of highly oriented specimens was the 'pulling' of fibres as they dried down. As suggested by Langridge et al. (1960), 'it is almost certain that parallelism of the molecules in a fibre is produced by shearing in the material during fibre drawing', and it was deemed very important (for clear diffraction spots, as will be discussed later) that the 'pulling' process was undertaken cautiously. During the process of drying, fibres were observed using an Olympus binocular microscope, and an Olympus BH polarising microscope.

The best exposures were obtained from fibres having diameters in the range of 100 μ → 400 μ, although thinner fibres were used (necessitating considerably longer exposures on conventional x-ray sets).

2.3 THE ACQUISITION OF DIFFRACTION DATA FROM FIBRES

Data was obtained using both conventional x-ray sources here in this laboratory, and the synchrotron radiation source (SRS) at Daresbury, Warrington, England, which is further described in Chapter Eight.

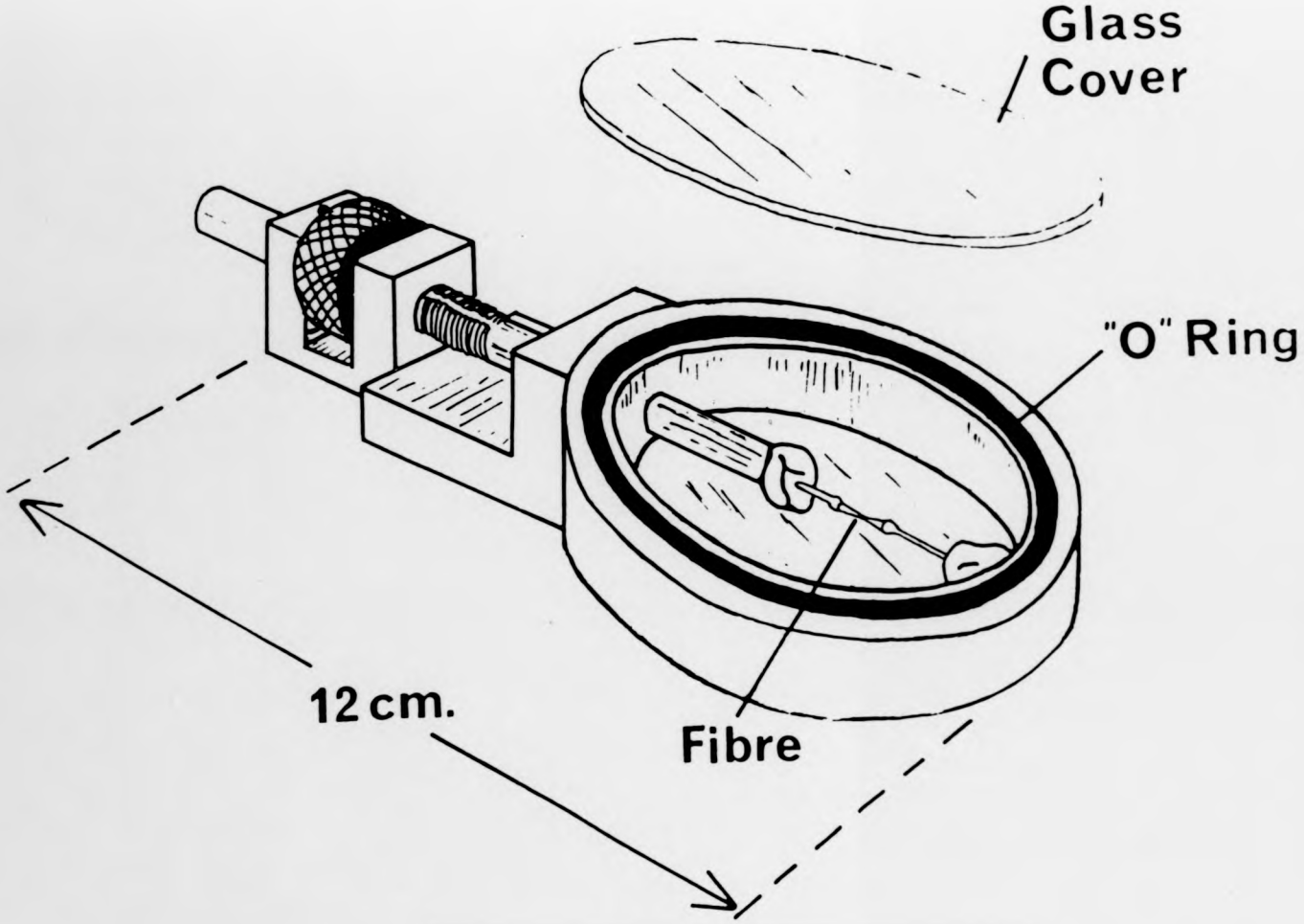


FIG 2.1 The apparatus used to draw fibres.

2.3.1 Conventional X-ray Sets

Two different types of X-ray set were used, each using nickel filtered copper K-α x-radiation having a wavelength of 1.5418Å. The first was a Hilger and Watts semi-micro focus unit based on the design of Ehrenberg and Spear (1951). This provided a focal spot having (roughly) a vertical half-width of 0.075 mm and a horizontal width of 0.1 mm. The direct voltage applied across the tube was usually in the region of 36 KV to 40 KV; higher voltages were found to produce an unacceptable level of white radiation. The level of white radiation was also increased by deposition on the anode, which ensued from 'cracking' of diffusion pump oil. The anode was cleaned regularly. Filaments carried between 3 and 10 mA and generally lasted about a week, but filament life was increased as vacuum conditions were improved.

Other generators used were Elliott GX6 and GX20 rotating anode machines. These sets were operated at 35 KV and at a tube current of ∿ 60 mA. (see Plate 2.1).

Two types of x-ray camera were used. Small pinhole cameras (very similar in design to those described by Langridge et al. (1960a)) were, in general, used on the Hilger and Watts generators. The larger GX6 and GX 20 machines were usually used in conjunction with Searle cameras. Searle cameras were arranged with either Elliott toroidal optics (Elliott, 1965) or Frank's optics (Franks, 1958) depending on the thickness of fibre in use and the desired exposure time.

2.3.2 Synchrotron Radiation Source (SRS)

Data was also obtained at the SERC Daresbury SRS. The use of this system eliminated many of the problems associated with the use of conventional sets. The flux was ∿ 80 times greater and consequently exposure times were very much shorter (3-15 minutes), hence enabling time-resolved studies to be undertaken. The radiation was highly monochromatic and diffraction patterns were very clear. Radiation damage to samples was not as severe as that

PLATE 2.1 An Elliott GX20 rotating anode generator with a Searle camera in position

PLATE 2.1 An Elliott GX20 rotating anode generator with a Searle camera in position

PLATE 2.1 An Elliott GX20 rotating anode generator with a Searle camera in position

encountered using conventional sources. The camera that was used for these studies at Daresbury was designed by Dr. W.J. Pigram and built in the Physics Department Workshop at the University of Keele. Further detail relating to the experimental set-up at Daresbury is given, along with an account of studies completed there, in Chapter 8.

2.3.3 Variation of the Water Content in Fibres

X-ray cameras were always filled with helium to avoid air scatter. The relative humidity of the fibre environment was varied to adjust the water content of fibres being studied. This was achieved by bubbling the helium through appropriate saturated salt solutions. The salt solutions used and the relative humidities that they provided are given below:

Salt	Relative Humidity (%)
Calcium Chloride	33
Potassium Carbonate	44
Potassium Bromide	58
Sodium Nitrite	66
Sodium Chlorate	75
Potassium Chloride	84
Sodium Tartrate	92
Sodium Sulphite	95
Potassium Chlorate	98

In later experiments at the SRS where closer control over relative humidity was needed, a small water bath was installed inside the camera to acquire high humidities as required. A Vaisala HMP 31VT probe was used to measure the ambient relative humidity and temperature of the fibre environment. Details of this time resolved work is given in Chapter 8.

2.4 MEASUREMENT OF DIFFRACTION PATTERNS

Fibres were calibrated by dusting with calcite powder. This gave a diffraction ring on the patterns corresponding to a spacing of 3.03Å. The

'flat film' coordinates of diffraction spots were measured with a travelling microscope; these were then converted to recriprocal space coordinates. The geometry of this conversion is described by Buerger (1942).

Intensities were assumed to be proportional to the trace area associated with a radial scan across a diffraction spot. Such scans were made using a Joyce-Loebl 3CS microdensitometer and also a software package called GUCKMAL (written by J.E. Ladner in Heidleburg and since modified by M. Elder, P. Machin and C. Nave at Daresbury Laboratory) in conjunction with diffraction data digitised on the rotating drum scanning densitometer at Daresbury. The C' pattern shown in Plate 4.3 was scanned with a raster size of 50 μ so that the entire pattern was made up of 900 x 900 records. Thepicture in Plate 2.2a shows the pattern displayed on the Sigma 'Args' graphics terminal at Daresbury by means of 255 grey levels which can be altered to highlight different regions of a pattern. A comparison of the digitised picture with a number of layer line traces made using 'GUCKMAL' suffices to illustrate some of the problems involved in computing intensities from fibre diagrams. Although this pattern is reasonably crystalline, thereare regions in which it is highly desirable to correlate discrete diffraction with continuous transform. The crystalline data itself (depending on the symmetry of the lattice) may be difficult to interpret where sampling seems in adjacent areas of reciprocal space (layer lines 1, 2, 3 in Plate 2.2). A further problem occurs in the correction of intensity data for a background profile of diffuse scatter.

Measured spot traces were corrected in the manner of Franklin and Gosling (1953b). The expression used for this correct was:

$$I = A_R \rho \xi \cos\theta \frac{2}{1 + \cos^2(2\theta)} \qquad 2.0$$

where A_R = radial trace area.

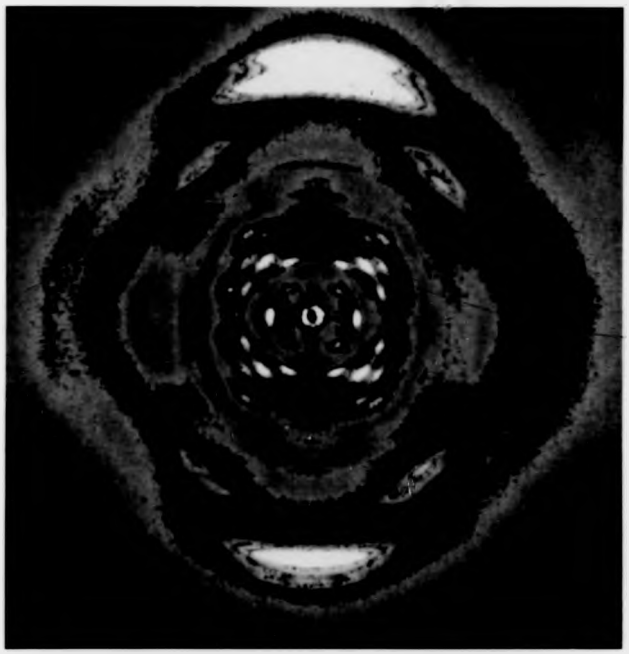


PLATE 2.2 The C' pattern of Plate 4.3 is shown here as displayed on an 'Args' terminal of the Daresbury VAX.

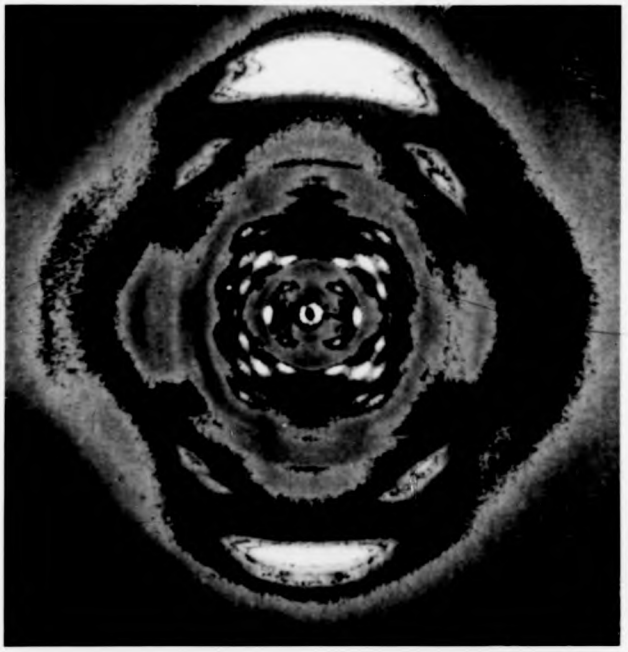


PLATE 2.2 The C' pattern of Plate 4.3 is shown here as displayed on an 'Args' terminal of the Daresbury VAX.

PLATE 2.2 The C' pattern of Plate 4.3 is shown here as displayed on an 'Args' terminal of the Daresbury VAX.

ρ = distance in reciprocal space from origin to reciprocal lattice point in question. This factor corrects for the broadening of spots into arcs.

ξ = reciprocal space ordinate. This factor in equation 2.1 applies the Lorentz correction.

θ = Bragg angle

$\cos\theta$ = factor necessary to correct observed intensities for absorption within a fibre

and

$$\frac{2}{1 + \cos^2(2\theta)} = \text{the polarisation factor}$$

2.5 MOLECULAR MODELBUILDING

Molecular modelbuilding was undertaken to generate models having satisfactory stereochemistry and good agreement with observed diffraction. Preliminary structures were constructed from wire models having a scale of 4 cm : 1Å but these were later refined so as to incorporate standard covalent bond lengths and angles by means of a modelbuilding program originally written by Pigram (1968) and modified by Goodwin (1977) and Greenall (1982). The major feature of this routine is in its constraint of a branched chain to specified helical symmetry and acceptable Van der Waals distances. Options exist which can apply other constraints (such as phosphate position) at run time. A flow chart of the modelbuilding program is shown in Figure 2.2. It currently runs on the CDC 7600 machine at the University of Manchester Regional Computer Centre (UMRCC).

During refinement, different models were varied so as to optimise agreement indices relating observed and calculated spot intensities or amplitudes:

$$R = \frac{\sum_i \left| |F_o| - |F_c| \right|_i}{\sum_i |F_o|_i} \qquad 2.1$$

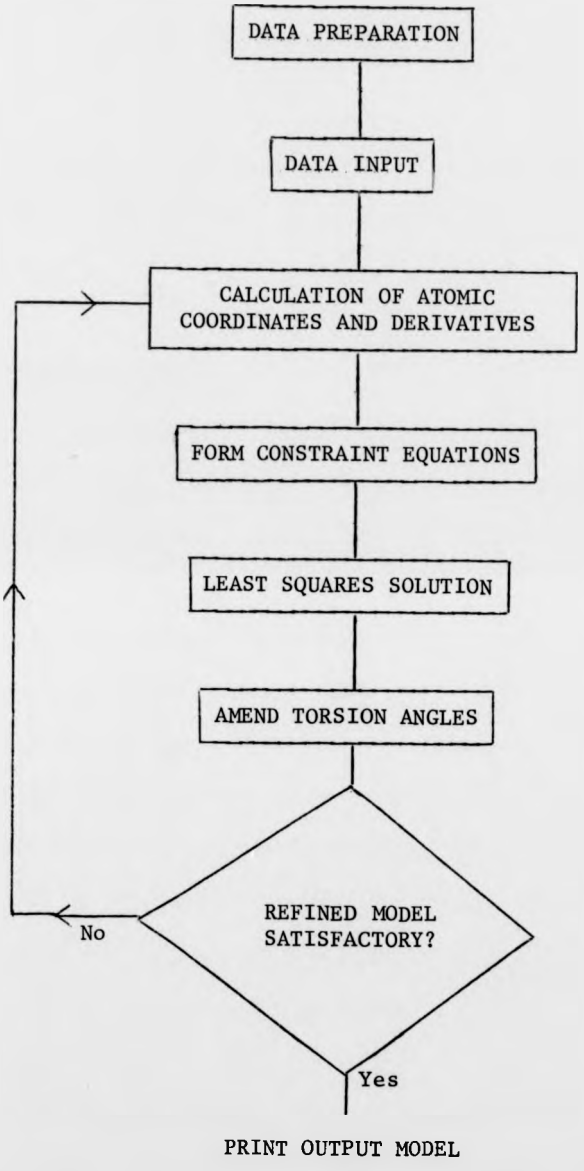


FIGURE 2.2 A flow chart of the modelbuilding program

$$R' = \frac{\sum_i |I_o - I_c|_i}{\sum I_{oi}} \qquad 2.2$$

where F_{oi} and F_{ci} are the observed and calculated structure amplitudes pertaining to the i^{th} reflection and I_{oi} and I_{ci} are the corresponding itensities.

2.6 COMPUTER PROGRAMS

A structure factor program was written by the author and computes the values of F_{ci} pertaining to equations 3.32 and 2.1. The methods used in this program are well documented and further detail will not be given here.

Another program was written to calculate and plot (using a GINO package of subroutines) the transform and squared transform of a helical molecule. This program has been used throughout the thesis to plot continuous squared transform calculations. It will also superimpose lines on the transform which indicate points of sampling by the lattice in question.

During the course of the work undertaken in Chapter Four, it also became necessary to write a program which would compute the variation of different classes of packing factor with molecular position (see Section 5.3).

A linked atom least squares program (LALS) which successively refines molecular models against diffraction data has been supplied by R. Chandrasekaran at the University of Purude, Indiana, and is currently being adapted for use on the Cyber 205 vector processing supercomputer that has just been installed at UMRCC. This program will also run on the CDC 7600 and the GEC computers at Keele. Other programs were written for use on the LSI 11/23 and ITT 2020 microcomputers in this laboratory.

CHAPTER THREE

DIFFRACTION OF X-RADIATION AND FIBRE CRYSTALLOGRAPHY

3.1 THE NATURE AND APPLICATION OF X-RAYS

X-rays are transverse waves in which the oscillation in the electric field vector is perpendicular to the direction of propagation of the wave: hard x-rays are characterised by wavelengths in the region of 1Å and as such can be used to determine atomic positions within a molecule. Although the atoms of a molecule will diffract x-rays, there is no analog of a lens to enable recombination of the scattered x-radiation hence forming an image. However, in certain circumstances this can be performed mathematically.

3.2 SCATTERING OF X-RAYS BE AN OBSTACLE

Diffraction of x-rays result from the interaction of the electric field component $\underline{E}$ of the radiation with the atomic electrons of the obstacle. The electric field oscillation of the incident wave causes a time dependent fluctuation of the electronic wave functions and the corresponding variation in current density is seen as the source of the scattered radiation. The component of this radiation which has the same frequency as the incident waves then gives rise to coherent scattering. Other frequency components scattered from the crystal will give rise to incoherent scattering, which appears as a continuous background which to a first order is independent of crystal orientation. If the incident wave vector is $\underline{K}_i$ then the wave vector of the elastically scattered radiation may be expressed as $\underline{K}_d$ where

$$|\underline{K}_i| = |\underline{K}_d|$$

CHAPTER THREE

DIFFRACTION OF X-RADIATION AND FIBRE CRYSTALLOGRAPHY

3.1 THE NATURE AND APPLICATION OF X-RAYS

X-rays are transverse waves in which the oscillation in the electric field vector is perpendicular to the direction of propagation of the wave: hard x-rays are characterised by wavelengths in the region of 1Å and as such can be used to determine atomic positions within a molecule. Although the atoms of a molecule will diffract x-rays, there is no analog of a lens to enable recombination of the scattered x-radiation hence forming an image. However, in certain circumstances this can be performed mathematically.

3.2 SCATTERING OF X-RAYS BE AN OBSTACLE

Diffraction of x-rays result from the interaction of the electric field component $\underline{E}$ of the radiation with the atomic electrons of the obstacle. The electric field oscillation of the incident wave causes a time dependent fluctuation of the electronic wave functions and the corresponding variation in current density is seen as the source of the scattered radiation. The component of this radiation which has the same frequency as the incident waves then gives rise to coherent scattering. Other frequency components scattered from the crystal will give rise to incoherent scattering, which appears as a continuous background which to a first order is independent of crystal orientation. If the incident wave vector is $\underline{K}_i$ then the wave vector of the elastically scattered radiation may be expressed as $\underline{K}_d$ where

$$|\underline{K}_i| = |\underline{K}_d|$$

and $\underline{K}_i$ differs from $\underline{K}_d$ in direction, but not magnitude.

If the scattering power of the object is describedby some function $f(\underline{r})$, then the diffracted radiation will be the Fourier transform of $f(\underline{r})$. Consider the general case of diffraction resultant from scattering at two points P_1 and P_2 of an obstacle (see Figure 3.1).

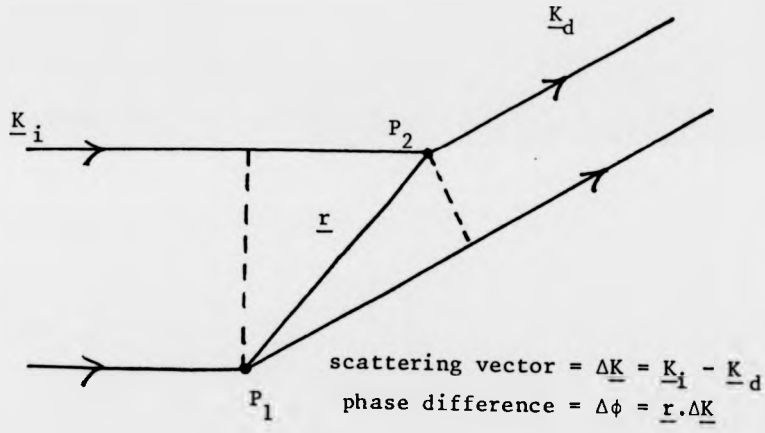


FIGURE 3.1 The geometry of the diffraction event.

The observed diffraction is given by:-

$$f(\Delta\underline{K}) = \int_{\text{all } \underline{r}} f(\underline{r}) \exp i\, \Delta\underline{K}.\underline{r}\, dr \qquad 3.1$$

and is thus described by a function in 'K-space' or 'Fourier space'. In principle the obstacle is reconstructed from the inverse transform:-

$$f(r) = \frac{1}{\sqrt{2\pi}} \int_{\text{all } \Delta K} F(\Delta K) \exp(-i\Delta\underline{K}.\underline{r})\, d(\Delta\underline{K}) \qquad 3.2$$

However, in real experiments the observed quantity is the intensity and the phase of the scattered radiation (the argument of the exponential function) is unknown.

A fully crystalline specimen is made up of a three dimensional lattice extending regularly over the entire sample. Each lattice point has an identical structure associated with it (the motif) and in the case of macromolecules this structure is very complicated. The scattering from such a crystal may be represented as the structure of the motif convoluted with a three dimensional set of δ functions (of unit scattering power) which represent the real lattice. By the Fourier convolution theorem the diffracted amplitude is then the product of the transform of the electron density within one unit cell with the transform of the set of δ functions representing the lattice which extends throughout the crystal:

i.e.

$$T\{f(\underline{r})\} = T\{\rho(\underline{r})\}.\ T\{L(\underline{r})\} \qquad 3.3$$

where $L(\underline{r})$ represents the lattice function and T denotes the Fourier transform operation.

The motif described by $\rho(\underline{r})$ is an irregular assembly of atoms and so its Fourier transform is continuous. The lattice function $L(\underline{r})$ can be written:

$$L(r) = \sum_{\text{all } pqr} \delta\,(\underline{r} - [p\underline{a} + q\underline{b} + r\underline{c}]) \qquad 3.4$$

where $\underline{a}$, $\underline{b}$, $\underline{c}$ are the crystallographic unit vectors, and p, q, r are integers (assuming a primitive lattice). The Fourier transform of $L(\underline{r})$ is

$$T\{L(\underline{r})\} = \sum_{\text{all } hk\ell} \delta(\Delta\underline{K}.\underline{a} - 2h\pi)\ \delta(\Delta\underline{K}.\underline{b} - 2k\pi)\ \delta(\Delta\underline{K}.\underline{c} - 2\ell\pi) \qquad 3.5$$

which in the case of an infinite lattice is non-zero only when

$$\Delta\underline{K} \cdot \underline{a} = 2h\pi \qquad 3.6a$$

$$\Delta\underline{K} \cdot \underline{b} = 2k\pi \qquad 3.6b$$

$$\Delta\underline{K} \cdot \underline{c} = 2\ell\pi \qquad 3.6c$$

Equations 3.6 are called the Laue equations and represent three sets of planes whose intersections specify the reciprocal lattice. Hence the amplitude diffracted from the crystal will be non-zero only when there exists a scattering vector:

$$\Delta\underline{K} = 2\pi\,(h\underline{a}^* + k\underline{b}^* + \ell\underline{c}^*) \qquad 3.7$$

satisfying the simultaneous equations 3.6. This is true when

$$\begin{array}{lll} \underline{a}^*.\underline{a} = 1 & \underline{a}^*.\underline{b} = 0 & \underline{a}^*.\underline{c} = 0 \\ \underline{b}^*.\underline{a} = 0 & \underline{b}^*.\underline{b} = 1 & \underline{b}^*.\underline{c} = 0 \\ \underline{c}^*.\underline{a} = 0 & \underline{c}^*.\underline{b} = 0 & \underline{c}^*.\underline{c} = 1 \end{array} \qquad 3.8$$

where a*, b*, c* are the unit vectors of the reciprocal lattice which are related to the unit vectors of the real lattice by

$$a^* = \frac{\underline{b} \times \underline{c}}{\underline{a}.(\underline{b}\times\underline{c})} = \frac{bc \sin\alpha}{V}$$

$$b^* = \frac{\underline{c} \times \underline{a}}{\underline{a}.(\underline{b}\times\underline{c})} = \frac{ca \sin\beta}{V} \qquad 3.9$$

$$c^* = \frac{\underline{a} \times \underline{b}}{\underline{a}.(\underline{b}\times\underline{c})} = \frac{ab \sin\gamma}{V}$$

and

$$\sin\alpha = \frac{V^*}{a^*\, b^*\, c^* \sin\beta^* \sin\gamma^*} \qquad 3.10a$$

$$\sin\beta = \frac{V^*}{a^*\, b^*\, c^* \sin\alpha^* \sin\gamma^*} \qquad 3.10b$$

$$\sin \gamma = \frac{V^*}{a^* b^* c^* \sin \alpha^* \sin \beta^*} \qquad 3.10c$$

where α, β, γ define the angles between the primitive vectors of the real lattice, and α^*, β^*, γ^* define the angles between those of the reciprocal lattice, as shown in Figure 3.2. It is now evident that the continuous Fourier transform associated with the motif is sampled at discrete reciprocal lattice points.

The amplitude and phase of the reflections observed from a crystalline lattice of molecules is given by the structure factor equation:-

$$F(hk\ell) = V \int_0^V \rho(\underline{r}) \exp 2\pi i (hx + ky + \ell z) \, dV_{\underline{r}} \qquad 3.11$$

and conversely the electron density can be obtained from:

$$\rho(r) = \frac{1}{V} \sum_h \sum_k \sum_\ell F_{hk\ell} \exp - 2\pi i (hx + ky + \ell z) \qquad 3.12$$

The electron density function $\rho(\underline{r})$ is in principle a continuous one, but it can be replaced by an 'atomic scattering factor', f_j, pertaining to a given atom, j. Equation 3.11 then becomes:-

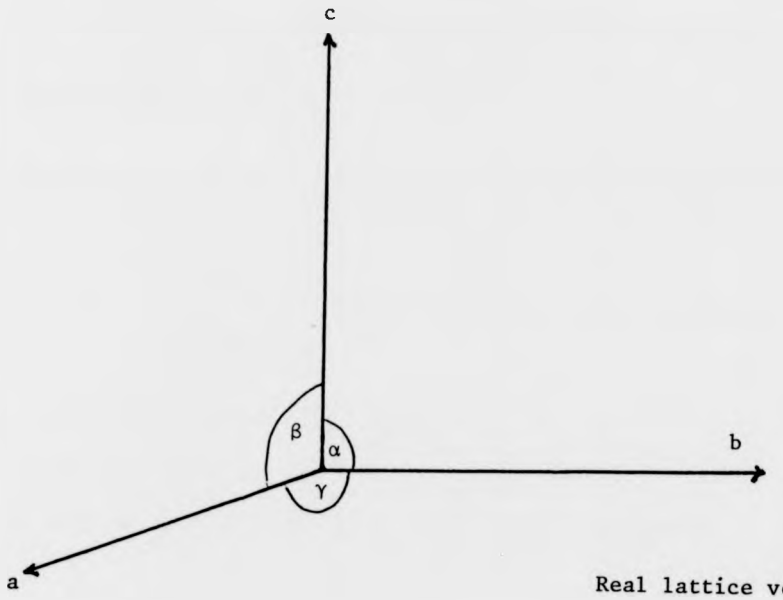
$$F(h,k,\ell) = \sum_{\text{all } j} f_j \exp 2\pi i (hx_j + ky_j + \ell z_j) \qquad 3.13$$

where the summation extends over all the atoms in the unit cell.

A complete wave mechanical treatment of the scattering process shows that the coherently scattered radiation from an atom can be found from first principles by assuming that the electronic charge is distributed about and not located at,a point. The wave function, $\psi(\underline{r})$ of an atomic electron is related to the charge distribution by

$$\rho = |\psi|^2 \qquad 3.14$$

and if it is assumed that the charge distribution is spherically symmetric



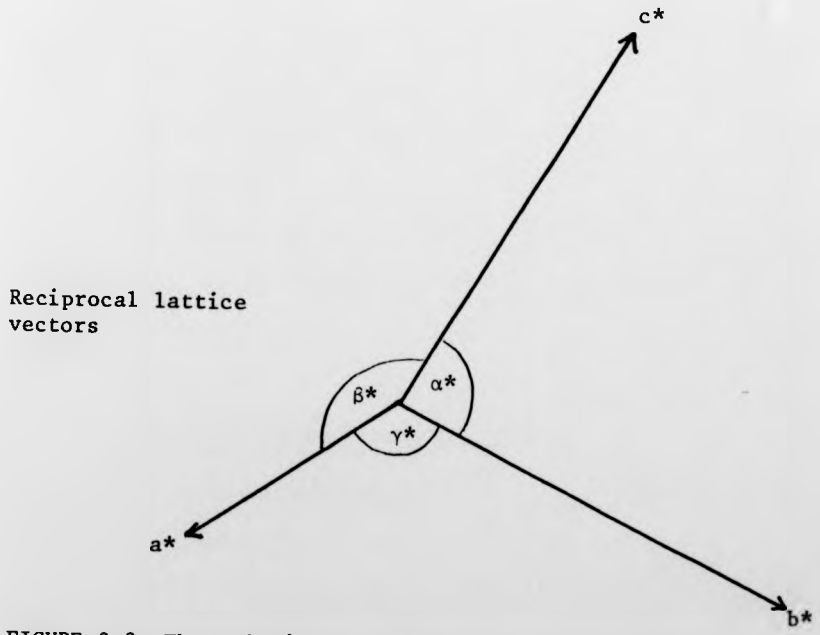


FIGURE 3.2 The relationship between the base vectors of the real and reciprocal lattices.

then the electron density distribution can be described by:-

$$\rho(r)r^2 \sin\psi \, dr \, d\psi \, d\phi \qquad 3.15$$

(see Figure 3.3). The total amplitude of coherently scattered radiation is:

$$f_j(\Delta\underline{K}) = \int_{r=0}^{\infty}\int_{\psi=0}^{\pi}\int_{\phi=0}^{2\pi} \rho(r)r^2 \exp \Delta Kr \cos\psi \sin\psi \, dr \, d\psi \, d\phi \qquad 3.16$$

where as shown in Figure 3.3, the coordinate system is arranged so that $\Delta\underline{K}$ coincides with the axis from which ψ is measured. Equation 3.16 can be simplified since for every point (r, ψ, ϕ) there is another $(r, \pi-\psi, \phi+\pi)$ and hence:-

$$f_j(\Delta\underline{K}) = \int_{r=0}^{\infty}\int_{\psi=0}^{\pi}\int_{\phi=0}^{2\pi} \rho(r)r^2 \cos(2\pi r\Delta K\cos\psi) \sin\psi \, dr \, d\psi \, d\phi \qquad 3.17$$

$$= 4\pi \int_{r=0}^{\infty} \rho(r) \, r^2 \, \frac{\sin\Delta Kr}{\Delta Kr} \, dr \qquad 3.18$$

However,

$$f_j(0) = 4\pi \int_{r=0}^{\infty} \rho(r) \, r^2 \, dr \qquad 3.19$$

which shows that at zero scattering angle the scattering factor is equal to the number of atomic electrons. As $\rho(r)$ increases f_j falls off in the manner depicted by Figure 3.4.

3.3 DIFFRACTION FROM HELICAL MOLECULES

An infinite smooth helix, $h_\infty(r,\theta,z)$ as shown in Figure 3.4, can be regarded mathematically as the convolution of a single turn of the helix $h_1(r,\theta,z)$ with a point lattice $L(x,y,z)$ having the periodicity of one pitch length along the z-axis

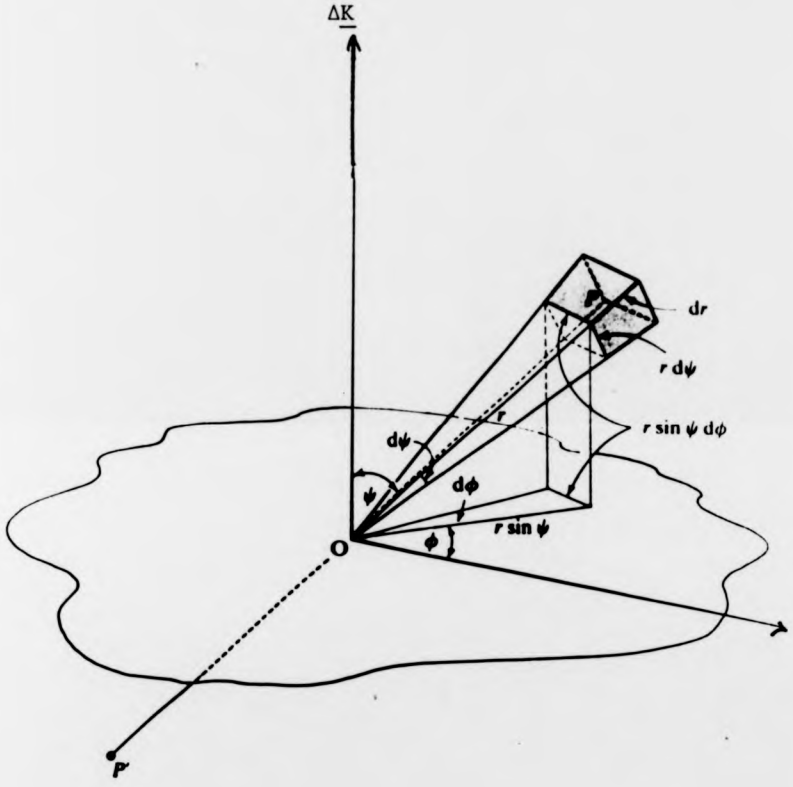


FIGURE 3.3 An elemental volume section seen in spherical polar coordinates

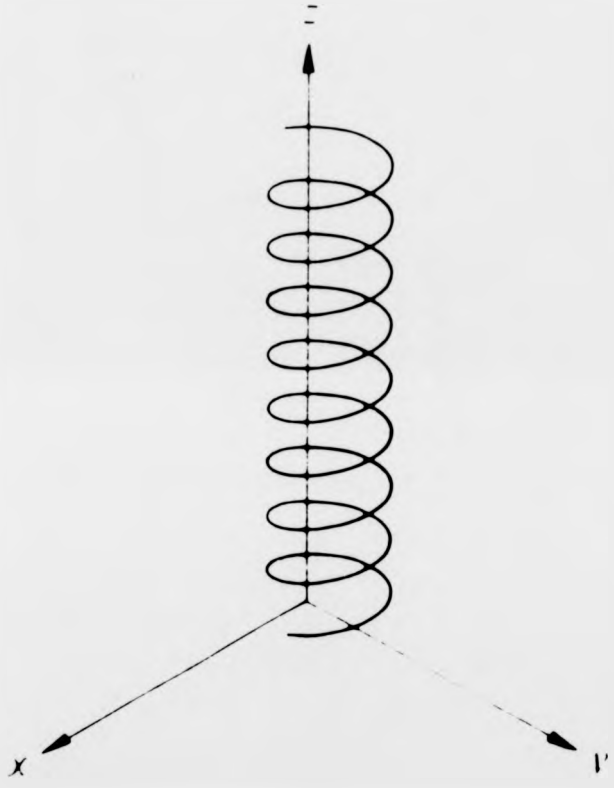


FIGURE 3.4 A long smooth helix being infinitessimally thin and having constant radius and scattering power.

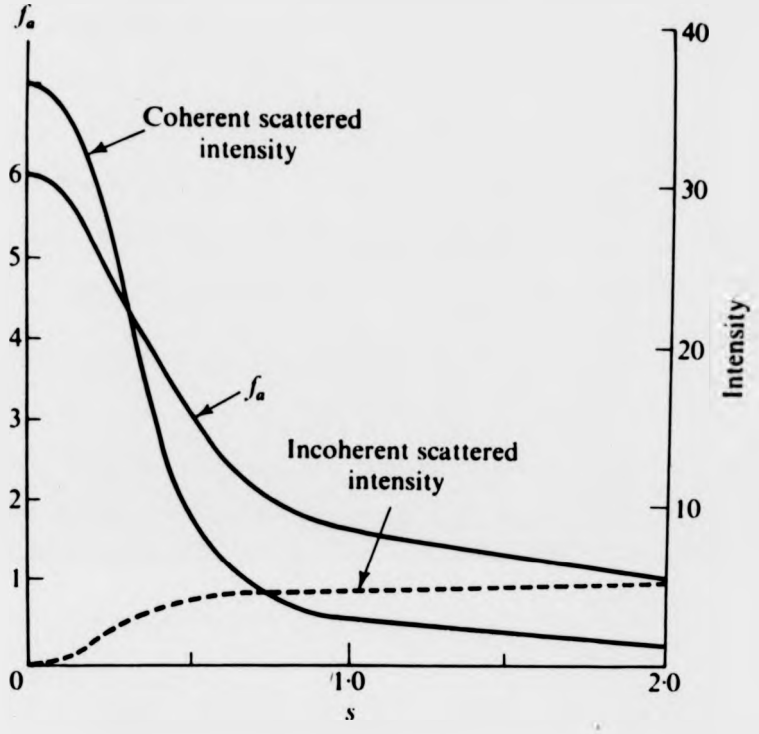


FIGURE 3.4 The atomic scattering factor, f_c, for the carbon atom. The corresponding coherent and incoherently scattered intensities are also shown. The quantity S is related to ΔK by $S = \Delta K/2\pi$.

i.e.

$$h_\infty(r,\theta,z) = h_1(r,\theta,z) \ * \ L(x,y,z) \qquad 3.20$$

so the scattered amplitude will be given by:-

$$T(h_\infty) = T(h_1) \ . \ T(L) \qquad 3.21$$

The point lattice L(x,y,z) can be expressed as:-

$$L(x,y,z) = \delta(x)\ \delta(y) \sum_{n=-\infty}^{n=\infty} \delta(z-nP) \qquad 3.22$$

and its Fourier transform is

$$T\ (L(x,y,z)) = \sum_n \delta \left(K_z - \frac{2n\pi}{P} \right) \qquad 3.23$$

so diffraction from an infinite smooth helix is confined to layer planes separated by ΔK_z where:-

$$\Delta K_z = \frac{2\pi}{P} \qquad 3.24$$

as shown in Figure 3.5.

The Fourier transform of such a continuous helix is given (in cylindrical polar coordinates) by:-

$$T\left[h_\infty(r,\theta,z)\right] = \sum_n r_0 \exp\left[in(\psi + \frac{\pi}{2} - \theta_0)\right] J_n(2\pi\xi r_0) \qquad 3.25$$

where the helix has radius r_0, pitch P and where z =0 when $\theta = \theta_0$. The cylindrical polar coordinate system applying to both real and reciprocal space are defined in Figure 3.6. The Bessel function orders that contribute to the summation in the above equation are determined by the helical selection rule:-

$$n = \frac{\ell - Nm}{K} \qquad 3.26$$

where N is the number of repeat units in K turns of the helix and m is an integer.

The diffraction from a continuous helix contains only those Bessel functions for which m = 0 in equation 3.26 (Cochran, Crick and Vand, 1952), which means that J_ℓ is the only Bessel function contributing to the ℓth layer plane. Such diffraction is schematically represented in Figure 3.7, and is seen to have cylindrical symmetry about the meridian. It also has mirror symmetry about the equator. The pattern has a highly characteristic 'cross' appearance due to the fact that the argument of the first peak of each function increases with increasing order. The cylindrically averaged intensity for a given layer line takes the form:-

$$I_\ell(\xi) = J_\ell^2(2\pi\xi r) \qquad 3.27$$

A discontinuous helix, $h_d(r,\theta,z)$, as depicted in Figure 3.8, can be

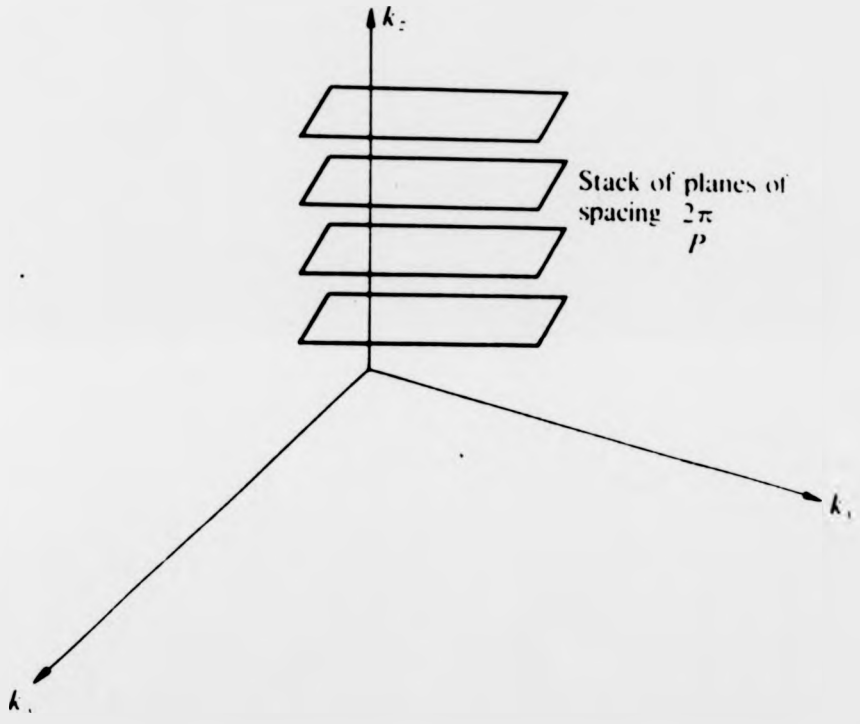


FIGURE 3.5 The Fourier transform of the function L(x,y,z) is a set of equally spaced planes in reciprocal space

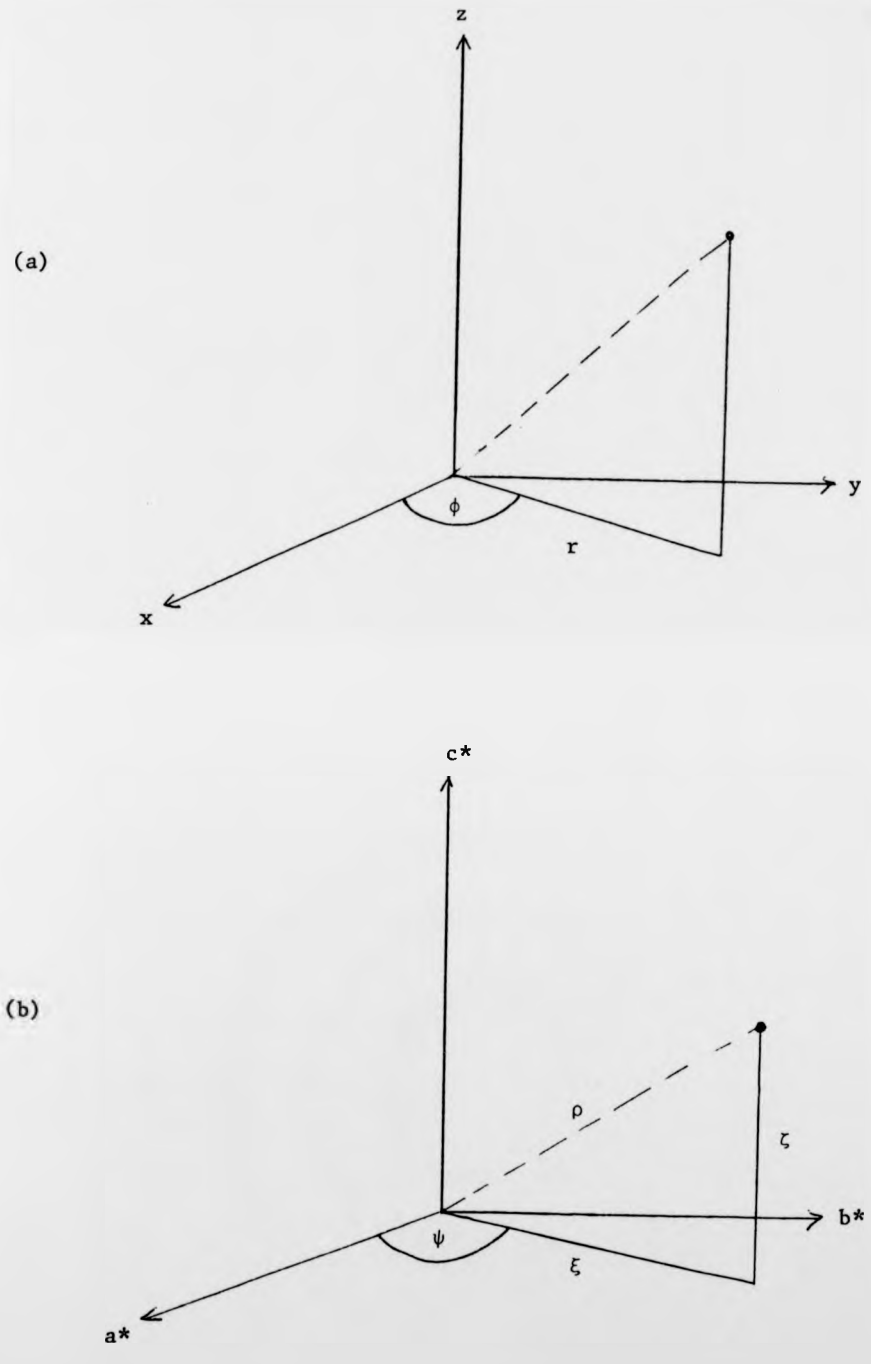
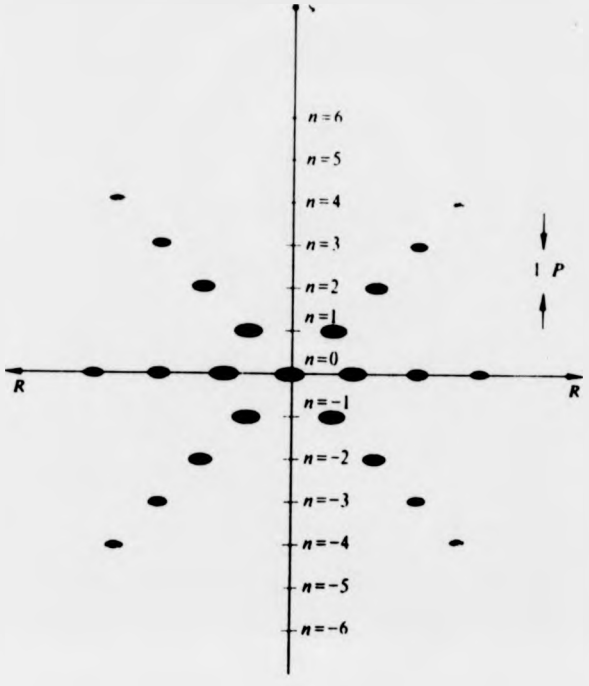
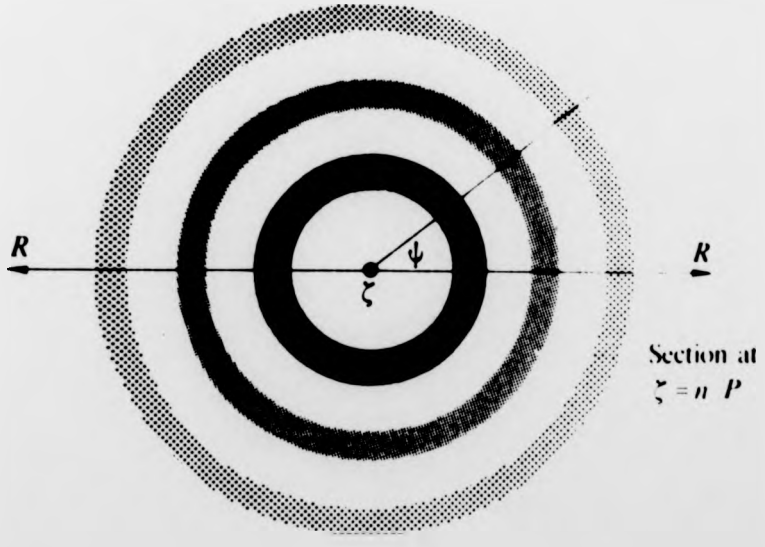


FIGURE 3.6 (a) Real space coordinate system used to describe helix

(b) Reciprocal space coordinate system



(a)



(b)

FIGURE 3.7 (a) the characteristic 'cross' diffraction from a continuous helix

(b) diffraction from a smooth helix has cylindrical symmetry about the K_z axis in reciprocal space.

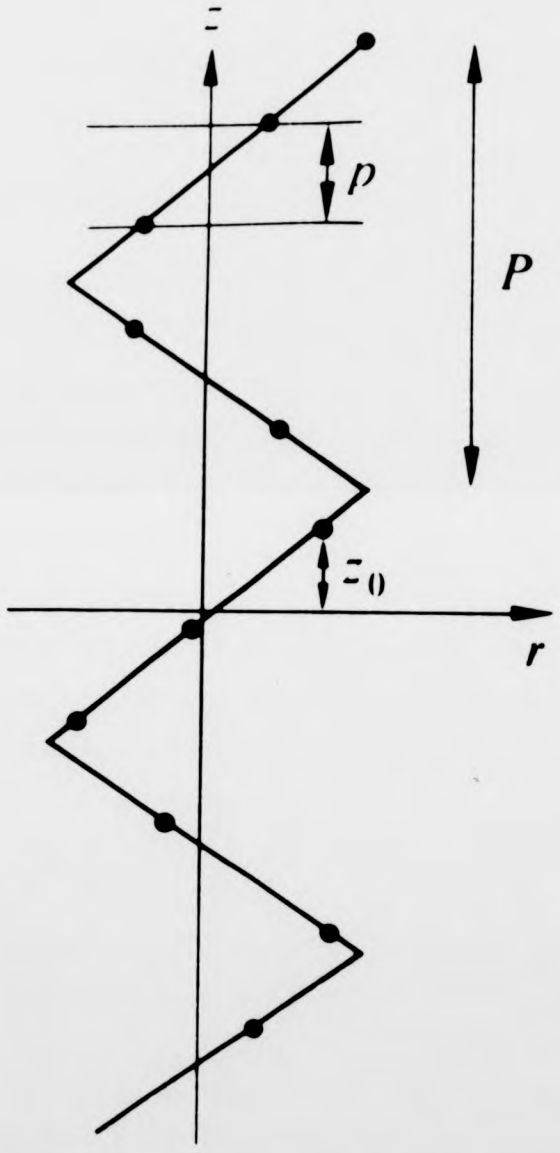


FIGURE 3.8 A 'discontinuous' helix composed of a helical array of points (seen in side elevation)

expressed as the product of a smooth helix function $h_\infty(r,\theta,z)$ with a point lattice $\ell(x,y,z)$ having a regular spacing along the helix axis, i.e.

$$h_d(r,\theta,z) = h_\infty(r,\theta,z).\ell(x,y,z) \qquad 3.28$$

and hence by the convolution theorem,

$$T\left[h_d(r,\theta,z)\right] = T\left[h_\infty(r,\theta,z)\right] * T\left[\ell(x,y,z)\right] \qquad 3.29$$

The function $T\left[\ell(x,y,z)\right]$ involves a function of the form $\delta(K_z - \frac{2m\pi}{P})$ and represents an array of δ functions along the K_z axis spaced at intervals of $^{2\pi}/_{P}$. In the case of a discontinuous helix, m is permitted to take a continuum of values, and further 'cross' shapes are manifest in the diffraction pattern of such a structure, with their centres alligned along the K_z axis and separated by $^{2\pi}/_{h}$ (see Figure 3.9). Instead of only one Bessel function contributing to each layer line as in the case of a continuous helix, an infinite number contribute (subject to the helical selection rule). However, in practice, only the low order Bessel functions are significant.

The Fourier transform of a helical array of points is found to be:

$$T\left[h_d(r,\theta,z)\right] = \sum_{\ell} f_j\, r_0 \exp\{2\pi i \ell z_0/c\} \sum_{n} \exp\left[in\,(\psi + {}^{\pi}/_{2} - \theta_0)\right] J_n(2\pi\xi r_0) \qquad 3.30$$

The Fourier transform of a helical polymer consisting of identical monomeric units all constrained to helical symmetry is

$$\sum_{j} f_j r_j \exp\frac{2\pi i \ell z}{c} \quad \sum_{n} \exp\left[in\,(\psi + \tfrac{\pi}{2} - \theta_j)\right] J_n(2\pi\xi r_j) \qquad 3.31$$

in which the summation over n allows only those values permitted by the helical selection rule (equation 3.26) and the summation over j is taken over all atoms in the monomer.

In the case of a fully crystalline array of such molecules, the structure factors are given by

$$F(h,k,\ell) = \sum_{p}\sum_{n}\sum_{j} f_j\, J_n\,(2\pi r_j \xi) \exp\left[i\{n(\psi + \tfrac{\pi}{2} - \theta_j)\} + \frac{2\pi\ell z_j}{c}\right]$$

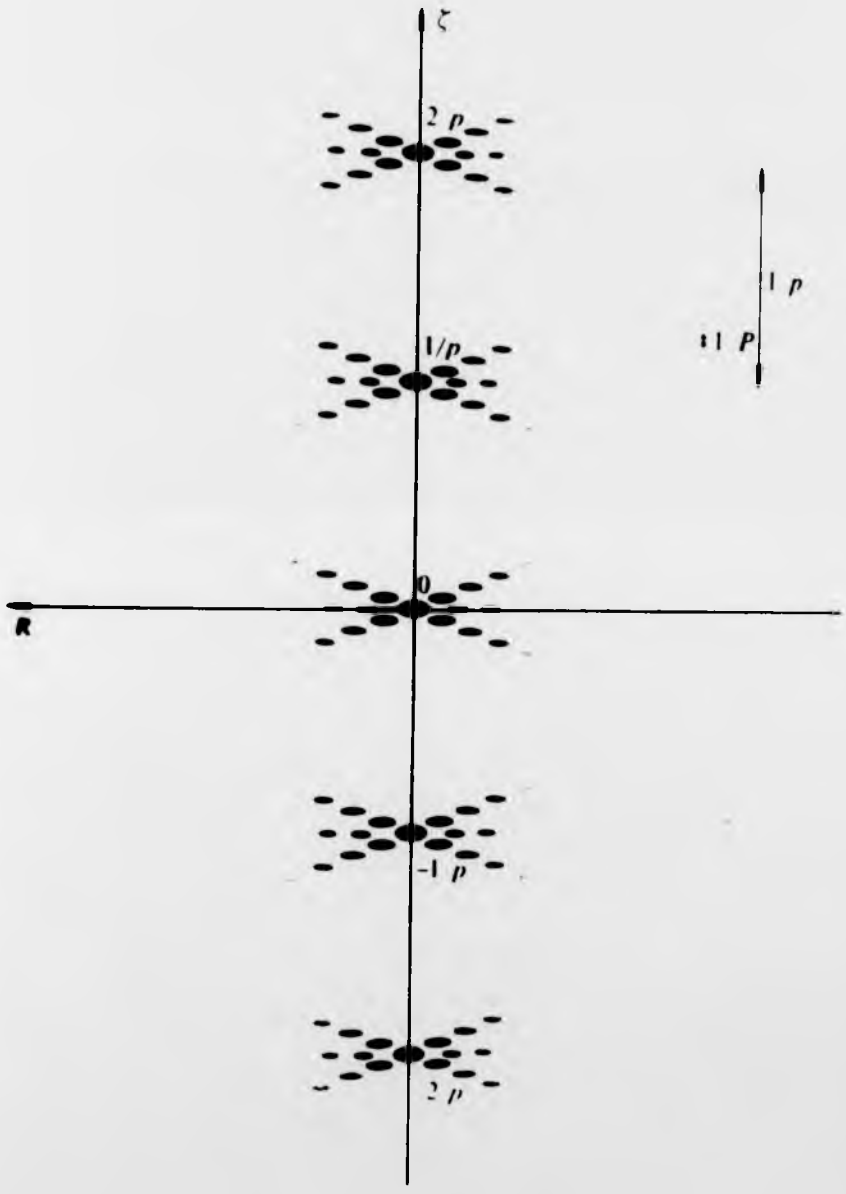


FIGURE 3.9 Diffraction from a helical array of points

$$\exp i\{2\pi(hx_p + ky_p + \ell z_p) - n\phi_p\} \qquad 3.32$$

where (x_p, y_p, z_p, ϕ_p) are (respectively) the fractional coordinates and azimuthal orientations of the p^{th} molecule in the unit cell.

3.4 THE CORRECTION TO CALCULATED MOLECULAR TRANSFORM FOR THE EFFECTS OF WATER

It has been estimated (Hodgson, 1969) that about $^1/_3$ of the contents of a nucleic acid crystallite is water. This figure will, of course, depend on the exact state of hydration of the sample in question. Water tends to interact with DNA in that it forms hydrogen bonds to accessible sites on bases and to sugar or phosphate oxygen atoms. It has been observed in many studies previously and will be seen in some pertaining to this work that such interaction is capable of influencing the conformation of nucleic acids in fibres. The water, which is commonly regarded as a 'sea' in which DNA molecules are immersed will tend to surround the hydrophilic regions (phosphates) of DNA more closely than the hydrophobic (sugar) regions. In terms of the diffraction observed from fibres of DNA and RNA, it will tend to have a subtractive effect on the molecular transform, especially in regions of the diffraction pattern that correspond to spacings greater than 7Å (Bragg and Perutz, 1952).

