TITLE X-RAY FIBRE DIFFRACTION STUDIES ON THE POLYMORPHISM OF DNA AND ITS SYNTHETIC ANALOGUES.

AUTHOR Arumugam MAHENDRASINGAM

DEGREE Ph.D

AWARDING BODY Keele University

DATE 1983

THESIS NUMBER DX187422

# X-RAY FIBRE DIFFRACTION STUDIES ON THE POLYMORPHISM OF DNA AND ITS SYNTHETIC ANALOGUES

by

Arumugam Mahendrasingam, B.Sc(Cey), M.Inst.P.

A Thesis submitted for the Degree of Doctor of Philosophy

to the University of Keele

1983

Department of Physics

University of Keele

Staffordshire

England

## ABSTRACT

The known conformations of DNA are briefly reviewed. Methods which are developed and used to prepare DNA fibres are described.

A left-handed model having $8_7$ symmetry has been proposed in an attempt to explain the D-DNA diffraction pattern.

Conditions are described which enable, the observation of a great variety of different conformations (C,A,B',B,α-D,β-D,D',D-type) from the synthetic polymer poly d(AT).poly d(AT). Many different conformations can be observed from suitably prepared single fibres of Na poly d(AT). poly d(AT). The D-conformation emerges as a particularly stable structure. Transitions between B and D forms are readily reversible, but those between A and D are more complex. Patterns indicating 'mixtures' of different conformations (e.g. A and C, A and B) are described. The possibility of an alternating structure for co-polymers is outlined.

Conditions are also described for the observation of a number of different conformations (C,A,B",B,S) from the synthetic polynucleotide poly d(GC).poly d(GC). Two different 'S' conformations have been detected and these are designated $S_I$ and $S_{II}$. Left-handed models giving satisfactory agreement indices between observed and calculated data are described for these two structures. Another unique observation relating to this polymer is the fact some of its transitions were not only sensitive to the relative humidity of the fibre environment, but also to the rate at which this humidity was changed.

The C→A→B sequence of transitions was observed for most natural DNAs (excepting $T_2$ DNA and SP-15 DNA), poly d(AT).poly d(AT) and poly d(AC). poly d(GT). The possible mechanism behind this sequence, especially with respect to hydration geometry has been given some speculation. Conditions necessary for the acquisition of the C,C',A,B and 'D-type' conformations are described.

Finally, conditions are also described by which the C, α-B' and β-B' conformations can be observed from poly d(A).poly d(T), and the possibility of 'heteronomous' structures occurring in such homopolymers is perused.

ACKNOWLEDGEMENTS

I would like to thank my supervisor, Professor Watson Fuller, for his supervision and provision of excellent facilities. I am grateful to other members of the group, Drs. W.J. Pigram, R.J. Greenall, N.J. Rhodes and Mr. V.T. Forsyth for constructive discussions, suggestions and prompt assistance when needed and for providing a pleasant and friendly atmosphere in which to work. I am also grateful to Mr. V.T. Forsyth for reading the manuscript. I am obliged to Professor K. Kunaratnam (University of Jaffna, Sri Lanka) for his help and continued interest.

I would like to express my gratitude to Professor J. Brahms (University of Paris, France) for the provision of synthetic polynucleotides, to Professor T. Warren (University of British Columbia, Canada) for provision of $\phi$w-14 and calf thymus DNA. Both these workers have been collaborating with this laboratory over the last few years. I would also like to thank Professor M. Mandel (University of Texas, U.S.A.) for provision of SP-15 DNA.

I would like to thank Mr. F. Rowerth and his technical staff for their excellent service. In particular, I would like to express my gratitude to the following members of the technical and secretarial staff:-

Mr. G. Dudley, Mr. G. Marsh and Mr. E.T.J. Greasley in the mechanical workshop for their excellent technical service, to

Mr. P. Clarke for prompt technical assistance in the maintenance of x-ray machines, to Mr. M. Daniels for producing the photographic prints in this thesis and technical advice, to Mr. Emley of the Geology Department for technical assistance in the use of the flame emission spectrometer.

I would like to thank Mrs. E.A. Durrant, Miss H. Martin and Mrs. K.B. Merifield for secretarial services. In particular, I am indebted to Miss H. Martin for fast and proficient typing of this thesis.

I would like to thank Ravi, Richard, David, Steve and all other friends for their help. In particular, I am obliged to Ravi for his unfailing assistance whenever it was needed.

I would like to record my gratitude to the University of Keele for financial support in the form of a Departmental studentship.

Finally, I am obliged to my family for their support and encouragement throughout my academic career.

CONTENTS

# CHAPTER 1

## Nucleic Acid Structure

### 1.1 Introduction

The genetic information of an organism is carried by its chromosomes. Chromosomes are highly organized structures made up of a nucleic acid component and a protein component. The nucleic acid component is usually deoxyribonucleic acid (DNA) and it is this molecule that carries the information responsible for the vast array of hereditary characteristics found in different lifeforms.

DNA is an unbranched polymer with a molecular weight of many millions. The repeating unit (monomer) is the nucleotide with a molecular weight of about 320. It consists of a deoxyribose sugar covalently bonded to both a phosphate group and a purine or pyrimidine base. There are four commonly occurring bases: two purines (adenine and guanine) and two pyrimidines (thymine and cytosine) (figure 1). In ribonucleic acid (RNA), an additional hydroxyl group is bonded to $C_2'$ (figure 3) and uracil (figure 1) is present instead of thymine. Successive nucleotides are joined by a phosphodiester linkage which joins the 3' hydroxyl group of the sugar of one nucleotide to the 5' hydroxyl group of the next (figures 2 and 3). It is the sequence of the bases on the polynucleotide chain that carries the genetic information. By convention the sequence of the bases is always written in a 5' to 3' direction. Thus in figure 2 and figure 3 the sequence of the polynucleotide chain in DNA is written as ApT or AT and in RNA ApU or AU.

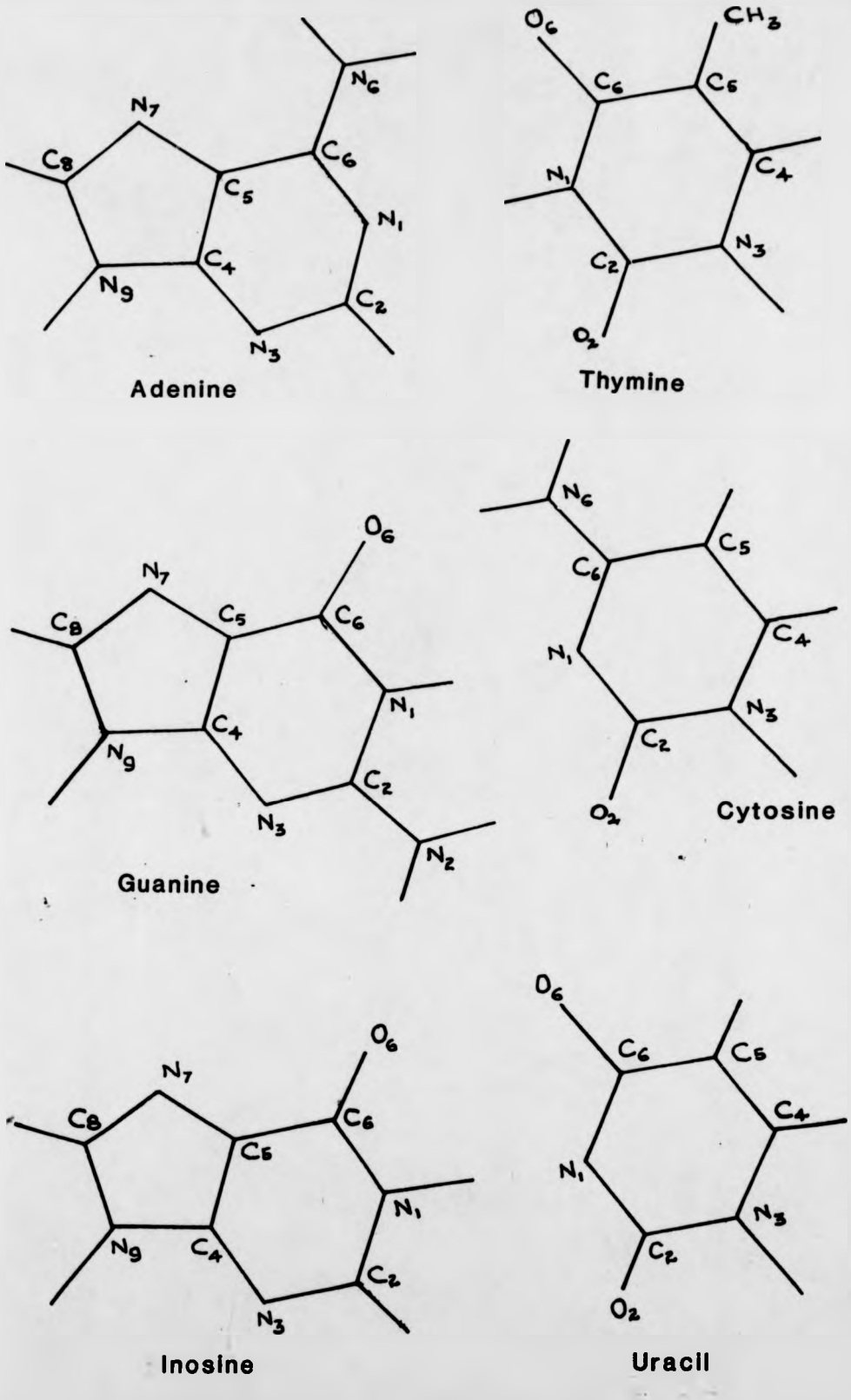



Figure 1 The Structure of Purine and Pyrimidine Bases

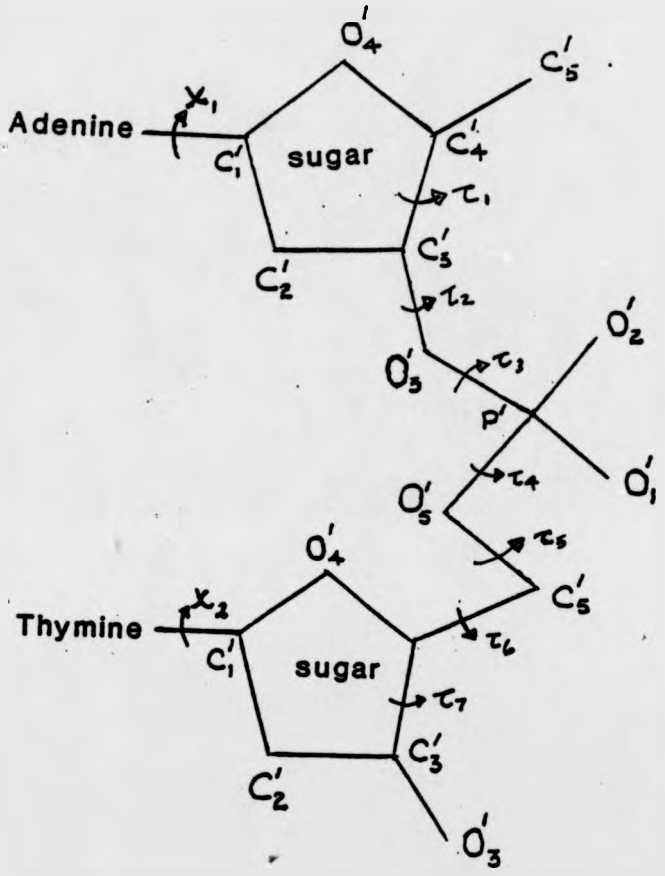



Figure 2 A Section of Polynucleotide Chain in DNA

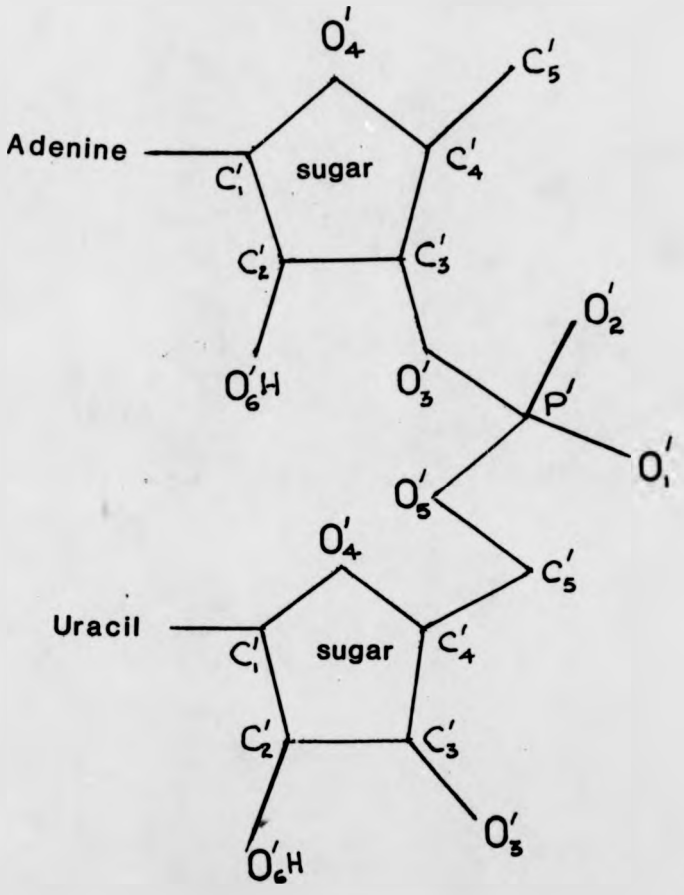



Figure 3 A Section of Polynucleotide Chain in RNA

Early X-ray diffraction studies of the molecular architecture of DNA was carried out by Astbury (1947) and later by Franklin and Gosling (1953a) and Wilkins. Using this information and other chemical observations, Watson and Crick (1953) proposed that the structure of the DNA molecule was double stranded and in the form of a right-handed helix with the two polynucleotide chains wound around the same axis and held together by hydrogen bonds between the bases (figure 4). The sugar and phosphate attached at each side of a base-pair are related by a two-fold rotation axis in the plane of the bases, and the helical symmetry of the molecule generates a further set of diad axes midway between the base planes.

Rodley et al (1976) suggested that fibre diffraction data for B-DNA could as readily fit a side-by-side (SBS) model having the chain changing from right- to left-handed helical sense at regular intervals. The main attraction of this model is that it would make the unwinding of the duplex during replication process easier to envisage. Greenall et al (1979) studied this model and concluded that the Watson-Crick model was preferable.

Recently, single crystal work, focussed on oligonucleotides, has provided further information about DNA structure to atomic resolution. The oligomer d(CpGpCpGpCpG), is known to crystallise in an unexpected left-handed conformation (Wang et al., 1979). This is called the Z-conformation and has antiparallel strands held together by Watson-Crick base-pairs.

The aim of this thesis is the study of the polymorphic behaviour of natural DNAs and its synthetic analogues. The remainder of this chapter is devoted to a discussion of the known conformations of DNA and an outline of the content of the project.

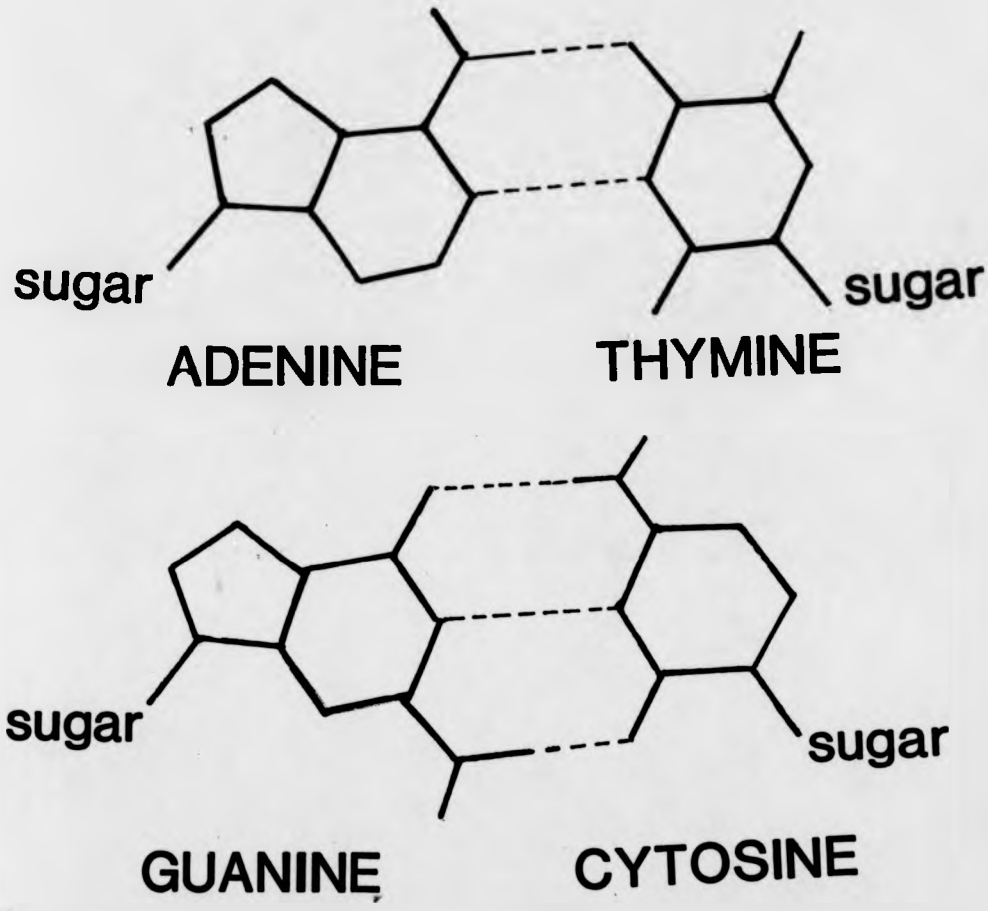



Figure 4 The Watson-Crick Base-Pairing Found in DNA

## 1.2 Structural Studies of DNA

The process by which fibre structures are determined involves trial and error. However, because we are dealing with a very long molecule and there is available a substantial amount of stereochemical information from single crystal studies of nucleosides and nucleotides, the number of possible solutions to the observed diffraction is considerably reduced. A survey of the bond lengths and angles observed in single crystal studies of nucleosides and nucleotides has been published by Arnott (1970), and his results were used in the structural studies by the author.

The conformational polymorphism of DNA is largely a result of variability in the backbone and glycosidic torsion angles, different possible sugar puckering arrangements, and different base positioning (Altona and Sundaralingam, 1972, Sundaralingam, 1969). The glycosidic torsion angle $\chi$, is defined by atoms $C_2'$, $C_1'$, N and $C_2$ in the case of pyrimidine or $C_4$ in a purine. When $\chi \simeq 90^o$ the conformation is said to be anti, when $\chi \simeq 300^o$ the conformation is syn. The backbone torsion angles fall into three sectors and these are conveniently used to describe the backbone conformation.

These sectors are:

| | | | |
|---|---|---|---|
| (i) | gauche$^+$ ($g^+$) | if | $0^o < \tau < 120^o$ |
| (ii) | gauche$^-$ ($g^-$) | if | $-120^o < \tau < 0^o$ |
| (iii) | trans (t) | if | $120^o < \tau < 240^o$ |

A theoretical study led Spencer (1959) to the conclusion that all the sugar ring atoms were unlikely to be coplaner, and that either the $C_2'$ or the $C_3'$ atom was likely to be displaced from the mean ring plane. Sundaralingam and Jenson (1965) confirmed this prediction from crystallographic studies of nucleotides and nucleosides. The sugar conformation is

## 1.2 Structural Studies of DNA

The process by which fibre structures are determined involves trial and error. However, because we are dealing with a very long molecule and there is available a substantial amount of stereochemical information from single crystal studies of nucleosides and nucleotides, the number of possible solutions to the observed diffraction is considerably reduced. A survey of the bond lengths and angles observed in single crystal studies of nucleosides and nucleotides has been published by Arnott (1970), and his results were used in the structural studies by the author.

The conformational polymorphism of DNA is largely a result of variability in the backbone and glycosidic torsion angles, different possible sugar puckering arrangements, and different base positioning (Altona and Sundaralingam, 1972, Sundaralingam, 1969). The glycosidic torsion angle $\chi$, is defined by atoms $C_2'$, $C_1'$, N and $C_2$ in the case of pyrimidine or $C_4$ in a purine. When $\chi \simeq 90^o$ the conformation is said to be anti, when $\chi \simeq 300^o$ the conformation is syn. The backbone torsion angles fall into three sectors and these are conveniently used to describe the backbone conformation.

These sectors are:

| | | | |
|---|---|---|---|
| (i) | gauche$^+$ ($g^+$) | if | $0^o < \tau < 120^o$ |
| (ii) | gauche$^-$ ($g^-$) | if | $-120^o < \tau < 0^o$ |
| (iii) | trans (t) | if | $120^o < \tau < 240^o$ |

A theoretical study led Spencer (1959) to the conclusion that all the sugar ring atoms were unlikely to be coplaner, and that either the $C_2'$ or the $C_3'$ atom was likely to be displaced from the mean ring plane. Sundaralingam and Jenson (1965) confirmed this prediction from crystallographic studies of nucleotides and nucleosides. The sugar conformation is

designated 'endo' if the atom most out of the plane $C_1'$ - $O'_4$ - $C_4'$ is on the same side as $C_5'$, otherwise the designation is 'exo'. The four conformations are shown in figure 5.

The position and orientation of the bases with respect to the helix axis is commonly defined by using three parameters (figure 6). These are:-

(i) The base-pairs displacement from the helix axis, denoted by the distance 'D'.

(ii) The 'tilt' of the base-pairs about the diad axis.

(iii) The 'twist' of the base-pairs about the twist axis, commonly the line connecting purine $C_8$ and pyrimidine $C_6$.

Several different forms of DNA have been observed. Natural DNAs are only observed in the A, B and C forms. Other forms are observed for synthetic polynucleotides. DNA conformations are commonly characterised by typical values for pitch, number of residues per turn, rise per residue, and the rotation per residue. The helical parameters of the observed DNA conformations in fibres are summarised in table 1. These DNA conformations are divided into three families. These are:

(i) A-family

Members of A-family are characterised by $C_3'$-endo sugars, positive base displacements and large positive base tilts. The sugar orientation is 'anti' with respect to the base.

(ii) B-family

Members of B-family (B,B',B",C,C',C",D,D' and E) are characterised by $C_2'$-endo (or $C_3'$-exo) sugars, small negative base displacement and small negative base tilts. The sugar orientation is 'anti' with respect to the base.

(iii) Z-family

Members of the Z-family are characterised by alternating

Table 1

The Conformational Parameters and Types of DNA Structures Observed in Fibres

| Conformational Type | Pitch (A ) | Helical Symmetry | Average Rise per Nucleotide residue (A ) | Average Turn angle per nucleotide residue (o ) | Reference |
|---|---|---|---|---|---|
| A | 28.2 | $11_1$ | 2.56 | 32.7 | Fuller et al (1965) |
| B | 33.8 | $10_1$ | 3.38 | 36.0 | Langridge et al (1960a) |
| B' | 32.9 | $10_1$ | 3.29 | 36.0 | Arnott and Selsing (1974a) |
| B" | 33.4 | $10_1$ | 3.34 | 36.0 | This Work |
| C | 31.0 | $28_3$ | 3.32 | 38.6 | Marvin et al (1961) |
| C' | 29.5 | $9_1$ | 3.28 | 40.0 | Leslie et al (1980) |
| C" | 58.2 | $9_2$ | 3.23 | 40.0 | Leslie et al (1980) |
| D | 24.3 | $8_1$ | 3.04 | 45.0 | Arnott et al (1974b) |
| D' | 24.0 | $4_1$ | 3.00 | 45.0 | This Work |
| E | 48.7 | $3_2$ | 3.25 | 48.0 | Leslie et al (1980) |
| S ($S_I$) | 43.5 | $6_5$ | 3.63 | -30.0 | Arnott et al (1980) |
| $S_{II}$ | 45.3 | $6_5$ | 3.78 | -30.0 | This Work |

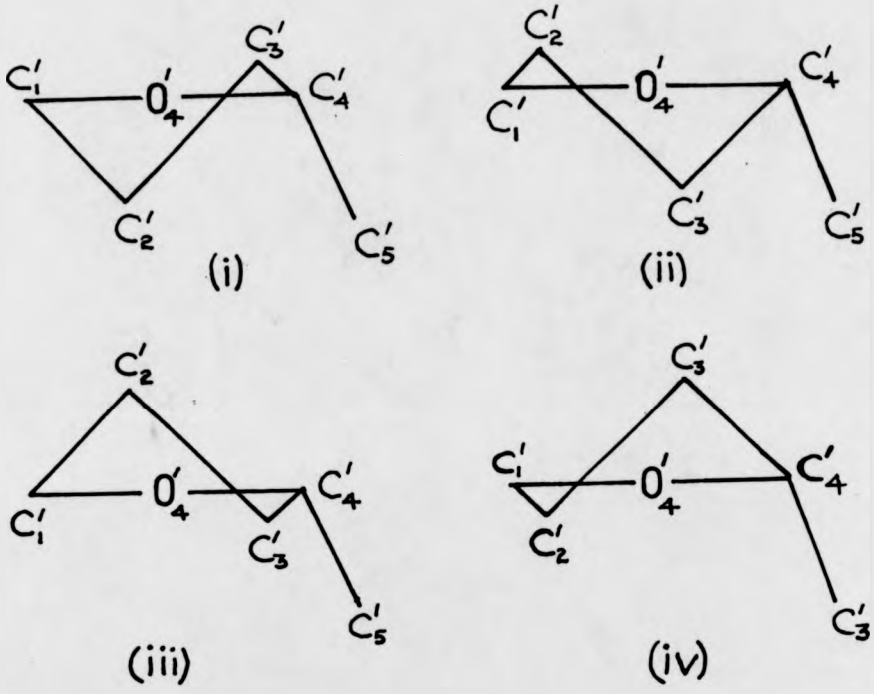



Figure 5 The Four Possible Sugar Puckers:

(i) $C_2'$-endo (ii) $C_3'$-endo

(iii) $C_2'$-exo (iv) $C_3'$-exo

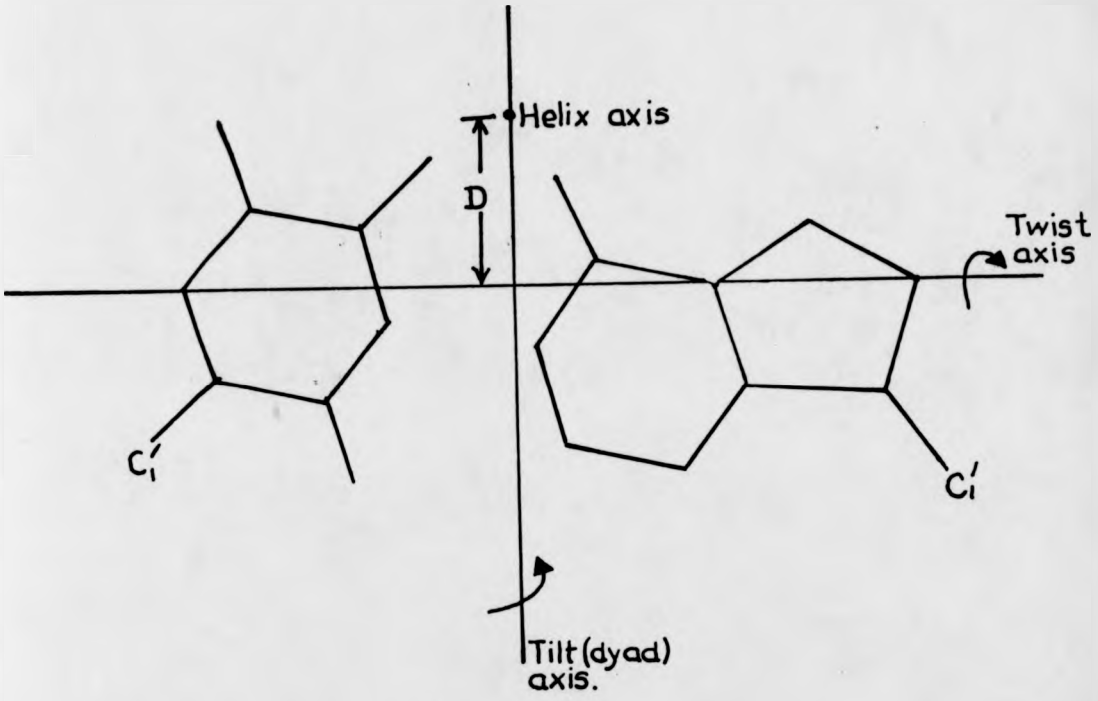



Figure 6 The Parameters Defining the Position of a Base-Pair Relative to the Helix Axis

sugar conformations. Deoxycytidine is observed in the 'anti' conformation with $C_2$'-endo sugar pucker (for Z and Z') and deoxyguanosine observed in 'syn' conformation with $C_1$'-exo (for Z') and $C_3$'-endo (for Z) sugar puckers (Wang et al., 1979 and Drew et al., 1980).

1.2.1 A-DNA

The first cyrstalline A-DNA pattern was obtained from NaDNA at low relative humidity (Franklin and Gosling, 1953b). An extensive study by Hamilton et al. (1959) showed that all natural DNAs adopt the A-conformation, irrespective of the base composition. The A conformation was observed with sodium, potassium and rubidium salts, but not with lithium salts. Leslie et al., (1980) reported after an extensive study of synthetic polynucleotides that all synthetic polynucleotides containing guanine could adopt the A-form with the exception of poly d(AG).poly d(CT).

A detailed conformational study was undertaken by Fuller et al. (1965) who reported that A-DNA is a $11_1$ helix with 28.15 $A^o$ pitch. The bases are displaced 4.25 $A^o$ in front of the helix axis. The tilt and twist of the bases are $20^o$ and $-8^o$ respectively. The sugar conformation was anti and had $C_3$'-endo sugar puckering. Arnott and his co-workers reported a computer refined A-structure, and it is a minor variant of the original structure (Arnott et al., 1972b, 1980).

1.2.2.1 B-DNA

At high relative humidity NaDNA exhibited a cross-shape semi-crystalline diffraction pattern (Franklin and Gosling, 1953a). It was this diffraction data that led Watson and Crick to their proposition of a double helical model for DNA. A similar, but highly crystalline pattern was observed for LiDNA and this was used in the detailed structural study

undertaken by Langridge et al. (1960a and b). B-DNA is a $10_1$ helix with a pitch of 33.8 $A^o$. Arnott et al. (1972b) reported a computer-refined model which has an 'anti' sugar conformation with $C_3$'-exo sugar puckering. An equally good model for B-DNA was reported by Arnott et al. (1973) with $C_2$'-endo sugar puckering.

The B-form has been observed for all natural DNAs and synthetic polynucleotides, irrespective of base composition and sequence. The highly crystalline B-form was commonly observed for the Li salt. However, recently the crystalline B-form has been observed from the sodium salt of poly d(AC).poly d(GT) (Leslie et al., 1980) and poly d(AT).poly d(AT) (Mahendrasingam, unpublished).

1.2.2.2 B'-DNA

Arnott et al (1974a) reported that poly d(A).poly d(T) exists in a $10_1$ double-helix with pitch of 32.4 $A^o$. This conformation has been nominated B'-DNA. At low relative humidity it was observed in an orthorhombic unit cell (β-B'-DNA), and high relative humidity it was observed in an hexagonal unit cell (α-B'-DNA). The B'-DNA conformation has also been observed in poly d(I).poly d(C), poly d(AI).poly d(CT) (Leslie et al., 1980) and poly d(AT).poly d(AT) (Mahendrasingam et al., 1983a).

1.2.2.3 B"-DNA

B"-DNA has been observed for poly d(GC).poly d(GC) at low relative humidity (Mahendrasingam et al., 1983b). The intensity distribution of this pattern is quite distinct from B'-DNA or B-DNA and the pitch is 33.4 $A^o$.

1.2.3.1 C-DNA

The C-form was observed for LiDNA at low relative humidity. Marvin et al. (1961) proposed that C-DNA was a $28_3$ helix with pitch of 93.0 $A^o$. It has been observed for the lithium salt of all natural DNAs irrespective of the base composition. For NaDNA, the C-form was regarded as an intermediate state in the A → B transition (Arnott et al., 1975). Recently, Rhodes et al. (1982) showed that the A-conformation is an intermediate state in the C → B transition. The C-conformation was observed for NaDNA at low salt and low relative humidity (Rhodes et al., 1982). The above results were confirmed by synthetic polynucleotide studies on poly d(AC).poly d(GT) (Rhodes et al., 1982) and poly d(AT).poly d(AT) (Mahendrasingam et al., 1983a).

1.2.3.2 C'-DNA

C'-DNA has only been observed for the synthetic polynculeotides poly d(AGC).poly d(GCT), poly d(GGT).poly d(ACC) (Leslie et al., 1980) and poly d(AC).poly d(GT) (Rhodes and Mahendrasingam, unpublished) in the presence of $Li^+$ ions. C'-DNA has $9_1$ helical symmetry with a pitch of 29.5 $A^o$. The crystallinity of these pattersn is much better than C-DNA patterns.

1.2.3.3 C"-DNA

C"-DNA has been observed for Na poly d(AG).poly d(CT) and has helical symmetry of $9_2$ with pitch of 58.2 $A^o$ (i.e. 9 dinucleotide pairs per pitch).

1.2.4.1 D-DNA

D-DNA patterns have been reported for Na poly d(AT).poly d(AT) by Davies et al. (1963), Arnott et al. (1974b) and Mahendrasingam et al. (1983a). A similar pattern was also reported for poly d(IC).poly d(IC)

(Mitsui et al., 1970 and Ramaswamy et al., 1982). The pattern was indexed on a tetragonal lattice and called α-D-DNA. Selsing et al. (1975) reported a similar conformation for poly d(ATT).poly d(AAT), but the molecules were arranged in the hexagonal unit cell (β-D-DNA). Further, they claimed that the molecular conformation is the same in α-D-DNA and β-D-DNA. It was observed that at low relative humidity, K poly d(AT).poly d(AT) fibres gave α-D-DNA and reversibly changed into β-D-DNA at 92% relative humidity (Mahendrasingam, unpublished). The D-form has also been observed in a hexagonal lattice for the polynucleotides poly d(AC).poly d(IT), poly d(AIT).poly d(ACT) and poly d(AIC).poly d(ICT).

1.2.4.2 D'-DNA

D'-DNA has been observed for high salt Cs poly d(AT).poly d(AT) with pitch of 24.0 $A^0$ (Mahendrasingam, unpublished). The helical symmetry of this conformation is $4_1$ with a dinucleotide as the repeating unit.

1.2.5 E-DNA

This conformation has been observed for poly d(IIT).poly d(ACC) by Leslie et al., (1980). The molecular asymmetric unit appears to contain five nucleotides with a mean rise per residue and rotation per residue per asymmetric unit of 3.25 $A^0$ and $48^0$ respectively. Hence the structure is approximately a $15_2$ helix.

1.2.6 S-DNA

The S-conformation has been observed for poly d(GC).poly d(GC) and poly d(AC).poly d(GT) by Arnott et al. (1980). It has a helical symmetry of $6_5$ with a pitch of 43.5 $A^0$, and the repeating unit is a dinucleotide. The molecules are packed in a hexagonal unit cell.

## 1.3 The Objective of the Present Project

The objective of this work is to investigate the effect of the following factors on DNA polymorphism:

(a) base composition
(b) base modification
(c) base sequence
(d) 'type' and 'concentration' of prevailing ions
(e) relative humidity of the fibre environment.

Further, an attempt has been made to obtain answers to the following questions:-

(i) What is the relationship between DNA polymorphism, salt concentration and relative humidity of the fibre environment?

(ii) How does the rate of change of the environmental relative humidity affect DNA polymorphism?

(iii) Do left-handed structures exist in DNA fibres?

Experimental techniques used in this study are discussed in Chapter 2. In view of the recent developments in our knowledge of DNA, it was decided to re-examine the assumed molecular structure of D-DNA. This is discussed in Chapter 4. An extensive study of the polymorphism of poly d(AT).poly d(AT), is discussed in Chapter 5. Mixtures of (A/C) and (A/B) conformations, as observed in natural and synthetic polynucleotides are discussed in Chapter 6. The possibility of an alternating molecular structure in various co-polymers is discussed in Chapter 7. An extensive study of poly d(GC).poly d(GC) and the possible left-handed models for S-DNA are discussed in Chapter 8. An extensive study of natural DNAs and the conformational transitions between the A, B and C forms are discussed in Chapter 9. The possibility of a heteronomous structure (Arnott et al., 1983) in homopolymers is discussed in Chapter 10. A polymorphic study of poly d(AC).poly d(GT) and the similarities it shows

with poly d(AT).poly d(AT), and poly d(GC).poly d(GC) are discussed in Chapter 11.

Finally, the last chapter concludes the thesis with a summary of the work undertaken and a few suggestions for further work.

# CHAPTER 2

## Materials and Methods

### 2.1 Purification of DNA

Both natural and synthetic DNAs have been obtained from a variety of sources. It is desirable to subject them to a purification process in order to extract RNA, protein and excess inorganic salt. This can be achieved either by precipitation or centrifugation. In some cases phenol extraction (Massie and Zimm, 1965) was performed in order to remove proteins from samples of DNA.

In the purification stage it is desirable to use low salt concentrations since DNA dissolves very slowly if the ionic strength is high. Nevertheless, DNA is denatured if the salt concentration is too low. Typical salt concentrations of such solutions were $\sim$ 2 mM, in which the concentration of DNA was approximately 1 mg/mℓ.

Precipitation of DNA from its solution was achieved by adding cold ethanol and salt solution (concentration $\sim$ 0.1 M). This method, however, has the following disadvantages:-

1. The amount of salt brought down with the DNA is difficult to control and estimation of salt concentration in fibres is difficult.
2. To increase the yield of DNA precipitation, concentrated salt solutions have to be used. As a result, fibres made by this method contain more salt than may be desired.
3. The precipitation method requires relatively large amounts of DNA. Whilst such quantities are readily available for native DNAs, synthetic forms are far more expensive and are supplied in far smaller quantities.

## 2.2 Extraction of DNA from solutions and ion exchange by precipitation

Most DNAs (natural and synthetic) obtained from commercial sources are precipitated from NaCℓ solutions. The exchange of sodium with another ion was accomplished in the following way:

(a) The DNA was dissolved in ∿ 2 mM NaCℓ solution.

(b) The salt solution containing the new ion (0.1 M) was added, along with cold ethanol (∿ $4^{o}$ C) to the above solution. The DNA precipitate was collected on a glass rod.

(c) The collected precipitate was redissolved in a 2 mM solution of the new salt and dialysed against a 0.1 M solution of the same salt.

(d) The DNA was finally precipitated using cold ethanol and salt solution.

(e) The purity of the material thus obtained was tested in a flame emission spectrometer.

## 2.3 Extraction of DNA from solutions and ion exchange by centrifugation

Having a large molecular weight, DNA can be extracted from solution by centrifugation. The method can also be used to change the excess ion content of DNA samples. The procedure adopted was:-

(a) DNA was dissolved in ∿ 2 mM NaCℓ solution or in deionized water.

(b) The solution was centrifuged for 12 hours at 50,000 r.p.m. During this process, the heavy DNA was brought down to the bottom of the tube in the form of a gel. A lot of the excess salt and low molecular weight impurity was left in the supernatant.

(c) The supernatant was decanted with care and some of the residual gel was used to make fibres.

(d) The salt concentration of the gel could, if necessary, be raised

by redissolving the gel in an appropriate salt solution. Subsequent centrifugation enabled re-extraction of the gel.

(e) The salt concentration of the gel could, if necessary be lowered by redissolving the gel in deionized water and by its subsequent centrifugation.

(f) Further ion exchange, if necessary, was accomplished by dialysis involving the appropriate solutions. Centrifugation was again used to extract the final gel.

This method enabled a variety of fibres having different salt contents to be made. The amount of DNA in the gel was determined by measurement of the optical density at 260 nm for the original solution and the supernatant. Subsequent calculations yielded a fair idea of the salt conditions inside the fibres made. Such calculations are described below.

2.4 Estimation of salt content in fibres

1. Before centrifugation, the UV absorption at 260 nm was measured. The concentration of DNA (say $C_1$ mM) was calculated from the following equation:-

$$\text{Concentration of DNA (in mM)} = \frac{\text{AMP x SCA x DIF}}{\text{ABCF x M}}$$

AMP - Amplitude, SCA - Scale, DIF - Dilution Factor, ABCF - Absorbance Coefficient, M - Molecular weight of nucleotide

It is important to note that DNA concentration implies the nucleotide molar concentration ($PO_4^-$ molar concentration). The above measurement was repeated for the supernatant: thus the DNA concentration in the supernatant was calculated (say $C_2$ mM).

2. The volumes of the starting solution (say $V_1$ mℓ) and the supernatant (say $V_2$ mℓ) were measured.

3. The amount of DNA in the gel was estimated as follows: the amount of DNA in the original solution is $C_1V_1$, the amount of DNA left in the supernatant is $C_2V_2$, and therefore the amount of DNA in the gel is $C_1V_1 - C_2V_2$.

4. The amount of salt in the gel was estimated by assuming that the salt is distributed uniformly in the solution before and after centrifugation.

Let the initial concentration of salt in the DNA solution be X mM.

Hence the amount of salt (salt per $PO_4^-$) in the gel is:

$$= \frac{X}{1000} \times \frac{(V_1 - V_2)}{(C_1V_1 - C_2V_2)}$$

The quantity X was determined either from original dialysis data or by use of a flame emission spectrometer. The use of the flame emission spectrometer was often made difficult by the non-linear response of the device over a wide range of concentrations. Two different calibration curves had to be constructed. These calibration curves are shown in figures 1 and 2.