The effect of scattering by the solvent can be incorporated into calculated diffraction by assuming that the region occupied by water is composed of an electron gas having an average density (ρw) characteristic of the solvent (in the case of water, $\rho = 0.33\ e^{-\text{Å}^{-3}}$). The 'water weighting' correction of Langridge et al. (1960b) is calculated on the basis of Babinets principle and is applied to the individual atomic scattering factors rather than to the total molecular transform. The corrected scattering factor is given by:-

$$f'(\sin\theta) = f(\sin\theta) - V_j\sigma_w\phi(\sin\theta) \qquad 3.33$$

where f' and f are respectively the corrected and uncorrected scattering

factors, v_j is the volume displaced by the j^{th} atom in the molecule, and where ϕ is the Fourier transform of the region from which solvent has been displaced. If the atoms of the molecule are assumed to be hard spheres of radius r_o then $\phi(\sin\theta)$ is given by:

$$\phi(\sin\theta) = \frac{3}{\alpha} \quad (\sin\alpha - \alpha\cos\alpha) \qquad 3.34$$

where

$$\alpha = \frac{2\pi \sin\theta \, r_o}{\lambda} \qquad 3.35$$

However, a number of different approaches have been developed towards the evaluation of ϕ (Langridge et al., 1960b; Fuller, 1961; Fraser, MacRae and Suzuki, 1978; Campbell-Smith and Arnott, 1978) and all of these methods show significant variation in the molecular transform calculated therewith. This problem has recently been reviewed by Greenall (1982) who concludes that the best approach to the correct of molecular transform for the effects of water is to use the original scheme of Langridge et al. (1960b).

3.5 DISORDER AND DIFFRACTION FROM FIBRES

An 'oriented' and 'crystalline' fibre consists of many long chain molecules which pack so that their long axes are roughly parallel to the fibre axis. The fibre is made up of 'crystallites' (small compared to the length of the molecules) which are interspersed by amorphous regions. A representation of this sort of system is given in Figure 3.10. Diffraction patterns recorded from such fibres (good examples are the DNA conformations pertaining to Plates 1.1, 1.4 and 1.6) are similar to single crystal rotation photographs since the individual crystallites all have random azimuthal orientation. The geometry of this sort of diffraction is outlined in Figure 3.11. The reciprocal lattice points are drawn out to form 'rings' confined to layer planes, and it is evident that diffraction data will have cylindircal symmetry about the meridian. Furthermore, these rings are broadened into 'belts' because of the average angular displacement

Figure 3.10 Representation of helices arranged in crystallites
(a) perfect crystallites disoriented about the fibre axis
(b) disordered crystallites (from Vainshtein, 1966)

of crystallites from the fibre axis. This means that diffraction peaks as specified by the Laue equations 3.6 become arcs, the exact shape of which will depend on the nature of the geometrical intersection of these layer rings with the Ewald sphere. The shape of the diffraction spots will also be affected by the average size of the crystallites constituting the fibre.

A variety of different forms of disorder may prevail in fibres. Samples, even if highly crystalline, may be disoriented, in which case diffraction diagrams are characterised by long arcs which in the limit become circles. Such behaviour is thought to be indicative of low molecular weight material (Leslie et al., 1980). Other effects which are noticeable in DNA fibres are 'shift', 'rotation' and 'screw' disorders. These all tend to disrupt the three dimensional lattice whilst preserving the c-axis ordering of the fibre (i.e. in the limit diffraction is discrete only for $\ell=0$). Pure 'shift' disorder causes diffuseness of the $\ell \neq 0$ reflections: this effect becomes more pronounced as ℓ increases, and eventually gives continuous diffraction over whole layer lines. The effect of shift is shown in

FIGURE 3.11 The reciprocal lattice 'rings' associated with a fibre containing randomly oriented crystallites (from Vainshtein, 1966).

Figure 3.12.

If the packing is close as in the case of C-DNA (Marvin, 1961), then the ridges of one molecule may intercalate with the grooves of nearest neighbours. Rotational disorder of molecules in such an array may then have two effects: firstly, it can disrupt the crystallinity of the lattice, and secondly, it can also introduce shift disorder of the type mentioned above. In the most general case both occur.

'Screw' disorder may also disrupt the three dimensional coherence of the fibre: However, if the molecules have a random displacement by a

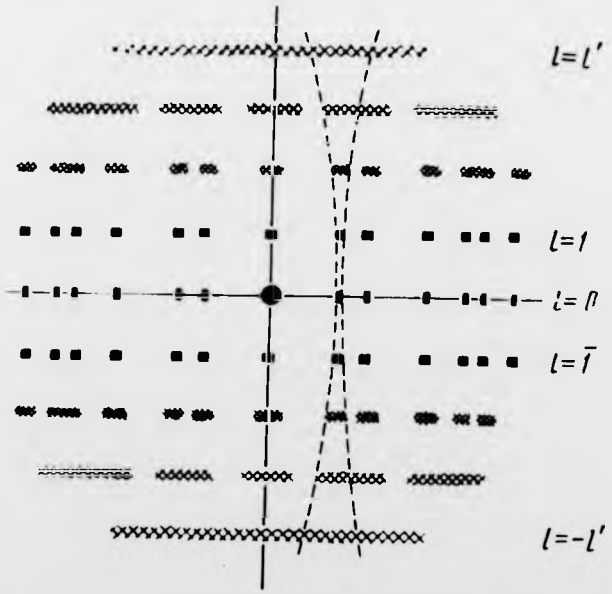


FIGURE 3.12 Shift distortion causing diffuse diffraction for $\ell \neq 0$. Layer lines eventually become continuous for $|\ell| > \ell'$ (from Vainshtein, 1966)

fixed amount, the fibre diagram may show spots on some layer lines and continuous streaks on others (Marvin, 1961; chapter four of this work).

The most disordered system is one in which the fibre is made up of an amorphous assembly of chain molecules. Such fibres usually show a few diffuse reflections on the meridian (relating to intra-molecular interference) and on the equator (relating to intermolecular interference).

3.6 THE APPROACH TOWARDS THE SOLUTION OF NUCLEIC ACID STRUCTURE FROM FIBRE DIFFRACTION DATA

As has become apparent from the theoretical considerations outlined in section 3.2 and 3.3, it is necessary to know both the amplitude and phase of diffracted x-radiation in order to directly solve a molecular structure. Since in any diffraction experiment the observed quantity is the intensity, it is evident that the amplitude of $F(hk\ell)$ in equation 3.32 (in the case of crystalline diffraction) can be determined but the phase is lost. This problem, known as the 'phase problem' is one to which many aspects of crystallography are directed. A number of methods are used to deal with the phase problem, amongst which are isomorphous replacement methods, heavy atom techniques and anomalous dispersion methods. It has been suggested by Marvin et al. (1966) that the phosphorous atoms of nucleic acids could be treated as heavy atoms and used to determine phases, but unfortunately the Fourier difference resolution of such a method is limited to $\sim 3\text{Å}$ for reasons outlined in section 3.4.

In the absence of any definitive methods to enable phase determination, the strategy that is adopted towards structure analysis in the case of fibre diffraction is a trial and error one that involves a prior knowledge of the likely overall stereochemistry of the molecule. Models are usually adjusted by considering the variation in position of the three main scattering groups (phosphate, sugar and base) until good agreement is obtained firstly on the equator, secondly on the inner part of the lower layer lines and lastly in the regions of the pattern for which ρ is largest. Molecular models are compared in terms of their overall agreement with the

observed diffraction by computing agreement indices R and R' as described in Chapter 2 and equations 2.1, 2.2.

Simplifying assumptions that are made during the course of such computation are firstly that the atoms are all constrained to the symmetry of a perfect helix, and secondly, that the effect of stabilizing counter-ions is negligible. The first assumption is reasonable within the available resolution of these fibre diagrams, and so is the second given that the ions in question are sufficiently light. However, ions such as caesium and rubidium are likely to have a profound effect on the observed diffraction. This effect can, of course, be put to good use in that a comparison of intensities observed from light and heavy atom salts of DNA may be used to determine the likely position of ions about the helix (Bartenev et al., 1983).

CHAPTER FOUR

THE CONFORMATION OF C'-DNA

4.1 INTRODUCTION

The C conformation was first described by Marvin et al., (1958, 1961). It was observed in X-ray fibre diffraction photographs as a semi-crystalline form of the lithium salt, when the prevailing humidity of the fibre environment was between 44% and 57%. Marvin et al. also reported that the appearance of the C-form depended on the amount of LiCl which had been precipitated with the DNA. For fibres containing less than 1% by weight of chloride, the hexagonal form was observed at RH's of 66%, or less. For fibres which contained between 1% and 6% by weight of chloride, an orthorhombic semi-crystalline C form was observed at 44% relative humidity (the same fibre was found to give a fully crystalline B form at 66% RH).

For many years the C form was generally regarded as being characteristic of Li DNA, and at best a poorly favoured form of Na DNA. The molecular conformation of C-DNA was seen as a distorted B-form and as a structural peculiarity associated with the Li^+ counterion, which has limited biological significance. However, infrared linear dichroism studies by Brahms et al. (1973) showed that a C-like conformation could occur in oriented films of Na DNA which contained very low amounts of NaCl. Arnott and Selsing (1975) and Leslie et al. (1980) reported, on the basis of X-ray fibre diffraction studies, that the C form of Na DNA is stable at salt concentrations and also at relative humidities which are intermediate between those that favour the A conformation and those that favour the B conformation.

Zimmerman and Pheiffer (1980) observed X-ray diffraction photographs of the C type from fibres of Na DNA immersed in t-butanol/water mixtures. Rhodes et al. (1982) have recently shown that the C form can be routinely observed for a wide variety of naturally occuring DNA's, and also the synthetic polynucleotide poly d(A-C).poly d(G-T) if both the salt content of the fibre and the relative humidity of the fibre are sufficiently low. Mahendrasingam et al. (1983, 1984) have subsequently reported that the C form is observed under comparable conditions for Na poly d(A-T).poly d(A-T) and for the sodium salt of poly d(G-C).poly d(G-C). Such observations now establish the C form as a major conformational possibility for the DNA double helix in biological environments.

The X-ray diffraction patterns that have so far been obtained indicate that there is a marked amount of variation associated with the structure of C-DNA (see Appendix). Leslie et al. (1980) have surveyed the observation of the C form in native and synthetic DNA double helices. They emphasise that the C form should be regarded as a family of closely related structures. The semi-crystalline forms from the lithium salt of naturally occuring DNA's are non-integral with approximately $9\frac{1}{3}$ nucleotide pairs per continuous helix pitch of 31Å and are classified as 28_3 helices and designated C. Fully crystalline C forms observed from Li poly (dG-dG-dT).poly (dA-dC-dC) and in less well defined patterns from Li poly (dA-dG-dC).poly (dG-dC-dT) were found to be integral with nine residues per helix pitch of 29.5Å and are classified as 9_1 helices and designated C'. Na poly (dA-dG).poly (dC-dT) was observed to adopt a 9_2 helix with a pitch of 58.2Å in which the repeating unit was a dinucleotide: this form was designated C".

Arnott and Selsing (1975) reported a reappraisal of the model of the C conformation proposed by Marvin et al. (1961), focussing in particular on the conformational differences between this C model and that assumed for

the B form of DNA. They generated models of the C type which more closely resembled the B conformation and compared the diffracted intensity calculated for models with 28_3 and 9_1 helical symmetry with that observed by Marvin et al. Arnott and Selsing claimed that the diffracted intensities calculated from their 9_1 helix are almost indistinguishable from those of the model described by Marvin et al. and further stated that they regarded their 28_3 model as the preferred model for the C set of structures.

In this study, a critical comparison of the data observed (both by Marvin et al., and in this laboratory) with various calculated data is undertaken. A careful examination of the crystalline x-ray diffraction patterns obtained in a variety of different conditions has been used to assess the complex packing arrangements that C'-DNA has been found to adopt. Three types of model have been considered and compared in terms of their agreement with the observed diffraction: firstly, the original Marvin 28_3 model and a similar structure distorted to meet the requirements of 9_1 helical symmetry; secondly, the models proposed by Arnott et al; and thirdly, a new model which differs significantly from either of the two above.

4.2 EXPERIMENTAL

Poly (dA-dC).poly (dG-dT) was synthesised for this study by Dr. J. Brahms at the University of Paris. Fibres were prepared by Drs. N.J. Rhodes and A. Mahendrasingam. In the final stages of purification the material was precipitated with ethanol from 0.1M LiCl and carefully washed in order to remove as much excess salt as possible. Fibres were drawn using standard techniques (Chapter Two of this work; Fuller et al., 1967) and x-ray patterns were recorded at relative humidities of the fibre environment ranging from 33% to 98%. Photographs were taken using both conventional Cu K-α x-ray sources (see section 2.3.1), and the rather more

rapid synchrotron radiation source (SRS) at Daresbury Laboratory.

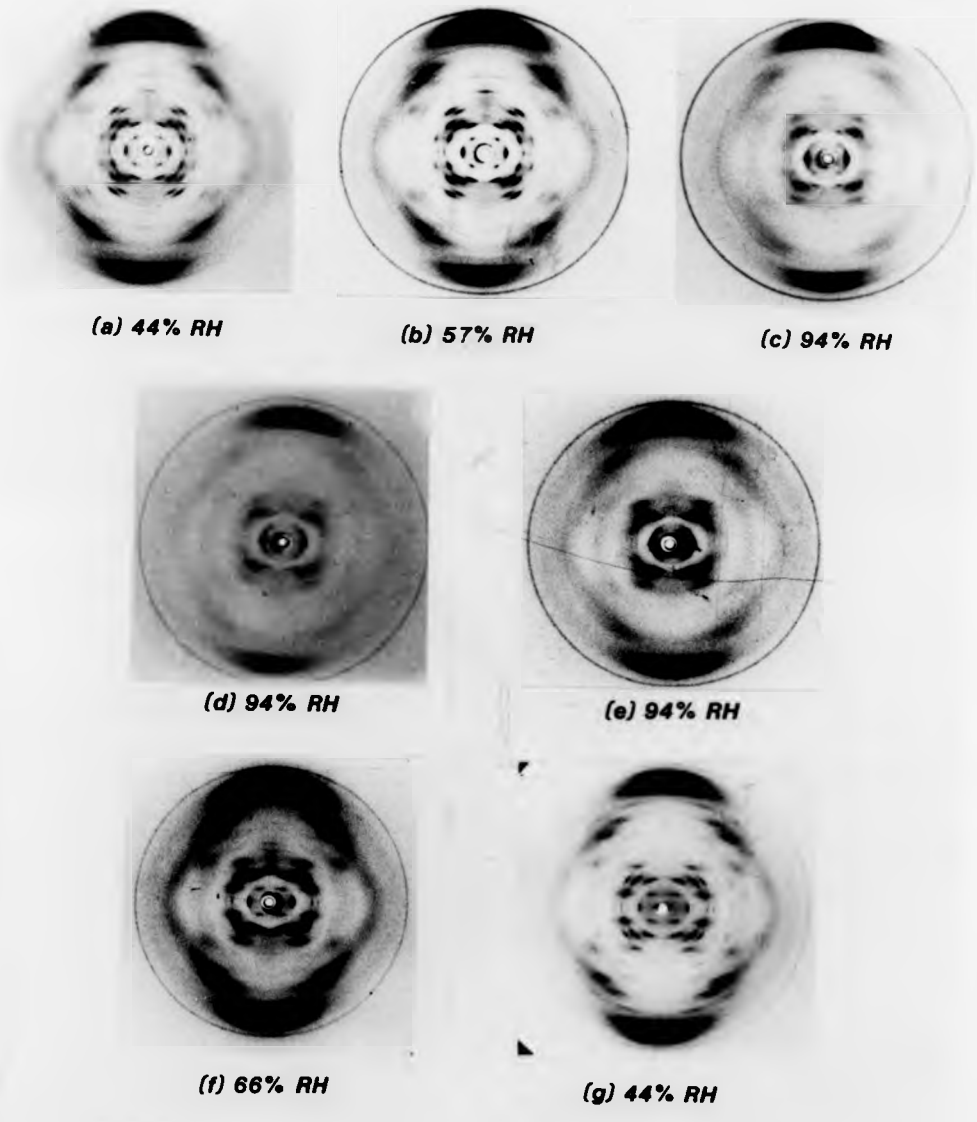
For the purposes of this analysis, intensities were determined from data produced by a Joyce-Loebl microdensitometer. A number of the C' patterns were also digitised using the rotating drum scanning densitometer at Daresbury Laboratory (see section 2.4), but at the time of analysis the software required to reduce this data was not fully operative and has yet to be used for intensity determination.

Molecular model building was undertaken in the manner of section 2.5, initially using scaled wire models and finally using a modelbuilding program to refine structures in terms of standard stereochemistry.

4.3 THE OBSERVATION OF X-RAY DIFFRACTION PATTERNS

A preliminary visual observation of the C form data obtained over a wide range of relative humidities (between 33% and 98%) revealed well defined patterns throughout the humidity range over which they were studied (see Plate 4.1). The highest degree of crystallinity was usually in the range 57% - 75% relative humidity (see Plate 4.2). Changes were observed in the detailed intensity distribution of the C patterns as a function of relative humidity, and it was thought that whilst some of these changes were typical of sampling effects due to variation of the lattice parameters as function of water content of the fibre, others (such as the variation in relative intensity of the eighth and ninth layer lines) indicated a change in the molecular conformation as a function of relative humidity.

Reduction of the relative humidity of the environment of a fibre which had assumed the semi-crystalline B form resulted in a transition (typically at 92% RH) to a C form which gave a pattern similar to the original crystalline form except that some crystallinity was lost and there were streaks rather than Bragg reflections on lower layer lines of odd order (see plate 4.1(g)). C patterns of this general form continued to be

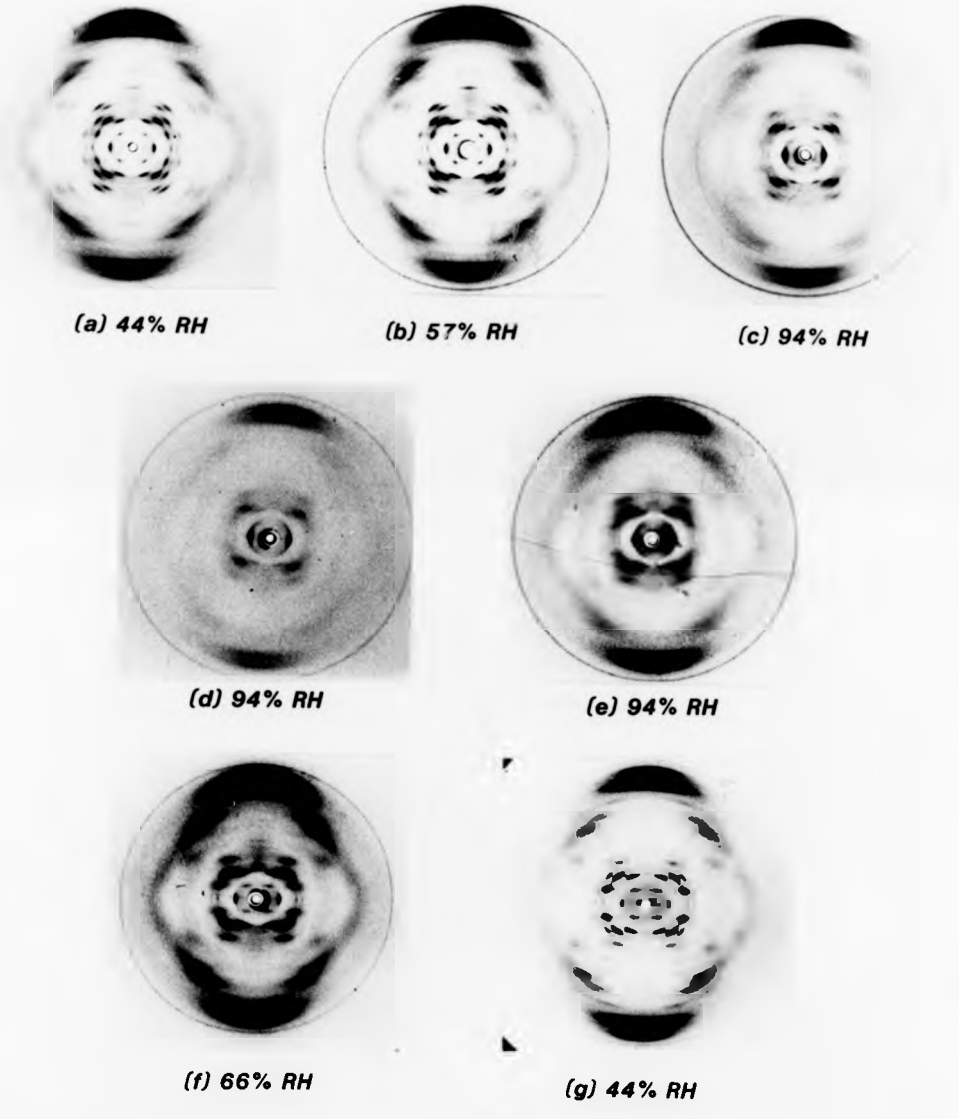


(a) 44% RH *(b) 57% RH* *(c) 94% RH*

(d) 94% RH *(e) 94% RH*

(f) 66% RH *(g) 44% RH*

PLATE 4.1:X-ray diffraction photographs obtained from a fibre of Li poly(dA-dC).poly(dG-dT) at various relative humidities.



(a) 44% RH (b) 57% RH (c) 94% RH

(d) 94% RH (e) 94% RH

(f) 66% RH (g) 44% RH

PLATE 4.1:X-ray diffraction photographs obtained from a fibre of Li poly(dA-dC).poly(dG-dT) at various relative humidities.

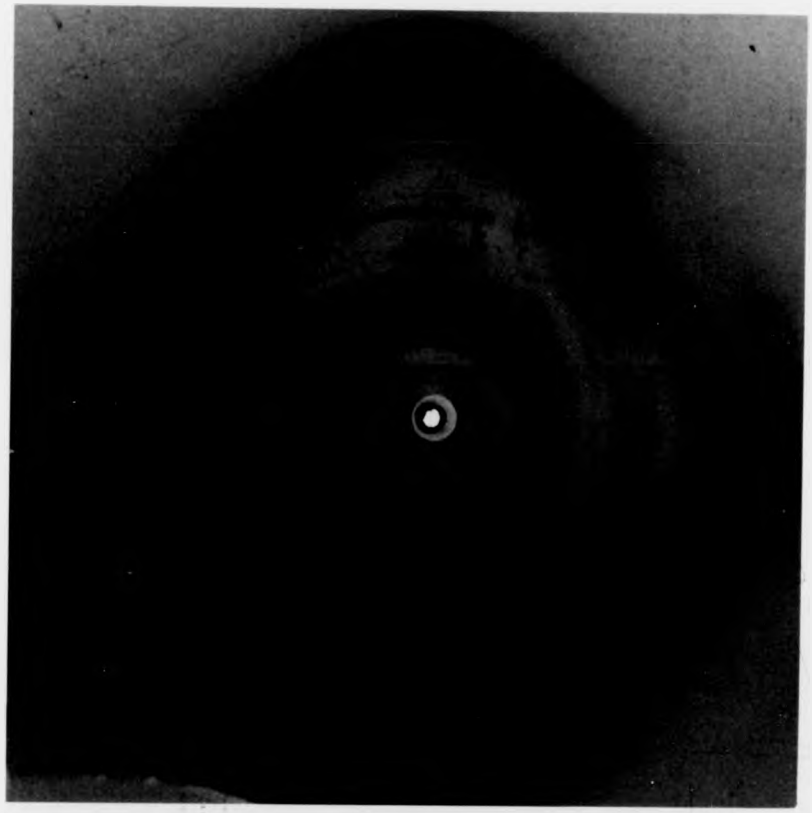


PLATE 4.2 The most crystalline diffraction pattern from the set shown in Plate 4.1 (RH = 66%) (pattern recorded by N.J. Rhodes, 1983).

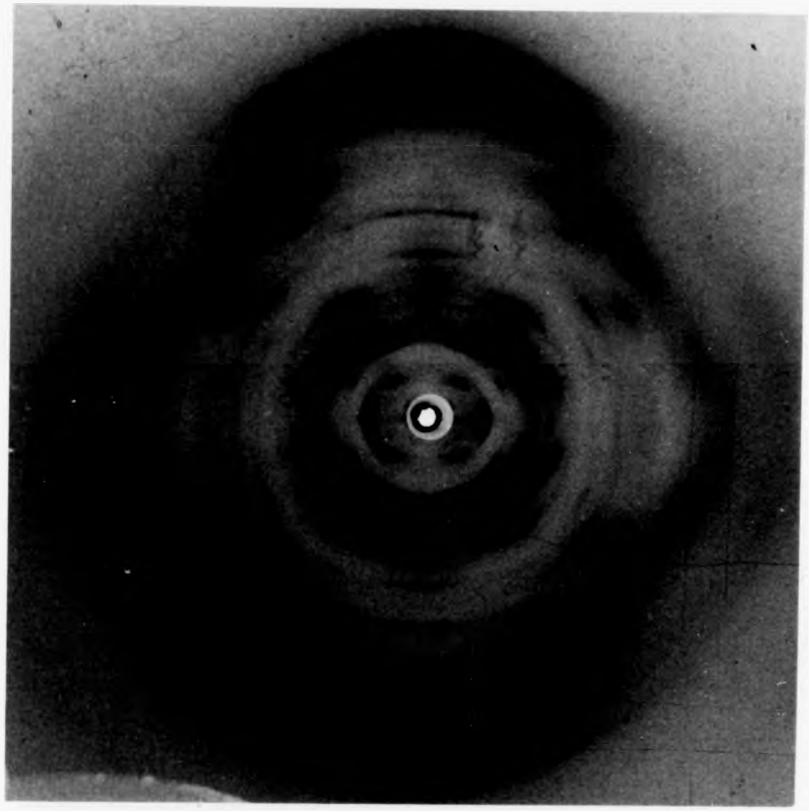


PLATE 4.2 The most crystalline diffraction pattern from the set shown in Plate 4.1 (RH = 66%) (pattern recorded by N.J. Rhodes, 1983).

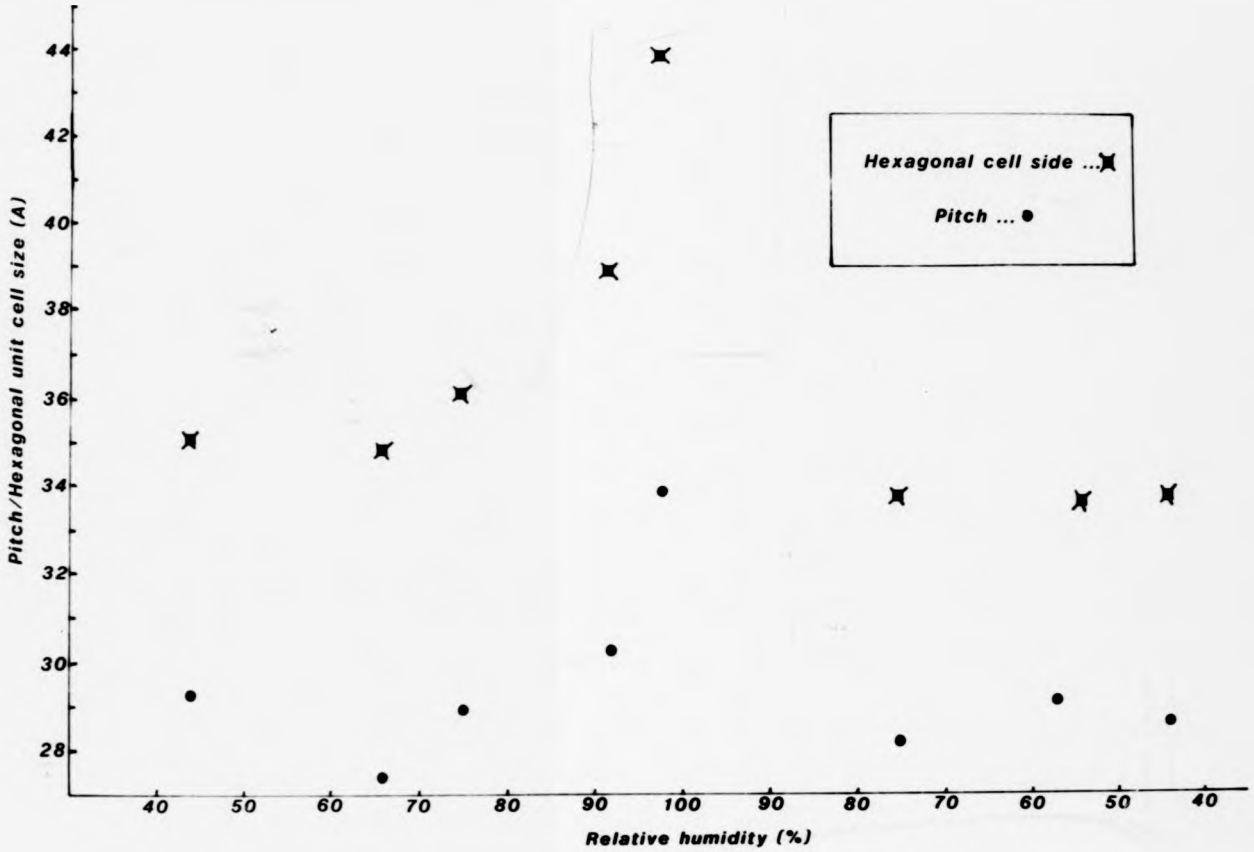


FIG 4.1: These measurements were taken from the sequence summarised in plate 4.1 and illustrate the variation of lattice parameters with relative humidity.

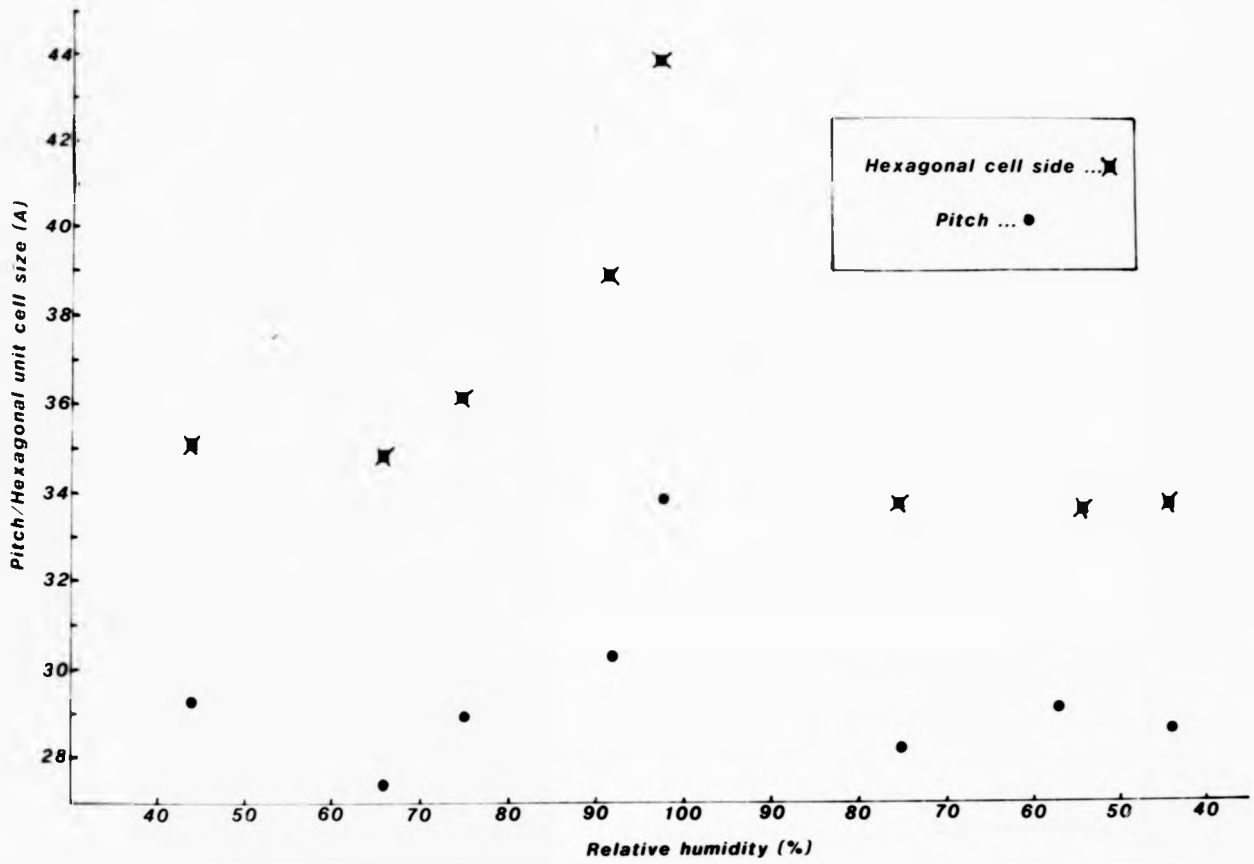


FIG 4 1: These measurements were taken from the sequence summarised in plate 4 1 and illustrate the variation of lattice parameters with relative humidity.

observed as the RH was reduced to 33%. If a fibre which had undergone the above mentioned sequence of transitions was rewet to form a gel, a new fibre could be pulled which was then fully crystalline. All of the C patterns showed uneven intensity distributions on (particularly) the first, second, third, fifth and sixth layer lines. A careful inspection of these layer lines as seen in Plates 4.2 and 4.4 shows rather dramatic differences between adjacent regions of reciprocal space which to a first approximation should produce similar diffracted intensities. Good examples are provided by a comparison of the 21, 30 and 22 reflections on the first layer line (see Plate 4.3). These packing effects are considered in detail in Sections 4.5 and 4.7.

4.4 THE LATTICE GEOMETRY AND ITS VARIATION WITH RELATIVE HUMIDITY

The patterns measured were all found to index onto hexagonal lattices (see Plate 4.3 and Tables 4.1, 4.2), the parameters of which varied depending on the prevailing relative humidity. The variation of lattice parameters with relative humidity for the sample associated with Plate 4.1 is shown in Figure 4.1. It seems likely that the increase of these lattice parameters with humidity is due to an increasing amount of water in the lattice as the RH of the fibre environment is increased. The parameters are, in general, similar to those reported by Leslie et al. (1980) for the C' form of Li poly (dG-dG-dT).poly (dA-dC-dC) (see Plate 4.5). They are, however, quite distinct from the hexagonal lattice with a = 22.1Å and c = 58.2Å which these authors observed from the C'' form of Na poly (dA-dG).poly (dC-dT). Leslie et al. have suggested that the assumption of 9_1 and 9_2 helices reflected the repetition of respectively a trinucleotide and a dinculeotide sequence. However, it has been observed in this work that Li poly (dA-dC).poly (dG-dT) has a unit cell characteristic of the 9_1 helix, and this suggests that the situation may well be more complicated than presumed by Leslie et al. Although the diffraction pattern shown in Plate

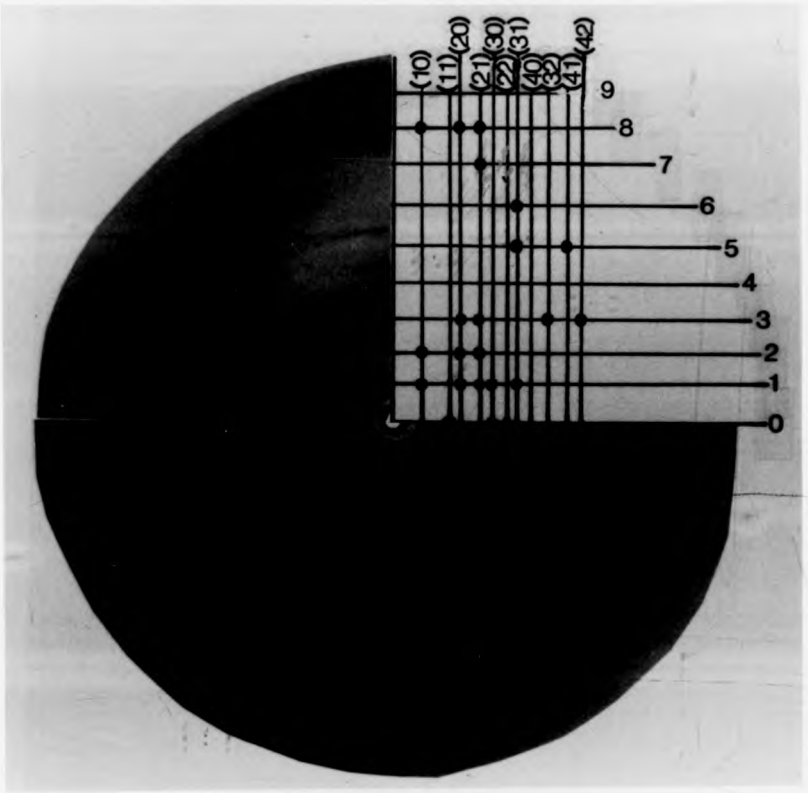


PLATE 4.3 The pattern shown in Plate 4.2 is here schematically indexed on a hexagonal grid having a=34A and c= 29.2A. The observed and calculated ρ values are given in Table 4.1

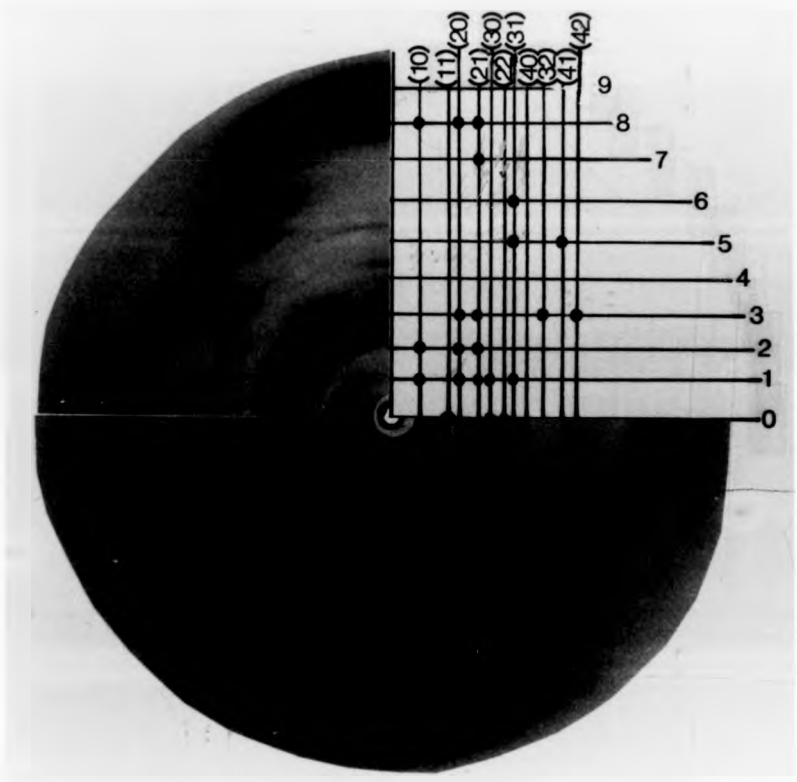


PLATE 4.3 The pattern shown in Plate 4.2 is here schematically indexed on a hexagonal grid having a=34A and c= 29.2A. The observed and calculated ρ values are given in Table 4.1

REFLECTION (HKL)	ρ_{obs} $\mathring{A}^{-1}$	ρ_{calc} $\mathring{A}^{-1}$
(110)	0.0586	0.0579
(300)	0.0998	0.1003
(220)	0.1162	0.1158
(101)	0.0472	0.0477
(201)	0.0746	0.0750
(211)	0.0942	0.0948
(301)	0.1056	0.1059
(311)	0.1235	0.1252
(102)	0.0758	0.0758
(202)	0.0950	0.0953
(212)	0.1119	0.1116
(203)	0.1202	0.1219
(213)	0.1333	0.1350
(323)	0.1795	0.1778
(423)	0.2035	0.2042
(315)	0.2083	0.2084
(415)	0.2288	0.2288
(316)	0.2374	0.2369
(217)	0.2542	0.2538
(108)	0.2740	0.2740
(208)	0.2798	0.2800
(218)	0.2880	0.2859
(009)	0.3103	0.3059

TABLE 4.1 The observed and calculated ρ values for the pattern shown in plates 4.2, 4.3. The lattice is hexagonal with a=b=34$\mathring{A}$ and c=29.9$\mathring{A}$.

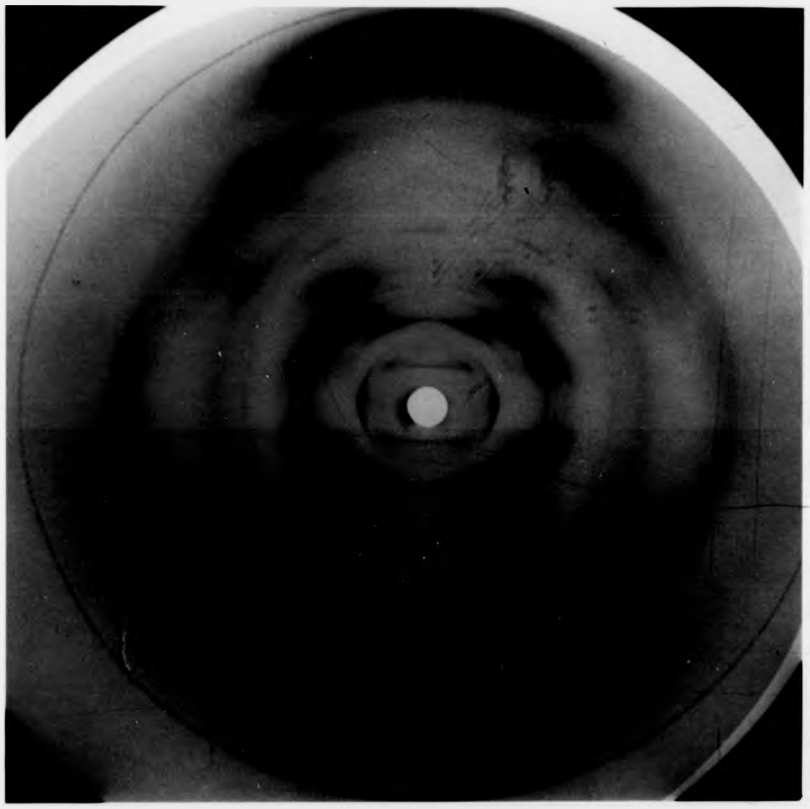


PLATE 4.4 Another C' pattern from a fibre having a slightly lower salt content. A comparison of Plates 4.4 and 4.3 (especially with reference to ℓ = 1, 3, 5, 6) shows that this small difference in salt content has caused significant variation in molecular packing. The observed and calculated ρ values for this pattern are given in Table 4.2. This picture was taken at the Daresbury SRS.

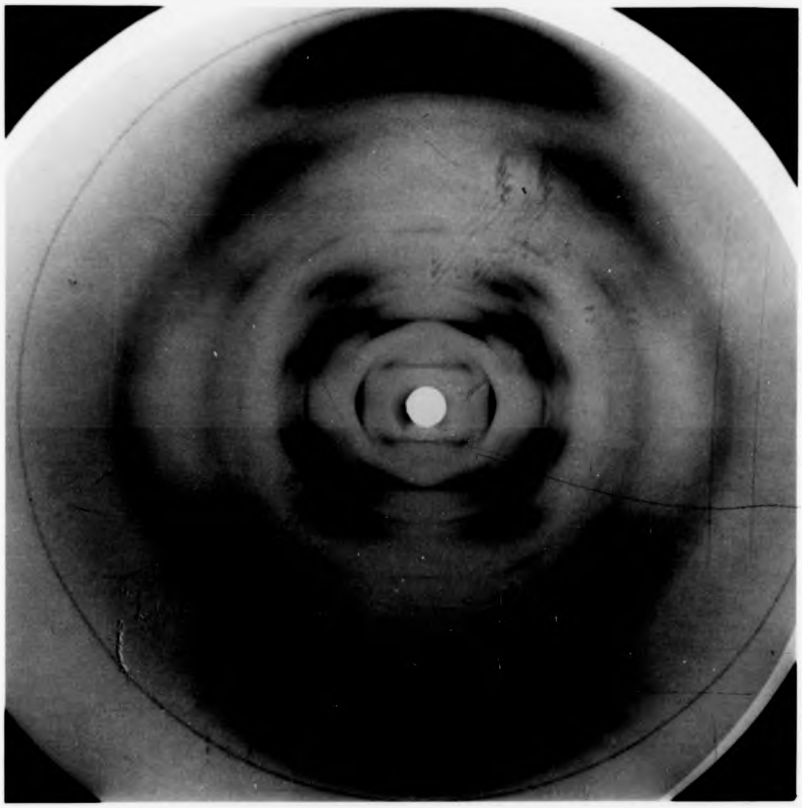


PLATE 4.4 Another C' pattern from a fibre having a slightly lower salt content. A comparison of Plates 4.4 and 4.3 (especially with reference to ℓ = 1, 3, 5, 6) shows that this small difference in salt content has caused significant variation in molecular packing. The observed and calculated ρ values for this pattern are given in Table 4.2. This picture was taken at the Daresbury SRS.

REFLECTION (HKL)	ρ_{obs} $\AA^{-1}$	ρ_{calc} $\AA^{-1}$
(110)	0.0644	0.0640
(300)	0.1101	0.1109
(220)	0.1274	0.1280
(410)	0.1683	0.1694
(101)	0.0506	0.0507
(211)	0.1040	0.1038
(301)	0.1162	0.1162
(221)	0.1326	0.1327
(102)	0.0784	0.0786
(202)	0.1015	0.1014
(212)	0.1200	0.1199
(322)	0.1760	0.1755
(103)	0.1106	0.1105
(113)	0.1223	0.1222
(203)	0.1279	0.1277
(213)	0.1427	0.1428
(323)	0.1937	0.1918
(423)	0.2238	0.2216
(214)	0.1706	0.1697
(305)	0.2069	0.2059
(315)	0.2181	0.2187
(316)	0.2493	0.2472
(406)	0.2569	0.2553
(108)	0.2785	0.2800
(208)	0.2880	0.2873
(218)	0.2957	0.2943
(009)	0.3110	0.3123

TABLE 4.2 The observed and calculated ρ values for the pattern shown in Plate 4.4. The lattice is hexagonal with a=b=31.2A and c=28.8A.

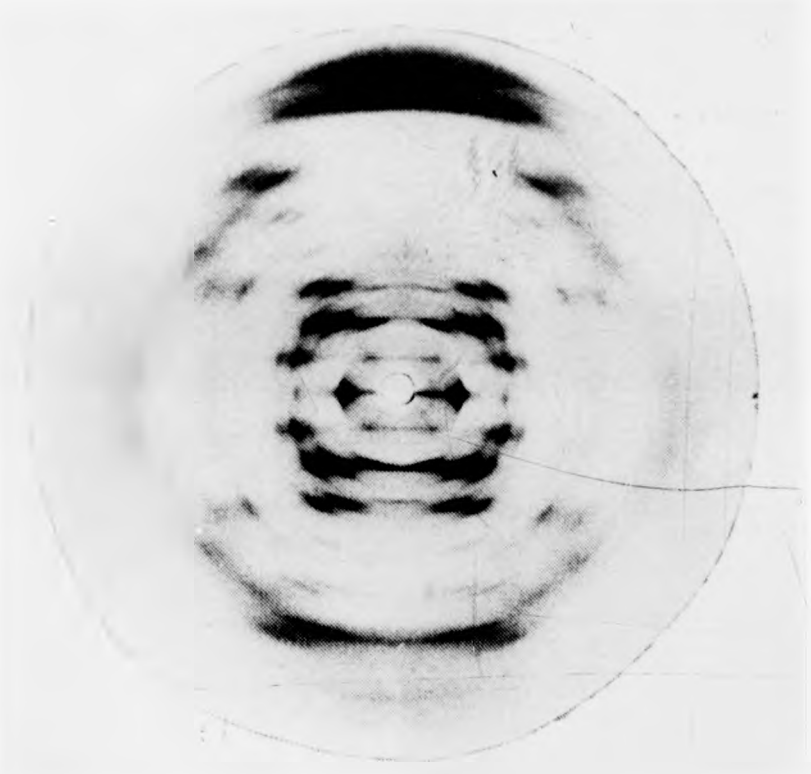


PLATE 4.5 The C' conformation from Li poly (dG-dG-dT).poly d(dA-dC-dC) as observed by Leslie et al. (1980). The lattice here is hexagonal such that a = b = 33.2 and c = 29.5Å.

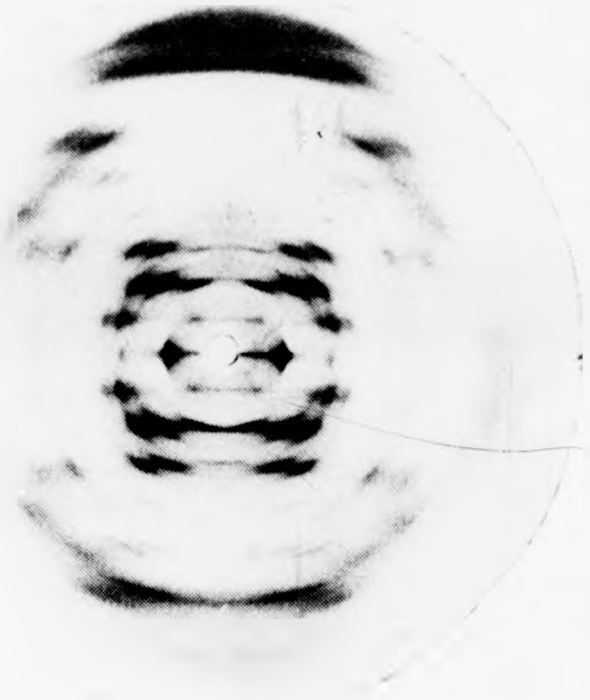


PLATE 4.5 The C' conformation from Li poly (dG-dG-dT).poly d(dA-dC-dC) as observed by Leslie et al. (1980). The lattice here is hexagonal such that a = b = 33.2 and c = 29.5Å.

4.2 is clearly distinct from the C" form described by Leslie et al., it does possess features which are characteristic of a 9_2 helix. There are weak meridional reflections half way between ℓ = 4 and 5 and also between ℓ = 5 and 6 which can be attributed to either small differences between successive residues along the double helix or to ions or water ordered to a periodicity of twice the translation per nucleotide. However, these additional reflections are weak and in the structural analysis described in this chapter, the pattern shown in Plate 4.2 was regarded as approximating to the C' type.

Leslie et al. did not report the determination of diffracted intensities for the C type patterns they observed. However, from a visual comparison there are distinct differences between the C' pattern from Li poly (dG-dG-dT).poly (dA-dC-dC) (see Plate 4.5) and that in Plate 4.2. For example, in addition to the extra meridional reflections (Plate 4.2) discussed above, the relative intensities of the (211), (301) and (221) reflections are quite different in the two patterns. Also the relative intensity of the 5th and 6th layer lines is reversed in the two patterns.

4.5 THREE MOLECULES IN THE UNIT CELL

The observation of sharp Bragg reflections on the higher layer lines in the diffraction pattern of Plates 4.2, 4.5 indicate that these fibres should be regarded as fully crystalline. The crystallinity is comparable to that in the C' diffraction pattern reported by Leslie et al. (1980) from Na poly (dG-dG-dT).poly (dA-dC-dC) although both patterns are significantly less crystalline than the best patterns reported for the crystalline B-form of naturally occuring DNAs (Langridge et al. 1960a). Although the C patterns of Li DNA analysed by Marvin et al. (1961) were all semi-crystalline, their approach to the characterisation of the molecular packing can be used as a basis for analysing the pattern in Plate 4.2.

From Cochran et al. (1952) the structure factor $F(h,k,\ell)$ of an array of helical molecules is given by equation 3.22, which is repeated here:

$$F(h,k,\ell) = \sum_p \sum_n \sum_j f_j J_n(2\pi r_j \xi) \exp i\{\psi + \frac{\pi}{2} - \phi_j) + 2\pi \ell Z_j/c\}$$

$$\exp i\{2\pi (hx_p + ky_p + \ell z_p) - n\phi_p\} \qquad 3.22$$

where r_j, ϕ_j, z_j are the coordinates of the j^{th} atom in the helical repeat, and x_p, y_p, z_p and ϕ_p are coordinates and orientation of the p^{th} molecule in the unit cell. $(R, \psi, \ell/c)$ are the cylindrical polar reciprocal space coordinates of the lattice point (h,k,ℓ). $J_n(2\pi r_j \xi)$ is an n^{th} order cylindrical Bessel function where for a helix with N residues per turn n takes values on the ℓ^{th} layer plane given by:-

$$n = \ell - mN \quad (m = 0, \pm 1, \pm 1, \pm 2 \text{ etc.}) \qquad 3.26$$

Following the analysis by Fuller et al. (1967) of an analogous structure of double-helical RNA, the intensity of the reflection (h,k,ℓ) can be expressed as:-

$$I(h,k,\ell) = Gn^2 \, P(h,k,\ell) \qquad 4.1$$

where

$$Gn = \sum_j f_j \, J_n \, (2\pi \xi r_j) \exp i \left[\frac{2\pi \ell z_j}{c} - n\phi_j\right] \qquad 4.2$$

and Gn^2 is the intensity scattered by a single molecule in the direction defined by (h,k,ℓ), and $P(h,k,\ell)$ is a packing factor whose magnitude depends on the arrangement of the molecules in the unit cell. For three molecules in the unit cell

$$P(h,k,\ell) = 3 + 2 \{\cos\left[2\pi \, (h(x_1-x_2) + k(y_1-y_2)) + \ell(\theta_1-\theta_2) - n(\phi_1-\phi_2)\right]$$

$$+ \cos\left[2\pi(h(x_2-x_3) + k(y_2-y_3)) + \ell(\theta_2-\theta_3) - n(\phi_2-\phi_3)\right]$$

$$+ \cos\left[2\pi(h(x_3-x_1) + k(y_3-y_1)) + \ell(\theta_3-\theta_1) - n(\phi_3-\phi_1)\right]\} \qquad 4.3$$

where $\theta_1 = 2\pi Z_1$ etc.

The pattern in Plate 4.2 is hexagonal and therefore in a fibre diagram there are 12 (6 if h = k or either h or k is zero) overlapping reflections. The packing factor is derived by dividing the observed intensity of a diffraction spot by Gn^2. $P(h,k,\ell)$ is therefore the sum of the packing factors for all the systematically overlapping reflections.

From Table 4.1 it can be seen that on ℓ=0 reflections are only observed for h-k = 3q where q is any integer including zero. Since for this region of the diffraction pattern m = 0 is the only significant component in the term Gn^2 (see Equation 4.2) then n = ℓ = 0 for these reflections. From Equation 4.3 it can be seen that the systematic absences on ℓ = 0 can be accounted for if the fractional unit cell coordinates for the three molecules are $(-\frac{1}{3}, -\frac{1}{3}, 0)$, $(0, \frac{1}{3}, Z_1)$, $(\frac{1}{3}, 0, Z_2)$ so that they are arranged as indicated in Figure 4.2. From Figure 4.2 it can be seen that each molecule has six nearest neighbours at a distance of $a/\sqrt{3}$. If the central molecule is at height 0 in the cell, three of these neighbours are at height Z_1 and three are at height Z_2. If we assume as in the models proposed for C-DNA by Marvin et al. (1961) and Arnott and Selsing (1975), that the strands of the DNA duplex are related by two-fold axes perpendicular to the helix axis through each nucleotide-pair then the interatomic contacts between the molecule at height 0 and those at height Z_1 will be identical to those between it and those at height Z_2 if $Z_1 = -Z_2$. Such maximisation of equivalence in intermolecular contacts is a common feature in crystalline fibres of nucleic acids (e.g. Arnott et al., 1967; Dover, 1977) and can be regarded as a reasonable assumption in the analysis of this structure. Such an assumption can, of course, be relaxed if it is found to be unable to account for the observed packing effects.