## 2.5 The Fibre Preparation

The fibres were prepared in the manner described by Fuller et al. (1967). Two glass rods of diameter $\sim$ 200 micro-metres are mounted using plasticine inside a circular frame as shown in figure 3. (The distance between the glass rods can be adjusted). A drop of DNA gel is then placed between the two glass rods, and allowed to dry to form a fibre. When the fibre had dried down to a slightly larger diameter than that of the glass rod ($\sim$ 300 micro-metres), tension was applied to the specimen by increasing the separation between the glass rods. The final diameter of fibres was $\sim$ 150 micro-metres. Birefringence was measured with Zeiss rotary compensator

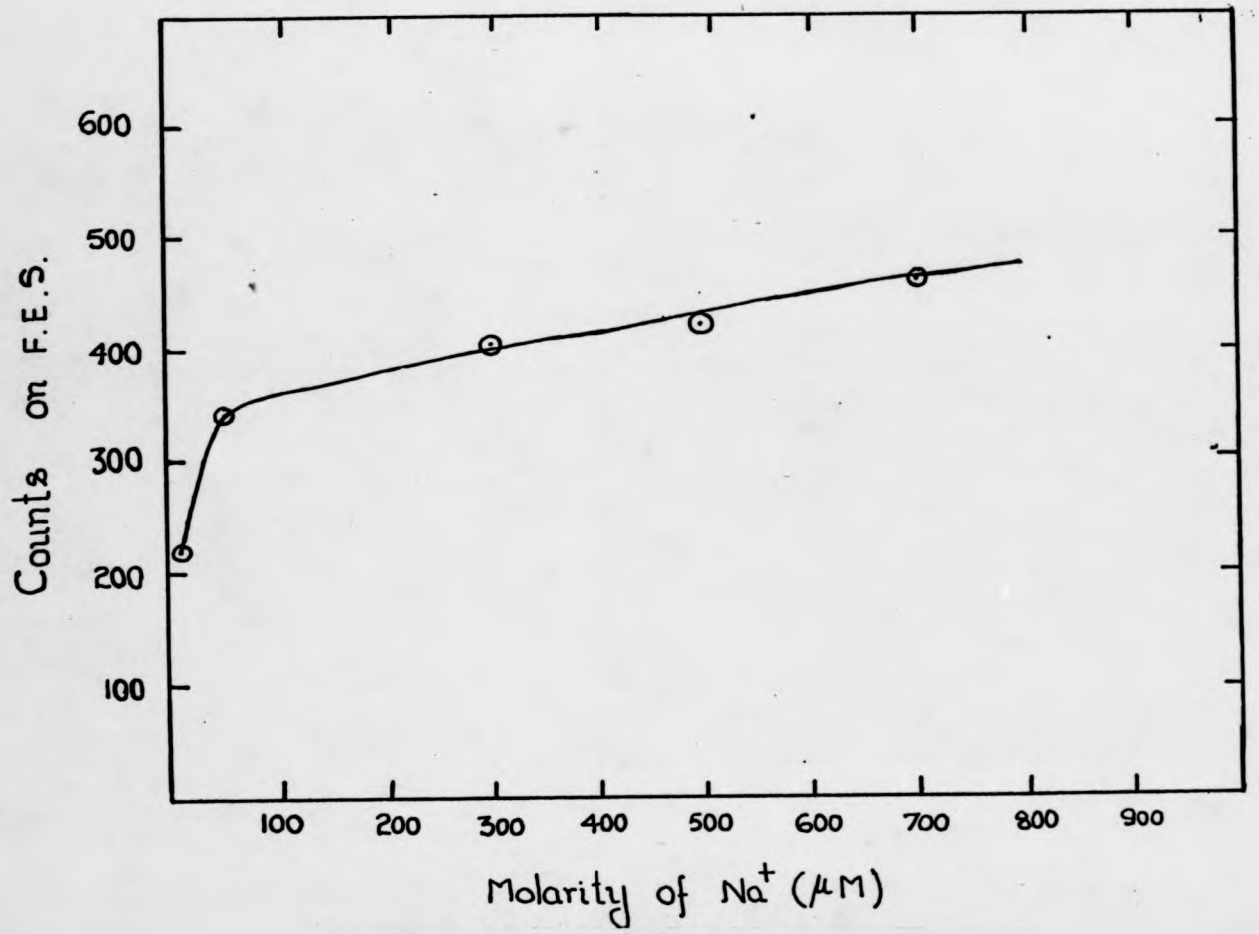



Figure 1 Typical Calibration Curve for the f.e. spectrometer (for lower range)

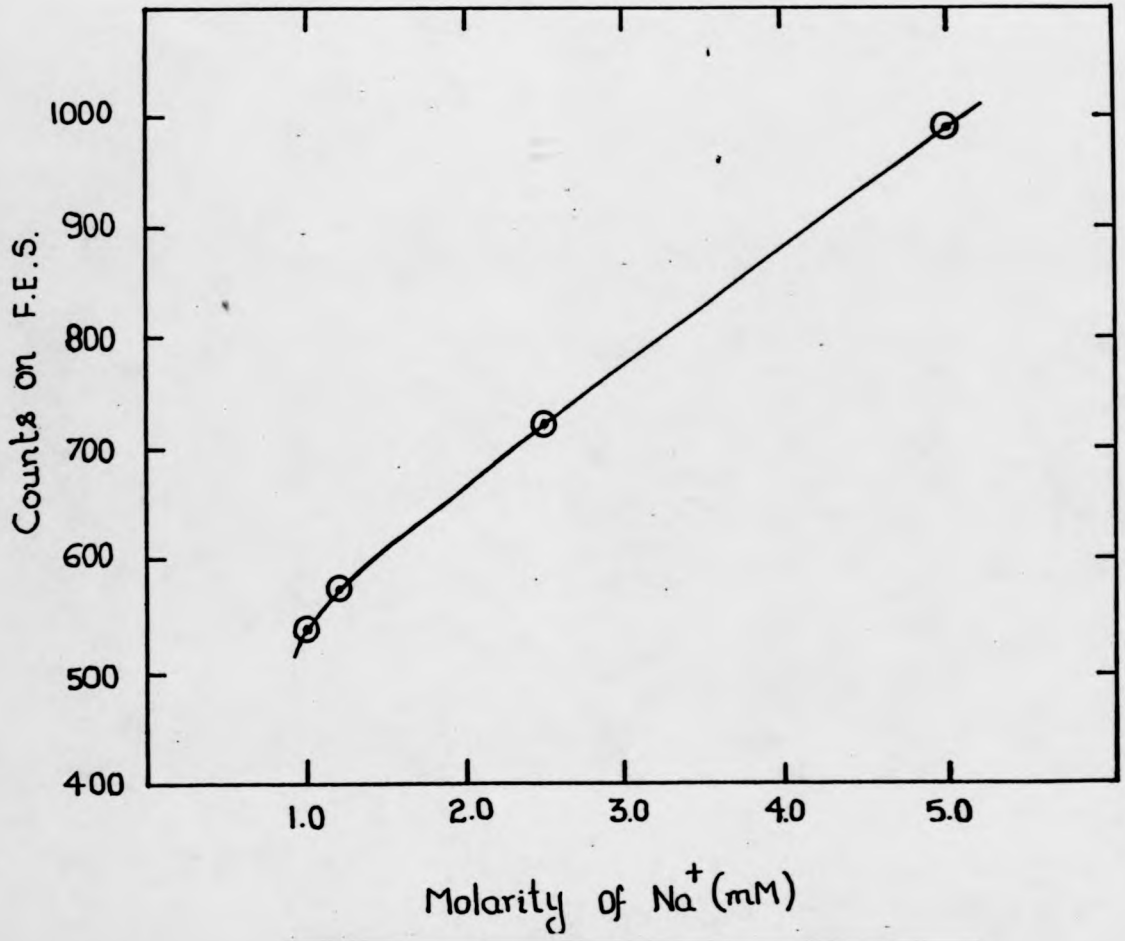



Figure 2 Typical Calibration Curve for the f.e. Spectrometer (for higher range)

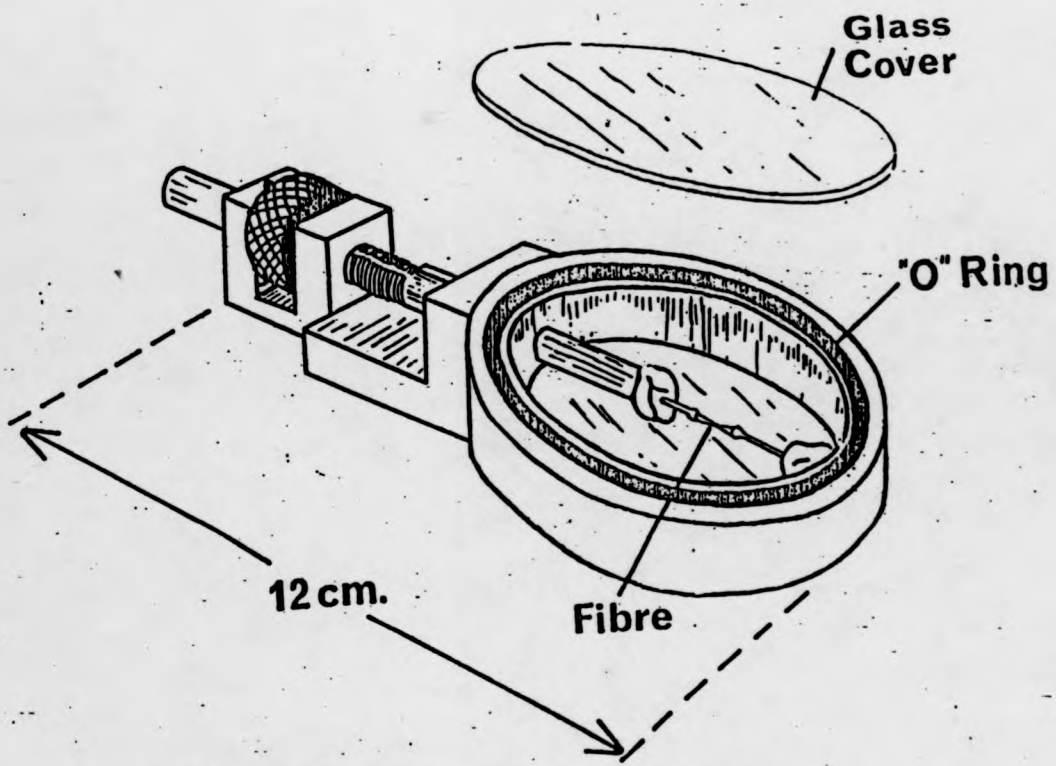



Figure 3 Fibre Pulling Cell (from Fuller et al., 1967)

in order to determine the degree of orientation in DNA fibres before taking X-ray diffraction photographs.

## 2.6 Obtaining the X-ray Fibre Diffraction Photographs

Pinhole and Searle cameras were used to obtain diffraction photographs. Pinhole cameras with a set of three gold apertures to collimate the X-ray beam, were made in the workshop of the Department of Physics, University of Keele and were similar in design to those described by Chesley (1947), and Langridge et al. (1960). The specimen to film distance in this camera was about 3 cm.

The Searle camera was operated using toroidal optics (Elliot, 1965) for the majority of the fibres. Franks optics (Frank, 1958) were used specially for thin fibres. The exposure time for the Frank's optics is larger than that for the toroidal optics. The specimen to film distance of about 4 cm was used in the Searle camera. All diffraction photographs were taken using nickel filtered Cu $K_\alpha$ radiation ($\lambda$ = 1.5418 $A^o$) generated either by a Hilger and Watts microfocus or by an Elliot Gx6 rotating anode generator. The Hilger and Watts machines were operated at 35 kv and at a tube current of 3.0 mA. The Elliot Gx6 was operated at 35 kv and at a tube current of 56 mA.

Alignment of fibres in pinhole camers was performed on an optical bench. The Searle camera was aligned in the manner described in the relevant manuals.

Before starting exposures cameras were flushed out with helium gas to eliminate air scatter. The relative humidity of the fibre environment was controlled by passing this helium through a saturated solution of a particular salt. The salts used and the relative humidity associated with their saturated solutions (at $20^oC$) are listed on the following page (O'Brien, 1948).

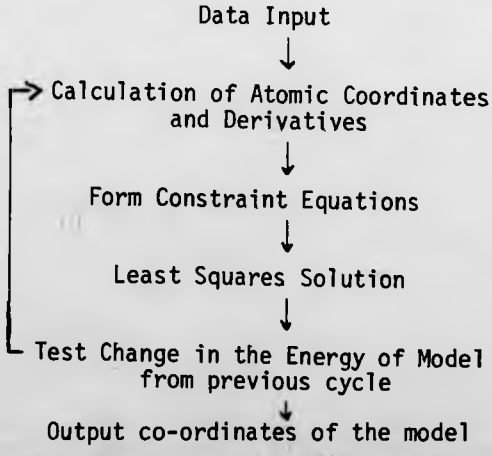
| Salt | R.H. (%) |
|---|---|
| Calcium Chloride | 33 |
| Potassium Carbonate | 44 |
| Sodium Bromide | 57 |
| Sodium Nitrite | 66 |
| Sodium Chlorate | 75 |
| Potassium Chloride | 86 |
| Sodium Tartrate | 92 |
| Sodium Sulphite | 95 |
| Potassium Chlorate | 98 |

## 2.7 Measurements of X-ray Fibre Diffraction Patterns

X-ray diffraction patterns were measured in order to obtain reciprocal space co-ordinates and diffracted intensities of the observed reflections. This is discussed in sections 3.2.3 to 3.2.5.

## 2.8 Computer Programs

The model building program described by Pigram (1968), and later modified by Goodwin (1977), was used in the present work. This program was executed on the CDC 7600 machine at the University of Manchester regional computer centre. The flow diagram of the model building program is shown below.



Other software was written on a variety of machines (ITT 2020, PDP 11/23, GEC 4190), but throughout the project, some effort was put into maintaining a 'user friendly' collection of packages. 'Fortran 66' was largely used because of its 'machine independence'. Programs are now being updated to 'Fortran 77'. Programs were written to calculate:-

(a) Cylindrically averaged squared fourier transform (CAST.FOR).

(b) Structure factors from an array of helical molecules (SF.FOR).

(c) Interhelical contact distances (IHC.FOR).

# CHAPTER 3

## X-ray Diffraction as a Means of Structure Determination

### 3.1 General Theory of X-ray Diffraction

#### 3.1.1 Introduction

X-ray diffraction is the most powerful technique currently available for study of the structure of large molecules, and has been used to establish many features pertaining to living systems the knowledge of which we now take for granted. These sort of advances could not have been made without crucial evidence obtained from X-ray crystallography. In many cases, X-ray diffraction studies on protein or nucleic acid crystals have yielded complete tertiary structures to 3 $A^{O}$ resolution or better. Even if the sample is not a very ordered one (such as an 'oriented' fibre), X-ray diffraction still provides a wealth of structural information. Though sometimes insufficient to determine a structure uniquely, the information in many cases can be used to provide decisive tests of structural models, but there are some limitations and shortcomings of the method. Normally, atoms are treated as motionless, even though in crystals there is appreciable motion at finite temperature. The crystals may actually be ordered only in local domains, but they are treated as perfectly ordered arrays. X-radiation is treated as monochromatic, even though a certain bandwidth is inevitable. The diffraction data are considered very precise even though experimental error is often a significant problem in reality.

#### 3.1.2 Theory of X-ray Scattering

In figure 1, $\underline{s}_0$ represents the incident beam, and $\underline{s}_1$ the diffracted beam. Two scattering centres P and Q are separated by a vector r. The

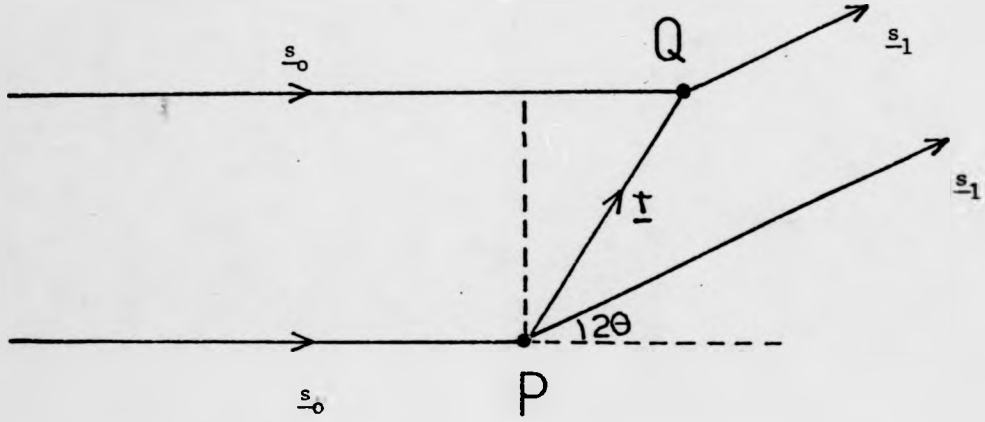



Figure 1 The Geometry of the Incident and Scattered Waves

path difference = $\underline{r}.\underline{s}_1 - \underline{r}.\underline{s}_0 = \underline{r}.(\underline{s}_1 - \underline{s}_0)$. If we consider elastic scattering of the X-rays only, the wavelength of the incident and scattered radiation is the same. Let us define a new vector $\underline{s}$ called the scattering vector, such that

$$\underline{s} = \frac{\underline{s}_1}{\lambda} - \frac{\underline{s}_0}{\lambda} \quad \text{and,} \quad |\underline{s}| = \frac{2}{\lambda} \sin\theta \qquad (1)$$

The vector $\underline{s}$ is defined in terms of reciprocal distance units, and corresponds to the reciprocal space vector necessary for Bragg reflection to occur. In figure 1, the incident beam $\underline{s}_0$ is scattered by an electron at P and by another one at Q such that P and Q are separated by $\underline{r}$. The path difference is $(\underline{r}.\underline{s}_0 - \underline{r}.\underline{s}_1) = \underline{r}.\underline{s}$. Thus if the radiation scattered by an electron at the origin is $E(\underline{s})$, moving the electron from the origin to a position $\underline{r}$ simply causes a phase shift of $\underline{s}.\underline{r}$. The scattered radiation is: $E(\underline{s}) \exp(2\pi i \underline{s}.\underline{r})$. In general, because electrons are not localised it is better to describe an electron density $\rho(\underline{r})$ in a volume element $d\underline{r}$ located at $\underline{r}$. The scattering is proportional to $\rho(\underline{r})\, d\underline{r}$. For a continuous electron density at position $\underline{r}$, the scattered radiation is described by:-

$$dE(\underline{s}) = \rho(\underline{r}) \exp(2\pi i \underline{s}.\underline{r})\, d\underline{r} \qquad (2a)$$

where $\rho(\underline{r})\, d\underline{r}$ is the number of electrons in the volume element. For a continous electron distribution the total scattered radiation is given by following equation:

$$E(\underline{s}) = \int_{\text{volume}} \rho(\underline{r}) \exp(2\pi i \underline{r}.\underline{s})\, d\underline{r} \qquad (2b)$$

The physical meaning of this equation is that the scattered radiation is the Fourier transform of the object. The Fourier transform of $\rho(\underline{r})$ is $E(\underline{s})$ and therefore the inverse Fourier transform gives the electron distribution function, $\rho(\underline{r})$, so that:

$$\rho(\underline{r}) = \int E(\underline{s}) \exp(-2\pi i\, \underline{s}.\underline{r})\, d\underline{s} \qquad (3)$$

The detailed pattern of electron density around an individual atom depends on the prevailing bonding conditions. However, in almost all X-ray diffraction experiments, the resolution is not high enough to detect this detailed electron distribution. Thus it is a good approximation to model the electron distribution of an atom as spherically symmetric. Then $\rho(\underline{r})$ becomes $\rho(r)$. If we express equation 2(b) in spherical polar coordinates (see figure 2)

$$E(s) = \int_0^{2\pi} d\phi \int_0^{\pi} \sin\theta\, d\theta \int_0^{rmax} \rho(r)\, r^2\, e^{2\pi i \underline{s}.\underline{r}}\, dr$$

$$= 2\pi \int_0^{rmax} \rho(r)\, r^2\, dr \int_0^{\pi} \sin\theta\, e^{2\pi i rs\cos\theta}\, d\theta$$

$$E(s) = 4\pi \int_0^{rmax} \rho(r)\, r^2\, dr \left| \frac{\mathrm{Sin}\,(2\pi rs)}{(2\pi rs)} \right| \equiv f(s) \qquad (4)$$

The function f(s) is defined as the atomic scattering factor (see figure 3). The atomic scattering factor given by equation (4) can be evaluated provided that the charge density $\rho(r)$ of the atom is known. Equation (4) is strictly correct only when the charge distribution is spherically symmetric and provided that the energy of the X-rays is not close to an absorption edge of the atom. A brief discussion of the procedures employed in evaluating scattering factors, along with complete tables of scattering factor values currently believed to be most reliable can be found in volume 3 of the International Tables for X-ray Crystallography.

Consider an atom position at $\underline{r}_n$ with respect to the coordinate

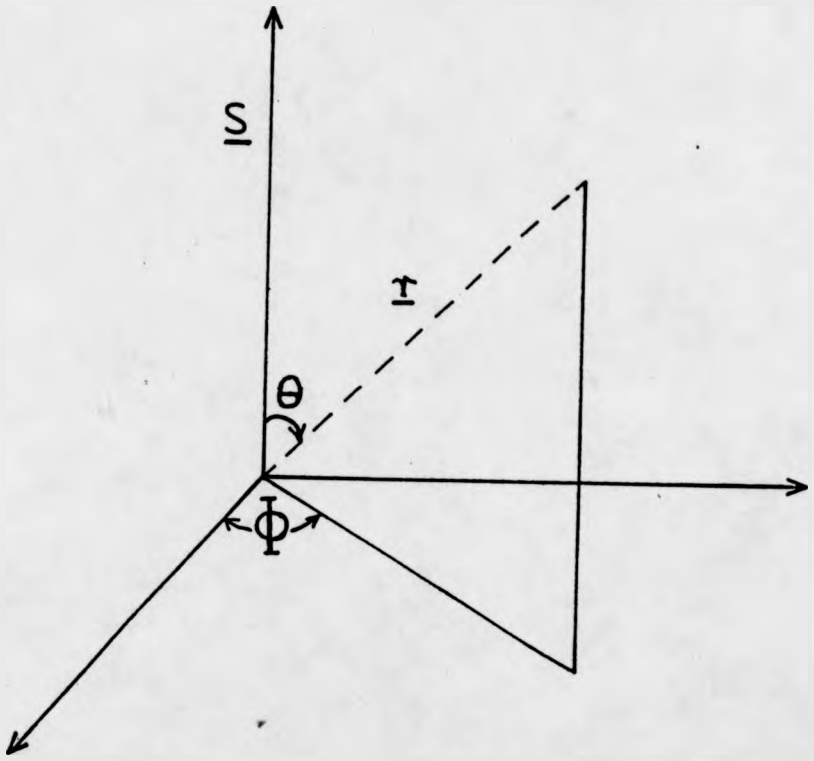



**Figure 2** Spherical polar coordinates in relation to an orthogonal cartesian system

system origin (see figure 4). Then the scattered radiation is given by:

$$E(s) = \int d(\underline{R} + \underline{r}_n) \cdot \rho(\underline{R} + \underline{r}_n)\, e^{2\pi i \underline{s}.(\underline{R}+\underline{r}_n)}$$

where $\underline{r}_n$ is a constant, since $d(\underline{R} + \underline{r}_n) = d\underline{R}\, e^{2\pi i \underline{s}.(\underline{R}+\underline{r}_n)}$ can be removed from the integral.

$$E(s) = e^{2\pi i \underline{s}.\underline{r}_n} \int dR \rho(\underline{R} + \underline{r}_n)\, e^{2\pi i \underline{s}.\underline{R}}$$

$$E(s) = f(\underline{s})\, e^{2\pi i \underline{s}.\underline{r}_n} \qquad (5)$$

where $f(s) = \int d\underline{R} \rho(\underline{R}+\underline{r}_n)\, e^{2\pi i \underline{s}.\underline{R}}$

is the atomic scattering factor. If N atoms are located at positions $\underline{r}_n$ with atomic scattering factors $f_n$, the structure factor is

$$F(\underline{s}) = \sum_{n=1}^{N} f_n(\underline{s})\, e^{2\pi i \underline{s}.\underline{r}_n} \qquad (6)$$

If the sample is a centre of symmetry and the centre is placed at the origin, equation (6) can be written as:

$$F(s) = \sum_{n=1}^{N/2} f_n(\mathbf{s}) \cos(2\pi \underline{s}.\underline{r}_n) \qquad (7)$$

3.1.3 X-ray Diffraction from One Dimensional Array of Atoms

Assume that we have a regular array of 2N+1 atoms and their translation vector is $\underline{a}$ so the position of the $n^{th}$ atom is $n\underline{a}$. The structure factor for the $n^{th}$ atom is

$$F_n(s) = f(s)\, e^{2\pi i n \underline{s}.\underline{a}}$$

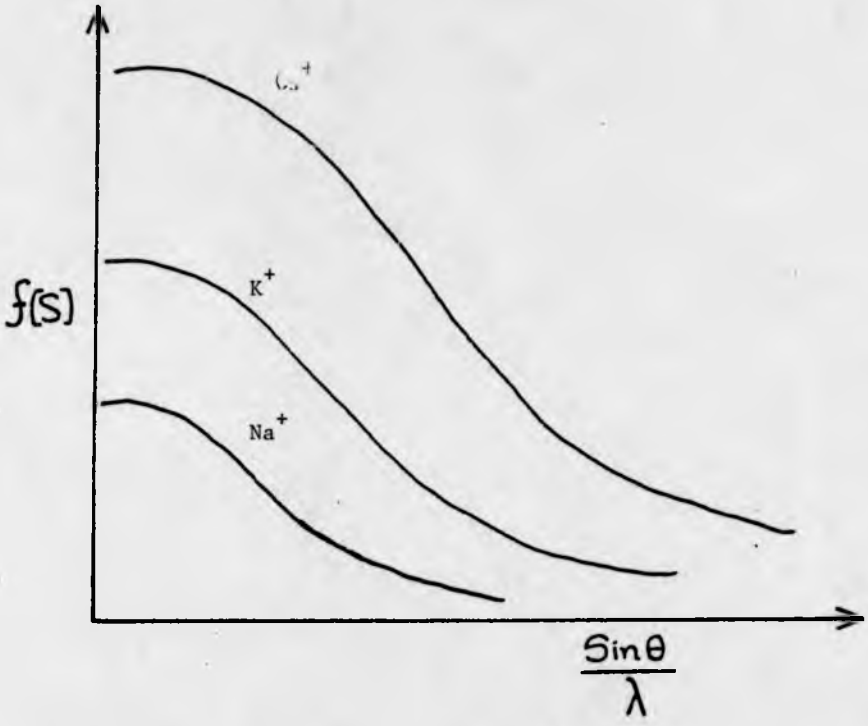



Figure 3 Typical Atomic Scattering Factor Curves

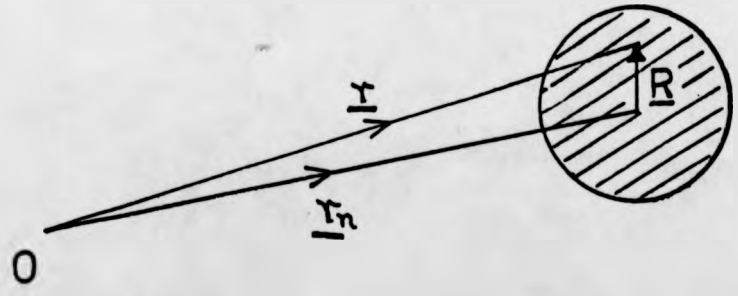



Figure 4 An Atom Position at $\underline{r}_n$ with respect to the Coordinate System Origin

The structure factor of the whole array can be written as:

$$F_{tot}(s) = f(s) \sum_{n=-N}^{N} e^{2\pi i n \underline{s}.\underline{a}}$$

Summation of this geometric series can be written as:

$$F_{tot}(s) = f(s) \frac{e^{2\pi i n \underline{s}.\underline{a}}(1-e^{2\pi(2N+1)\underline{s}.\underline{a}})}{1 - e^{2\pi i \underline{s}.\underline{a}}}$$

This equation can be simplified to:-

$$F_{tot}(s) = f(s) \frac{\sin ( (2N+1) \pi \underline{s}.\underline{a})}{\sin (\pi \underline{s}.\underline{a})}$$

The intensity of the scattering from the array will be:-

$$I_{tot}(s) = \left|F_{tot}(s)\right|^2 = f(s)^2 \left| \frac{\sin( (2N+1) \pi \underline{s}.\underline{a})}{\sin(\pi \underline{s}.\underline{a})} \right|^2 \qquad (8)$$

If N is large the structure factor becomes significant only when $\sin(\pi \underline{s}.\underline{a})$ approaches to zero. This may be expressed analytically by the Laue equation:-

$$\underline{s}.\underline{a} = h \quad \text{where } h = 0, \pm 1, \pm 2 \qquad (9)$$

3.1.4 X-ray Diffraction From a Three Dimensional Array of Atoms

The periodicity of a three dimensional array is defined by three vectors $\underline{a}$, $\underline{b}$ and $\underline{c}$. In general, $\underline{a}$, $\underline{b}$ and $\underline{c}$ are not orthogonal nor are they of equal length. The periodicity along $\underline{a}$ will cause a set of scattering fringes at $\underline{s}.\underline{a} = h$. The additional periodicity along $\underline{b}$ and $\underline{c}$ causes similar sets of fringes defined by $\underline{s}.\underline{b} = k$ and $\underline{s}.\underline{c} = \ell$ where k, $\ell$ intersects the parallel lines generated by $\underline{s}.\underline{a} = h$ and $\underline{s}.\underline{b} = k$. The result is a three dimensional lattice with points spaced evenly by $\frac{1}{\underline{a}}$ in the direction parallel to $\underline{a}$, by $\frac{1}{\underline{b}}$ parallel to $\underline{b}$ and by $\frac{1}{\underline{c}}$ parallel to $\underline{c}$.

The lattice which describes the allowed scattering geometries is not the same lattice of points that represents the position of the atoms in the array. The position lattice has unit vectors $\underline{a}$, $\underline{b}$ and $\underline{c}$ where as the spacings in the diffraction lattice a*, b* and c* are related to the inverse of these. This scattering lattice is called "the reciprocal lattice".

### 3.1.5 The Relationship Between the Crystal Lattice and Reciprocal Lattice

Real crystals are three-dimensional. The reciprocal lattice as seen in an X-ray diffraction pattern is also three-dimensional. It is related in a simple way to the actual crystal lattice. By measuring the spatial position of the diffracted spots, it is possible to compute the cell dimensions and the shape of the reciprocal lattice. From this, the corresponding dimensions and shape of the unit cell of the actual crystal lattice can be derived. We shall now solve the von Laue equations,

$$\underline{s}.\underline{a} = h,\ \underline{s}.\underline{b} = k,\ \underline{s}.\underline{c} = \ell \qquad (10)$$

for scattering vector $\underline{s}$ which satisfies all three equations. Assume now that the vector $\underline{s}$ takes the form $\underline{s} = \chi(H\underline{a}^* + K\underline{b}^* + L\underline{c}^*)$ where $\chi$ is an arbitrary scaler and H, K and L are undetermined constants. Then:-

$$\underline{s}.\underline{a} = \chi(H\underline{a}^* + K\underline{b}^* + L\underline{c}^*) \,.\, \underline{a} = h \qquad (11)$$

$$\chi\, H\underline{a}^*.\underline{a} + \chi K\underline{b}^*.a + \chi L\underline{c}^*.a = h \qquad (12)$$

$$\chi\, H\underline{a}^*.\underline{b} + \chi K\underline{b}^*.\underline{b} + \chi L\underline{c}^*.\underline{b} = k \qquad (13)$$

and

$$\chi\, H\underline{a}^*.\underline{c} + \chi K\underline{b}^*.\underline{c} + \chi L\underline{c}^*.\underline{c} = \ell \qquad (14)$$

where a*, b*, c*, $\chi$, H, K and L are constants, and may be chosen as found appropriate for the solution of equations 12, 13 and 14. Let $\chi$=1, this leads 12, 13 and 14 to:-

$$\underline{Ha}^*.\underline{a} + \underline{Kb}^*.\underline{a} + \underline{Lc}^*.\underline{a} = h \tag{15}$$

$$\underline{Ha}^*.\underline{b} + \underline{Kb}^*.\underline{b} + \underline{Lc}^*.\underline{b} = k \tag{16}$$

$$\underline{Ha}^*.\underline{c} + \underline{Kb}^*.\underline{c} + \underline{Lc}^*.\underline{c} = \ell \tag{17}$$

The equations have a solution if we select a*, b* and c* such that

$$\left.\begin{array}{lll} \underline{a}^*.\underline{a} = 1 & \underline{a}^*.\underline{b} = 0 & \underline{a}^*.\underline{c} = 0 \\ \underline{b}^*.\underline{a} = 0 & \underline{b}^*.\underline{b} = 1 & \underline{b}^*.\underline{c} = 0 \\ \underline{c}^*.\underline{a} = 0 & \underline{c}^*.\underline{b} = 0 & \underline{c}^*.\underline{c} = 1 \end{array}\right\} \tag{18}$$

Therefore H = h, K = k, L = $\ell$ and the scattering vector

$$\underline{s} = h\underline{a}^* + k\underline{b}^* + \ell\underline{c}^*$$

satisfies the Laue equations on condition that equations(18) hold if h, k and $\ell$ are integers then this term will generate a lattice, which we have called the recriprocal lattice. In equations (18) $\underline{a}^*.\underline{a} = 1$, $\underline{a}^*.\underline{b} = 0$ and $\underline{a}^*.\underline{c} = 0$. Therefore a* must be orthogonal to $\underline{b}$ and $\underline{c}$ simultaneously and hence a* can be written as:-

$$\underline{a}^* = \xi(\underline{b}\times\underline{c})$$

$$\underline{a}^*.\underline{a} = \xi\underline{a}.(\underline{b}\times\underline{c}) = 1$$

$$\xi = \frac{1}{\underline{a}.(\underline{b}\times\underline{c})} \tag{19}$$

Therefore:

$$\underline{a}^* = \frac{\underline{b}\times\underline{c}}{\underline{a}.(\underline{b}\times\underline{c})} \tag{20}$$

similarly b* and c* can be written as:-

$$\underline{b}^* = \frac{\underline{c}\times\underline{a}}{\underline{a}.(\underline{b}\times\underline{c})} \tag{21}$$

$$\underline{c}^* = \frac{\underline{a}\times\underline{b}}{\underline{a}.(\underline{b}\times\underline{c})} \tag{22}$$

where $\underline{a}.(\underline{b}x\underline{c})$ is the volume of the unit cell in the real space; $v=\underline{a}.(\underline{b}x\underline{c})$. The reciprocal lattice volume is $v^* = a^*.(\underline{b}^*.\underline{c}^*)$. It can be shown that $v^* = \frac{1}{v}$ by using the equations 20, 21 and 22.

### 3.1.6 The Structure Factor

It is convenient to write out the unit-cell structure factor, $F_m(s)$, explicitly in terms of the positions of each atom in the unit cell, and of the corresponding atomic scattering factors. Using equation (2) for the molecular structure factor, we choose a co-ordinate system based on the unit-cell vectors $\underline{a}$, $\underline{b}$ and $\underline{c}$. The position of the $j^{th}$ atom in the unit cell is then $\underline{r}_j = x_j\,\underline{a} + y_j\underline{b} + z_j\underline{c}$ where $x_j$, $y_j$ and $z_j$ are now fractions of the corresponding unit cell dimensions. Then equation (2) becomes

$$F_m(s) = \sum_j f_j(s)\, e^{2\pi i(x_j\underline{s}.\underline{a} + y_j\underline{s}.\underline{b} + z_j\underline{s}.\underline{c})} \qquad (23)$$

where the sum is taken over all the atoms in the unit cell. However, $F_m(s)$ can be sampled only at geometries allowed by the von Laue conditions: when we apply these, equation (23) simplifies to:-

$$F_m(h,k,\ell) = \sum_j f_j(s)\, \exp|2\pi i(hx_j + ky_j + \ell z_j)| \qquad (24)$$

This equation is called the structure factor equation. It represents the unit cell X-ray scattering sampled at the reciprocal lattice points defined by the Miller indices, $(hk\ell)$.

## 3.2 Theory of X-ray Fibre Diffraction

### 3.2.1 Introduction

Fibrous proteins and DNA, are long rod-shaped molecules and frequently reluctant to crystalize into perfect single crystals. Concentrated gels can, however, be used to form partially oriented fibres by precipitation

of such samples. These long, rod-like molecules can pack together with their axes parallel, or nearly parallel to the fibre axes. The degree of order within the fibres may vary considerably. In some fibres the molecules are regularly oriented about the fibre axes. The diffraction pattern from such fibres are similar to single crystal rotation photographs, although spots may be drawn out into arcs due to the disorientation, and may be rather broad if the crystalline regions are small. The crystallinity of the fibres may be divided into three major categories: crystalline, semi-crystalline and non-crystalline. To explain the types of crystallinity the molecular spatial arrangement is shown schematically in figure 5. $\theta$ is the orientation angle of the molecule with respect to the fibre axes, $\phi$ angle of rotation about fibre axes, and z is the displacement along the fibre axes.

(a) A Crystalline Fibre

Each molecule has the same orientation angle $\theta$. The fibre consists of individual microcrystals also ordered along z and $\phi$, but are, packed in a random way about $\phi$(sometimes also about z).

(b) A Semi-crystalline Fibre

Each molecule has the same orientation angle $\theta$, but there are no microcrystals. There is a fairly regular lattice perpendicular to z, but individuals disordered in the z direction as well as the $\phi$ direction. If molecules are randomly displaced relative to each other in the direction of the fibre axes, discrete spots only are observed along the equator, and the higher layer-lines have a continuous distribution of intensity along them. A random displacement by a fixed amount can give rise to discrete spots on some layerlines and continuous streaks of intensity along others. Thus, random displacement by half the repeat period along the fibre axis will produce spots on the even numbered layerlines, and streaks on the odd ones.

(c) A Non-crystalline Fibre

There is no regular lattice in the fibre. The only order is that long axes of the molecules share a general common direction.

3.2.2 Diffraction by Helical Structures

Helical diffraction theory was first developed by Cochran, Crick and Vand (1952) and Stokes (unpublished). Further detailed analysis was done by Klug et al., (1958) and Ramachandran (1960). For the DNA double helix, the theory was described by Langridge et al., (1960b).

If there are j atoms in the repeating unit of a helical molecule, the structure factor equation can be written thus:-

$$F(\ell, \psi, \xi) = \sum_n \sum_j f_j J_n(2\pi R_j \xi) \exp i\{n(\psi - \phi_j + \pi/2) + \frac{2\pi \ell z_j}{c}\}$$

where $\ell$ = layerline number

$\psi$ = angular co-ordinate of a lattice point in reciprocal space

$\xi$ = radial co-ordinate of a lattice point in reciprocal space

$R_j, \phi_j, z_j$ = real space cylindrical polar co-ordinates of the $j^{th}$ atom

$f_j$ = scattering factor of the $j^{th}$ atom

c = length of the unit cell axis

$J_n$ = Bessel function of the first kind of order n, where n is defined as the integral solution of:-

$$\frac{n}{P} = \frac{\ell}{c} - \frac{m}{p}$$

P = pitch of the helix

p = axial separation of the repeating units and m can take any integral value:

$\frac{c}{P}$ and $\frac{c}{p}$ must be integral because c is the crystallographic repeat, but $\frac{p}{P}$ is not necessarily integral; it can be written as $\frac{c}{P} = K$ and $\frac{c}{p} = N$ we have

$$n = (\ell - mN)/K \qquad (26)$$

For an integral helix K=1, and for a non integral helix K≠1. If the

azimuthal orientation of the molecule in the unit cell is unknown, $\psi$ cannot be determined, and a more useful function in this case is the Fourier transform of a single molecule calculated for each value of n

$$G(\ell,n,\xi) = \sum_j f_j J_n(2\pi R_j\xi) \exp i \left(\frac{2\pi\ell z_j}{c} - n\phi_j\right) \quad (27)$$

If there is a diad axis in each helical repeat perpendicular to helix axis, then for every atom at $(R_j,\phi_j,z_j)$ there is an equivalent one at $R_j$, $-\phi_j$, $-z_j$ so that the expression $G(\ell,n,\xi)$ need only involve the real (cosine) part of the exponential term. For D-DNA N=8, K=1. Therefore, the selection rule is $n=\ell-8m$ where m=0, ±1, ±2 .....