If reflections only occurred on ℓ=1 and 2 when (h-k) = 3q and on ℓ=3 when (h-k) = 3q, then by equation 4.3 packing factors which accounted for these systematic absences would be obtained if $Z_1 = +\frac{1}{3}$ and $Z_2 = -\frac{1}{3}$

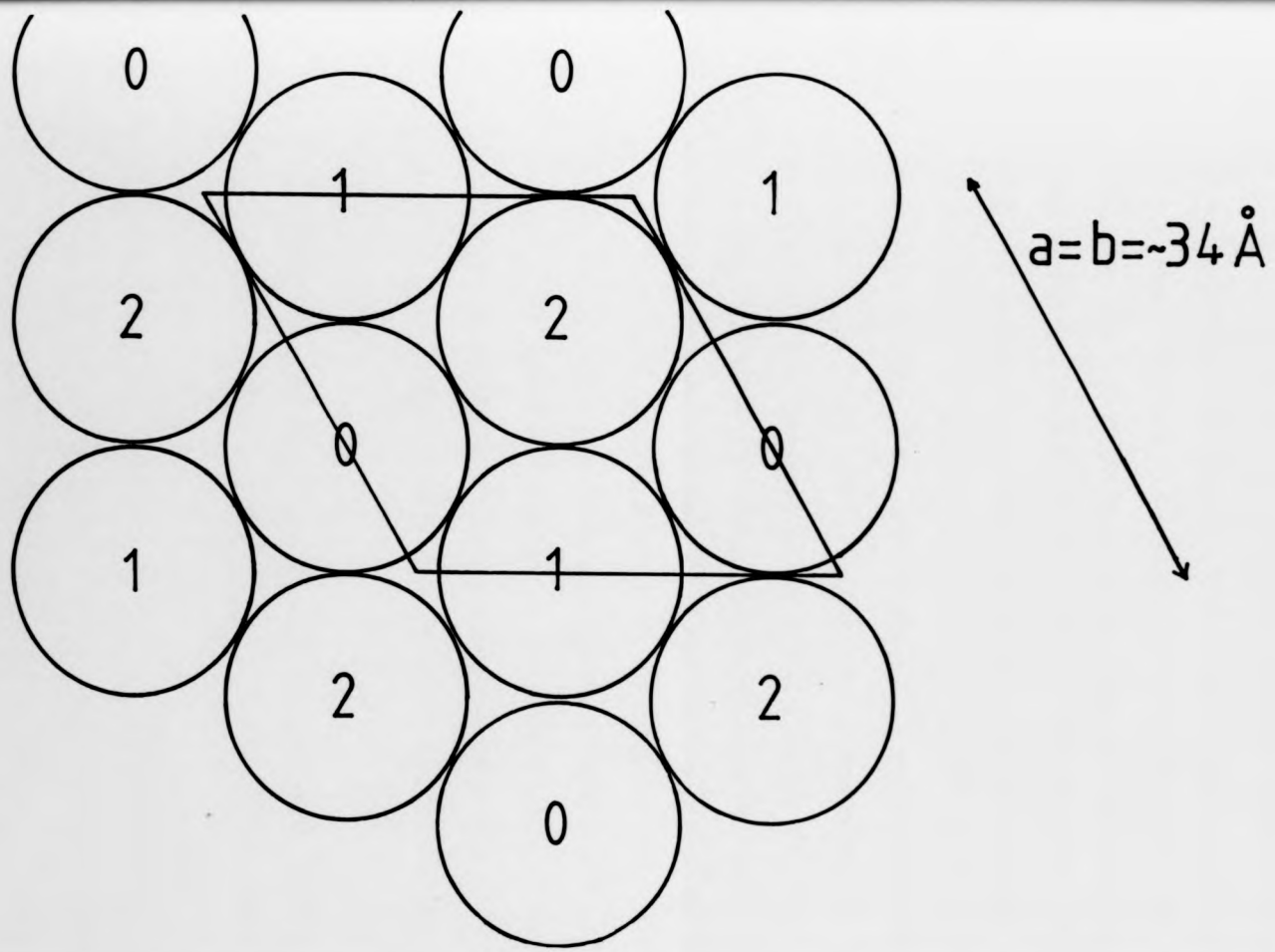


FIGURE 4.2 Arrangement of the molecules in the unit cell of the C' form of Li poly (dA-dC).poly (dG-dT). Molecules labelled 0 are all at the same height in the unit cell and similarly molecules labelled 1 and 2.

or if $Z_1 = -\frac{1}{3}$ and $Z_2 = \frac{1}{3}$. These two alternatives correspond to the three molecules passing through each cell being arranged according to a 3_1 axis or a 3_2 axis. In fact although some reflections on ℓ=1 and 2 for which (h-k) = 3q and on ℓ=3 for which (h-k) = 3q are weak, they are not all absent. These weak reflections occur in regions of the diffraction pattern where the molecular transform can be expected to be strong, e.g. the regions associated with the 301, 112, 203 and 213 reflections. Marvin et al. (1961) calculated the packing factors on ℓ=0 to 3 for various values of Z_1 and showed that their data was best accounted for if $Z_1 = 0.16 = -Z_2$. In Figure 4.3 the variation of the packing factor for the various classes of relfections is plotted as Z_1 varies from 0.00 to 0.50. From this variation it can be seen that the relative intensities of the spots of ℓ=1 to 3 are best accounted for if Z_1 is close to 0.25. The effect of varying Z_1 significantly from 0.25 can be investigated in terms of its effect on the relative intensity of neighbouring spots since the molecular transform can be assumed to a first approximation not to differ significantly at these two points. The following comparisons are of particular help in determining Z_1:

i) The (211) is clearly very strong whereas (301) is barely observed. As Z_1 is decreased below 0.25, P(301) increases rapidly and P(211) decreases rapidly.

ii) The (102), (202), (212) reflections are clearly very strong whereas the (112) and (302) are barely observed. In the region of $Z_1 = 0.25$ P(112) and P(302) are generally small and do not therefore place a strong constraint on Z_1 although they do require $0.12 < Z_1 < 0.38$. Similarly P(102), P(202) and P(212) are close to a maximum over a broad range about $Z_1 = 0.25$ and restrict Z_1 so that $0.12 < Z_1 < 0.38$.

iii) The (203) and (213) are strong, whilst the (113) and (303) are weak.

The packing factors are found to exist in several different families. The curves shown in figure 4.3 show P(hkℓ) for:

(a) the 101, 201, 211, 311, 321 reflections (green)

(b) the 111, 221, 301, 331 reflections (green)

(c) the 102, 202, 212, 312, 322 reflections (black)

(d) the 112, 222, 302, 332 reflections (red)

(e) the 103, 203, 213, 313, 323 reflections (blue)

(f) the 113, 223, 303, 333 reflections (red)

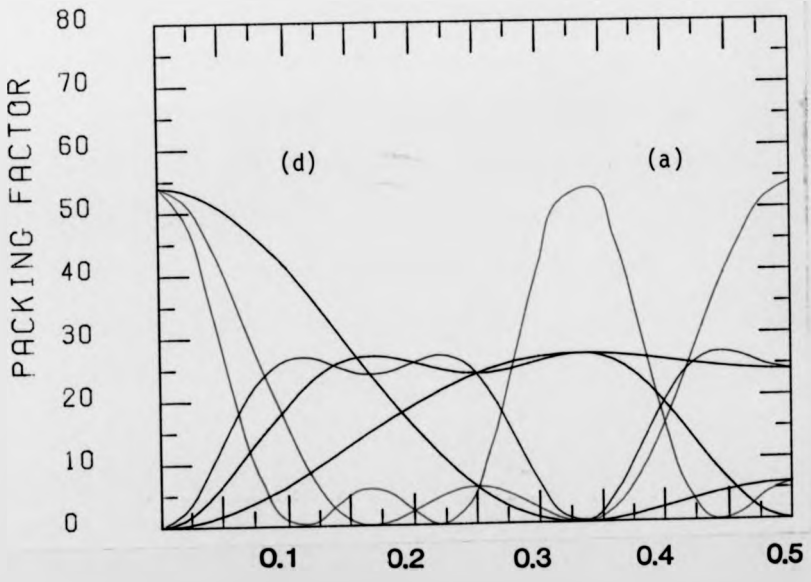


FIG 4.3 Illustration showing variation of different classes of packing factor with molecular displacement.

The packing factors are found to exist in several different families. The curves shown in figure 4.3 show P(hkℓ) for:

(a) the 101, 201, 211, 311, 321 reflections (green)

(b) the 111, 221, 301, 331 reflections (green)

(c) the 102, 202, 212, 312, 322 reflections (black)

(d) the 112, 222, 302, 332 reflections (red)

(e) the 103, 203, 213, 313, 323 reflections (blue)

(f) the 113, 223, 303, 333 reflections (red)

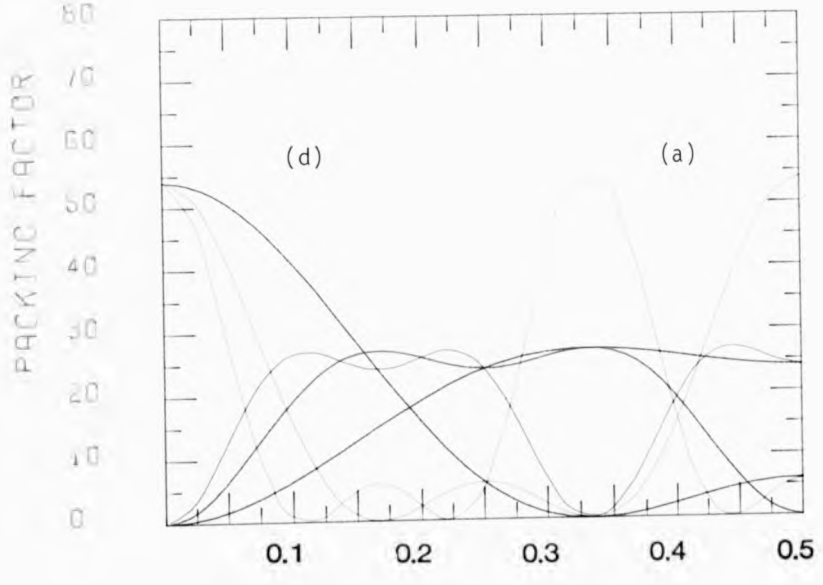


FIG 4.3 Illustration showing variation of different classes of packing factor with molecular displacement.

The packing factors at Z_1 = 0.25 reflect these relationships but if Z_1 increases above 0.25 P(113) and P(303) increase rapidly and P(203) and P(213) decrease rapidly.

The presence in the pattern shown in Plate 4.1(g) of streaks on ℓ=1 and 3 and Bragg reflections on ℓ=2 indicates that the molecules in the unit cell are randomly displaced with a screw disorder of C/2 and π from the relative positions determined above. This screw disorder was found to occur whenever a sample had undergone a sequence of transitions to the B form and back, and could only be eliminated by rewetting the fibre (see Section 4.3).

4.6 THE MOLECULAR CONFORMATION OF C'-DNA

Initially, the observed molecular transform was derived using equation 4.3 by dividing the observed spot intensities by the packing factors calculated assuming that the parameter Z_1 defining the relative height of adjacent molecules in the unit cell was 0.25 as determined in the previous section. This transform is compared in Figure 4.4 with that determined by Marvin et al. for a non integral C form. The data from this study and that of Marvin et al. are very similar on layer lines 0 to 4. However, there are differences on higher layer lines. As well as the additional meridional diffraction between ℓ=4 and 5 and between ℓ=5 and 6 the diffraction of ℓ=5 is marginally stronger than on ℓ=6 whereas in the patterns studied by Marvin et al. ℓ=6 is slightly higher. As was noted earlier, the relative intensity of layer lines 8 and 9 in the patterns described here vary with the relative humidity of the environment. The observed data described above is reasonably accurate in terms of a low resolution molecular model. However, for dealing with reflections on higher layer lines, the use of packing factors in which 'm' is assumed to be zero is not satisfactory. In the final stages of analysis the observed data was

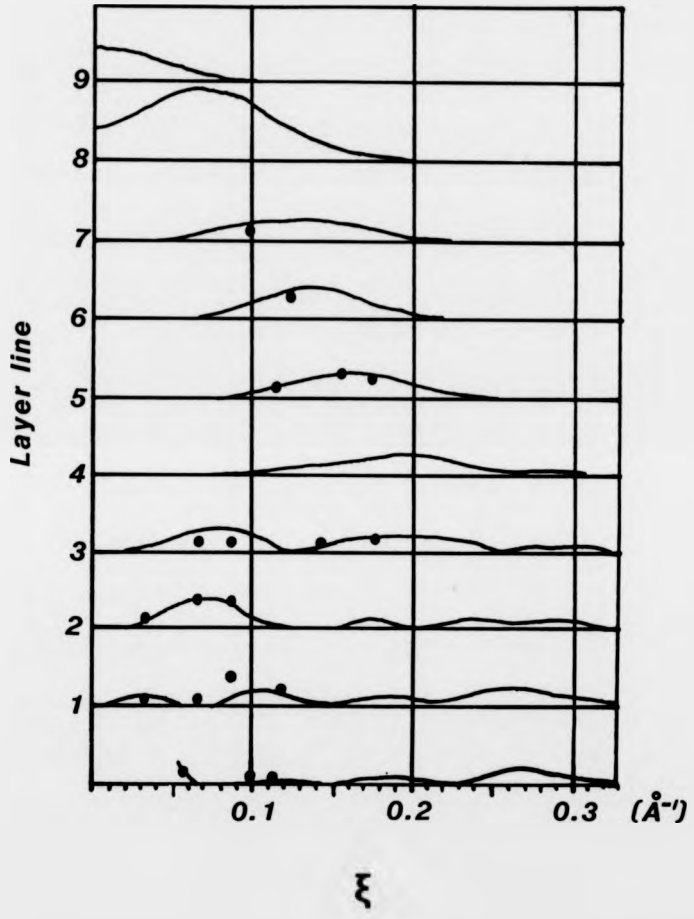


FIG 4.4 Comparison of the observed transform from Marvin et al.(1961) (continuous lines) with that obtained during this work. The dots shown superimposed on the continuous data indicate the spot intensities obtained from the pattern shown in plate 4.2 and corrected as per section 4.5 and fig 4.3.

compared with calculated structure factors and an agreement index was computed.

The agreement between the data determined here and that in Marvin et al. on the lower layer lines suggested that the Marvin model was a good starting point for the derivation of a satisfactory C' model. The Marvin model was thus distorted from a helix having 28_3 symmetry to one having 9_1 symmetry. Such a distortion is in fact quite small since it involves a change of rotation per residue from 38.6° to 40° and a change in translation per residue from 3.23Å to 3.32Å. Therefore, the original Marvin model was distorted, modifying the coordinates of all the atoms in the helical repeat so that r was unchanged but ϕ was multiplied by 40/38.6 and Z by 3.23/3.32.

The diffraction calculated from this model is shown in Figure 4.5 and Table 4.3. The agreement index of 41% is rather high but it was hoped that this would be reduced during computer refinement. Attempts were then made to adjust the stereochemistry of this model using the modelbuilding program described in section 2.5. These attempts failed to produce a model that satisfied the molecular constraints necessary for good agreement between observed and calculated diffraction. This was mainly a result of the stereochemical crudity of the wire models used by Marvin et al. during their analysis of the C structure. The intramolecular stereochemistry of the Marvin C model contains many bond lengths and angles which could not now be thought of as standard. Permissible variation in bond lengths and angles are in the region of ±0.03Å and ±0.3Å respectively. The author has made attempts to refine the Marvin model to incorporate standard stereochemistry but none of these converged to a satisfactory final model.

A number of 'B-type' models of the sort described by Arnott and Selsing (1975) were generated but none produced satisfactory agreement between observed and calculated diffraction. The cylindrically averaged

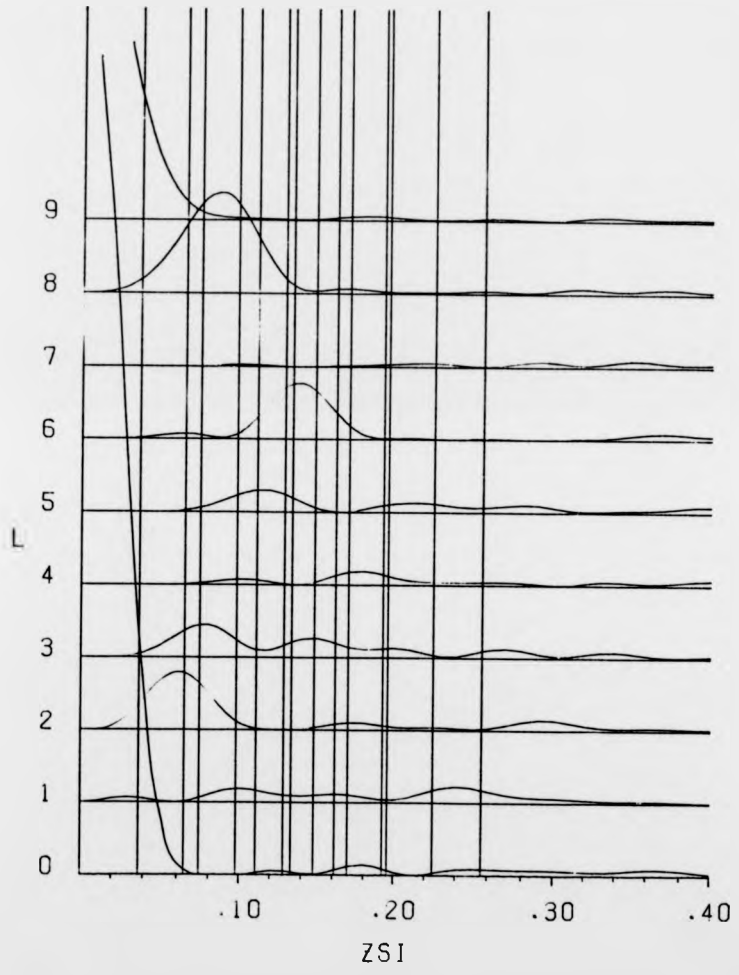


FIGURE 4.5 The cylindrically averaged squared transform of the 'distorted' Marvin structure (MARDIS). Points of sampling are indicated by the superimposed lattice, which is hexagonal and has a = b = 34A, c = 29.9A.

Reflection (HKL)	F_{OBS}	F_{CALC}
(110)	3250	3702
(300)	2667	698
(220)	1951	1879
(101)	1325	1255
(201)	1528	964
(211)	4250	3223
(301)	2644	1197
(311)	3163	2666
(102)	2400	3024
(202)	4004	4887
(212)	3620	3952
(203)	2458	3432
(213)	3421	4524
(323)	2728	4041
(423)	2649	2673
(315)	2353	4334
(415)	5839	594
(335)	3528	334
(316)	4178	6140
(217)	1929	710

Residual, R = 42%

TABLE 4.3 The observed structure factors along with those calculated from the 'distorted Marvin' model

TABLE 4.4 Coordinates relating to the 'distorted' Marvin model for C'-DNA. Here the original model of Marvin et al. (1961) has been distorted to meet the requirements of 9_1 symmetry. See section 4.6. The torsion angles pertaining to this model are given in the Appendix.

	ATOM	R(Å)	φ(°)	Z(Å)	X(Å)	Y(Å)
PHOSPHATE	P	8.98	111.4	2.87	-3.28	8.36
	O1	7.90	114.4	3.81	-3.26	7.19
	O2	9.12	117.7	1.79	-4.24	8.07
	O3	10.25	109.2	3.61	-3.37	9.68
	O4	8.52	102.0	2.22	-1.77	8.33
DEOXYRIBOSE	C1	5.48	91.2	0.86	-0.11	5.40
	C2	6.81	91.8	0.07	-0.21	6.81
	C3	7.77	85.1	0.89	0.66	7.74
	C4	7.29	86.3	2.30	0.47	7.27
	C5	8.19	94.2	3.05	-0.60	8.17
	O5	5.91	89.8	2.19	0.02	5.91
PURINE	N1	1.60	163.4	0.12	-1.53	0.46
	C2	1.12	107.0	0.27	-0.33	1.07
	N3	2.35	91.7	0.43	-0.07	2.35
	C4	3.34	110.0	0.44	-1.14	3.14
	C5	3.58	133.3	0.28	-2.46	2.61
	C6	2.90	155.6	0.12	-2.64	1.19
	N7	4.93	132.4	0.32	-3.32	3.64
	C8	5.38	118.8	0.50	-2.59	4.71
	N9	4.67	105.5	0.57	-1.25	4.50
(ADENINE)	N6	3.86	170.7	-0.03	-3.81	0.62
(GUANINE)	O6	3.86	170.7	-0.03	-3.81	0.62
	N2	0.83	178.2	0.27	-0.83	0.03
PYRIMIDINE	N1	2.98	131.5	0.28	-1.97	2.23
	C2	3.30	106.7	0.46	-0.95	3.16
	O2	2.79	84.6	0.51	0.26	2.78
	N3	4.67	105.5	0.58	-1.25	4.50
	C4	5.54	117.2	0.53	-2.53	4.93
	C5	5.34	132.2	0.35	-3.59	3.96
	C6	4.17	141.6	0.22	-3.27	2.59
(THYMINE)	Me	6.74	138.8	0.29	-5.07	4.44
	O4	4.51	158.5	0.06	-4.20	1.65
(CYTOSINE)	N4	4.51	158.5	0.06	-4.20	1.65

squared transform of the best such model is shown in Figure 4.6. A comparison of Figures 4.4 and 4.6 shows numerous flaws in the x-ray fit, most notably on layer lines 1, 3 and 5.

A new structure was then developed (initially using wire models) after a careful consideration of the way in which the different scattering components could be varied to give best fit firstly on the equator and later on higher layer lines. The various components contributing to the molecular transform of the 'distorted' Marvin model are shown for the inner region of the pattern in which an m=0 approximation is adequate (Figure 4.7). The most significant components of this transform are those due to the bases and to the phosphate. It was estimated that a shift in the radial position of the phosphorous atoms such that $r_{phos} \sim 8\text{Å}$ and a corresponding displacement of the bases further from the helix axis would maintain a satisfactory equatorial fit whilst taking some of the strain out of the sugar-phosphate backbone. Subsequent modification to the relative 'height' of the phosphorous atom in the chain was then applied to optimise the fit on layer lines one to three. In the final stages of analysis very small variations in the positioning of the different groups were made to vary the diffraction calculated for regions of the pattern where crystalline data was not available. In these instances agreement between observed and calculated data was largely qualitative in nature. The molecular transform calculated from the components of the final model are shown in Figure 4.8, and the cylindrically averaged squared transform (calculated for all components from m=-2 to m=3) in Figure 4.10. For the inner part of the pattern (up to ℓ=4), a comparison of the observed and calculated data yields a crystallographic residual of 17%. The coordinates of this model are given in Table 4.5 and Figure 4.9 provides a graphical comparison of the author's model and the distorted Marvin helix.

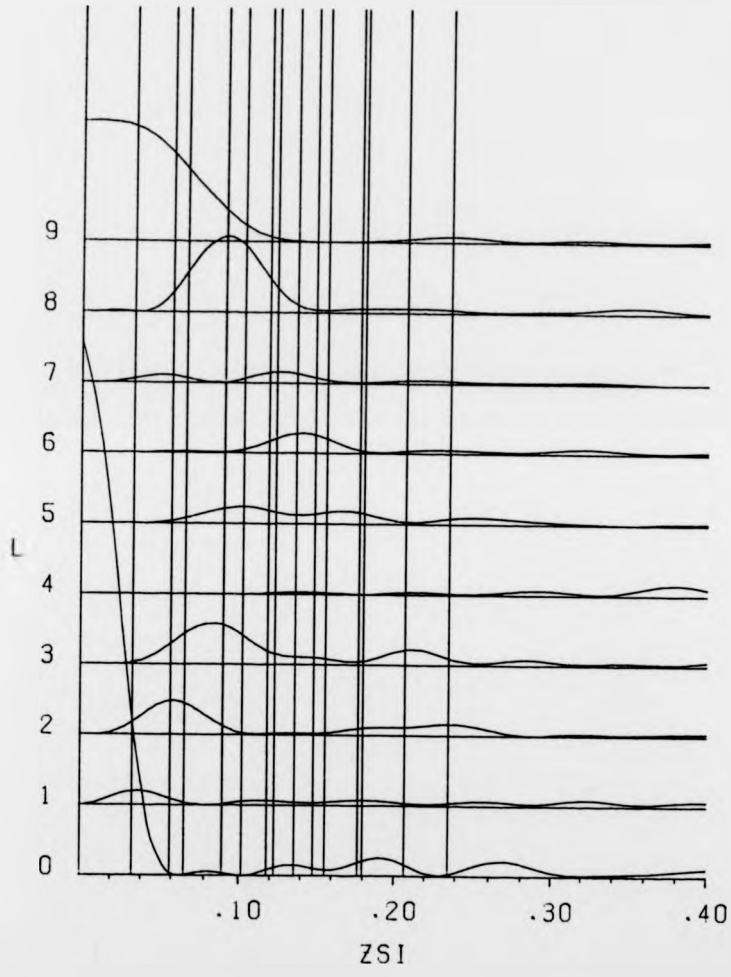


FIGURE 4.6 The cylindrically averaged squared transform of the 'Arnott type' model. The superimposed grid indicates the points at which the observed lattice (see Table 4.1) samples the transform.

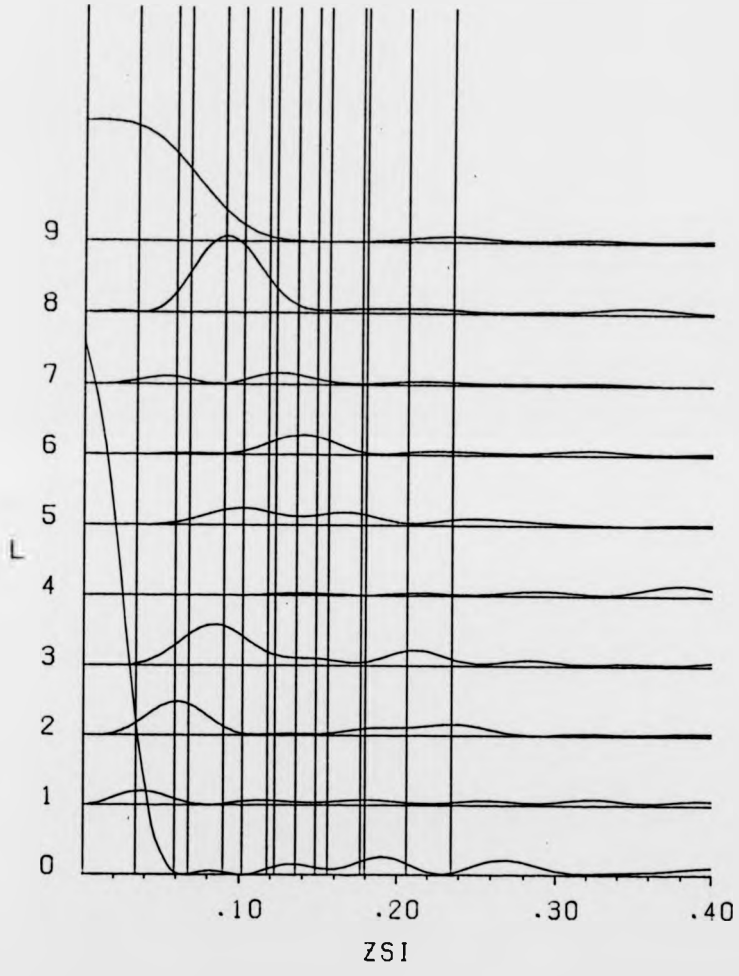
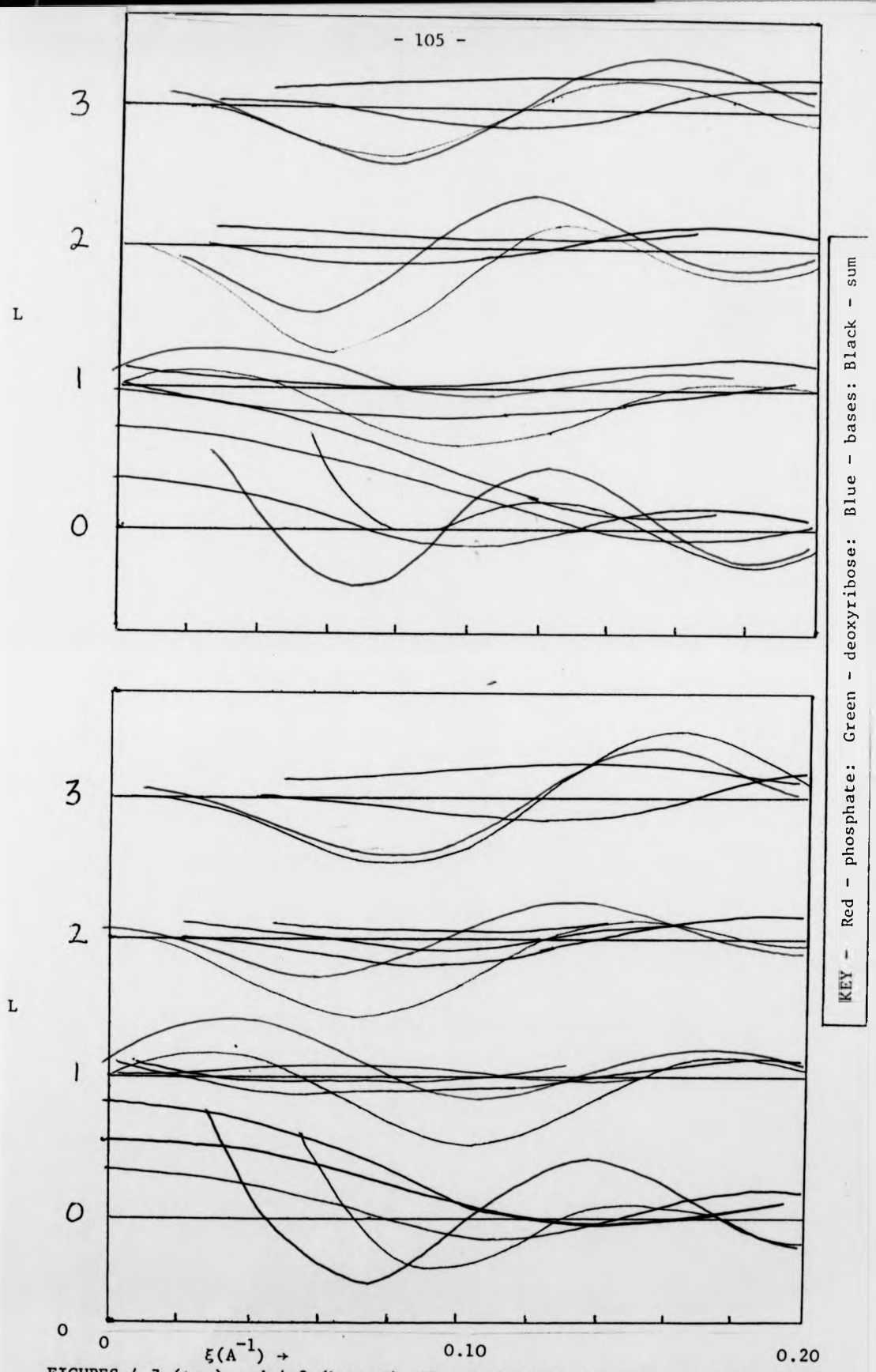
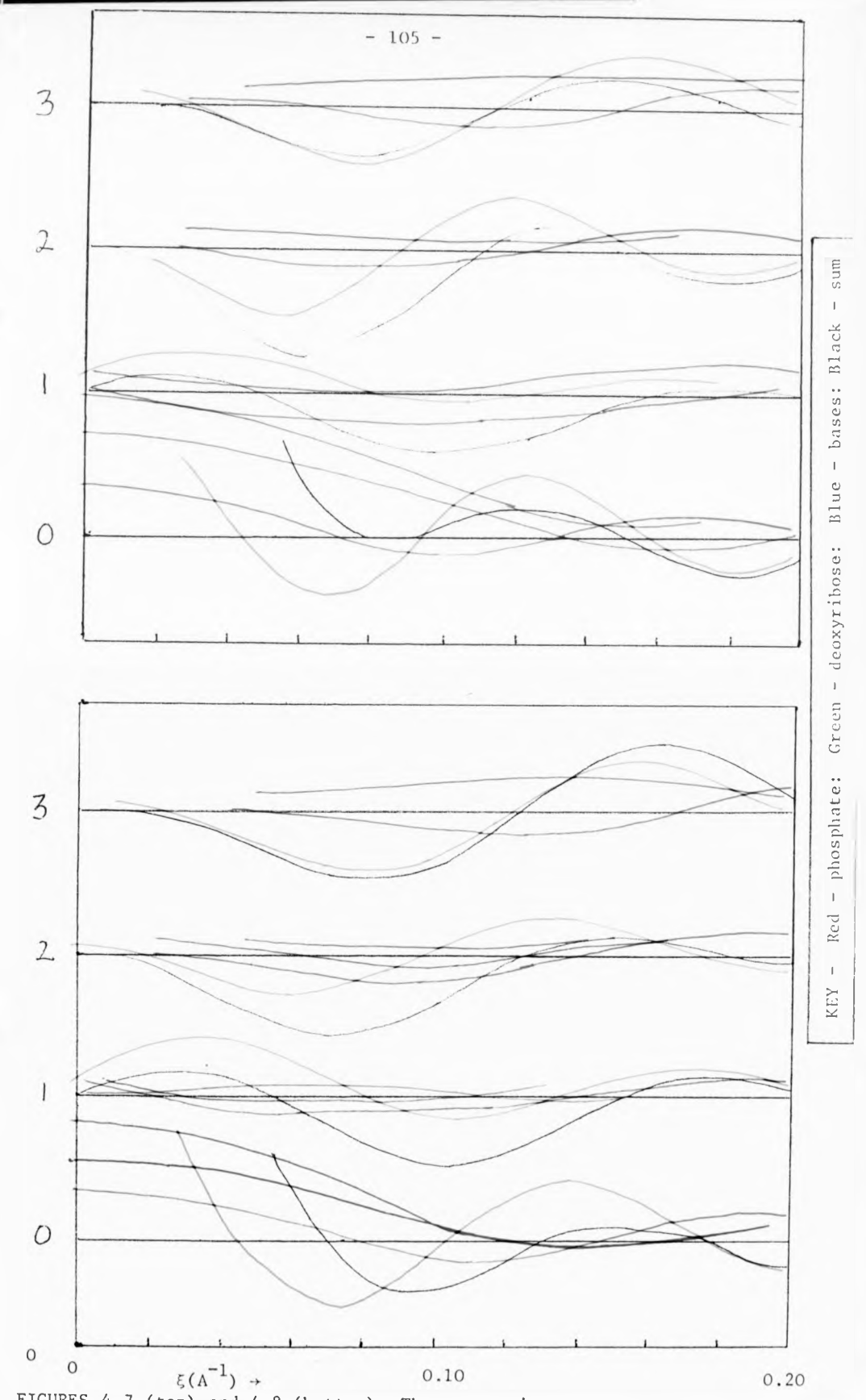


FIGURE 4.6 The cylindrically averaged squared transform of the 'Arnott type' model. The superimposed grid indicates the points at which the observed lattice (see Table 4.1) samples the transform.



FIGURES 4.7 (top) and 4.8 (bottom) The scattering components calculated from the m=0 molecular transform of the distorted Marvin model and the author's new model (respectively)



FIGURES 4.7 (top) and 4.8 (bottom) The scattering components calculated from the m=0 molecular transform of the distorted Marvin model and the author's new model (respectively)

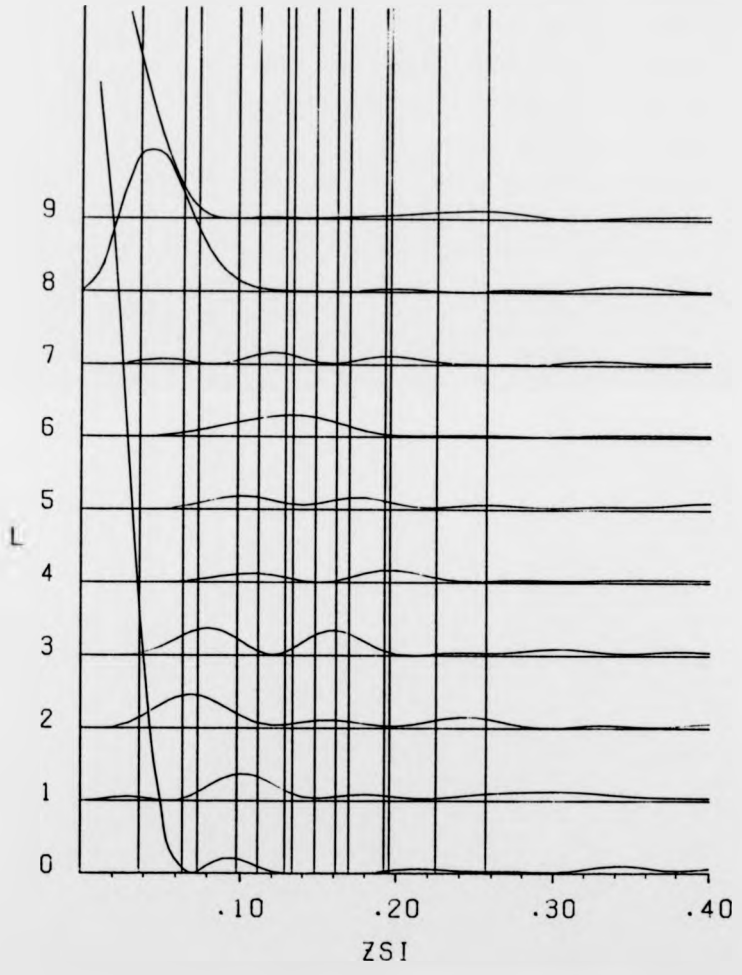


FIGURE 4.9 The cylindrically averaged squared transform calculated from the author's model for C'-DNA. The superimposed lattice is again that associated with the pattern shown in Plate 4.2.

	ATOM	R(Å)	φ(°)	Z(Å)	X(Å)	Y(Å)
PHOSPHATE	P	8.10	110.4	4.16	-2.82	7.59
	O1	7.74	120.6	3.47	-3.94	6.67
	O2	7.41	109.0	5.45	-2.41	7.00
	O3	9.57	110.2	4.21	-3.30	8.99
	O4	7.64	102.3	3.11	-1.62	7.46
DEOXYRIBOSE	C1	5.39	97.3	0.30	-0.69	5.35
	C2	6.92	98.4	0.23	-1.01	6.85
	C3	7.44	89.1	1.06	0.11	7.44
	C4	6.39	87.1	2.15	0.32	6.38
	C5	6.52	95.0	3.36	-0.57	6.50
	O5	5.11	89.5	1.48	0.05	5.11
PURINE	N1	2.48	170.0	-0.02	-2.42	0.52
	C2	1.60	138.4	-0.09	-1.20	1.06
	N3	2.45	109.8	-0.01	-0.83	2.30
	C4	3.64	121.9	0.19	-1.90	3.11
	C5	4.22	139.6	0.29	-3.21	2.73
	C6	3.72	158.8	0.18	-3.47	1.35
	N7	5.54	136.3	0.49	-4.01	3.83
	C8	5.80	123.5	0.51	-3.20	4.83
	N9	4.86	112.9	0.33	-1.89	4.47
ADENINE	N6	4.78	169.9	0.26	-4.71	0.84
GUANINE	O6	4.78	169.9	0.26	-4.71	0.84
	N2	0.30	132.0	0.31	-0.25	0.10
PYRIMIDINE	N1	4.88	112.8	0.34	-1.88	4.50
	C2	3.57	118.3	0.17	-1.69	3.14
	O2	2.73	102.4	0.01	-0.59	2.67
	N3	3.68	140.0	0.21	-2.82	2.37
	C4	4.98	145.6	0.40	-4.11	2.81
	C5	6.00	135.1	0.56	-4.25	4.24
	C6	5.93	121.9	0.52	-3.13	5.03
THYMINE	Me	7.37	140.0	0.76	-5.65	4.74
	O4	5.47	158.2	0.42	-5.08	2.03
CYTOSINE	N4	5.47	158.2	0.42	-5.08	2.03

TABLE 4.5 Coordinates of the author's model for C'-DNA. The helix has 9_1 symmetry, a translation of 3.22A per base pair and a rotation per residue of 40°. The torsion angles relating to this model are given in the Appendix.

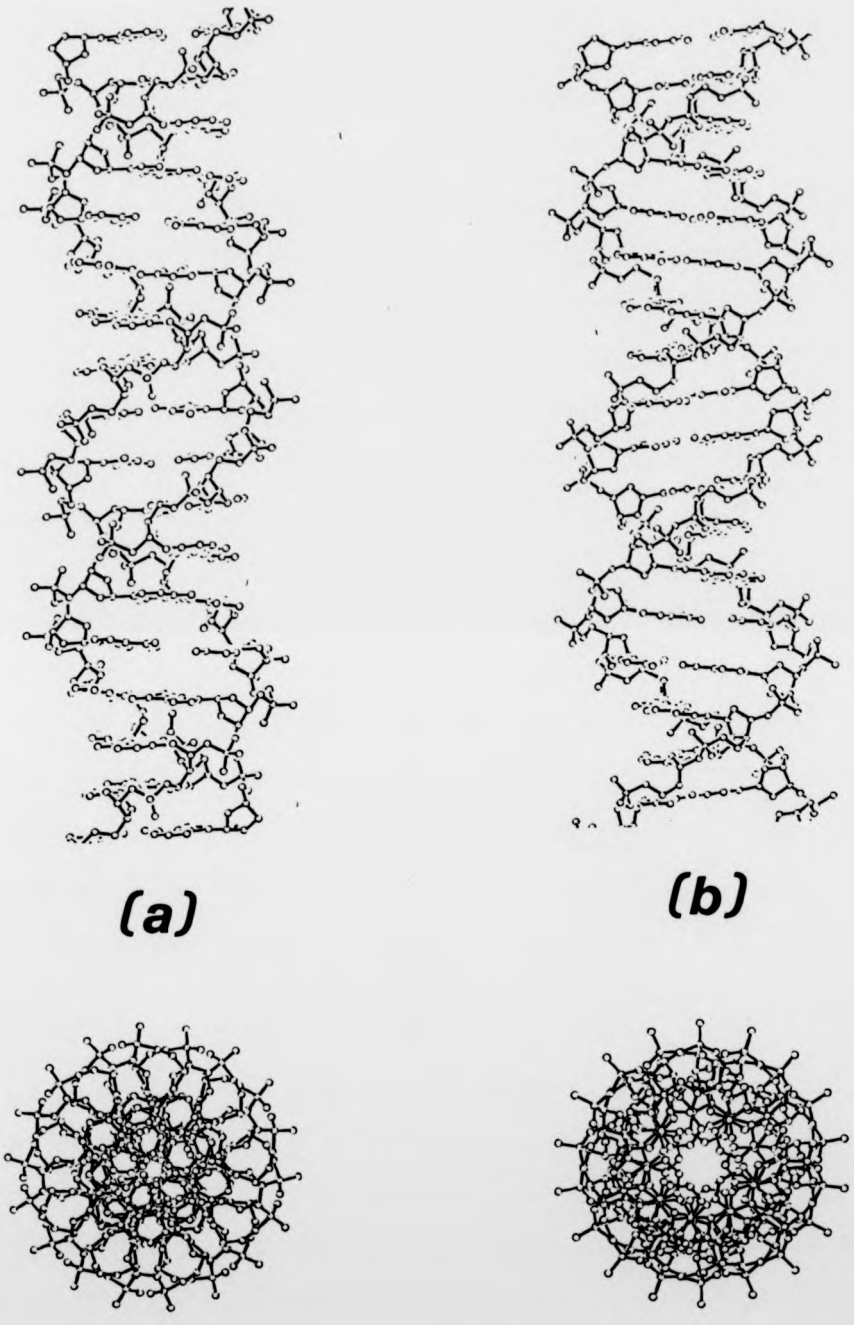


FIGURE 4.10 Two projections of (a) the 'distorted' Marvin model and (b) the modified model produced by the author during this study.

4.7 MOLECULAR ORIENTATION WITHIN THE UNIT CELL

A preliminary analysis of the packing effects manifest in Plate 4.2 has been described in Section 4.5, in which the C helices were regarded as being 'smooth' so that any molecular displacement along the Z axis was indistinguishable from an equivalent molecular rotation. Packing factors were calculated from equation 4.3 and applied to the intensity data relating to the central 'm=0' part of the pattern. This simplified 'low resolution' approach was used to determine the relative 'heights' of the three 'smooth' molecules in the unit cell, but for reasons just mentioned could not distinguish between molecular rotation and molecular displacement along the Z axis.

However, the few crystalline data that are available in regions for which the m=1 component of the pattern is significant can be used in conjunction with the packing factor analysis to obtain the most likely molecular orientation of the three helices within this unit cell. For a satisfactory m=0 fit to the pattern, it is required that $Z_1 = -Z_2 = 0.25$, or strictly speaking, that

$$(\theta_1-\theta_2) - (\phi_1-\phi_2) \approx 90^o \qquad 4.4$$

$$(\theta_3-\theta_1) - (\phi_3-\phi_1) \approx 90^o \qquad 4.5$$

As noted in Section 4.5, it is reasonable to presume that each of the three molecules in this unit cell will have equivalent intermolecular contacts. This can only occur if the molecules are packed as shown in Figure 4.12 so that each of the molecules have an equal angular displacement from the sides of an equilateral triangle formed by the group of three.

It is thus evident that

$$\begin{aligned} \phi_1 - \phi_2 &= -120^o \\ \phi_2 - \phi_3 &= -120^o \\ \phi_3 - \phi_1 &= 240^o \end{aligned} \qquad 4.6$$

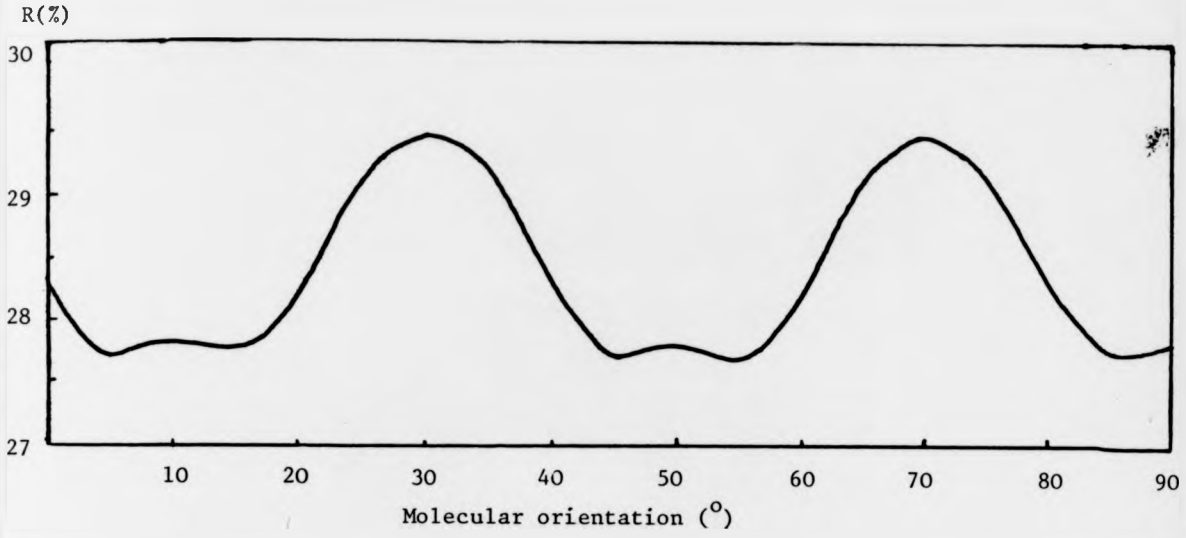


FIGURE 4.11 Graph showing variation of crystallographic residual as a function of molecular orientation within the hexagonal unit cell (see also Figure 4.12)

no matter how the helices are otherwise positioned. If θ_1 is assumed to be 0 then θ_2 is calculated to be 30^o and θ_3 to be 330^o and the molecules can be rotated freely subject to the constraints imposed by equations 4.6 without disrupting the m=0 x-ray fit.

The structure factors and crystallographic residual were thus calculated as a function of molecular orientation and this variation is depicted graphically in Figure 4.11.

The plot shown in Figure 4.11 is however ambiguous in that R has a minimum at 15^o and at 55^o. This situation arises because of the lattice symmetry. It is noteworthy firstly that R does not vary dramatically with orientation and secondly that the molecular rotations implied by this analysis do not pertain to the space group having the highest possible symmetry (i.e. P321 or $P3_221$ where the molecular diads would lie along the unit cell edges).

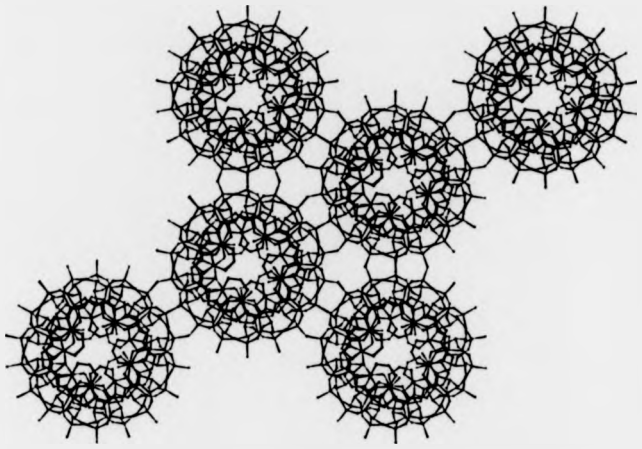


FIGURE 4.12 A view of the C' helices in the hexagonal cell (as seen down the c-axis)

The calculated structure factors associated with a molecular orientation angle of 15° are compared with the observed in Table 4.6. The agreement index R (see Equation 2.1) for this data is 28%.

4.8 CONCLUSION

Diffraction data obtained from the lithium salt of poly d(A-C). poly d(G-T) has been analysed in terms of a regular 9_1 duplex. The models of Arnott and Selsing (1975) do not give satisfactory agreement between observed and calculated data. A structure very similar to the original Marvin model but distorted to meet the requirements of 9_1 helical symmetry provides good overall fit to the observed data but has considerable stereochemical problems. The author's model was developed from the Marvin structure and refined to minimise an agreement index which relates observed and calculated structure amplitudes. This agreement index of 28% is good considering that the pattern of Plate 4.2 is not crystalline to the extent

Reflection (HKL)	F_{OBS}	F_{CALC}
(110)	3250	4221
(300)	2667	3305
(220)	1951	1557
(101)	1325	1140
(201)	1528	1117
(211)	4250	4336
(301)	2644	1624
(311)	3163	3198
(102)	2400	1852
(202)	4004	3757
(212)	3620	4372
(203)	2458	3002
(213)	3421	5151
(323)	2728	3852
(423)	2649	3861
(315)	2353	2464
(415)	5839	2079
(335)	3528	1741
(316)	4178	4357
(217)	1929	1092

Residual, R = 27.7%

TABLE 4.6 The observed structure factors along with those calculated from the new model. The torsion angles for this model are given in Appendix, The relative orientation of the three molecules in the unit cell is as described in Section 4.7.

of lithium B patterns (Langridge et al., 1960a, b). The model has C3 exo sugar puckering, bases which are displaced away from the helix axis by 1Å further than their position in the Marvin model, and a chain whose conformation angles (Appendix) are similar to those of the Marvin models. The dihedral angle 'α' (see Figure 1.7) has changed from $\sim -148^{o}$ (t) in the Marvin double helix to -40^{o} (g^{-}) in the author's model. This change was necessary to maintain the phosphorous atom in roughly the position required to satisfy observed lower layer line diffraction given a modified base position. The variation of intensity on ℓ = 8 and 9 as a function of humidity is seen as resulting from slight changes in tilt affecting the m = 1 region of the transform without greatly altering the predominantly m = 0 region of the pattern.

As mentioned in 4.4, the meridional reflections that can be observed between the 4^{th} and 5^{th} layer lines of Plates 4.7, 4.3 and 4.4 are thought to be due to either small differences between successive residues which hence strictly speaking constitutes a dinucleotide repeat, or to the presence of ions or water ordered to a periodicity of twice the translation per nucleotide. Similar observations have also been noted by Leslie et al. (1980) for the paracrystalline C" form of Na poly d(A-G). poly d(C-T) which these authors classify as a 9_2 molecule. It is interesting that evidence of such a periodicity is not found in the C' pattern of Li poly d(G-G-T).poly d(A-C-C) (Plate 4.5) presented by the same authors.

Two further variants of the preferred C model are currently being investigated:firstly, a model having 9_2 symmetry and a dinucleotide repeat; and secondly a 28_3 model is also being refined against the crystalline and continuous data obtained from native DNA's. The scaling of crystalline intensities with continuous transform has been described theoretically by Wilkins (unpublished).

The structure of C-DNA is also of interest because of the nature of its transition to the A and B forms. A preliminary investigation of these transitions is described in Chapter Eight and further experiments are scheduled at the SRS.

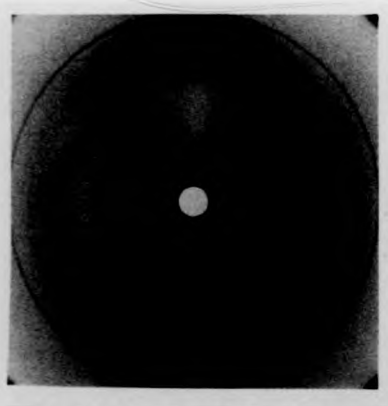
CHAPTER FIVE

THE STRUCTURE OF α-B'-POLY d(A-T).POLY d(A-T)

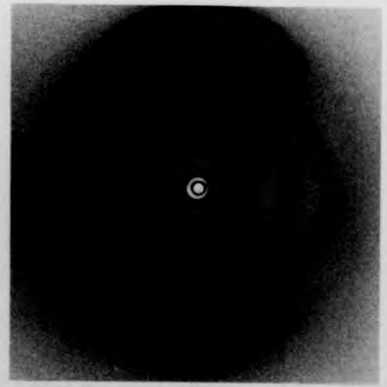
5.1 INTRODUCTION

The B conformation of DNA, as noted in Chapter 1, was the first DNA structure to be observed by means of X-ray diffraction (Astbury, 1947) and it was data recorded from this conformation that led Watson and Crick (1953) to the first satisfactory molecular model for double helical DNA. Evidence obtained since (Wilkins and Randall, 1953; Finch et al., 1981; Bentley, Finch and Lewitt-Bentley, 1981) has suggested that this conformation prevails widely in the nucleosome structure and in intact cells. Since the first observation of the crystalline and semi-crystalline B forms (see section 1.4.2), a great variety of synthetic polynucleotides have become available (Leslie et al., 1980) and experimental work has shown that a family of B structures exist. In addition to the classic crystalline and semi-crystalline B forms of native DNA, a number of other crystalline B patterns have been obtained from the lithium and sodium salts of poly d(A-T).poly d(A-T) and from the lithium salt of poly d(G-C). poly d(G-C) (Arnott et al., 1980; Mahendrasingam, 1984). These patterns are all slightly different and show that the B helix is capable of some variation.

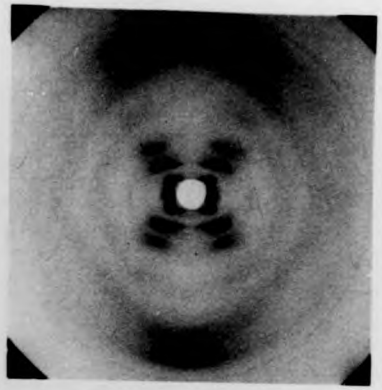
Native DNA's having high proportions of either (A-T) (as in Cℓ. perfringens DNA) or (G-C) (as in M.lysodeickticus) base pairs are known to exhibit distinctive structure features. Plate 5.1 shows diffraction patterns obtained from six different types of DNA. On the top row ((a) → (c)) a selection of native B forms are shown, whilst on the



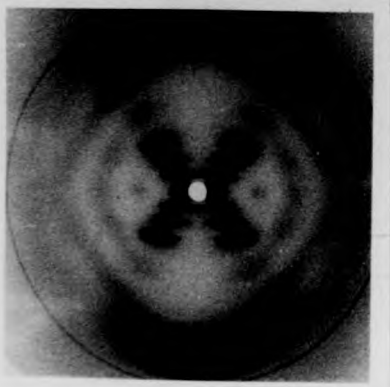
(a) Cl. perfringens B-DNA(A-T rich)



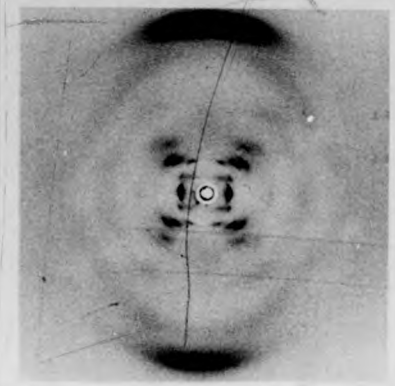
(b) Calf thymus B-DNA(approx. equal no. of A-T/G-C pairs)



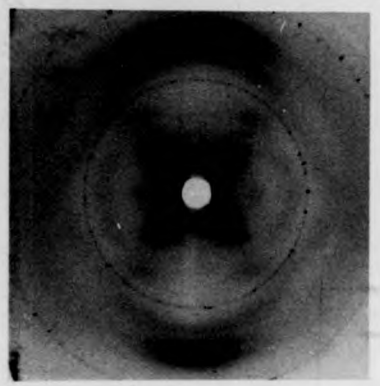
(c) M. lysodeikticus B-DNA(G-C rich)



(d) poly d(A-T).poly d(A-T) B-DNA

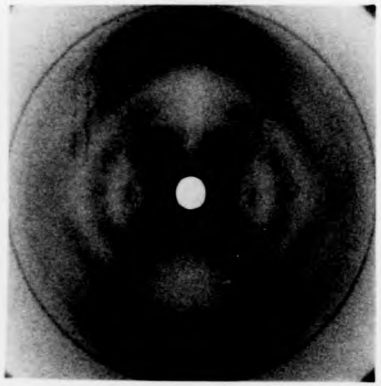


(e) poly(dA-dC).poly(dG-dT) B-DNA



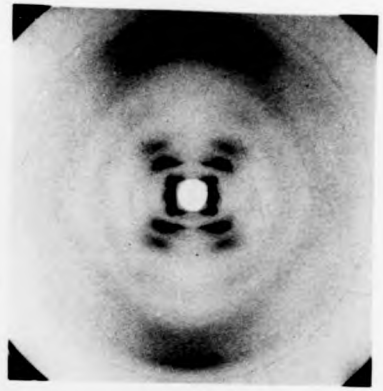
(f) poly d(G-C).poly d(G-C) B-DNA

PLATE 5.1 Patterns showing the variation of B-DNA diffraction with base content.

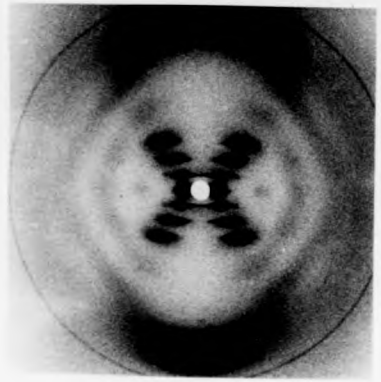


(a) Cl. perfringens B-DNA(A-T rich)

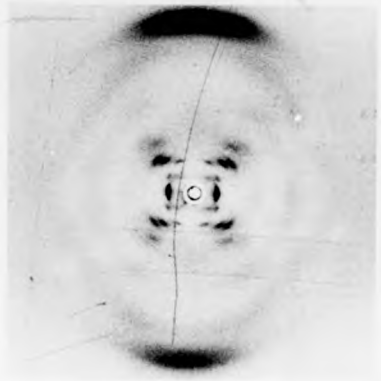
(b) Calf thymus B-DNA(approx. equal no. of A-T/G-C pairs)



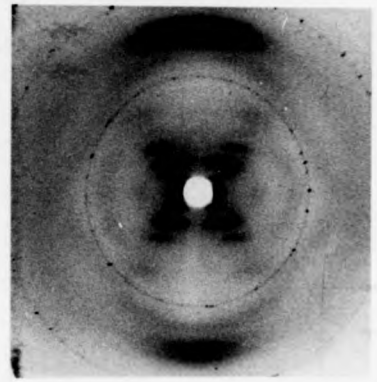
(c) M. lysodeikticus B-DNA(G-C rich)



(d) poly d(A-T).poly d(A-T) B-DNA



(e) poly(dA-dC).poly(dG-dT) B-DNA



(f) poly d(G-C).poly d(G-C) B-DNA

PLATE 5.1 Patterns showing the variation of B-DNA diffraction with base content.

bottom row ((d) → (f)) a corresponding selection of 'synthetic' B patterns are given . It can be noted from an inspection of Plate 5.1 that the B forms pertaining to samples having high A-T contents have strong diffraction on layer lines 1, 2 and 3. In contrast, those from 'high G-C' fibres have strong diffraction on ℓ=2 but weak on ℓ=1 and ℓ=3. B forms recorded from calf thymus DNA (which has a roughly equal proportion of A-T and G-C base pairs) show intermediate behaviour in this respect, as does B-DNA observed from poly d(A-C).poly d(G-T). It is thought likely that the above observations relate to a change in the value of $(\theta-\phi)$ from $\sim 70^{o}$ in the case of helices having high (A-T) content to something nearer $\sim 90^{o}$ in the case of samples having high G-C content. A number of different models are now being constructed in this laboratory to investigate such a variation.

A novel form of B-DNA has been obtained from fibres of high salt K poly d(G-C).poly d(G-C) and has been called B"-DNA (Mahendrasingam et al., 1983 - see Plate 1.5). This molecule is thought to have a dinucleotide repeat and 5_1 symmetry. A B form has been obtained from poly (dA-IU). poly (dA-IU) by Mahendrasingam (unpublished) and appears to have similar symmetry.

A further variant of the B form has been observed in fibres of Na poly dA.poly dT (Arnott et al., 1974) and is called B'-DNA. The B' form has been observed in two different lattices: the α (hexagonal) and the β (orthorhombic) forms. The α form can also be obtained from Li poly dA. poly dT and from the sodium salt of poly d(A-T).poly d(A-T) fibres (see Plate 5.2) (Rhodes and Mahendrasingam, unpublished). In their original study of the homopolymer, Arnott et al. found that at relative humidities of below 77% fibres gave a B' diffraction pattern that could be indexed on an orthorhombic grid such that a = 17.8Å, b = 20.0Å and c = 32.4Å. Above 85% relative humidity, these workers obtained a diffraction pattern which indexed on a hexagonal lattice having a = b = 22.8Å and

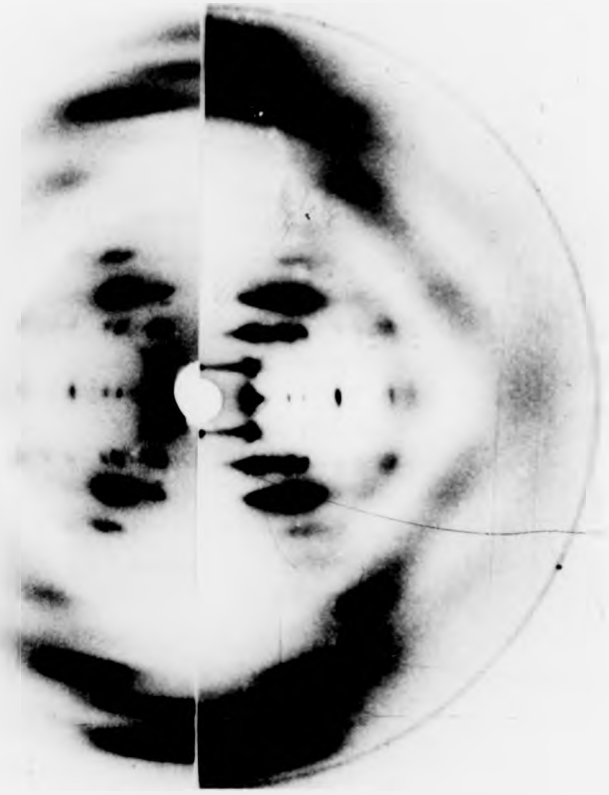


PLATE 5.2 'Split' photographs showing a comparison of the diffraction obtained from α-B'-poly dA.poly dT (left) and that from α-B'-poly d(A-T).poly d(A-T) (right) fibres

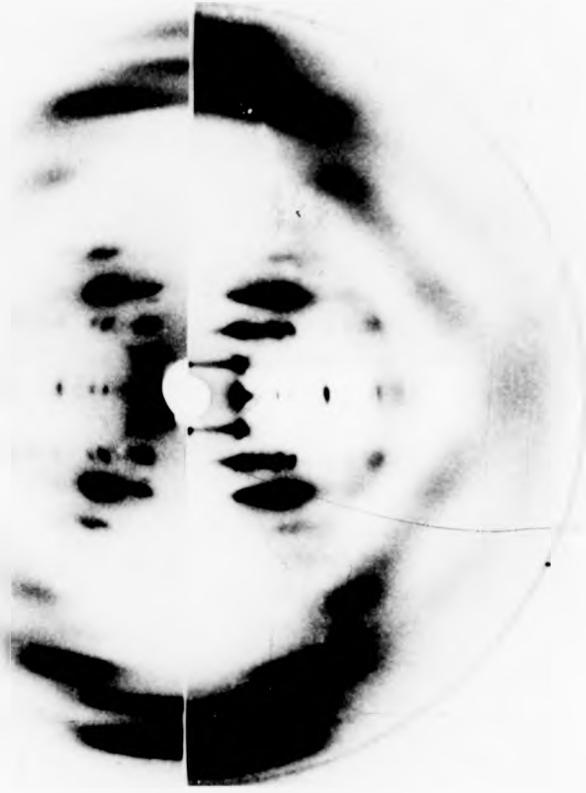


PLATE 5.2 'Split' photographs showing a comparison of the diffraction obtained from α-B'-poly dA.poly dT (left) and that from α-B'-poly d(A-T).poly d(A-T) (right) fibres

c = 32.9Å. Arnott et al. assumed (on the basis of similar pitch and overall intensity distribution) that the two patterns resulted from the same molecular structure packing into different lattices. A structure analysis by these authors of the α-B' pattern yielded a molecular model (similar to the classic B structure) whose calculated diffraction was in good agreement with that observed. However, in a recent publication, Arnott et al. (1983) retracted this model in favour of a unique structure which instead of having conventional diad symmetry had two morphologically distinct chains, one made up with 'A-type' steroechemistry and the other with 'B-type' stereochemistry. Their main reasons for favouring this model were:

(a) it gave acceptably good fit to the x-ray data

(b) the structure evolved spontaneously during the course of an 'unbiased' linked-atom least squares refinement procedure'

(c) it relieved some fairly chronic intermolecular contact problems when it came to packing into the orthorhombic unit cell of the β form.

Rhodes and Mahendrasingam (unpublished) have undertaken work relating to both these polymers and have found that whilst the lithium salts of both poly d(A-T).poly d(A-T) and poly dA.poly dT give the α-B' conformation, it does not appear to adopt the β-B' form. It is instructive here to summarise the appearance of the B' forms throughout the different polymer salts that have thus far been examined:

Salt	α-B'	β-B'
Sodium poly dA.poly dT	✓	✓
Lithium poly dA.poly dT	✓	X
Sodium poly d(A-T).poly d(A-T)	✓	X
Lithium poly d(A-T).poly d(A-T)	✓	X

TABLE 5.1 A Summary of the conformations observed in two salts of the polymers poly dA.poly dT and poly d(A-T). poly d(A-T) (✓ = observed; X = unobserved as yet)

There are two points of interest that emerge from an inspection of this table. Firstly, the β-B' conformation has not been observed in any salt of the alternating copolymer poly d(A-T).poly d(A-T). This observation is sharply contrasted by the fact that the α-B' form occurs in both copolymer and homopolymer, regardless of whether the counterion is lithium or sodium. As has been noted earlier, the α-B' diffraction patterns obtained from the two different polymers are virtually indistinguishable (see Plate 5.2), and it would seem reasonable to assume that the fibre structures are very similar as well. It is hard to envisage the alternating copolymer as a heteronomous molecule - the chemical repeat is such that one would expect a diad in between every second base pair (or some multiple thereof), and each chain should therefore be identical. If, on such a basis, the α-B' form of poly d(A-T).poly d(A-T) is not heteronomous, then the α-B' molecule in poly dA.poly dT cannot be either, since its fibre diagram is identical.

Assuming that α-B'-DNA is regular and not heteronomous, then where does this leave the concept of 'heteronomy' in B'-DNA? If, for a moment, we abandon Arnott's assumption that α and β are 'isomolecular', then (refer to Table 5.1) it still remains possible that the β form is heteronomous since it occurs only in the homopolymer (see Table 5.1). This idea is lent some support by the fact that the 1983 analysis of B'-DNA in terms of a heteronomous model was based on data observed from a crystalline β form, although the authors implied that the derived structure satisfied observed diffraction from both α and β.

The second point to note from Table 5.1 is that the lithium salt of the homopolymer does not appear to adopt the β form. This is interesting because to date no lithium salt of DNA has been found to adopt the A form and it is possible that the lithium ion directly suppresses this conformation. Whilst the A form itself is not specifically of interest here, the heteronomous

molecule proposed by Arnott et al. does contain 'A' stereochemistry in one chain and could possibly be expected to exhibit some of the features of A-DNA (such as being suppressed by the lithium ion).

The above observations tend to implicate the β-B' conformation as the heteronomous structure rather than both the α and β forms as presumed by Arnott and coworkers. The α-B' form occurs in situations which would not favour such a chain and, indeed, in which the β-D' conformation does not occur.

The fibre diagram from α-B'-poly d(A-T).poly d(A-T) is now analysed in the remainder of this chapter in an effort to quantitatively compare regular and heteronomous models for α-B'-DNA.

5.2 Experimental

Poly d(A-T).poly d(A-T) was supplied by Dr. J. Brahms and Mr. J. Vergne at the University of Paris. This material had been precipated from a 0.1 M sodium chloride solution, dialysed against a 10^{-4}M sodium chloride solution and then freeze dried. Although no flame emission spectroscopy (FES) studies were performed on this material, it has been estimated (Rhodes, 1983; Mahendrasingam, 1984) that such fibres of sodium poly d(A-T).poly d(A-T) contain ∿ 0.6-0.7 Na^+ and Cl^- ion pairs per DNA phosphate. Fibres were made by Professor W. Fuller and Dr. N.J. Rhodes using methods described in Chapter Two. X-ray diffraction photographs were recorded using a GX6 rotating anode generator and a Searle camera with Elliott toroidal optics. Pictures were also recorded at the Daresbury SRS. Diffraction pictures were measured up using a travelling microscope and intensities were calculated from traces produced by a Joyce-Loebl microdensitometer.

The overall distribution of intensity throughout the pattern shown in Plate 5.3 is very similar to that of the crystalline and semi-crystalline B forms as seen in Plate 1.3 and 1.4, and this suggested that

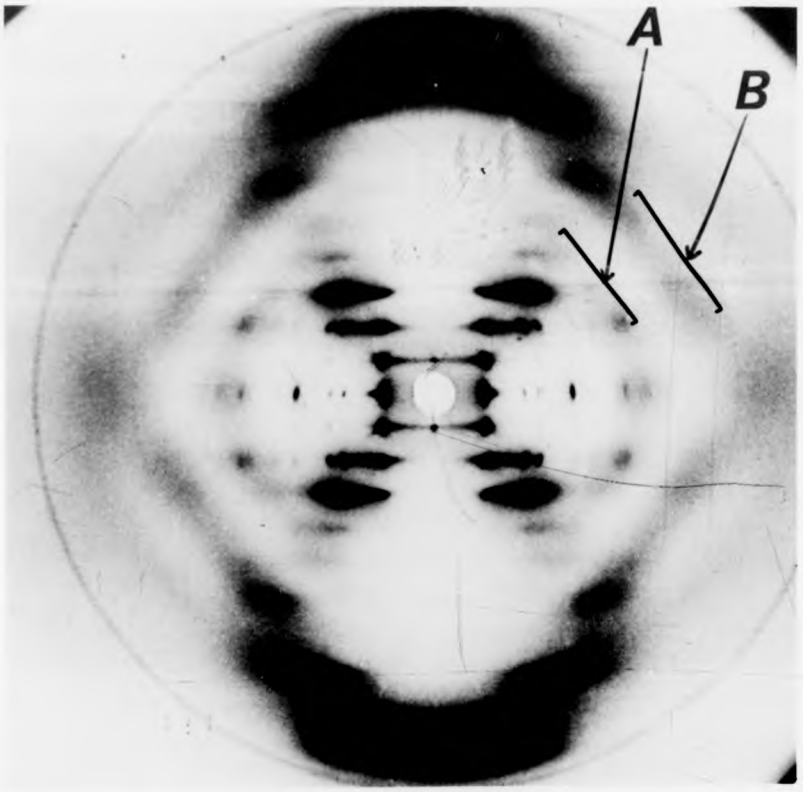


PLATE 5.3 The α-B' conformation of Na Poly d(A-T).poly d(A-T)

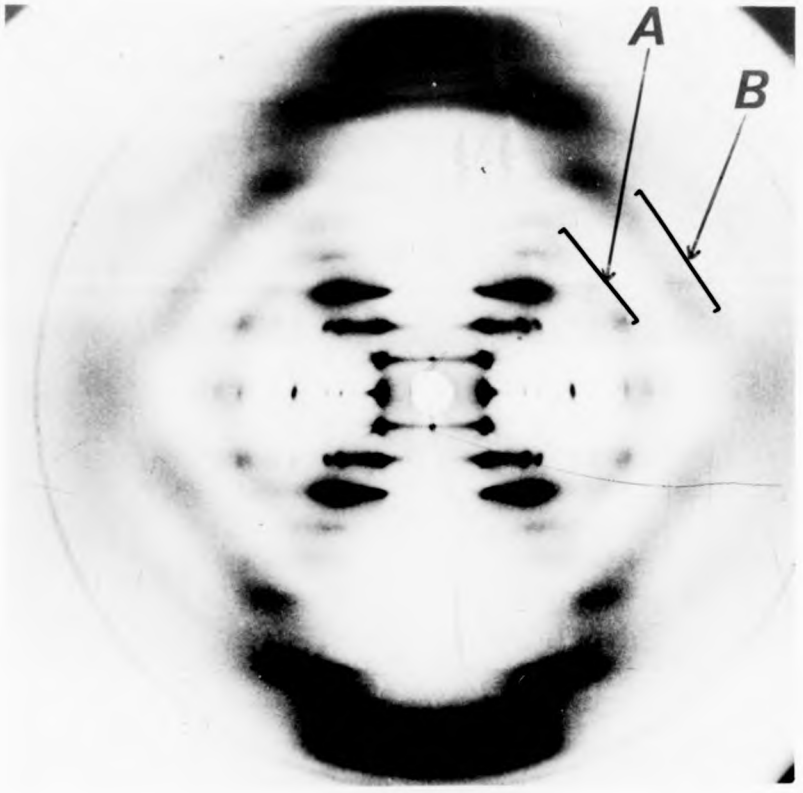


PLATE 5.3 The α-B' conformation of Na Poly d(A-T).poly d(A-T)

the B-DNA model of Arnott and Hukins (1973) might provide a reasonable starting model. There are, however, noticeable differences between the pattern of Plate 5.3 and diffraction from other B forms pertaining to A/T rich polymers. The α-B' pattern although still semi-crystalline is very much more crystalline than any of the classic semi-crystalline photographs seen in Plate 5.1. It does not, however, compare in crystallinity to the lithium pattern (Plate 1.4) analysed by Langridge et al. (1960). The B' pattern also differs to some extent in molecular transform. On the equator, the (110) and (200) reflections occur in a region where nothing is observed in other B-DNA patterns and where the transform calculated from the B-DNA model is weak. On the first layer line of this pattern there are also differences in the region for which $\xi \approx 0.10$ $\mathring{A}^{-1}$. Weak reflections occur here where the transform from the B model predicts them to be unobservable. On the second and third layer lines, the diffraction observed from the α-B' form compares well with other B forms and with the diffraction calculated from the B-DNA model. Layer lines four to ten are all more clearly resolved on the α-B' pattern than on semi-crystalline B patterns and it appears that the intensity distribution in this region of the pattern is very similar to that of other A/T rich B forms.

These preliminary observations were used during the development of various α-B' models. Such models were generated using the modelbuilding program as described in Chapter Two and adapted to provide satisfactory stereochemistry and best fit with the observed diffraction (see equations 2.1, 2.2).

5.3 Results and Discussion

The data used in this analysis was obtained from a conventional Cu K-α x-ray source, but diffraction patterns were also recorded at the Daresbury SRS.

The original α-B' pattern obtained from this sodium poly d(A-T). poly d(A-T) salt was indexed on a hexagonal grid with a = b = 24.4 Å and c = 32.8 Å. The observed and calculated ρ values are given in Table 5.2. The intensities calculated from radial densitometer traces across diffraction spots are also shown and compared with those obtained by Arnott et al. (1974) for Na poly dA.poly dT. This comparison shows two main regions where there are discrepancies between Arnott's observed data and this: firstly, on the second layer line the (102), (112) and (202) reflections of the poly d(A-T).poly d(A-T) pattern appear to be significantly stronger than their counterparts in poly dA.poly dT; secondly, on the fourth layer line diffraction appears to be much weaker in the copolymer than in the homopolymer. In other respects the intensity distribution is similar in both patterns.

A number of molecular models were generated; in the first instance these were adjusted to give best agreement with the measured data, and in the second to give qualitative agreement with non-crystalline data in other regions of the pattern. The best model had bases sitting right on the helix axis so that the distance 'D' (see Figure 1.10(c)) was 0.0 Å. The bases were twisted by 1° and tilted by -6°. The other torsion angles fit into the class ($tg^{-}g^{-}tg^{+}$) as is characterised by the first B' model of Arnott et al. (1974). The residual R (see equation 2.1) relating the observed data and that calculated from this model is 19.5%. The cylindrically averaged squared transform of the model is shown in Figure 5.1(a). The problems of the equatorial transform have been solved: the shift in base position relative to the other scattering groups has resulted in a swelling of the first subsidiary maximum ($\xi \approx 0.08$ Å^{-1}) and the final model can thus account for the two reflections observed in this region. (see observed and calculated structure factors in Table 5.3). Qualitatively, agreement between the observed and calculated transform

Reflection (HKL)	ρ_{OBS} (A^{-1})	ρ_{CALC} (A^{-1})	F_{OBS} (this work)	F_{OBS} (Arnott et al., 1974)
(100)	0.0470	0.0473	3660	3390
(110)	0.0823	0.0819	1610	1296
(200)	0.0950	0.0946	1428	1321
(210)	0.1254	0.1251	3925	2139
(300)	0.1416	0.1419	1926	
(220)	0.1418	0.1420	2233	1395
(310)	0.1710	0.1705	3291	2483
(400)	0.1886	0.1892	3279	
(101)	0.0550	0.0563	3933	3850
(111)	0.0880	0.0875	442	
(201)	0.1000	0.0994	1308	
(211)	0.1288	0.1289	860	
(102)	0.0755	0.0772	3913	2905
(112)	0.1022	0.1021	3710	2125
(202)	0.1130	0.1126	2962	1684
(212)	0.1390	0.1393	1267	1850
(302)	0.1542	0.1545	1424	2036
(222)	0.1748	0.1749	1635	1977
(312)	0.1808	0.1812	3180	2855
(103)	0.1031	0.1030	2752	3223
(113)	0.1232	0.1228	4717	5614
(203)	0.1326	0.1316	4013	5110
(114)	0.1480	0.1469	1291	3073
(204)	0.1550	0.1544	1839	3821

Hexagonal lattice : a = b = 24.4Å, c = 32.8Å

TABLE 5.2 The observed and calculated ρ values along with the observed structure factors and those of Arnott et al. (1974) for the hompolymer

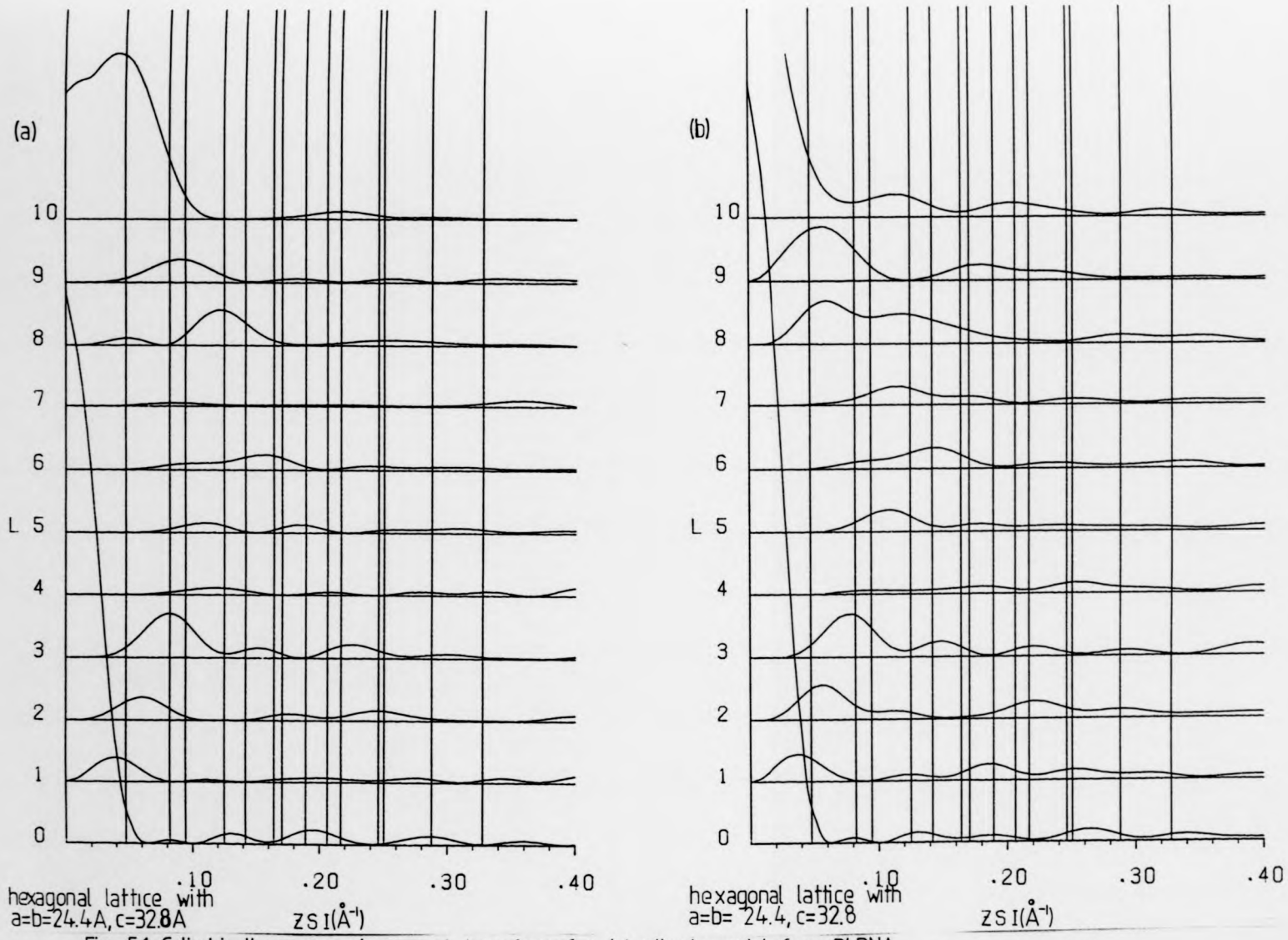


Fig 5.1: Cylindrically averaged squared transforms for (a) author's model for ∝-B'-DNA (see section 5.2)
(b) Arnott's 'heteronomous' model

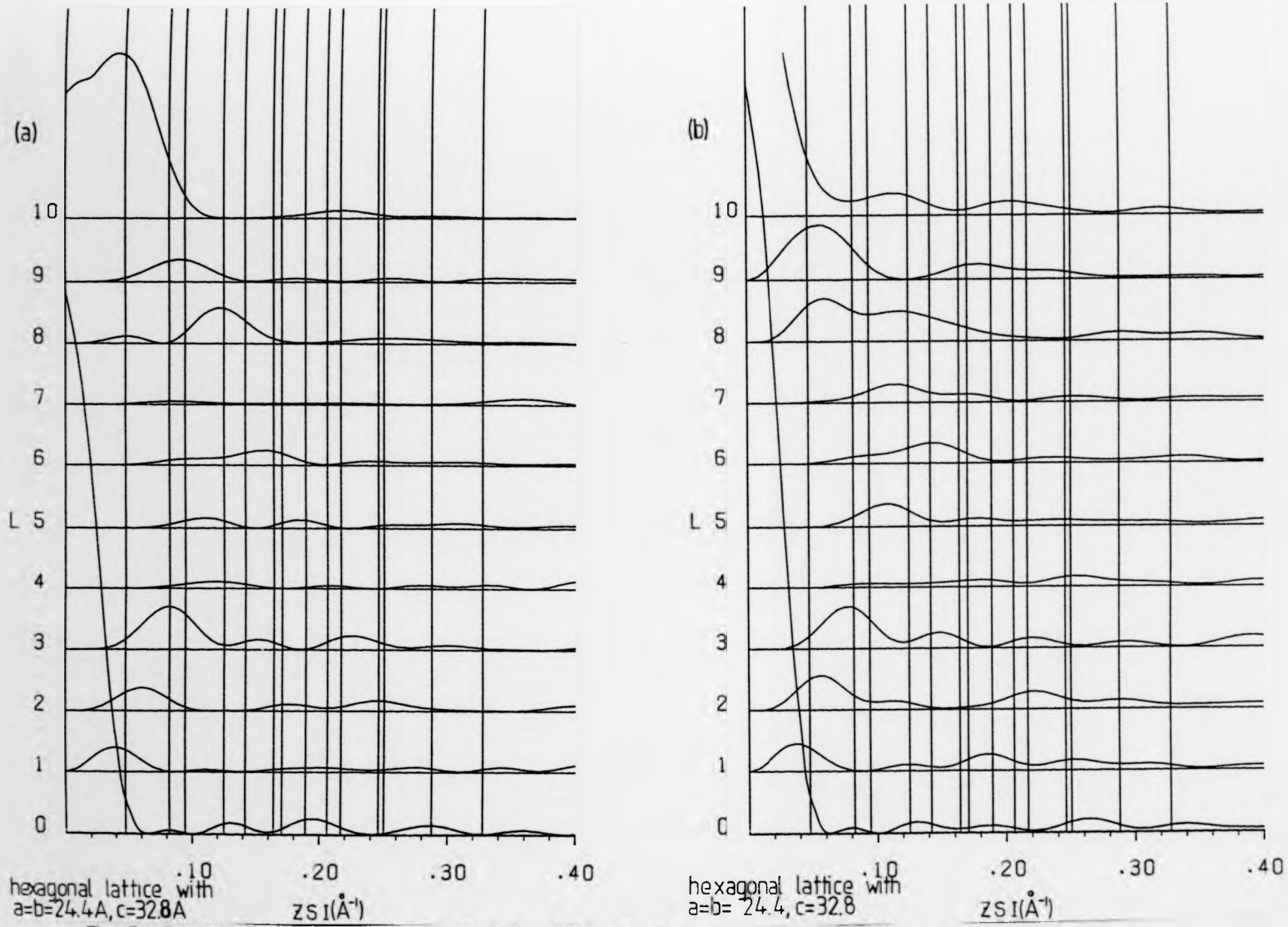


Fig 5.1: Cylindrically averaged squared transforms for (a) author's model for ∝-B'-DNA (see section 5.2)
(b) Arnott's 'heteronomous' model

REFLECTION	F_{OBS}	F_{CALC} (AUTHOR'S MODEL)	F_{CALC} (HETERONOMOUS MODEL)
(100)	3660	5087	5187
(110)	1610	1549	1803
(200)	1428	613	732
(210)	3925	3724	3460
(300)	1926	2273	2228
(220)	2233	1344	1460
(310)	3291	2744	2323
(400)	3279	3179	2045
(101)	3933	3888	3854
(111)	442	287	529
(201)	1308	805	621
(211)	860	1240	2881
(102)	3913	3577	4425
(112)	3710	2788	2975
(202)	2962	1466	2204
(212)	1267	661	2908
(302)	1424	700	1043
(222)	1635	1950	1047
(312)	3180	3055	1743
(103)	2752	2631	2784
(113)	4717	5453	5110
(203)	4013	4895	4259
(114)	1291	1351	1518
(204)	1839	1811	1604

Residual, R for Author's Model = 19.5%
Residual for Heteronomous Model = 27%

TABLE 5.3 Observed structure factors compared with those calculated from Author's model and Heteronomous model.

on higher layer lines is good.

The presentation of the data in Figure 5.1 and Table 5.3 affords a comparison between this 'regular' model for α-B'-DNA and the'heteronomous' one proposed by Arnott et al. (1983). Although the x-ray fit in the region for which crystalline data was obtained is good, there are other regions in which diffuse scatter in the observed data does not tally with the molecular transform calculated from the heternomous model. There are two ridges that are marked on Plate 5.3 as 'A' and 'B'. These ridges are well explained in the regular model but not in the heteronomous (see Figures 5.1(a) and (b)) Ridge 'A' is too weak on the equator and second layer lines and too strong on the first. Ridge B is too strong on the seventh layer line. Another problem with the heteronomous model is the intensity distribution on the fifth layer line, which is stronger than it should be, as is that from $\ell = 8$ in the region of $\xi \simeq 0.05$ $\mathring{A}^{-1}$.

The coordinates of the final model are given in Table 5.4. The torsion angles and associated parameters are given in Appendix. The α-B' molecule is very similar to the B' structure proposed by Arnott et al. (1974), differing only in base position and orientation. Whereas the Arnott model has tilt, twist and base displacement of -7.9°, -1.0° and -0.1 $\mathring{A}$, the authors model has -6.0°, 1.0° and 0.0 $\mathring{A}$.

The hexagonal cell side of $24.4\mathring{A}$ is large enough to accommodate molecular packing such that there are no unacceptable short contacts between adjacent helices.

This analysis thus tends to suggest that the α-B' conformation is a regular molecule very much like B-DNA structure. The torsion angles and other conformation parameters are all very similar to those of the model of Arnott and Hukins (1972b) although they differ substantially from those of later models (e.g. Arnott and Hukins, 1973; Arnott et al., 1980). There is not enough evidence to suggest a heteronomous model for α-B' DNA.

	ATOM	R(Å)	φ(°)	Z(Å)	X(Å)	Y(Å)
PHOSPHATE	P	8.74	91.5	2.28	-0.225	8.74
	O1	8.74	94.3	3.82	-0.65	9.71
	O2	9.93	96.5	2.01	0.611	9.92
	O3	8.71	99.5	1.44	-1.43	8.59
	O4	7.44	84.9	2.18	0.67	7.41
DEOXYRIBOSE	C1	5.84	64.2	0.60	2.55	5.25
	C2	7.07	70.3	0.01	2.38	6.66
	C3	8.16	66.6	0.95	3.23	7.49
	C4	7.44	66.2	2.30	3.01	6.80
	C5	7.47	76.1	3.08	1.79	7.25
	O5	6.09	62.1	1.97	2.85	5.38
THYMINE	N1	4.60	73.3	0.49	-1.52	4.35
	C2	3.40	63.8	0.35	-0.58	3.35
	O2	3.63	44.2	0.31	0.62	3.58
	N3	2.33	81.4	0.25	-1.07	2.07
	C4	2.94	108.6	0.28	-2.40	1.70
	O4	2.80	133.3	0.19	-2.75	0.52
	C5	4.35	104.0	0.43	-3.33	2.79
	Me	5.37	116.8	0.46	-4.78	2.46
	C6	4.98	89.1	0.52	-2.87	4.07
ADENINE	N1	0.86	33.3	0.07	0.72	0.47
	C2	2.19	26.7	0.14	1.96	0.98
	N3	3.23	43.2	0.28	2.35	2.21
	C4	3.31	67.1	0.35	1.29	3.05
	C5	2.71	90.8	0.29	-0.04	2.71

ATOM	R(Å)	φ(°)	Z(Å)	X(Å)	Y(Å)
C6	1.37	103.6	0.14	-0.32	1.33
N6	1.80	151.34	0.67	-1.58	0.86
N7	3.92	102.1	0.39	-0.82	3.83
C8	4.81	89.9	0.51	0.01	4.81
N9	4.60	73.3	0.49	1.32	4.41

TABLE 5.4 The coordinates of the final α-B' model.

The dyad related residue can be generated by negating Y and Z whilst coordinates of successive nucleotides may be obtained by applying a right handed screw operation involving a rotation of 36° about the Z axis and a translation of 3.28A along it. The torsion angles associated with this model are given in Appendix 1.

There is, however, every reason to believe that the heteronomous model holds good as an explanation for the β-B' diffraction data from Na poly dA.poly dT, and as such the model is of considerable value in assessing the nature of sequence dependent local variation. The structure of β-B'-DNA (be it heteronomous or not) could also be of significance in endeavours to establish the way in which the lithium ion exerts its effect in DNA.

It would be interesting to determine whether the possibility of a heteronomous duplex exists in the homopolymer poly dG.poly dC. Extensive studies on this polymer have so far not been published although one report by Arnott and Selsing (1974b) suggests that no novel conformations are obtained from the sodium salt of this polynucleotide.

FIGURE 5.1 (a) The α-B' form of DNA as seen in sections looking perpendicular and parallel to the helix axis.

(b) The heteronomous structure proposed by Arnott et al. (1983)

CHAPTER SIX

THE STRUCTURE AND SIGNIFICANCE OF β-D-POLY d(A-T).POLY d(A-T)

6.1 INTRODUCTION

In this study, the evidence for the existence of left-handed helices in polynucleotide duplexes is reviewed, and a comparison of the activity of poly d(A-T).poly d(A-T) and poly d(G-C).poly d(G-C) has prompted a reassessment of the current assumption that D-DNA is a right-handed double helix. An analysis of the x-ray diffraction obtained from the β-D form of poly d(A-T).poly d(A-T) has been undertaken in an effort to determine whether a right or a left-handed model is to be preferred.

6.1.1 Studies relating to left-handed helices

Within the last few years the field of nucleic acid research has been dramatically influenced by the realisation that particular sequences of DNA may adopt a left-handed helix sense. Studies on the short oligomers d(CpGpCpGpCpG) and d(CpGpCpG) have revealed the very unique ZI, ZII and Z' structures (Wang et al., 1979; Wang et al., 1980; Crawford, et al., 1980; Drew et al., 1980), and a corresponding examination of the synthetic polynucleotide poly d(G-C).poly d(G-C) by Arnott et al. (1980) has provided sufficient evidence to throw into contention the whole question of DNA handedness.

Further interest has been evoked by circular dichroism (CD) and ^{31}P-NMR studies on synthetic oligomers of (dC-dG) inserted within DNA restriction fragments (Klysik et al., 1981). These have illustrated not only the presence of left-handed fragments of (dC-dG) and the ability of

these fragments to change helix sense (as a function of salt concentration), but also that the nozel Z helix can exist in very close proximity to supposedly 'B-like' helical segments. These authors note a 'junction' between Z and B helices is possible, and such a region appears to be approximately 11 base pairs in extent. They also assert that the presence of the (dC-dG) segment in the recombinant plasmid is sufficient to have a profound influence on the supercoiling properties of the DNA. The possible biological significance of all this is manifold and of great intrigue - for example, a segment of left-handed DNA in viral or cellular DNA which is principally composed of right handed conformations could provide a difference in continuity which was sufficiently dramatic as to be an easily recognisable locus for binding of specific proteins or lingands (Sage and Leng, 1980).

The B form has for some time been acclaimed as a conformation which prevails extensively throughout chromatin and the intact cell (see sections 1.4.2 and 5.1). However, work undertaken by Milman et al. (1967), O'Brien and MacEwan (1970), Chandrasekaran et al. (1980) and Arnott et al., (1980) has implicated the A conformation (see section 1.4.1) as the one favoured by DNA/RNA hybrids and as such that the A form may be involved in the process of transcription. Z-DNA has also been found to exist in the cell. Z-DNA, in contrast to B-DNA, is strongly immunogenic and antibodies have been produced which react against it (Lafer et al., 1981; Malfoy et al., 1981; Pohl et al., 1982; Moller et al., 1982): this fact has enabled immunohistochemical studies to be undertaken. The presence of Z-DNA has been established in two dipteran species (Nordheim et al., 1982; Lemeunier et al., 1982), in some plant nuclei, and more recently in a variety of different types of rat tissue (Morgenegg et al., 1983).

The evidence thus far obtained now firmly establishes the

significance of the Z conformation in vivo. The left-handed form is easily induced in the solid and liquid states of poly d(G-C).poly d(G-C) (Pohl and Jovin, 1972; Pohl et al., 1972; Pohl et al., 1973; Pohl, 1976; Arnott et al., 1980; Zimmerman, 1982; Neidle, 1983; Mahendrasingam et al., 1983) but it appears more difficult to induce the Z conformation in poly d(A-C).poly d(G-T) and to date it has not been observed at all in poly d(A-T).poly d(A-T).

The transition from what is assumed to be the 'B' form to the Z form can occur as a result of changing solvent composition, or by modification with certain large substituents or by the torsional strain associated with supercoiling (Pohl et al., 1973; Zimmerman, 1982; Neidle, 1983). The implication of supercoiling in these studies is significant and is of special concern when it is realised that large stretches of poly d(A-C).poly d(G-T) and of poly d(G-C).poly d(G-C) have frequently been observed as periodic elements in the composition of mammalian DNA. It is conceivable that these regions play a part in the packaging of DNA into its final nucleoprotein structure as seen in the intact cell. Another possible role of Z-DNA in the nucleus is that relating to specific Z-DNA binding proteins. Morgenegg et al., (1983) claims to have detected such proteins in the nuclei or rat cells and bovine brain cells. Nordheim et al. (1982) has observed Z-DNA binding proteins in Drosphila nuclei.

6.1.2 Can S-DNA and D-DNA be regarded as analogues?

At about the time that this work started, it became apparent (from studies undertaken by Dr. A. Mahendrasingam in this laboratory) that there were certain features relating to the scheme of transitions as seen in two synthetic alternating purine/pyrimidine polynucleotides (namely poly d(A-T).poly d(A-T) and poly d(G-C).poly d(G-C)) that merited some attention.

The sequence of transitions observed in fibres of KF poly d(G-C).poly d(G-C) is summarised in Figure 6.1 and Plate 6.1 shows the diffraction photographs corresponding to the various conformations. At relative humidities of up to 75%, a B" diffraction pattern is observed. This then changes to an A pattern which prevails to 92% RH. The fibre then undergoes a transition to the S_{II} form. A semi-crystalline B form is obtained after prolonged exposure of the fibre to 98% RH. The

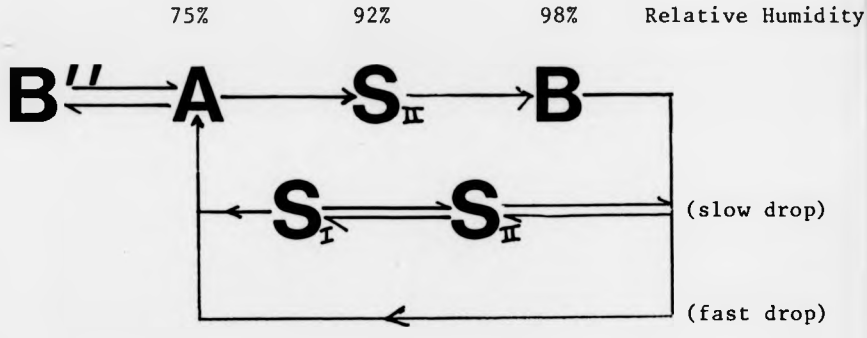
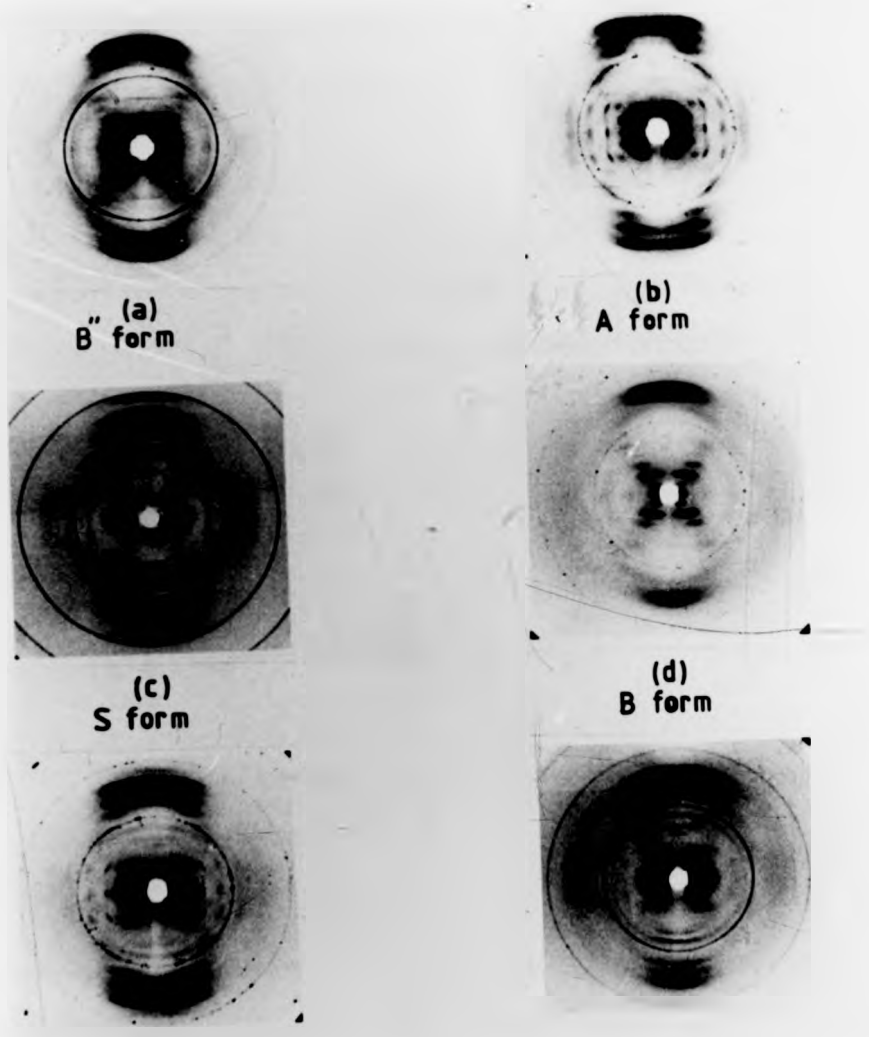


FIGURE 6.1 The sequence of transitions observed in fibres of KF poly d(G-C).poly d(G-C).

reverse sequence, as can be seen from Figure 6.1, is somewhat more complicated; a gradual reduction of the RH from 98% results firstly in the reacquisition of the S_{II} form, and then in the appearance of what is clearly an S form but which differs from S_{II}-DNA in lattice parameters and in molecular conformation. The structure is called S_I-DNA (see Figure 6.1). The S-conformations obtained during this procedure were found to be very stable, and the lower humidity A and B" conformations could not be recovered by subsequent reduction of the RH. If, however, in the initial stages of reversing the sequence, the humidity was suddenly changed from



(a)
B" form

(b)
A form

(c)
S form

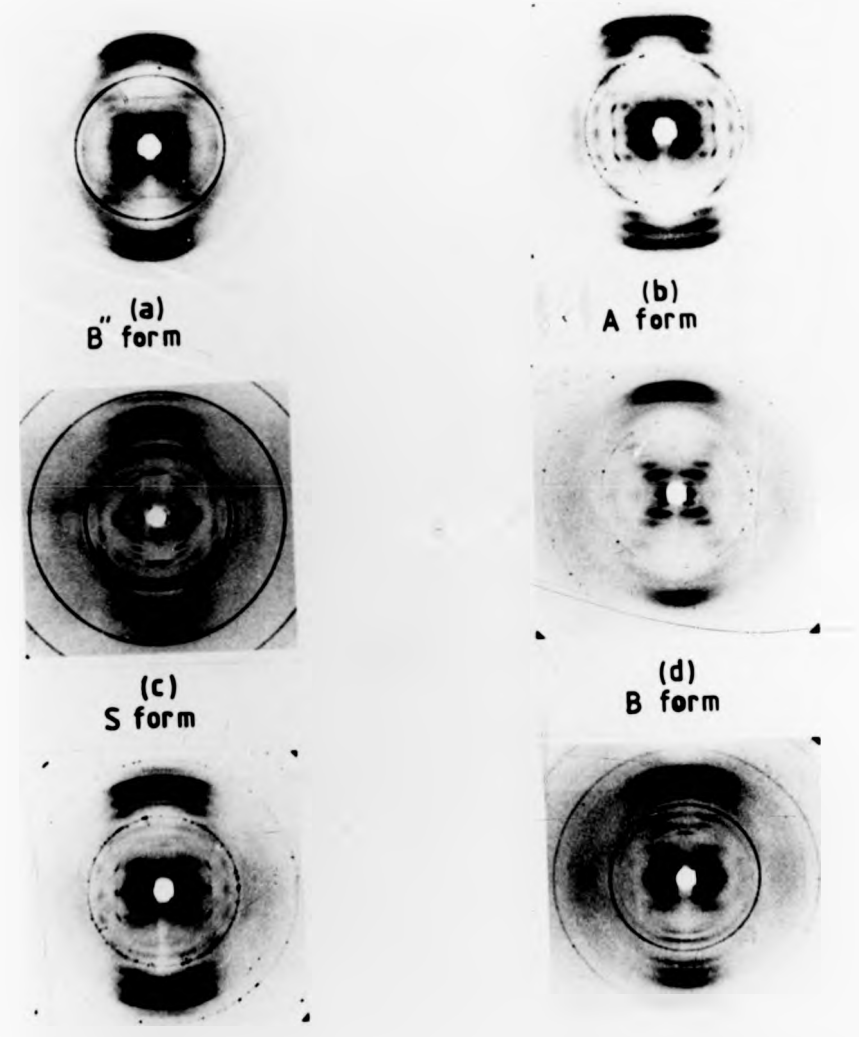
(d)
B form

(e)
mixture of A and S forms

(f)
mixture of S and B forms

PLATE 6.1:

X-ray fibre diffraction patterns of poly d(G-C). poly d(G-C) at various relative humidities



(a)
B" form

(b)
A form

(c)
S form

(d)
B form

(e)
mixture of A and S forms

(f)
mixture of S and B forms

PLATE 6.1:

X-ray fibre diffraction patterns of poly d(G-C). poly d(G-C) at various relative humidities

98% to 33%, the S form was not observed at all, and a transition occurred from the semi-crystalline B form to the A conformation and later to the B" form. This study (Mahendrasingam et al., 1983) is believed to be the first account of experiments in which structural transitions are seen to depend on the rate at which the RH of the fibre environment is changed.

The overall sequence of transitions found to occur in Na poly d(A-T).poly d(A-T) can be summarised as shown in Figure 6.2. Plate 6.2 shows the diffraction patterns corresponding to the different structures. The sodium salt of this alternating polynucleotide is found to start off (at low humidity) in a semi-crystalline C form which changes to a classical A form at 75% RH. An α-D conformation is observed at 92% RH,

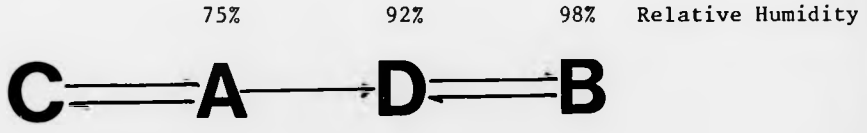
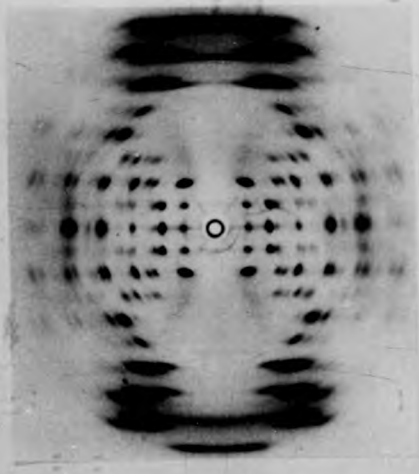


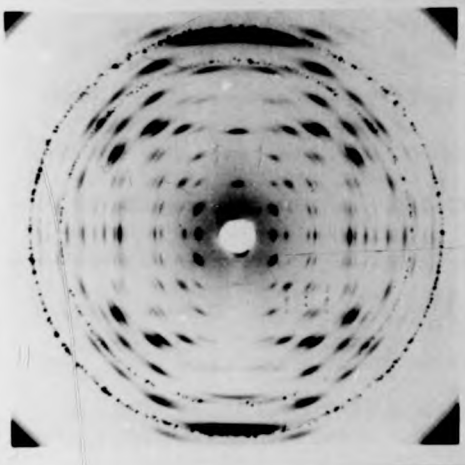
FIGURE 6.2 The sequence of transitions observed in the sodium salt of poly d(A-T).poly d(A-T).

after a period of annealing at this humidity. A semi-crystalline B form is induced at 98% RH. The reverse sequence, as indicated in Figure 6.2 is not identical to the forward sequence: the D form can be recovered from the semi-crystalline B but it seems that the A and C conformations can by no means (even by rewetting of the fibre) be recovered from fibres which have undergone the forward sequence into the D conformation. The D form thus emerges as a particularly stable conformation which is interconvertible with a high humidity B but not with the lower humidity A and C forms. Another interesting feature of the behaviour of poly d(A-T).poly d(A-T) is found

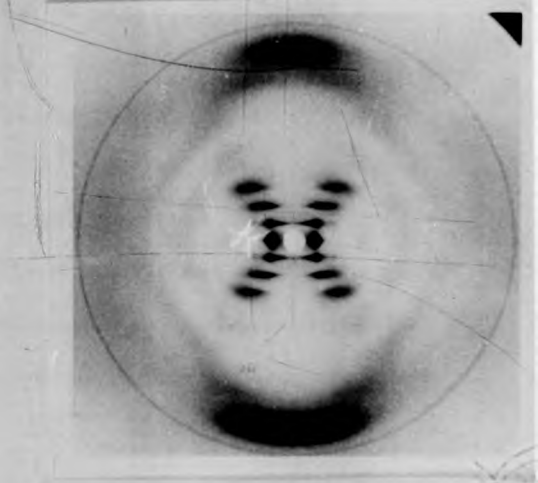
(a) C-form



(b) A-form

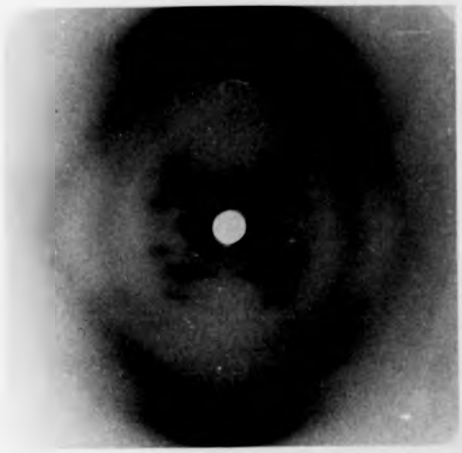


(c) D-form

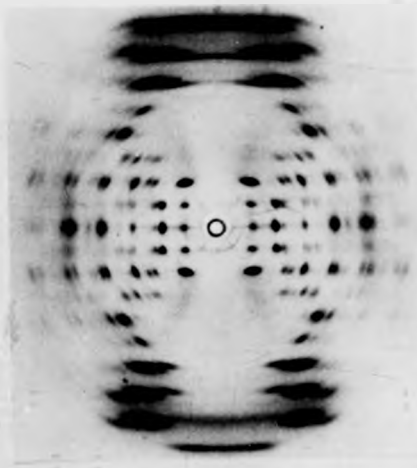


(d) B-form

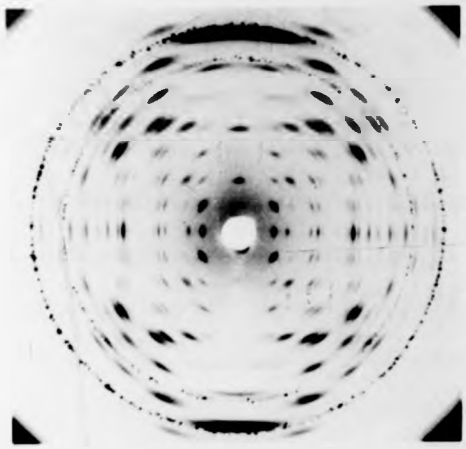
PLATE 6.2: The various conformations seen to occur in a single fibre of KF poly d(A-T).poly d(A-T) as a function of relative humidity.



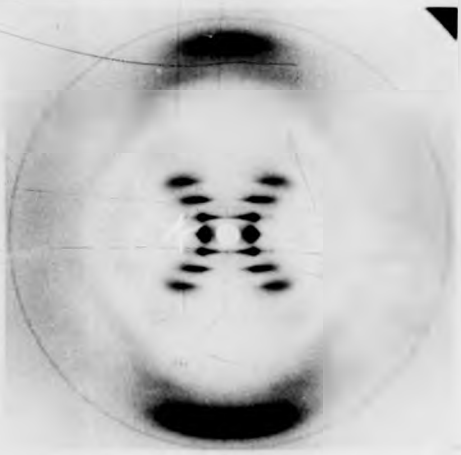
(a) C-form



(b) A-form

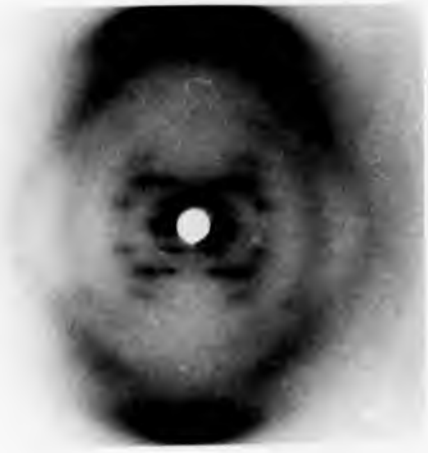


(c) D-form

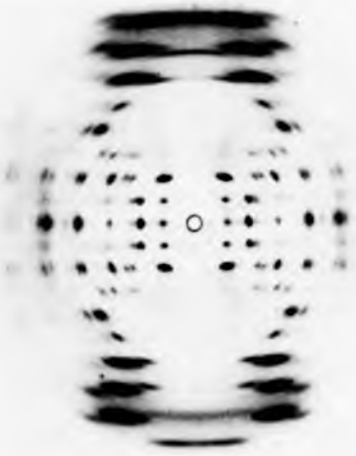


(d) B-form

PLATE 6.2: The various conformations seen to occur in a single fibre of KF poly d(A-T).poly d(A-T) as a function of relative humidity.



(a) C-form



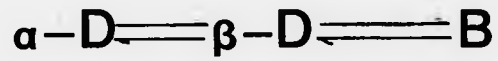
(b) A-form

(c) D-form

(d) B-form

PLATE 6.2: The various conformations seen to occur in a single fibre of KF poly d(A-T).poly d(A-T) as a function of relative humidity.

in the polymorphism of its potassium salt. This has been observed to undergo a sequence of transitions that is schematically represented by:



so that poly d(A-T).poly d(A-T) is evidently capable of adopting two distinct D conformations.

The fact that the C and A forms of the sodium salt cannot be retrieved upon rewetting is puzzling - the conformation present in the rewet gel is evidently not the same as that in the original. The assumption that DNA usually adopts a 'B-like' conformation in solution has been based on many studies that relate to various native DNA's (Bram, 1971, a,b,c; 1972; Griffith, 1978, Levitt, 1978; Hogan et al., 1978), but recent work has shown not only the presence of left-handed DNA in some crystalline oligomers and in solution, but also evidence of its existence in the cell (see Section 6.1.1 and associated references). It is tempting to suggest that the rewet gel may contain left-handed molecules in contrast to the original which is presumed to contain right-handed helices. However, this idea is obscured by the observation that such a gel produces a fibre which is composed of D molecules: these are supposedly right handed (Arnott et al., 1974). It is almost totally inconceivable that left-handed molecules prevailing in solutions should crystallise into an array of right-handed helices in a fibre. However, if D-DNA was itself left-handed then the behaviour of poly d(A-T).poly d(A-T) as outlined in Figure 6.2 would almost make sense. Interms of this hypothesis, one final question then remains: what is the nature of the high humidity B form which should in principle be very similar to the gel state which ensues upon rewetting?

Whilst all these observations and ideas relating to the structure and transitions of poly d(G-C).poly d(G-C) and those of poly d(A-T). poly d(A-T) are of great interest in themselves, there is a further interesting point to be made from a comparison of the two sequences as outlined in Figures 6.1 and 6.2. It can be noted that apart from the details of the reverse sequence, the two polynucleotides show similar behaviour: the S and the D conformations emerge (by comparison) as analogues and have many features in common. Both S and D are difficult to acquire from conformations assumed to be right-handed, and when they are finally obtained, both are very stable. Another fairly striking similarity is the fact that both S and D have distinguishably different variants although in D these are observed in the presence of a different cation (potassium). The diffraction patterns obtained from the S_I and S_{II} structures are compared in Plate 6.3, and those obtained from the α-D and β-D conformations in Plate 6.4. It is possible that the S_I and S_{II} structures may correspond to the Z_I and Z_{II} structures of Wang et al., (1981). The S_I and S_{II} molecules are both observed in hexagonal lattices, but the pitch of S_{II} (45.3Å) is 7% larger than that of S_I (42.8Å). The α-D and β-D structures pack into different lattices (the former tetragonal with a = b = 17.0Å, c = 24.1Å and the latter hexagonal with a = b = 21.2Å and c = 25.4Å) and there is again a 7% difference between the pitches of the two conformations.