For a left handed helix the values of n are replaced by -n in the equation (25) and (26) (Klug et al., 1958). It can be shown that the cylindrically averaged squared transform (CAST) is:

$$\text{CAST}\ (\ell,\xi) = \sum_n G^2(\ell,n,\xi) \quad (28)$$

The most general crystal structure will contain p equivalent molecules in the unit cell with fractional unit cell co-ordinates of the helix axis $(x_p, y_p, z_p)$ and relative orientation $\phi_p$. By using equation (24) the general expression for structure factors of helical molecules in a regular array can be written as:

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

$$\cdot \exp i\{2\pi(hx_p + ky_p + \ell z_p)\} \quad (29)$$

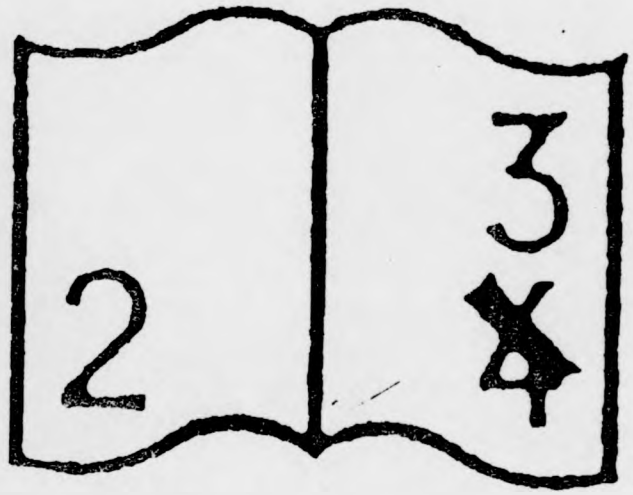
The atomic scattering factors $f_j$ were calculated using the expression given by Vand, Eiland and Pepinsky (1957)

$$f(\sin\theta) = A \exp(-a \sin^2\theta) + B \exp(-b \sin^2\theta) \quad (30)$$

where $\sin\theta = \lambda\rho/2$ and A, a, B, b are constants. Since DNA fibres generally contain an appreciable proportion of water, it makes a significant contribution to the diffraction intensity particularly at small angles of diffraction (Bragg and Perutz, 1952), and allowance must be made for this

# PAGINATION ERROR

p29



contribution when calculating structure factors intended for use in the refinement of the trial structures. The effect of the solvent can be simulated by supposing that the volume occupied by the solvent is filled with an electron gas with a density of the solvent (Hrinch, 1950). The corrected structure factor can be calculated, using Babinet's principle, (Langridge et al., 1960b) from the expression:-

$$F' = F - v\rho_w F_v \qquad (31)$$

F is the structure factor calculated in the usual way, v is the volume of the water displaced by the molecules, $\rho_w$ is the mean electron density of water, and F is the Fourier transform of a solid uniform density with a shape corresponding to the displaced water. A simpler approach can be used by incorporating the correction term into the individual atomic scattering factors. If $v_j$ is the volume displaced by the $j^{th}$ atom then the modified, scattering factor is:-

$$f'_j(\sin\theta) = f_j(\sin\theta) - v_j \rho_w \phi_j(\sin\theta) \qquad (32)$$

where

$$\phi_j(\sin\theta) = 3(\sin(\tfrac{4\pi R\sin\theta}{\lambda}) - \tfrac{4\pi R\sin\theta}{\lambda} \cos(\tfrac{4\pi R\sin\theta}{\lambda})) / (\tfrac{4\pi R}{\lambda}\sin\theta)^3$$

and the value of R is taken as a constant radius of 2A (Langridge et al., 1960b, Fuller 1961). Arnott et al. uses the Van der Waals radius for R. There have been a number of other methods suggested by other workers for calculation of atomic scattering factors. This is discussed in detail by Greenall (1982).

### 3.2.3 Lattice Measurements

X-rays diffracted by calcite-dusted fibres were recorded on flat films. The calcite gives a diffraction ring of spacing 3.029A (In some cases the NaCℓ in the fibre gives a diffraction ring of spacing 2.81A). Either way, the spacing can be used in the determination of the specimen to film

distance D. The co-ordinates of the diffraction spots in the diffraction pattern were measured with a two-dimensional travelling microscope. A computer program (written by W.J. Pigram) was used to convert the film co-ordinates into reciprocal space parameters. By minimising a function

$$\phi = \sum_{j=1}^{m} (r_j - R)^2$$

the radius R and the centre $x_o$, $y_o$ of the calibration ring were obtained where:-

$$r_j = \{(x_j - x_o)^2 + (y_j - y_o)^2\}^{\frac{1}{2}}$$

$x_j$, $y_j$ are the co-ordinates of $j^{th}$ point on the calibration ring. D was determined by using the equation:-

$$\frac{1}{d} = \frac{2}{\lambda} \sin\left(\tfrac{1}{2} \tan^{-1} E/D\right) \quad \text{(M.J. Buerger, 1945)}$$

where $\lambda$ is the wavelength of the CuK$\alpha$ radiation (i.e. 1.5418A) and d is the spacing of the calibration ring. E is the distance of a spot from the centre of the diffraction pattern. The co-ordinates x, y of the diffraction spots are converted into reciprocal space co-ordinates $\zeta$, $\xi$ and reciprocal radius $\rho$ using the following equations:-

$$\zeta = \frac{Y}{\lambda(D^2 + E^2)^{\frac{1}{2}}} \tag{33}$$

$$\xi = \left(\frac{1}{\lambda}\right) \{2 - \lambda^2 \zeta^2 - 2 (1-\lambda^2\zeta^2)^{\frac{1}{2}} D/(D^2 + X^2)^{\frac{1}{2}} \tag{34}$$

$$\rho^2 = \zeta^2 + \xi^2 \tag{35}$$

### 3.2.4 The refinement of the unit cell parameters

The approximate values of the lattice parameters were estimated from measurements of the photographs and these were refined using a least squares procedure (Fuller, 1961). By minimising the function:-

$$\phi = \sum_{\ell=1}^{n} (\rho_{iobs} - \rho_{ical})^2$$

where $\rho_{iobs}$ = the $i^{th}$ observed $\rho$ value

$\rho_{ical}$ = the $i^{th}$ calculated $\rho$ value

n = total number of reflections

The final lattice parameters can thus be obtained.

### 3.2.5 Intensity Measurements

Methods to obtain diffracted intensities from fibre diffraction photographs have been discussed by Franklin and Gosling (1953c), and Langridge et al. (1960a). The integrated intensity $I_o$ is proportional to A, where A is the area under the radial densitometer trace across the spot in the fibre diagram. The $I_o$ is multiplied by the Lorentz correction factor, which is proportional to $\xi$, and for the drawing out of the spots into arcs of constant width, the intensity is further multiplied by $\rho$. The intensities must also be corrected for the absorption effect; which is proportional to cos $\theta$, and polarisation correction, which is equal to:-

$$2/(1 + \cos^2 2\theta)$$

The corrected integrated intensity is

$$I = A\rho\xi \times \cos\theta \frac{2}{(1+\cos^2 2\theta)} \qquad (37)$$

### 3.2.6 Scaling of the intensities

The scaling factor,

$$S_I = \frac{\sum_{hk\ell} I_o(hk\ell)}{\sum_{hk\ell} I_c(hk\ell)} \qquad (38)$$

where $I_o(hk\ell)$ = the observed (corrected) intensity of the $(hk\ell)$ reflection

and $I_c(hk\ell)$ = the calculated intensity of the $(hk\ell)$ reflection

The scaling factor for structure factor amplitudes is $S_F = (S_I)^{\frac{1}{2}}$

3.2.7 Comparison of agreement obtained from different models

It is very useful to define a method to compare different models obtained from the same diffraction data.

The residual

$$R_1 = \frac{\sum_{hk\ell} \left| |kF_c| - |F_o| \right|}{\sum_{hk\ell} |F_o|} \qquad (39)$$

and

$$R_2 = \frac{\sum_{hk\ell} \left| \alpha I_c - I_o \right|}{\sum_{hk\ell} I_o} \qquad (40)$$

3.2.8 The packing of the molecules in the unit cell

(1) Consider a molecule of orientation $\theta$ in the unit cell (which is shown schematically in figure 6). During molecular model building the orientation $\theta$ is varied by small angle such that the steric interaction with the neighbouring molecules is minimised, and this enables best molecular packing within the cell to be determined. If there is more than one molecule in the unit cell each molecule is rotated about its axis and displaced along the c-axis in small steps so as to obtain the best orientation.

(2) To obtain the best agreement between the observed and calculated structure factors (once the best molecular transform is obtained from model building studies) the value of $R_1$ (see equation (39)) has to be minimised. The structure factors for different orientations of the molecule (or molecules)can be calculated by rotating the molecule in the unit cell. Hence, the value of $R_1$ is obtained as a function of orientation $\theta$. As the molecular orientation is changed, $R_1$ varies and will pass through a minimum. For a good model, the value of $\theta$ from (1) and $\theta_{min}$ from (2) must be equal or nearly equal.

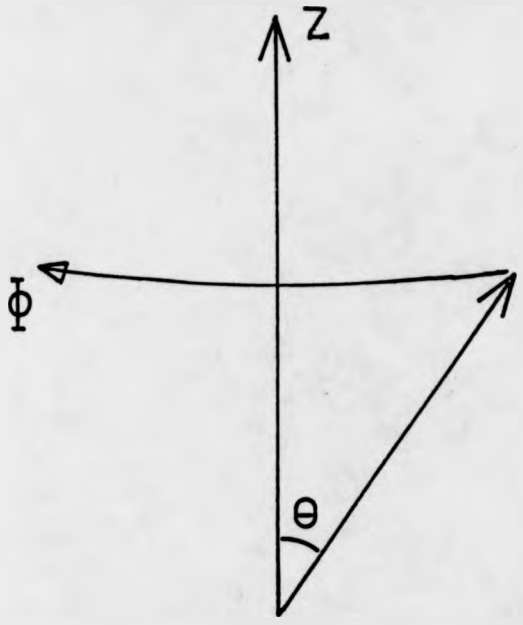



Figure 5 Molecular Spatial Arrangement in a Fibre

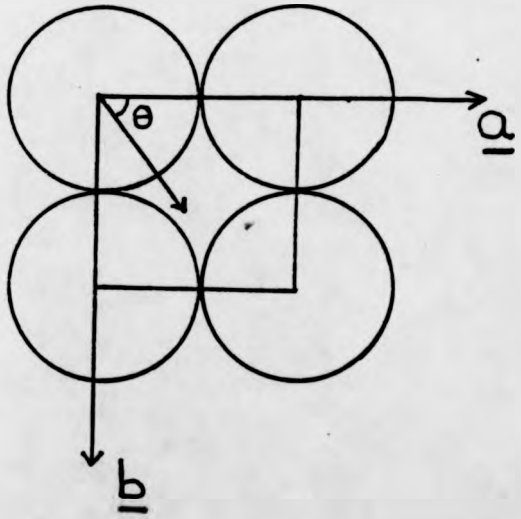



Figure 6 Molecular Arrangement in a Unit Cell

# CHAPTER 4

## The D Conformation

### 4.1 Introduction

Synthetic DNA polymers are of great interest in the study of DNA as a whole, since they provide consistently repeating nucleotide sequences which may well behave in much the same way as constituent parts of native DNAs; for example satellite DNAs, and possibly regions which are believed to be actively involved in transcription.

The various synthetic polymers of DNA that have been isolated are often associated with characteristic molecular geometries, and it is for this reason (amongst others) that their study is of particular importance; it provides a starting point for a correlation between the structural and functional roles of the macromolecule.

Davies and Baldwin (1963) observed that the sodium salt of poly d(AT).poly d(AT) could assume not only the crystalline A and semi-crystalline B forms observed for naturally occurring DNAs (Fuller et al., 1965, Langridge et al., 1960b, Arnott et al., 1972b, 1973, Rhodes et al., 1982) but also a new conformation which was designated the 'D' form. The diffraction patterns were of poor quality, and only the axial periodicity (24.5 A) and the general intensity distribution was described. They also reported a poor quality diffraction pattern of the D conformation for $NH_4$dAT. Further, it was reported that an unusual feature of the pattern was the appearance of reflections near the meridian on the third and fourth layerlines. It was observed that some of the fibres gave mixtures of A and D patterns which after a time generally changed completely to the D form (Davies et al., 1963). These observations suggested that the D

form is a particularly stable conformation for this polynucleotide. The diffraction patterns from Na dAT and $NH_4$dAT closely resemble each other, apart from the fact that the Na dAT material was more crystalline (Davies, et al., 1963).

A similar diffraction pattern was obtained from a fibre of Na poly d(IC).poly d(IC). This was indexed on a tetragonal lattice with a = 17.5 A and c = 25.0 A (Mitsui, et al., 1970).

In the same publication they also suggested that the patterns from poly d(IC).poly d(IC) were best accounted for by a left-handed double helix with Watson-Crick base-pairing, and eight nucleotide paris per helix pitch of 25.01 A. In this analysis they systematically varied the position of the base-pairs with respect to the helix axis, the orientation of the sugar base link, and the sugar puckering. They also investigated right-handed models, but found that these were inferior in terms of stereo-chemistry and agreement with X-ray data, when compared with the best left-handed models. However, since neither atomic co-ordinates nor torsion angles were published for any of these models, it was not possible to see the extent to which the left-handed model was to be preferred.

For the polynucleotide poly d(AT).poly d(AT) Arnott et al., (1974b) reported that the D form was observed under conditions which would normally favour the A form of natural DNAs, ie a minimum of retained salt in the fibre. Arnott et al. (1974b) also emphasised that the D pattern from poly d(AT).poly d(AT) was very similar to patterns obtained by Mitsui et al. (1970) from Na poly d(IC).poly d(IC) and by themselves from poly d(GC).poly d(GC).

On the basis of a linked atom least squares refinement technique, they dismissed the left-handed model of Mitsui et al. (1970) as bizarre, and proposed a model for all three complementary D-DNA structures (poly d(AT).poly d(AT), poly d(IC).poly d(IC), poly d(GC).poly d(GC)). This was a right-handed eight-fold double helical model which used Watson-Crick base

pairing, along with sugar puckering and conformational angles reminiscent of the conformation of natural DNAs.

The D conformation was also observed for Na poly d(ATT).poly d(AAT) and such patterns from this material were indexed on a hexagonal lattice with a = b = 19.8 A and c = 24.1 A (Selsing et al., 1975). Leslie et al. (1980) reported the D conformation for the following polynucleotides:

(a) Na poly d(AC).poly d(IT); indexed on a hexagonal lattice with a = b = 20.4 A and c = 25.1 A

(b) Na poly d(AIT).poly d(ACT); indexed on a hexagonal lattice with a = b = 20.2 A and c = 24.3 A

(c) Na poly d(AIC).poly d(ICT); indexed on a hexagonal lattice with a = b = 20.4 A and c = 24.5 A.

Leslie et al. (1980) prepared one fibre from Na poly d(AT).poly d(AT) for which the A conformation was stable over a period of six months, even at very high relative humidity (95%). However, they were not able to identify the conditions which favoured the A and D forms although for one fibre, which they described as having been made from 'impure synthetic' DNA, they observed a D → A → B transition by raising the relative humidity of the fibre environment. No indication was given either of the nature of the impurities in the fibre, or of the quality of the patterns which were obtained.

Gupta, Bansal and Sasisekharan (1980) have recently described a survey of right-handed and left-handed models for A, B and D forms of DNA. In particular, they claim that both the right- and left-handed models which they propose for the D conformation are in qualitative agreement with the observed diffraction pattern, the residuals being very similar for both RH and LH, and very close to the values calculated for the right-handed models proposed by Arnott et al. (1974b). Drew et al. (1982) proposed a new model for D-DNA which is very different from the earlier

models. They build a left-handed helix with seven residues per turn by incorporating Hoogsteen base-pairs into a Z helix framework. They claim that the X-ray intensities calculated from this novel left-handed D model provide a better fit to the D-DNA diffraction pattern than do intensities calculated from previously proposed D helix structures (Drew et al., 1982).

The main themes of this chapter are:

(a) The humidity and salinity condition under which the high quality crystalline D-form is routinely observed for Na poly d(AT).poly d(AT).

(b) The question of whether the D-form is a sevenfold helix (as proposed by Drew et al., 1982) or an eight-fold helix.

(c) A re-appraisal of the models previously proposed for the D conformation and a description of a left-handed model for which the calculated diffraction is in substantially better agreement with the observed than that calculated for any other previously published model for D-DNA, whether of the right- or left-handed type.

A detailed study of the polymorphism of poly d(AT).poly d(AT) will be discussed in the following chapter.

## 4.2 Materials and Methods

### (a) Fibre Preparation

The polynculeotides used in this study were obtained from Boehringer. Samples were also provided by Dr. J. Brahms of the University of Paris. Fibres were prepared from ethanol precipitated polynucleotides or centrifuged gels, as discussed in Chapter 2.

### (b) Model Building

It is not possible to solve the phase problem directly for fibre diffraction patterns from nucleic acids. The method of

analysis used is to build a good trial model from general considerations and then to refine it by a detailed comparison of the observed and calculated intensities.

The preliminary structures, prior to detailed refinement were built using wire models, which used 'shrunken' atoms, and a representative scale of 4 cm per A unit. A variety of such models were built and systematically changed until there was good general agreement between the cylindrically averaged square transform of the model, and the observed intensities. The best wire model was then used to provide torsion angles (see figure 1) for a starting model, which was refined using a linked atom least squares computer program. This model building program was originally written by Pigram (1968), and later modified by Goodwin (1977) and is briefly described in Chapter 2. The model building routine also incorporates relevant stereochemical data, such as standard bond lengths and angles obtained from single crystal work on nucleotides. Standard sugar puckering conformations (Arnott and Hukins,1972a) have been used, and as such, the torsion angles $\tau_5$, $\tau_{10}$ have been fixed in agreement with C2-endo sugar puckering. The twist and tilt (dyad) axes are as shown in more detail in figure 1. The rotation and rise per residue are shown in figure 1 by $\tau_{10}$, $\rho_{15}$ respectively, and the base displacement from the helix axis is shown as 'D' in figure

During the early stages of the refinement, only the 'backbone' torsion angles were allowed to vary, but later on twist, tilt and base displacement were allowed to vary as a means of eliminating bad interbase, intra helical, and inter helical contacts. The final model was thus optimised in terms of agreement with X-ray data, good stereochemistry and satisfactory puckering into the tetragonal lattice. Having obtained a final model, the structure factors and residuals were calculated and compared with other published models.

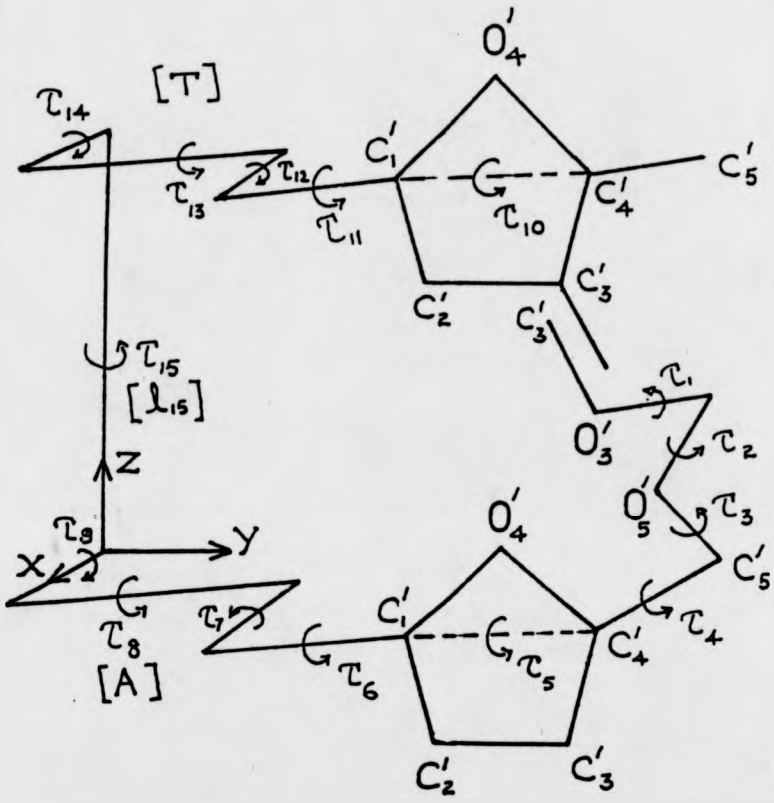



Figure 1 Model Building of the Dinucleotide using the two-branch method

4.3 Results and Discussions

The fibres used were estimated to contain a salt concentration corresponding to about 0.5 $Na^+$ and $C\ell^-$ ions for every phosphate group.

At low relative humidities (33% → 66%), the diffraction patterns obtained showed the A form of DNA to be predominant, but there was, however, evidence of C-DNA impurity (the presence of C-DNA was indicated by the existence of a strong meridional reflection at 3.3 A; see chapter 6). As the humidity was increased to 75%, a highly crystalline A-form diffraction pattern was obtained, and this pattern persisted right up until a relative humidity of 98%. However, after prolonged exposure at 98% RH diffraction patterns were obtained which indicated first an A-DNA/B-DNA mixture, and later on, highly crystalline D-DNA. After further exposure at 98% RH, a B-form pattern was observed. However, as the humidity environment of the same fibre was decreased to 95%, the D-form pattern returned and remained, right down to relative humidities as low as 33% (figure 2). It is noteworthy that the A-conformation was never recovered, either by remaking the fibre or by increasing the salt content. It thus seems that the A-DNA to D-DNA transition is irreversible by any means, and in agreement with previous reports, that the D-form is a particularly stable conformation of Na poly d(AT).poly d(AT). These observations indicate a marked tendency for the polynucleotide to become "locked" into the D-form, which thereafter appears to be resistant to large changes in the RH of the environment.

Other fibres were prepared from the same centrifuged gel, and the diffraction patterns so obtained indicated A/C mixture (see detailed discussion in Chapter 6) as above. However, in contradiction to earlier reports by Davies et al. (1963) this conformational mixture remained, even after several months at room humidity. Davies et al. (1963) had claimed that the A-form obtained from this polynucleotide changed from A to D conformation within a few days. The transitions observed from all

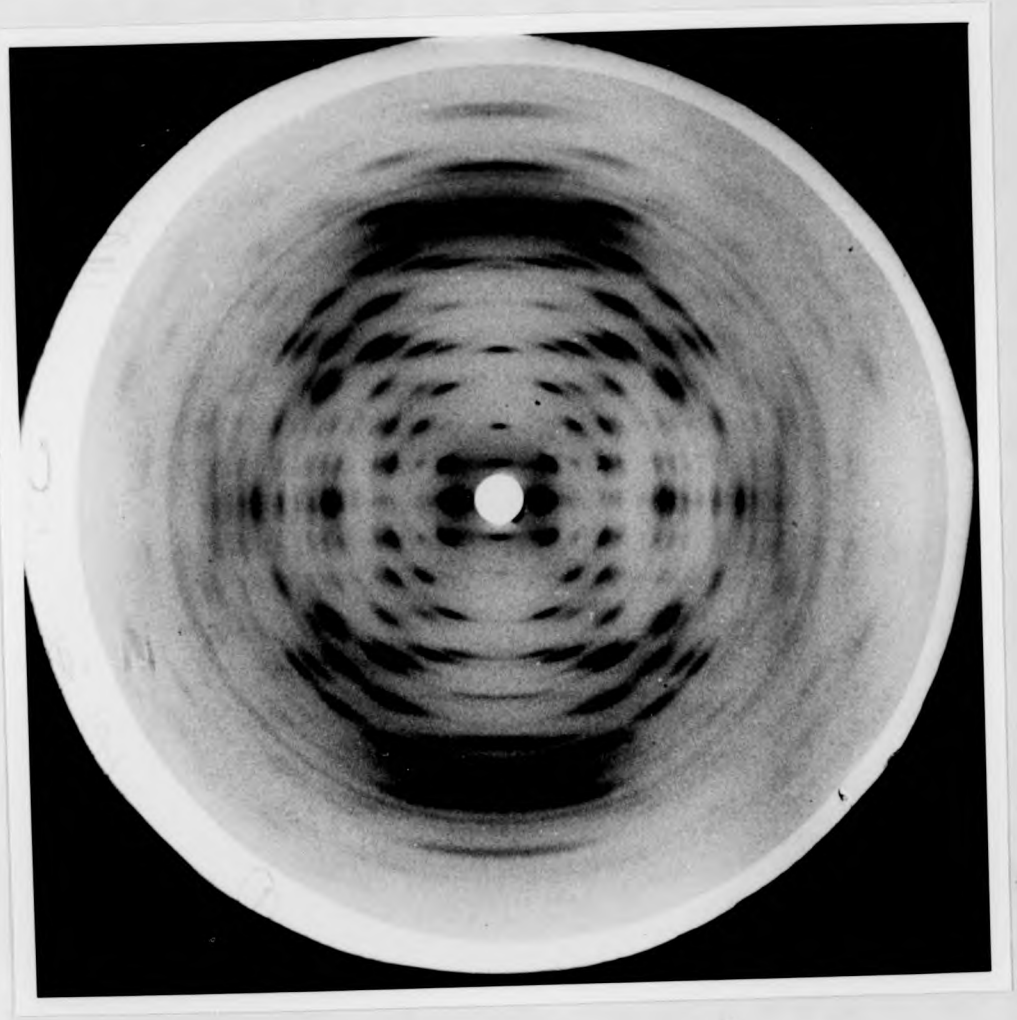


Figure 2 D-form Diffraction Pattern of Napoly d(AT).poly d(AT) at 75% R.H.

fibres prepared subsequently from this material showed 100% reproducibility when the fibres were treated in the same way even when they were several months old.

When the salt concentration was increased slightly, and the same sequence of the humidities used, a semi-crystalline B-DNA pattern was obtained after the A/B mixture (instead of the D-form).

The above experimental observations of fibres prepared from Na poly d(AT).poly d(AT) at low salt concentrations lead one to two significant conclusions:-

(i) Induction of the A → B transition is very difficult at low salt concentrations. However, an intermediate A/B mixture is obtained when a fibre is maintained at high relative humidity for a prolonged period. Then the D-form was obtained before semi-crystalline B form.

(ii) At slightly higher salt conditions the B-form is obtained (without difficulty) upon increase of relative humidity.

A major question relating to the structure of D-DNA, as pointed out earlier, is that of whether the molecular contains seven or eight base pairs per full turn of the helix. Evidence obtained in this study shows, for the first time, that the eight-fold structure is to be preferred. By suitable tilting of the fibre, an eighth layer line meridional reflection was clearly distinguished from the neighbouring reflection (108) (figure 3). Evidence supporting this comes from Mitsui et al. (1970), and Arnott et al. (1974b). Mitsui et al. (1970) studied D-DNA from Na poly d(IC). poly d(IC). Arnott et al. (1974b) studied D-DNA from Na poly d(AT).poly d(AT). Both authors favoured an eight-fold helical structure with strong seventh layerline reflections accouted for by highly tilted bases. A recent model proposed by Drew et al. (1982) shows a seven fold structure. However, none of these workers have published any clear experimental evidence to support

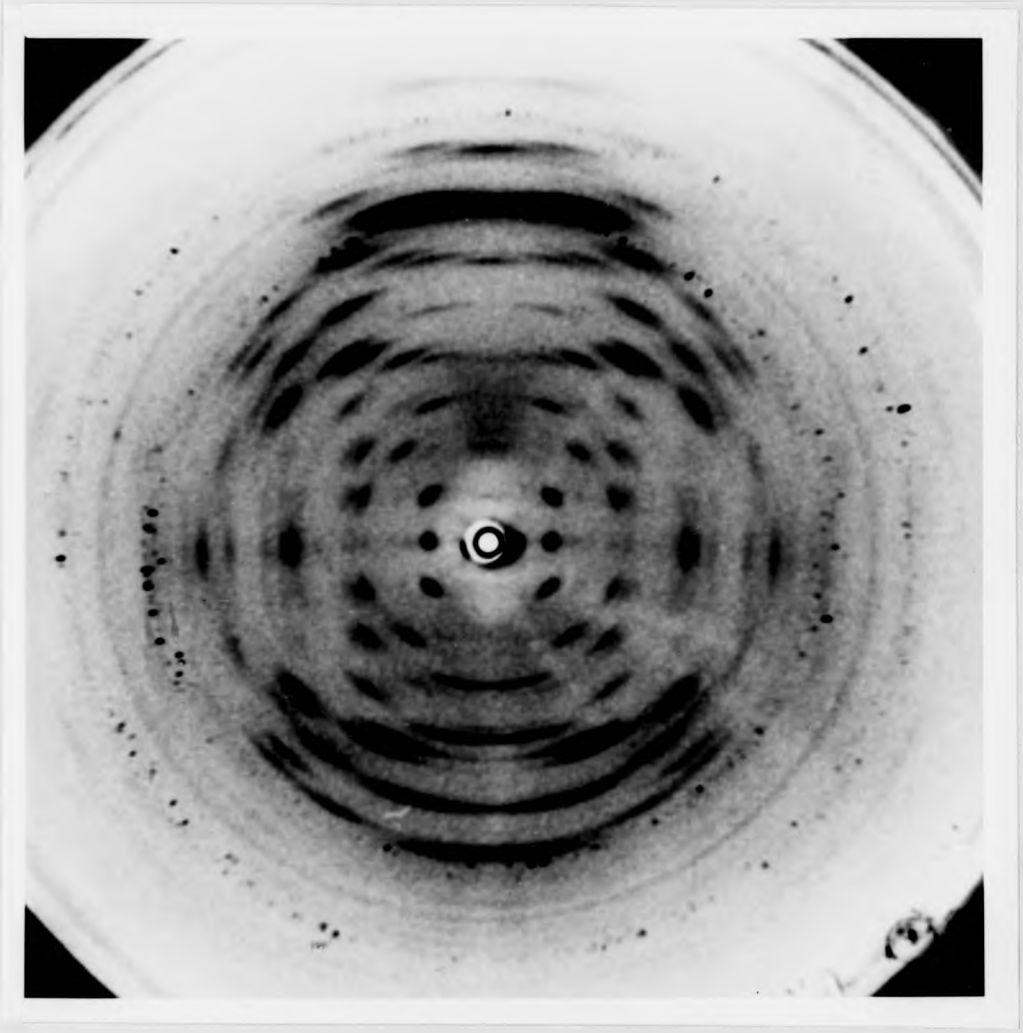


Figure 3 D-form Diffraction Pattern of Na poly d(AT).poly d(AT) at 75% R.H. Fibre is appropriately tilted to resolve the meridional reflection (008) and (108) reflection

their contentions.

In this study, diffraction from tilted fibres shows clearly that there is no meridional reflection on the seventh layerline (figure 4), and that it is the 107 and 117 reflections which have been mistakenly taken for meridionals from less oriented poor quality diffraction patterns (these reflections are well resolved on weakly exposed patterns, and on the underfilms of normally exposed pictures (figure 4)). These considerations thus contradict the proposal of Drew et al. (1982).

The D-form diffraction pattern shown in figure 4 was measured up, and the helical parameters and unit cell dimensions determined. The pitch and the base separation are 24.1 A and 3.01 A respectively. The unit cell parameters are given by a = b = 17.1 A, c = 24.1 A. The calculated and observed $\rho$ values along with the RMS deviation in $\rho$, are shown in table 1. The pitch and the unit cell parameters are somewhat less than those previously published in Arnott et al. (1974b).

Fibres prepared from Na poly d(IC).poly d(IC) gave D form diffraction patterns (figure 5) up to 92% RH; it was noted that the crystallinity of these fibres was improved by thermal annealing. Although as pointed out by Arnott et al. (1974b) the D patterns obtained from Na poly d(AT).poly d(AT) and Na poly d(IC).poly d(IC) are similar, it is important to emphasise that there are differences evident from a visual inspection. The 102 reflection is essentially absent or very weak in Na poly d(AT).poly d(AT), and the 112 is of intermediate strength; the reverse is true for poly d(IC).poly d(IC) patterns. These striking differences, which are quite reproducibly observed, occur in a region of the diffraction pattern for which the molecular transform can be expected to be cylindircally symmetrical. It is thus not possible to attribute them either to differences in molecular orientation within the unit cell, or to differences in lattice parameters. It seems possible that these anomalies may occur as a result of different distribution of ions in the two molecular species. It is unlikely that

Table 1

Observed and calculated $\rho$-values for the D-pattern in figure 4

RH = 75% a = b = 17.1 A c = 24.1 A

| h | k | ℓ | $\rho_o(A^{-1})$ | $\rho_c(A^{-1})$ |
|---|---|---|---|---|
| 1 | 0 | 0 | 0.0574 | 0.0579 |
| 1 | 0 | 1 | 0.0709 | 0.0711 |
| 2 | 0 | 1 | 0.1228 | 0.1229 |
| 1 | 1 | 2 | 0.1181 | 0.1164 |
| 2 | 0 | 2 | 0.1425 | 0.1423 |
| 1 | 0 | 3 | 0.1372 | 0.1369 |
| 2 | 1 | 3 | 0.1766 | 0.1793 |
| 1 | 0 | 4 | 0.1756 | 0.1753 |
| 1 | 1 | 4 | 0.1847 | 0.1846 |
| 2 | 0 | 4 | 0.2018 | 0.2019 |
| 2 | 1 | 4 | 0.2129 | 0.2101 |
| 1 | 0 | 6 | 0.2542 | 0.2549 |
| 1 | 1 | 6 | 0.2614 | 0.2614 |
| 2 | 0 | 6 | 0.2739 | 0.2739 |
| 2 | 1 | 6 | 0.2798 | 0.2799 |
| 1 | 0 | 7 | 0.2953 | 0.2953 |
| 1 | 1 | 7 | 0.3009 | 0.3009 |
| 2 | 1 | 7 | 0.3166 | 0.3172 |
| 1 | 0 | 8 | 0.3362 | 0.3360 |

The rms deviation in $\rho = 6.2 \times 10^{-3}$

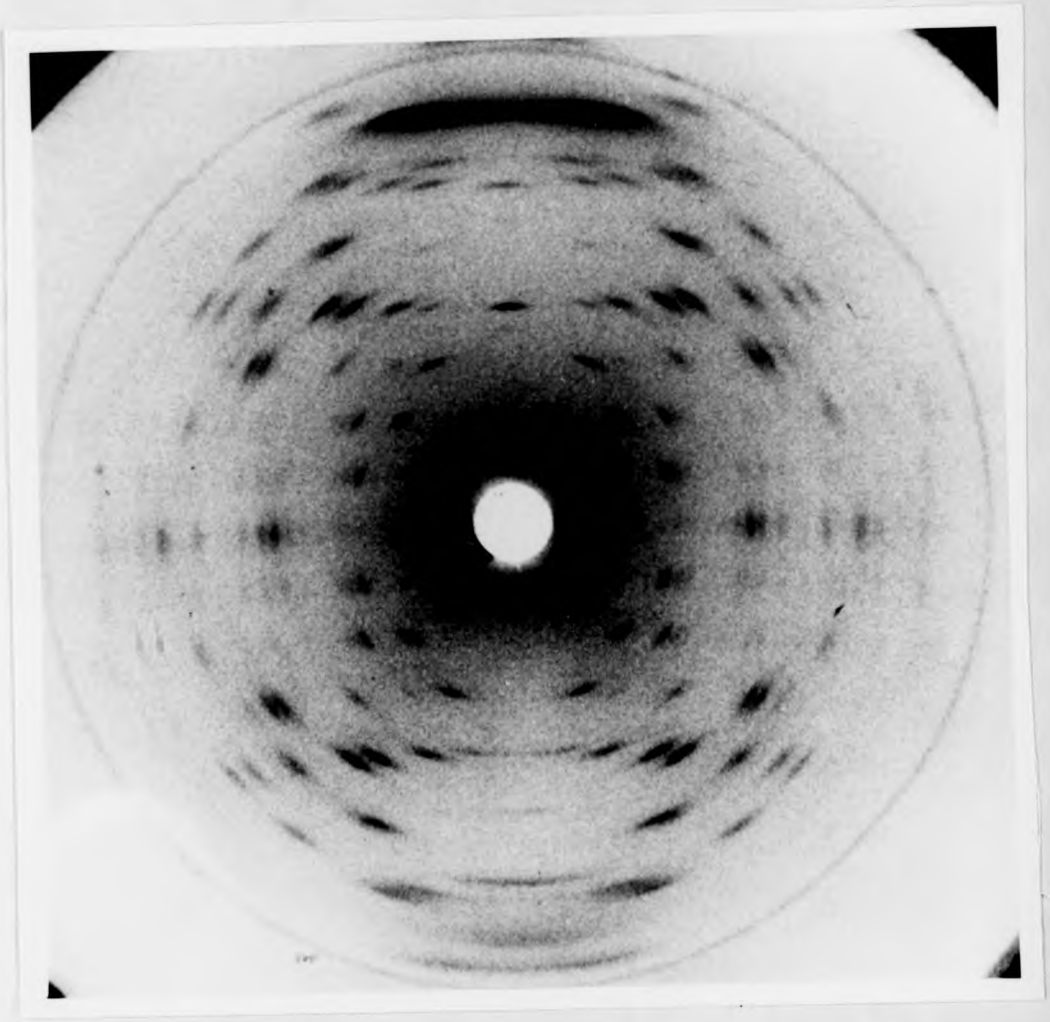


Figure 4 D-form Diffraction Pattern of Na poly d(AT).poly d(AT) at 75% R.H.
Fibre is appropriately tilted to resolve the strong seventh layerline reflections

there is any major structural difference between the two polynucleotides. In the light of these differences, it was thought preferable to study models based around diffraction data obtained from Na poly d(AT).poly d(AT).

Despite the fact that Arnott et al. (1974b) claim as good an agreement between observed and calculated diffraction for D-DNA as has been obtained for other nucleotide structures, there are a number of ways in which their D-DNA model is inferior (in terms of agreement) to the best models proposed for crystalline A and B-DNA. This is illustrated in figure 6, where the calculated diffraction is shown along with the observed. It should be noted that, as a result of the fact that Arnott et al. (1974b) did not observe a number of reflections, most importantly (102), (113) and (105) reflections (as in figure 5), the table comparing the observed and calculated structure factors (Arnott et al., 1974b) is now revealed to have a number of major discrepancies. The stereochemistry of Arnott's model is satisfactory, with the exception of the one short contact of 2.4 A between deoxyribose $C_2$ and phosphate $O_3$.

Both the left-handed and right-handed models proposed for the D-form Gupta et al. (1980) give poor agreement between observed and calculated diffraction (figure 7) and whilst the intramolecular stereochemistry is acceptable, the molecular model ends up in serious interhelical contact trouble when one attempts to pack the molecule into the observed tetragonal lattice.

The problems and discrepancies outlined above relating to the various proposed models are of particular concern since all three models were obtained using sophisticated refinement procedures, and might therefore have been expected to yield slightly better models. The question arises, in passing, as to whether problems have arisen as a result of the intrinsic nature of the models, or technique used to refine them.

It was, thus thought very important to have a fresh look at the X-ray data and investigate a new unbiased analysis. Firstly, it is clear

Figure 5 D-form Diffraction Pattern of Na poly d(IC).poly d(IC) at 44% R.H.

from figure 3 and figure 4 and previous discussion, that the D-form of the DNA is an eight-fold helix having a pitch of 24.1 A. Secondly, since D-DNA has only been observed for DNA polymers containing exclusively A/T or I/C base pairs, it seems unlikely that the base pairing should be any other than the Watson-Crick type. Non Watson-Crick pairing would require protonation of cytosine $N_3$ in order to establish effective hydrogen bonding. Mitsui et al. (1970), in order to justify their use of Watson-Crick base-pairing, prepared fibres of Na poly d(IC).poly d(IC) in differing pH conditions. The fibres initially used in this study were prepared from an alochol precipitated material. They were rewetted with distilled water (pH-6) and diffraction patterns taken. They also took X-ray pictures of fibres prepared by precipitation from two other solutions (0.01 M sodium phosphate buffer of pH 7.2, and 0.001 M tris-maleate buffer of pH 8.2). The diffraction patterns thus obtained, showed that although the fibres were not as crystalline and not as well oriented as previous samples, they gave essentially the same pattern. It therefore seems highly unlikely that the bases in this polymer are protonated at pH 8.2. On the basis of this argument, the Watson-Crick base pairing is much favoured to the Hoogsteen base-pairing as used by Drew et al. (1982).

Having thus established very fundamental features of D-DNA, the analysis then proceeded to the stage of formulating a preliminary model, bearing in mind that the main contributions the molecular transform would be from the phosphate and bases. The sugar atoms have a much weaker contribution.

The diffraction pattern of D-DNA (figure 3 and figure 4) shows a very strong 100 reflection, indicating that it falls on the principle maximum of the transform, rather than on the 1st subsidiary maximum. At the same time a relative weakness of the subsidiary maximum is suggested by the fact that the 110, 200 and 210 reflections are all absent. This

transform profile implies that the molecular structure is more of a solid cylinder type (as in B-DNA) than a 'hollow' cylinder type (as in A-DNA) and that the bases are quite close to the helix axis.

The next observed equatorial reflections are the 310 and the 300 reflections; these are both fairly strong as a result of contribution from the $J_8$ Bessel function component (the selection rule is $n = \ell - 8m$ $m = \pm 1$ implies along the equator $n = \pm 8$). The phase term associated with this is $(\ell\theta - n\phi)$ (following equation 27 in Chapter 3) reducing to $n\phi$ for $\ell = 0$ on the equator. Thus, for strong phosphate contribution in this region, one would expect $8\phi$ to be of the order of $360^{o}$, suggesting a phosphate phase angle of $\sim 45^{o}$.

The phosphate radius was easily fixed by adjusting the molecular transform to fit the equatorial reflections. A number of different trial models were constructed with phosphate radii ranging from 7.5 A to 9.5 A ($\sim$ 0.5 A steps). The best fit was obtained for a phosphate radius of 8.5 A.

The approximate azimuthal height of the phosphate, Zph, was chosen by adjusting the transform to fit the lower layer line reflections. The selection rule for an integral helix is $n = \ell - Nm$. For an eight-fold D-DNA molecule, this reduces to $n = \ell - 8m$. For lower layer lines, the dominant Bessel functions are going to be given by the above equation. Thus when $m = 0$, $n = \ell$. For a left-handed helix, the phase term is $(\ell\theta + n\phi)$ (for right-handed helix $(\ell\theta - n\phi)$), but since $n=\ell$ for $m=0$, this then becomes $\ell(\theta + \phi)$, where $\theta = {}^{2\pi z}/_{c}$.

The observed diffraction shows a strong 1st layer line in comparison to the 2nd and 3rd, very much like the diffraction expected from a single stranded helical molecule. The inference is that the two chains are fairly close together. This is shown schematically in figure 8.