Whilst the observations expressed in this section suggest that the D and S forms could be thought of as analogues, there is no definitive evidence to prove that these alternating purine/pyrimidine duplexes have any more in common than the features of their chemical behaviour that have already been outlined. However, the experiments do provide enough stimulus to consider the possibility (given that S-DNA is thought to be left-handed) of D-DNA being left-handed.

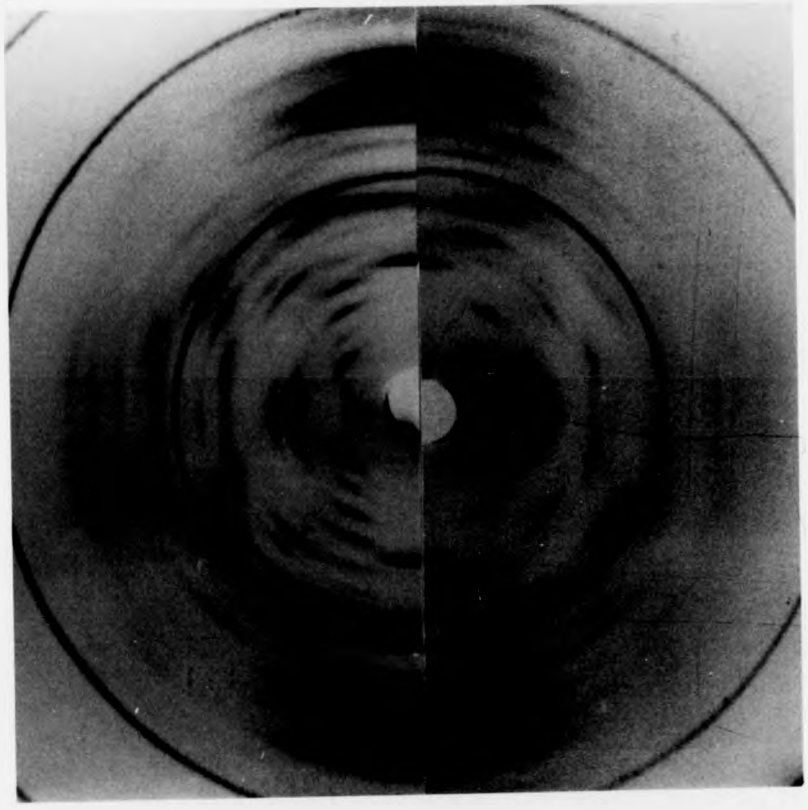


PLATE 6.3 Photographs comparing the S_I(left) and S_{II}(right) diffraction diagrams as observed from a fibre of KF poly d(G-C).poly d(G-C)

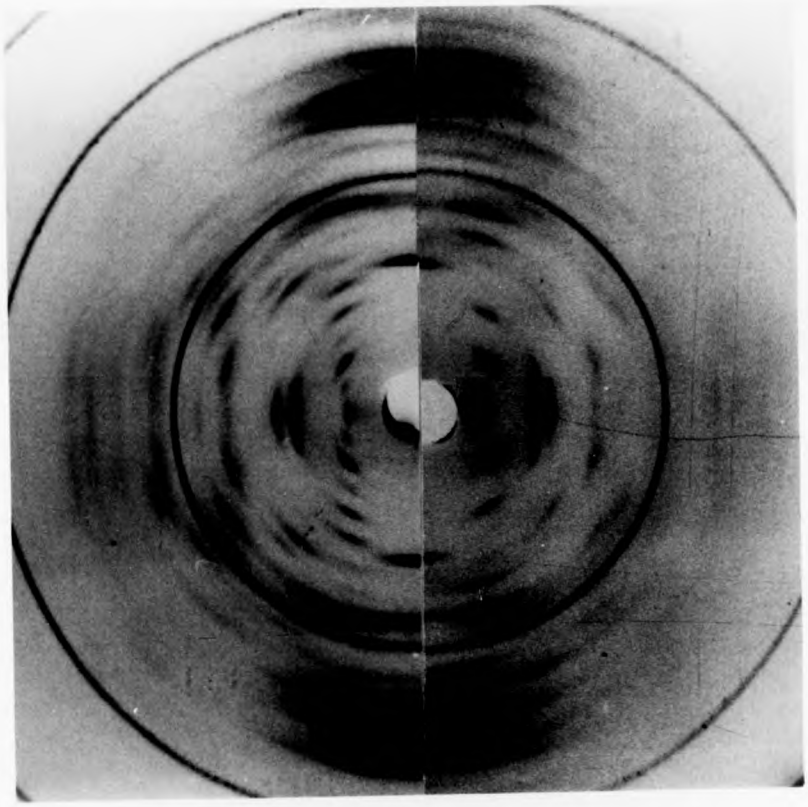


PLATE 6.3 Photographs comparing the S_I(left) and S_{II}(right) diffraction diagrams as observed from a fibre of KF poly d(G-C).poly d(G-C)

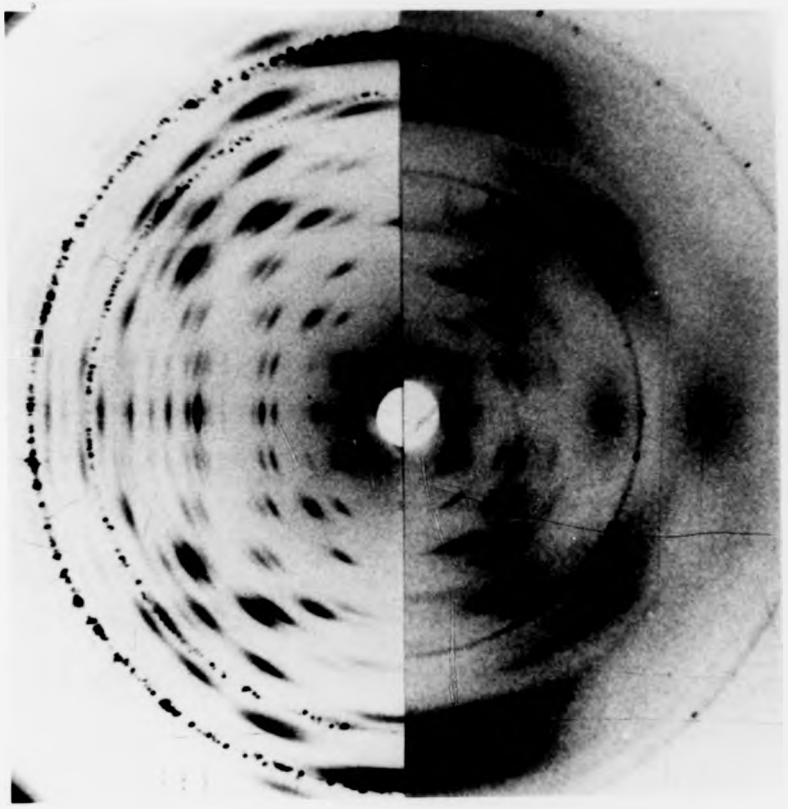


PLATE 6.4 Photographs comparing the α-D (left) and β-D (right) diffraction patterns taken from a fibre of KF poly d(A-T).poly d(A-T)

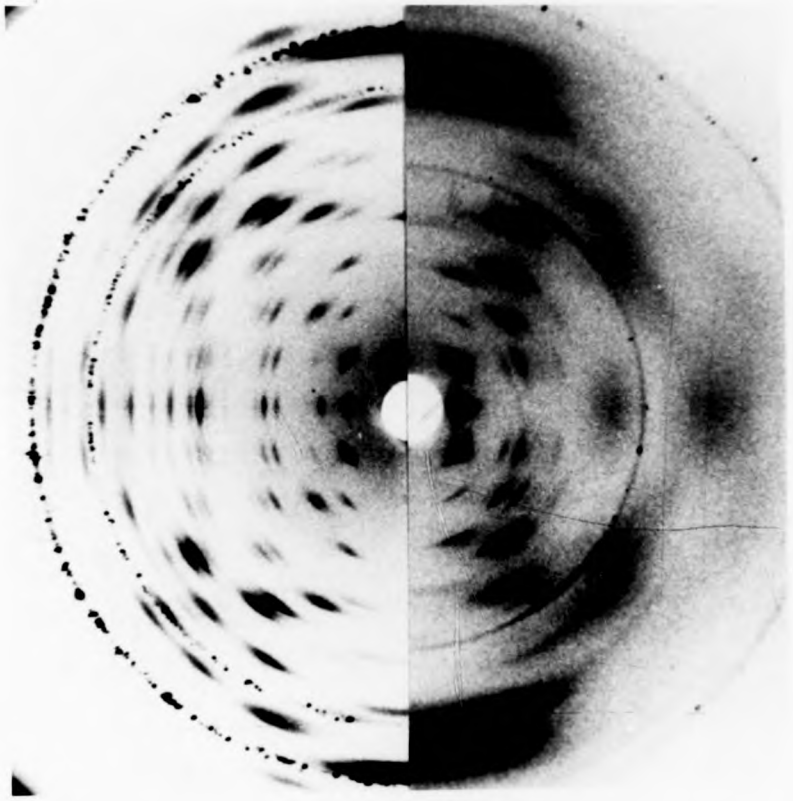


PLATE 6.4 Photographs comparing the α-D (left) and β-D (right) diffraction patterns taken from a fibre of KF poly d(A-T).poly d(A-T)

6.1.3 Models that have been suggested for D-DNA

The first D conformation to be observed was the α form in the sodium salt of poly d(A-T).poly d(A-T) (Davies and Baldwin, 1963). These authors observed that many different conformations (A, D, A/D mixtures, B) could be obtained from one fibre and that transitions could occur between these forms as a function of RH. A report by Arnott etal. (1974) suggested that the D form was observed in conditions of minimal retained salt. In the same article, Arnott et al. emphasised the similarity between the patterns obtained from Na poly d(A-T).poly d(A-T) and those obtained by Mitsui et al. (1970) from Na poly d(I-C).poly d(I-C). They also asserted that they had obtained similar diffraction data from poly d(G-C).poly d(G-C). Mitsui et al. (1970) proposed a left-handed double helical model with Watson-Crick base pairing as a structure compatible with the observed data. The model had 8_7 symmetry and a helix pitch of 25.01Å. The authors claimed that all attempted right-handed models were inferior both in terms of agreement with the x-ray data and overall stereochemistry. Arnott et al. (1974) rejected this model and used a least-squares linked-atom refinement procedure (LALS) to generate right-handed models having 8_1 symmetry and resembling the B form. The conformational parameters of such B and D models are supplied in the Appendix. This D model was proposed to explain diffraction from all the alternating purine/pyrimidine polynucleotides mentioned above.

Gupta et al. (1980) and Ramaswamy et al. (1982) have surveyed a variety of right and left-handed models for α-D-DNA and conclude that on the basis of fibre diffraction data alone, it is not possible to express strong preference for either the best left or the best right-handed models that have so far been generated. Drew and Dickerson (1982) have published a left-handed model for the α-D data which has 7_6 symmetry and Hoogsteen base pairs, and which they claimed to provide better agreement than any

models hitherto published. However, in a recent publication, Mahendrasingam et al. (1983) gave an account of a systematic study of Na poly d(A-T). poly d(A-T) in which they make it clear that the diffraction observed on the seventh layer line is due to the 107 and 117 reflections rather than the 007 (see Plate 1.5) and that, as such, there must be eight residues per helix pitch rather than the seven supposed by Drew and Dickerson.

Mahendrasingam et al. (unpublished) have also generated 8_7 models which are in better agreement with the observed diffraction than the currently accepted right-handed model of Arnott et al. (1972). These authors emphasise the need to exercise caution in asserting the handedness of this molecule and also suggest that the structure is likely to have a more complex asymmetric unit than the mononucleotide repeat that has thus far been assumed. A molecular model for D-DNA having a dinucleotide repeat is currently being developed in this laboratory.

The first report of the β-D conformation came from Arnott et al. (1975) in relation to a study of the sodium salt of poly d(A-T-T). poly d(A-A-T). These workers found that the β-D form occured in fibres when the humidity was between 85% and 0% on the return half of a humidity cycle. Above 85% they observed a semi-crystalline B conformation. The authors gave no indication of what happened in the first half of the humidity run. Arnott et al. (1975) claimed that α-D-DNA and β-D-DNA were isomolecular and differed only in lattice packing. They generated another 'B-like' model to fit the β-D data (see Appendix for model parameters).

The remainder of this chapter is now used to describe the results and conclusions drawn from an analysis of a β-D diffraction pattern obtained from KF poly d(A-T).poly d(A-T) by Dr. A. Mahendrasingam in this laboratory. The best right and left-handed models for β-D-DNA are

compared with those for α-D-DNA, and finally a comparison is drawn between the left-handed D models and the left-handed S models in a bid to pursue the question of analogy between the two.

6.2 THE STRUCTURE OF β-D-POLY d(A-T).POLY d(A-T)

6.2.1 Experimental

The poly d(A-T).poly d(A-T) was purchased from Boehringer, and the material was prepared by precipitation from 0.1 M KF solution. X-ray diffraction photographs were recorded using an Elliott GX6 rotating anode generator in conjunction with a Searle camera and Elliott toroidal optics. The lattice was measured using a travelling microscope and intensities were calculated from traces produced by a Joyce-Loebl microdensitometer.

A variety of left- and right-handed models were generated: initially measurements were taken from scaled models, but these were later refined to include precise covalent stereochemistry using a model-building program as described in Chapter 2. These models were varied so as to minimise agreement indices R and R' as per equations 2.1 and 2.2

6.2.2 Results

The full β-D-DNA pattern obtained from KF poly d(A-T).poly d(A-T) is shown in Plate 6.5. The lattice was found to be hexagonal with $a = b = 21.2\text{Å}$, $c = 25.4\text{Å}$. These figures were obtained by incorporating the reciprocal space measurements of all clear Bragg reflections and refining the lattice parameters to give best fit with this data. The observed and calculated ρ values are given in Table 6.1 along with the observed relative intensities and the corresponding structure amplitudes.

6.2.2.1 Best Right Handed (RH) Models

A number of right handed models were generated such that base displacement, base orientation and conformation angles were varied to

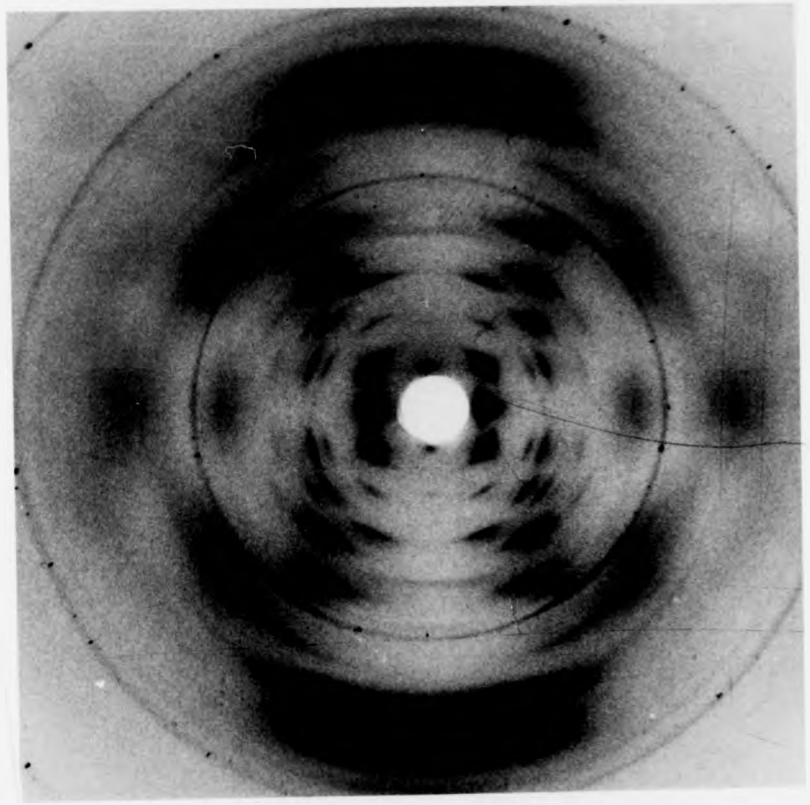


PLATE 6.5 A β-D diffraction Pattern obtained from a fibre of KF Poly d(A-T).poly d(A-T).

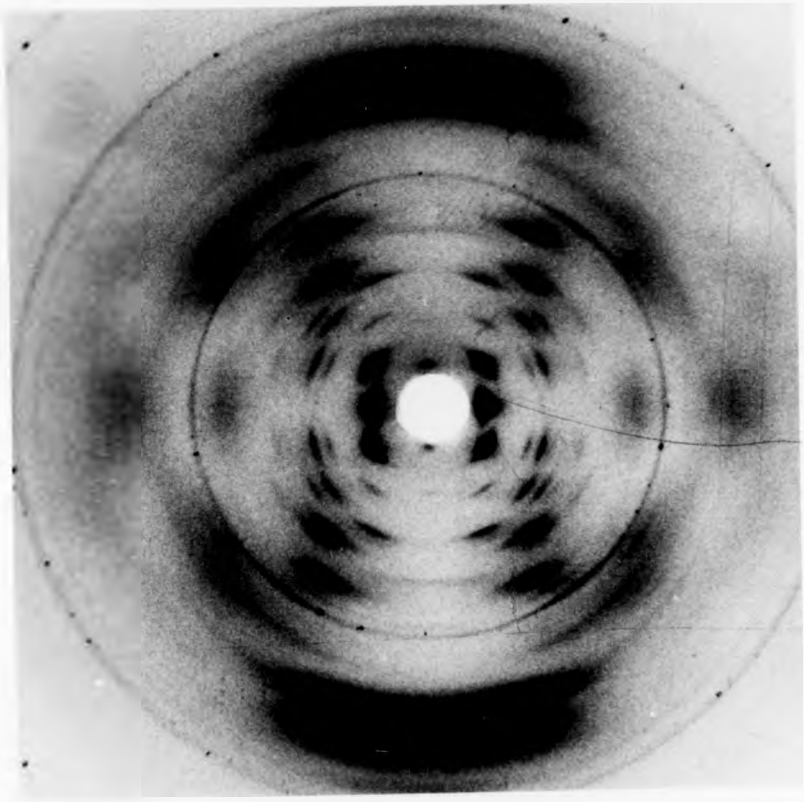


PLATE 6.5 A β-D diffraction Pattern obtained from a fibre of KF Poly d(A-T).poly d(A-T).

REFLECTION (HKL)	ρ_{obs} (A^{-1})	ρ_{calc} (A^{-1})	I_{obs} (Arbitrary units)	F_{obs} (Arbitrary units)
100	0.0543	0.0547	4473	3861
101	0.0675	0.0674	3073	3200
111	0.1021	0.1026	394	1147
201	0.1169	0.1162	1036	1858
211	0.1500	0.1500	297	996
102	0.0955	0.0959	426	1192
112	0.1229	0.1232	2544	2912
202	0.1343	0.1347	1814	2459
103	0.1311	0.1302	1754	2418
113	0.1507	0.1514	4343	3805
203	0.1603	0.1610	2432	2847
114	0.1840	0.1838	2915	3117
204	0.1920	0.1917	3523	3427

Lattice hexagonal with $a = b = 21.2\text{Å}$

$c = 25.4\text{Å}$

TABLE 6.1 The observed and calculated ρ values, observed intensities and observed structure amplitudes measured from the pattern shown in Plate 6.5.

optimise agreement with the observed diffraction. The best agreement index was obtained from the Arnott model itself, which has a structure reminiscent of B-DNA but in which the helices are very much more tightly wound about base pairs displaced 1.8Å from the helix axis. Although the bases alternate between A and T down a given chain, the model maintains a regular phosphodiester backbone characteristic of a mononucleotide repeat. Figure 6.3 shows the cylindrically averaged squared transform of this model. The observed and calculated structure factors are shown in Table 6.2. The overall stereochemistry is satisfactory in all respects other than a rather short intrahelical contact of 2.4Å between the deoxyribose C_2 and phosphate O_3.

Whilst the residual, R, of 28% for this model is reasonably low, there are some problems with the x-ray fit in regions corresponding to the (102), (112), (202) and (212) reflections. This is a defect that was acknowledged by Arnott et al. (1975) who claimed that the fit could be drastically improved in this region by making trivial changes to the furanose ring geometry. A variety of changes to the ring puckering were attempted by the author, but it was found that the overall fit could not be improved without severely deleterious stereochemical effects.

6.2.2.2 Best Left-Handed (LH) Models

Much the same approach towards molecular modelbuilding was applied to the construction of left-handed D models, starting from general deductions relating to an overview of the intensity distribution as a whole, followed by gradual refinement to improve the x-ray fit thereafter. The first LH models had 5'3' polynucleotide backbones: no model having satisfactory overall stereochemistry provided adequate agreement with the x-ray diffraction data.

Following the work of Mahendrasingam (1984), Greenall (1982) and that of Hopkins (1981), models were then developed such that the chain

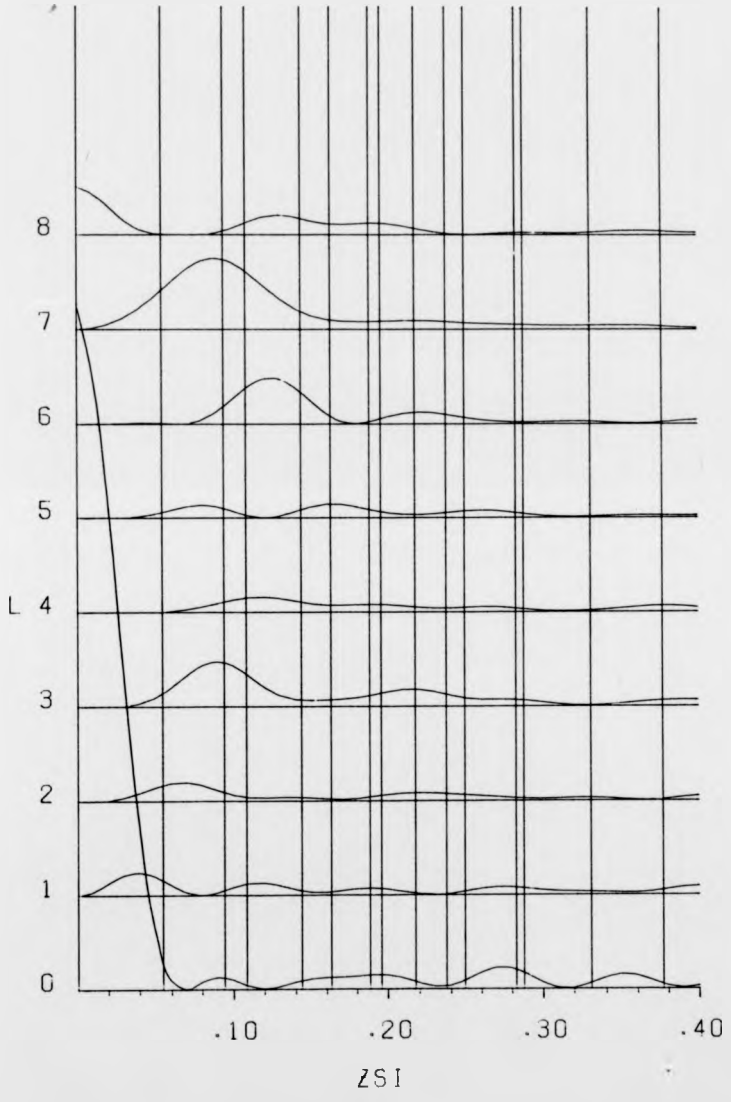


FIGURE 6.3 Cylindrically averaged squared transform of the Arnott D model (Arnott et al., 1974) shown here with a superimposed index grid such that a = b = 21.2A, c = 25.4A.

REFLECTION (HKL)	F_{obs} (ARBITRARY UNITS)	F_{calc} (ARBITRARY UNITS)
(100)	3861	3678
(101)	3201	2710
(111)	1147	1515
(201)	1858	2362
(211)	996	1985
(102)	1192	2539
(112)	2912	1852
(202)	2459	1226
(103)	2418	2431
(113)	3805	4446
(203)	2847	3796
(114)	3117	2306
(204)	3427	2698

R = 28%

TABLE 6.2 The structure factors calculated from the Arnott D model (Arnott et al., 1974) compared with those observed from the β-D pattern shown in Plate 6.5

direction was 3'5' (referred to by Greenall as β stacking) instead of 5'3' (referred to as α stacking). Several very good models were produced, the best of which had bases displaced so that D = 0.2Å, tilted about the diad by $8\frac{1}{2}^{o}$ and twisted by 10^{o}. The residual of 16.2% compares very favourably with all the other right and left-handed models thus far mentioned. Figure 6.4 shows the cylindrically averaged squared transform of the model, and the observed and calculated structure factors are shown in Table 6.3. The discrepancy between F_{obs} and F_{calc} for the (211) reflection is noticeably large, but an inspection of Plate 6.5 indicates clearly that this reflection should (if such a scheme were used) be given a very low 'weighting' and as such, the discrepancy is not serious. The coordinates of the model are given in Table 6.4 (see Appendix for torsion angles). The overall stereochemistry of this model is satisfactory, with no intrahelical or interhelical contacts shorter than 2.65Å. Figure 6.5 shows representations of the best RH and LH models in different perspectives.

6.3 DISCUSSION AND CONCLUSION

The behaviour of poly d(A-T).poly d(A-T) has been reviewed and compared with that of poly d(G-C).poly d(G-C). Although both polymers show comparable behaviour in a number of respects, that of poly d(A-T).poly d(A-T) is outstanding in the total irreversibility of one section of its polymorphism (see section 6.1.2). It has become apparent very recently that results from scanning microcalorimetry studies of poly d(A-T).poly d(A-T) in solution provide a very interesting corollary with this behaviour. Hinz and co-workers (private communication) have noted that there is an exceptionally high heat of enthalpy (ΔH) associated with the first transitions observed (presumably as a function of alcohol content in the solution - the details of this work have not yet been

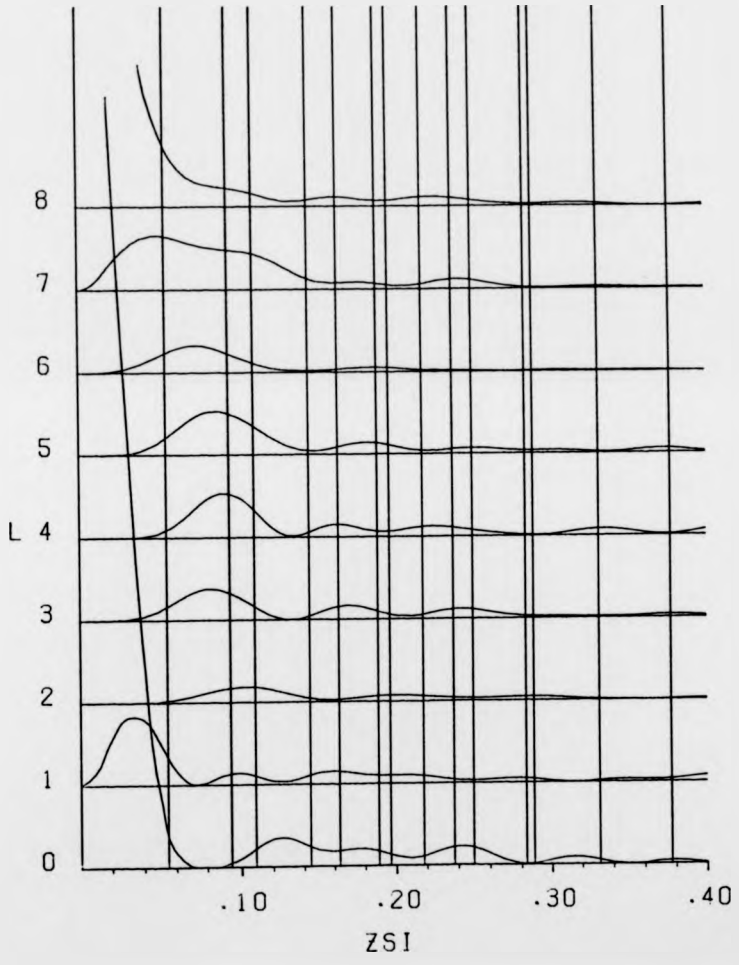


FIGURE 6.4 Cylindrically averaged squared transform of author's best LH model. The superimposed index grid is hexagonal such that a = b = 21.2A, c = 25.4A.

REFLECTION (HKL)	F_{obs} (ARBITRARY UNITS)	F_{calc} (ARBITRARY UNITS)
(100)	3861	3694
(101)	3201	2873
(111)	1147	1822
(201)	1858	1652
(211)	996	2073
(102)	1192	1785
(112)	2912	2240
(202)	2459	2056
(103)	2418	2640
(113)	3805	3720
(203)	2847	2633
(114)	3117	3698
(204)	3427	3261

Residual, R, = 16.2%

TABLE 6.3 The observed and calculated structure factors for the best LH model

	ATOM	R(Å)	φ(°)	Z(Å)	X(Å)	Y(Å)
PHOSPHATE	P	8.95	42.0	0.35	6.66	5.99
	O1	8.63	52.1	0.15	5.30	6.82
	O2	8.98	37.6	-0.95	7.11	5.47
	O3	10.23	41.8	1.09	7.63	6.81
	O4	7.80	37.8	1.27	6.16	4.78
DEOXYRIBOSE	C1	5.96	63.1	0.56	2.70	5.32
	C2	6.29	52.5	-0.38	3.83	4.99
	C3	7.70	55.7	-0.81	4.33	6.36
	C4	7.86	66.8	-0.74	3.10	7.23
	C5	7.62	71.9	-2.05	2.37	7.25
	O5	7.01	71.3	0.26	2.24	5.64
ADENINE	N1	1.10	27.1	-0.02	0.98	0.50
	C2	2.44	25.9	-0.11	2.20	1.07
	N3	3.44	41.9	0.05	2.56	2.30
	C4	3.42	64.1	0.35	1.50	3.10
	C5	2.69	85.9	0.48	0.19	2.68
	C6	1.31	92.8	0.28	-0.06	1.31
	N7	3.80	99.1	0.79	-0.60	3.75
	C8	4.77	87.6	0.84	0.20	4.76
	N9	4.68	71.3	0.62	1.50	4.43
	N6	1.51	148.8	0.37	-1.29	0.78
THYMINE	N1	4.68	71.3	0.62	1.50	4.43
	C2	3.53	61.1	0.32	1.71	3.09
	O2	3.84	43.1	0.07	2.80	2.62
	N3	2.37	75.7	0.35	0.59	2.30
	C4	2.81	104.5	0.62	-0.70	2.72
	C5	4.21	101.5	0.89	-0.84	4.13
	C6	4.94	87.0	0.86	0.26	4.93
	ME	5.14	115.4	1.18	-2.20	4.64
	O4	2.54	130.7	0.63	-1.66	1.93

TABLE 6.4 Coordinates of the best LH model for β-D-DNA. The associated torsion angles may be found in the Appendix. The diad related residue can be acquired by negating Y and Z. Successive nucleotides on one chain can be generated by applying a left-handed screw operation of 45° and a translation of 3.18A.

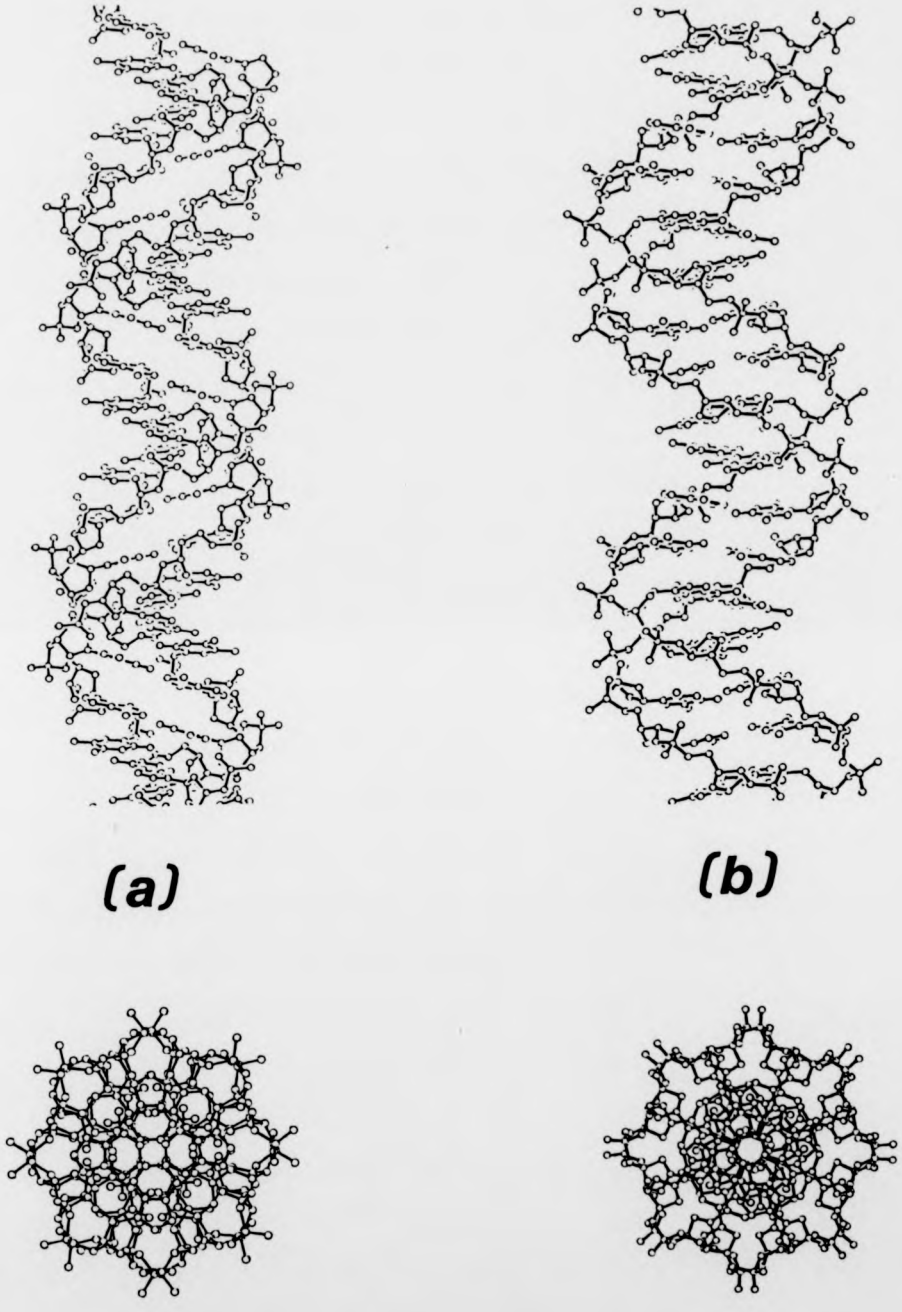


FIGURE 6.5 (a) The RH Arnott D model viewed perpendicular to, and directly down the helix axis
(b) The author's model seen in the same perspective.

obtained). Transitions involving a similar value of ΔH are never recorded thereafter in much the same way that the C and A forms of this polymer are never recovered once the D form is established. Whatever resemblance D-DNA bears to the classical right-handed C and A forms, it is thus separated from them by some sort of energy of activation: the question of helix chirality again springs to mind.

It is pertinent at this point to introduce the problem of the dynamics and stereochemistry of transitions involving right-handed α stacked helices and left-handed β stacked molecules. Wang et al., (1979) visualised the B $\rightarrow$ Z transition as involving a 'flipping' of bases through approximately 180^{o} along with changes in the positioning of the base pairs relative to the helix axis and compatible modification in chain conformation angles. A similar scheme is readily visualised for transitions involving D-DNA, but as yet it is puzzling that D-DNA is more readily interconvertible with B than A. One (albeit tentative) explanation would be that the high humidity B form is also left-handed: the diffraction data obtained from this B are sufficiently semi-crystalline and ill defined to allow for such a possibility without necessarily destroying the concept of classical right-handed B-DNA in other polymers. Whether or not a similar argument could be applied to the polymorphism of the poly d(G-C).poly d(G-C) double helix is, at this stage, conjecture.

The fibre data and associated analyses conducted here leave little doubt that the left handed model for β-D-DNA is superior in terms of its agreement with the observed diffraction. Further support to the concept of left-handed D models is provided by a previous analysis of the α-D conformation (Mahendrasingam, 1984). The Mahendrasingam model is very similar conformationally to the LH model described in this work (see Appendix). Whilst the β-D data only provides 13 measurable reflections, the α-D pattern analysed by Mahendrasingam contains 39 reflections and the

residual of 29% is thus very respectable.

The emphasis throughout this study has been on the handedness of D-DNA as a whole and β-D-DNA in particular. Whilst the analysis of β-D-DNA in this work and of α-D-DNA in that of Mahendrasingam (1984) supports a left-handed model for D structures, and as such, adds weight to the hypothesis extended in section 6.2, it is difficult to suggest a basis for analogy between S and D structures at a molecular level. A comparison of the two D models with the two S models (Mahendrasingam, 1984) reveals little bar handedness to show any direct structural analogy. The D models have an axial translation per dinucleotide of $\sim 6.2\text{Å}$ whilst the S has $\sim 7.4\text{Å}$. The rotation per dinucleotide in D-DNA is 90° whilst in S-DNA it is 60°. Both S-DNA and D-DNA have a novel β chain sense and although S-DNA has distinctive dinucleotide base stacking as compared with the mono-nucleotide of both LH D models, it is possible that future D models having a dinucleotide repeat may show similar features. In D-DNA the bases are displaced from the helix axis by distance of $\sim 0.2\text{Å}$ where in the S form they are arranged so that D $\sim -3.0\text{Å}$. The phosphorous atom lies $\sim 9\text{Å}$ distant from the helix axis in both D forms, whilst in S-DNA successive phosphorous atoms lie $\sim 6\text{Å}$ and $\sim 7\text{Å}$ radially away from the axis. The tilt and twist parameters (see Appendix) are totally different in the D and S forms. Figure 6.6. shows projections of the S_I, S_{II}, α-D and β-D helices as seen in two different perspectives.

If the D forms of poly d(A-T).poly d(A-T) are left-handed, then it is possible that as in the case of S-DNA in G-C rich regions (see section 6.1), there may be a specific role for D-DNA in A-T rich regions of chromatin. Further work on the D conformation is currently being undertaken in this laboratory by the author and his colleagues. Dinucleotide models for both α and β forms are being refined. Recent experimental work at the Daresbury SRS has produced a D pattern the quality of which is unsurpassed

FIGURE 6.6 Projections of four different helices as seen in profile (above) and down the helix axis (below)

(a) S_I-DNA (Mahendrasingam, 1984) (b) S_{II}-DNA (Mahendrasingam, 1984)

(c) α-D-DNA (Mahendrasingam, 1984) (d) β-D-DNA (Forsyth, this work)

by any nucleic acid fibre diagram (see Plate 1.10) and which will, without a doubt, enable the handedness of the conformation to be irrefutably determined. The pattern has crystalline reflections out to less than 2Å and should even permit some of the water in the structure to be resolved.

The evidence for the left-handed helices in the alternating co-polymers poly d(G-C) and poly d(A-T) is now strong enough to speculate on existence of left-handed structures in the polynucleotide poly d(A-C). poly d(G-T). As mentioned in the introduction, there is some evidence that S-DNA occurs in this duplex, although the quality of the data is poor and it seems equally possible that the conformation is actually D or something distinct from either S or D. The main problem with this investigation is in the acquisition of high quality material: that which was used to obtain the lithium C' patterns described and analysed in Chapter Four was obtained from Dr. J. Brahms at the University of Paris, but unfortunately, supplies appear to be limited. The corresponding polymer as obtained from the Sigma Chemical Company is not of sufficient quality for this sort of work, and although that supplied by PL Chemicals Limited has not previously merited detailed study, their material will shortly be sampled again.

CHAPTER SEVEN

THE EFFECT OF PROFLAVINE ON THE TRANSITIONS OF THREE DIFFERENT POLYNUCLEOTIDES

7.1 INTRODUCTION

The acridines dyes are cationic planar aromatic compounds and have attracted much attention because of their ability to bind to nucleic acids and to act as frame shift mutagens. The structures of these molecules are of further interest in that they resemble a number of powerful carcinogens. Many acridine drugs are bacteriostatic and this has found extensive clinical application. Proflavine and acriflavine are still used in the treatment of minor wounds. The ability of the acridines to stain nucleic acids has been, and still is, exploited by cytogists. Whilst many of the properties and applications of these molecules have been well known for some years, it was only the advent of the double helical model for DNA that provided a molecular basis for the interpretation of such phenomena.

One of the first studies of proflavine/DNA interaction was that of Peacocke and Skerratt (1956) who used spectrophotometric and equilibrium dialysis methods to investigate the binding processes involved. The spectrophotometric method has been particularly productive. These authors suggested the existence of two binding methods and hinted strongly that proflavine intercalation was one of them.

These results were put on a firm foundation when the work of Lerman (1961, 1963) and Luzzati et al. (1961) suggested a theory of interaction in which local unwinding of the sugar phosphate backbone accommodated for the intercalation of proflavine between bases

of the double helix (see Figure 7.1). This model was based on in vitro studies of proflavine and acridine orange in which binding to nucleic acids was noted to cause observable changes in their physical properties - most notably an increase in viscosity (due to lengthening and stiffening of the molecules) and also a decrease in sedimentation coefficient. The amino group reactivity of proflavine was found to diminish upon binding and this further supported the intercalation model. This theory has greatly contributed to the understanding of the binding behaviour of many drugs and of their biological properties. Intercalation models have since been constructed to explain the binding to DNA of other molecules such as the trypanocidal drug ethidium bromide (Fuller and Waring, 1963) and the anthrocycline antibiotic daunomycin (Pigram, Fuller and Hamilton, 1973). More direct evidence for intercalation has been supplied by the autoradiography studies of Cairns (1962) in which it was observed that the length of T2 phage DNA molecules increased from $\sim$ 50 μm to $\sim$ 72 μm so that $\sim$ 44% of the potential spaces between base pairs were occupied. Temperature jump relaxation studies have detailed the kinetics of proflavine/nucleic acid binding and have shown that the initial reaction is the one external to the helix and that this is followed by intercalation (Li and Crothers, 1969). The intercalation model is also supported by the x-ray diffraction work of Neville and Davis (1966). The work of Blake and Peacocke (1968) confirmed the involvement of two binding methods - one (the stronger) which was found to occur at relatively low drug levels in which the phosphate to drug ratio (P/D) was $\sim$ 5 and another weaker interaction which was observed right up to the electroneutrality limit at which P/D = 1. Both binding methods were found to diminish with ionic strength, but the second (weaker) interaction more so than the first. This implied that the second method of binding had a greater electrostatic component than the first.

Acridines also self-aggregate extensively around nucleic acids (Bradley and Wolf, 1959) and it has been shown that many of them dimerise

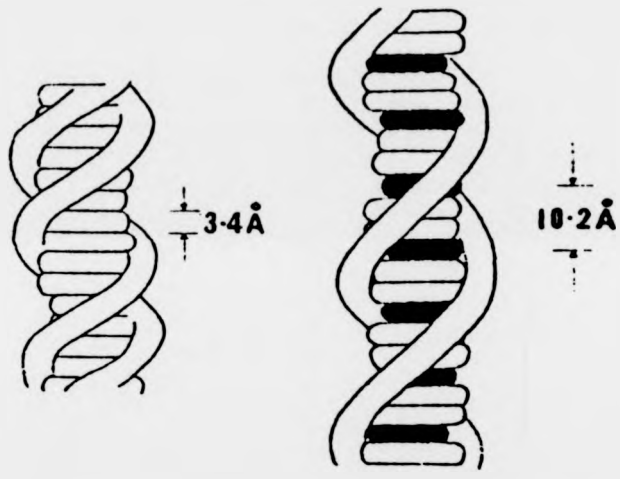


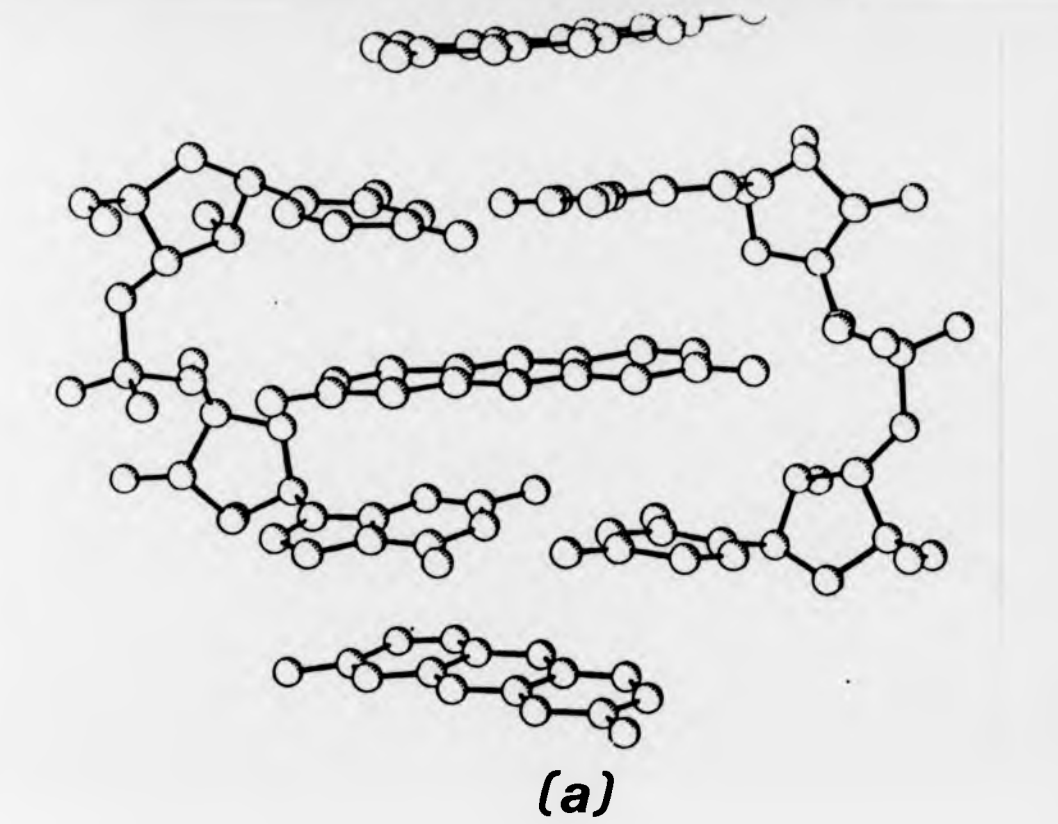
FIGURE 7.1 The intercalation model of Lerman. The left-hand side shows native DNA, the base pairs being represented as discs. The right-hand side represents drug molecules intercalated between base pairs (from Waring, 1968).

in solution and in the solid state. Studies on T2 bacteriophage DNA tend to corroborate this idea. The DNA of this virus has a much weaker external binding with proflavine and this is presumed to be a result of the fact that T2 DNA is glucosylated and hence prevents the stacking of proflavine that normally occurs in the major grooves of other DNA (Li and Crothers, 1969).

The influence of base content in the binding of acridines is still not conclusively established. Fluorescence quenching analysis of acriflavine bound to different DNA's has shown that the binding constant appears to increase with increasing A-T base content (Tubbs et al., 1964). Viscometric and temperature jump relaxation work by Ramstein et al. (1972) also tends to confirm this. These authors showed that the increase in length of the DNA molecule upon intercalation was greater for A-T rich DNA's (e.g. Clostridium perfringens DNA having $\sim$ 70% A-T) than it was for G-C rich DNA's (e.g. Micrococcus lysodeikticus DNA having 72% G-C). Kinetic analysis work (Li and Crothers, 1969) suggested that the external binding of proflavine to DNA is more prevalent in G-C rich regions than it is in A-T ones.

Neidle et al. (1977) have undertaken a single crystal study of a proflavine-cytidylyl-3'-5'-guanosine complex, and have detailed the intercalation and external binding in these crystals at atomic resolution (see Figure 7.2). Berman et al. (1979) verify this model and note that the intercalation of proflavine between bases seems to involve minimal unwinding. These workers maintain that only two torsion angles (δ and χ - see Figure 1.7) need change appreciably to allow inter-base separation to be extended from 3.4$\mathring{A}$ to 6.8$\mathring{A}$.

A variety of different models have been suggested to explain the frame shift mutagenic activity of the acridines. It was originally suggested by Lerman (1964) that intercalation occurred on one strand only during replication. However, this model has been largely abandoned on the



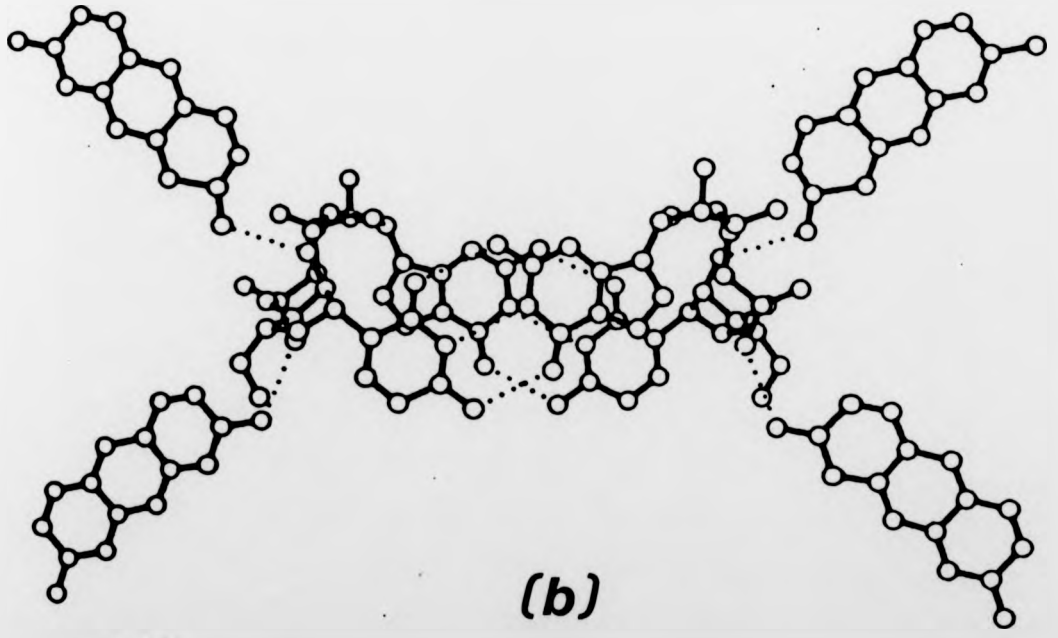


FIGURE 7.2 Two views of the proflavine-cytidylyl-3'-5'-guanosine complex viewed (a) perpendicular to the helix axis and (b) down the helix axis (from Neidle and Berman, 1983).

basis that there is no correlation between mutagenic activity and replication (Drake, 1964). The approach of Stersinger et al. (1966) has been more successful. By this model, acridine molecules stabilize 'looped-out' regions of DNA which occur in the first instance as a result of a strand break (a strand break could occur by misrepair or during replication). Peacocke (1969) has suggested that the acridines cause miscopying of DNA by intercalation into either the old or new strand causing respectively insertion or deletion of a base pair, and claimed that his modified model of intercalation (Pritchard, et al., 1966) provided a molecular basis for such action. Crick and Brenner (1967) have produced genetic evidence to show that acridine induced nutations do indeed correspond to insertion or deletion. Another model for mutagenesis is due to Streiber and Duane (1974) and suggests that only a G-C pair can initiate a frameshift, this premise then governing the occurrence of additions and deletions.

7.2 THIS WORK

Proflavine is thus an acridine that has been extensively studied in relation to native DNA by means of a variety of different techniques. Furthermore, its implication as a frame shift mutagen provides a unique approach to the problem of relating molecular structure to molecular function. This project aimed to survey the effect of the dye on the transitions of KF poly d(G-C).poly d(G-C) and those of KF poly d(A-T). poly d(A-T). Both these polymers are normally found to undergo a fairly dramatic sequence of transitions as a function of relative humidity and are here examined in the presence of proflavine. The activity of a Micrococcus lysodeikticus DNA/proflavine fibre is also examined. The emphasis throughout the study is largley qualitative, and relates more to the effect of provflavine on the scheme of transitions rather than direct structural effects.

The work was undertaken mainly to establish the possibilities that exist for future definitive investigations.

No ultraviolet absorption studies have been undertaken to establish phosphate/drug ratios, and when quoted, P/D values reflect crude calculations and are supplied only as guidelines. The rarity and expense of high quality poly d(G-C).poly d(G-C) has justified the reuse of material from previous studies and this has inevitably had an effect on the quality of diffraction data obtained from such samples. Proflavine was supplied in the form of proflavine hemisulphate. Diffraction patterns were recorded at each of the humidities 33%, 44%, 58%, 66%, 75%, 84%, 92%, 98% for every sample except where stated. No underfilms were used: this has resulted in some loss of data, and is described where appropriate but does not affect the conclusions reached concerning the effect of proflavine in the transitions. Fibres of DNA/drug complexes usually produce diffraction patterns that have considerably less detail than those of the polymer by itself, and the pictures shown here bear ample testimony to this. The molecular disorder that is evident in such patterns can be attributed either to effects of 'random' intercalation or to significant external binding, or both (Neville and Davies, 1966).

7.3 FIBRES MADE FROM M.LYSODEIKTICUS DNA/PROFLAVINE

7.3.1 Experimental

Micrococcus lysodeikticus DNA (originally stabilized with the sodium ion) was purchased from Boehringer. 2 mg of this DNA were allowed to dissolve in 2 ml of water over a period of 24 hours. 2 ml of 5 mM proflavine hemisulphate solution were then added. The resulting mixture was shaken thoroughly and left to stand overnight, after which it was centrifuged at 60,000 rpm for 12 hours. A fibre was then drawn from the concentrated gel using methods described in Chapter 2.

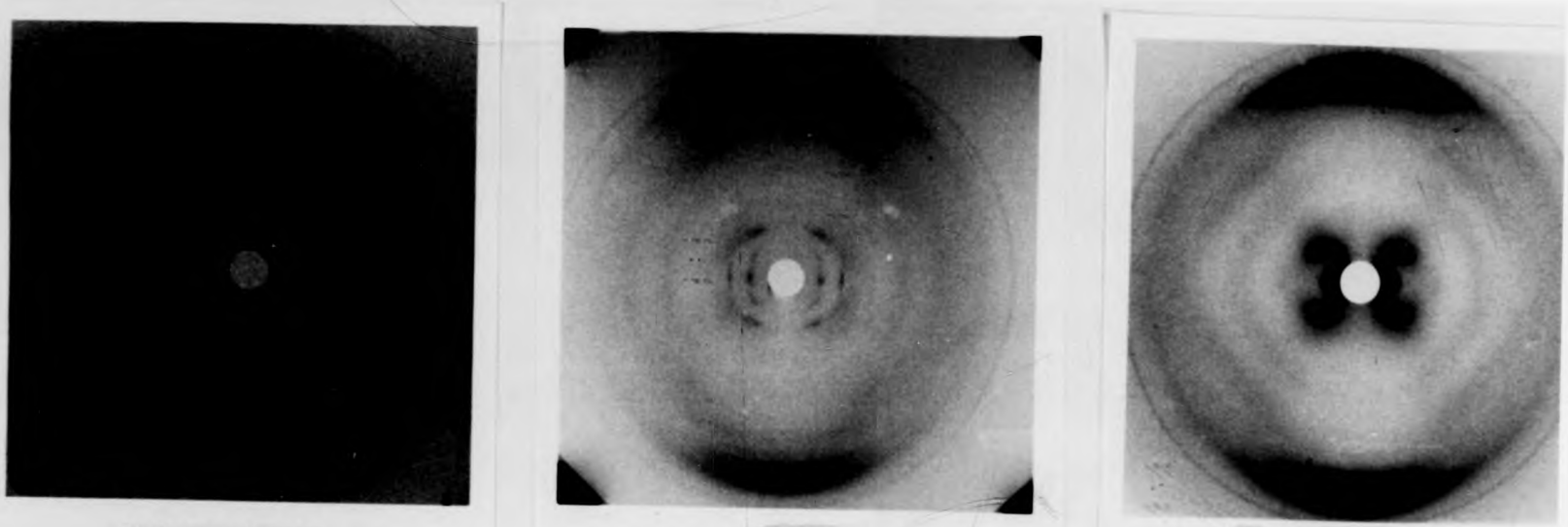
7.3.2 Results and Discussion

Fibres made from the sodium salt of native DNA's are ordinarily found to undergo a sequence of transitions as represented below:

$$C \overset{75\%RH}{\rightleftharpoons} A \overset{98\%RH}{\rightleftharpoons} B$$

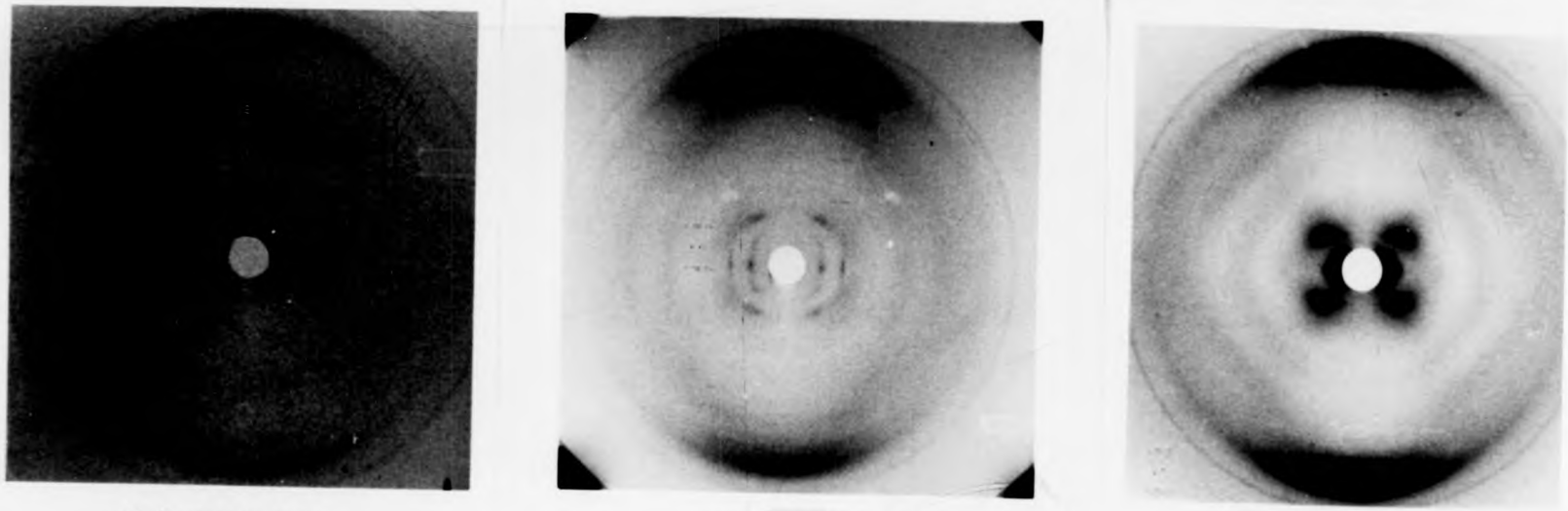
The incorporation of proflavine into this material resulted in an expected decrease in the internal ordering of the fibre, and made quantitative (and even qualitative!) assessment of the effect of the drug quite difficult. The most representative data of the set are displayed in Plate 7.1. Plate 7.1(a) was recorded at 33% RH and shows what should normally be at this humidity be a C pattern. However, the absence of detail in this picture demands further clarification. The layer lines are very badly defined and the pitch cannot be determined without making unwarranted assumptions about the symmetry of the molecule. The translation per residue is, however, clear and is calculated to be 3.33 Å. The position of what should be the (110) reflection of a hexagonal lattice can also be measured and turns out to be such that $\rho_{110} \sim 0.064$ Å^{-1} implying that the hexagonal cell side is ~ 31.2Å. The ratio of ρ_{merid}/ρ_{110} is ~ 4.7: this value is very close to that measured from C patterns and totally distinct from such a ratio in hexagonal or orthorhombic B patterns. On the basis of the features mentioned above and of the position of this conformation in the humidity run, it thus seems safe to assume that this structure is similar to C-DNA. It is not possible to determine the extent or nature of any bonding that may be occurring between this molecule and proflavine.

An increase of relative humidity results in the emergence of some of the features normally associated with A-DNA (see Plate 7.1(b) for the most clearly defined pattern). The pitch, as determined from the first two layer lines, is 28.1Å. However, the pattern lacks the characteristic higher layer line definition of A-DNA (see Plate 1.1). The badly defined



(a) 33% RH (b) 66% RH (c) 98% RH

PLATE 7.1: Diffraction patterns recorded from a M.Lysodeikticus/proflavine fibre at different humidities.



(a) 33% RH (b) 66% RH (c) 98% RH

PLATE 7.1: Diffraction patterns recorded from a M.Lysodeikticus/proflavine fibre at different humidities.

meridional occurs in a region corresponding to a spacing of $\sim 3.2\text{Å}$ and not the $\sim 2.6\text{Å}$ expected for a normal A form. Although the higher layer-line diffraction could, in fact, correspond to that from B-DNA, it has no corrollary on lower layerlines (especially the equator), as would be expected if some sort of mixture was occurring. It is possible that the pattern of Plate 7.1(b) represents an unusual distortion of the A helix induced by effects exterior to the molecule. Another possibility is that it represents a C/A mixture or intermediate (see 8.4.1).

A semicrystalline B form (as in Plate 7.1(c)) is observed at 98% RH: this helix has a pitch of approximately 41Å (as determined from the second layer line) and a rise per residue of $\sim 3.4\text{Å}$. This extended pitch suggests that intercalation has occurred so that there are two base pairs for every one turn of the helix. The intermolecular separation is difficult to measure from this pattern since an underfilm was not used during the exposure, but it is approximately 23Å. Wycoff (1955) has suggested that acridine dyes may bind in the large groove of DNA and maintains that this could cause a lateral shrinkage of the lattice: this behaviour has also been noted by Neville and Davies (1966). There is no evidence of such a shrinkage in this example, although the high water content of the fibre may prevent it anyway.

The results of this study illustrate a number of points that could be of value in relation to more definitive future studies. Firstly, it seems that proflavine, without actually intercalating, affects the acquisition of the A form. This presumably occurs by electrostatic effects outside the helix, and it is tempting to compare this observation with the effect of the lithium ion, which is suspected of preventing the occurrence of A-DNA altogether. This study has suggested that the $C \rightleftharpoons A \rightleftharpoons B$ transitions of DNA are significantly perturbed by the presence of proflavine. The A → B section of this polymorphism which is thought to be significant

in replication processes (Arnott, Fuller, Hodgson and Prutton, 1968) is of special interest in the context of different models for frame-shift mutagenesis (see section 7.1).

7.4 FIBRES MADE FROM KF POLY d(G-C).POLY d(G-C)/PROFLAVINE

7.4.1 Experimental

The poly d(G-C).poly d(G-C) was obtained in gel form from Dr. J. Brahms at the University of Paris. It was originally prepared by precipitation from 0.1 M KF. Fibres drawn from this material by Dr. A. Mahendrasingam indicated that it was of exceptionally high quality. The B"$\rightleftharpoons$A$\rightarrow$S$\rightleftharpoons$B transitional sequence (as mentioned in 6.1.2) was investigated in the first instance and (as a result of the scarcity of the material) fibres that had been used to maximum advantage for that purpose were then re-used for proflavine studies. In this work, the existing fibre was rewet with 10 µℓ of 1mM proflavine hemisulphate. The low concentration (1 mM) was chosen so as to avoid local precipitation effects. The low volume (10 µℓ) was used out of practical necessity. The phosphate to drug raio (P/D) was estimated as being $\sim$ 50. The fibre was re-drawn as described in Chapter Two.

7.4.2 Results and Discussion

A fibre of KF poly d(G-C).poly d(G-C) prepared in the right conditions of ionic strength (see Mahendrasingam, 1984) will give the sequence of transitions as outlined in Figure 6.1 and repeated here:

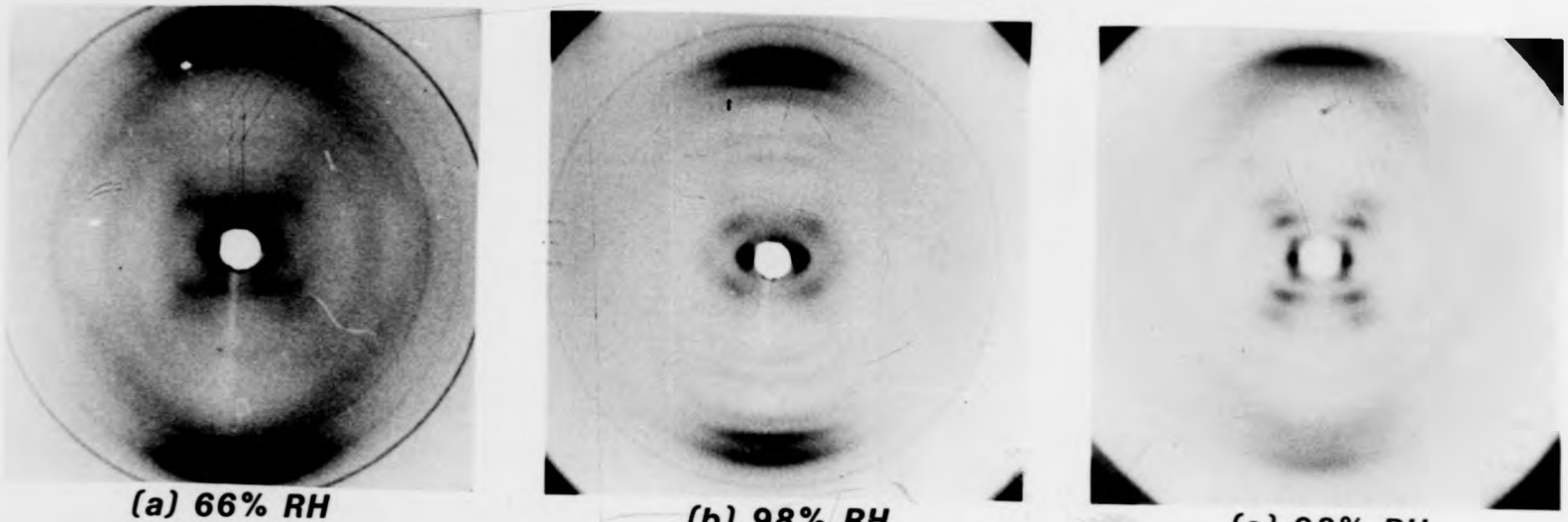
$$B'' \rightleftharpoons A + S \rightleftharpoons B$$

Fibres having slightly higher levels of KF are found to omit the A form so that the sequence is then:

$$B'' \rightleftharpoons S \rightleftharpoons B$$

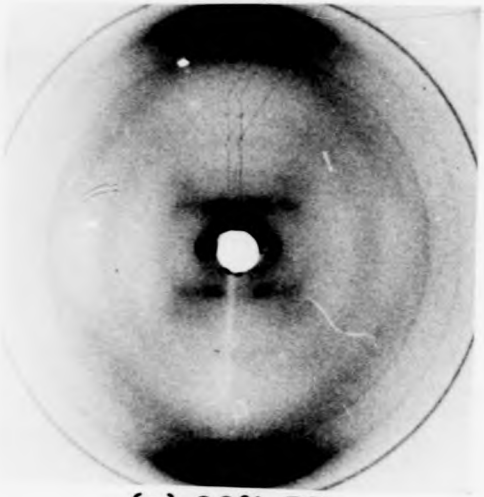
Fibre diffraction data recorded from the fibre used in this study is summarised in Plate 7.2. Three distinguishably different conformations are observed, corresponding to the B", S and B forms respectively. No A pattern was recorded at any stage in the sequence. Measurements from these photographs indicate that the B" molecule has a pitch of 33Å, a separation per residue of ∿ 3.3Å and that adjacent molecules are separated by ∿ 18Å. The S form appears to have a pitch of ∿ 43Å, a translation per residue of ∿ 3.4Å and an intermolecular separation of ∿ 22.4Å. The high humidity B form has a pitch of ∿ 33.5Å, a translation per residue of ∿ 3.3Å and its molecules are separated by ∿ 19Å. Within the range of experimental error, all these parameters are the same as those normally expected from the B", S and B conformations. It is pertinent to note here that the P/D ratio is such that intercalation effects could only be expected to produce (at the most) a pitch increase of ∿ 0.2Å. The diffraction data recorded here does not enable resolution of such a small change. It thus seems that the addition of a small amount of proflavine to this fibre has had a similar effect to that observed upon a slight increase in the amount of KF, and as such implies an external and electrostatic mode of interaction.

This study of poly d(G-C).poly d(G-C) in relation to proflavine was restrained by the scarcity of the material, and is clearly deficient in that it does not investigate drug interaction at lower P/D ratios. It is possible that the P/D ratio can be decreased without eliminating the S form or causing an unacceptable level of disorder in the fibre. Any evidence of interaction of acridines with S-DNA will be of great interest firstly from a structural point of view and secondly in that it may further illuminate the functional role of this conformation in vivo (see section 6.1.1).



(a) 66% RH (b) 98% RH (c) 98% RH

PLATE 7.2: Diffraction patterns recorded from a fibre of poly d(G-C)/proflavine at different humidities.



(a) 66% RH

(b) 98% RH

(c) 98% RH

PLATE 7.2: Diffraction patterns recorded from a fibre of poly d(G-C)/proflavine at different humidities.

7.5 FIBRES OF KF POLY d(A-T).POLY d(A-T)/PROFLAVINE

7.5.1 Experimental

The poly d(A-T).poly d(A-T) used here was purchased from Boehringer and was known to be stabilized by sodium chloride. About 0.9 mg of this material was mixed with ∿ 4 mℓ of a 10 mM solution of potassium fluoride and allowed to dissolve overnight. The solution was then centrifuged for 12 hours at 60,000 rpm and the sedimented gel used to prepare a fibre. This fibre was found to undergo the sequence of transitions that has been mentioned in 6.1.2, i.e.

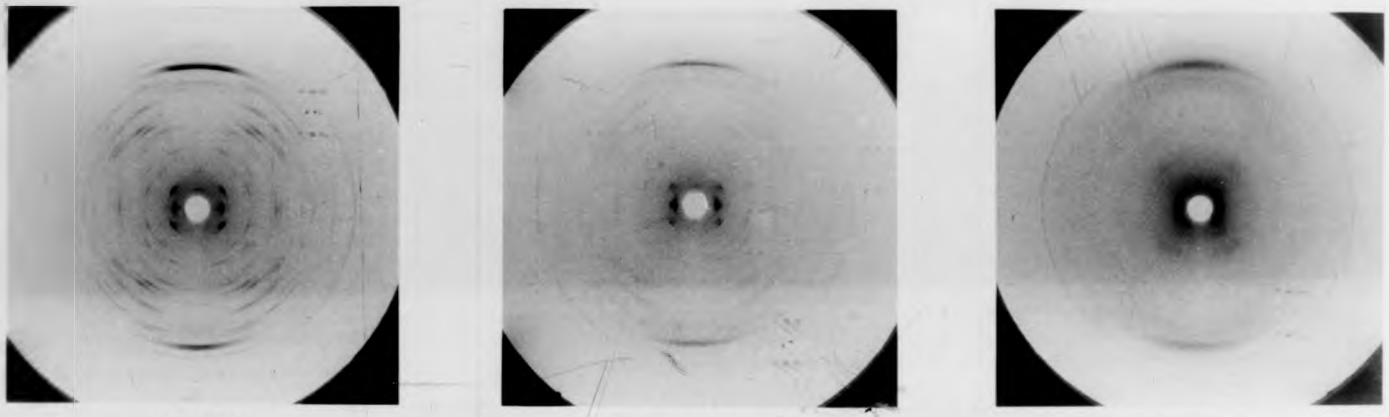
$$\alpha\text{-D} \underset{}{\overset{95\%\ RH}{\rightleftharpoons}} \beta\text{-D} \overset{98\%\ RH}{\rightleftharpoons} B$$

as a function of relative humidity. It was desired to investigate the effect of proflavine on this sequence and therefore the remaining gel (which was estimated to contain ∿ 0.7 mg of DNA) was redissolved in 2 mℓ of water (overnight) and then 0.5 mℓ of 1 mM proflavine hemisulphate was added. The rest of the tube was topped up in preparation for centrifugation and the effective concentration of the proflavine was calculated to be ∿ 0.1 mM. The solution was centrifuged as described above and a fibre was drawn which was thought to have a P/D ratio of ∿ 35. Diffraction patterns were recorded as described in Chapter Two and 7.2.

The gel remaining from above was again redissolved but this time with a greatly increased proflavine concentration (4 mℓ of 2.5 mM proflavine added to ∿ 0.5 mg of DNA). The mixture was allowed to dissolve overnight and then centrifuged as described previously. The P/D ratio here was estimated to be ∿ 2. Diffraction patterns were again recorded.

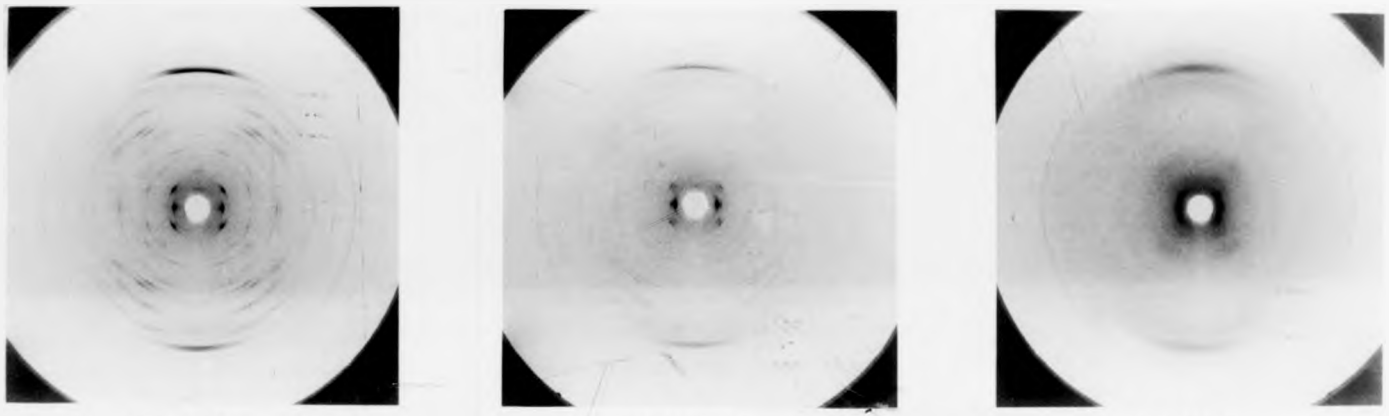
7.5.2 Results and Discussion

The first fibre produced a set of diffraction patterns as summarised in Plate 7.3 and showed no sign at all of being perturbed by the presence of proflavine in these concentrations. Plate 7.3(a) shows an



(a) 66% (b) 95% (c) 98%

PLATE 7.3: Diffraction patterns from a poly d(A-T)/proflavine fibre (P/D approx. 35) at different humidities.

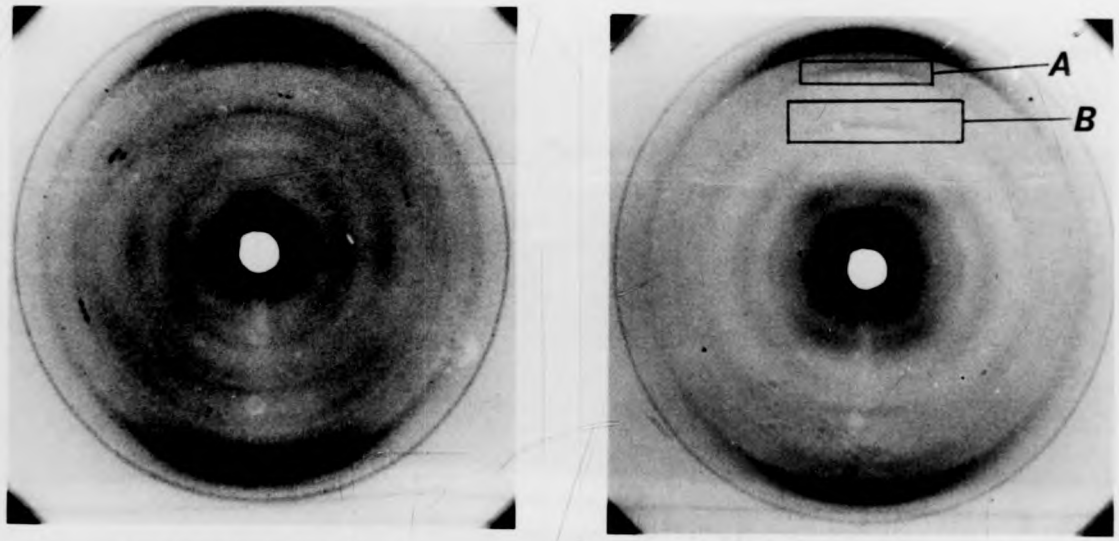


(a) 66% (b) 95% (c) 98%

PLATE 7.3: Diffraction patterns from a poly d(A-T)/proflavine fibre (P/D approx. 35) at different humidities.

α-D pattern, 7.3(b) a mixture of α-D/β-D diffraction, and 7.3(c) a semi-crystalline B form. It is believed that a good β-D pattern could have been obtained from this fibre but the humidity conditions necessary for its stabilization are delicate and were not obtained here. Plate 7.3(a) shows the α-D pattern to have a pitch of 23.2Å and an intermolecular separation of 17.4Å. The higher layer line meridional corresponded to a base separation of 2.9Å. All these values compare well with other α-D forms having a = 17.2Å, c = 24.0Å (Mitsui et al., 1970; Arnott et al., 1974; Mahendrasingam, 1983). Plate 7.3(b) shows the α-D/β-D mixture obtained at 75% RH: all the parameters associated with the lattices evident on this pattern are standard (see 6.1.2 and table 6.1) and there are no detectable signs of drug interaction. At 98% RH, a semi-cyrstalline B pattern was obtained: this was found to have a characteristic B intensity distribution with a pitch of 34.1Å, an intermolecular separation of 23.4Å and a residue separation of ∿ 3.4Å.

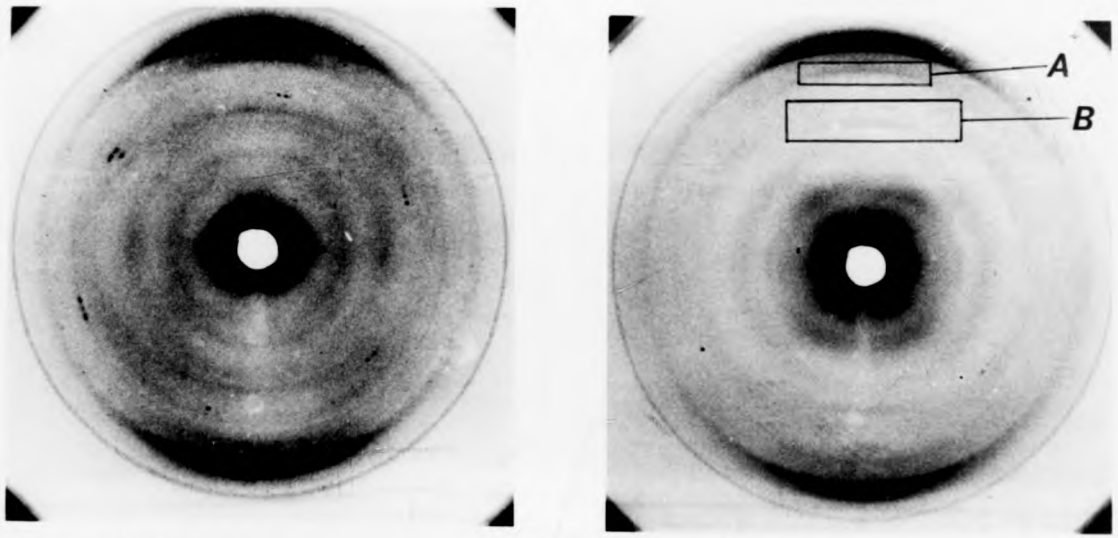
The second fibre, believed to have a P/D ratio of ∿ 2 produced diffraction patterns two of which are shown in Plate 7.4. The immediately noticeable features are the almost total loss of crystallinity and a considerable loss of orientation. Plate 7.4(a) shows what should correspond to the D pattern of Plate 7.3(a). As far as can be discerned, this conformation has a pitch of ∿ 25Å, an intermolecular separation of ∿ 18Å and a rise per residue of ∿ 3.5Å. There are some features of this pattern that are possibly more distinctive of the F pattern mentioned in Chapter 1 and displayed in Plate 1.8, although the quality of the data makes this very little more than speculation. The B pattern (Plate 7.4(b)) that was obtained at 98% RH was again poorly defined but appeared to have a helix pitch of ∿ 34Å, an intermolecular separation of ∿ 23Å and a residue separation of ∿ 3.4Å. In addition, there are features (see regions marked 'A' and 'B' on Plate 7.4(b)) that pertain to the previous



(a) 66% RH

(b) 98% RH

PLATE 7.4: Diffraction patterns obtained from a poly d(A-T)/proflavine fibre at different humidities(P/D approx. 2)



(a) 66% RH

(b) 98% RH

PLATE 7.4: Diffraction patterns obtained from a poly d(A-T)/proflavine fibre at different humidities(P/D approx. 2)

D form of plate 7.4(a) but which persist even after long periods at this high humidity. Given the high drug content of this fibre it is also unusual for no intercalation to have occurred.

7.6 FURTHER STUDIES

The studies described here were preliminary experiments to assess the value of a more extensive systematic investigation into the effect of acridines in general, and proflavine in particular, on a selection of polynucleotides. To date, no detail has been provided on the interaction of acridines with D and S-DNA and whilst this work has not provided such detail, it does provide a basis for further experiments in more controlled conditions. The results from poly d(G-C).poly d(G-C) are interesting but inconclusive since the drug content of the fibre was too low to produce detectable effects on the observed diffraction other than possibly inhibiting transitions. It is believed that the quality of the diffraction outlined in 7.4.2 and shown in Plate 7.2 can be improved by using fresh material. A systematic examination of the polymorphism of this polynucleotide as a function of proflavine content could further elucidate the biological significance of (G-C) rich regions of DNA. The results obtained from poly d(A-T).poly d(A-T) also warrant further study. The fact that the B form of this polymer did not allow intercalation with proflavine at such high drug levels is puzzling since it has been asserted that proflavine binds preferentially to A-T rich regions (Tubbset al., 1964).

There is a large amount of information available on the molecular conformation of nucleic acids and this is a unique situation in the study of drug receptors. Because most of the definitive fibre diffraction data on acridine/nucleic acid complexes has been recorded from random sequence DNA, molecular models have therefore tended to represent an 'average' polynucleotide chain in which subtle differences in conformation are not

detected. The availability of synthetic polynucleotides such as the two described here has illustrated that sequence effects may well promote differences in structural behaviour and it is likely that these differences will be of significance to the way in which acridine complexes are formed.

CHAPTER EIGHT

DYNAMIC TRANSITIONS WITHIN FIBRES AND TIME RESOLVED X-RAY MEASUREMENTS

8.1 INTRODUCTION

Work outlined elsewhere (see sections 1.4.1, 1.4.2, 6.1) has indicated that some of the conformations identified in single crystal, fibre and solution scattering studies of DNA are observed in nucleosome structures and in the intact cell. It is possible that the extensive polymorphism of nucleic acids is implicated in the remarkable structural behaviour of chromatin throughout the cycle of the mitotic (and meiotic) cell. The dynamics of DNA in the fibrous and liquid states is thus of crucial interest to our understanding of the biomolecular operation of genetic material *in vivo*. Whilst x-ray fibre diffraction has isolated the physical conditions necessary to observe certain discrete conformations, little is known about the transitions that occur between these structures and the role that water and ions must play therein. Such transitions in fibres are likely to depend not just on effects prevailing at the molecular and crystallite levels of organisation (see section 3.5) but also on features relating to the fibre and its surroundings as a whole. It is of obvious interest to investigate the stereochemical pathways associated with transitions known to occur in random- and repetitive-sequence DNA's, especially those that may involve a change in helix sense.

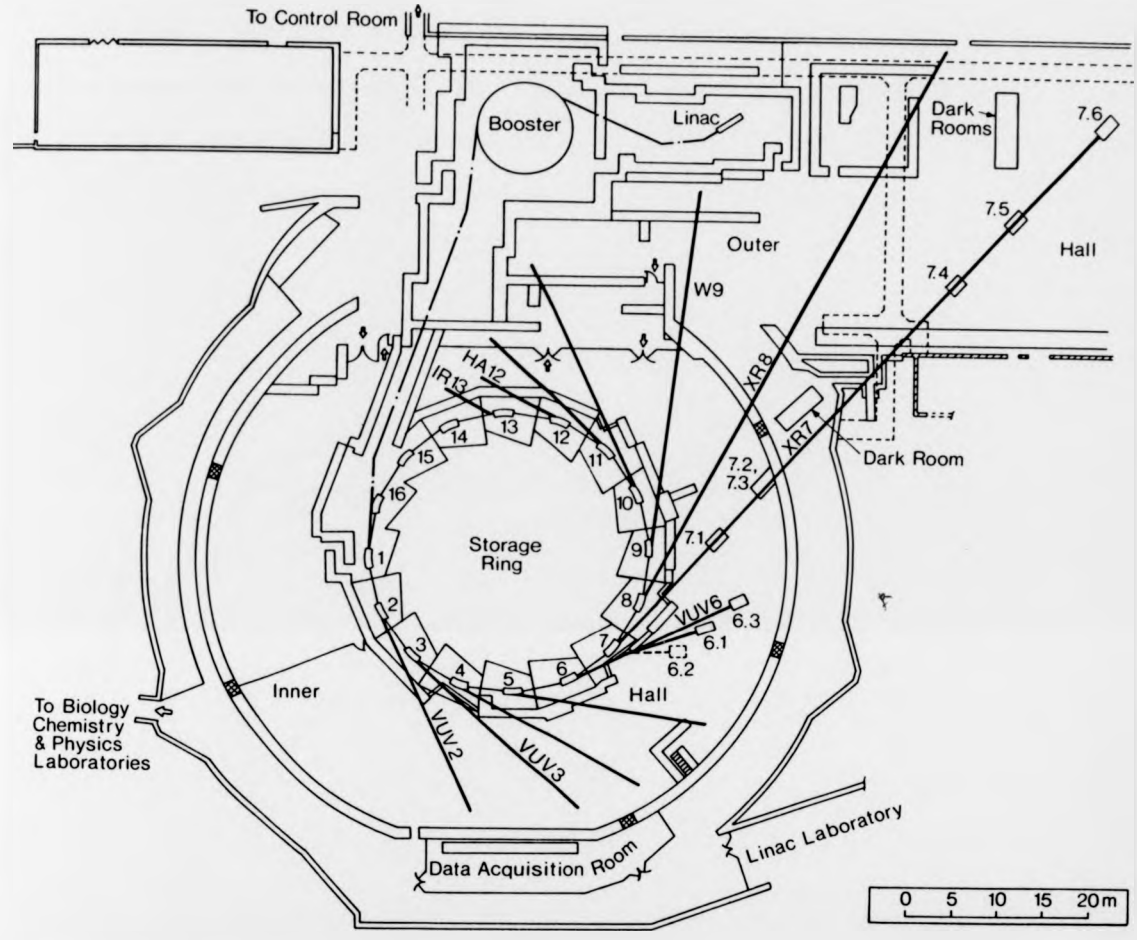
It has not hitherto been possible to observe transitions directly by x-ray diffraction since the use of low intensity conventional x-ray sources and relatively slow films has required long exposure times. However,

the recent availability of very high intensity synchrotron and storage ring sources, fast data acquisition systems and electronic x-ray detectors has reduced to the order of seconds the exposure times needed to identify the DNA conformation within a fibre. Since these samples are made up of many crystallites each containing an array of molecules (see section 3.5). and since the fibre as a whole will require time to 'conduct' water through its cross section, it is clear that diffraction changes observed as a function of time should be interpreted with some caution: for example, the stereochemistry of a transition could occur either as a result of the whole molecular population changing cooperatively in a particular way, or as a result of changes in the number of molecules (or even crystallites) undergoing rapid 'flips' from one molecular structure, (say A-DNA) to the next (say B-DNA). Hence it is quite feasible that many different effects are manifest in the diffraction data so obtained, and it is thus of great important that the independent physical variable associated with a given transition (usually relative humidity) is very carefully controlled. The implications of such time-resolved studies are of intense interest: carefully recorded data can be used firstly to assess the dynamics of various transitions and secondly to detail plausible molecular models which can be generated to satisfy 'intermediate' diffraction data.

The study that is described in this chapter reports the preliminary results from the first real-time studies on the conformational changes within DNA fibres as a function of time, using both film and a two-dimensional x-ray detector.

8.2 THE SYNCHROTRON RADIATION SOURCE (SRS) AT SERC DARESBURY LABORATORY, WARRINGTON, ENGLAND

The overall layout of the SRS is shown in Figure 8.1, together with its operational characteristics. The experimental apparatus was situated



INJECTOR
S-band electron Linac
10 – 15 MeV, 20 mA
3 GHz r.f.
10 Hz rep. rate

BOOSTER SYNCHROTRON
600 MeV, 20 mA
500 MHz r.f.
10 Hz rep. rate
B_{max} = 0.786 T
$r_{bending}$ = 2.55 m
$r_{average}$ = 5.06 m
f_{rev} = 9.43 MHz
τ_{rev} = 106 ns
No. of electron bunches = 53

STORAGE RING
2 GeV, 0.37 A initially;
later 1.0 A (2×10^{11} electrons)
500 MHz r.f.
B = 1.2 T in 16 dipole bending magnets
$r_{bending}$ = 5.56 m
$r_{average}$ = 15.28 m
$\lambda_{characteristic}$ = 3.88 Å
f_{rev} = 3.12 MHz
τ_{rev} = 320 ns
No. of electron bunches = 160 (max.)
Bunch length = 0.12 ns
Bunch spacing = 2.0 ns
Lifetime $\geq$ 8h when pressure $\leq 10^{-9}$ torr
S.R. power = 255 kW at 1 A

BEAM LINES
Total S.R power ~ 40 W/mrad
~ 7 kW/m at orbit chamber wall
Beam height (above floor) = 1.25 m
Source size ~ 7.0 × 0.7 mm²
(See Table 1)

WIGGLER LINE W9
B = 4.5 T
S.R. power = 19 kW at 1 A
$\lambda_{characteristic}$ = 1.03 Å
Aperture = 64 mrad (horizontally)

EXPERIMENTAL STATIONS

FIGURE 8.1 Plan of the Daresbury Synchrotron Radiation Source and summary of operational parameters (from Lea and Munro, 1980)

at the small angle scattering (SAS) station which, on different occasions, was beam ports 7.2 and 7.3. The optical system whereby highly monochromatic and well focussed x-radiation is produced is shown in Figure 8.2.

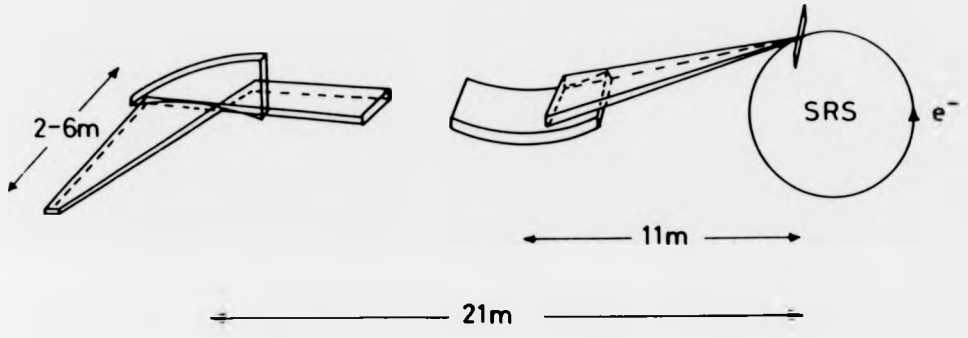


FIGURE 8.2 Schematic diagram showing the vertically focussing mirror (11 m from the source) and the Ge or Si crystal monochromator (21 m from the source) which performs 10:1 focussing in the horizontal direction and yields highly monochromatic ($\Delta\lambda/\lambda = 10^{-3}$) radiation. (From Greenhough and Helliwell, 1983).

8.3 EXPERIMENTAL

8.3.1 The Material

Calf thymus DNA was obtained from the Sigma Chemical Company. The poly d(A-T).poly d(A-T) and poly d(G-C).poly d(G-C) were purchased from Boehringer. Gels were prepared by centrifugation from solutions having known ionic strength. All fibres used here were prepared by Dr. A. Mahendrasingam using methods that are well documented (Fuller et al., 1967; Chapter Two of this work) and in conditions cited where appropriate.

8.3.2 The Diffraction Camera

The x-ray camera was designed by Dr. W.J. Pigram and built in the Physics Department at the University of Keele. A photograph is shown in Plate 8.1. It is constructed with pinhole collimation and has a melinex near window, on which the backstop is mounted. This system enables the use of either film (single cartridge/rotating carousel) or the area detector (see section 8.3.3), or both simultaneously. The camera and associated equipment is available to other users on the SAS beamline. The design at the moment will have to be modified to accommodate a smaller specimen to detector distance when used with the MWPC. The camera itself was filled with helium to avoid air scatter and the relative humidity of the fibre environment was adjusted by passing this helium through appropriate saturated salt solutions (see 2.3.3). A small water bath in the camera with an immersion heater was used to attain high humidities as required. The temperature and relative humdiity in the region of the fibre were measured from a Vaisala HMP 31UT probe positioned in the service collar of the camera. Such measurements were not taken as accurate on an absolute scale since the effects of the helium inside the camera on the operation of this device was uncertain.

8.3.3 The Two Dimensional X-ray Detector

The multi-wire proportional chamber was built in the Rutherford Appleton Laboratory by Bateman et al. (1982) and has a spatial resolution of approximately 0.3 mm x 1.0 mm. Although worse than film it is adequately matched to the beam size and diffraction broadening due to fibrous samples. Digital data from the detector is stored in a fast histogramming memory which at present may be sliced into 32 sections each containing 256 x 512 pixels. The section to which data is mapped may be changed as the experiment progresses thus allowing 32 "time-slices" to be collected. A recognisable

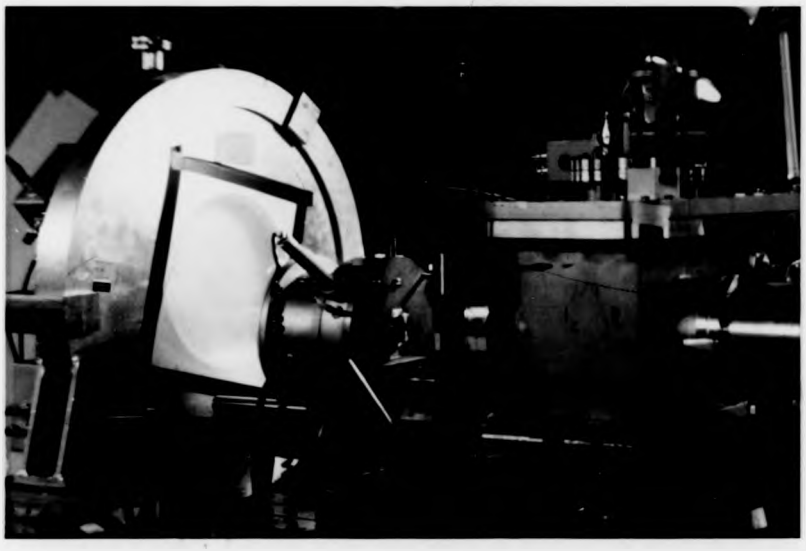


PLATE 8.1 Photograph showing the MWPC and diffraction camera in position on beam port 7.3 at the Daresbury SRS.

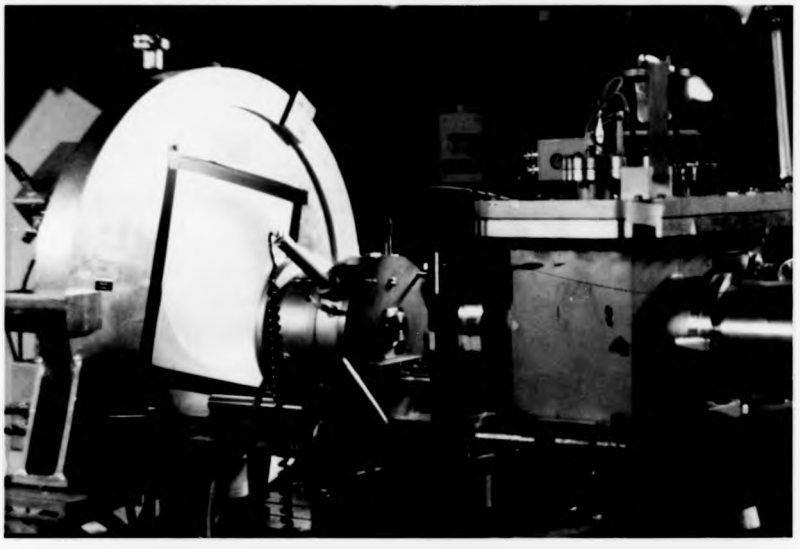


PLATE 8.1 Photograph showing the MWPC and diffraction camera in position on beam port 7.3 at the Daresbury SRS.

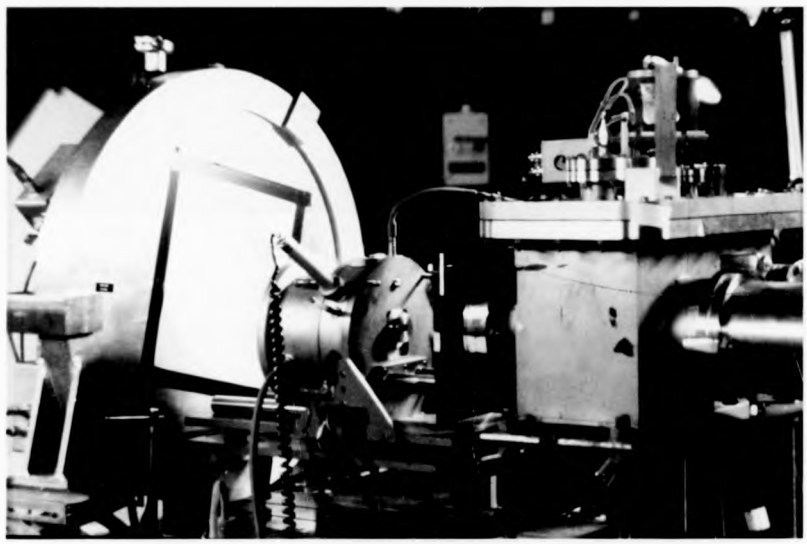


PLATE 8.1 Photograph showing the MWPC and diffraction camera in position on beam port 7.3 at the Daresbury SRS.

image could be collected in about 10 seconds but in general each frame was 60 seconds long. The data may be displayed as it is collected on a continuously refreshing t.v. screen and thus recorded on videotape. Plate 8.1 shows the MWPC positioned behind the camera on beam port 7.2. The video monitor for the device is shown in Plate 8.2, along with the LSI 11/22 microcomputer.

It is important to emphasise that the diffraction data described here was for the most part recorded before the MWPC was cleaned at the Rutherford Laboratory. The spatial resolution of the chambers was hence significantly lower than specified. The correction for the response of the MWPC has not been incorporated into the results shown in this work, nor has any image processing been applied to the data. It was accepted at the outset that the results obtained would at first only supply an idea of the MWPC's potential and of some of the problems related to its efficient use rather than any definitive data. Recorded data is displayed so that the scale of the 'x-axis' is twice that of the 'y-axis'. This results in some geometrical distortion of the image as observed on the monitor, but is not a problem inherent in the stored data itself. One of the problems that was encountered very quickly during these experiments was that of dealing with old data while new data was accumulating. When full, the 16 M byte Winchester disc had to be offloaded down a transfer line to the VAX 760 and this incapacitated further data storage for a period of three hours. As a result of this, although some of the data displayed was recorded on videotape it was not collected in digital form on magnetic tape. The camera design as of present means that the specimen-to-detector distance is necessarily larger than desired and photographs do not show complete diffraction patterns. However, the detector was on occasions rearranged so that data from one complete quadrant of a fibre diagram was collected.

A full account of the detector, associated electronics and software

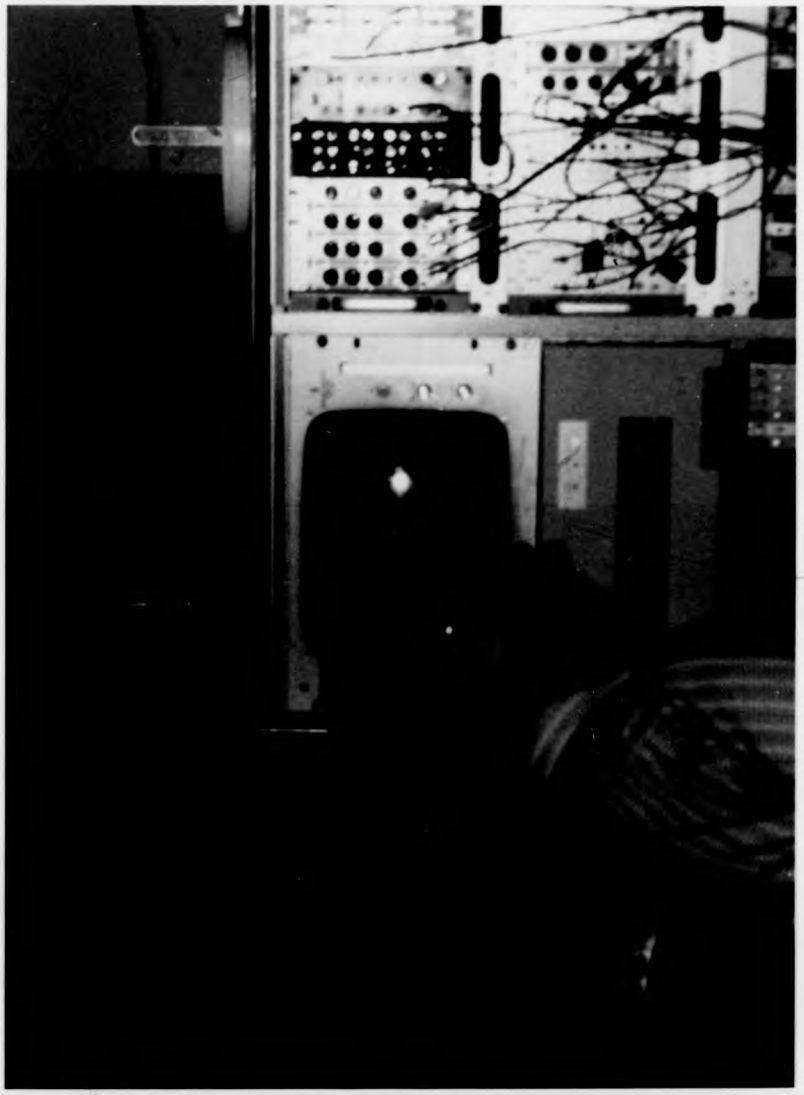


PLATE 8.2 The video monitor for the MWPC about which is seen various items of hardware connected to an LSI 11/23 microcomputer which drives the device.

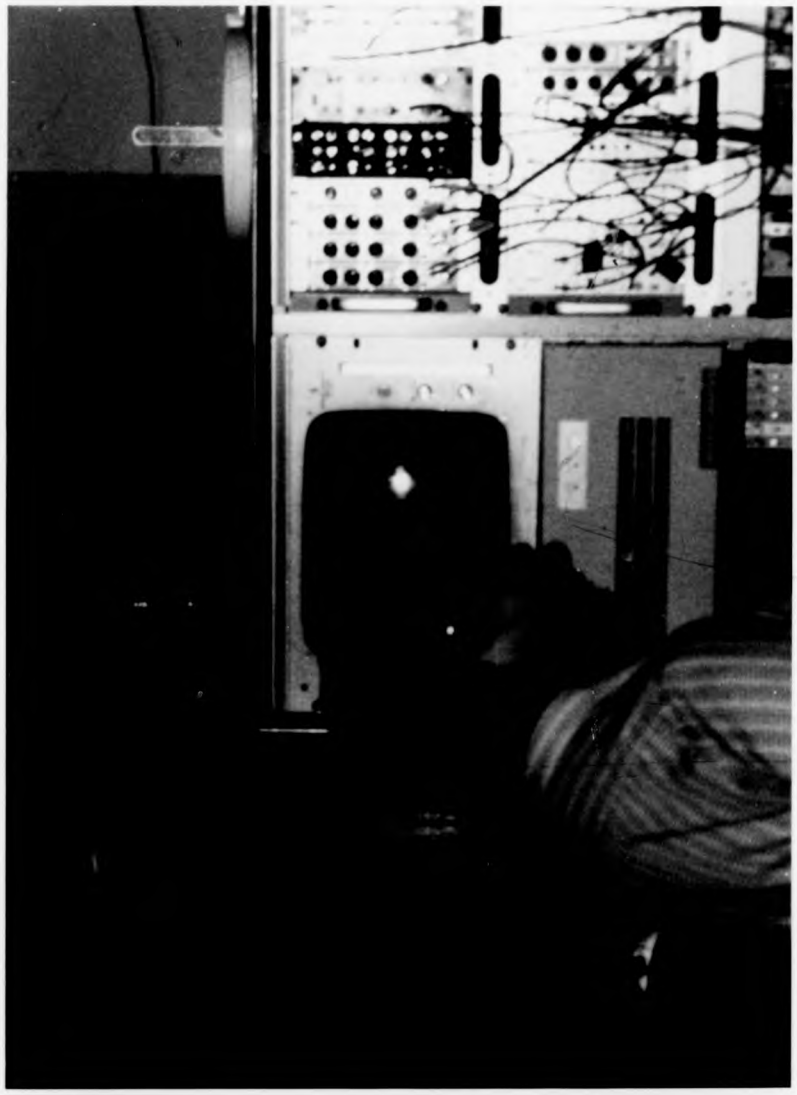


PLATE 8.2 The video monitor for the MWPC about which is seen various items of hardware connected to an LSI 11/23 microcomputer which drives the device.

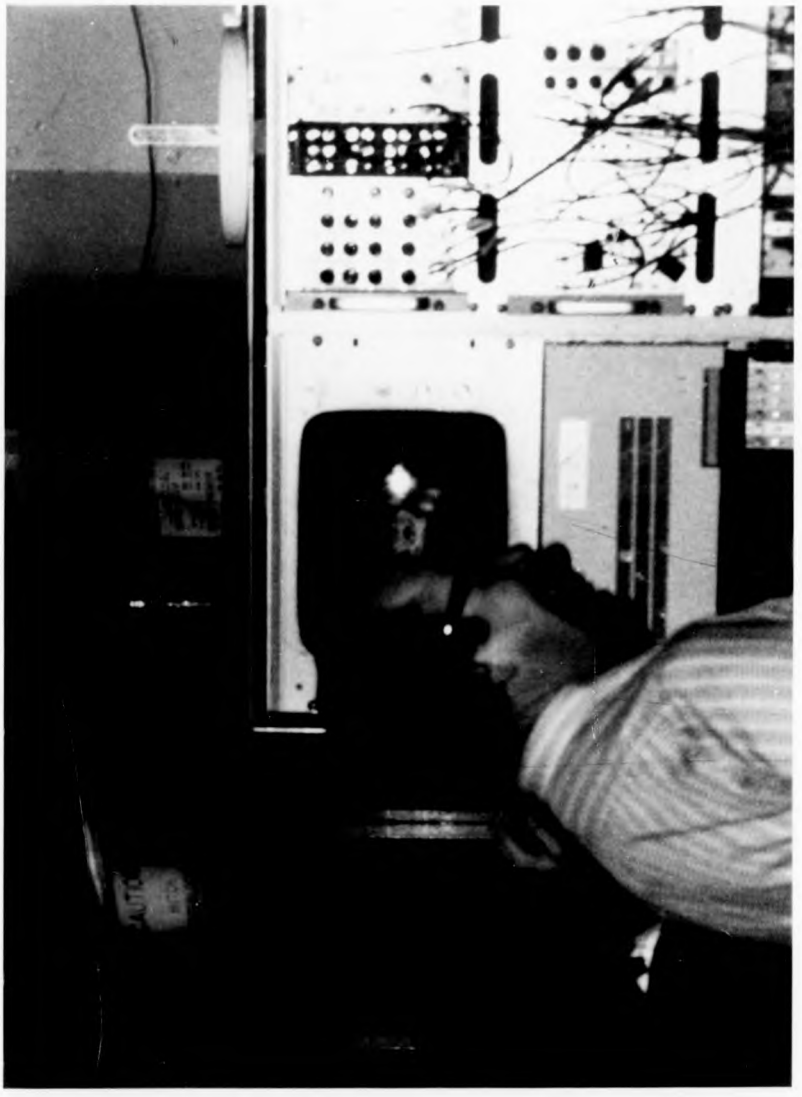


PLATE 8.2 The video monitor for the MWPC about which is seen various items of hardware connected to an LSI 11/23 microcomputer which drives the device.

will be given elsewhere (Greenall et al., in preparation).

8.4 TRANSITIONS IN POLY d(A-T).POLY d(A-T)

8.4.1 The Sodium Salt

The conditions in which the sodium salt of poly d(A-T).poly d(A-T) undergoes the reversible sequence C⇌A⇌B are now reasonably well established (Mahendrasingam et al., 1983).

Two separate experiments were performed to assess the dynamics of these transitions in fibres. In the first instance the whole sequence was observed both on film and on the area detector. This data has since been used to develop a preliminary model of the C⇌A part of the sequence. In the second experiment, the A⇌B transitions were examined on the MWPC. In each case the experiments proceeded by gradual changes in the relative humidity of the fibre environment (see 2.3.3 for details). In the case of diffraction recorded on film, exposures were completed everyten minutes. The area detector was used as described in 8.3.3.

8.4.1.1 The Complete C⇌A⇌B sequence

The whole sequence as observed on detector and film is shown in Plates 8.3 and 8.4 respectively. It is evident that whilst the C⇌A transition occurs over a fairly wide range of RH and shows evidence of structures that could be regarded as 'intermediates', the same cannot be said of the A⇌B transition. The A⇌B part of the sequence occurs over a very small range of RH and at no stage shows evidence of anything other than 'mixtures' of the initial and final conformation. (See Plate 8.3(h). Some preliminary measurements on the data relating to the C⇌A transitions have been made and are outlined in the following section.

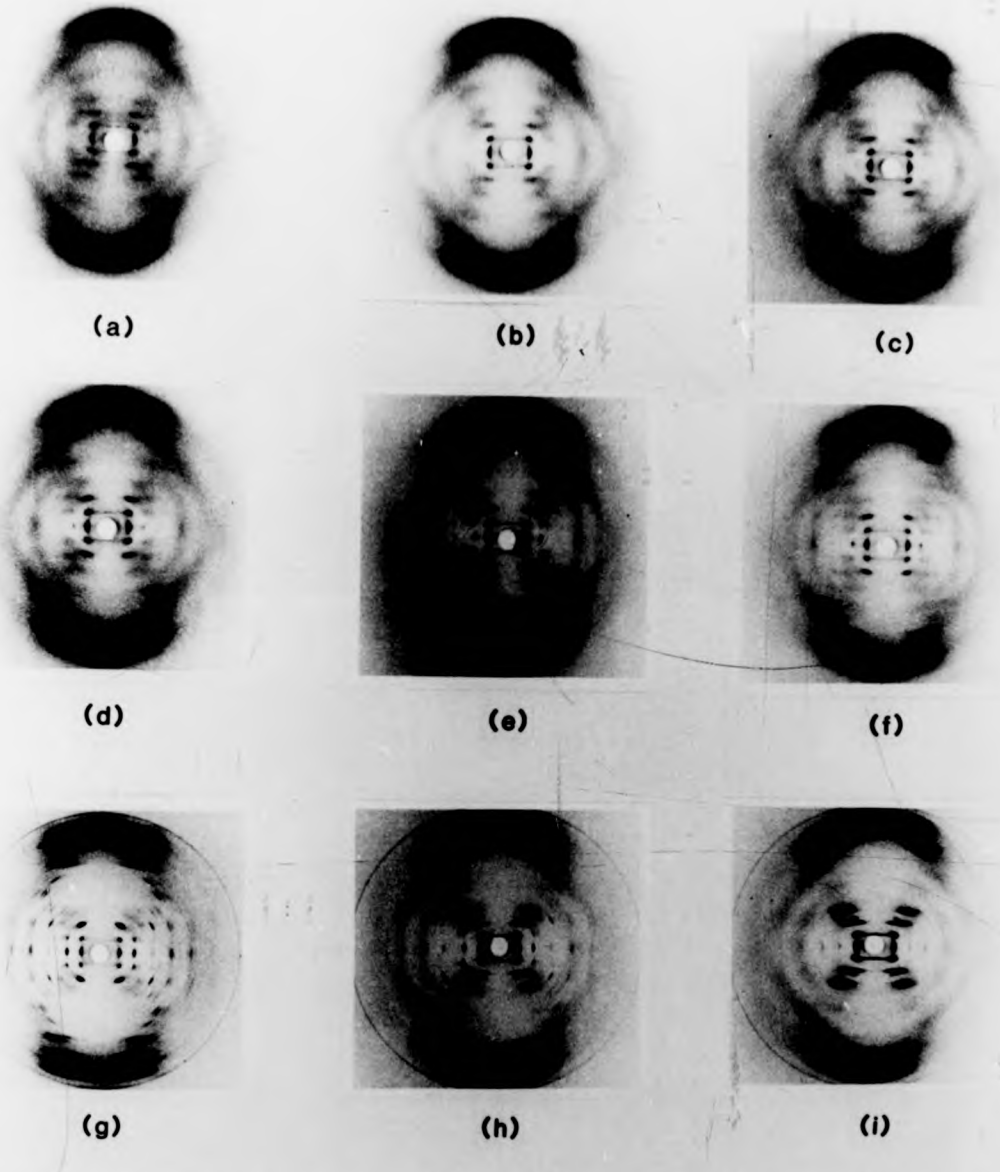


PLATE 8.3 A selection of patterns observed from a fibre of poly d(A-T).poly d(A-T). The C to A transition is evident, and thereafter the A to B transition can be seen. The diffraction photographs were recorded at the Daresbury SRS.