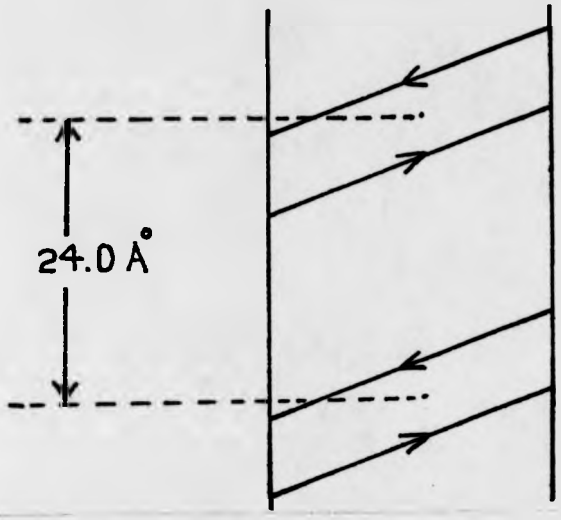



Figure 8

On the 1st layerline, the 101 reflection is strong whereas the 111 is absent. On the second layer line, however, the 102 reflection is absent (or very weak), but the 112 is present. Since the phosphate contribution to the transform is very significant (and sensitive to the phase variation) in the region of the 10 reflections, it might be expected that the phase term cos $(\theta + \frac{1}{2}\phi)$ is $\sim 160^{o}$. This will give a negative base contribution and tend to cause some cancellation of the phosphate component in the same region, thus providing the right relative magnitude for the 112 reflection. The above mentioned phase conditions also satisfy the 3rd and 4th layer line intensities.

The intensities of higher layer line reflections are very sensitive to the base tilt. From the helical selection rule and the transform equations, the phase term for base contribution is $(\ell\theta + n\phi)$ i.e. for strong 7th layer line, $7\theta - \phi$ should be approximately zero (this is for a left-handed helix). Thus $7\theta - \phi = 0$, and the expected tilt can then be calculated: e.g. for $\phi \sim 80^{o}$:

$$\theta \simeq {}^{80}/_{7} \simeq 11^{o} \qquad \theta = \frac{2\pi z}{c}$$

$$z = \frac{c\theta}{360} = \frac{24 \times 11}{360} = 0.7 \text{ A}$$

Therefore the tilt = $\tan^{-1}\left(0.7/5\right) \sim 8^{o}$.

During the preliminary molecular model building, it was noted that tilting of the bases affected the position of the phosphate. In order to attain the required tilt and maintain the phosphate position, it was found necessary to twist the bases. Base twisting has a very much smaller effect on the transform than does tilting, and as such, was not only used to restore and maintain phosphate position, but was also found valuable in eliminating short intrahelical contacts in the final stages of modelbuilding.

By applying these considerations to the D-DNA model, several trial models were built and adjusted to obtain the best fit to the X-ray data. Once this had been achieved, the models were refined to optimise the stereochemistry (i.e. to remove the bad inter- and intra-helical contacts). The results of the analysis has led to the derivation of the left-handed model for the D-form in which the agreement with the X-ray data is significantly better than that of the best right-handed model proposed by Arnott et al. (1974b), and substantially better than that of either the right- or left-handed models proposed by Gupta et al. (1980). It has not proved possible to build a right-handed model which gives better agreement with the observed data than that proposed by Arnott et al. (1974b). The calculated diffraction of all four models is compared with the observed diffraction in figure 6 and figure 7. The structure factors have been calculated for the left-handed model proposed here as a function of the orientation of the molecule in the unit cell. From these calculations the best agreement between observed and calculated structure factors is the dyad axis is at an angle of $15^{o}$ to the unit cell 'a' axis. In order to make as meaningful a comparison as possible of the four models, all calculations used the atomic scattering factors, temperature factor and the scaling procedure employed by Arnott et al. (1974b). The structure factors of these

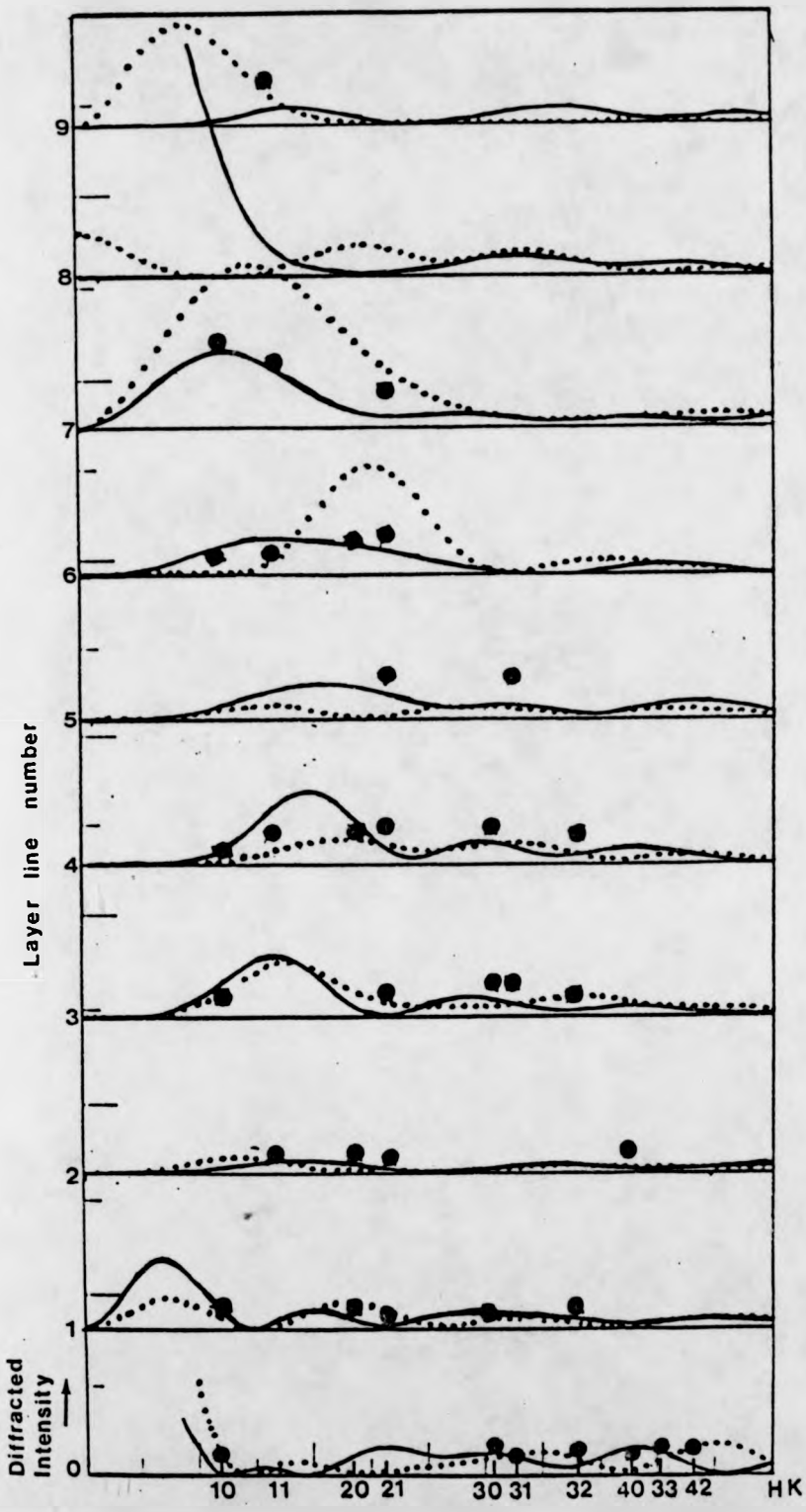



Figure 6 The cylindrically averaged squared molecular transform and the observed intensities (represented by filled circles) for D-DNA. The right-handed model proposed by Arnott et al (1974) and that proposed in this study are in dotted-line & full-line respectively.

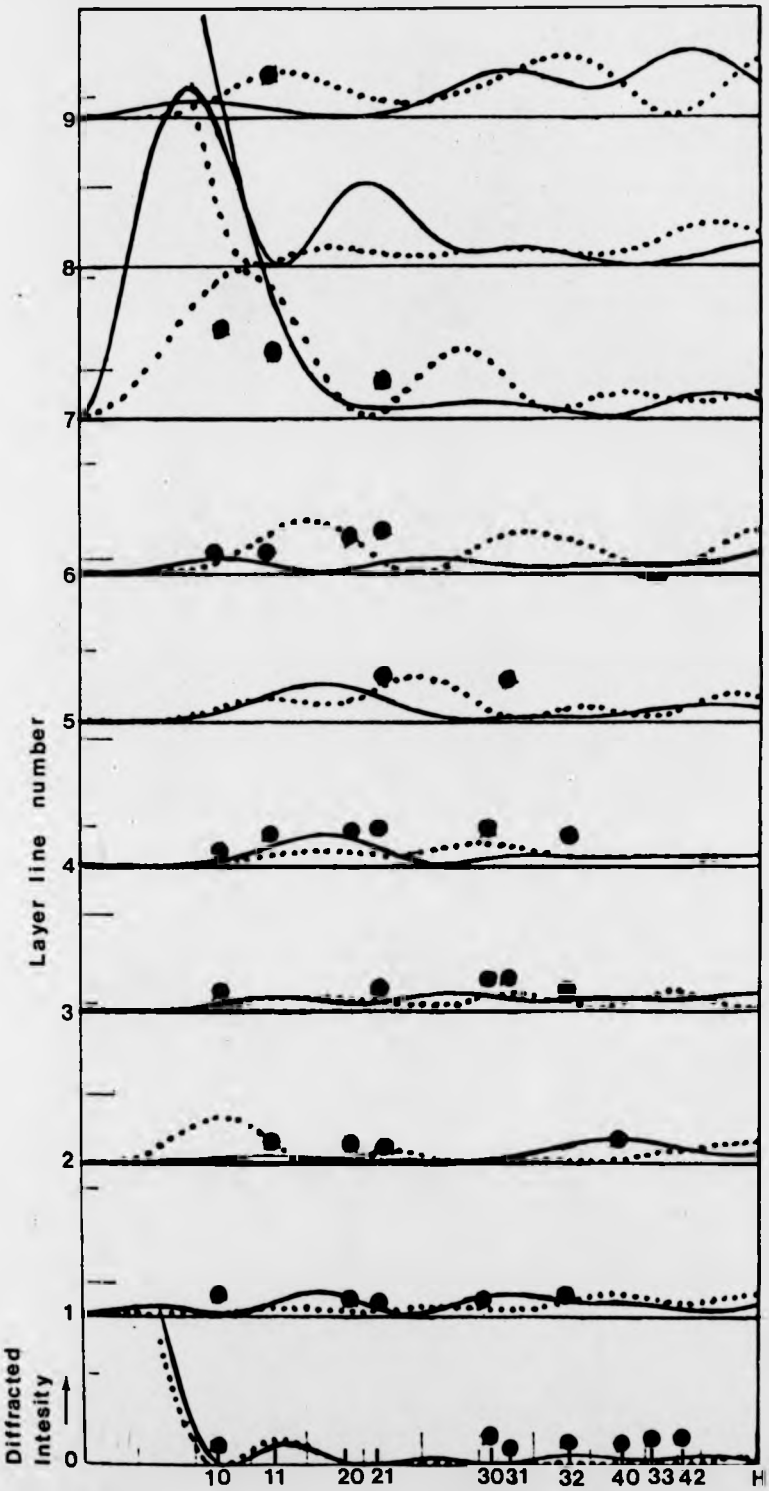



Figure 7 The cylindrically averaged squared molecular transform and the observed intensities (represented by filled circles) for D-DNA. The right-handed and left-handed model proposed by Gupta et al. (1980) are in dotted-line and full-line respectively.

four models are tabulated in table 2. The residuals for the left-handed model developed here, the right-handed model proposed by Arnott et al. (1974b), and right- and left-handed models proposed by Gupta et al. (1980) are respectively 29%, 33%, 32% and 35%. Atomic co-ordinates of the model proposed here are shown in table 3. The torsion angles for the four models are tabulated in table 4. The base-pairs lie close to the helix axis and have tilt and twist of $-6^{o}$ and $8.5^{o}$ respectively. The adenine and thymine nucleotides have identical conformations with an anti-orientation about the glycosidic link and a $C_2$-endo pucker of the sugar ring. The stereochemistry of the model is satisfactory, with no intra- or inter-molecular non-bonded interatomic contacts more than 0.4 A less than standard van der Waals distances. The preferred molecular orientation from an analysis of contacts between atoms in neighbouring molecules coincides with that determined from an analysis of X-ray data.

The fact that a left-handed model for the D-pattern of Na poly d(AT).poly d(AT) can be constructed, having satisfactory intra-molecular and inter-molecular stereochemistry, and being in better agreement with the X-ray data than the currently accepted right-handed model for this structure raises a number of issues of major conformational and biological importance for the role of regions of DNA containing alternating adenine, thymine sequences. While there is no great difficulty in visualising transitions within orientated fibres between conformations of two-stranded polynucleotides of the same helix sense, there are difficulties in such visualisation when the transition is between a right-handed and left-handed form or vice versa. However, encouragement of the consideration of such a possibility comes from the observation that the oligonucleotides $d(CpGCpCpG)_2$ (Wang et al., 1979) and d(CpGpCpG) (Drew et al., 1980) both crystallise as two stranded complexes with a left-handed helical conformation. Arnott et al. (1980) reported that a structure very similar to that

TABLE 2

Indices of the Observed Bragg Reflections with the Observed ($F_{obs}$) and Calculated ($F_{calc}$) Values of the Structure Amplitudes for D-DNA Models

| h | k | ℓ | $F_{obs}$ | $F_{calc}$(RH) Arnott et al | $F_{cal}$(RH) Gupta et al | $F_{cal}$(LH) Gupta et al | $F_{cal}$(LH) (Present Model) |
|---|---|---|---|---|---|---|---|
| 1 | 0 | 0 | 2571 | 2426 | 408 | 290 | 839 |
| 3 | 0 | 0 | 3641 | 2948 | 458 | 361 | 4225 |
| 1 | 3 | 0 | 2460 | 4812 | 859 | 696 | 2895 |
| 2 | 3 | 0 | 3088 | 3048 | 1487 | 1301 | 3463 |
| 4 | 0 | 0 | 2895 | 1755 | 381 | 351 | 3074 |
| 3 | 3 | 0 | 3679 | 2336 | 891 | 873 | 3183 |
| 2 | 4 | 0 | 3743 | 5701 | 1644 | 1690 | 1864 |
| †4 | 3 | 0 | | | | | |
| 5 | 0 | 0 | | | | | |
| 1 | 5 | 0 | 4427 | 3357 | 4578 | 5497 | 5287 |
| 1 | 0 | 1 | 2743 | 2593 | 2100 | 1457 | 3479 |
| 2 | 0 | 1 | 2576 | 2767 | 1286 | 944 | 2149 |
| 1 | 2 | 1 | 1594 | 2819 | 3567 | 2649 | 2098 |
| 3 | 0 | 1 | 2054 | 1422 | 2568 | 2081 | 2323 |
| 2 | 3 | 1 | 2460 | 1833 | 3072 | 2730 | 4007 |
| 1 | 1 | 2 | 2306 | 2349 | 4973 | 3621 | 1878 |
| 2 | 0 | 2 | 2322 | 1428 | 1663 | 1247 | 2244 |
| 1 | 2 | 2 | 1599 | 1512 | 4831 | 3700 | 3005 |
| 4 | 0 | 2 | 2877 | 1039 | 1722 | 1687 | 393 |
| 1 | 0 | 3 | 2453 | 2626 | 2215 | 1648 | 3949 |
| †2 | 0 | 2 | | | | | |
| 1 | 2 | 3 | 2900 | 4004 | 5902 | 4725 | 3002 |
| 3 | 0 | 3 | 3720 | 547 | 3783 | 3354 | 3735 |
| 1 | 3 | 3 | 3957 | 4016 | 4232 | 3818 | 2783 |
| 2 | 3 | 3 | 2455 | 3900 | 2575 | 2474 | 1695 |

TABLE 2 (Cont.)

| h | k | ℓ | $F_{obs}$ | $F_{calc}$(RH) Arnott et al | $F_{cal}$(RH) Gupta et al | $F_{cal}$(LH) Gupta et al | $F_{cal}$(LH) (Present Model) |
|---|---|---|---|---|---|---|---|
| 1 | 0 | 4 | 1640 | 989 | 1358 | 1083 | 1866 |
| 1 | 1 | 4 | 3407 | 2381 | 2978 | 2433 | 4092 |
| 2 | 0 | 4 | 4306 | 2870 | 2762 | 2355 | 4194 |
| 1 | 2 | 4 | 4563 | 4536 | 4523 | 3959 | 3610 |
| 3 | 0 | 4 | 4952 | 2910 | 4421 | 4216 | 3229 |
| †1 | 3 | 4 | | | | | |
| 2 | 3 | 4 | 3854 | 3806 | 4781 | 4834 | 5799 |
| 1 | 2 | 5 | 5381 | 1983 | 7226 | 6958 | 3994 |
| †3 | 0 | 5 | | | | | |
| 1 | 3 | 5 | 5548 | 3825 | 3549 | 3765 | 5345 |
| 1 | 0 | 6 | 2602 | 889 | 2271 | 2264 | 4145 |
| 1 | 1 | 6 | 2478 | 1229 | 4461 | 4446 | 5536 |
| 2 | 0 | 6 | 4440 | 4215 | 4340 | 4590 | 4115 |
| 1 | 2 | 6 | 5453 | 8323 | 4009 | 4319 | 3682 |
| 1 | 0 | 7 | 9893 | 6035 | 7457 | 8512 | 6377 |
| 1 | 1 | 7 | 7790 | 6013 | 7240 | 8450 | 5609 |
| †2 | 0 | 7 | | | | | |
| 1 | 2 | 7 | 5512 | 7373 | 2275 | 2783 | 5697 |
| †1 | 0 | 9 | | | | | |
| 1 | 1 | 9 | 5327 | 6527 | 3520 | 5738 | 2575 |

TABLE 3

Cylindrical Polar Coordinates for the Proposed Left-Handed D-DNA Model

The dyad axis relating to the two nucleotides within a base-pair is at (r,0,0)

| Atom | r(A ) | $\phi$(o) | Z(A ) |
|---|---|---|---|
| $ADO'_4$ | 7.59 | 38.0 | 0.65 |
| $ADO'_2$ | 9.17 | 37.9 | -1.30 |
| $ADO'_3$ | 10.00 | 42.2 | 0.95 |
| ADP' | 8.90 | 42.3 | -0.03 |
| $ADO'_1$ | 8.62 | 52.5 | -0.33 |
| $ADC'_5$ | 7.48 | 72.0 | -2.55 |
| $ADC'_4$ | 7.80 | 67.0 | -1.23 |
| $ADO'_5$ | 7.02 | 71.8 | -0.20 |
| $ADC'_2$ | 6.24 | 52.8 | -0.72 |
| $ADC'_3$ | 7.63 | 55.9 | -1.25 |
| $ADC'_1$ | 5.98 | 63.7 | 0.19 |
| THN1 | 4.74 | 27.1 | 3.32 |
| THC2 | 3.56 | 17.3 | 3.15 |
| THO2 | 3.84 | -0.8 | 2.93 |
| THN3 | 2.43 | 32.3 | 3.23 |
| THC4 | 2.91 | 60.0 | 3.47 |
| THO4 | 2.65 | 85.3 | 3.53 |
| THC5 | 4.32 | 56.8 | 3.64 |
| THME | 5.27 | 70.3 | 3.90 |
| THC6 | 5.04 | 42.6 | 3.56 |
| ADN9 | 4.71 | 72.0 | 0.28 |
| ADC8 | 4.84 | 88.2 | 0.51 |
| ADC4 | 3.44 | 65.1 | 0.14 |
| ADN3 | 3.42 | 42.7 | -0.11 |
| ADC2 | 2.40 | 26.6 | -0.18 |
| ADN1 | 1.07 | 29.2 | -0.06 |
| ADC6 | 1.35 | 94.8 | 0.19 |
| ADN6 | 1.58 | 148.3 | 0.32 |
| ADC5 | 2.74 | 87.1 | 0.29 |
| ADN7 | 3.88 | 99.7 | 0.53 |

Table 4 : The torsion angles for the four D-DNA models

| Torsion Angle | Arnott et al RH | Gupta et al RH | Gupta et al LH | Present model LH |
|---|---|---|---|---|
| $C_4'$-$C_3'$-$O_3'$-p' | 141 | 208 | 225 | 188 |
| $C_3'$-$O_3'$-p'-$O_5'$ | 260 | 209 | 215 | 64 |
| $O_3'$-p'-$O_5'$-$C_5'$ | 298 | 302 | 263 | 211 |
| p'-$O_5'$-$C_5'$-$C_4'$ | 208 | 146 | 156 | 247 |
| $O_5'$-$C_5'$-$C_4'$-$O_3'$ | 69 | 61 | 32 | 201 |
| $C_5'$-$C_4'$-$C_3'$-$O_3'$ | 156 | 154 | 154 | 153 |
| $O_4'$-$C_1'$-$N_9'$-$C_4'$ | 82 | 72 | -4 | 205 |

observed in crystals of d(CpG(pCpG)$^2$) could, if extended to a two-stranded polynucleotide, account for a diffraction pattern observed from Na poly d(GC).poly d(GC) which was then designated as S. They also observed a transition within oriented fibres from the B-form (assumed to be right-handed) to the S-form.

The left-handed D-conformation described here differs from the Z conformation in that both the adenine and thymine nucleotides are in the anti conformation whereas in the Z conformation the purine nucleotide is syn. Transition from a right-handed helix of either the DNA A or B type to this left-handed D conformation could occur by a simple modification of the conformation as proposed by Wang et al. (1979) for the B to Z transition. This would require every base-pair to rotate through 180$^o$ about a line parallel to the line joining $N_1$ thymine to $N_9$ adenine, coupled with changes in the relative positioning of base pairs about the helix axis.

Finally, it is important to emphasise that the above discussion does not claim to prove conclusively that the D conformation is a left-handed helix. However, it does show that a satisfactory left-handed model can be proposed which is in better agreement with the observed diffraction than the hitherto accepted right-handed models.

# CHAPTER 5

## Polymorphism of Poly d(AT).Poly d(AT)

### 5.1 Introduction

The synthetic polynucleotide poly d(AT).poly d(AT) is of fundamental interest in studies which aim to elucidate the relationship between nucleic acid structure and function. In particular A-T rich regions in DNA double-helices have been cited as centres for the control of transcription of genetic information. Davies and Baldwin (1963) observed that the sodium salt of poly d(AT).poly d(AT) could assume not only the crystalline A and semi-crystalline B forms observed for naturally occurring NaDNAs (Fuller et al., 1965, Langridge, et al., 1960b, Arnott et al., 1972b, 1973), but also a new conformation which was designated as D. These authors also studied Li poly d(AT).poly d(AT) for which they observed crystalline B patterns, but not the semi-crystalline C patterns observed for naturally occurring LiDNAs (Marvin et al., 1961).

Furthermore, these authors reported that fibres of the potassium and rubidium salts of DNA, which exist in a semi-crystalline B form at high relative humidites, exhibit transformation into a crystalline A form when the humidity is reduced to 75% (Franklin and Gosling, 1953b). They claimed that the A form patterns (from certain NadAT fibres) contain little intensity on the meridian at 3.4 A suggesting that the material contains only a very small amount of B form structure. Further they reported these fibres remained in the A conformation even at 98% relative humidity.

For $NH_4$ poly d(AT).poly d(AT) Davies et al., (1963) observed a semi-crystalline C conformation and D conformation, and reported the appearance of intensity near the meridian of third and fourth layerlines.

Arnott et al. (1974b) reported that the D form was observed under conditions which would have favoured the A-form of natural DNAs. Leslie et al. (1980) did however prepare one fibre for which the A conformation was stable over a period of six months even at very high relative humidity (95%). However, they were not able to identify the conditions which favoured the A and D forms, although for one fibre which they described as 'impure synthetic DNA', they observed the D → A → B transition by raising the relative humidity of the fibre environment. No indication was given of the nature of the impurities in the fibre nor of the quality of the patterns which were obtained.

In this chapter, a study of the polymorphism of poly d(AT). poly d(AT) in the presence of LiCℓ, NaCℓ, KF, RbCℓ and CsF and varying conditions of humidity is described.

## 5.2 Materials and Methods

The poly d(AT).poly d(AT) material precipitated from different salt solutions (LiCℓ, NaCℓ, KF and CsF) was provided by Dr. J. Brahms of the University of Paris. The Na poly d(AT).poly d(AT) was obtained from Boehringer. From previous work on the polynucleotide there was evidence that the conformation assumed was influenced by the ionic content of the fibre. Therefore, preparative procedures were used which allowed the salt concentration of the fibre to be systematically varied. While the salt content of the solution of the polynucleotide is readily varied by dialysis, it is more difficult to control the level of salt in precipitated material since there is variability in the amount of salt which is "brought down" with polynucleotide. The process of centrifugation was thought to allow more quantitative control over the salt associated with the extracted material. For both precipitation and centrifugation the yield of polynucleotide decreased as the ionic strength of the original solution is reduced. In this study

the poly d(AT).poly d(AT), initially precipitated from 0.1 M (or less) salt solutions, was centrifuged at 50,000 r.p.m. from salt solutions whose ionic strength ranged from 0.001 M to 0.05 M. From measurements on the supernatant it was possible to estimate the amount of salt per $PO_4^-$ in the extracted material (as discussed in Chapter 2). There are likely to be considerable uncertainties in these estimates mainly because of non-uniformity in the distribution of salt in the gel and precipitate. Nevertheless, despite these reservations, this study and a recent parallel study on natural DNAs (see Chapter 9) represents a much more extensive systematic analysis of the effect of salt on the polynucleotide conformation than has been hitherto reported (with the exception of studies in which very much larger amounts of material were available than is normally the case for synthetic polynucleotides (Cooper and Hamilton, 1966). X-ray fibre diffraction patterns were recorded over a range of relative humidities from 0% to 98%.

## 5.3 Results and Discussion

In this section the X-ray fibre diffraction patterns obtained from the polynucleotide in the presence of various 'excess' ions are described and discussed.

### 5.3.1 Fibres which contained $Li^+$ and $C\ell^-$

Fibres prepared from low salt LiCℓ precipitate or centrifuged gel gave a semi-crystalline C-form diffraction pattern (Marvin et al., 1961) up to 92% relative humidity (Figure 1). The semi-crystalline B-form was then obtained at 98% relative humidities. These diffraction patterns are very similar to those observed from natural DNAs. A new fibre was made by adding small amounts of LiCℓ solution (~ 10 μℓ of 0.02 M LiCℓ) to these fibres. These fibres gave a "D-type" conformation (Figure 2) up to 92% relative humidities. The diffraction pattern was highly crystalline and fairly

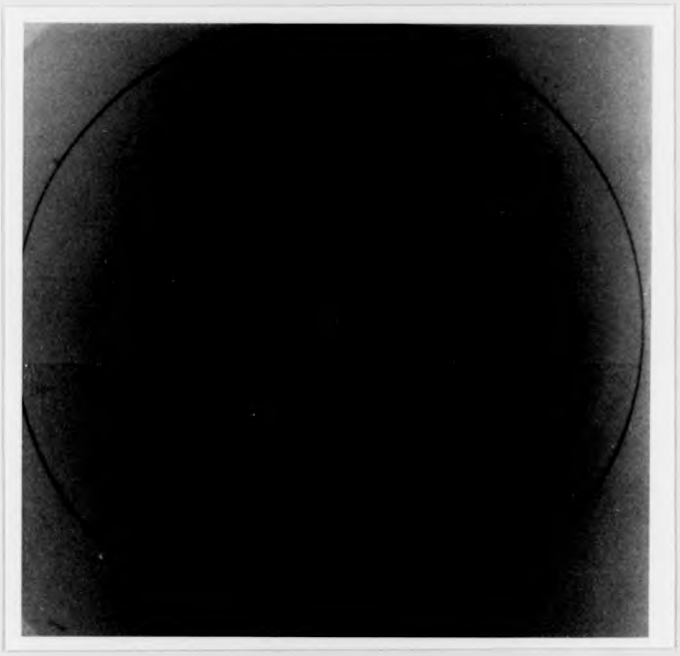


Figure 1: Semi-crystalline C-form Diffraction Pattern of Li poly d(AT).poly d(AT) at 33% R.H.

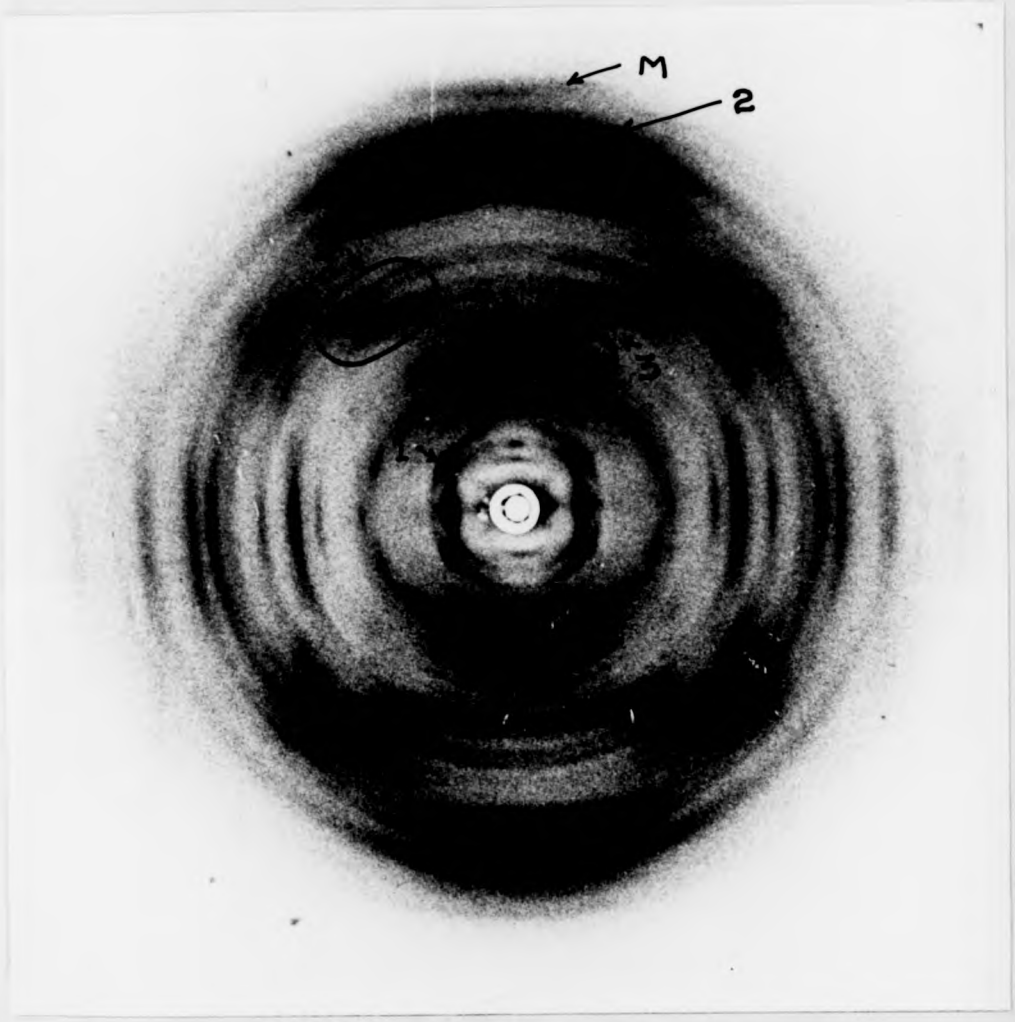



Figure 2: "D-type" form Diffraction Pattern of Li poly d(AT).poly d(AT) at 92% R.H.

well oriented. The strong reflection (1) and the near meridional reflection (2) appear to be "D-like" (see Figure 2, Chapter 4). However, apart from above similarities to the D pattern the overall intensity distribution of this pattern is quite distinct from D-DNA. Particularly the set of reflections (3) and (4) are uncharacteristic of the D conformation. The set of reflections (3) are not from the same layerline, but probably from two different layerlines. Similarly, the set of reflections (4) is probably from three different layerlines. These observations strongly suggest that the diffraction pattern on figure (2) indicates a non-integral helix similar to the 'D-type' structure. The reflections (5) and (6) in the diffraction pattern (Figure 2) show similarities to patterns described by Davies et al. (1963) for $NH_4$dAT. Their D-pattern reported for $NH_4$dAT is of poor quality, and is very difficult to compare with the pattern in figure 2. However, in their description of the D-form of $NH_4$dAT, they reported an unusual feature in that there was a near meridional reflection between the third and fourth layerlines. This is similar to the reflections (5) and (6) in figure 2. Figure 2 thus bears a greater similarity to the ammonium patterns than it does to the sodium patterns of Arnott et al. (1974). In the present work on NadAT a very similar pattern to figure 2 was observed (figure 3). The detail of this work will be discussed later in the chapter. It is most likely that the pattern reported by Davies et al. (1963) for NadAT (Plate IX, Davies, et al. 1963) is very similar to that seen in figure 3. Even though the quality of the pattern reported by Davies et al. (1963) for NadAT is poor, the set of reflections (3) can easily be identified in the Plate IX of their report. This set of reflections is not observed in the D-pattern reported by Arnott et al. (1974b) nor in those presently described for NadAT (see Chapter 4). However, since the pattern on the Plate IX (Davies, et al., 1963) is of poor qualtiy, it is difficult to come to a firm conclusion whether their pattern is the D-form or the "D-type" form.

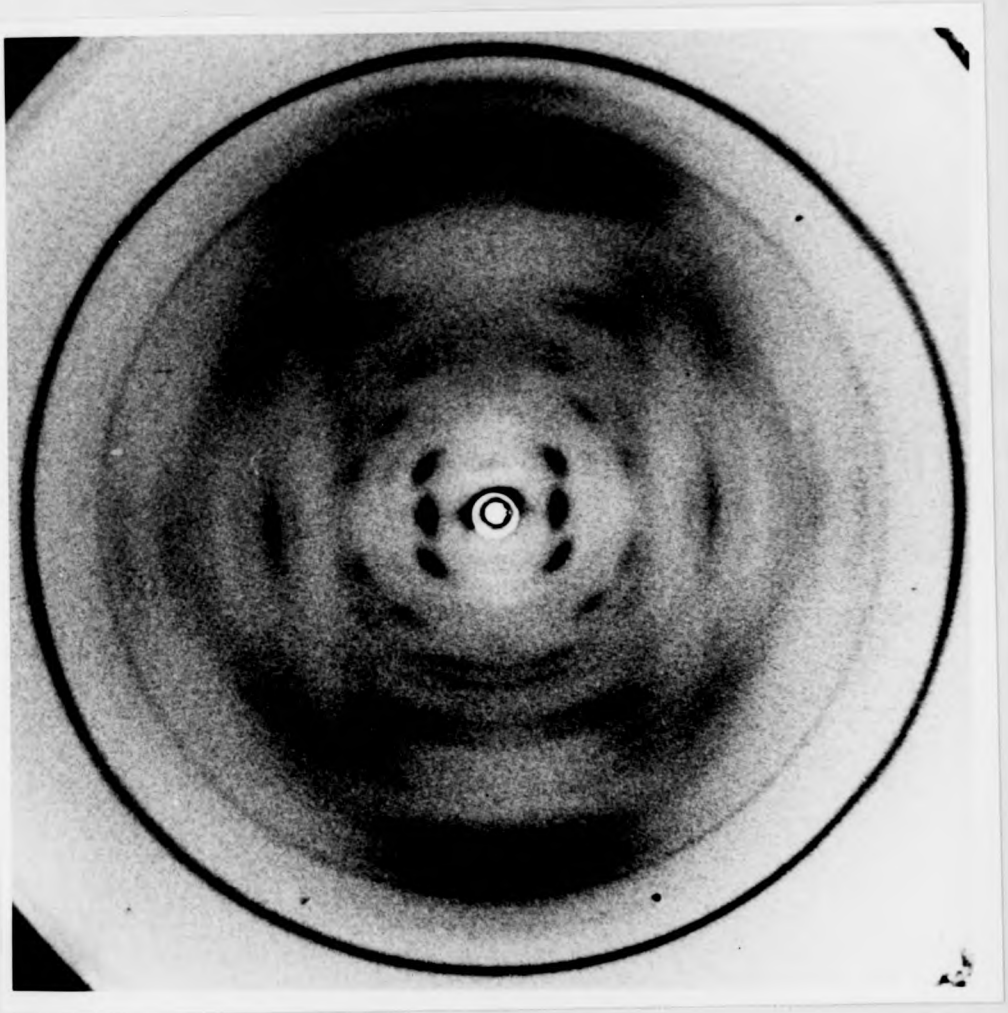


Figure 3: "D-type" form Diffraction Pattern of Na poly d(AT).poly d(AT) at 86% R.H.

The strong region (2) is resolved in weakly exposed tilted fibres as the (10) and (11) reflections rather than a meridional (00). The reflection M is most probably the meridional reflection corresponding to the one seen in D-DNA. Preliminary measurements showed that this is a non-integral helix, but is close to being an eight-fold helix. The spacing of the meridional reflection M is 3.04 A. The layerline spacing of the reflection (1) is 23.5 A. Therefore, the number of turns is $\simeq {}^{23.5}/_{3.04} \simeq 7.7$. The helix was thus determined to have $23_3$ symmetry, and the reflection A indexed as (103). The reflection (5) and (6) are meridional reflections on the 8th and 12th layerlines respectively. Similar meridional reflections are seen on even layerlines in diffraction patterns from this sequence. From the $\rho$ values of the equatorial reflection the unit cell was decided to be tetragonal. The $\rho$ values were used to refine the unit cell parameters. The calculated and observed values are given in table (1). The unit cell parameters are a = 16.9 A, c = 69.9 A and the r.m.s. deviation in the $\rho$ values is 0.01 $A^{-1}$.

In conclusion the diffraction pattern of figure 2 indicates a $23_3$ helix with a 69.9 A pitch. These parameters are very close to those obtained for the D-conformation. Neither this pattern, nor any D-patterns have been reported for Li poly d(AT).poly d(AT) previously.

Further increases in the salt content of the above figures gave a pattern which was found to be very similar to the pattern in figure 2 at low relative humidities. However, the reflection (1) is weak (see figure 2) compared to the equatorial reflection (100) from this diffraction pattern (similar observations were noted in the diffraction photographs obtained from Na poly d(IC).poly d(IC) fibres (see Chapter 4, figure 5)). At higher humidities (i.e. above 92%) a semi-crystalline B pattern was obtained. Further additions of salt to this fibre gave a semi-crystalline B-form even at lower humidities. 'Medium' salt fibres gave a crystalline

Table 1

Observed and calculated ρ values for the "D-type" pattern used in figure 2

RH = 92% a = 16.9 A c = 69.9 A

| h | k | ℓ | $\rho_o(A^{-1})$ | $\rho_c(A^{-1})$ |
|---|---|---|---|---|
| 1 | 0 | 0 | 0.059 | 0.059 |
| 2 | 1 | 0 | 0.135 | 0.133 |
| 2 | 2 | 0 | 0.170 | 0.168 |
| 3 | 1 | 0 | 0.190 | 0.188 |
| 3 | 2 | 0 | 0.216 | 0.214 |
| 4 | 1 | 0 | 0.240 | 0.245 |
| 3 | 3 | 0 | 0.254 | 0.252 |
| 4 | 2 | 0 | 0.268 | 0.265 |
| 5 | 1 | 0 | 0.301 | 0.303 |
| 1 | 0 | 3 | 0.073 | 0.073 |
| 2 | 0 | 3 | 0.127 | 0.126 |
| 1 | 1 | 6 | 0.119 | 0.120 |
| 1 | 0 | 9 | 0.139 | 0.142 |
| 1 | 0 | 10 | 0.150 | 0.155 |
| 2 | 0 | 11 | 0.193 | 0.197 |
| 1 | 1 | 12 | 0.186 | 0.191 |
| 2 | 0 | 12 | 0.203 | 0.209 |
| 2 | 1 | 12 | 0.214 | 0.217 |
| 1 | 1 | 13 | 0.198 | 0.204 |
| 2 | 1 | 13 | 0.224 | 0.229 |
| 1 | 0 | 17 | 0.259 | 0.250 |
| 1 | 1 | 17 | 0.265 | 0.257 |
| 2 | 1 | 17 | 0.282 | 0.277 |
| 2 | 1 | 21 | 0.331 | 0.329 |

B form (figure 4) at 92% relative humidity. The crystalline B-pattern was very similar to the crystalline B-form of Li DNA. The B pattern obtained at 98% relative humidity was a semi-crystalline B-form (figure 5). The detailed analysis of the crystalline B pattern is in progress in this laboratory.

5.3.2 Fibres which contained $Na^+$ and $C\ell^-$

The general method of preparation and estimation of $Na^+$ $C\ell^-$ ions per phosphate in the fibres is described in Chapter 2.

Very low salt fibres prepared from centrifuged gel (typically containing less than 0.2 $Na^+$, $C\ell^-$ ions per $PO_4$) gave the C conformation (figure 6) at lower relative humidities (typically up to 75% or 95%), above which they changed reversibly into the B-form. Fibres containing slightly higher than 0.2 $Na^+$ + $C\ell^-$ per $PO_4$ gave a C pattern at lower relative humidities (typically from 33% to 92%). At higher humidities (typically 95%) these fibres gave an A-form and then a crystalline B-form (figure 7). Further humidification of these fibres gave the semi-crystalline B-form at 98% relative humidity. However, some of these fibres did not give an A-form, but the following sequence of conformations:

C form → crystalline B-form → semi-crystalline B-form

All these transitions were reversible with respect to relative humidity. This is the first time the crystalline B-form has been observed in fibres of Na poly d(AT).poly d(AT) - the only other report was published by Leslie et al. (1980) for fibres obtained from poly d(AC).poly d(GT). The packing of the crystalline B-form for Na poly d(AT).poly d(AT) appears to be different from crystalline B-forms of Li DNA. Further analysis of this pattern is in progress in our laboratory. Fibres estimated to contain about 0.6 $Na^+$ and $C\ell^-$ pairs per $PO_4$ gave an A pattern at low relative

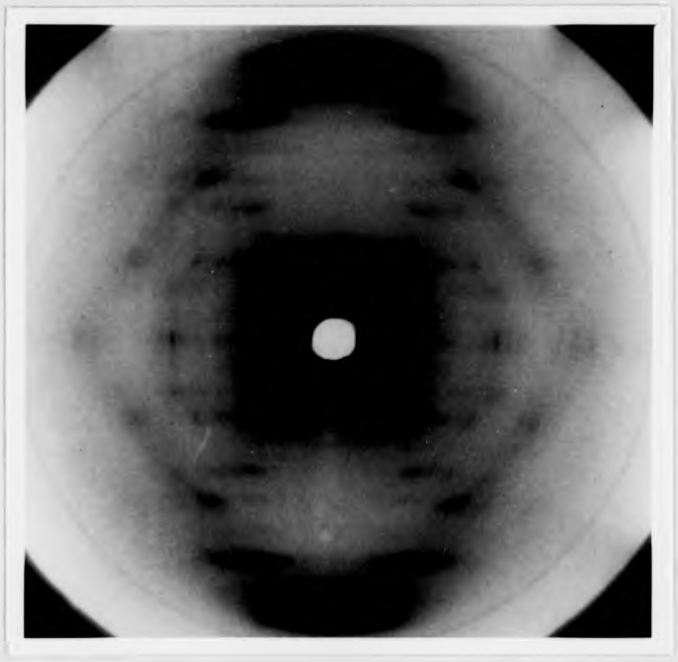


Figure 4: Crystalline B-form Diffraction Pattern of Li poly d(AT).poly d(AT) at 92% R.H.