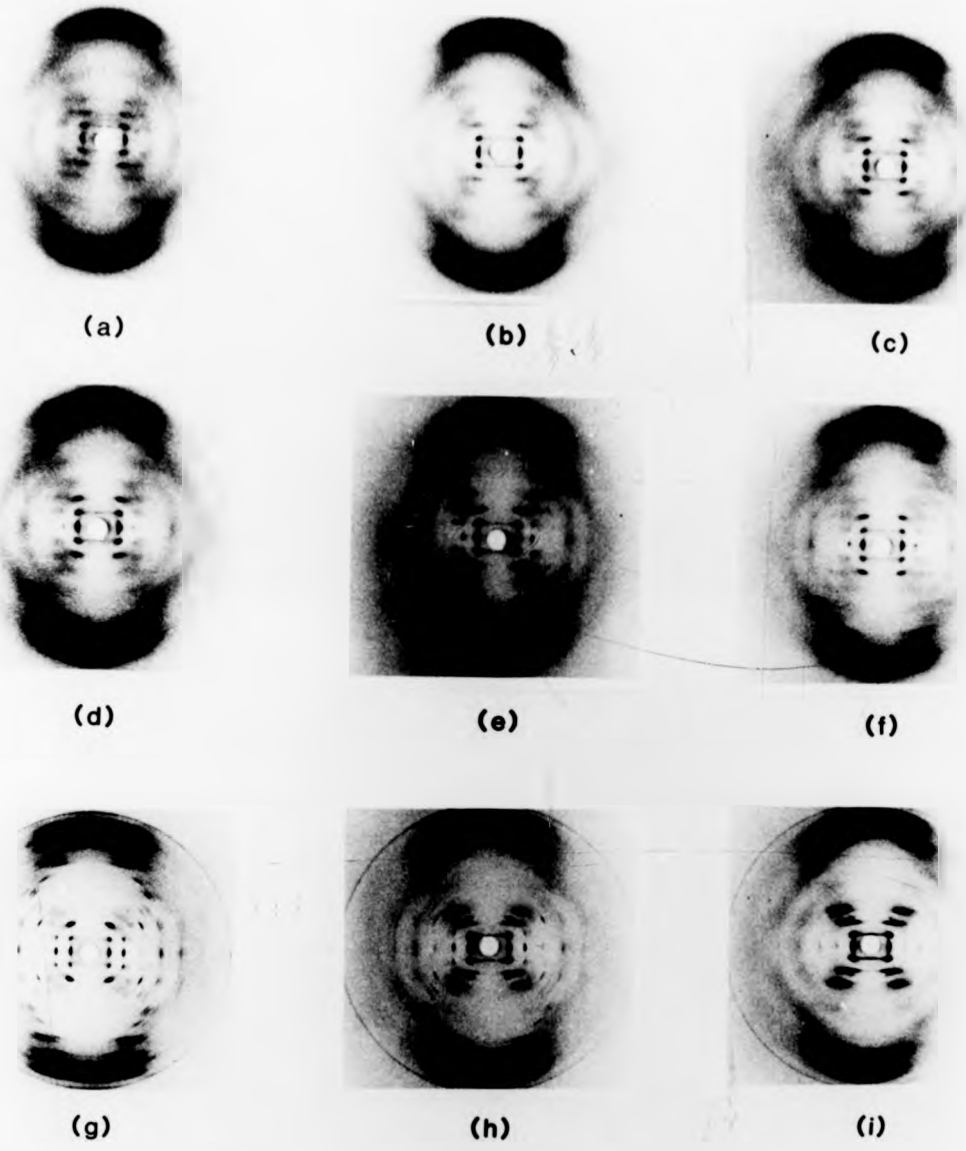


PLATE 8.3 ***A selection of patterns observed from a fibre of poly d(A-T).poly d(A-T). The C to A transition is evident, and thereafter the A to B transition can be*** *seen. The diffraction photographs were recorded at the Daresbury SRS.*

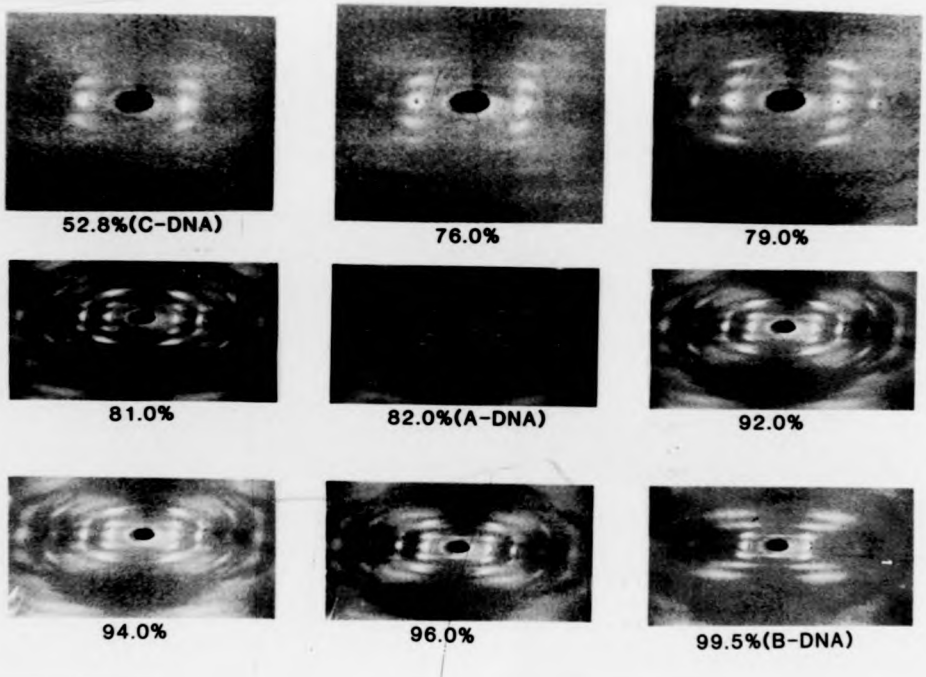


PLATE 8.4 Photographs from the video display of the area detector showing the C → A → B transitions in the sodium salt of poly d(A-T).poly d(A-T) (See section 8.4.1)

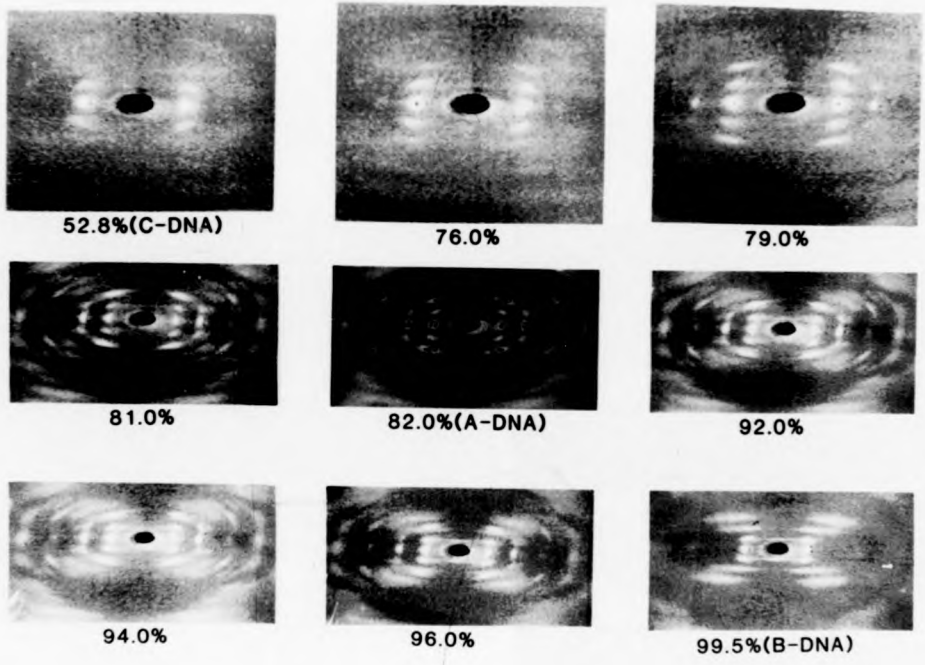


PLATE 8.4 Photographs from the video display of the area detector showing the C → A → B transitions in the sodium salt of poly d(A-T).poly d(A-T) (See section 8.4.1)

8.4.1.2 The C⇌A Transition

The data shown in Plates 8.3, 8.4 indicate that the transition from C → A involves a relatively smooth variation of pitch from ∿ 29/30Å to ∿ 28Å. The change of pitch is plotted as a function of time in Figure 8.3. It is appreciated that this graphical profile is obviously one that should be expressed as a function of relative humidity and as such the time variation should not be taken too literally. Layer line traces (made using a Joyce-Loebl microdensitometer) were recorded for the equator and the first two layer lines of each of four patterns and are shown in Figure 8.4

Whilst the change in pitch was an expected feature of this transition, the results seen in Figure 8.3 do show that there is significant decrease in pitch well before the A form becomes established. This accords well with the general belief that the C-conformation is capable of some variation (Leslie et al., 1980). What is not clear is whether or not this flexibility plays any part in transitions from the C to other forms.

Table 8.1 shows the observed ρ values from the initial C pattern recorded. Calculated ρ values are also shown for an orthorhombic lattice having a = 20.0Å; b = 14.8Å; c = 28.7Å. The pattern is indexed on the basis of the first, second and third layer line reflections. The equatorial reflection does not index onto the orthorhombic lattice which is quite different in dimensions from the one reported by Marvin et al. (1961) for lithium C-DNA. It is thought possible that this reflection is actually due to diffraction from B-DNA; it remains throughout the sequence even when the A-form materialises into its fully crystalline monoclinic lattice (whose 100 reflection should be systematically absent) and can be clearly seen in Plate 8.3(g). Plate 8.3(g) also shows evidence of meridional diffraction from B-DNA. In other respects diffraction on the equator shows a change from that characterised by C-DNA to the more 'hollow cylinder' type of diffraction exhibited by A-DNA. This change appears to be closely

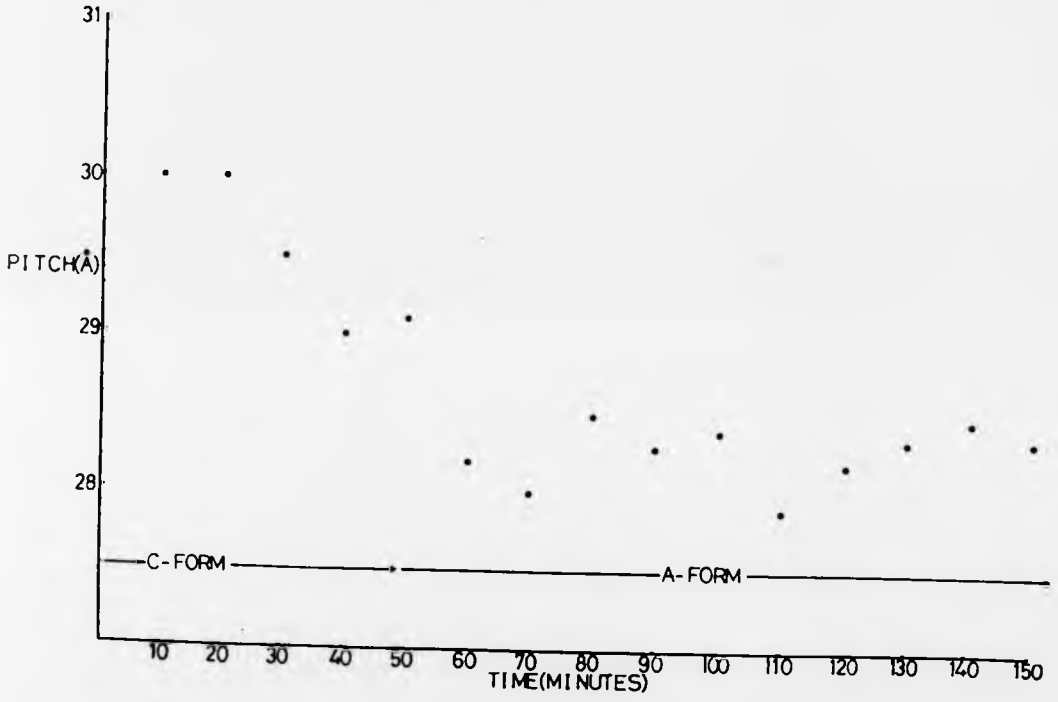


FIGURE 8.3 Change of helix pitch plotted as a function of time during the C → A transition.

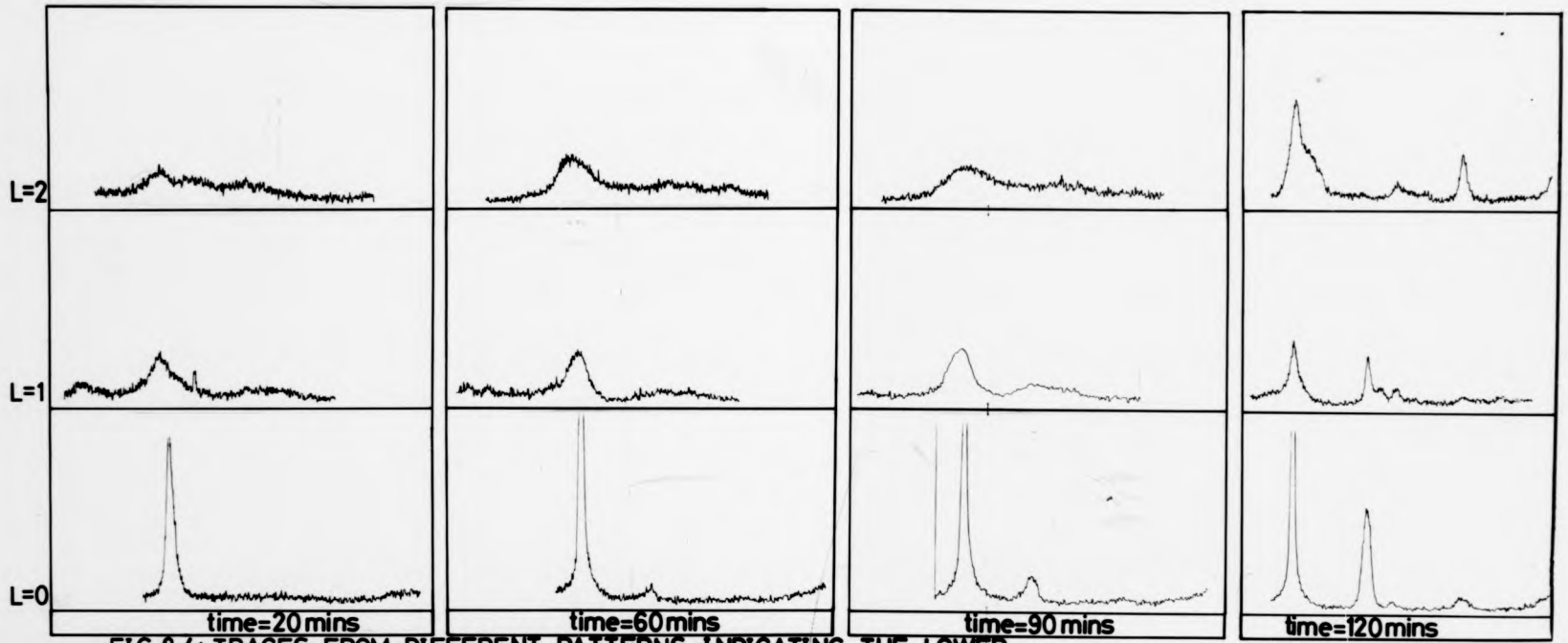


FIG 8.4: TRACES FROM DIFFERENT PATTERNS INDICATING THE LOWER LAYER LINE CHANGES DURING THE C→A TRANSITION

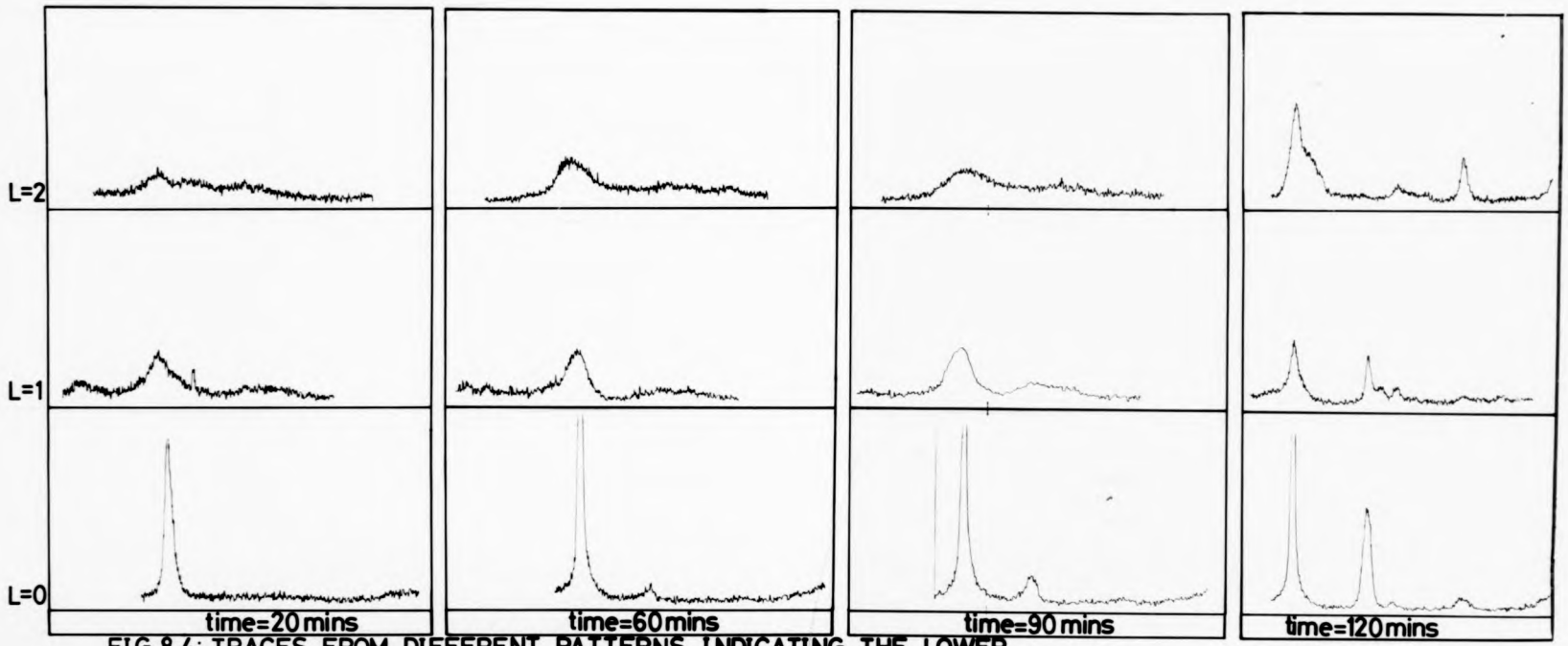


FIG 8.4: TRACES FROM DIFFERENT PATTERNS INDICATING THE LOWER LAYER LINE CHANGES DURING THE C→A TRANSITION

DESCRIPTION	Possible Indices For ORTHO. C (hkℓ)	ρ_{calc} for ORTHO. C ($\AA^{-1}$)	ρ_{obs} ($\AA^{-1}$)	Index Well? (Y/N)	COMMENT
Equator			0.058	N	from B?
ℓ = 1 1st reflection	(101)	0.061	0.064	Y	
ℓ = 1 2nd reflection			0.071	N	from A?
ℓ = 2 1st reflection	(102)	0.086	0.086	Y	
ℓ = 2 2nd reflection	(012)	0.097	0.098	Y	
ℓ = 2 3rd reflection	(112)	0.109	0.123	Y	
ℓ = 3 1st reflection	(013)	0.125	0.121	Y	

TABLE 8.1 Measurements made from a C pattern similar to the one seen in Plate 8.3(a)

accompanied by the shift in base tilt from $\sim -6^{o}$ to $\sim 20^{o}$, as is evident from an examination of the higher layer line structure of the patterns seen in Plates 8.3(a) → (g). On the first layer line, Figure 8.4 and plate 8.3 show the change in the region of $\xi = 0.05A^{-1}$ which initially contains the (101) reflection of the C form, but which eventually comes to contain the (111), (021), and (111) reflections of A-DNA. There is also a parallel swelling in the region for which $\xi \simeq 0.1\overset{o}{A}^{-1}$ where the (131), (201), (131), (201), (041) and (221) reflections of the A form occur. The second layer line shows an increase in the magnitude of both the first and second maxima. The observation of these effects is enhanced by greatly increased crystallinity

The results obtained during this experiment show that the humidity of the fibre environment was inadequately controlled and that the transition was probalby induced too quickly. Although the change from C to A can be established quickly, it is doubtful that useful information relating to intermediate structures can be obtained in this way.

Despite the fact that the humidity control during this experiment was a little cumbersome, the results have enabled a preliminary model for the C⟶A transition to be developed. It is thought that the first part of the change involves the shifting and tilting of the bases. Thereafter, observations from the first, second and third layer lines indicate a change in the value of $(\theta-\phi)$ from $\sim 60^{o}$ in C-DNA to $\sim 117^{o}$ in A-DNA. Developments in humidity control and measurement are now in progress and it is hoped that high quality data will soon be recorded which will enable this transition to be accurately detailed.

8.4.1.3 The A⇌B transition

An investigation of the A⇌B transition was undertaken at a later stage. This transition occurs over a very narrow range of relative

humidity ($\sim$ 99.0% $\rightarrow$ $\sim$ 99.5% as measured in helium with the Vaisala probe and details of any molecular change are thus difficult to establish. As can be seen from Plates 8.5(a) - (e), the first sign of change from A to B is the appearance of the (100) reflection from B-DNA. This reflection increases in intensity as the transition proceeds and is accompanied by loss of the crystallinity normally associated with the A-DNA lattice (see Plates 8.5(e) - (g)). Plate 8.5(h) shows an A/B mixture of the sort previously seen in plate 8.3(h). Plates 8.5(i)-(m) show the remainder of the transition: the 3.4Å meridional from B-DNA emerges and the higher layer line definition of A-DNA disappears. The reverse sequence is seen in Plates 8.5(n)-(r) and shows a very similar behaviour as the B form reverts to the A. The higher layer line structure of A-DNA reappears followed by the lower layer line definition and overall crystallinity. The last part of the B pattern to disappear is the (100) reflection.

8.4.2 The Lithium Salt

It is well known that the lithium salts of many DNA's are characterised by polymorphism in which the crystalline A form has not been observed, but in which crystalline B conformations are frequently adopted (see 1.4.2). The lithium salt of poly d(A-T).poly d(A-T) can be made to undergo the C$\rightleftharpoons$crystalline B$\rightleftharpoons$semicrystalline B sequence as a function of relative humidity. The exact conditions of ion concentration in the fibre that are required to obtain this sequence are not well described although Mahendrasingam (private communication) has estimated that such samples contain $\sim$ 0.6 Li^+ and $C\ell^-$ ion pairs per DNA PO_4^-. The fibre used in this study was prepared by centrifugation from a 5 mM lithium fluoride solution. containing 1.2 mg of poly d(A-T).poly d(A-T). The fibre was initially humidified at 33% RH and this was gradually increased to $\sim$ 100% RH. The reverse sequence was also observed. This work was done before the detector

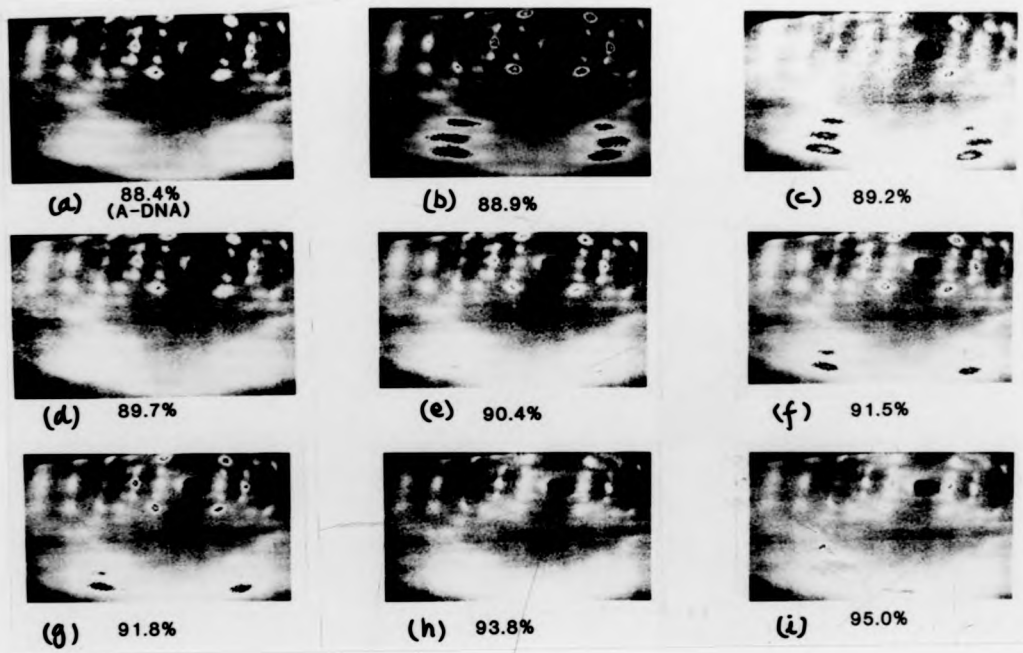


PLATE 8.5 The A⇌B transition as observed in the sodium salt of poly d(A-T). poly d(A-T).

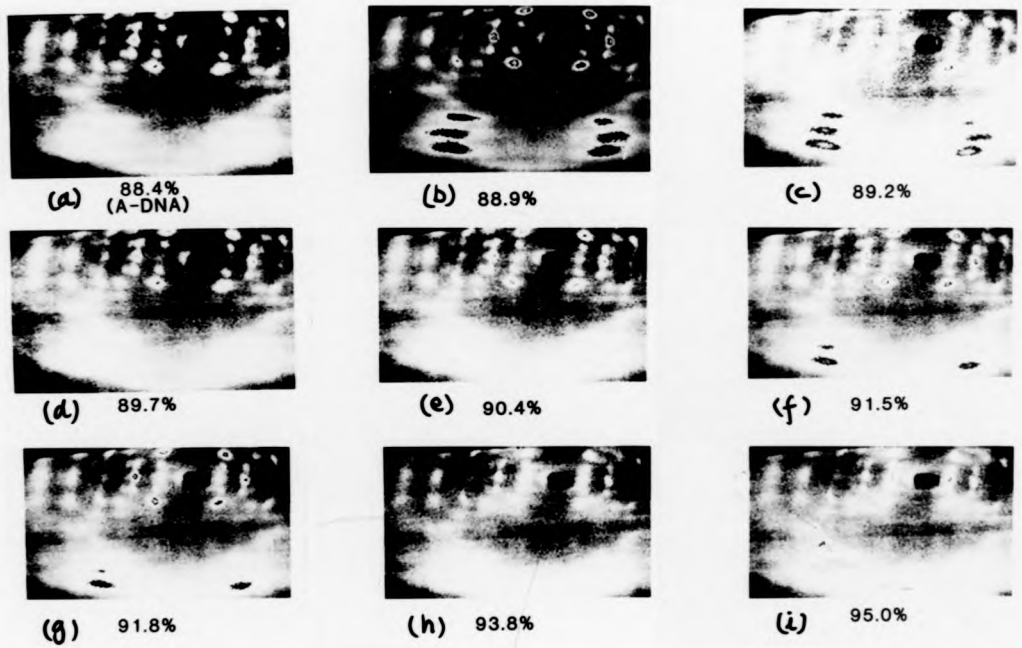


PLATE 8.5 The A⇌B transition as observed in the sodium salt of poly d(A-T). poly d(A-T).

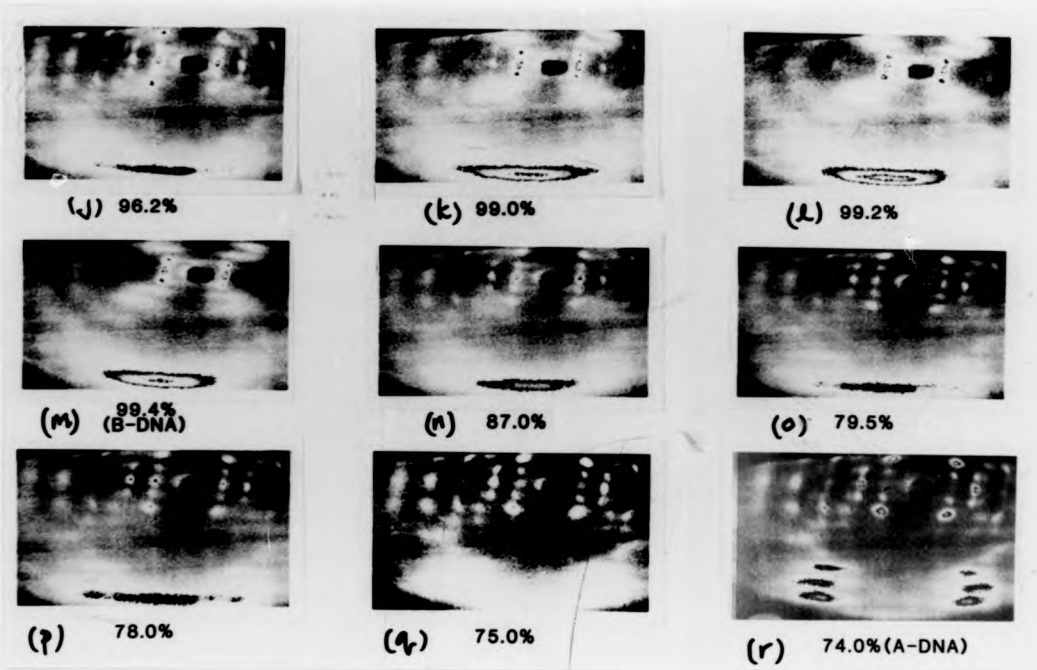


PLATE 8.5 (Cont.) The A⇌B transition as observed in the sodium salt of poly d(A-T).poly d(A-T)

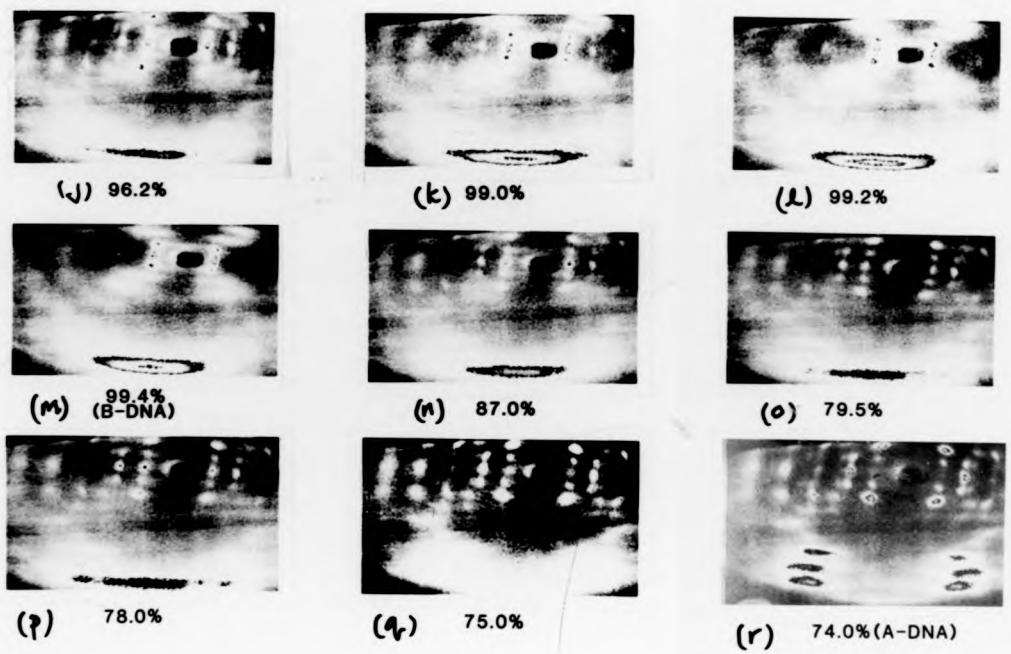


PLATE 8.5 (Cont.) The A⟷B transition as observed in the sodium salt of poly d(A-T).poly d(A-T)

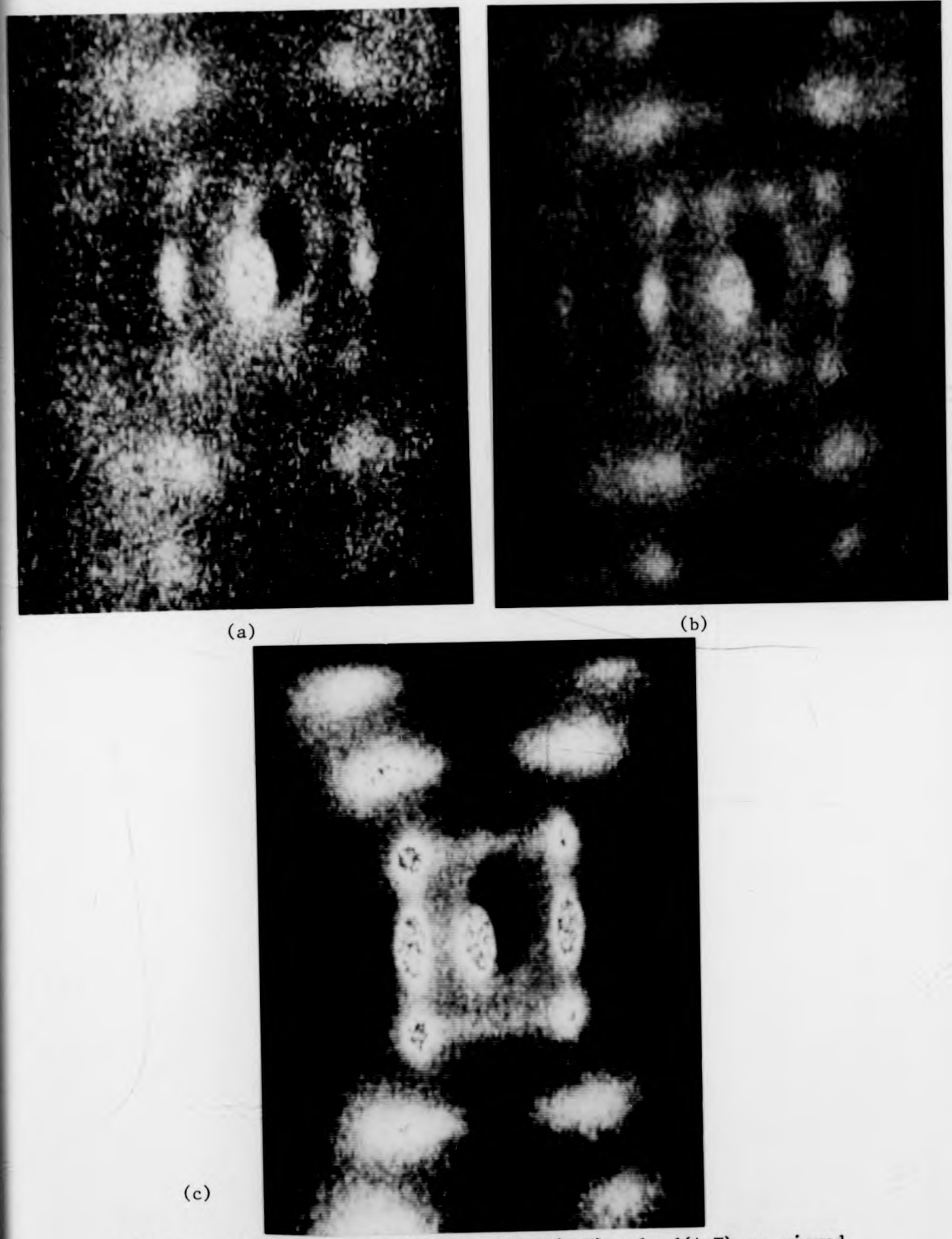
chamber had been cleaned and the diffraction data shown here illustrates the MWPC operating well below specification. At low relative humidity, a C conformation was clearly identified (see Plate 8.6(a)). As the relative humidity was increased, a crystalline B form was observed (Plate 8.6(b)) and finally at ∿ 100% RH, a semicrystalline B form was seen (Plate 8.6(c)). An attempt was made to reduce the humidity so as to recover the crystalline B in more detail, but shortage of beam time made for a clumsy return to the low humidity C.

8.4.3 The Rubidium Salt

The polymorphism of fibres made from Rb poly d(A-T).poly d(A-T) is distinct from that of either the sodium or the lithium salt of this polynucleotide. Neither the C nor the A conformations are obtained, but instead a very crystalline D form prevails up to ∿ 95% whereupon it undergoes a transition to semi-crystalline B-DNA. The sequence is not totally reversible in that the structure recovered upon reduction of relative humidity is not fully crystalline - instead a poorly defined D pattern is observed.

Owing to shortage of beam time the transition from D to B was not observed on the MWPC. However, a few timeslices of the crystalline D form at lower humidities were stored on magnetic tape and suffice to illustrate the chamber working at its best. Plate 8.7 shows a pattern of the rubidium D form as photographed from the Sigma 'Args' terminal of the VAX 760 computer at Daresbury.

Although incidental to the issue of transitions the acquisition of a high quality D pattern from a heavy atom salt of poly d(A-T) comes at a time when concern over the handedness of D-DNA is greatest (Mahendrasingam et al., 1983; Mahendrasingam, 1984; Fuller et al., 1984; Forsyth, unpublished) It is hoped that a comparison of this D data with similar quality patterns



(a) (b)

(c)

PLATE 8.6 Polymorphism of lithium poly d(A-T).poly d(A-T) as viewed on the video monitor of the MWPC. (a) C-DNA (∿ 33% RH); (b) Crystalline B-DNA (∿ 75% RH); (c) Semicrystalline B-DNA (∿ 98% RH)

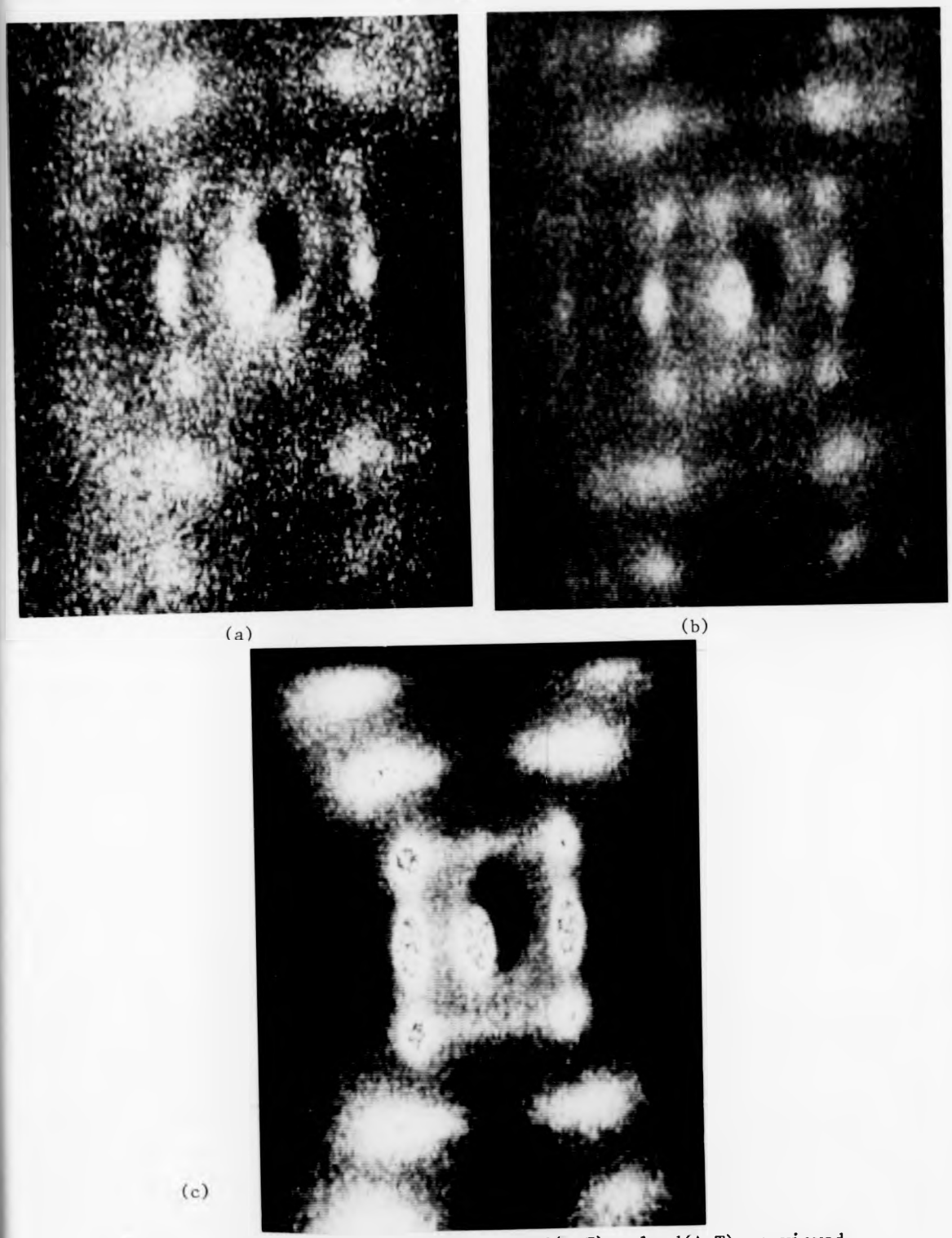
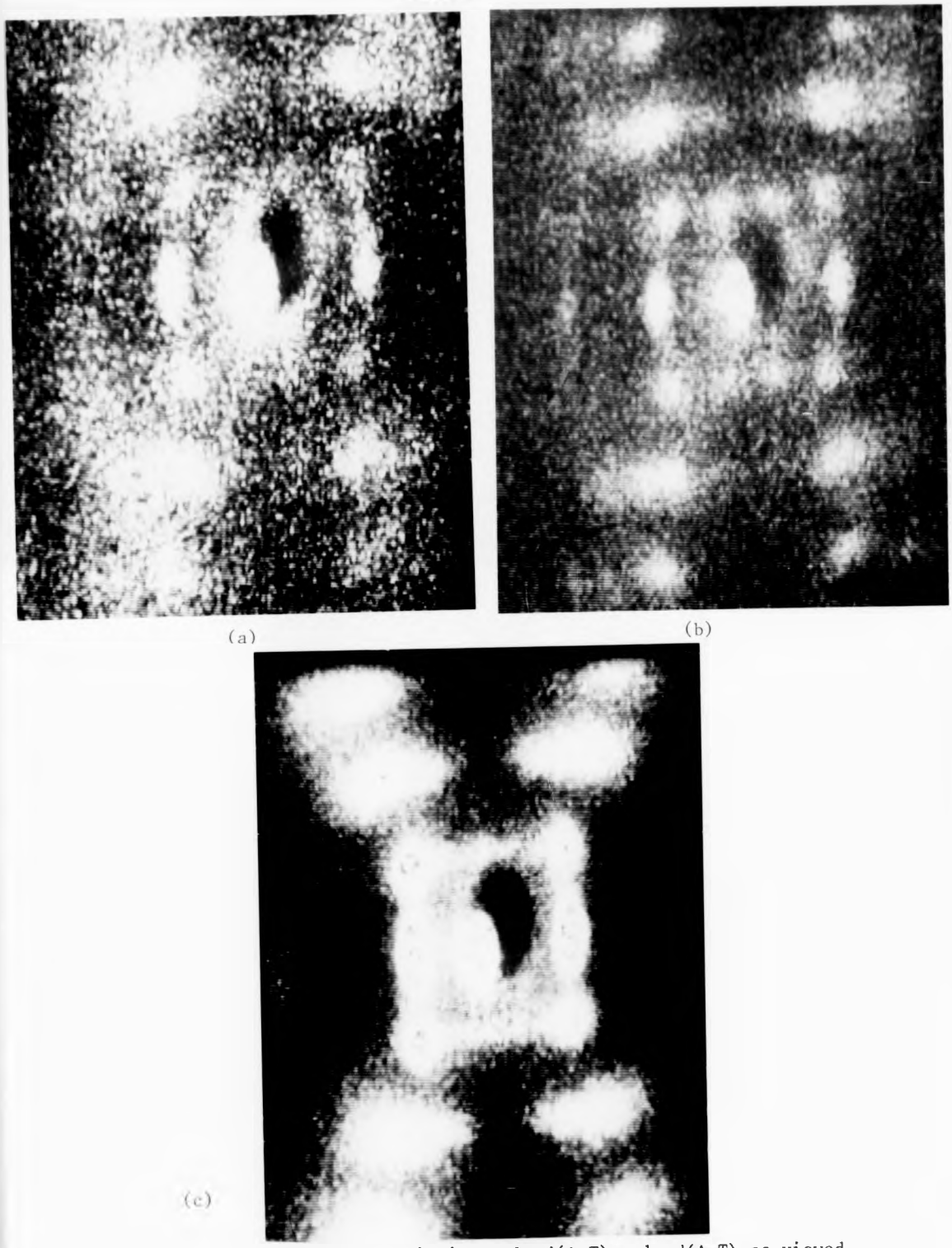


PLATE 8.6 Polymorphism of lithium poly d(A-T).poly d(A-T) as viewed on the video monitor of the MWPC. (a) C-DNA (∿ 33% RH); (b) Crystalline B-DNA (∿ 75% RH); (c) Semicrystalline B-DNA (∿ 98% RH)



(a) (b)

(c)

PLATE 8.6 Polymorphism of lithium poly d(A-T) . poly d(A-T) as viewed on the video monitor of the MWPC. (a) C-DNA (∿ 33% RH); (b) Crystalline B-DNA (∿ 75% RH); (c) Semicrystalline B-DNA (∿ 98% RH)

PLATE 8.7 The D conformation of Rb poly d(A-T).poly d(A-T) as photographed from the Sigma 'Args' graphics terminal in the data acquisition room at Daresbury. This can be compared with the pattern shown in Plate 1.7.

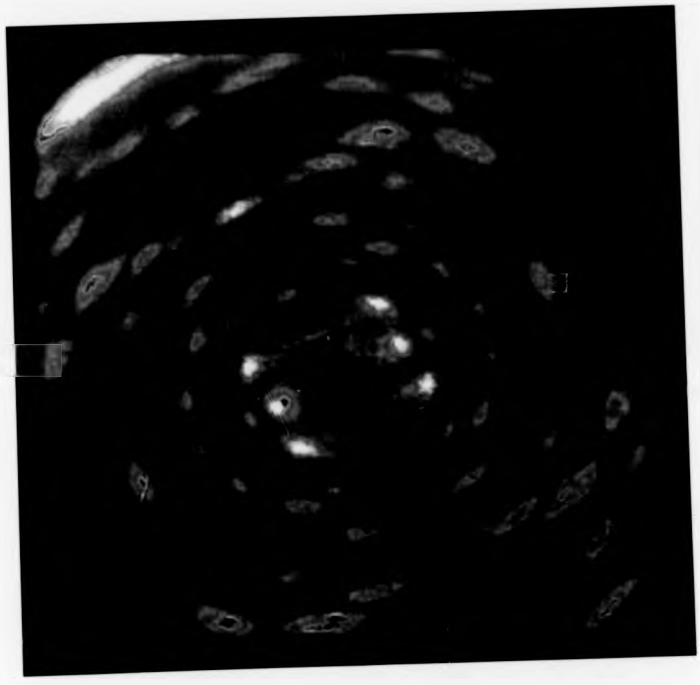


PLATE 8.7 The D conformation of Rb poly d(A-T).poly d(A-T) as photographed from the Sigma 'Args' graphics terminal in the data acquisition room at Daresbury. This can be compared with the pattern shown in Plate 1.7.

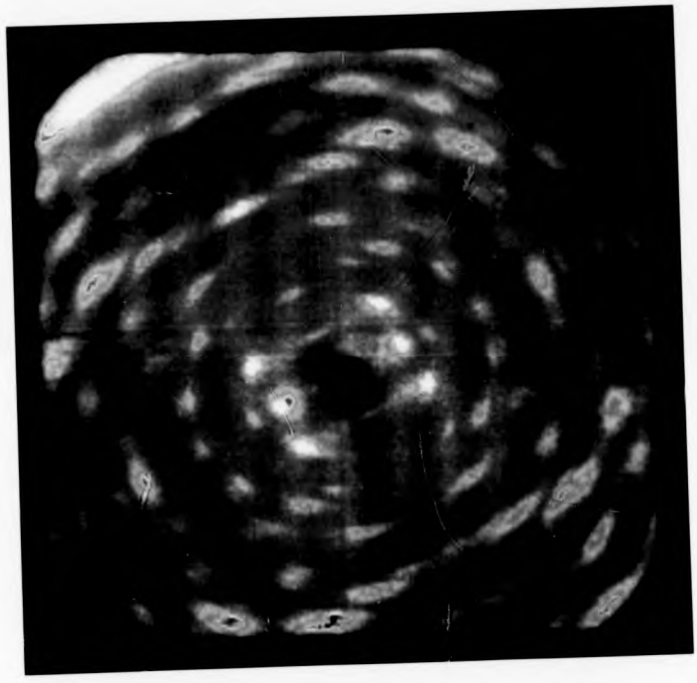


PLATE 8.7 The D conformation of Rb poly d(A-T).poly d(A-T) as photographed from the Sigma 'Args' graphics terminal in the data acquisition room at Daresbury. This can be compared with the pattern shown in Plate 1.7.

obtained from the sodium salt (Mahendrasingam et al., 1983; Mahendrasingam 1984) will enable the rubidium ions about the D helix to be located and as such to further detail the structural role that ions play in the transitions of DNA molecules. Work of this sort has already been described (Skuratovskii et al., 1979; Bartenev et al., 1983) in which attempts have been made to locate the positions of the caesium ions around B-DNA. The D data from the MWPC and the data analysis software that is currently being developed by the Daresbury Laboratory computing staff will greatly facillitate this analysis.

Whilst it is very likely that the accurate analysis proposed above will be based on data from high resolution x-ray film, this particular experiment has been used to assess the resolution capacity of the area detector, and in particular its performance in discerning variation in diffraction as a function of time.

8.5 TRANSITIONS IN POLY d(G-C).POLY d(G-C)

8.5.1 The Lithium Salt

Fibres can be drawn from the lithium salt of poly d(G-C).poly d(G-C) such that they exhibit polymorphism involving the C and B conformations (Mahendrasingam, 1983). The fibre used in this study was the same one that was used by Mahendrasingam during his work on this salt. It was originally perpared by centrifugation from a 5 mM solution of lithium fluoride and was estimated to contain $\sim$ 0.6 Li^+ and F^- ion pairs per DNA nucleotide. The X-ray fibre diagrams obtained by Mahendrasingam are shown in Plate 8.8. A semi-crystalline C form occurs up to relative humidities of $\sim$ 75%. At 85% RH, a very crystalline B form is obtained but the crystallinity is lost at about 95% RH and the pattern then obtained is almost identical to the other semi-crystalline B forms obtained from poly d(G-C).poly d(G-C) salts (see Plate 5.1(f)).

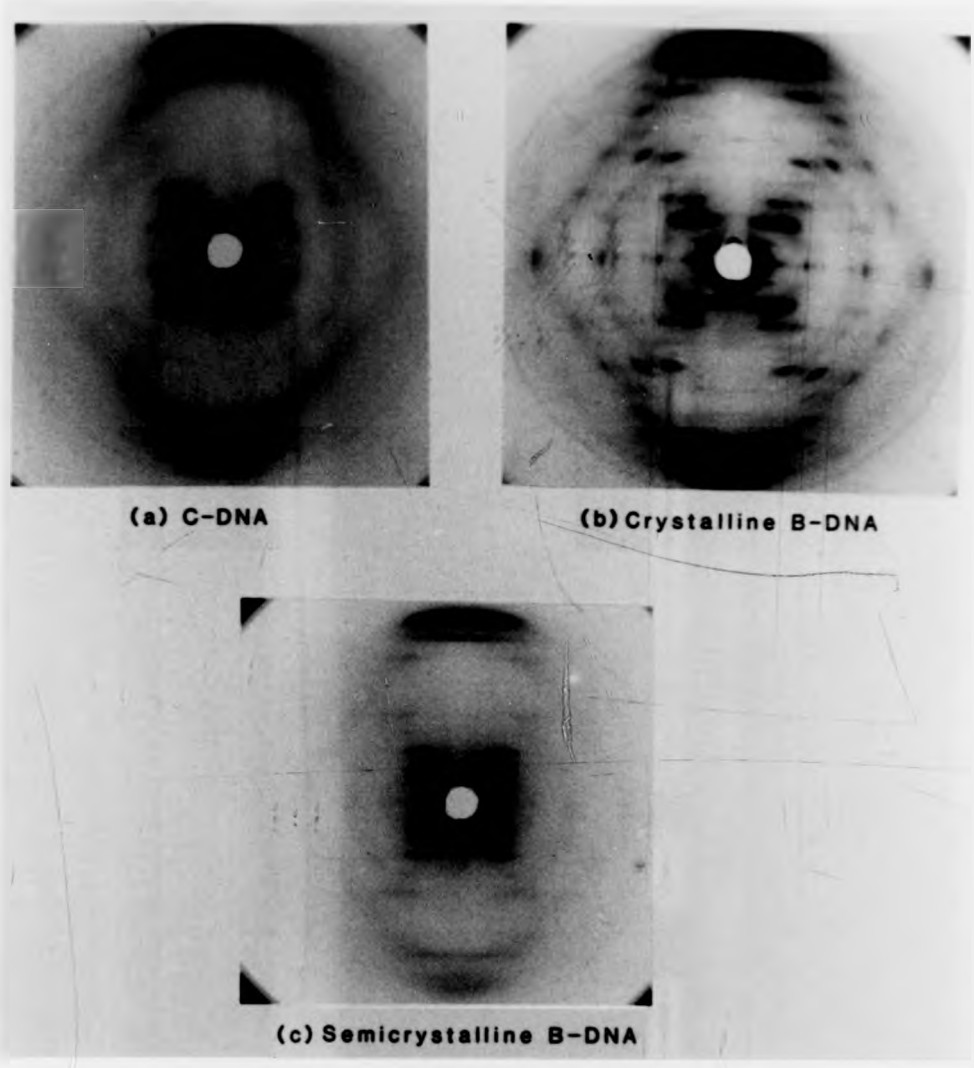


PLATE 8.8 Fibre diagrams obtained from the lithium salt of poly d(G-C). poly d(G-C): (a) C-DNA; (b) Crystalline B-DNA; (c) Semicrystalline B-DNA (from Mahendrasingam, 1984)

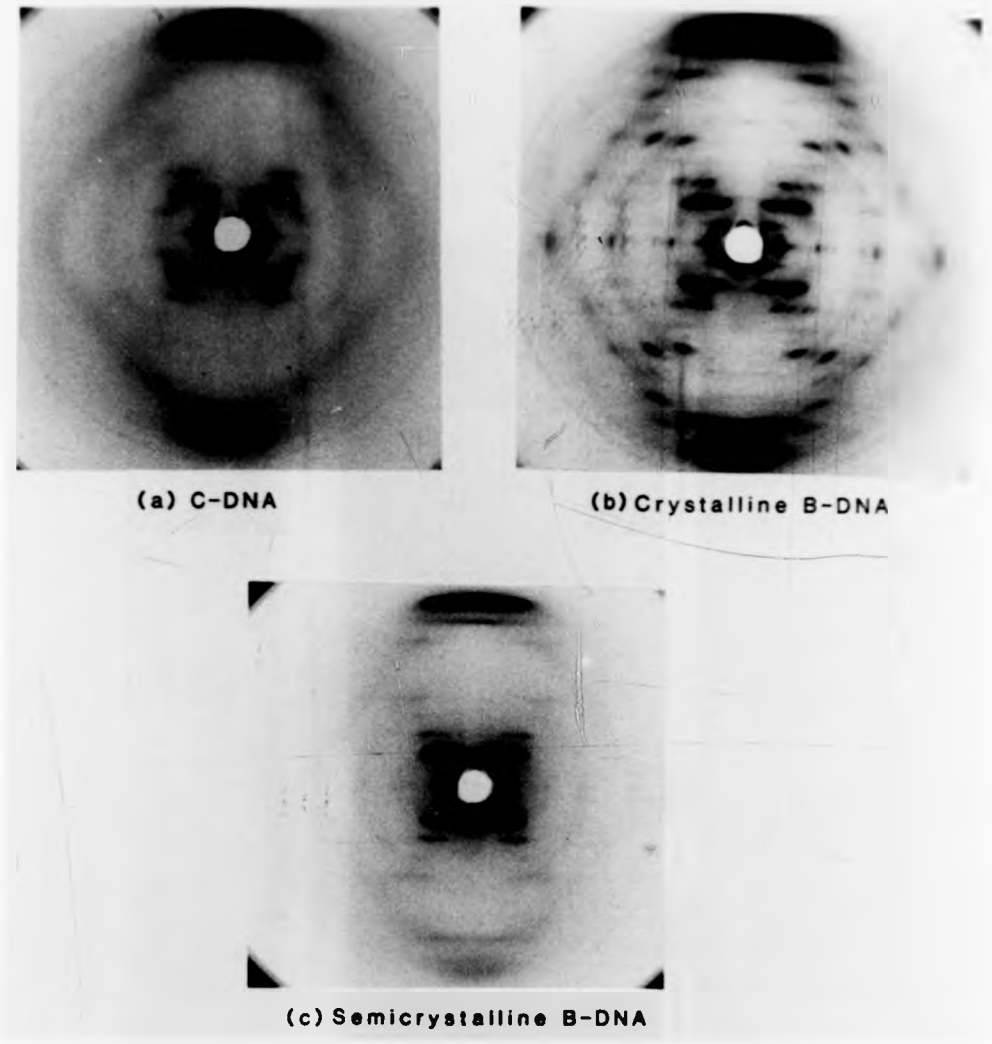


PLATE 8.8 Fibre diagrams obtained from the lithium salt of poly d(G-C). poly d(G-C): (a) C-DNA; (b) Crystalline B-DNA; (c) Semicrystalline B-DNA (from Mahendrasingam, 1984)

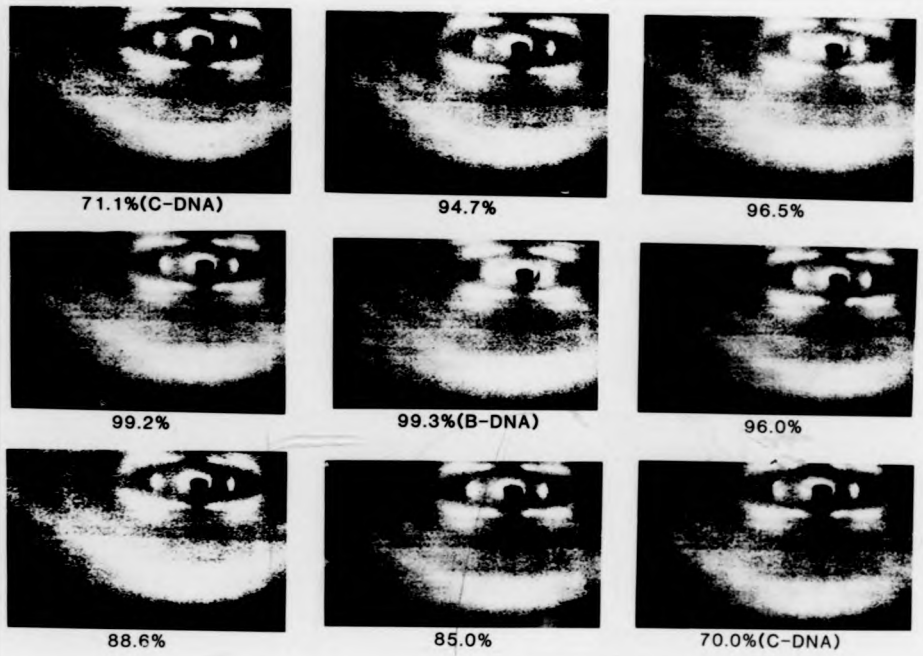


PLATE 8.9 The C⇌B transitions in the lithium salt of poly d(G-C).poly d(G-C) as recorded on the area detector.

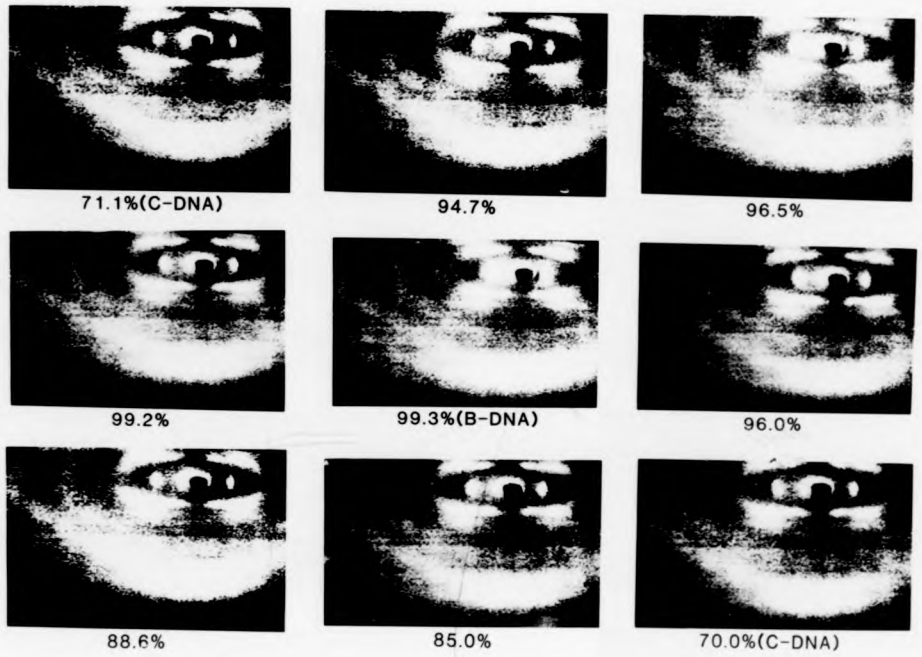


PLATE 8.9 The C⟷B transitions in the lithium salt of poly d(G-C).poly d(G-C) as recorded on the area detector.

The relative humidity of the fibre environment was gradually increased from ∿ 66% to ∿ 95%. The observed sequence was found to differ from that described by Mahendrasingam in that no crystalline B conformation was observed at intermediate relative humidities. The humidity 'run' was repeated but at no stage was there any sign of the conformation corresponding to Plate 8.8(b). A few selected time frames photographed from the MWPC videotape are shown in Plate 8.9. It thus seems that either some change has occurred within this fibre during the period separating the two experiments, or that the acquisition of this crystalline form requires long periods of annealing at intermediate humidities.

One feature that is very apparent from cursory inspections of Plates 8.8, 8.9 is the similarity between the semi-crystalline C and B forms of this polymer. The two patterns are distinguishably different on $\ell = 1$ in the region for which $\xi \simeq 0.10 \text{Å}^{-1}$ but on the whole are very similar in other regions which there is definitive diffraction. The small changes involved in the sequence from C-DNA to this 'G-C rich' (see 5.1) B form provide an ideal starting point for molecular modelbuilding studies of this transition. Such studies are now in progress.

8.6 TRANSITIONS IN CALF THYMUS DNA

8.6.1 The Lithium Salt

The transitions found to occur in the lithium salt of calf thymus DNA are well described (Marvin, 1960; Marvin, 1961). Fibres that have ∿ 0.5 Li^+ ions per nucleotide are found to adopt a C conformation at low humidities. This C conformation can pack into either an orthorhombic lattice (at lower humidities ∿ 44%) or a hexagonal lattice (∿ 66%). At higher relative humidities (∿ 98%) a semicrystalline B form is obtained.

A fibre of lithium calf thymus DNA was examined with the area

detector over a period of two hours. The diffraction obtained from this sample at various relative humidities is shown in Plate 8.10. All the C patterns seen here appear to be hexagonal and it seems that the starting humidity in this run was too high to observe the orthorhombic form. The first sign of a change to the semi-crystalline B form occurs at $\sim$ 88% RH and a well characterised pattern is observed at 99.4% RH.

In section 8.4.1 the C$\rightleftharpoons$A and the A$\rightleftharpoons$B transitions were compared. Whereas the C$\longrightarrow$A transition showed evidence of intermediate structures, the A$\longrightarrow$B only showed 'mixtures' of A and B. The experiment outlined in this section (and also the one describing Li d(G-C) in the previous section) examines the C$\rightleftharpoons$B transitions, and tends to suggest a relatively smooth transition from C$\rightarrow$ B. As the RH of the fibre environment was increased, there were changes in crystallinity and molecular transform: these were most noticeable on the first layer line but could also be seen on the second and third layer lines.

The C$\rightleftharpoons$B transitions are of particular interest to the author since there is some disagreement between the C models of Arnott et al., and those proposed by Marvin et al., (1958) and the author in Chapter Four of this work. Although the C and B forms are more closely related than any other two conformations found to exist in DNA, the analysis of Chapter Four suggests that these conformations are markedly different. The data obtained here provide an opportunity to generate a model for this transition. Such a model would presumably involve a gradual change in base position so that D' (see Figure 1.10c) moved from $\sim$ 3Å in the C form to $\sim$ 0.2Å in B-DNA, coupled with a corresponding alteration in chain dihedral angles.

8.7 CONCLUSION

These studies have clealy emphasised a number of points. The use of the SRS has undoubtedly enabled a closer examination of transitions within fibres than has hitherto been available. In addition to providing

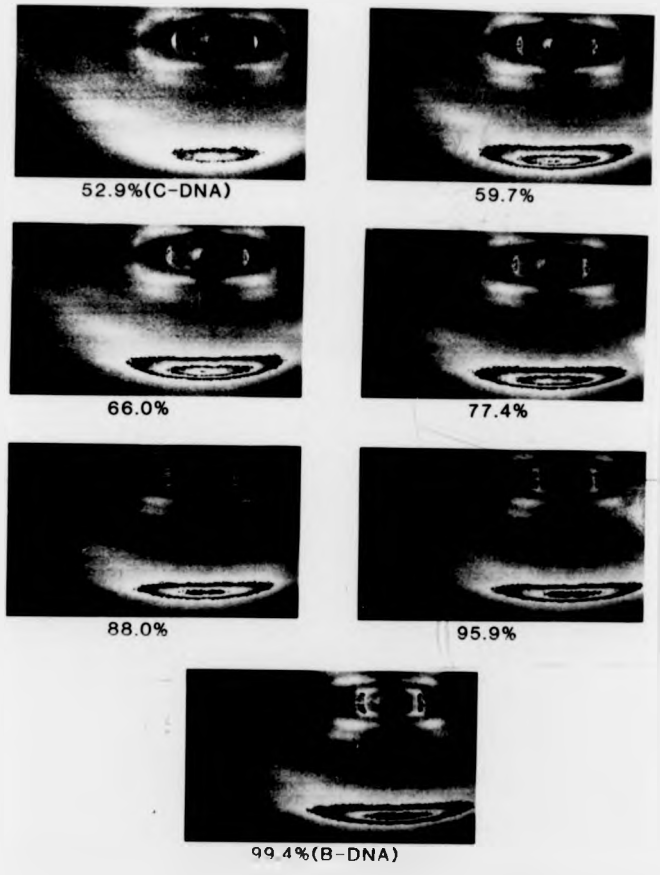


PLATE 8.10 The C → B transition in the lithium salt of calf thymus DNA

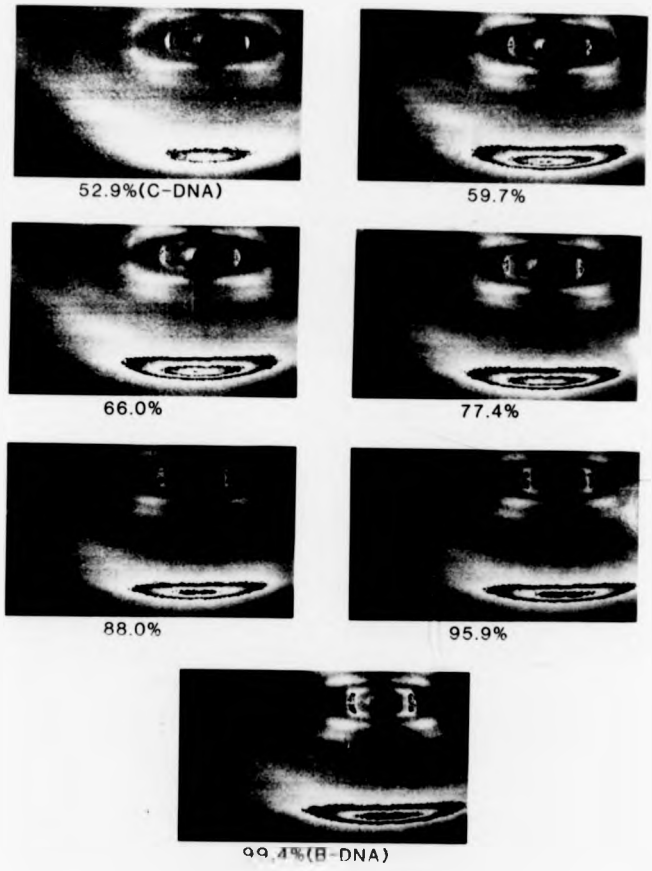


PLATE 8.10 The C → B transition in the lithium salt of calf thymus DNA

an added insight into already well known molecular structures, this sort of work furnishes invaluable detail relating to the conformational flexibility of DNA. It has also produced evidence that indicates the existence of 'intermediates'. The MWPC has here been used to assess the details of particular transitions and to determine the conditions (and their stability) which are likely to be necessary to describe the stereochemical pathways for these different transitions. In section 8.6, the C → A transition is seen to occur relatively slowly with evidence of intermediate structures. By contrast, the A → B transition seems to occur over a narrow range of relative humidity and therefore reaffirms the need for very well controlled conditions of humidity in the environment. The MWPC has thus illustrated that the dynamics of these fibre transitions necessitate the implementation of a humidity system that is both stable and 'tunable'. The speed and ease with which different transitions are induced varies enormously and in all likelihood depends on many different factors operating at different levels of organisation.

The results of future studies will depend on the effectiveness of the experimental strategy used. The relative humidity of the fibre environment must be closely controlled and accurately measured. The large amount of two dimensional time resolved data needs to be reduced and processed efficiently. The problem of data transfer from the experimental workstation to the VAX computer is being investigated. The existing diffraction camera is being modified to accommodate for a somewhat shorter specimen to detector distance.

CHAPTER NINE

CONCLUSION

The fibre studies described here and elsewhere exemplify the conformational flexibility of a large variety of native and synthetic polynucleotides (see section 1.4). The structures of the classical A and B forms of DNA are well established and have not been examined in this work, but in Chapter Four the C' form has been analysed and is exposed as a robust and distinctive molecular structure which differs significantly from B-DNA (in contrast to the assertions of Arnott et al.). It is believed that C-DNA is capable of a 'smooth' and 'time resolvable' interconversion with the A and B fibre structures (see sections 8.4.1.2 and 8.6 respectively) and there is scope for molecular modelling of these transitions.

Work on random sequence DNA (all the nucleotides of which can be expected to have a very similar structure) and on highly repetitive sequence polymers (in which consecutive nucleotides have different characteristic configurations) has illustrated the problems associated with assessing nucleic acid interaction with water and various types of ion. Although it is almost inevitable that experiments of this type tend to monitor the effects induced by one independent variable (such as relative humidity), it is very important to emphasise that no one condition should be considered in isolation. There are numerous examples that illustrate this point: the effect of hydrating a fibre depends not only on polynucleotide sequence but on the ionic strength within the sample; the effect of drugs on DNA depends on both ionic conditions and hydration

(see Chapter Seven), and there is every reason to believe that the behaviour of other groups (such as amino acids, histone proteins etc.) will, in part, be influenced in a similar fashion.

The stereochemical basis for the interaction of such groups with nucleic acids will depend on the exact nature of the exposed hydrophilic and hydrophobic regions of the molecule, along with the availability of hydrogen bond donor and acceptor groups. In so far as conformational change affects these molecular properties it is obviously very important to establish the conditions necessary for the acquisition of all known nucleic acid conformations. The question of chirality in DNA is of significance since the existence of right- and left-handed helices in vivo would appear to suggest distinguishably different functional roles for the double helix in biological environments. The D and S forms analysed and discussed in this work (Chapter Six) are of special relevance here. That the S form is left-handed is widely asserted but, as yet, not irrefutably confirmed. The D form, by way of contrast, is generally thought of as being a right-handed helix: the x-ray fit obtained in section 6.2 and that of Mahendrasingam (1983) are thus both of direct interest to sequence related studies of DNA.

Molecular modelling has, over the years, developed in strides and bounds since the days in which the first wire models were constructed to explain diffraction from the A, B and C forms of native DNA. Computerised techniques now make the exploration of many different models fairly routine. Although the best molecular models (in terms of agreement with diffraction data and overall stereochemistry) which have so far been produced are left- and right-handed double helices having Watson-Crick base pairs, there are a number of other models that merit serious attention. Some incorporate different base-pairing schemes (e.g. Hoogsteen, 1959) whilst others are not double helical but triple and even four stranded molecules. A model

due to Rodley et al. (1976) involves short stretches of left- and right-handed DNA so that the whole molecule is side-by-side (SBS) in nature. The main advantage of this model as compared to others is that it overcomes the topological problem of strand separation that is inherent in the Watson-Crick scheme for DNA replication. Whilst these models do not provide the sort of agreement between observed and calculated diffraction (e.g. Greenall et al., 1979) that is afforded by the best double helical models, they are still of considerable stereochemical interest and cannot be discounted as possibilities for DNA in certain cytological situations. Examples of rather less dramatic departures from the conventional double helical models are the 'heteronomous', 'wrinkled' and 'pleiomeric' structures proposed by Arnott and his colleagues who suggest that DNA 'surfaces' may signal what base sequences lie beneath them. The 'heteronomous' model for B'-DNA has been examined in Chapter Five. Although the 'F' form (recognised as similar to the D form - see section 1.4.4) has not been closely scrutinised in this thesis, Arnott et al. assert that this molecule is 'pleiomeric' in nature, has 4_3 screw symmetry and a hexanucleotide repeat. Whilst refreshingly exotic in concept, this right-handed 'pleiomeric' molecule does not tally with the arguments and modelling described in Chapter Six.

It is likely that the experimental approach presently being adopted towards ascertaining the molecular structure and function of DNA will eventually provide enough information to secure a comprehensive knowledge of the way in which various chemical groups bind to and influence the behaviour of DNA. Such knowledge will not only provide greater detail relating to chromatin, sequence recognition, transcription and translation but will also provide a rational basis for the structured design of drugs which may specifically interfere with the operation of nucleic acids (see Chapter Seven).

The scope for future work is extensive. In addition to refining the existing techniques associated with conventional structure analysis and time resolved fibre studies, there also exists possibility of studying DNA transitions in solutions as a function of ionic concentration. It is very difficult to alter the ionic concentration within a fibre without disrupting the three dimensional order of the system; this sort of experiment can only be performed with solutions. Whilst x-ray scattering from concentrated solutions does not offer as much definitive information as is obtained from fibres, it does mean that measurements relate to DNA in what could be regarded as a more realistic environment. The application of the TV detector on the Wiggler at Daresbury Laboratory will mean that time resolved measurements on both fibres and solutions should enable a comprehensive picture of transitions in DNA to be established.

There also exist a number of experimental techniques that can be applied to the study of DNA and which have not, to date, been fully pursued. As propounded earlier, the exact position of ions and water in relation to the double helix is of great interest to workers in this field. The possibility of analysing diffraction data from heavy atom salts of DNA is one that has been investigated by Bartenev et al. (1983). Bartenev et al. maintain that cationic stabilization of Cs B-DNA occurs by the distribution of caesium ions equally between the large and small grooves of DNA so that there is a ratio of 1:1 between Cs^+ and PO_4^- groups. Such a method can equally well be applied to the C and A forms of heavy metal salts of DNA. Recently, Mahendrasingam (unpublished) has obtained the D conformation from caesium and rubidium salts of poly d(A-T).poly d(A-T). This data affords a unique opportunity to further investigate the D helix (see Chapter Six) amidst current concern over its handedness.

At the SRS it should also be possible to utilize soft x-rays in the region of the phosphorous absorption edge (5.787Å) to calculate anomalous difference intensities and hence locate the phosphorous atoms in nucleic acid structures.

The availability of such innovative methods provides added impetus to the science of nucleic acid crystallography, and stands to produce a wealth of structural information. The interactive data analysis techniques that are currently being developed for fibre diffraction have been mentioned in passing (Chapters Two and Four) and will greatly facilitate structure determination.

APPENDIX

TABLE A.1

AN ANTHOLOGY OF MOLECULAR AND CRYSTAL PARAMETERS FOR CONFORMATIONS OF RELEVANCE TO THIS WORK

Torsion angles in this table are calculated in the manner of section 1.3 of this work.

The angle ζ defines the sugar pucker and is C3'-endo if $\zeta \sim 80^{\circ}$, C2'-endo if $\zeta \sim 147^{\circ}$, and C3'-exo if $\zeta \sim 156^{\circ}$ (see Arnott and Hukins, 1972).

χ is sometimes expressed in terms of the atoms O5', C1', N9, C8 (angles marked with an '*'), or alternatively in terms of O5', C1', N9, C4 (angles marked '**').

γ is the angle between the perpendicular to a base plane and the helix axis.

STRUCTURE	CONFORMATIONAL PARAMETERS RELATING TO VARIOUS PROPOSED MODELS (°)											RISE/RESIDUE h(A)	TURN/RESIDUE c(°)	MOLECULAR SYMMETRY	REPEATING UNIT	LATTICE AND PARAMETERS					REFERENCES
	α	β	γ	δ	ε	ζ	χ	TWIST	TWIST	'γ'	D(A)	CLASS					TYPE	a(A)	b(A)	c(A)	MOLS/CELL	CRYSTALLINITY	
FULLER A	-139.3		-81.2	-76.9	167.4	67.5	76.5	76.7	20.0	-8.0	—	4.25	$tg^-g^-tg^+$	2.56	32.7	11_1	Mononucleotide	Monoclinic (β=97°)	22.30	40.60	28.15	2	Crystalline	74
ARNOTT A1	178.1		-46.9	-84.7	-151.6	45.3	83.2	86.4	20.0	-1.2	—	4.72	$tg^-g^-tg^+$	2.56	32.7	11_1	Mononucleotide	Monoclinic (β=97°)	22.30	40.60	28.15	2	Crystalline	4
ARNOTT A2	175.0		-45.0	-80.0	-149.0	47.0	83.0	-154.0	—	—	—	—	$tg^-g^-tg^+$	2.56	32.7	11_1	Mononucleotide	Monoclinic (β=97°)	22.30	40.60	28.15	2	Crystalline	17
A-RNA	-151.1		-74.3	-62.1	179.8	48.1	83.2	74.9	16.0	-6.9	—	—	$tg^-g^-tg^+$	2.81	32.7	11_1	Mononucleotide	Hexagonal	39.7	39.7	30.9	3	Crystalline	10
A'-RNA	-162.5		-59.0	-65.1	-167.0	43.0	83.2	79.5	10.0	-7.6	—	—	$tg^-g^-tg^+$	3.02	30.0	12_1	Mononucleotide	Hexagonal	39.4	39.4	36.2	3	Crystalline	10
POLY rI. POLY dC	-160.1		-64.7	-75.5	-179.1	61.9	77.7	75.0	6.2	-2.0	—	4.85	$tg^-g^-tg^+$	2.97	30.0	12_1	Mononucleotide	Hexagonal	39.0	39.0	35.6	3	Crystalline	134
WATSON-CRICK B	-174.0		-64.3	-178.1	179.1	165.3	99.8	—	0	0	90	-3.5	tg^-ttt	3.38	36.0	10_1	Mononucleotide	Hexagonal	27.4	27.4	33.8		Semi-Crystalline	176,177
LANGRIDGE B	147.2		-78.0	-79.3	-147.7	33.7	129.8	138.2	-2.0	0.0	—	-0.69	$tg^-g^-tg^+$	3.37	36.0	10_1	Mononucleotide	Orthorhombic	22.7	31.2	33.7	2	Crystalline	100
ARNOTT B1	154.7		-93.2	-46.1	-146.5	36.4	136.5	162.5	-5.9	-2.1	—	-0.16	$tg^-g^-tg^+$	3.38	36.0	10_1	Mononucleotide	Orthorhombic	22.5	30.8	33.8	2	Crystalline	6,9

STRUCTURE		CONFORMATIONAL PARAMETERS RELATING TO VARIOUS PROPOSED MODELS (°)										RISE/RESIDUE h(A)	TURN/RESIDUE τ(°)	HELICAL SYMMETRY	REPEATING UNIT	LATTICE AND PARAMETERS					REFERENCES
		α	β	γ	δ	ε	ζ	χ	TILT	TWIST	'γ'	D(A)	CLASS					TYPE	a(A)	b(A)	c(A)	MOLS/CELL	CRYSTALLINITY	
ARNOTT B2		-165.7	-135.9	-25.3	139.9	26.8	146.2	138.6	-0.1	-3.8	—	-0.14	ttg⁻tg⁺	3.38	36.0	10_1	Mononucleotide	Orthorhombic	22.5	30.8	33.8	2	Crystalline	9
ARNOTT B3		-133.0	-157.0	-41.0	136.0	38.0	139.0	-107.0	—	—	—	—	ttg⁻tg⁺	3.38	36.0	10_1	Mononucleotide	Orthorhombic	22.5	30.8	33.8	2	Crystalline	17
ARNOTT WRINKLED B-d(GC)	CpG	-85	-169	--	145	22	148	164	—	—	—	—	g⁻tg⁻tg⁺	6.80	72.0	5_2	Dinucleotide	Orthorhombic	37.9	36.1	68.0	4	Crystalline	21
	GpC	-136	-158	-30	126	47	143	143	—	—	—	—	ttg⁻tg⁺											
LEVITT B		178.0	-85.0	-45.0	170.0	65.0	168.0	-119.0	-17.0	-14.0	—	—	tg⁻g⁻tg⁺	3.38	36.0	10_1	Mononucleotide	Orthorhombic	22.5	30.8	33.8	2	Crystalline	109
KLUG ALTERNATING B	TpA	-160.2	-82.1	-47.2	171.6	64.2	143.7	140.3	—	—	—	—	tg⁻g⁻tg⁺	6.80	72.0	5_2	Dinucleotide	Potential energy minimisation study designed to explain enhanced repressor binding						94
	ApT	-159.4	-134.0	-40.2	151.2	59.1	99.2	-87.9	—	—	—	—	tg⁻g⁻tg⁺											
DREW & DICKERSON DODECAMER B (average τ's shown)		-169	-108	-43	171	54	123	-117	—	—	—	—	tg⁻g⁻tg⁺	3.33	36.9	9.75 base pairs per helix turn		$P2_12_12_1$	24.87	40.39	66.20		Single crystal study of the dodecamer d(CpGpCpGpApApT pTpCpGpCpG)	51,56
ARNOTT α-B'		145.3	-86.9	-52.1	-136.4	39.4	156.5	-144.5	-7.9	-1.0	—	-0.01	ttg⁻tg⁺	3.29	36.0	10_1	Mononucleotide	Hexagonal	22.79		32.9	1	Para-crystalline	12
ARNOTT B-B' (first model)		145.3	-86.9	-52.1	-136.4	39.4	156.5	144.5	-7.9	-1.0	—	-0.01	tg⁻g⁻tg⁺	3.24	36.0	10_1	Mononucleotide	Orthorhombic	17.82	20.00	32.44	1	Crystalline	12
ARNOTT B-B' (2nd model)	ApA	-170	-59	-80	172	54	81	89	—	—	—	—	tg⁻g⁻tg⁺	3.23	36.0	10_1	Dinucleotide. No dyad symmetry.	Orthorhombic	18.7	35.5	32.3	2	Crystalline	20
	TpT	171	-121	-41	-174	43	152	147	—	—	—	—	ttg⁻tg⁺											

STRUCTURE	CONFORMATIONAL PARAMETERS RELATING TO VARIOUS PROPOSED MODELS (°)											RISE/RESIDUE h(A)	TURN/RESIDUE τ(°)	MOLECULAR SYMMETRY	REPEATING UNIT	LATTICE AND PARAMETERS						REFERENCES
	α	β	γ	δ	ε	ζ	χ	FILE	TWIST	'γ'	D(A)	CLASS					TYPE	a(A)	b(A)	c(A)	MOLS/CELL	CRYSTALLINITY	
FORSYTH α-B'	141.6	-84.9	-55.6	-132.2	42.2	156.6	-31.6	-6.0	-1.0	—	0.00	$tg^-g^-tg^+$	3.28	36.0	10_1	Mononucleotide	Hexagonal	24.4	24.4	32.8	1	Para-crystalline	Chapter Five (this work)
B"-DNA	—	—	—	—	—	—	—	—	—	—	—	—	6.68	72.0	5_1	Dinucleotide	Orthorhombic	18.57	27.88	33.4	2	Para-crystalline	115, Chapter One (this work)
MARVIN C	212.4	-148.1	45.8	-216.7	48.0	140.4	139.2	-6.0	5.0	—	-2.24	ttg^-tg^+	3.32	38.6	28_3	Mononucleotide	Hexagonal	35.0	35.0	30.8	3	Semi-crystalline	121
																	Orthorhombic	32.2	20.2	31.1	2	Semi-crystalline	121
ARNOTT C	160.6	-105.8	-38.7	-160.0	37.2	156.5	143.1	-8.0	1.0	—	-0.9	$tg^-g^-tg^+$	3.32	38.6	28_3	Mononucleotide	No unique x-ray data. Model was qualitatively refined against Marvin data (above)						15
ARNOTT C'	167.1	-114.3	-33.6	-172.1	37.9	156.3	143.6	-8.1	1.8	—	-0.15	$tg^-g^-tg^+$	3.31	40.0	9_1	Mononucleotide	No x-ray data at all. Model refined from Arnott C helix to permit 9_1 symmetry.						15
FORSYTH C'	-39.7	138.0	-96.8	170.6	10.3	155.3	-11.1	-6.0	6.0	—	-3.2	$g^-tg^-tg^+$	3.31	40.0	9_1	Mononucleotide	Hexagonal	34.8	34.8	29.8	3	Crystalline	Chapter 4 (this work)
DISTORTED MARVIN C'	-139.3	-148.8	-45.4	142.0	33.5	140.4	132.0	-6.0	5.0	—	-2.2	ttg^-tg^+	3.31	40.0	9_1	Mononucleotide	Marvin 28_3 C helix distorted to meet the requirements of 9_1 molecular symmetry						Chapter 4 (this work)
ARNOTT C"	—	—	—	—	—	—	—	—	—	—	—	—	6.47	80.0	9_2	Dinucleotide	Hexagonal	22.1	22.1	58.2		Crystalline	107
MITSUI α-D	—	—	—	—	—	—	05' endo	—	—	—	—	—	3.13	-45.0	8_7	Mononucleotide	Tetragonal	17.52	17.52	25.01	1	Crystalline	126

STRUCTURE	CONFORMATIONAL PARAMETERS RELATING TO VARIOUS PROPOSED MODELS (°)												RISE/RESIDUE h(Å)	TURN/RESIDUE r(°)	MOLECULAR SYMMETRY	REPEATING UNIT	LATTICE AND PARAMETERS					REFERENCES
	α	β	γ	δ	ε	ζ	χ	TILT	TWIST	'γ'	D(Å)	CLASS					TYPE	a(Å)	b(Å)	c(Å)	MOLS/CELL	CRYSTALLINITY	
ARNOTT α-D	141.4	-100.5	-61.8	-152.0	68.6	156.5	143.9	-16.0	5.6	—	-0.18	$tg^-g^-tg^+$	3.04	45.0	8_1	Mononucleotide	Tetragonal	17.0	17.0	24.3	1	Crystalline	13
ARNOTT WRINKLED α-D-d(AT) ApT	-157	-145	-45	142	55	142	128	—	—	—	—	ttg^-tg^+	6.08	90.0	4_1	Dinucleotide	Tetragonal	17.0	17.0	24.3	1	Crystalline	21
TpA	-133	-153	-72	143	69	154	138	—	—	—	—	ttg^-tg^+											
ARNOTT WRINKLED α-D-d(IC) IpC	-130	-168	-63	136	60	149	136	-13.4	6.8	16.2	-2.1	ttg^-tg^+	6.13	90.0	4_1	Dinucleotide	Tetragonal	17.1	17.1	24.5	1	Crystalline	21
CpI	-109	-152	-81	134	65	151	146	-13.4	1.1	15.0	-2.1	$g^-tg^-tg^+$											
DREW α-D	-168.4	-143.9	-154.9	75.4	153.9	131.9	30.8	~16	—	—	—	$tttg^+t$	6.94	-102.9	7_5	Dinucleotide	Tetragonal	—	—	48.6	1	Crystalline	58
	72.3	159.1	33.7	-175.5	-87.9	151.3	-114.3	~16	—	—	—	$g^+tg^+tg^-$											
MAHENDRA-SINGAM α-D	173.0	-63.7	148.2	114.0	159.4	153.0	-148.1	-6.0	8.5	—	0.17	tg^-tg^+t	3.01	-45	8_7	Dinucleotide	Tetragonal	17.1	17.1	24.1	1	Crystalline	115
ARNOTT β-D	140.4	-104.1	-56.2	-158.8	62.8	—	140.4	—	—	—	-0.16	$tg^-g^-tg^+$	3.01	45	8_1	Mononucleotide	Hexagonal	19.8	19.8	24.1	1	Para-Crystalline	13
FORSYTH β-D	175.1	-79.3	147.0	126.7	163.4	153.6	-148.6	-10.0	8.5	—	0.20	tg^-ttt	3.18	-45	8_7	Mononucleotide	Hexagonal	21.2	21.2	25.4	1	Para-crystalline	Chapter 6 (this work)
γ-D	—	—	—	—	—	—	—	—	—	—	—	—	~3	45/-45(?)	$8_1/8_7$ (?)	Dinucleotide	Approx. Tetragonal	~17	~17	~24	1	Crystalline	Mahendrasingam (unpublished), Chapter One (this work)
F-DNA	—	—	—	—	—	—	—	—	—	—	—	—	3.04	47.0	2_3 (?)	?	Tetragonal	16.9	16.9	69.9	1	Crystalline	115, Chapter One (this work)

STRUCTURE	CONFORMATIONAL PARAMETERS RELATING TO VARIOUS PROPOSED MODELS (°)													RISE/RESIDUE h(A)	TURN/RESIDUE τ(°)	MOLECULAR SYMMETRY	REPEATING UNIT	LATTICE AND PARAMETERS					REFERENCES
																		TYPE	a(A)	b(A)	c(A)	MOLS/CELL	CRYSTALLINITY	
ARNOTT 'PLEIOMERIC' F-DNA	1 TpA	-152	-149	-64	138	78	841	120	-17.5	2.8	20.1	-3.9	ctg⁻tg⁺	18.5	49.5	4_3	Hexanucleotide	Tetragonal	17.2		74.0	1	Crystalline	124
	2 ApT	-169	-104	-84	-178	74	140	137	-15.8	-2.8	19.1	-2.3	tg⁻g⁺tg⁺		36.5								
	3 TpA	-96	-160	-103	124	85	143	137	-18.7	2.3	23.3	-0.9	g⁻tg⁻tg⁺		54.4								
	4 ApT	172	-101	-78	-176	74	138	137	-18.7	-3.3	20.0	-0.9	tg⁻g⁺tg⁺		38.7								
	5 TpA	-114	-160	-83	126	81	141	140	-15.8	2.8	22.7	-2.3	g⁻tg⁻tg⁺		54.4								
	6 ApT	179	-96	-72	171	68	130	129	-17.5	-2.8	18.0	-3.9	tg⁻g⁻tg⁺		36.5								
E-DNA													ttttt	3.25	48.7	3_2	Pentanucleotide	Orthorhombic	36.3	21.0	48.7	—	Crystalline	44
ARNOTT SI	GpC	-103	-91	-110	-168	54	147	-159	5	—	7	—	g⁻g⁻g⁻tg⁺	7.25	-60.0	6_5	Di-nucleotide	Hexagonal	19.1	19.1	43.5	1	Crystalline	17
	CpG	-72	102	52	-153	176	76	89	5	—	18	—	g⁻g⁺g⁺tt										
MAHENDRASINGAM SI	GpC	-85.0	-75.0	13.0	-164.0	-144.0	100.0	120.4	-2.0	6.0	—	-3.2	g⁻g⁻g⁺tt	7.13	-60.0	6_5	Dinucleotide	Hexagonal	17.9	17.9	42.8	1	Crystalline	115
	GpG	-100.0	98.0	-151.0	-158.0	77.0	139.0	91.9	2.0	6.0	—	-3.2	g⁻g⁺ctg⁺										
MAHENDRA-SINGAM SII	GpC	-144	-50.0	169.0	-174.1	76.3	92.0	-177.4	8.0	4.0	—	-3.0	tg⁻ctg⁺	7.55	-60.0	6_5	Dinucleotide	Hexagonal	20.0	20.0	45.3	1	Crystalline	115
	GpG	-100.0	85.0	-132.0	-136.0	27.0	146.0	-84.8	8.0	4.0	—	-3.0	g⁻g⁺ctg⁺										
WANG ZI	CpC	-93.7	80.5	-136.7	-139.0	55.5	138.0	-96	-5	—	7	-3.2	g⁻g⁺ctg⁺	7.43	-60.0	6_5	Dinucleotide	$P2_12_12_1$	17.9	31.6	44.6	Single crystal study of hexamer d (CpGpCpGpCpG)		174
	CpC	-104.7	-68.4	47.3	178.6	-169.1	99.7	128.3	-5	—	18	-3.2	g⁻g⁺tt											
WANG ZII	GpC	-103.4	77.4	143.7	162.8	68.2	146.7	-84.4	-5	—	18	-3.2	g⁻g⁺ctg⁺	7.43	-60.0	6_5	Dinucleotide							
	CpC	-179.5	56.3	90.9	-167.3	156.9	93.6	122.0	-5	—	18	-3.2	tg⁺g⁺tt											

STRUCTURE		CONFORMATIONAL PARAMETERS RELATING TO VARIOUS PROPOSED MODELS (°)											RISE/RESIDUE h(A)	TURN/RESIDUE t(°)	MOLECULAR SYMMETRY	REPEATING UNIT	LATTICE AND PARAMETERS						REFERENCES
		α	β	γ	δ	ε	ζ	χ	TILT	TWIST	'γ'	D(A)	CLASS				TYPE	a(A)	b(A)	c(A)	MOLS/CELLS	CRYSTALLINITY	
Poly d(As^{4}T)·poly d(As^{4}T)	As^{4}TpA	-92	-106	-77	-96	-148	36	135	-154	—	—	—	—	7.71	-31.4	$\frac{7}{6}$	Dinucleotide	Hexagonal	23.4	23.4	56.0	—	Crystalline	156,17
		-158	161	80	85	—	—	—	—	—	—	—	—											

BIBLIOGRAPHY

-A-

1. Arnott, S., and Wonacott, A.J., (1966), Polymers 7, 157.

2. Arnott, S., Hutchinson, F., Spencer, M., Wilkins, M.H.F., Fuller, W. and Langridge, R., (1966), Nature 211, 227.

3. Arnott, S., Wilkins, M.H.F., Fuller, W., Venable, J.H., and Langridge, R., (1967), J. Mol. Biol. 27, 549.

4. Arnott, S., Fuller, W., Hodgson, A.R., Prutton, I., (1968), Nature 220, 561.

5. Arnott, S. and Hukins, D.W.L., (1972a), Biochem. J. 130, 453.

6. Arnott, S. and Hukins, D.W.L., (1972b), Biochem. Biophys. Res. Commun., 47, 1504.

7. Arnott, S. and Bond, P.J., (1973a), Science 181, 68.

8. Arnott, S. and Bond, P.J., (1973b), Nature New Biol. 244, 99.

9. Arnott, S. and Hukins, D.W.L., (1973a), J. Mol. Biol. 81, 93.

10. Arnott, S. and Hukins, D.W.L., (1973b), J. Mol. Biol. 81, 107.

11. Arnott, S. and Selsing, E., (1974a), J. Mol. Biol. 88, 501

12. Arnott, S. and Selsing, E., (1974b), J. Mol. Biol. 88, 509.

13. Arnott, S., Chandrasekaran, Hukins, D.W.L., Smith, P.J.C., Watts, L., (1974), J. Mol. Biol. 88, 523.

14. Arnott, S., Chandrasekaran, R., Marttila, C.M., (1974d), Biochem. J. 141, 537.

15. Arnott, S. and Selsing, E., (1975), J. Mol. Biol. 98, 265.

16. Arnott, S., Bond, P.J., Selsing, E., Smith, P.J.C., (1976a), Nuc. Acids Res. 3, 2459.

17. Arnott, S., Chandrasekaran, R., Birdsall, D.L., Leslie, A.G.W., Ratliff, R.L., (1980), Nature 283, 743.

18. Arnott, S., (1980), in Fibre Diffraction Methods eds. A.D.French and K.H. Gardner), A.C.S. Symp., Series 141.

19. Arnott, S., (1980), Report at British Biophys. Soc. Christmas Conf., Imperial College, London.

20. Arnott, S., Chandrasekaran, R., Hall, I.H., Puigjaner, L.C., (1983), Nucl. Acds. Res. 11, 4141.

21. Arnott, S., Chandrasekaran, R., Puigjaner, L.C., Walker, J.K., Hall, I.H., Birdsall, D.L., and Ratliff, R.L., (1983), Nucl.

Acids Res. 11, 1457.

22. Astbury, W.T., (1947), Symp. Soc. Exp. Biol. (Nucleic Acids) 1, 66, Cambridge University Press.

-B-

23. Bak, A.L., Zeuthen, J. and Crick, F.H.C., (1977), Proc. Nat. Acad. Sci. 74, 4, 1595.

24. Barrell, B.G., Air, G.M., Hutchinson, C.A. III, (1976), Nature 264, 34.

26. Bartenev, V.N., Golovanov, Eu.I., Kapitonova, K.A., Mokulskii, M.A., Volkova, L.I., Skuratovskii, Ya., (1983), J. Mol. Biol. 169, 217.

27. Bateman, J.E., Connolly, J.F., Sawyer, E.C., and Stephenson, R., (1982), Nuclear Inst. and Methods 196, 515.

28. Bentley, G.A., Finch, J.T., Lewitt-Bentley, A., (1981), J. Mol. Biol., 145, 771.

29. Berman, H.M. and Neidle, S., (1979a), in Stereodynamics of Molecular Systems ed. R.H. Sarma (Pergamon, New York).

30. Berman, H.M., Stallings, W., Carrell, H.L., Glusker, J.P., Taylor, G., Achari, A., (1979), Biopolymers 18, 2045.

31. Blake, A. and Peacocke, A.R., (1968), Nature (London) 219, 1320.

32. Bradley, D.F., and Wolf, M.K., (1959), Proc. Nat. Acad. Sci. U.S.A. 45, 944.

33. Bragg, W.L. and Perutz, M.F., (1952), Acta. Cryst. 5, 277.

34. Brahms, J., Pillet, J., Tan, T.P.L. and Hill, L.R., (1973), Proc. Nat. Acad. Sci. U.S.A. 70, 3352.

35. Bram, S., (1976a), J. Mol. Biol. 58, 277.

36. Bram, S., (1976b), Nature New Biol. 232, 174.

37. Bram, S., (1976c), Nature New Biol. 232, 161.

38. Buerger, M.J., (1942) in 'X-ray Crystallography; an introduction to the investigation of crystals by their diffraction of monochromatic radiation', (New York).

39. Butler, J.A.V., Johns, E.W., Phillips, D.M.P. (1969), Prog. Biophys. molec. Biol. 18, 211.

-C-

40. Cairns, J., (1962), Cold Spring Harbour Symp. Quant. Biol. 27, 311.

41. Campbell-Smith, P.J. and Arnott, S., (1978), Acta. Cryst. A34, 3.

42. Chambers, R., (1924), Cellule 35, 107.

43. Chandrasekaran, R. and Leslie, A.G.W., (1976).

44. Chandrasekaran, R., Arnott, S., Banerjee, A., Campbell-Smith, P.J., Leslie, A.G.W., Puigjaner, L., (1980), in 'Fibre Diffraction Methods' (eds. A.D.French and K.H.Gardner), A.C.S. Symp. Series 141.

45. Cheng, P.Y., (1965), Biochem. Biophys. Acta 102, 314.

46. Cochran, W., Crick, F.H.C. and Vand, V., (1952), Acta. Cryst. 5, 581.

47. Crawford, J.L., Kolpack, F.J., Wang, A.H.J., Quigley, G.J., van Broom, J.H., van der Marel, G., Rich, A., (1980), Proc. Nat. Acad. Sci. U.S.A. 77, 4016.

48. Crick, F.H.C. and Watson, J.D., (1954), Proc. Roy. Soc. 223A, 80.

49. Crick, F.H.C., and Brenner, S., (1967), J. Mol. Biol. 26, 361.

-D-

50. Davies, D.R. and Baldwin, R.L., (1963), J. Mol. Biol. 6, 151.

51. Dickerson, R.E. and Drew, H.R., (1981), J. Mol. Biol. 149, 761.

52. Donohue, J. and Trueblood, K.N., (1960), J. Mol. Biol. 2, 363.

53. Dover, S.D., (1977), J. Mol. Biol. 110, 699.

54. Drake, J.W., (1964), J. Cell. Comp. Physiol. 64, 19.

55. Drew, H.R., Dickerson, R.E., Itakura, K., (1978), J. Mol. Biol. 125, 535.

56. Drew, H.R. and Dickerson, R.E., (1981), J. Mol. Biol. 151, 535.

57. Drew, H.R. and Dickerson, R.E. (1981), J. Mol. Biol., 152, 723.

58. Drew, H.R. and Dickerson, R.E., (1982), EMBO Journ. 1, 663.

-E-

59. Ehrenberg, W. and Spear, W.E., (1951), Proc. Phys. Soc. Lond. B64, 67.

60. Elliott, A., (1965), J. Sci. Inst. 42, 312.

61. Erdmann, V.A., (1976), Nucl. Acids Res. 6, r29.

62. Erfurth and Peticolas, (1975), Biopolymers 14, 247.

-F-

63. Finch, J.T. and Klug, A., (1976), Proc. Nat. Acad. Sci. U.S.A. 73, 6, 1897.

64. Finch, J.T., Brown, R.S., Rhodes, D., Richmond, T., Rushton, B.,Lutter, L.C., Klug, A., (1981), J. Mol. Biol. 145, 757

65. Forsyth, V.T., Mahendrasingam, A., Pigram, W.J., Greenall, R.J., Nave, C. and Fuller, W. (1984) The 8th International Biophysics Congress.

66. Forsyth, V.T., Greenall, R.J., Mahendrasingam, A., Nave, C., Pigram, W.J. and Fuller, W. in 'Biological Systems: Structure and Analysis' Proc. Daresbury Study Weekend 1984 (in press).

67. Franklin, R. and Gosling, R.G., (1953a), Acta Cryst. 6, 671.

68. Franklin, R. and Gosling, R.G., (1953b), Acta Cryst. 6, 678.

69. Franklin, R. and Gosling, R.G., (1955), Acta Cryst. 8, 151.

70. Franks, A., (1958), Brit. J. Appl. Phys. 9, 349.

71. Fraser, R.D.B., MacRae, T.P., Suzuki, E., (1978), J. Appl. Cryst 11, 693.

72. Fuller, W., (1961), PhD Thesis, University of London.

73. Fuller, W. and Waring, M.J., (1964), Ber. Bunsenges. Phys. Chem. 68, 805.

74. Fuller, W., Wilkins, M.H.F., Wilson, H.R. and Hamilton, L.D., (1965) J. Mol. Biol. 12, 60.

75. Fuller, W., Hutchinson, F., Spencer, M. and Wilkins, M.H.F., (1967) J. Mol. Biol. 27, 507.

76. Fuller, W., Pigram, W.J., Mahendrasingam, A., Forsyth, V.T., Nave, C., Greenall, R.J. in 'Biological Systems: Structure and Analysis', Proc. Daresbury Study Weekend 1984 (in press).

77. Fuller, W., Mahendrasingam, A., Pigram, W.J., Greenall, R.J., Forsyth, V.T. and Nave, C. (1984), The International Biophysics Congress.

-G-

78. Goodwin, D., (1977), PhD Thesis, University of Keele.

79. Greenall, R.J., Pigram, W.J., Fuller, W., Nature 282 5741, 880.

80. Greenall, R.J., (1982), PhD Thesis, University of Keele.

81. Greenall, R.J., Mahendrasingam, A., Forsyth, V.T., Pigram, W.J., and Fuller, W. (1984) 'Biological Systems: Structure and Analysis', Proc. Daresbury Study Weekend 1984.

82. Greenall, R.J., Nave, C., Helliwell, J.R., Burley, S.K., Miller, A., Sumner, I., Berry, A., Przybylski, M., Rake, F.M., Ridley, P.A., Worgan, J.S., (1984) to be subm. J. Appl. Cryst.

83. Greenhough, T., and Helliwell, J.R., (1983), Prog. Biophys. molec. Biol. 41, 67.

84. Griffith, J.D., (1978), Science 201, 525.

85. Gupta, G., Bansal, M., Sasisekaran, B., (1980a), Proc. Nat. Acad. Sci. U.S.A., 77, 6486.

86. Gupta, G., Bansal, M., Sasisekaran, B., (1980b), Int. J. Biol. Macromol., 2, 368.

-H-

87. Hanlon, S., Brundo, S., Wu and Wolf, B., (1975), Biochemistry 14, 8.

88. Hogan, M., Dattagupta, N., Crothers, D.M., (1978), Proc. Nat. Acad. Sci. U.S.A., 75, 195.

89. Holley, R.W., (1968), in 'The Molecular Basis of Life', eds. R.H.Haynes and P.C.Hanawalt, Freeman, San Francisco.

90. Hoogsteen, K., (1959), Acta Cryst., 12, 822.

91. Hopkins, R.C., (1981), Science 211, 289.

-J-

92. Johnson, C.K., (1965), ORTEP: Oak Ridge National Laboratory report, ORNL-3794, Oak Ridge, Tennesse.

-K-

93. Kim, S.H., Suddarth, F.L., Quigley, G.J., McPherson, A., Sussman, J.L., Wang, A.J., Seeman, N.C. and Rich, A., (1974), Science 185, 435.

94. Klug, A., Jack, A., Viswamitra, M.A., Kennard, O., Shakked, Z., Steitz, T.A., (1979), J. Mol. Biol.

131, 669.

95. Klysik, J., Stirdivant, S.M., Larson, J.E., Hart, P.A., Wells, R.D., (1981), Nature 290, 672.

96. Kornberg, R.D. and Thomas, J.D., (1974), Science 184, 865.

-L-

97. Ladner, J.E., Jack, A., Robertus, J.D., Brown, R.S., Rhodes, D., Clark, B.F.C. and Klug, A., (1975), Proc. Nat. Acad. Sci. U.S.A. 72, 4414.

98. Lafer, E.M., Moller, A., Nordheim, A., Stollar, B.D. and Rich, A., (1981), Proc. Nat. Acad. Sci. U.S.A., 78, 3546.

99. Langridge, R., Seeds, W.E., Wilson, H.R., Hooper, C.W., Wilkins, M.H.F. and Hamilton, L.D., (1960a), J. Mol. Biol. 2, 19.

100. Langridge, R., Marvin, D.A., Seeds, W.E., Wilson, H.R., Hooper, C.W., Wilkins, M.H.F. and Hamilton, L.D., (1960b), J. Mol. Biol. 2, 38.

101. Langridge, R. and Gomatos, P.J., (1963), Science 141, 694.

102. Lea, K.R. and Munro, I.H., (1980), Daresbury Laboratory publiaction.

103. Lemunier, F., Derbin, C., Malfoy, B., Leng, M. and Tailandier, E., (1982), Expl. Cell Res. 141, 508.

104. Lerman, L.S., (1961), J. Mol. Biol. 3, 634.

105. Lerman, L.S., (1963), Proc. Nat. Acad. Sci. U.S.A. 49, 84.

106. Lerman, L.S., (1964), J. Cell. Comp. Physiol. 64, 1.

107. Leslie, A.G.W., Arnott, S., Chandrasekaran, R. and Ratliff, R.L., (1980), J. Mol. Biol. 143, 49.

108. Levitt, M., (1974), J. Mol. Biol. 82, 393.

109. Levitt, M., (1978), Proc. Nat. Acad. Sci. U.S.A. 75, 641.

110. Li, H.J. and Crothers, D.M., (1969), J. Mol. Biol. 39, 461.

111. Li, H.J. and Crothers, D.M., (1969), Biopolymers 8, 217.

112. Luzzati, V., Masson, F., Lerman, L.S., (1961), J. Mol. Biol. 3, 634.

-M-

113. Mahendrasingam, A., Rhodes, N.J., Goodwin, D.C., Nave, C., Pigram, W.J., Fuller, W., Brahms, J. and Vergne, J., (1983), Nature 301, 5900. 535.

114. Mahendrasingam, A., Pigram, W.J., Fuller, W., Brahms, J., Vergne, J., (1983), J. Mol. Biol. 168, 897.

115. Mahendrasingam, A., (1983), PhD Thesis, University of Keele.

116. Mahendrasingam, A., Pigram, W.J., Greenall, R.J., Forsyth, V.T., Nave,C., Fuller, W., (1984)8th International Biophysics Congress.

117. Mahendrasingam, A., Pigram, W.J., Fuller, W., Forsyth, V.T. and Greenall, R.J., (1984), in 'Biological Systems: Structure and Analysis', Proc. Daresbury Study Weekend

118. Malfoy, B., Hartmann, B. and Leng, M., (1981), Nuc. Acds. Res. 9, 5659.

119. Marvin, D.A., Spencer, M., Wilkins, M.H.F. and Hamilton, L.D., (1958), Nature 182, 387.

120. Marvin, D.A., (1960), PhD Thesis, University of London.

121. Marvin, D.A., Spencer, M., Wilkins, M.H.F. and Hamilton, L.D., (1961), J. Mol. Biol. 3, 547.

122. Marvin, D.A., Wilkins, M.H.F., Hamilton, L.D., (1966), Acta Cryst. 20, 663.

123. Meselson, M. and Stahl, F.W., (1958), Proc. Nat. Acad. Sci. 44, 671.

124. Millane, R.P., Walker, J.K., Arnott, S., Chandrasekaran, R., Birdsall, D.L., Ratliff, R.L., (1984), submitted to Nuc. Acids Res.

125. Milman, G., Langridge, R., Chamberlain, M.J., (1967), Proc. Nat. Acad. Sci. U.S.A., 57, 1804.

126. Mitsui, Y., Langridge, R., Shortle, B.E., Cantor, C.R., Grant, R.C., Kodama, M., Wells, R.D., (1970), Nature 288, 1166.

127. Moller, A., et al., (1982), J. Biol. Chem. 257, 12081.

128. Morgenegg, G., Celio, M.R., Malfoy, B., Leng, M. and Kuenzle, C.C., (1983), Nature 303, 540.

-N-

129. Nave, C. and Greenall, R.J., (1984), in 'New Methods in X-ray Absorption, Scattering and Diffraction' eds. R.Britton-Chanance and H.D. Bartunik, Academic Press (in press).

130. Neidle, S., Achari, A., Taylor, G.L, Berman, H.M., Carrell, H.L., Glusker, J.P., Stallings, W.C., (1977), Nature 269, 304.

131. Neidle, S. and Berman, H.M., (1983), Prog. Biophys. Molec. Biol. 41, 43.

132. Neville, D.M. Jr. and Davies, D.R., (1966), J. Mol. Biol. 17, 57.

133. Nordheim, A., et al., (1982), Cell. 31, 309.

-O-

134. O'Brien, E.J. and MacEwan, A.W., (1970), J. Mol. Biol. 48, 243.

135. Olins, A.L. and Olins, D.E., (1974), Science 183, 330.

-P-

136. Peacocke, A.R. and Skerrett, J.N.H., (1956), Trans. Farad. Soc. 52, 261.

137. Peacocke, A.R., (1969), in Chemistry and Industry, 642.

138. Pigram, W.J., (1968), PhD Thesis, University of London.

139. Pigram, W.J., Fuller, W. and Hamilton, L.D., (1972), Nature New Biol. 235, 17.

140. Pohl, F.M., (1971), First European Biophysics Congress, p343.

141. Pohl, F.M. and Jovin, T.M., (1972), J. Mol. Biol. 67, 375.

142. Pohl, F.M., Ranade, A., Stockburger, M., (1973), Biochem. Biophys. Acta 333, 83.

143. Pohl, F.M., Thomas, R., DiCapua, E., (1982), Nature 300, 545.

144. Pritchard, N.J., Blake, A., and Peacocke, A.R., (1966), 212, 1360.

-R-

145. Ramaswamy, N., Bansal, M., Gupta, G., (1982), Proc. Nat. Acad. Sci. U.S.A. 79, 6109.

146. Ramstein, J., Dourlent, M., Leng, M., (1972), Biochem. Biophys. Res. Commun. 47, 874.

147. Rhodes, N.J., Mahendrasingam, A.,Pigram, W.J., Fuller, W., Brahms, J., Vergne,J. and Warren, R.A.J., (1982), Nature 296, 267.

149. Rhodes, N.J., Mahendrasingam, A., Pigram, W.J., Fuller, W., Brahms, J., Vergne, J. and Warren, R.A.J., (1982), Nature 296, 267.

150. Rhodes, N.J., (1982), PhD Thesis, University of Keele.

151. Rich, A. and RajBhandary, U.L., (1976), Ann. Rev. Biochem. 45, 805.

152. Robertus, J.D., Ladner, J.E., Finch, J.T, Rhodes, D., Brown, R.S., Clark, B.F.C. and Klug, A., (1974), Nature 250, 546.

153. Rodley, G.A., Scobie, R.S., Bates, R.H.T., Lewitt, R.M., (1976), Proc. Nat. Acad. Sci. U.S.A. 73, 2959.

154. Rodley, G.A. and Bates, R.H.T., (1980), T.I.B.S. 5, 231.

-S-

156. Saenger, W., Landmann, Lezius, A.G., (1973), Jr. Symp. Quantum Chemistry and Biochemistry 5, 457.

157. Sage, E. and Leng, M., (1980), Proc. Nat. Acad. Sci. U.S.A. 77, 4597.

158. Selsing, E., Arnott, S., Ratliff, R.L, (1975), J. Mol. Biol. 98, 243.

159. Shaw, D.C., Walker, J.E., Northrop, F.D., Barrell, B.G., Godson, G.N., and Fiddes, J.C., (1978), Nature 272, 510.

160. Skuratovskii, I.Ya., Volkova, L.I., Kapitonava, X.A. and Bartenev, V.N., (1979), J. Mol. Biol. 134, 369.

161. Smith, M., Brown, N.L., Air, G.M., Barrell, B.G., Coulson, A.R., Hutchinson, C.A., Sanger, F., (1977), Nature 265, 702.

162. Spencer, M., (1959), Acta Cryst. 12, 59.

163. Streiber, J.P. and Daune , M.P., (1974), J. Mol. Biol. 83, 487.

164. Streisinger, G., Okada, Y., Emrich, J., Newton, J., Tsugita, A., Terzaghi, E., Inouye, M., (1966), Cold Spring Harbour Symp. Quant. Biol. 31, 77.

165. Sundralingam, M. and Jensen, L.H., (1965), J. Mol. Biol. 13, 94.

-T-

166. Thamann, T.J., Lord, R.C., Wang, A.H.J., Rich, A., (1981), Nuc. Acds. Res. 9, 5443.

167. Tomita, K.I. and Rich, A., (1964), Nature 201, 1160.

168. Tubbs, R.K., Ditmars, W.E.Jr. and Van Winkle, Q., (1964), J. Mol. Biol. 9, 545.

169. Tunis-Schneider, M.-J.B. and Hearst, J.E., (1968), Biopolymers 6, 1218.

170. Tunis-Schneider, M.-J.B. and Maestre, M.F., (1970), J. Mol. Biol. 52, 521.

-V-

171. Vainshtein, B.K., (1966), in 'The Diffraction of X-rays by Chain Molecules', Elsevier Publishing Company.

-W-

172. Wang, J.C., (1979), Proc. Nat. Acad. Sci., U.S.A. 76, 200.

173. Wang, A.H.J., Quigley, G.J., Kolpack, F.J, Crawford, J.L., van Boom, J.H., van der Marel, G., Rich, A., (1979), Nature 282, 680.

174. Wang, A.H.J., Quigley, G.J., Kolpack, F.J., van der Marel, G., van Boom, J.H., Rich, A., (1981), Science 211, 171.

175. Waring, M.J., (1968), Nature (London) 219, 1320.

176. Watson, J.D. and Crick, F.H.C., (1953a), Nature 171, 737.

177. Watson, J.D. and Crick, F.H.C., (1953b), Nature 176, 965.

178. Watson, J.D., (1976), in 'The Molecular Biology of the Gene', (Addison-Wesley).

179. Wilkins, M.H.F. and Randall, J.T, (1953), Biochem. Biophys. Acta 10, 192.

180. Wycoff, H.W., (1955), PhD Thesis, Massachusetts Institute of Technology.

-Z-

181. Zimmerman, S.B. and Pheiffer, B.H., (1980), J. Mol. Biol. 142, 315.

182. Zimmerman, S.B., (1982), Ann. Rev. Biochem. 51, 395.

180. Wycoff, H.W., (1955), PhD Thesis, Massachusetts Institute of Technology.

-Z-

181. Zimmerman, S.B. and Pheiffer, B.H., (1980), J. Mol. Biol. 142, 315.

182. Zimmerman, S.B., (1982), Ann. Rev. Biochem. 51, 395.