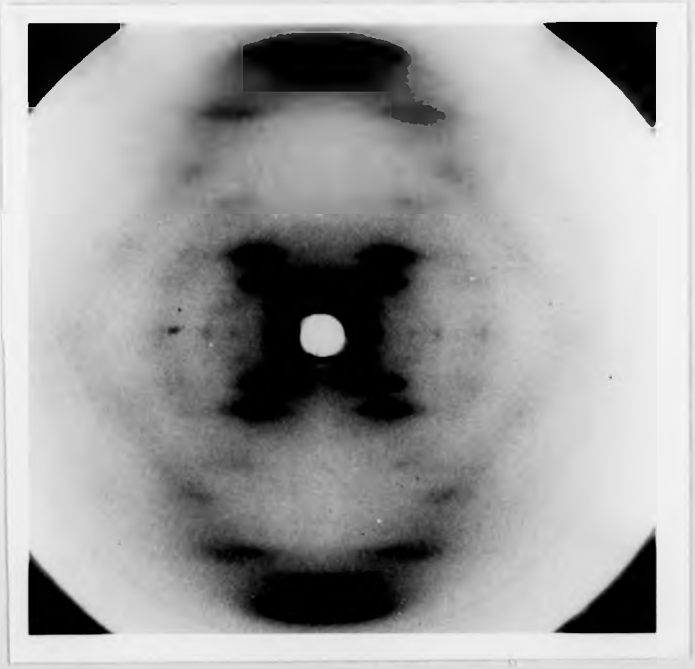


Figure 5: Semi-crystalline B-form Diffraction Pattern of Li poly d(AT).poly d(AT) at 98% R.H.

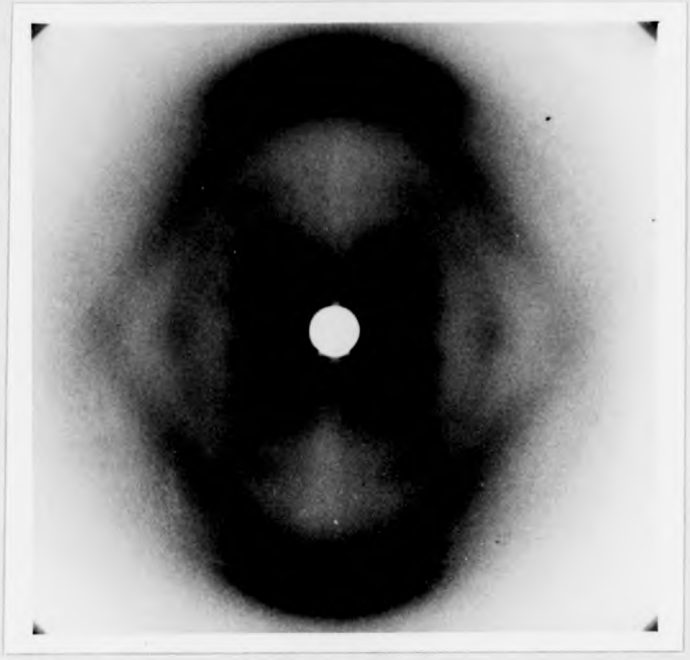


Figure 6: Semi-crystalline C-form Diffraction Pattern of Na poly d(AT).poly d(AT) at 75% R.H.

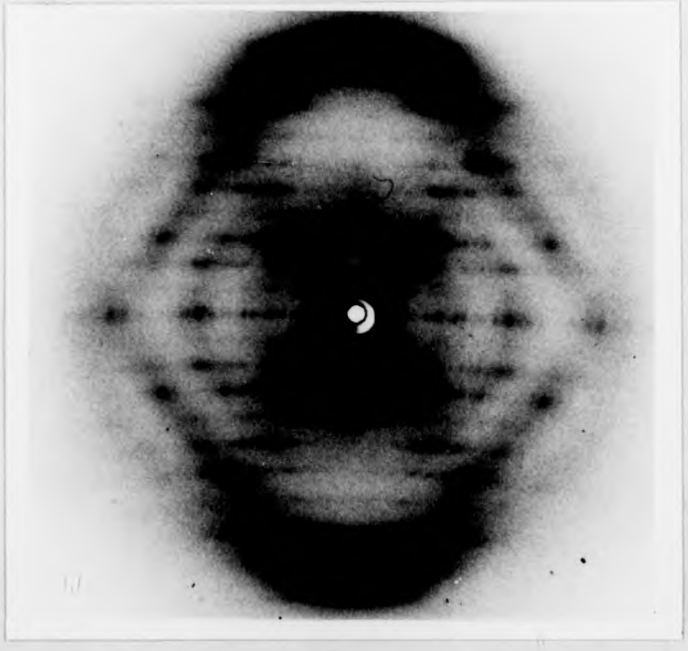


Figure 7: Crystalline B-form Diffraction Pattern of Na poly d(AT).poly d(AT) at 75% R.H.

humidities (typically from 33% to 75%), but such patterns did show evidence of the C conformation. At higher relative humidities (typically 86% to 95%), patterns showed a mixture of the A and B forms (figure 8). In these patterns, the two components of the mixture are both very much more clearly defined within the one diffraction pattern than has been reported previously. The B component on these patterns is somewhat different from the B form observed from calf thymus DNA. The base separation or rise per nucleotide, the pitch and the inter-molecular separation are 3.28 A, 32.8 A, 24.2 A respectively. These parameters are very similar to those of B'-DNA which was reported for the homopolymer poly (dA).poly (dT) (Arnott et al., 1974a and this work (Chapter 10)).

Prolonged exposure at higher relative humidities (typically 95%) resulted in the polynucleotide undergoing a transition to the D-form. Exposure at humidities (98%) resulted in a transition from the D-form to the B-form. Reduction of the relative humidity below 98% resulted in a reversion to the D-form which then persisted even when the relative humidity was reduced to 33%. Other fibres with concentrations of NaCℓ in this range assumed the D-form over a period of a few months without being exposed to high humidities. For fibres estimated as containing rather smaller amounts of NaCℓ (typically 0.4 $Na^+$ and $Cℓ^-$ ions per $PO_4$) the sequence of transitions as a function of relative humidity was rather similar, except that at low relative humidities the pattern obtained was a mixture of A and C types (figure 9a) which then changed to an A pattern (figure 9b) at relative humidities in the range of 66% to 75%. A transition from an A pattern to a mixture of A and B patterns also occurred at rather higher relative humidities. As stated, at lower humidities (typically 33% to 57%) an A/C mixture type diffraction pattern was obtained (figure 10) from low salt fibres. The diffraction photographs were thought to indicate two 'phases' (A and C in roughly equal amounts within the fibre, see

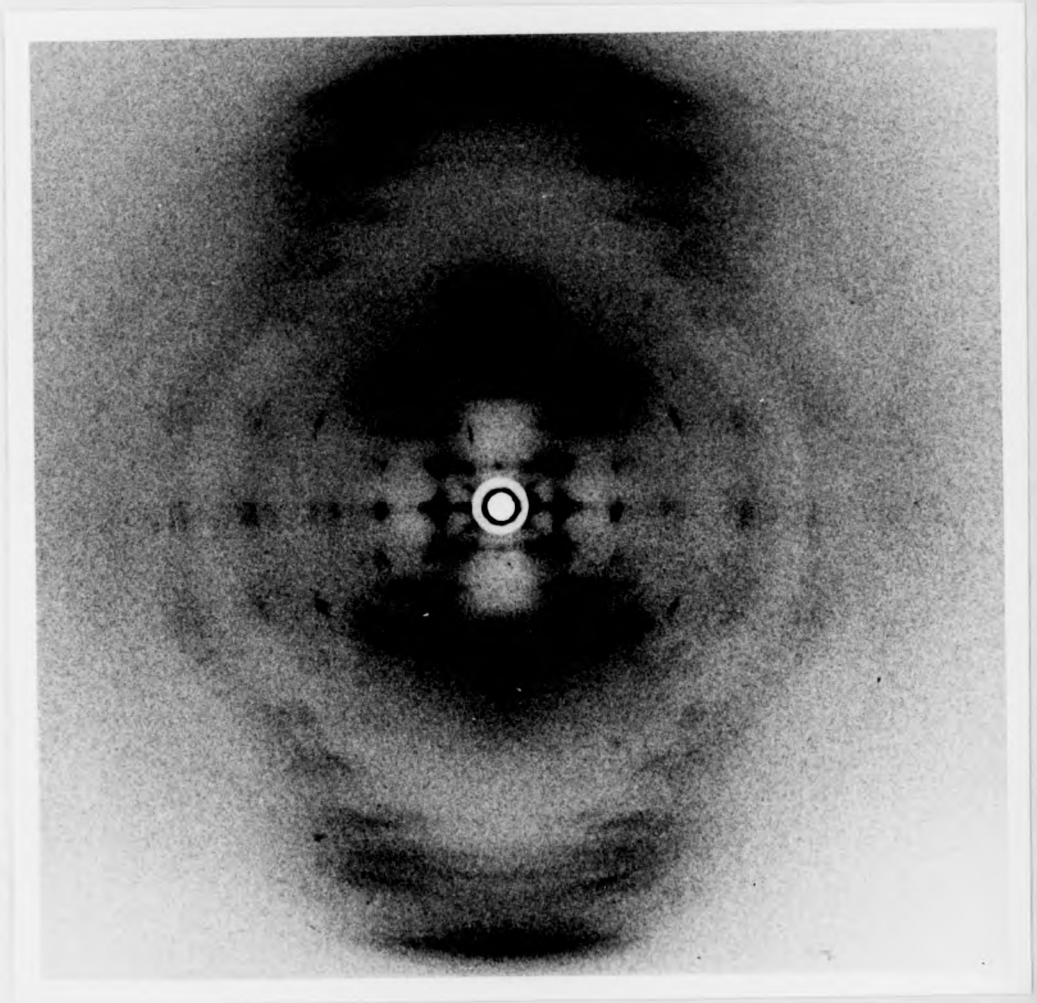


Figure 8: Mixture of A and B forms. Diffraction Pattern of Na poly d(AT).poly d(AT) at 95% R.H.

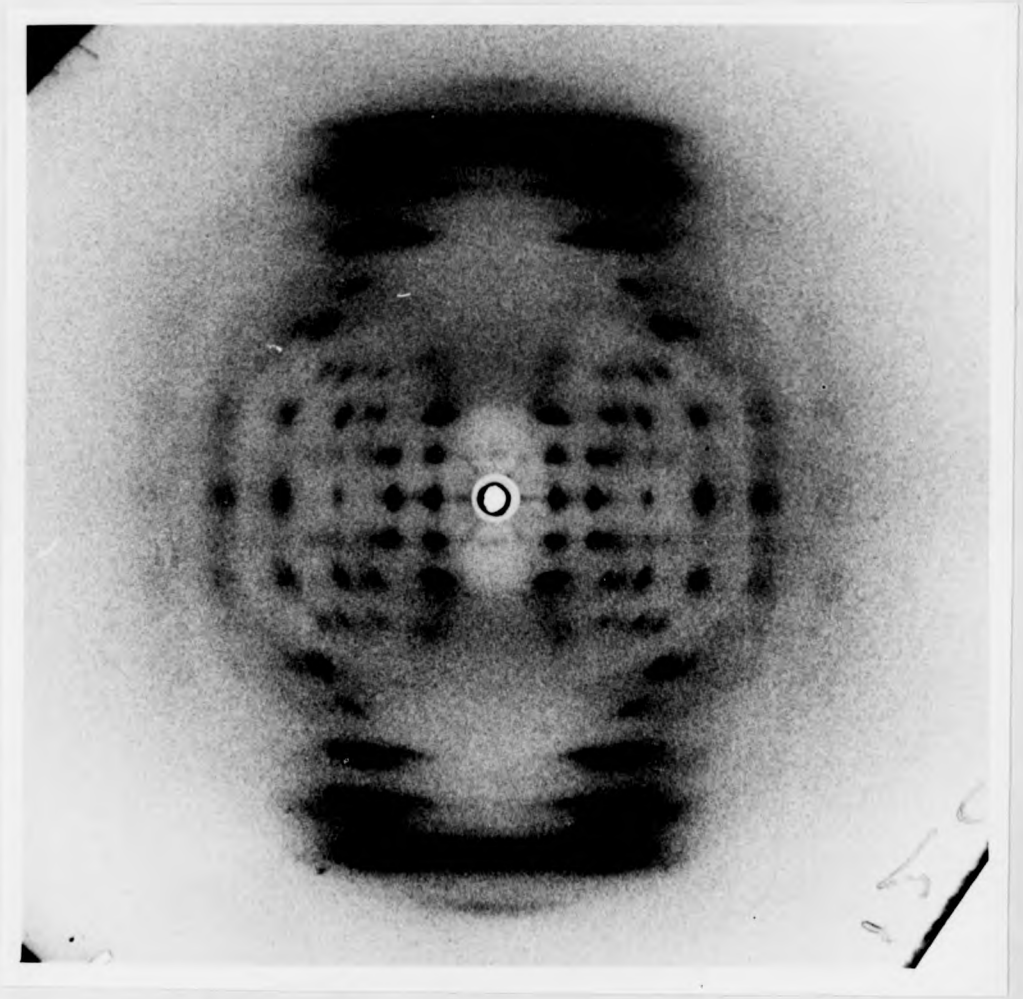


Figure 9a: Mixture of A and C forms. Diffraction Pattern of Na poly d(AT).poly d(AT) at 57% R.H.

Chpater 6). The relative 'concentration' of the A-form was seen to increase (at the expense of the C-form) as the salt concentration was increased. The effect is clear when one compares figure 10 (low salt) to figure 9a (slightly higher salt).

One of the fibres prepared from low salt concentrations (just above 0.3 $Na^+$, $C\ell^-$ per $PO_4$) gave very interesting conformational changes with relative humidity. The C-form diffraction (figure 11a) pattern was observed at 33% relative humidity. This pattern indicates a pitch of 31.0 A, and the innermost region of the first layerline is weak as expected for the C-conformation. As the relative humidity of the environment was increased, a continuous reduction in the pitch of the helix was noticed, along with a simultaneous increase in the intensity of the first layerline reflection. At 92% relative humidity the pitch of the pattern is 26.5 A (figure 11b), and the pattern is semi-crystalline. The next exposure at 98% relative humidity gave a diffraction pattern indicating an A/B mixture (similar to figure 8). A well oriented and highly crystalline D-form was obtained in the very next exposure at 95% relative humidity (see figure 4, Chapter 4). The fibre remained in the D-form even when the relative humidity was reduced from 92% to 33%. The diffraction pattern of figure 11b shows yet another morphological strain associated with poly d(AT).poly d(AT) fibres. It is very closely related to the C or D conformations, although the information provided by this pattern does not afford a clear classification of the conformation. A similar pattern to figure 8 was also obtained by Dr. N.J. Rhodes in this laboratory (figure 12) for Na poly d(AT).poly d(AT) at 98% relative humidity. However, this pattern has much more detail than the previous pattern (figure 8) and simultaneously shows the presence of the A, B and B' conformations reference to the equator of this pattern shows reflections indicating spacings of 21.5 A (characteristic of B'-DNA), and of 24.6 A (characteristic

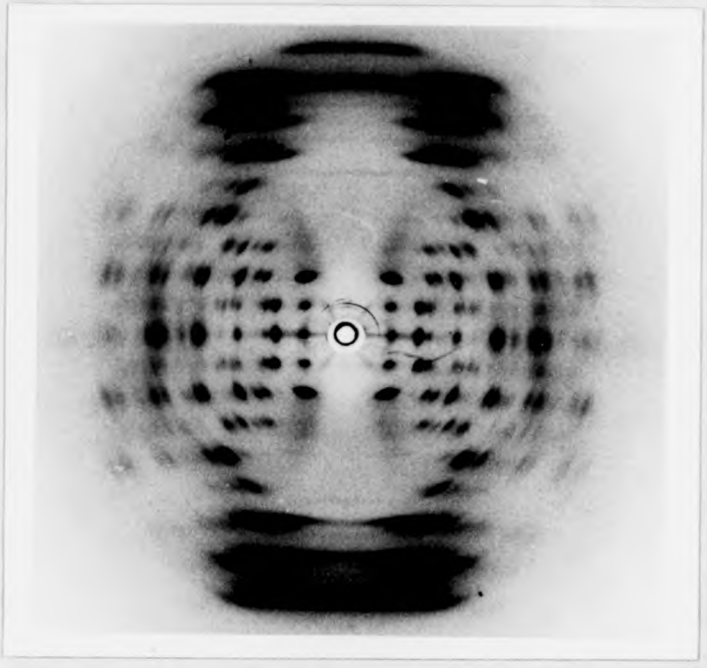


Figure 9b: Crystalline A-form Diffraction Pattern of Na poly d(AT).poly d(AT) at 92% R.H.

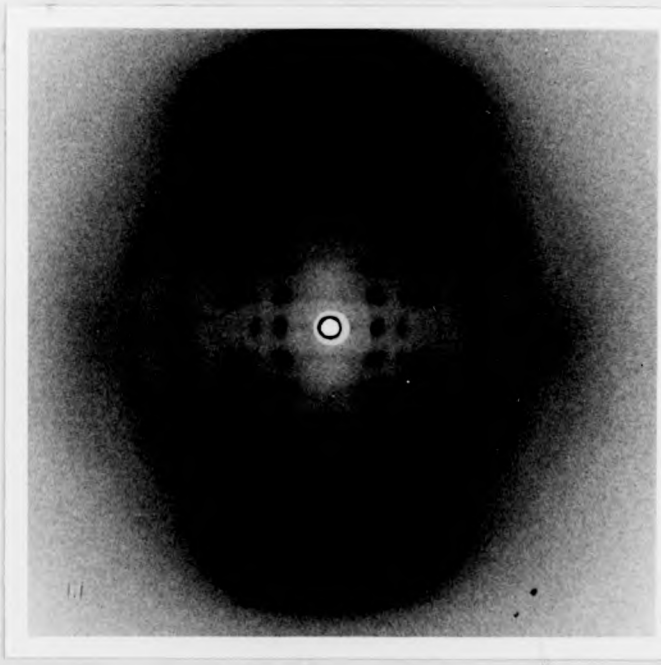


Figure 10: Mixture of A and C forms. Diffraction Pattern of Na poly d(AT).poly d(AT) at 57% R.H.

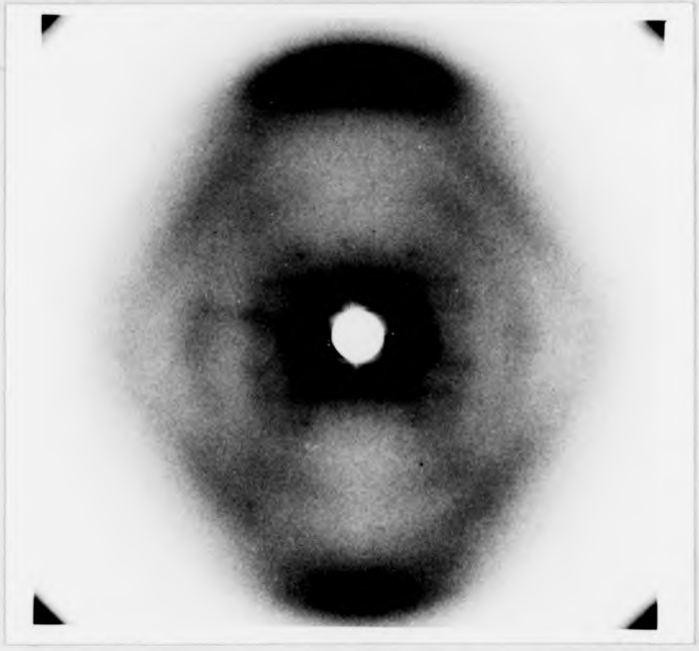


Figure 11a: Semi-crystalline C-form Diffraction Pattern of Na poly d(AT).poly d(AT) at 33% R.H.

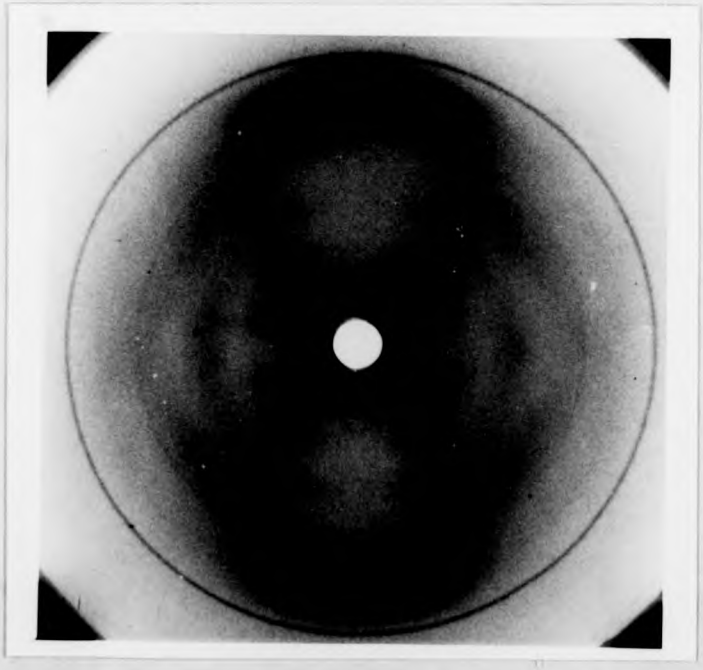


Figure 11b: Semi-crystalline C or D type form Diffraction Pattern of Na poly d(AT).poly d(AT) at 92% R.H.

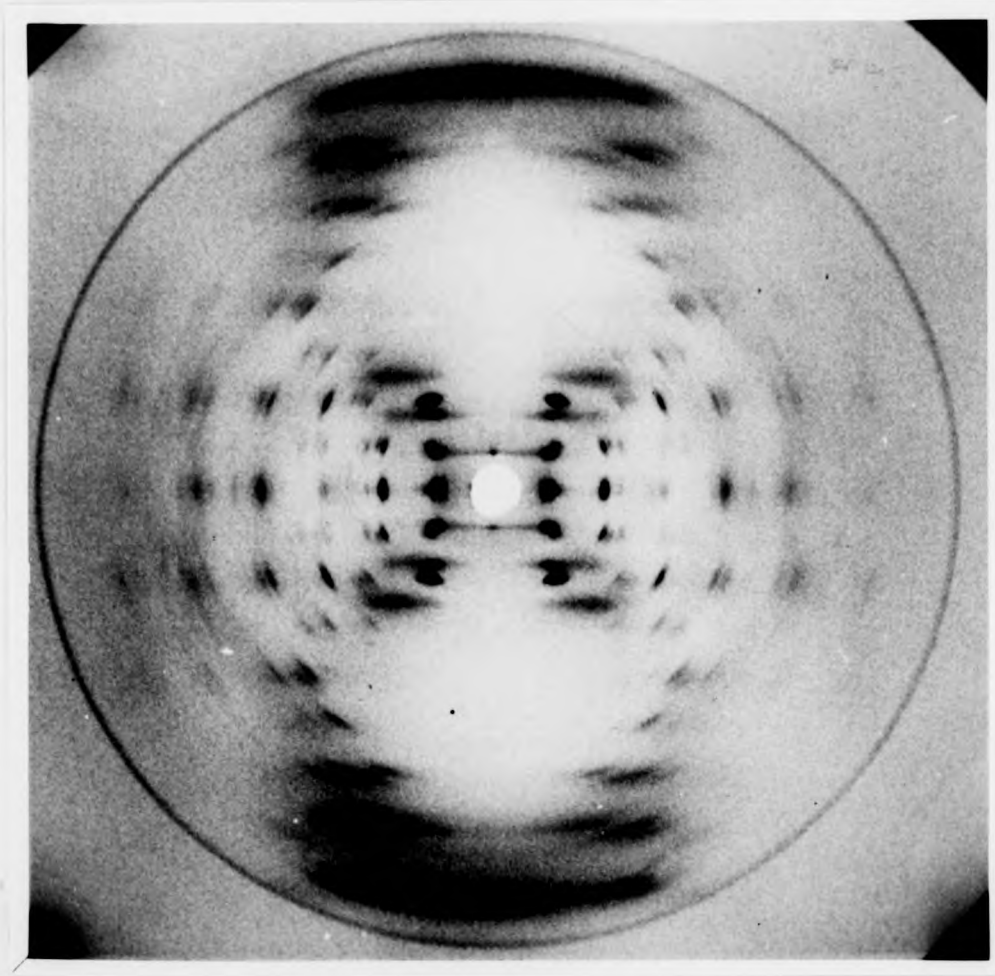


Figure 12: Mixture of A and B forms. Diffraction Pattern of Na poly d(AT).poly d(AT) at 95% R.H.

of B-DNA). The next exposure obtained from this fibre (SP16) at 98% relative humidity gave a semi-crystalline B-form. The equatorial spacing on this pattern is 25.7 A. When the relative humidity of the fibre was reduced to 95% a well characterised B'-pattern was obtained (figure 13). Reduction of the relative humidity to 92% gave an A/B mixture diffraction pattern very similar to the pattern in figure 8. In the following exposure an A pattern was obtained at the same relative humidity. This fibre stayed in the A conformation even after a few months. The fibre was lost during rewetting. No D-form was observed from the above fibre (this material was precipitated by ethanol from 0.1 M NaCℓ solution). Another fibre was prepared from material which was dialysised against $10^{-4}$ M NaCℓ solution. A D-pattern was obtained from this fibre (SP14) after a few months. Therefore, it is reasonable to say that the A-form is stable in the salt conditions which prevailed for the fibre SP16.

It is thought possible that the A-form changed to the D-form in the fibre SP14 because the amount of salt was reduced in dialysis, causing destabilisation of the A-form in the fibre. The following experiment was carried out to verify this idea. A small portion of the material was left after making the fibre SP16. Since the material was just enough to make a fibre it is difficult to remove the excess salt by precipitation or centrifugation. Therefore, a new fibre SS15 was made by adding a low salt centrifuged gel from a different batch of Na poly d(AT). poly d(AT). Thus the $Na^+ + C\ell^-$ per $PO_4$ in the fibre was thought to be lower in this sample as compared with the original material. The conformational transitions seen in the fibre (observed as a function of relative humidity) were noted to be very similar to those associated with the fibres described earlier in this section, which were believed to contain $\sim$ 0.5 $Na^+ + C\ell$ per $PO_4$, i.e. the transitions seen were:-

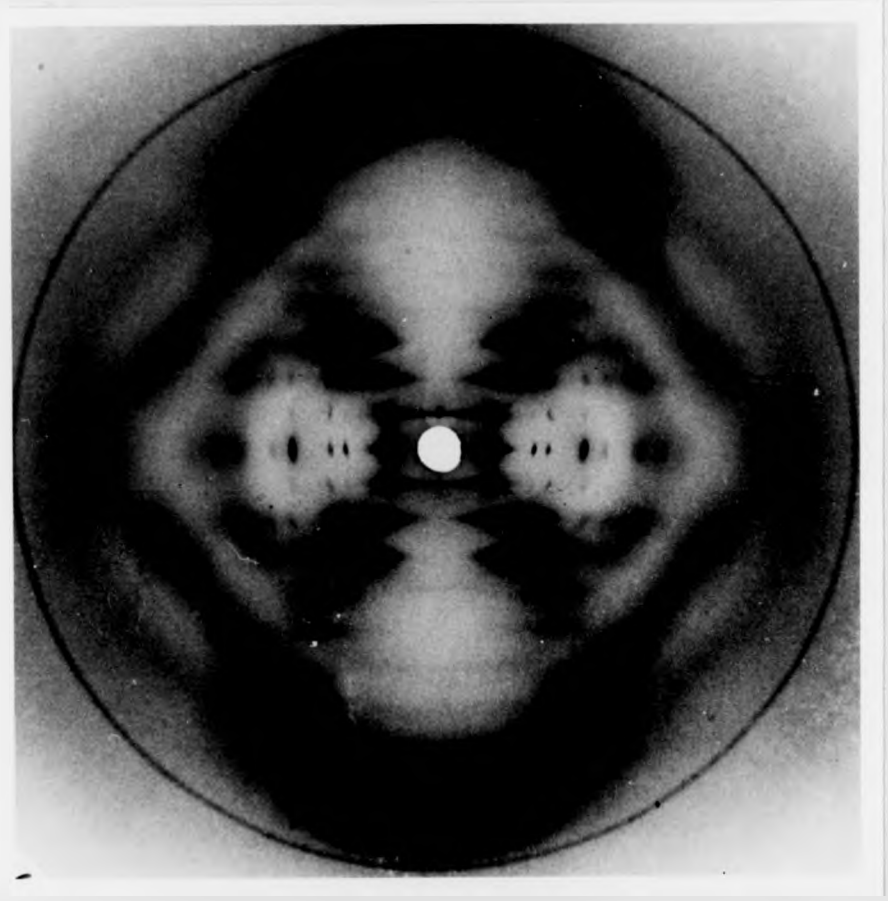


Figure 13: Semi-crystalline B'-form. Diffraction Pattern of Na poly d(AT).poly d(AT) at 95% R.H.

$$(A/C) \rightarrow A \rightarrow (A/B) \rightarrow D \rightarrow B$$

The observation of the B' conformation for poly d(AT).poly d(AT) is of significance in the explanation of the semi-crystalline B-form associated with A-T rich DNAs. The intensity of the first and third layerlines on the B' pattern observed for Na poly d(AT).poly d(AT) (figure 13) are enhanced as compared with the second layerline but this is reversed for the semi-crystalline B pattern of calf thymus DNA . The variation in the X-ray pattern intensities for the B-form have been correlated with the base composition (Bram (1973)). After a detailed study Premilat et al. (1975) concluded that the conformation of the sugar-phosphate chain is not affected by the base composition and that the proportion of (G+C) bases cannot be detected by X-ray analysis. Leslie et al. (1980) reported that they did observe intensity variation analogous to those reported by Bram and Tougard (1972) in B patterns from specimens which are oriented. Further, they explained that the different semi-crystalline packing arrangements of isomorphous molecules would produce the observed modulation of intensity. However, from the present work, it is possible to provide an explanation for these observations. The A-T rich DNAs probably have long segments of alternating sequence, and thus are more likely to attain the B' conformation in these regions, as occurs in Na poly d(AT).poly d(AT). Therefore this sort of pattern could end up being superimposed on the normal B pattern. As the A-T content in the DNA increases the contribution from the B' conformation to the semi-crystalline B patterns (from these DNAs) is likely to increase. Therefore, it is conceiveable that as the A-T content in these DNAs increase, the intensity on the first and third layerlines, as compared to the second layerline will also increase with a simultaneous decrease in the pitch towards $\sim$ 33 A.

In conclusion, the intensity variation in A-T rich DNAs is more likely to be due to the structural variation associated with alternating

A-T rich regions as reported by Bram (1973), Bram and Tougard (1972), David Goodwin (1977) and Dr. C. Nave (unpublished) rather than to the molecular packing as suggested by Leslie et al. (1980).

Further work was done on Na poly d(AT).poly d(AT) fibres with increased salt content in the fibres. A new fibre was made by adding 5 ml of 0.01 M NaCℓ solution to a fibre which was prepared from ethanol precipitated material (from 0.1 M NaCℓ solution). The original fibre gave an A conformation at low humidity, but the new fibre gave a "D-type" diffraction pattern with a strong NaCℓ diffraction ring (figure 3). As discussed earlier this diffraction pattern is very closely related to D-pattern. However, the whole intensity distribution is quite different from D, and suggests that structure is a non-integral helix. The helical parameters of these patterns are very similar to the "D-type" pattern which is observed for Li poly d(AT).poly d(AT) (figure 2). The crystallinity of the Li dAT pattern is much better than the NadAT pattern. The fibre remained in the "D-type" pattern up to 86% relative humidity. As the relative humidity was increased the intensity of the diffracted salt ring decreased. This observation strongly suggests that the excess salt crystallites on the fibre surface gradually migrated into the fibre and thus increased the salt distribution inside the fibre. At 92% and 98% relative humidities a mixture of A and "D-type" patterns were observed (figure 14) with no salt ring. Further humidification at 98% gave an A pattern and then a B pattern. As the relative humidity was reduced gradually to 33%, neither the salt ring nor the "D-type" diffraction pattern were restored. The fibre remained in the A-conformation. This provides evidence that the "D-type" conformation is favoured in the lower salinity conditions than the A conformation. The "D-type" pattern and the salt ring were recovered when the fibre was re-wetted and the new fibre was drawn. However, this time the fibre was not taken to higher humidities,

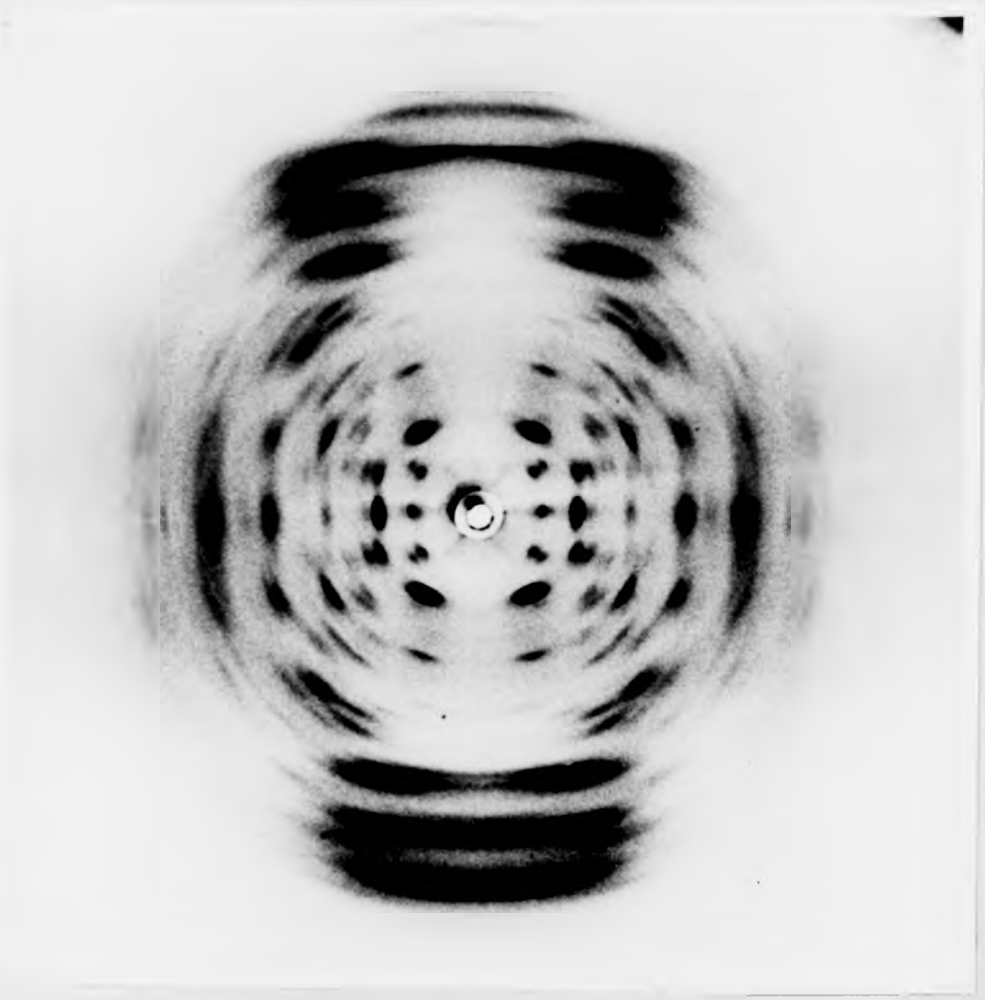


Figure 14: Mixture of A and "D-type" forms. Diffraction Pattern of Na poly d(AT).poly d(AT) at 92% R.H.

but left for a few days at room humidity before a diffraction pattern was taken. The pattern was an A pattern with no salt ring. This observation implies that even at room humidity the excess salt slowly migrated into the fibre and induced the transition from "D-type" conformation to "A" conformation. Another fibre was prepared from a small fraction of the same precipitated material by adding 2.5 μℓ of 0.1 M NaCℓ solution. The diffraction pattern from this fibre was taken after a few days. It showed a mixture of A-pattern and "D-type" pattern, and had a weak NaCℓ salt ring. As the relative humidity was increased the crystallinity of the pattern improved and a well characterised, crystalline mixture of the A-pattern and the "D-type" pattern was obtained at 98% relative humidity (figure 14). In this pattern the strong near meridional reflection of the "D-type" pattern is well resolved and clearly indicates that this reflection is the (10) rather than (00). Further humidification of the fibre gave an A-pattern and then a semi-crystalline B pattern.

### 5.3.3 Fibres which contain $K^+$ and $F^-$

The general method of preparation of fibres has already been described in Chapter 2.

The fibres were prepared from material which was precipitated from 0.1 M KF solution. A crystalline A pattern was observed up to 75% relative humidity and further humidification at 92% and 95% gave a pattern showing an A/B mixture. A semi-crystalline B form was observed at 98% relative humidity. All these transitions were found to be reversible with relative humidity. The general polymorphic behaviour of K poly d(AT).poly d(AT) is very similar to that described for NaCℓ fibres. However, some fibres gave a crystalline D-form having a tetragonal unit cell (figure 15) for relative humidities up to 75%. At 92% relative

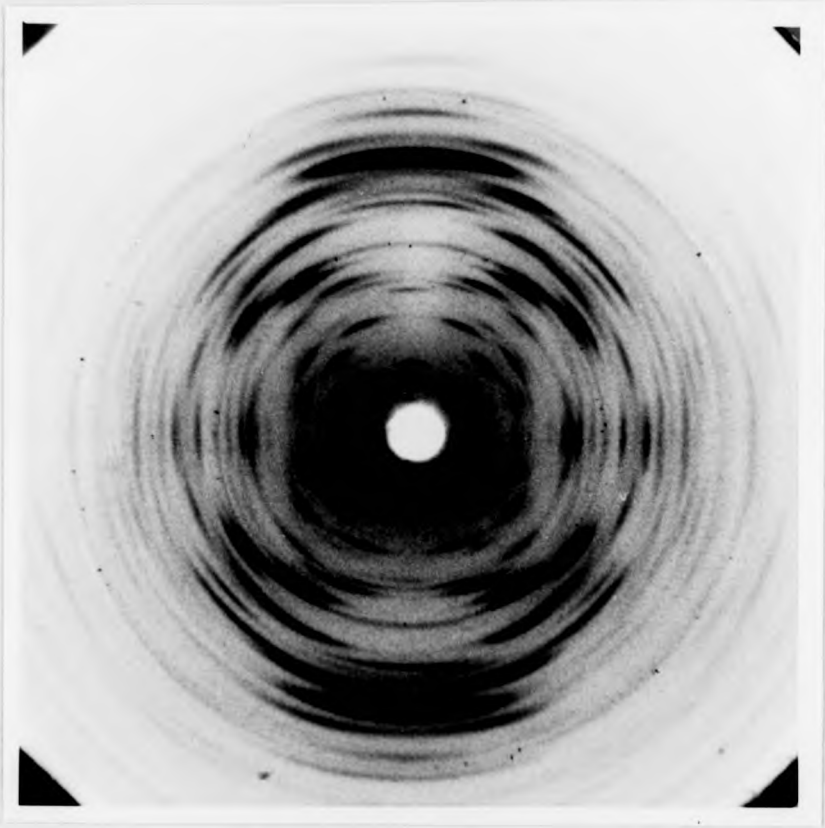


Figure 15: Tetragonal D-form. Diffraction Pattern of K poly d(AT).poly d(AT) at 75% R.H.

humidity, the D-form was observed with hexagonal unit cell (figure 16). A semi-crystalline B pattern (figure 17) was observed at 98% relative humidity. All these transitions were found to be reversible. Apart from the hexagonal D pattern, all the other patterns are identical to patterns obtained from Na poly d(AT).poly d(AT). At 84% relative humidity a mixture of two types of D pattern was obtained, and it was found that the layerline spacing of these two D patterns are slightly different. The two types of pattern were measured up.

The pitch and the unit cell parameters of the tetragonal D-form were found to be 24.0 A, a = 17.2 A and c = 24.0 A respectively. This is the first report of a hexagonal D-form from poly d(AT).poly d(AT). Earlier it has been reported for poly d(AAT).poly d(ATT), (Selsing et al., 1975), poly d(AC).poly d(IT), poly d(AIT).poly d(ACT) and poly d(AIC).poly d(ICT) (Leslie et al., 1980). The pitch and the unit cell parameters of the hexagonal D-form are 25.7 A a = 21.4 A, c = 25.7 A respectively. The observed and calculated $\rho$ values for the tetragonal D-form and the hexagonal D-form are given in tables 2 and 3 respectively. Further detailed analysis is needed to establish the structural similarities of these two D-patterns.

### 5.3.4 Fibres which contain $Rb^+$ and $C\ell^-$

These fibres were prepared as described in Chapter 2.

The polymorphic behaviour of the fibres prepared from RbCℓ were remarkably different from that of the Na poly d(AT).poly d(AT) fibres. Neither the C nor the A forms were observed for Rb poly d(AT).poly d(AT) fibres. Only the D-pattern (figure 19) was observed for relative humidities up to 95% and a semi-crystalline B-form (figure 18) was observed at 98% relative humidity. However, the qualtiy of the D-pattern was reduced upon reversal. The D-pattern was found to be stable in a wide range of salt concentrations. It was evident that some of the D-patterns contained

Table 2

Observed and calculated ρ-values for the tetragonal D-DNA pattern used in figure 15

RH = 75% a = 17.2 A c = 24.0 A

| h | k | ℓ | $\rho_o(A^{-1})$ | $\rho_c(A^{-1})$ |
|---|---|---|---|---|
| 1 | 0 | 0 | 0.057 | 0.058 |
| 2 | 0 | 0 | 0.116 | 0.116 |
| 1 | 0 | 1 | 0.071 | 0.071 |
| 1 | 1 | 1 | 0.094 | 0.092 |
| 2 | 0 | 1 | 0.123 | 0.123 |
| 1 | 2 | 1 | 0.136 | 0.136 |
| 0 | 0 | 2 | 0.083 | 0.083 |
| 1 | 0 | 2 | 0.102 | 0.101 |
| 1 | 1 | 2 | 0.118 | 0.117 |
| 2 | 0 | 2 | 0.142 | 0.143 |
| 4 | 0 | 2 | 0.243 | 0.247 |
| 1 | 0 | 3 | 0.137 | 0.138 |
| 1 | 1 | 3 | 0.149 | 0.149 |
| 3 | 0 | 3 | 0.217 | 0.214 |
| 3 | 1 | 3 | 0.225 | 0.222 |
| 0 | 0 | 4 | 0.166 | 0.167 |
| 1 | 0 | 4 | 0.175 | 0.176 |
| 1 | 1 | 4 | 0.185 | 0.186 |

Table 3

Observed and calculated ρ-values for the hexagonal D-DNA pattern used in figure 16

RH = 92% a = 21.4 A c = 25.7 A

| h | k | ℓ | $\rho_o(A^{-1})$ | $\rho_c(A^{-1})$ |
|---|---|---|---|---|
| 1 | 0 | 0 | 0.053 | 0.054 |
| 1 | 0 | 1 | 0.066 | 0.066 |
| 1 | 1 | 1 | 0.101 | 0.101 |
| 2 | 0 | 1 | 0.115 | 0.115 |
| 1 | 0 | 2 | 0.093 | 0.095 |
| 1 | 1 | 2 | 0.121 | 0.121 |
| 2 | 0 | 2 | 0.136 | 0.132 |
| 1 | 0 | 3 | 0.129 | 0.128 |
| 1 | 1 | 3 | 0.149 | 0.149 |
| 2 | 0 | 3 | 0.158 | 0.159 |

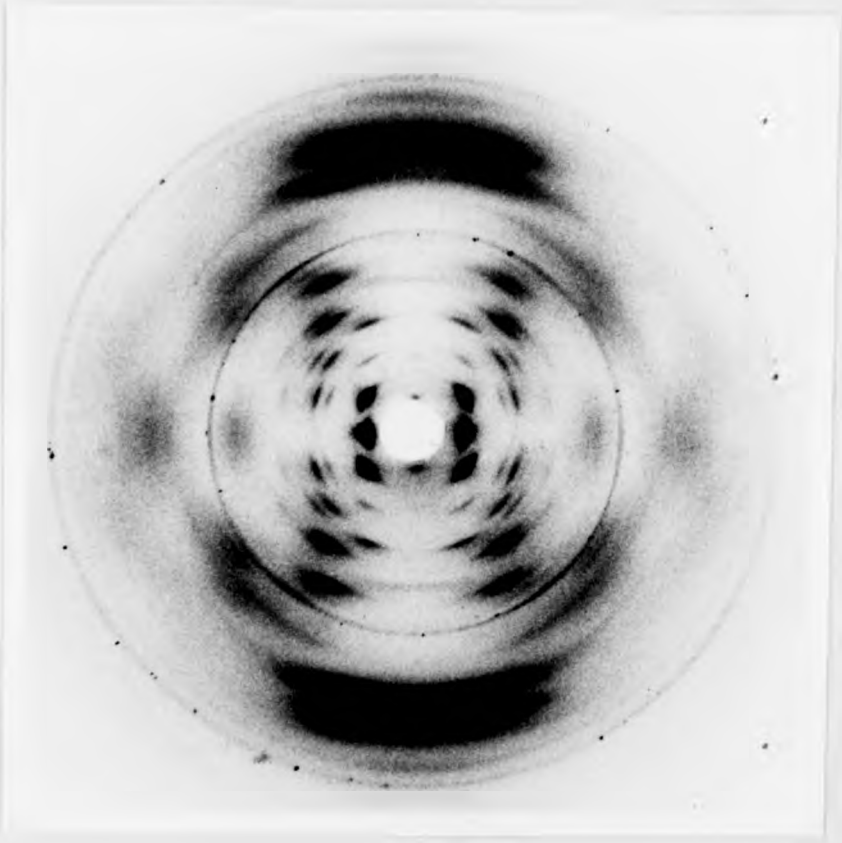


Figure 16: Hexagonal D-form. Diffraction Pattern of K poly d(AT).poly d(AT) at 92% R.H.

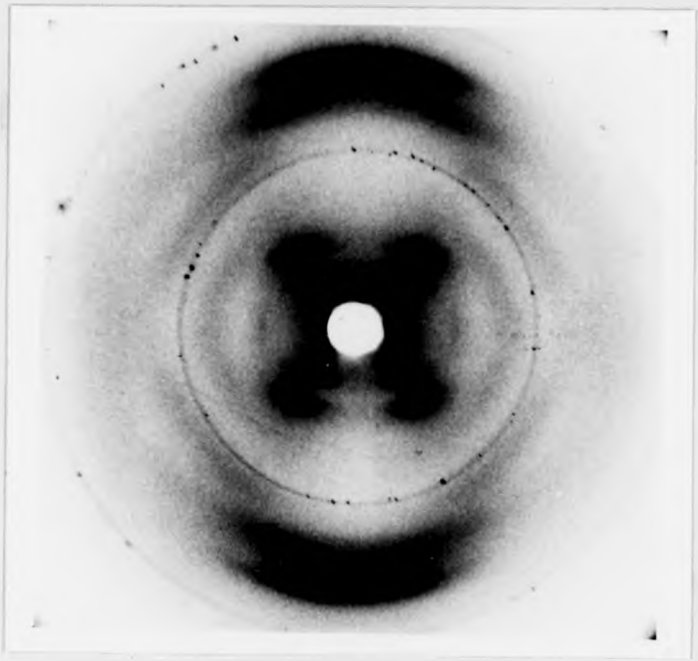


Figure 17: Semi-crystalline B-form. Diffraction Pattern of K poly d(AT).poly d(AT) at 98% R.H.

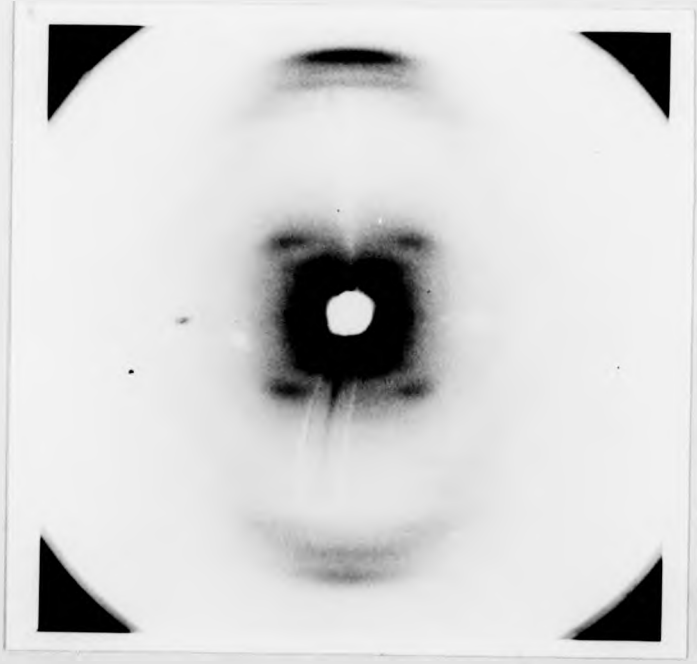


Figure 18: Semi-crystalline B-form. Diffraction Pattern of Rb poly d(AT).poly d(AT) at 98% R.H.

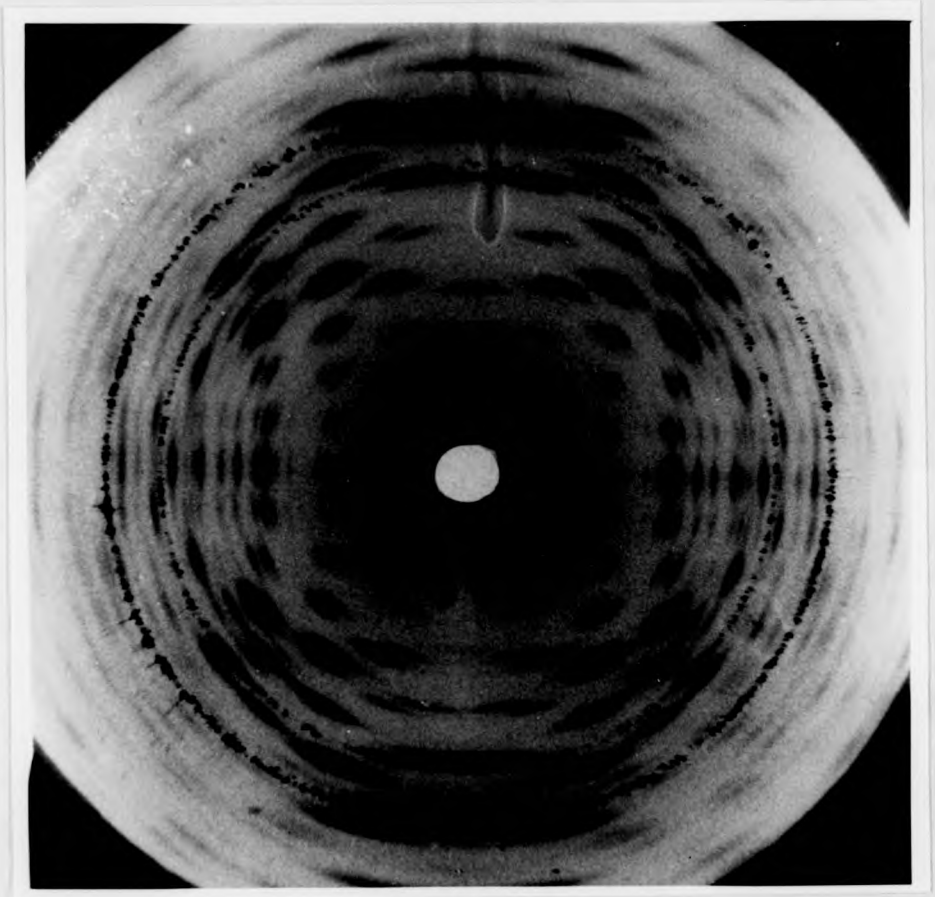


Figure 19: D-form. Diffraction Pattern of Rb poly d(AT).poly d(AT) at 66% R.H.

the RbCℓ diffraction rings even at the higher relative humidities. The crystallinity and the orientation of the D pattern was very good and it is possible to see Bragg reflections up to 2.5 A. The D-patterns from Na poly d(AT).poly d(AT) and Rb poly d(AT).poly d(AT) are very similar. Therefore, it may be possible to locate the position of $Rb^+$ in the D-structure by comparing the diffracted intensities of the two D-patterns. A similar technique was used to study the position of $Cs^+$ in the B-DNA structure by Skuratovskii and his co-workers (Skuratovskii et al., 1979).

5.3.5 Fibres which contain $Cs^+$ and $F^-$

These fibres were prepared as described in Chapter 2.

The polymorphic behaviour of Cs poly d(AT).poly d(AT) fibres was found to be very similar to Rb poly d(AT).poly d(AT). However, high salt fibres gave a new type of D-pattern for relative humidities up to 92%, and a semi-crystalline B pattern was observed at 98% relative humidity, and this transition was found to be reversible. The new D-pattern has been called D'-DNA (figure 3, Chapter 7) and is discussed in Chapter 7.

The low salt fibres gave D patterns which are similar to the D-pattern obtained for Na poly d(AT).poly d(AT) up to relative humidities of 92%. A semi-crystalline,B pattern was obtained at 98% relative humidity. This transition was found to be reversible.

In the present work, the A form was not observed for either Rb poly d(AT).poly d(AT) or Cs poly d(AT).poly d(AT). It was found that the D-form was stable in the presence of the heavy ions such as $Rb^+$ and $Cs^+$. Therefore, it may be possible that in biological systems the positively charged large protein groups provide similar environments in such a way that the molecules adopt the D-conformation. Therefore, the D-form may be a biologically important conformation.

# CHAPTER 6

## The Mixture of (A/C) and (A/B) Conformations

### 6.1 Introduction

DNA and its synthetic analogues can assume a number of distinct and regular conformations. X-ray diffraction studies on oriented fibres have shown the conformational variations induced by changes in ambient relative humidity, the cation species present and the amount of retained salt. In the first systematic study of the effect of ambient relative humidity on the X-ray diffraction patterns of fibres of NaDNA (Franklin and Gosling, 1953b), it was found that the patterns of the sodium salt of DNA changed reversibly with changes in the water content of the fibres. A highly crystalline form (figure 1) was observed at 75% relative humidity and a semi-crystalline B-form was observed at 92% relative humidity and above. The first molecular model of the A-form was also observed for the synthetic polynucleotides containing guanine except poly d(A-C).poly d(C-T) (Leslie et al., 1980). Most of the A patterns published up to now are very similar to figure 1.

In figure 1 the equatorial reflection X (at $\sim 0.054\ A^{-1}$) and meridional reflection M (at $\sim 0.3\ A^{-1}$) were regarded as being due to an impurity (or mixture) of the B-form (which is a higher humidity and higher salt form than the A-form). Recently, systematic studies on a wide variety of natural DNAs and synthetic polynucleotide poly d(AC).poly d(GT) revealed that three conformations C → A → B could be induced in the same fibre by increasing the humidity environment of the fibre when the prevailing counter-ion was sodium (Rhodes et al., 1982). It has also been observed that the C-form is a lower salt conformation than the A (Rhodes et al.,

1982). The C-form has a meridional reflection M at $\sim 0.30\ A^{-1}$ and an equatorial reflection (110) at $\sim 0.054\ A^{-1}$ (for hexagonal packing).

The observations now raise the question of whether these reflections X and M on figure 1 are from either the B-form or the C-form. Was the first sign of the C-form for NaDNA seen twenty-five years ago? To answer this question, systematic observations and analyses have been made using a wide variety of natural DNAs, as well as poly d(AC).poly d(GT) and poly d(AT).poly d(AT), and the results are discussed in this chapter.

## 6.2 Materials and Method

All the natural DNAs were obtained from Sigma. $\phi$w-14 was provided by Dr. R.A.J. Warren from the University of British Columbia, Canada.

The synthetic polynucleotides poly d(AC).poly d(GT) and poly d(AT).poly d(AT) were provided by Dr. J. Brahms at the Universite Paris VII, France and also bought from Boehringer. The final stage of purification of these materials was either by ethanol precipitation from solutions containing 0.1 M or less NaCℓ, or by centrifugation at 50,000 r.p.m. from solutions for which the NaCℓ concentration was 0.005 M. Fibres were drawn from either the precipitated material or from the concentrated gel at the bottom of the centrifuge tube. X-ray diffraction patterns were obtained over a range of relative humidities of the fibre environment from 0% to 98%.

## 6.3 Results and Discussion

X, $X_1$ are the equatorial reflections (see figure 1). M, $M_1$ are the meridonal reflections (see figure 1). A, $A_1$ are the (130) reflections of the A-form (see figure 1). The reciprocal space radii of these reflections are $\sim 0.0860\ A^{-1}$ (Fuller, 1961). $\rho_x$ is the reciprocal space

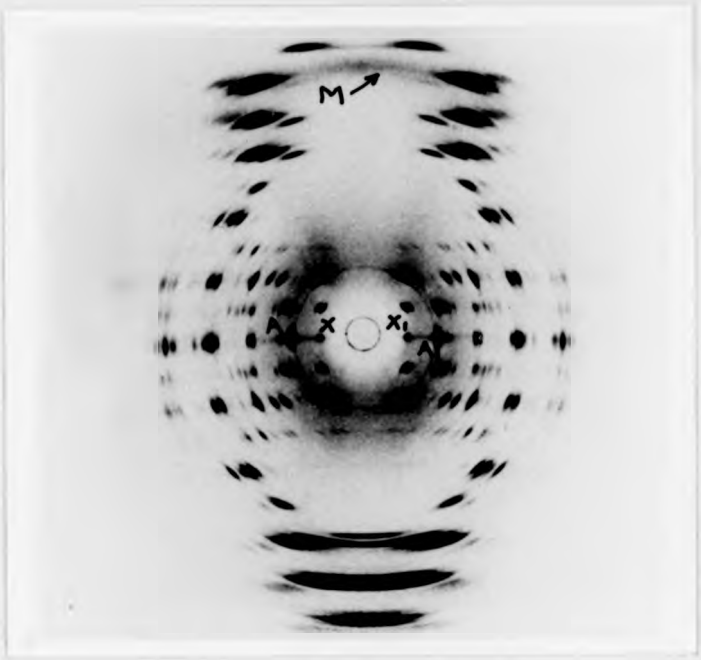



Figure 1: A-form Diffraction Pattern of NaDNA at 75% R.H. M and $M_1$ (not shown) are the meridional reflections from B or C forms. X and $X_1$ are the first equatorial reflections (from the centre of the pattern) from B or C forms.

A and $A_1$ are the first equatorial reflections (from the centre of the pattern) from A-form

(From Fuller et al., 1965)

radius of the reflections X, $X_1$.

$\rho_A$ is the reciprocal space radii of the reflections A, $A_1$.

$I_x$ is the intensity of the reflections X, $X_1$.

$I_A$ is the intensity of the reflections A, $A_1$.

$I_M$ is the intensity of the reflections M, $M_1$.

Intensities were obtained by measuring a radial densitometer trace across each diffraction arc (Langridge et al., 1960a). The trace area is corrected using the equation $I = A\rho\xi$ where A is the area under the radial densitometer trace across the spot in the fibre diagram and $\xi$ is the distance of the reciprocal lattice point from the c* axis. $\rho$ is the reciprocal radii.

Low humidity diffraction patterns from poly d(A-T).poly d(A-T), poly d(A-C).poly d(G-T) and natural DNAs are shown in figure 2, figure 3 and figure 4 respectively. If the reflections X, $X_1$ and M are due to the B-form (which is a high humidity and high salt form) we would expect an increase in the ratios $\frac{I_x}{I_A}$ and $\frac{I_m}{I_A}$ with an increase in the relative humidity on the fibre environment. The experimental observations are shown schematically in figure 5. There is a decrease in the intensity ratio, and then an increase at an intermediate relative humidity ($\sim$ 75%) as the relative humidity increases from 33%. The calculated intensity ratios are tabulated (see table 1 and table 2) for two different fibres from poly d(AT).poly d(AT). The diffraction patterns used in table 2 were taken by Dr. N.J. Rhodes in this laboratory.

The relative humidity for the minimum point is dependent on the amount of salt retained in the fibre. The lower the salt content in the fibre the higher the relative humidity required to obtain the pure crystalline A-form.

The above experimental observation could be explained either by a change in molecular transform or by a gradual change in dimensions of

TABLE (1)

| Relative Humidity | $\rho_X/\rho_A$(equatorial) | $I_x/I_A$(equatorial) | $\frac{I_M}{I_A}$ (meridional) | $\rho_X$ ($A^{-1}$) |
|---|---|---|---|---|
| 57% | 1.61 | 0.33 | 20.92 | 0.0556 |
| 75% | 1.60 | 0.28 | 18.95 | 0.0542 |
| 92% | 1.63 | 0.05 | 2.20 | 0.0530 |
| 98% | 1.86 | 3.37 | 147.01 | 0.0459 |
| 98% | 2.2 | - | - | 0.0390 |

TABLE (2)

| Relative Humidity | $\rho_X/\rho_A$(equatorial) | $I_X/I_A$(equatorial) | $\frac{I_M}{I_A}$ (meridional) | $\rho_X$ ($A^{-1}$) |
|---|---|---|---|---|
| 66% | 1.57 | 1.20 | 140.0 | 0.0553 |
| 75% | 1.70 | 0.00 | 0.00 | 0.0509 |
| 98% | B' 1.90 | | | B' 0.0465 |
| 98% | B 2.16 | 0.90 | 24.6 | B 0.0407 |

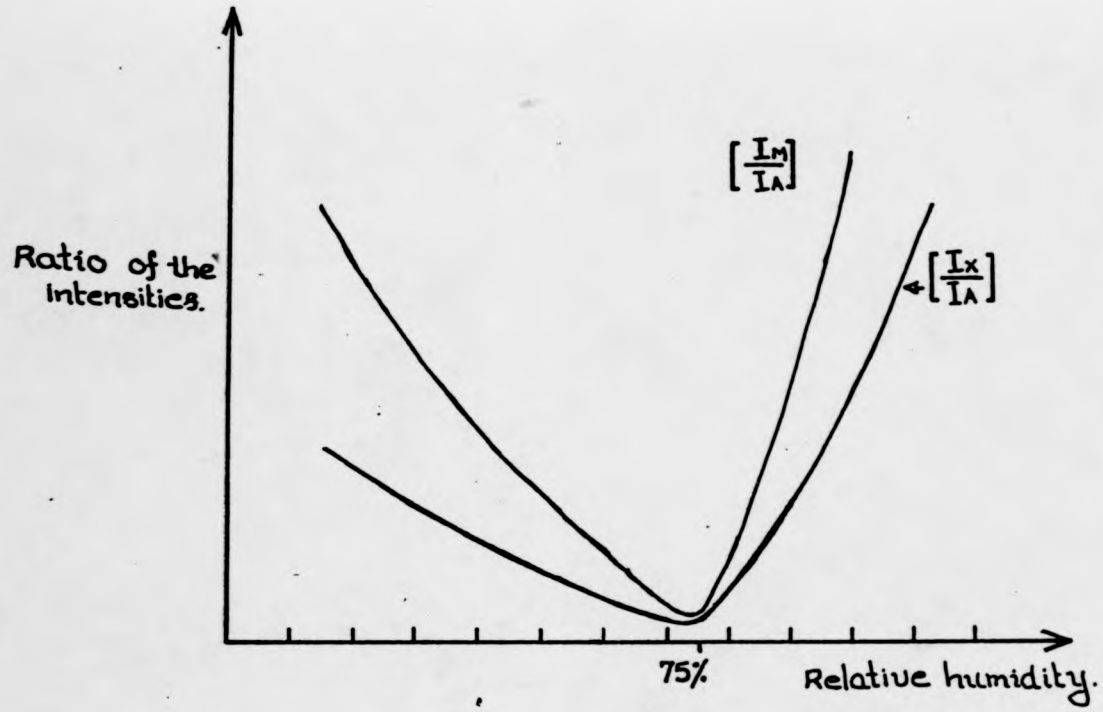



Figure 5: Ratio of the intensities Vs relative humidity is shown schematically.

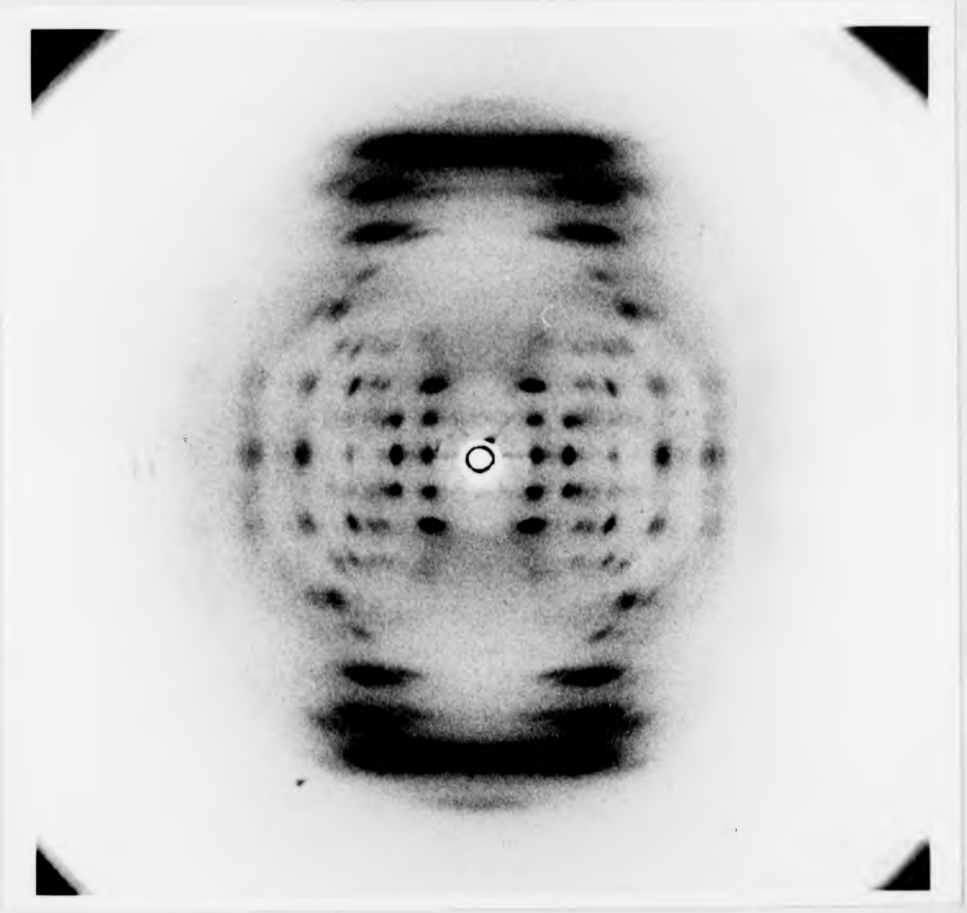


Figure 2: Mixture of A and C forms. Diffraction Pattern of Na poly d(AT).poly d(AT) at 57% R.H.

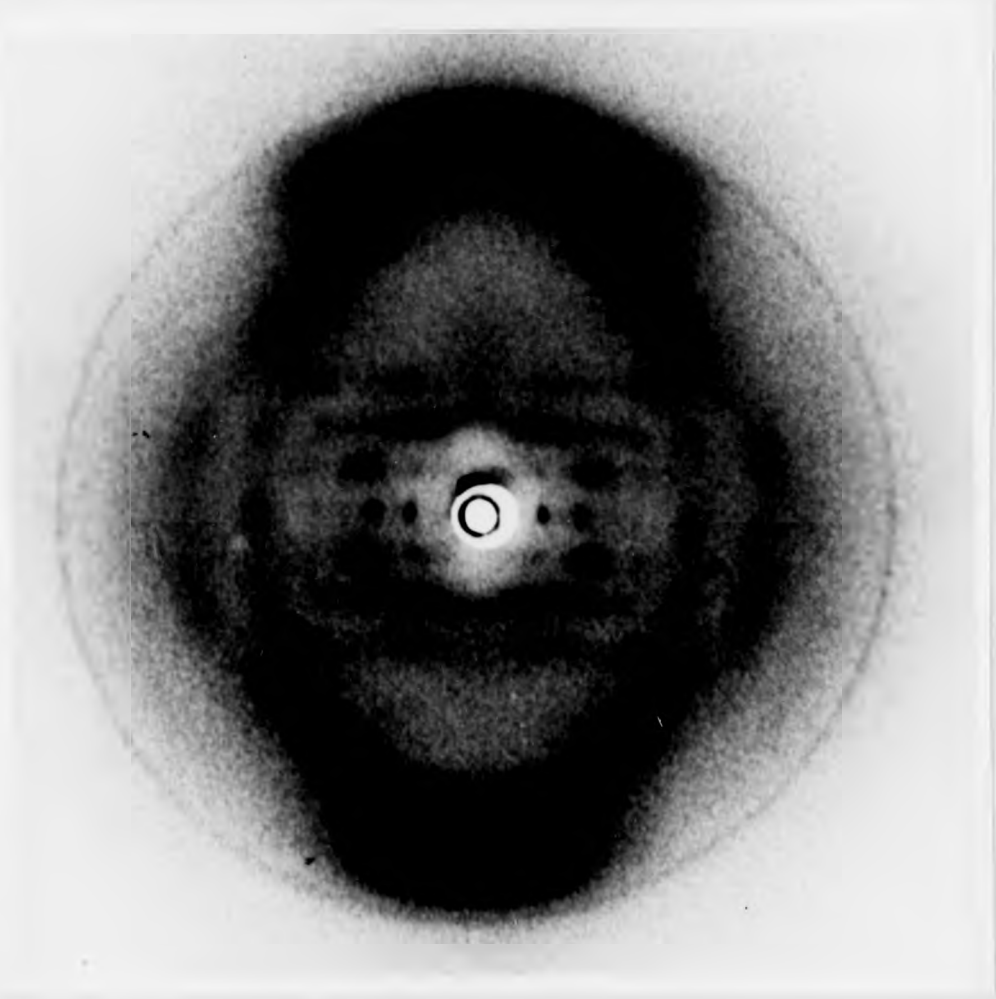


Figure 3: Mixture of A and C forms. Diffraction Pattern of Na poly d(AC).poly d(GT) at 66% R.H.

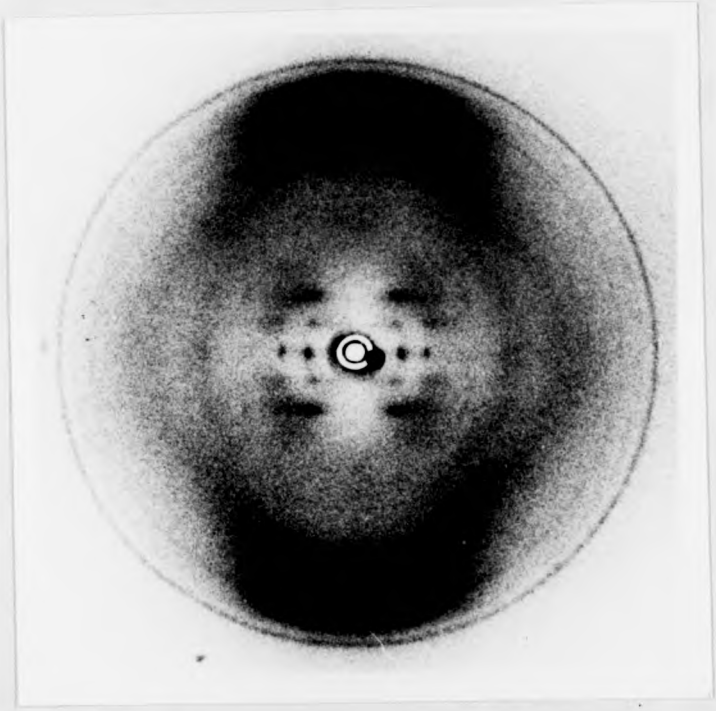


Figure 4: Mixture of A and C forms. Diffraction Pattern of Na DNA at 66% R.H.

the unit cell. This may be schematically shown in figure 6.

Supposing, for the sake of argument, we assume as shown in figure 6, that at low humidity the (110) plane samples the molecular transform at point (1) (see figure 6). This will give a medium-strong intensity $I_1$ to the (110) reflection. Further increase in humidity moves the (110) plane towards the centre of the transform and at an intermediate humidity it samples the molecular transform at point (2) (see figure 6) of the transform. This will give a zero or very weak intensity for the (110) reflection. At high humidities the (110) plane samples the transform in the principle maximum. This will give a very strong (110) reflection. This explains the intensity changes of the 110 reflection with humidity.

The intensity changes of the meridional reflections $M_1$ and M are schematically shown in figure 5 and tabulated in table 1 and table 2. The intensities of the meridional reflections are not going to be affected by sampling as for the equatorial reflections. Thus it must be due to a change in the molecular transform. That is, the low humidity diffraction pattern is from an (A/C) mixture. As the humidity is increased, at an intermediate relative humidity all the residual 'C' form changes to the crystalline A form. Further increases in the relative humidity of the fibre environment results in the crystalline A form changing to the semi-crystlaline B-form via an (A/B) mixture state. This explains the observed intensity variations of the reflections X, $X_1$, M and $M_1$ with humidity.

The transitions between the A, B and C conformations and their intermediate states may be explained by either "temporal" effects or "spatial" effects or both. The term "temporal" refers to the process whereby humidity changes exert their effects more rapidly on the fibre surface than they do on the interior. The phenomenon is of increasing significance as the humidity increases. A schematic representation is shown in figure 7,

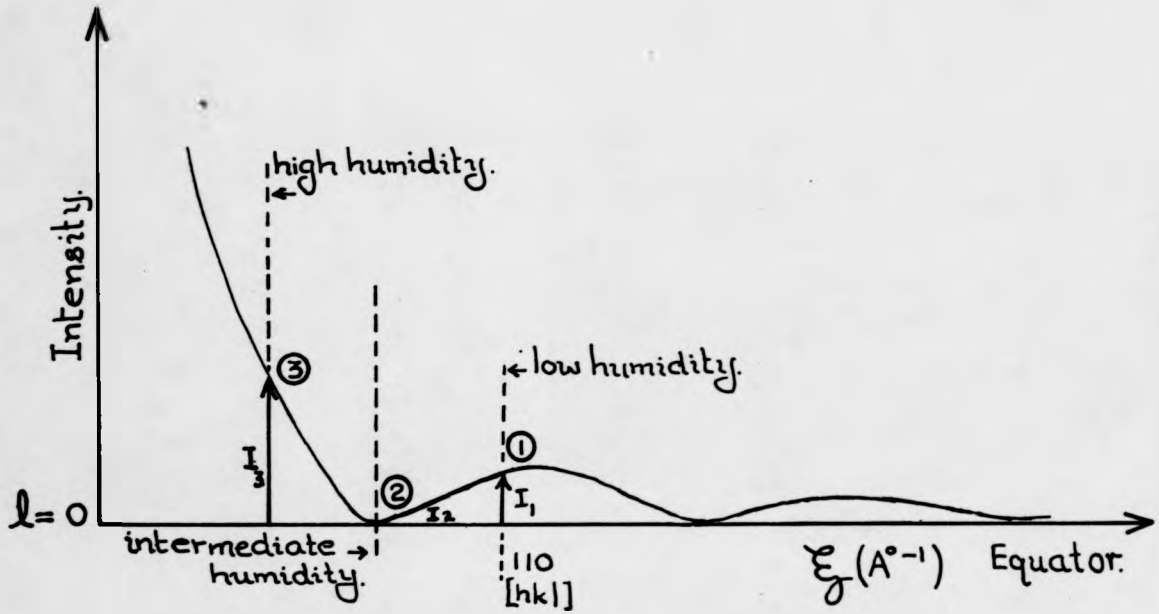



Figure 6: The cylindrically averaged squared transform (for ℓ=0) and (110) lattice plane are shown schematically.

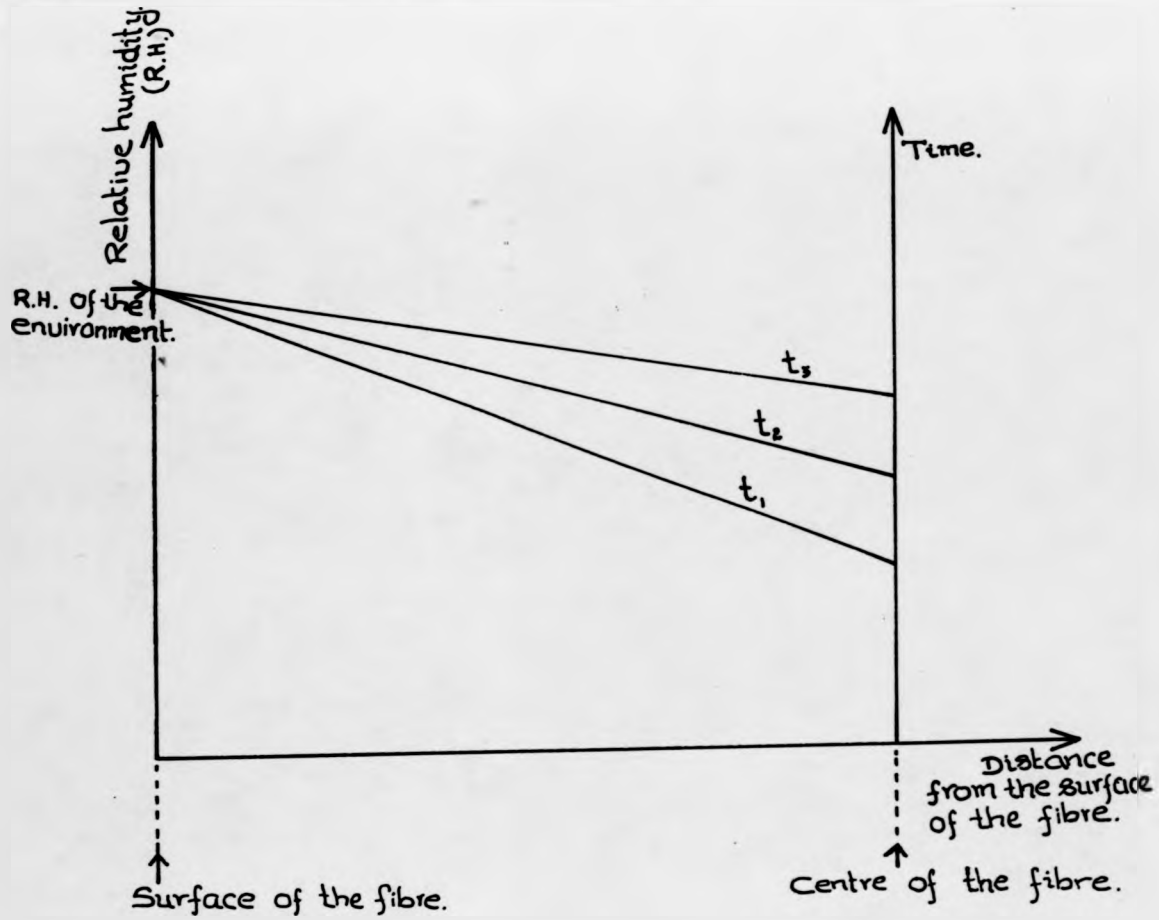



Figure 7: A schematic representation of the hydration of a fibre with time.

where $T_1$, $T_2$ and $T_3$ are indicating the time duration. Thus, conformational changes 'migrate' from the surface of the fibre towards the centre of the fibre. Such 'migration' within the fibre is schematically illustrated in figure 8. With reference to figure 8, it is noted that the formation of a pure crystalline intermediate A form (see table 2) crucially depends on the amount of salt present in the fibre. For a lower salt concentration than the above case (table 2) the intermediate A form may contain a very small fraction of B or C conformations. This is shown schematically in figure 9.

When pulling a fibre from the material in use, the piece of DNA was added to the solute (of known concentration) which was held between two very fine glass rods. A gel was thus formed. Sometimes, whilst observing the suspension under a microscope, it was noted that two phases of the DNA solute mixture seemed to be present, and as the gel dried to a fibre, the microcrystals of salt were observed on the fibre surface. This effect was observed to be prominent in poly d(AC).poly d(GT). The diffraction pattern from this type of fibre, at low humidity, was similar figure 10 and had a strong salt ring. When the humidity was increased, the salt ring disappeared from the diffraction pattern and did not re-appear when the environmental humidity of the fibre was again lowered. The above observations suggest that the distribution of the excess salt in the fibre may not be uniform. In the discussion this will be referred to as "the spatial effect". As a result of this some DNA molecules (or different sections of the same DNA molecule) have more neighbouring ions present than the rest. This may lead to a non-uniform hydration of the fibre and thus a non-uniform distribution of the polar environment in the fibre. Under these circumstances the A conformation contains an impurity (or mixture) of either B or C or both. It was also observed that the low salt very high humidity and high salt low humidity fibres both gave the semi-

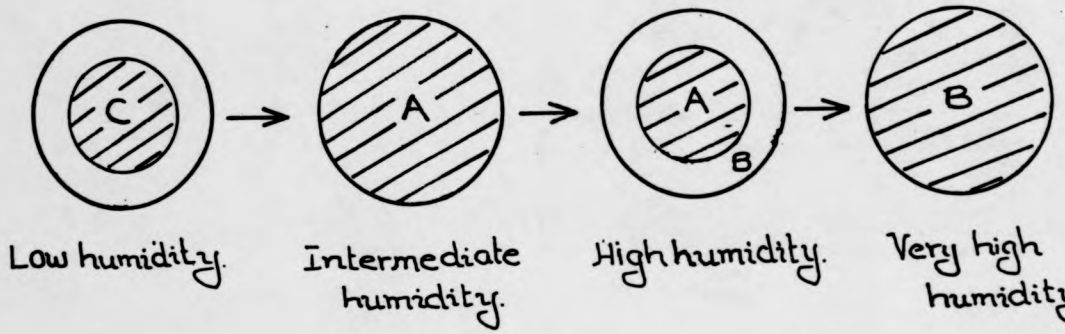



Figure 8: A schematic represenation of the migration of the conformational change in a fibre

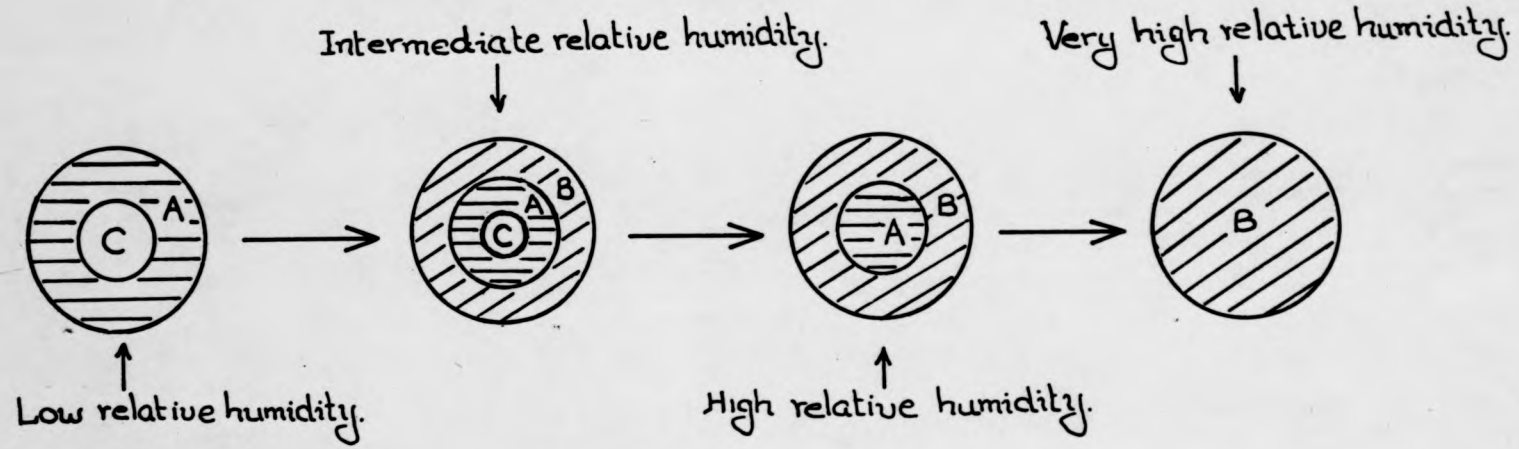



Figure 9: A schematic representation of the migration of the conformational change in a fibre (the salt concentration is lower than in the case of figure 8)

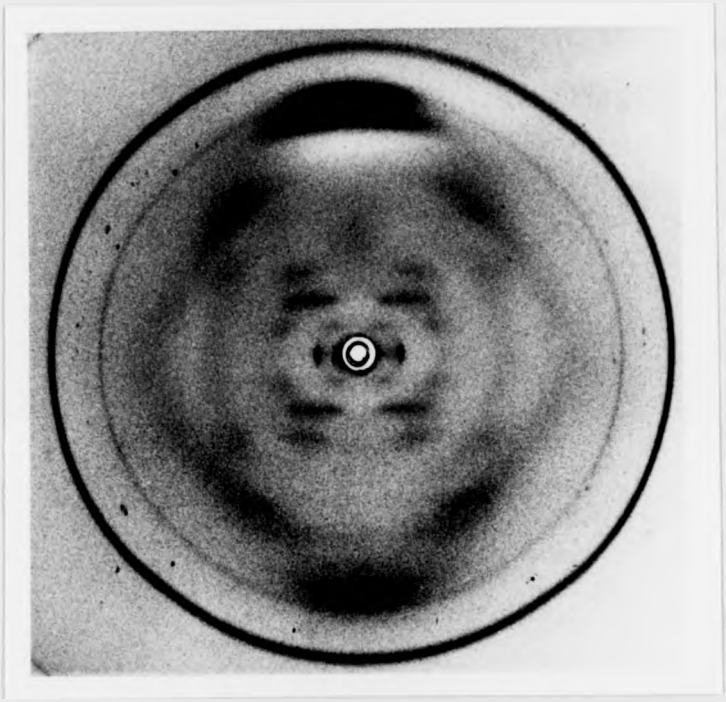


Figure 10: Semi-crystalline C-form Diffraction Pattern of Na poly d(AC).poly d(GT) at 57% R.H. Strong NaCℓ diffraction ring of 2.81A$^{o}$ spacing.

crystalline B form. This may be explained by suggesting that the hydration mechanism in the two cases is different. In high salt fibres at low humidity there will be more water "grasping" centres present than in low salt fibres at the same humidity. In time the high salt fibre will grasp a sufficient amount of water molecules to produce a transition to the B form. The low salt fibre, however, may remain in the C-form, or in an (A/C) mixture for many months, if not permanently. Water taken into the fibre is limited by the number of water "grasping" centres (i.e. ions) within fibre. However, low salt fibres gave a B conformation at very high humidity. This is possibly a result of migration of water molecules from the environment into the fibre and in this case the 'temporal' effect may be the predominant one. Further evidence relating to these transitions has been obtained by measurement of the change in $\rho$-values of the reflections X, $X_1$, $M_1$ and M with relative humidity.

At low humidity for the C conformation the ratio $\rho_A/\rho_X$ is

$$\frac{0.0866}{0.0341} = 1.6 \text{ (57\% RH)}$$

for the semi-crystalline B-form the ratio is:

$$\frac{\rho_A}{\rho_X} \text{ is } \frac{0.0866}{0.0375} = 2.3 \text{ (98\% RH)}$$

Some examples of the experimental observations of these ratios are tabulated in table 1 and table 2. These can be schematically shown in figure 11, which indicates a small increase in the ratio $\rho_A/\rho_X$ up to the intermediate humidity and then a sharp increase as higher humidities are approached.

The low humidity $\rho_A/\rho_X$ ratio indicates that the reflection X must be from the C conformation. The ratio $\rho_A/\rho_X$ was calculated for the A pattern published by previous workers (e.g. Langridge et al., 1957, see figure 1, Franklin and Gosling, 1953b). In most of the cases the ratio is 1.6-1.7. This suggests that the reflection X is probably from the C conformation rather than B-conformation. The ratios $\rho_A/\rho_X$ for an (A/C) mixture from

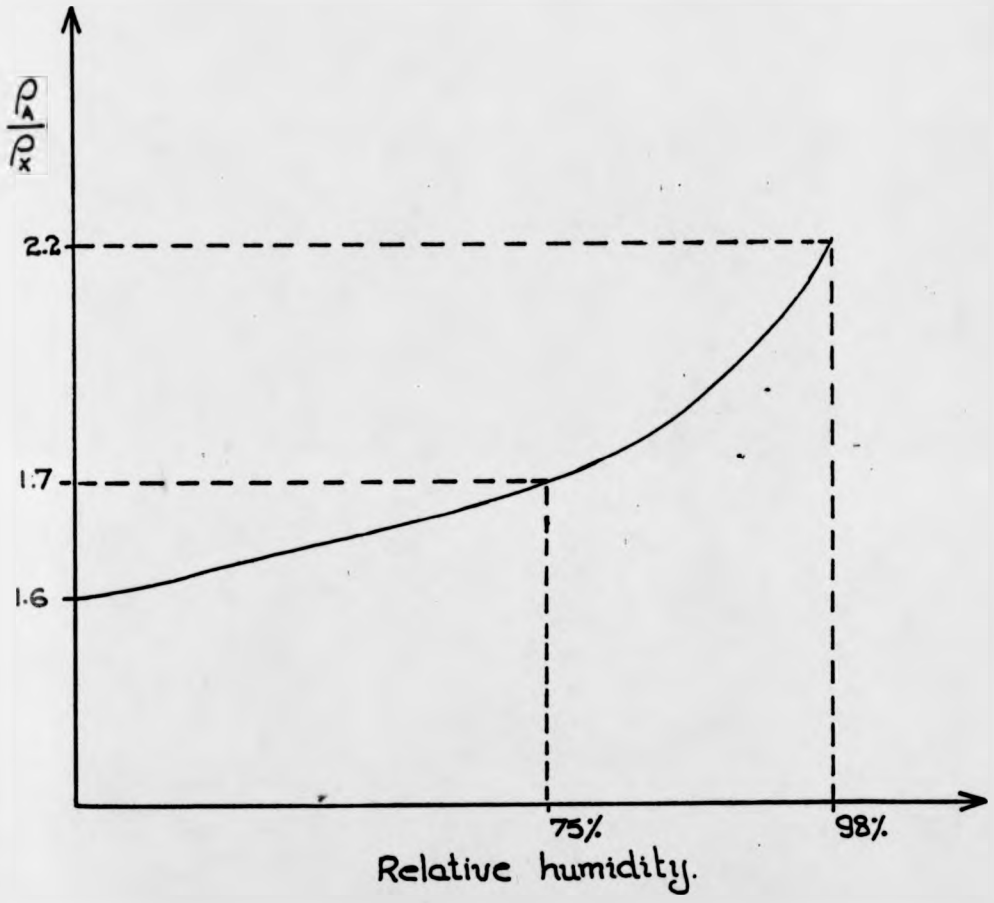



Figure 11: $\rho_A/\rho_X$ Vs relative humidity is shown schematically

natural DNAs and poly d(AC).poly d(GT) are also consistent with above results.

In poly d(AC).poly d(GT) at low humidity an (A/C) mixture was observed (see figure 3). When the humidity was increased the reflections X and M became weaker. At an intermediate humidity a fully crystalline A form was observed with no equatorial reflection X, $X_1$ and meridional reflections M, $M_1$ (see figure 12). A further increase in the environmental humidity gave a semi-crystalline B-form via an (A/B) mixture. When reducing the humidity the crystalline A-form was observed. There was no sign of either the meridional reflection M or equatorial reflection X even at very low humidities. From the work in this laboratory on poly d(AC). poly d(GT) (Rhodes et al., 1982) it was found that under the salt conditions which have been studied the C → A transition was irreversible and the A → B transition was reversible for this polynucleotide. These observations strongly suggest that at low humidities the reflections X and M must be from the C-conformation rather than the B conformation.

The layerline spacings for the A and C forms are very close, and both are distinct from more than that of the B-form. Thus in the (A/C) mixture diffraction pattern at low humidity the $1^{st}$ and $2^{nd}$ layerline reflections from both components tend to overlap, and are hence difficult to resolve (see figure 3). However, the equatorial reflection X and meridional reflection M are visible on the A-form patterns. It is observed that the central part of the C-form pattern is much more crystalline than the semi-crystalline B-form. That is, the reflections in the C-form are much sharper (well sampled) than in the B-form. The reflection X on the A form (see figure 1) at low humidity is highly crystalline and shows no sign of the broadening so characteristic of continuous diffraction. However, at high humidity this reflection X is broadened (that is, the diffraction tends to become more continuous). Again these observations suggest that the reflections X and M at low humidity are more likely to originate from the

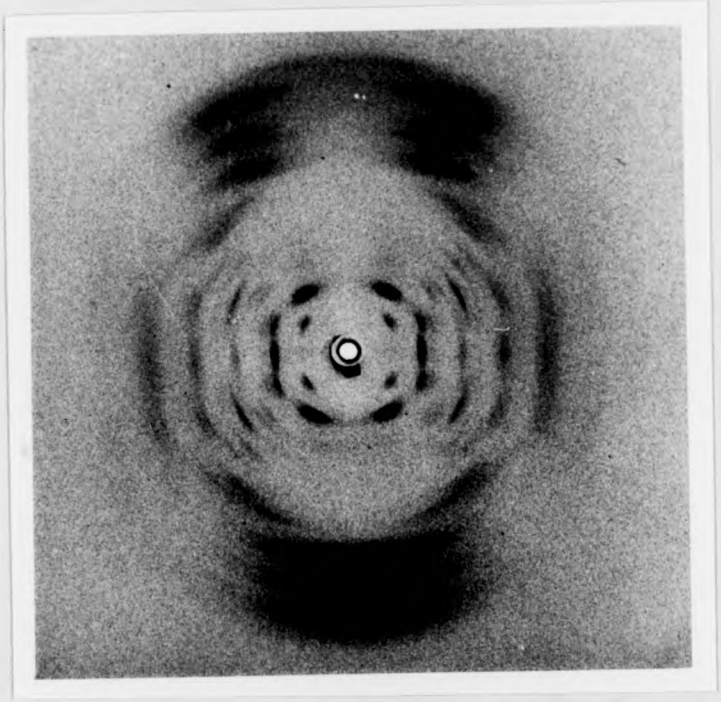


Figure 12: A-form. Diffraction Pattern of Na poly d(AC).poly d(GT) at 66% R.H.

C-form rather than the B-form.

## 6.4 Conclusion

The discussion in this chpater clearly indicates that the reflections X and M at low humidity for low salt fibres are from the C-form rather than the B-form. At high humidities corresponding reflections are from the B-form and hence it seems that the first sign of the C-form of NaDNA was actually observed twenty years ago!!!

C-form rather than the B-form.

## 6.4 Conclusion

The discussion in this chpater clearly indicates that the reflections X and M at low humidity for low salt fibres are from the C-form rather than the B-form. At high humidities corresponding reflections are from the B-form and hence it seems that the first sign of the C-form of NaDNA was actually observed twenty years ago!!!

# CHAPTER 7

## Alternating DNA Conformations

### 7.1 Introduction

Up till fairly recently, the study of DNA has been restricted to natural 'random sequence' DNA. The 'A' and 'B' forms of DNA were solved and there was speculation as to the possible significance of the two forms. What is now proving to be increasingly significant is the examination of the various synthetic polynucleotides that have since become available. The polynucleotides have been made in such a way as to have a regularly repeating sequence of bases as well as a faithfully repeating asymmetric unit (which may be mono, dinucleotide, or other).

Natural DNA is known to have regions where the base coding repeats itself over a considerable length of the molecule, and it now becomes apparent that the study of synthetic polynucleotides is of great importance in the development of contemporary ideas about native DNA as a whole.

The possibility of DNA exisiting in a variety of different forms, each suited to a particular function, or biochemical environment is evident as is the possibility of DNA structure being not only a function of salt and water content, but also of base sequence. Frams (1971) has explored the idea of recognition of certain regulatory proteins, such as repressors by special types of DNA.

An assumption that has been made widely throughout the earlier considerations of DNA is that the asymmetric unit is one nucleotide in extent, repeating itself along the molecule with identical sugar puckering and conformational angles. Recent experimental work on synthetic polynucleotides has, however, produced evidence to support the idea of

slightly more complex asymetric units within the structure of the macromolecule. Viswamitra et al. (1978) have reported an alternating variation in the sugar puckering and phosphodiester geometry of the tetranucleotide $(dA\text{-}dT)_2$. The sugar puckering was proposed to be $C_3'$-endo when attached to purine and $C'_2$-endo when attached to pyrimidine. Although this tetranucleotide forms a 'bent' structure rather than a complementary duplex, Klug et al., (1979) speculated that such an alternating backbone structure could exist in double helical poly d(AT).poly d(AT), and that this structure might have specific biological properties. They point out two observations in support of their hypothesis. Firstly, that the 'lac' repressor protein of E.Coli binds about 100-1000 times more strongly to poly d(AT).poly d(AT) than to calf thymus DNA (Riggs et al., 1972). Secondly they note that when synthetic poly d(AT).poly d(AT) was digested (Scheffler et al., 1968) with pancreatic DNA-ase I, it was found that the oligonucleotides thus formed had thymine at their 5' ends, and that successive oligomers differed in length by two nucleotides (or a multiple of thereof). No fragments were detected that either consisted of odd numbers of nucleotides, or that started with adenine. It was later shown that the enzyme caused cleavage at the $O_3'$-P bond, leaving a 5' terminal phosphate (Laskowski, 1971). Recent experiments on random sequence DNA have tended to support these results (Lutter, 1977).

Another alternating structure has been described by Wang et al., (1979). They have studied an hexanucleotide d(CpGpCpGpCpG) which is believed to crystallise into a left-handed double helix with Watson-Crick base pairing but with a 'zig-zag' phosphate backbone having a dinculeotide asymetric unit. They termed this structure 'Z-DNA'. Leslie et al. (1980) have reported a left-handed alternating structure (called S-DNA) for fibres obtained from poly d(GC).poly d(GC) and poly d(AC).poly d(GT). They also reported a C-form structure from poly d(AG).poly d(CT), which has a dinucleotide repeat.

The remainder of this chapter will be devoted to assessing the evidence for the existence of alternating structures as related to X-ray diffraction photographs taken from poly d(AT).poly d(AT), poly d(GC).poly d(GC), poly d(A-BrU).poly d(A-BrU) and poly d(A-IU).poly d(A-IU).

## 7.2 Materials and Methods

All the polynucleotides were provided by Dr. J. Brahms of the University of Paris. The polynucleotides poly d(AT).poly d(AT), poly d(GC).poly d(GC) and poly d(AC).poly d(GT) were obtained from Boehringer. The purification of the polynucleotides and the preparation of the fibres was performed in the same manner as described in Chapter 2. All the fibres were prepared from the sodium salt unless otherwise stated.

## 7.3 Results and Discussion

The poly d(AT).poly d(AT) fibres were prepared by precipitation from 0.1 M CsF, 0.2 M CsF solutions. The estimated values of the $F^-/PO_4^-$ ratio in these fibres were 0.3 and 0.6 respectively. The diffraction patterns were obtained in an environment having known humidity. The diffraction pattern shown in figure 1, D-DNA, was observed at humidities ranging from 33% to 92%, and was very similar to patterns previously obtained for Na poly d(AT).poly d(AT) (Davies et al., 1963, Arnott et al., 1974b, the present work by the author). The same fibre gave the classical B form (i.e. tenfold helix with pitch of 24.0 A) at 98% relative humidity (figure 2). The D to B transition was found to be reversible. At higher ionic concentrations the diffraction pattern obtained at low relative humidities (figure 3) was very different to the D-form described above. The observed diffraction indicated an eightfold helix with 24.0 A pitch, designated as the D' conformation. The same fibre gave a B' conformation at relative humidities of 92% and 98%. The B' conformation was also

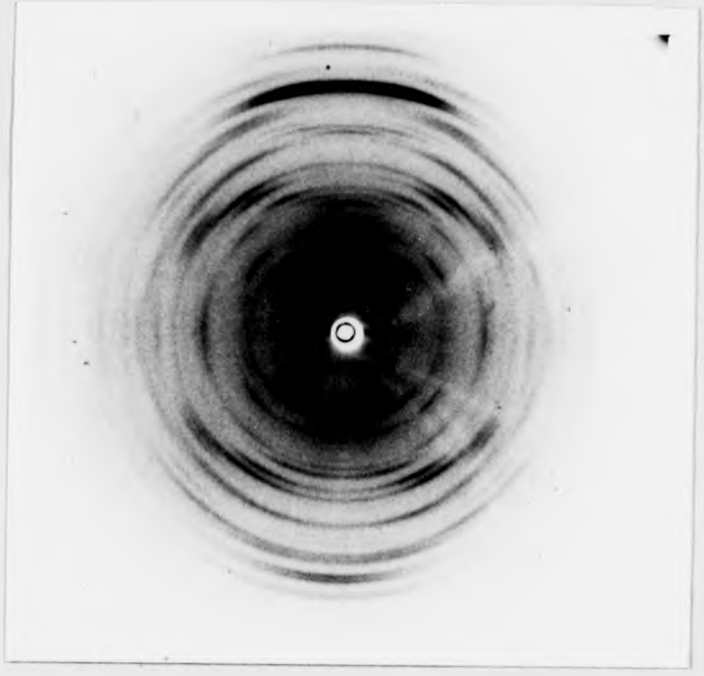


Figure 1 : D-form Diffraction Pattern of Cs poly d(AT).poly d(AT) at 66% R.H.

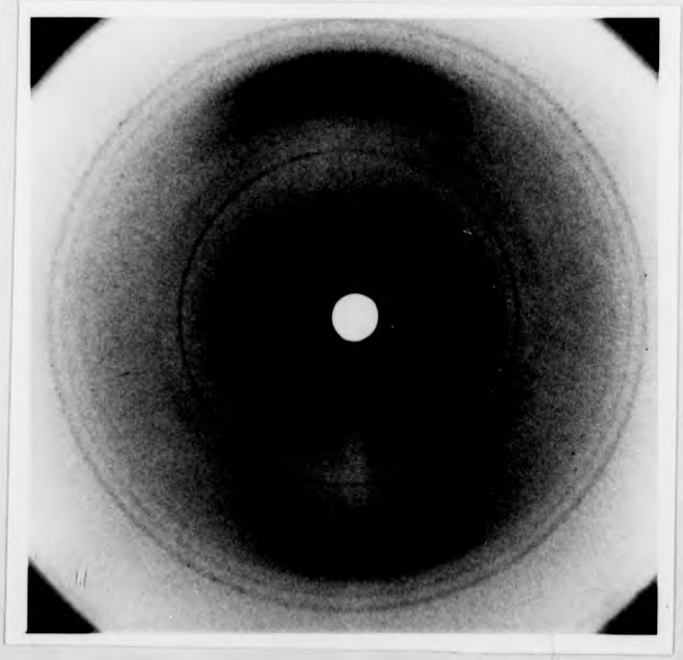


Figure 2 : Semi-crystalline B-form Diffraction Pattern of Cs poly d(AT).poly d(AT) at 98% R.H.

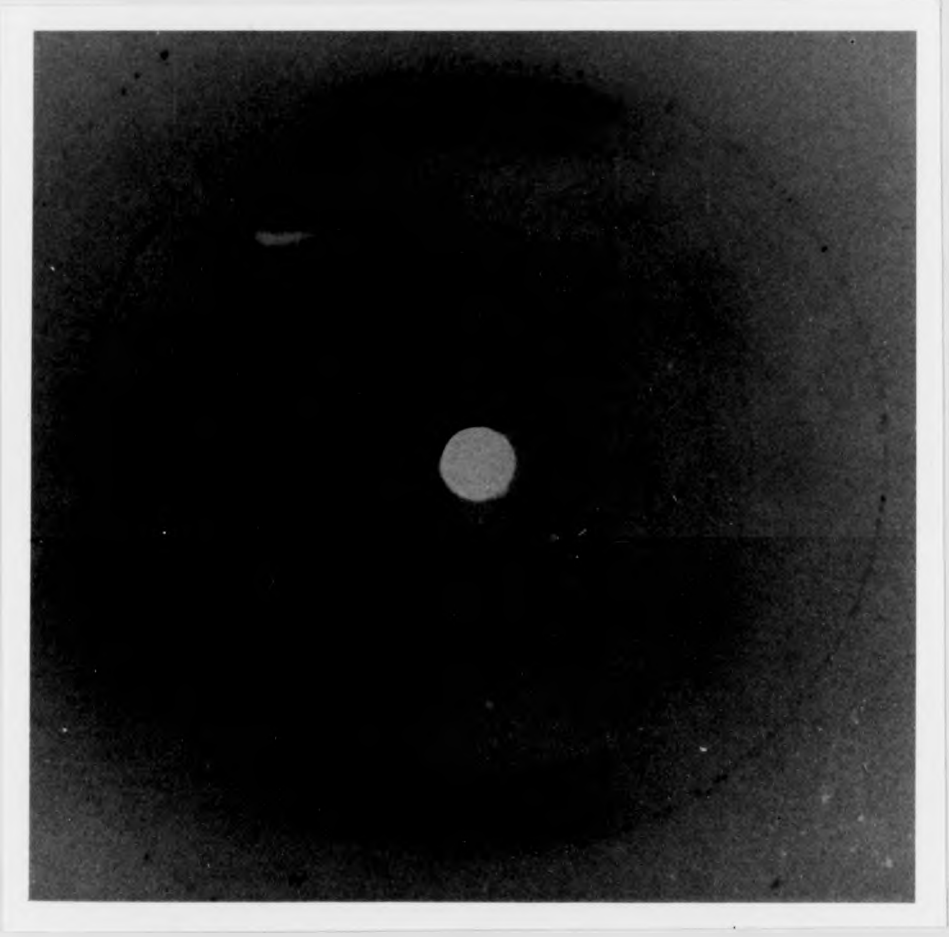


Figure 3 : D'-form Diffraction Pattern of Cs poly d(AT).poly d(AT) at 66% R.H.

observed for NadAT fibres, although this was not a crystalline form as observed for the homopolymer poly d(A).poly d(T) (Arnott et al., 1974a, and in the present work by the author). A well characterised B' pattern was also observed for Na poly d(AT).poly d(AT) by Dr. N.J. Rhodes in this laboratory (figure 4). Further humidification of this fibre induced a transition from B' to B-DNA. The important differences between D'-DNA and D-DNA are:

(a) In D-DNA the molecules are packed in the tetragonal unit cell with cell dimensions:

a = 17.1 A c = 24.1 A

In D'-DNA the molecules are packed in the hexagonal unit cell with cell dimensions:

a = 20.5 A c = 24.0 A

(b) The overall intensity distribution in each diffraction pattern is very different, even when lattice sampling differences have been accounted for. The above observation suggests that these molecular conformations of D-DNA and D'-DNA are distinct from each other. The interesting feature of D-DNA is the strong meridional reflection on the even numbered layerlines. These are probably due to the alternating base sequence of the polynucleotide. The strong meridional reflection of the 4th layerline of D'-DNA strongly suggests that the sugar-phosphate backbone conformation is alternating rather than regular. It thus seems that D'-DNA has $4_1$ symmetry and a dinucleotide asymetric unit. This is convincing experimental evidence from the fibre diffraction studies that there is an alternating backbone structure in some poly d(AT).poly d(AT) samples.

The fact that the D' → B' transition was reversible with humidity coupled with the supposed alternating structure of D'-DNA suggests that the

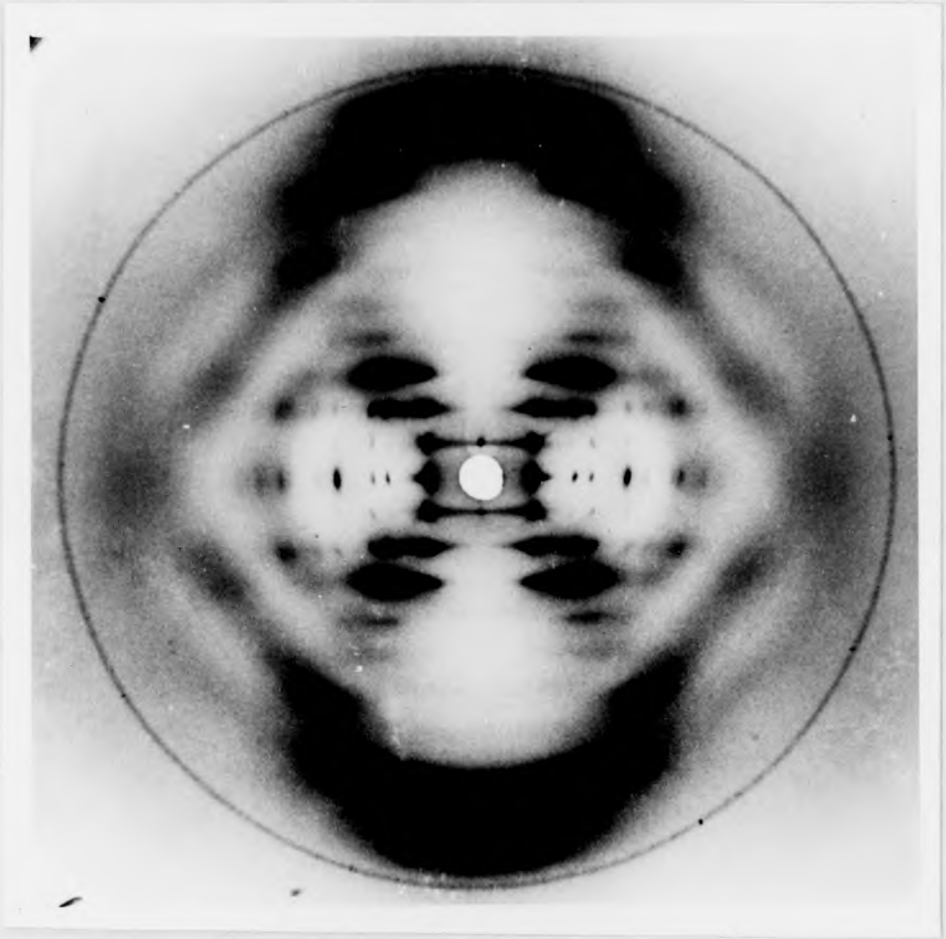


Figure 4 : Semi-crystalline B'-form Diffraction Pattern of Na poly d(AT).poly d(AT) at 95% R.H.

B'-form may also be an alternating structure. The proposition creates a problem - where is the fifth layerline meridional reflection on the B'DNA patterns? One possible reason is that, since crystallinity is poor in the B' pattern, the reflection may be masked by diffuse scattering. Another reason is evident from the work of Klug et al, (1979), who proposed that in poly d(AT).poly d(AT) thymine is more tightly stacked over adenine than it is in natural DNA, and coversely adenine stacked less over thymine. Further they point out that this would explain the n.m.r. spectra obtained by Patel et al. (1974). Because of the different stacking between thymine and adenine it is thought that the backbone may become distorted. The extent of this distortion depends, however, on the rotation per residue; the larger the rotation, the more significant will be the distortion. This then implies that the nucleotide repeat in B'-DNA (rotation = $36^{o}$) may well be less pronounced than that of D'-DNA (rotation = $45^{o}$). If this assertion is true then the fifth layerline meridional may be present, but is possibly so weak that it is very easily obscurred by non-crystalline diffraction.

The comparison of the diffraction patterns from B'-DNA (from Na poly d(AT).poly d(AT)) and B-DNA (random sequence) shows that in B'-DNA the higher order even numbered layerlines are generally stronger in intensity ($\ell$=6, $\ell$=8) than the odd-numbered layerlines. This fits with the idea of a dinucleotide repeat.

It may be argued that this effect is actually introduced by the presence of an alternating A-T sequence, rather than the presence of an alternating sugar phosphate backbone conformation. However, if this was the case, one might expect the same sort of intensity enhancement on the higher layerlines of the A patterns, as seen on those of the B' patterns. These observations, along with the fact that in strongly exposed A patterns (figure 5) there is evidence of dinucleotide repeat (5.15 A meridional and weak reflections between 7/8, 8/9, 9/10 layerlines) not only suggest that it

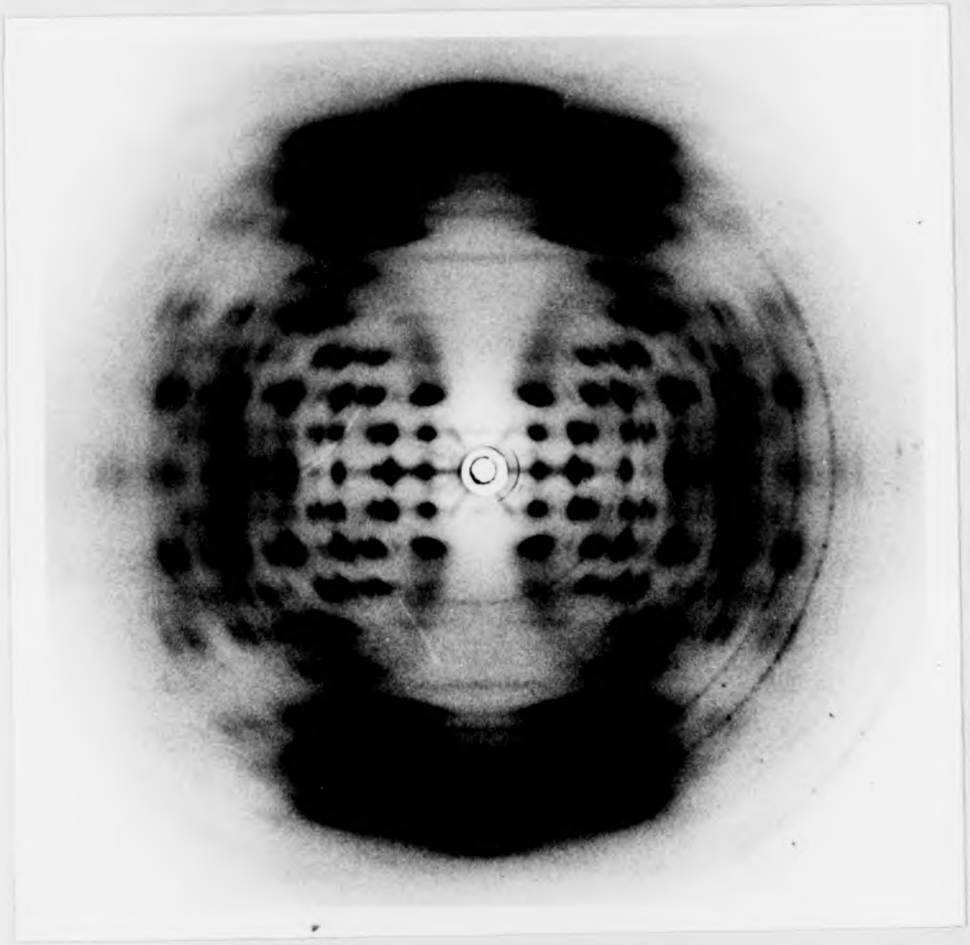


Figure 5 : A-form Diffraction Pattern of Na poly d(AT).poly d(AT) at 92% R.H.

is an alternating backbone conformation which is responsible for the observed B' patterns, but also that there is some sort of conformational dinucleotide irregularity (however mild) in the A-structures observed from poly d(AT).poly d(AT).

The fact that the pitch of B'-DNA is smaller than B-DNA (from random sequence DNA) may also be due to the stacking system envisaged in the alternating structure. Further support for this idea of an alternating backbone structure in poly d(AT).poly d(AT) is obtained from Raman spectroscopic studies. Shindo et al. (1979) studied a 145 base-pair length of poly d(AT).poly d(AT) using n.m.r. The results showed two distinct phosphorous-31 n.m.r. signals of approximately equal area at low salt concentrations (up to 0.1 M). This appears to be further direct evidence that there are two conformations in the phosphate diester backbone. Further, they point out that the calculation of magnetic shielding from the bases show that the bases alone cannot be held responsible for these results. The two signals collapsed into a singlet when the 145 base pair oligomer was heated above its melting temperature, but the doublet was regenerated on cooling to the room temperature, and it was concluded that in solution poly d(AT).poly d(AT) does exhibit an alternating sequence dependent on conformationa variation (Shindo et al., 1979). Since the properties of B-DNA (as implied by fibre diffraction studies) are thought likely to be very similar in outline of those of DNA found in free solution (Finch et al., 1977), the above n.m.r. solution study further supports the alternating structure. Also, solid state phosphorous-31 n.m.r. techniques were used by Shindo et al. (1981) to study DNA fibres and poly d(AT).poly d(AT) (Shindo et al., 1980) and provide further support for this conclusion. Similar fibres were prepared from the same precipitate and subjected to X-ray diffraction and Raman spectroscopic studies. The Raman spectroscopice studies were carried out in the University of Paris by Dr. J. Brahms and

his colleagues. Their experimental observations of fibres and solutions further supports the idea of an alternating conformation for co-polymers poly d(AT).poly d(AT), poly d(A-BrU).poly d(A-BrU) and poly d(A-IU). poly d(A-IU). Their work implies that the alternating conformation is formed by alternate $C_2$-endo and $C_3$-endo sugar ring puckering conformations.

The diffraction work on poly d(AT).poly d(AT) has been discussed in the previous sections and work on poly d(A-BrU).poly d(A-BrU) and poly d(A-IU).poly d(A-IU) is going to be discussed in the following section.

Fibres were made from precipitated or centrifuged gel of poly d(A-BrU).poly d(A-BrU) and poly d(A-IU).poly d(A-IU) (both were from NaCℓ solution). X-ray diffraction patterns were obtained in the relative humidity range from 33% to 98%.

The information given by the low humidity diffraction patterns was too poor even to clearly identify the conformation of the DNA. The medium strong equatorial and the strong meridional reflections are very similar to those from the C-form, and measurements confirm that these are C patterns. The same fibre gave well characterised B-form diffraction patterns at 98% relative humidity (figure 6). The diffraction pattern obtained from poly d(A-BrU).poly d(A-BrU) and poly d(A-IU).poly d(A-IU) were very similar. The B conformation observed for these polynucleotides has a meridional reflection on the fifth layerline, and the intensities of even numbered layerline reflections is much enhanced. This suggests that the B-DNA conformation may have $5_1$ symmetry with a dinucleotide asymmetric unit. For the lithium salt of the poly d(A-BrU).poly d(A-BrU) a B conformation has been observed (Davies and Baldwin, 1963) but the quality of the pattern is not sufficient to obtain detailed information about the structure. The diffraction patterns at 98% relative humidity from poly d(AT).poly d(AT) and poly d(A-BrU).poly d(A-BrU) or poly d(A-IU).poly d(A-IU)

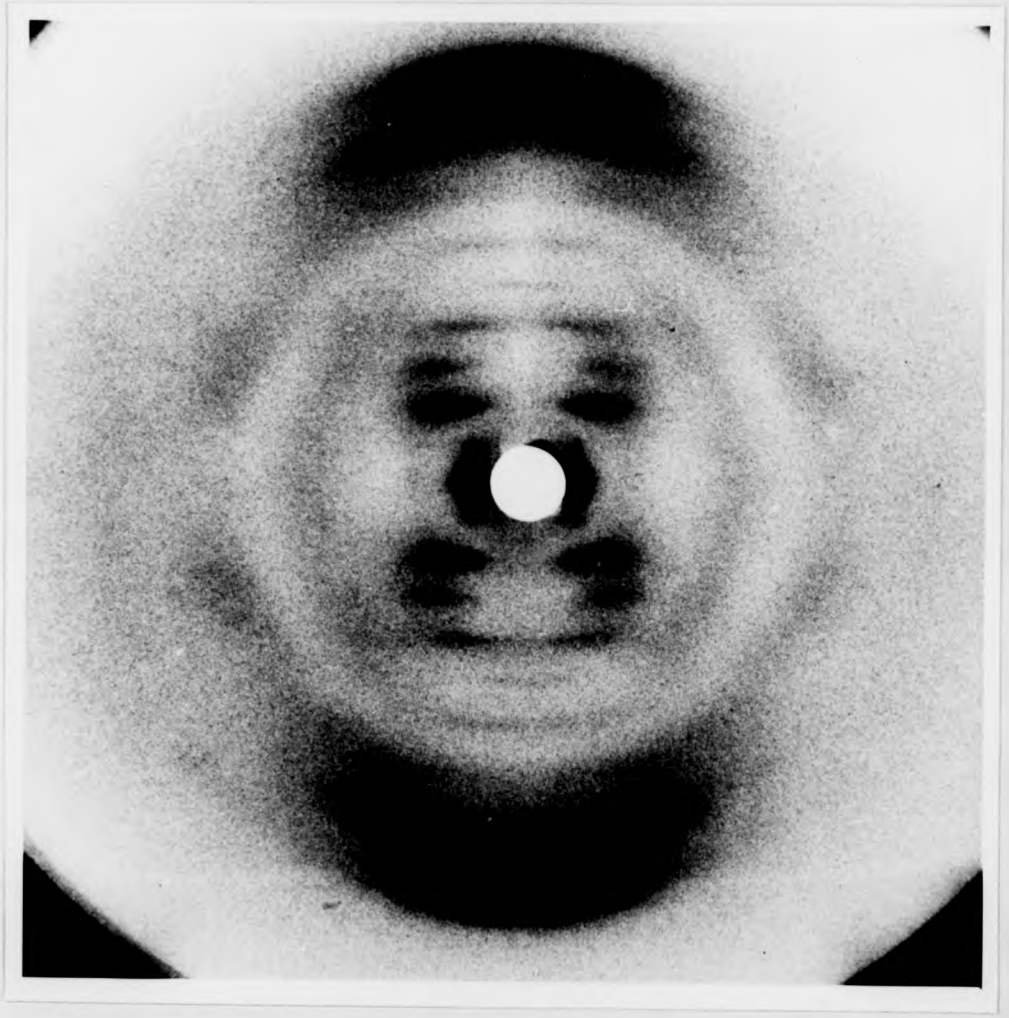


Figure 6 : Semi-crystalline B-form Diffraction Pattern of Na poly d(A-BrU).poly d(A-BrU) at 98% R.H.

were all semi-crystalline. Even though the diffraction pattern of poly d(A-BrU).poly d(A-BrU) fibre is semi-crystalline at 98% relative humidity, it shows a meridional reflection clearly on the fifth layerline. This was not observed for poly d(AT).poly d(AT). This suggests that the backbone variation in poly d(A-BrU).poly d(A-BrU) is much greater than in poly d(AT).poly d(AT).

Riggs and his colleagues (Lin and Riggs, 1971, 1972, Riggs et al., 1972) have shown that lac repressor protein can bind with much higher affinity to certain synthetic DNAs of defined sequence than it does to natural non-operator DNA. For example there is a 100 fold difference in the affinity of lac repressor for poly d(AU).poly d(AU) derivatives substituted at the fifth position of Uracil. This cannot be explained in terms of specific interactions in the major (or minor) groove of the DNA double-helix. The replacement of the 5th hydrogen of Uracil by a fifth-methyl group (i.e. in thymine) enhances repressor affinity twenty-fold, and replacement of the fifth-$CH_3$ by a fifth-bromine increases affinity by a similar factor (Lin and Riggs, 1971). The binding studies on poly d(A-U(Hg-DTT)) (where DTT is, dithiothreitol) (Richmond and Steitz, 1976) ruled out the possibility of an alternative explanation in terms of a favourable hydrophobic interaction between the 5th position of the substituted Uracil and the repressor. Therefore, the difference in the affinity of the lac repressor to these co-polymers must be due to the variation in the backbone conformation. The enhanced behaviour of poly d(A-BrU).poly d(A-BrU) in lac repressor binding as compared with poly d(AT).poly d(AT) is probably due to the enhanced backbone variation in poly d(A-BrU).poly d(A-BrU). (This is seen in figures 4 and 6).

Furthermore Klug et al. (1979) suggested that it is the overlap which makes the alternating structure particularly stable (by considering the cases where methyl group is absent or replaced it with a group such as

bromine atom, which would enhance the stacking). This poly d(A-BrU).poly d(A-BrU) melts at a temperature 9°C higher than does poly d(AT).poly d(AT) and the increased stability cannot be attributed to an increase in the net hydrogen bonding strength (Inman and Baldwin, 1962).

The next section describes a study on poly d(GC).poly d(GC).

A small fraction of the original material poly d(GC).poly d(GC) was used to make a fibre (i.e. before the centrifugation procedure, which is discussed in Chapter 2). When the fibre was viewed through the microscope, the excess salt crystallites were visible on the fibre surface. The diffraction pattern obtained from this fibre only contained salt rings. Therefore, the centrifugation process was used to remove excess salt. The fibre made from centrifuged gel was used to take the diffraction photograph for the relative humidity range from 33% to 98%. The diffraction pattern obtained at 33% was not strongly characterised because of the poor quality of the material. It was, however, sufficient to measure the helical parameters of the conformation. The base separation is 3.32 A and the pitch is 33.4 A. The values of the base separation and pitch are very similar to the B-type conformation. In a systematic polymorphic study of K poly d(GC).poly d(GC), a very similar, but well characterised pattern was observed (detail will be discussed in Chapter 8). This pattern (see figure 7) has a meridional reflection on the fifth layerline which suggests that the sugar phosphate backbone conformation is an alternating conformation rather than a regular conformation. It thus seems that the DNA conformation has '$5_1$' symmetry and a dinucleotide asymmetric unit. This is another piece of convincing experimental evidence from fibre diffraction studies that indicates that there is an alternating backbone structure in poly d(GC).poly d(GC) conformations. Further humidification of these fibres gave an S conformation (Arnott et al., 1980), which is regarded as a left-handed conformation. There is strong meridional reflection on layerline 6,

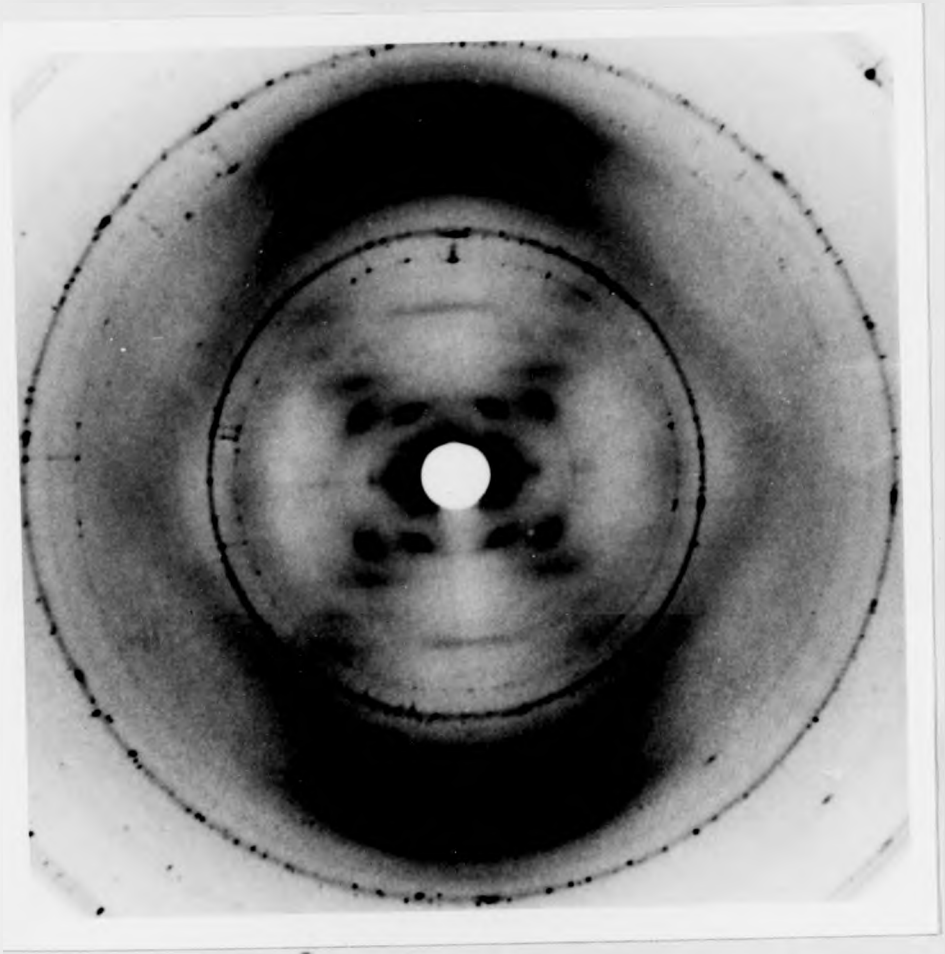


Figure 7 : Semi-crystalline B"-form Diffraction Pattern of K poly d(GC).poly d(GC) at 66% R.H.

reflecting the dinucleotide repeat, and a very strong meridional reflection on the 12th layerline, due to the stacking of bases (discussed in detail in Chapter 8). Therefore, the S conformation has a $6_5$ symmetry with a dinculeotide repeat.

Finally, in conclusion, the above ideas are summarised in terms of the features of polynucleotide fibres which are most likely to affect backbone conformation.

1. The base sequence:

It is evident from the above discussion the (GC) alternating sequence gave the S conformation and B" conformation; both have an alternating backbone conformation. In (AT) alternating sequence the B' and D' conformations are also likely to have an alternating conformation.

2. The ion concentration of the fibre is an important factor relating to backbone conformation. High salt concentrations are one of the favourable conditions for the alternating conformation.

3. The fifth position of thymine might be an important factor in alternating (AT) sequence. When this is replaced by H, $CH_3$ and Br (or I), the base stacking between T and A might be enhanced respectively. This might be the reason for the afinity of protein binding, or why enzyme activity is enhanced when the fifth position of thymine is replaced by H, $CH_3$ and Br (or I) respectively.

4. Both the sequence of the polynucleotides and the conformational variation of the backbone might be key factors relating to the different behaviours of the alternating copolymers in their activities.

# CHAPTER 8

## Polymorphism of Poly d(GC).poly d(GC)

### 8.1 Introduction

The DNA polymer containing guanine and cytosine residues, poly d(GC).poly d(GC) can exist in different forms. These conformational variations in the structure of the nucleic acid duplex may play a fundametnal role in the recognition of specific DNA sequences by proteins. The observation by Wang et al. (1979) that the oligonucleotide dCpG$(pCpG)_2$ crystallises as a two-stranded complex with a left-handed helical conformation has had a major impact on thinking about the structure of DNA. The three dimensional structure of the above crystal was solved to atomic resolution and it revealed an unusual conformation of the DNA double helix in which the guanine and crytosine residues form the usual Watson-Crick base-pairs, and the sugar phosphate backbone pursues an irregular, 'zig-zag' anti-parallel chain. Thus this is called Z-DNA. Studies by Pohl and co-workers (Pohl and Jovin, 1972; Pohl et al., 1972, Pohl et al., 1973; Pohl, 1976) had previously shown that in solution the poly d(GC).poly d(GC) double-helix underwent a cooperative transition with increasing ionic strength from a conformation designated as R which was identified with the righ-handed B form, to a quite distinct form designated as L. The symbols R and L reflect the fact that there is a near inversion of the circular dichroism spectrum following the R → L transition. A number of possible origins for this inversion were suggested including a change from anti- to syn in the sugar-base conformation and a change in handedness of the helix from right to left (Pohl and Jovin, 1972). The studies by Wang et al.,(1979) and subsequent studies by Drew et al., (1980) of the

crystal structure of the oligonucleotide d(CpGpCpG) allowed the direct visualisation of the left-handed double helices, and as such, could be used to account for observations from the poly d(GC).poly d(GC) duplex. In particular, Arnott et al. (1980) reported that a structure very similar to that observed in crystals of $dCpG(pCpG)_2$ could, if extended to form a two-stranded polynucleotide, account for an X-ray fibre diffraction pattern observed from Na poly d(GC).poly d(GC) which they designated as the S form. Wang et al. (1981) emphasised that the Z form should be regarded as a family of similar structures rather than a single conformation. They analysed the various Z type structures observed in oligonucleotide single crystals in terms of varying contributions from two conformations which they nominated $Z_I$ and $Z_{II}$. They suggested that the relative contribution of the two forms was related to the ionic environment of the oligonucleotide and showed that the Fourier transform calculated for the $Z_I$ conformation was in reasonable agreement with the observed fibre diffraction for the S-form. Parallel laser-Raman studies (Thamann et al., 1981) on crystals of d(CpGpCpGpCpG) and on solutions of poly d(GC).poly d(GC) in low and high ionic strength provide direct evidence for identifying the Z-form in crystals with the L-form in solution.

X-ray fibre diffraction studies of poly d(GC).poly d(GC) have concentrated on determining which conformations can be assumed by this polynucleotide in fibres and also in identifying the structural transitions which can occur. Arnott et al., (1980) reported that they observed the S form after prolonged annealing of fibres of Na poly d(GC).poly d(GC) which had previously been given the B form. These fibres contained 3.6% retained NaCℓ. For fibres containing a minimum of retained salt the A form was observed; this form persisted without change to the S form even after annealing for 30 months. An earlier report by Arnott et al. (1974b) that the D form was assumed by poly d(GC).poly d(GC) was qualified by

Leslie et al. (1980). Sasisekharan and Brahmarchari (1981) have recently reported that fibres of poly d(GC).poly d(GC) drawn from 1:1 (V/V) water-ethanol initially gave a B pattern at 40% relative humidity then underwent a transition through a modified B-type pattern to a mixture of B and S, and ultimately to an S pattern. This transition was achieved by increasing the relative humidity to 92% and then reducing it to 40% and also by exposing the fibre to room humidity for a few days. X-ray fibre diffraction patterns of the S form have also been reported by Behe et al., (1981) from fibres of Na poly d(GC).poly d(GC) and Na poly d(G-5MeC). poly d(G-5MeC). These two polymers gave identical S patterns although fibres of the former contained 4.5% excess NaCl by weight and the fibres of the latter had no excess salt. These workers also reported that they observed a transition within oriented fibres from the B form (assumed to be right-handed) to the left-handed S form.

The rest of the chapter is devoted to a description of the conditions for observing the various forms of the poly d(GC).poly d(GC) double helix, and for inducing transitions between them. Attention has also been focussed on those conformations and transitions which are characteristic of poly d(GC).poly d(GC) and in particular the relationship of the S form of this duplex to the D form of poly d(AT).poly d(AT).

## 8.2 Materials and Method

Poly d(GC).poly d(GC) was provided by Dr. J. Brahms, University of Paris and also bought from Boehringer. The fibres were prepared as described in Chapter 2. X-ray fibre diffraction patterns were recorded in a range of relative humidities from 33% to 98%.

## 8.3 Results and Discussion

The fibres prepared from the material precipitated from 0.1 M $KF$ solution were first studied. These fibres gave, at low humidity (typically

between 33% and 66%), a pattern which had not previously been described, but which was somewhat similar to the semi-crystalline B-patterns observed from fibres of the sodium salt of naturally occurring DNAs at high humidity (figure 1). It was designated B"-DNA. When the relative humidity was increased (typically from 66% to 75%) a structural transition to the A-form was observed (figure 2). Further increase in the relative humidity (typically from 75% to 92%) led to a transition to the S form (figure 3). On some occasions patterns were observed which showed a mixture of the A and S forms (figure 4). Once the fibre had fully assumed the S form no further transitions were observed unless the fibre was exposed to a very high humidity (typically 95% or greater). Under these conditions a transition occurred to a semi-crystalline B form (figure 5). On some occasions, patterns were observed which indicated a mixture of the S and semi-crystalline B-form (figure 6). Gradual reduction of the relative humidity from 98% to 92% resulted in a transition back to the S-form. On some occasions a mixture of the S and a semi-crystalline B form was observed, as seen in figure 6. Once again this form was stable and persisted even when the relative humidity was reduced to 33%. However, if the relative humidity of the fibre environment was reduced, more rapidly from 98% to 75% the S form was not observed. Following further reduction in the relative humidity into the range 33% to 66% the B" pattern was recovered. If the fibre was re-wet so that a gel was formed and a new fibre drawn, a B" pattern was again observed at low humidity and subsequent variation in the relative humidity reproduced the sequence of changes detailed above. The sequence of these transitions is schematically shown in figure 7.

This is the first report in which structural transitions within a nucleic acid fibre have been shown to depend on the rate at which the relative humidity of the fibre environment is varied. It should be emphasised that these transitions involve major structural changes in the poly d(GC).poly d(GC) double-helix in which the high humidity semi-crystalline

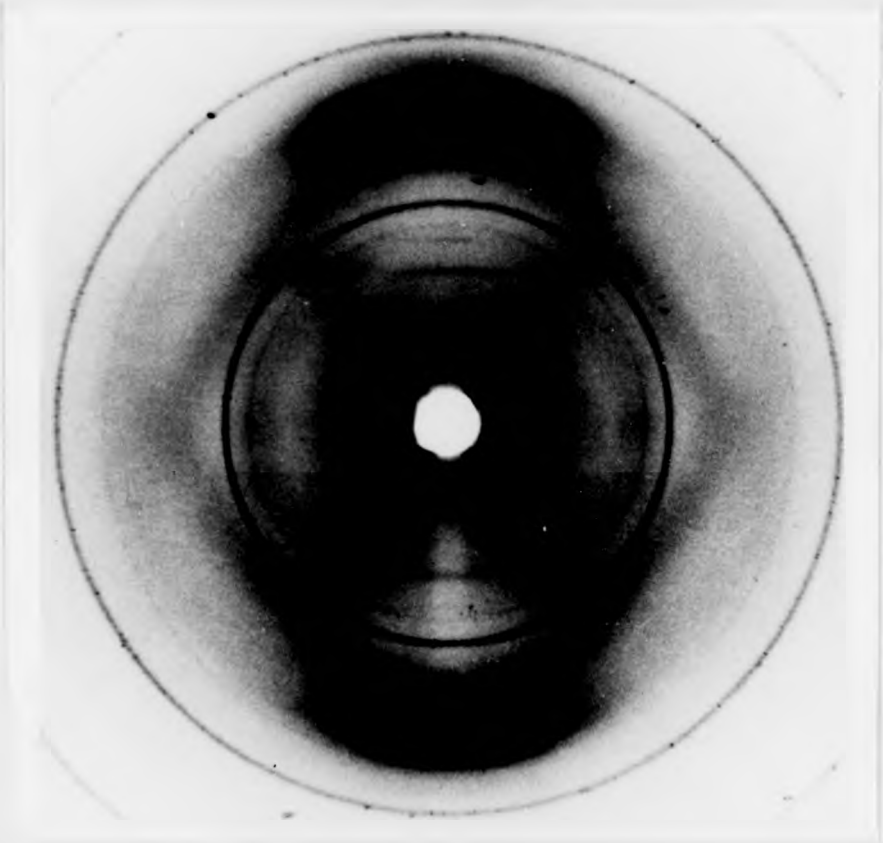


Figure 1 : Semi-crystalline B"-form Diffraction Pattern of K poly d(GC).poly d(GC) at 66% R.H.

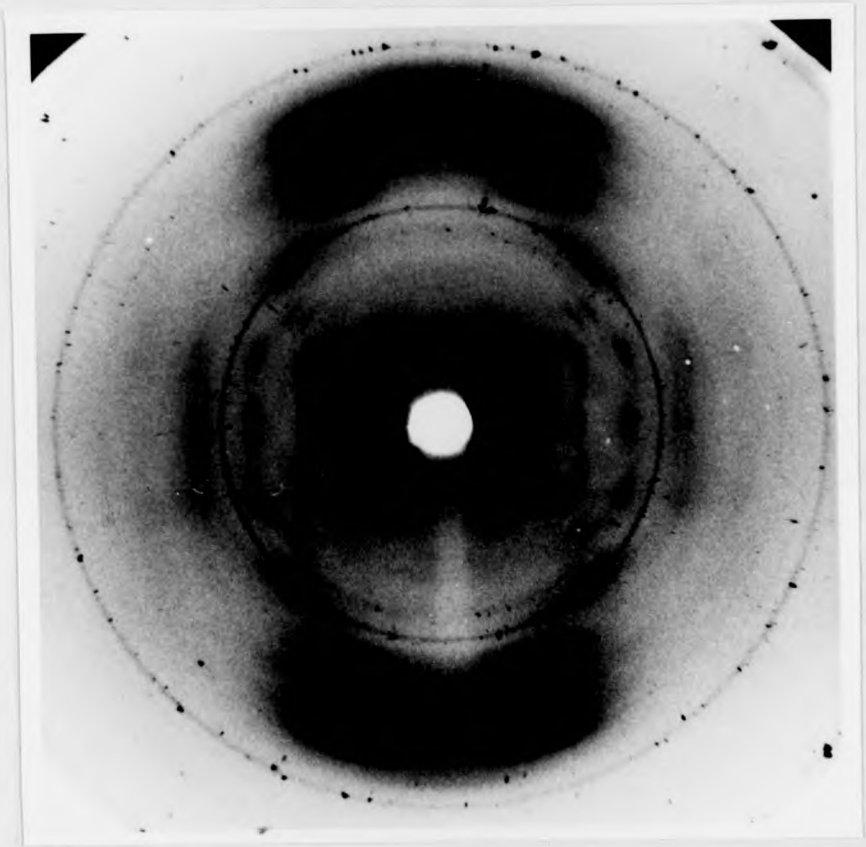


Figure 2 : A-form Diffraction Pattern of
K poly d(GC).poly d(GC) at 75% R.H.

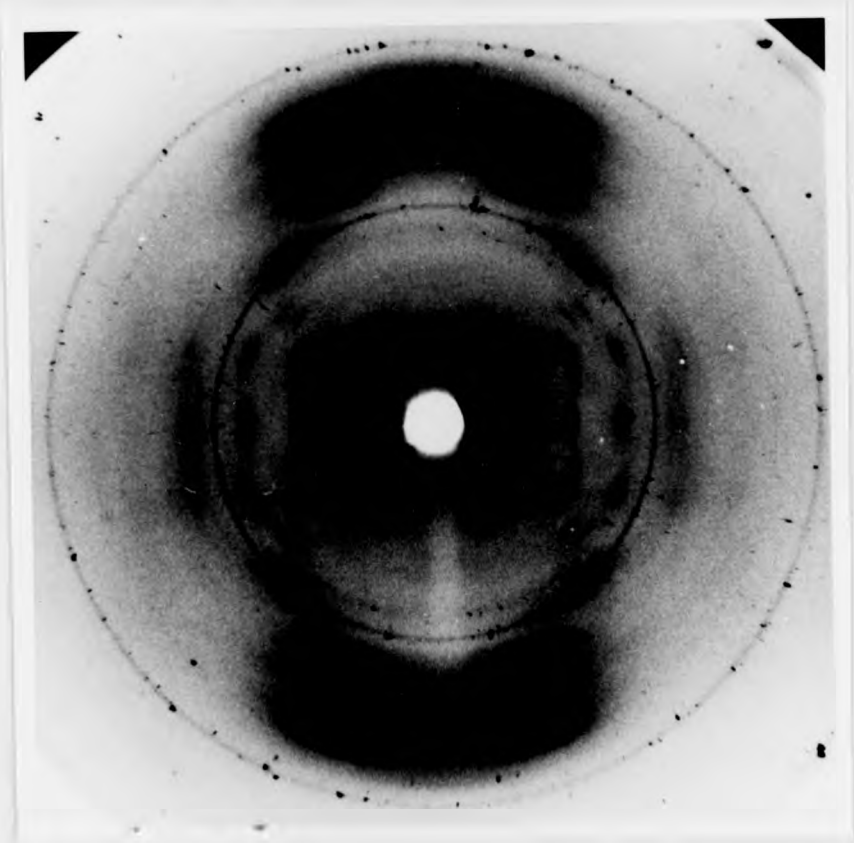


Figure 2 : A-form Diffraction Pattern of
K poly d(GC).poly d(GC) at 75% R.H.

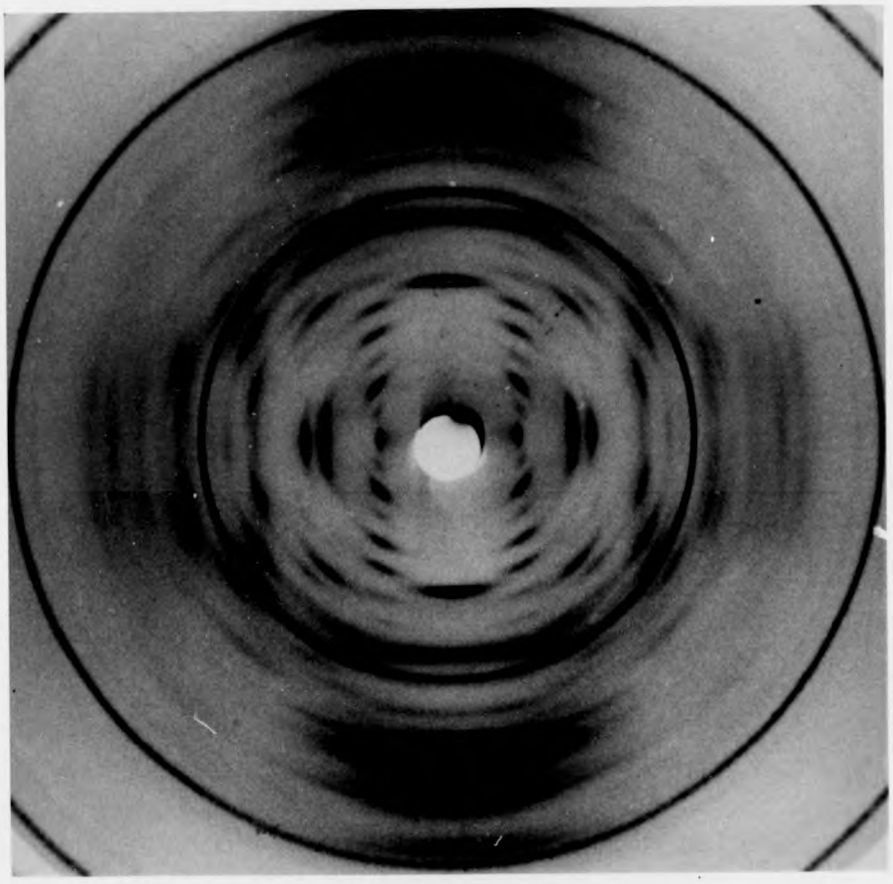


Figure 3 : Crystalline S-form Diffraction Pattern of K poly d(GC).poly d(GC) at 66% R.H.

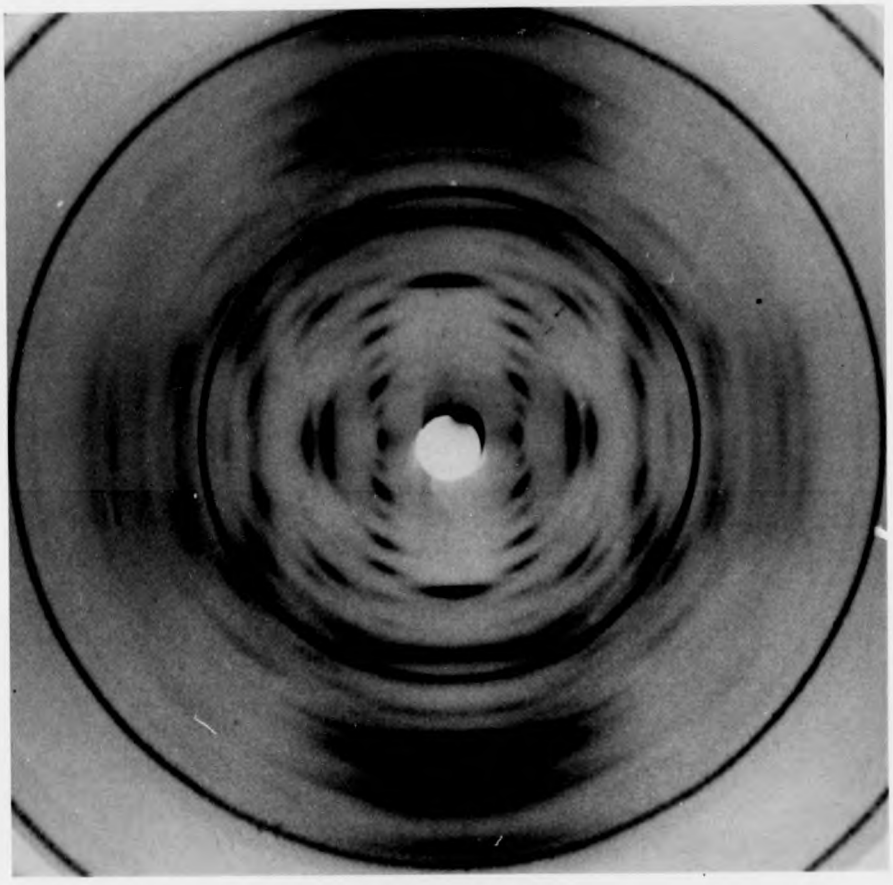


Figure 3 : Crystalline S-form Diffraction Pattern of K poly d(GC).poly d(GC) at 66% R.H.

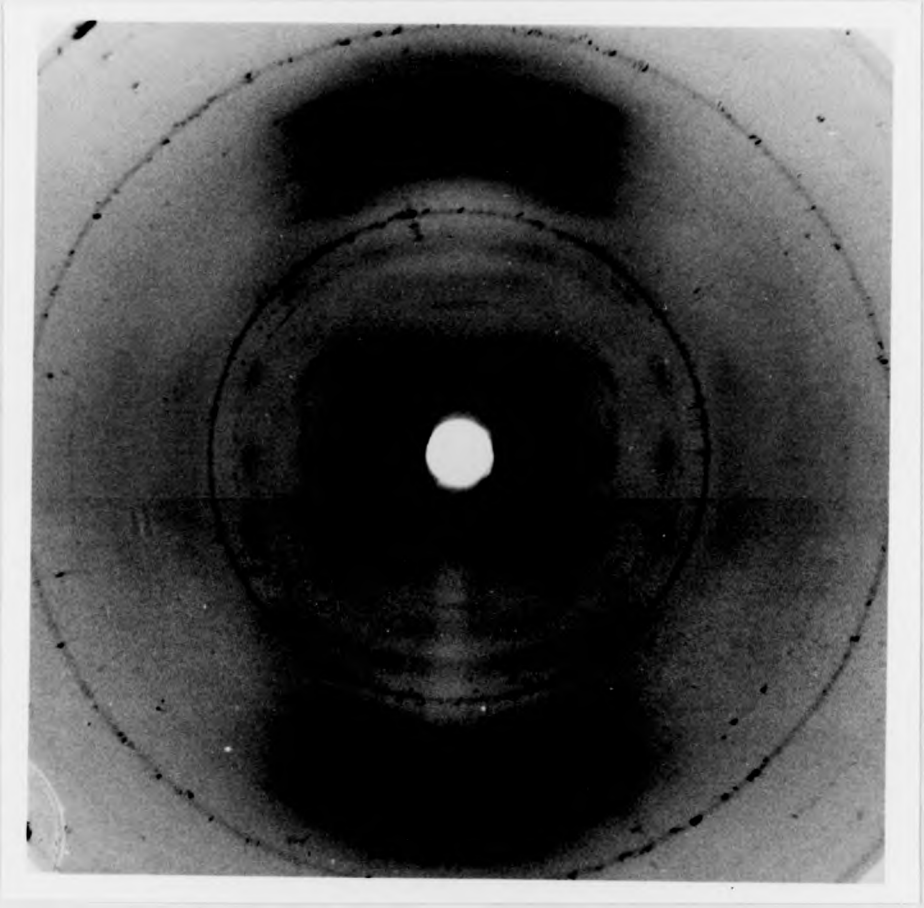


Figure 4 : Mixture of A and S forms. Diffraction Pattern of K poly d(GC).poly d(GC) at 86% R.H.

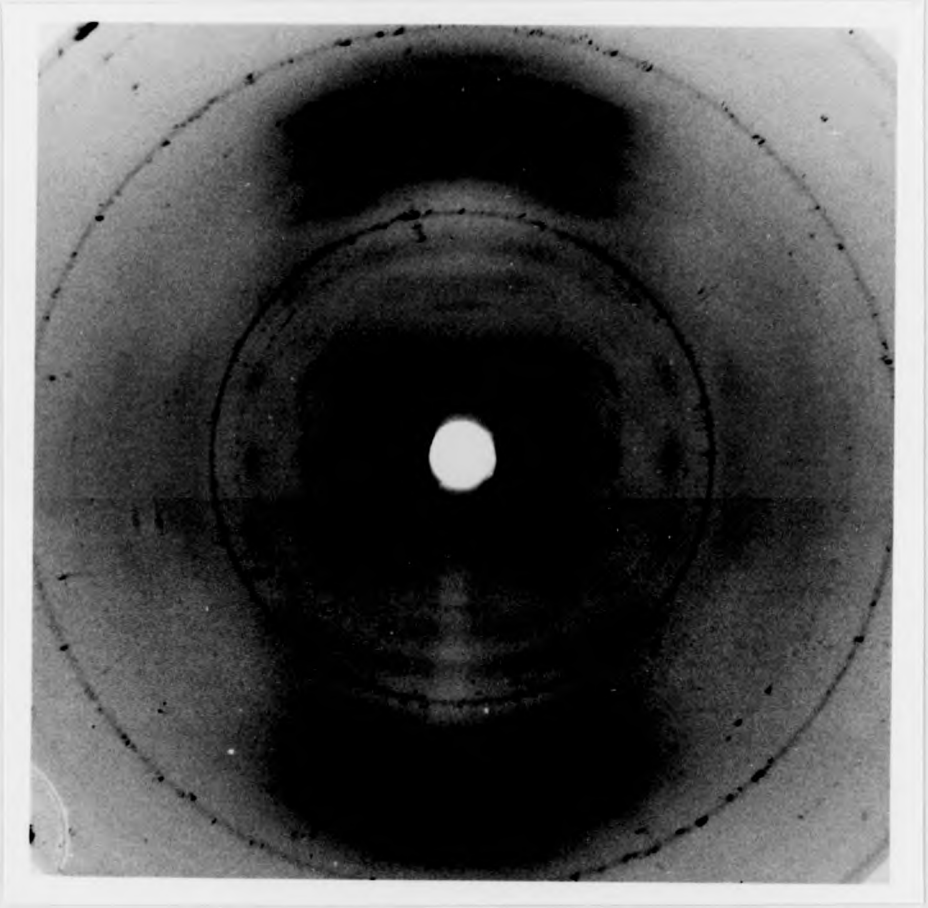


Figure 4 : Mixture of A and S forms. Diffraction Pattern of K poly d(GC).poly d(GC) at 86% R.H.

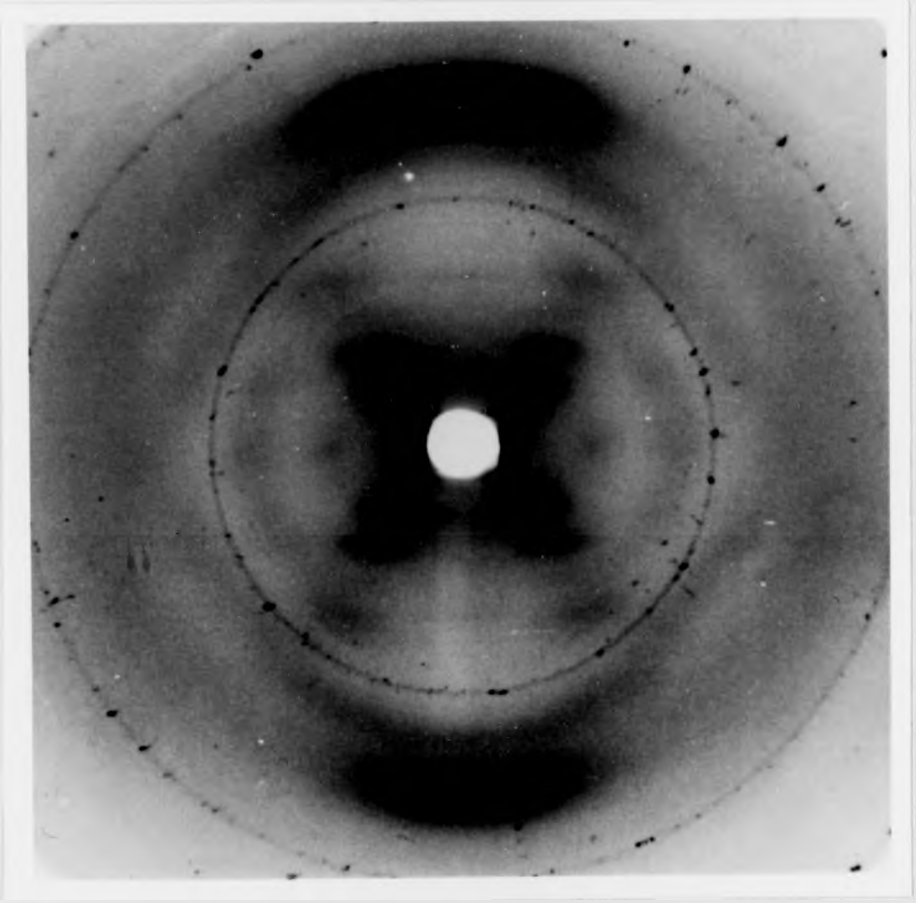


Figure 5 : Semi-crystalline B-form. Diffraction Pattern of K poly d(GC).poly d(GC) at 98% R.H.

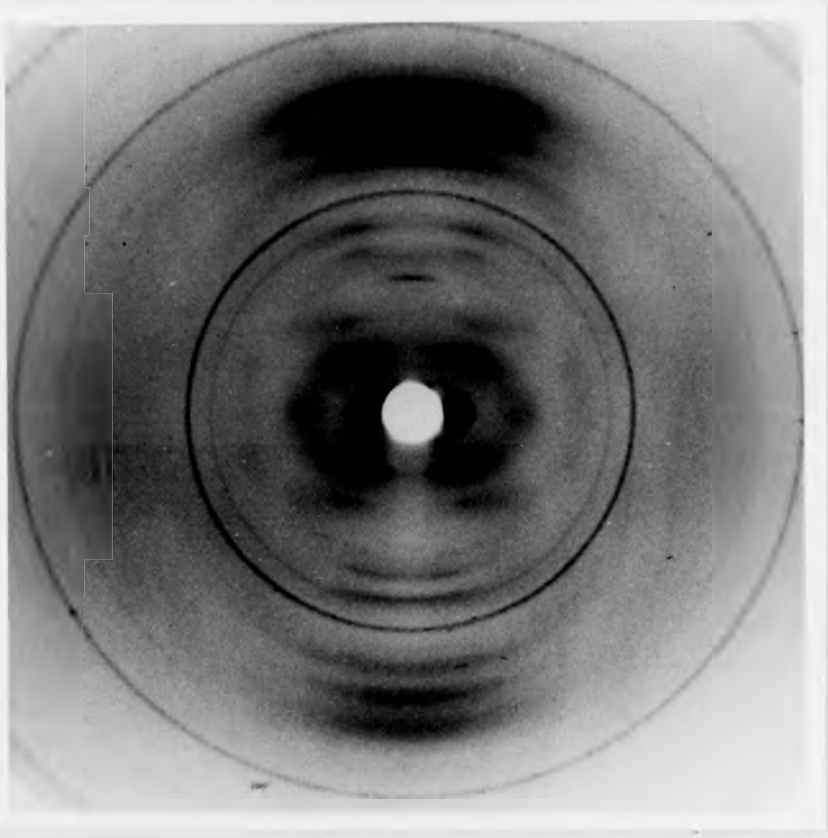


Figure 6 : Mixture of B and S forms. Diffraction Pattern of K poly d(GC).poly d(GC) at 92% R.H.

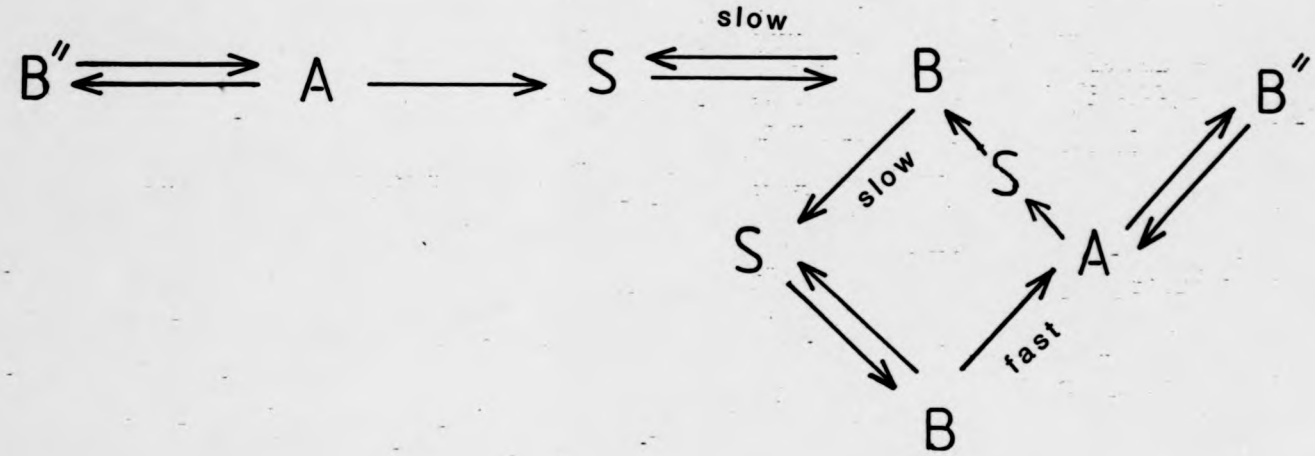



Figure 7 : Conformational transitions in K poly d(GC).poly d(GC) fibre is shown schematically

B form changes rapidly to the A-form or more slowly to the S-form. While the preliminary analysis of the X-ray fibre diffraction pattern of the S-form is not definitive, a left-handed model derived from the Z conformation observed in single crystals of $dCpG(pCpG)_2$ does appear to account in a satisfactory way for the observed fibre diffraction data. If the S form is indeed a left-handed helix, and assuming (as is generally accepted) that the A and B conformations are right-handed, then both the B → S and A → S transitions involve a change in helix sense. In this context it should be noted that whilst the possibility of the A and B forms having a left-handed helical sense has not seriously been canvassed (Gupta et al., 1980), models of this type having a good stereochemistry and as good a fit with the observed data have yet to be obtained. The problems of constructing a left-handed model for the A form are particularly acute (Fuller et al., 1965). While transitions between left- and right-handed helices in solution are readily visualised, there is much greater difficulty in envisaging such changes within a fibre. However, it should be recognised that at the relative humidities at which the A → S and B → S transitions occur, the fibre contains between a third and a half by volume of water, so that there may well be sufficient spatial freedom for transitions involving a change as profound as reversal of helix sense to occur. It may well be, however, that a transition involving a change in helix sense will require more time to occur, and this might conceivably explain the dependence of transitions on the rate of change of humidity. It may also be that spatial freedom is more limited when the water content of the fibre is being reduced as compared with situations in which it is being increased. This difference may provide a reason why within approximately the same relative humidity range the A → S transition occurs with increasing relative humidity, but the S → A transition does not occur when the relative humidity is decreased.

These observations can be related to studies by Ivanov and Minyat (1981) of transitions between the A, B and Z forms of poly d(GC).poly d(GC) in water/trifluorethanol solutions. They found conditions under which there were reversible transitions between all pairs of these three forms. However, transitions which involved the Z form were slower by a factor of $\sim$ 100 than those which did not involve this form.

Sasisekharan and Brahmachari (1981) have taken the view that the observation of B $\rightarrow$ S transition within a fibre (which they identify with the 'solid state') clearly demonstrates "the transition cannot involve a change of handedness of the duplex". They propose, therefore, that the overall handedness of the duplex must be the same in both the B and S forms. They claim that the transition can be readily understood in terms of the so-called RL model of DNA(Sasisekharan et al., 1978) in which short regions of left- and right-handed helices alternate along the length of the molecule. While such models may well be important as a feature of DNA structure in vivo it is important to emphasise that if such a model is to account for the above observations on B $\rightarrow$ Z and A $\rightarrow$ Z transitions in fibres it is necessary for models of the RL type to be constructed which will account for the A and B X-ray fibre diffraction patterns at least as well as the currently generally accepted right-handed models. To date this has not been achieved (Greenall et al., 1979).

Arnott et al. (1980) noted that the S patterns they had obtained correspond to a "statistical" crystal structure in which there were both Bragg reflections and continuous streaks. Such disorder is less apparent in the S patterns obtained in the present work (see figure 3) which could be regarded as fully crystalline. The degree of crystallinity and orientation in these patterns is sufficiently high for the (0,0,12) and (1,0,12) reflections to be resolved. This resolution is important for establishing the number of residues per turn of the helix and in particular

excluding models with 11 residues per turn. Further, the relative intensity these two reflections have is likely to be particularly sensitive to the tilt of the base-pairs and their distance from the helix axis. In general the quality of the X-ray diffraction data now available for the S form is comparable with that for the crystalline A and B forms offers the prospect of a definitive structure determination.

The pattern in figure 3 was recorded when the relative humidity of the fibre environment was 66%. However, S patterns were recorded at relative humidities from 33% to 92% with the lattice variation indicated in figures 8a and 8b. The high degree of stability of the S conformation is indicated by very small variation in these parameters for relative humidities from 75% down to 44%. However, at 92% relative humidity the lattice parameters are observed to have a sharp increase. This also corresponds to a change in the overall intensity distribution on the S pattern. As the relative humidity was increased from 44% to 98% the following intensity changes were observed in the diffraction patterns (see figure 3 and figure 9a, b):-

1. The 11th layerline became weaker, suggesting that the tilt of the base is probably changing with increasing relative humidity (see figures 9a, b).
2. The S pattern observed at relative humidities of 75% and below have a strong 108 reflection, and a very weak 107 reflection. However, the S pattern observed at a relative humidity of 92% and above have approximately equal intensity for 108 and 107 reflections (see figures 9a, b).
3. Low humidity S patterns have a strong 105 reflection, and higher humidity S patterns have a weak 105 reflection (see figures 9a, b).
4. Low humidity S patterns have a strong (006) reflection and high humidity S patterns have a weak (006) reflection. However, both patterns have strong (00 12) reflections. This observation suggests

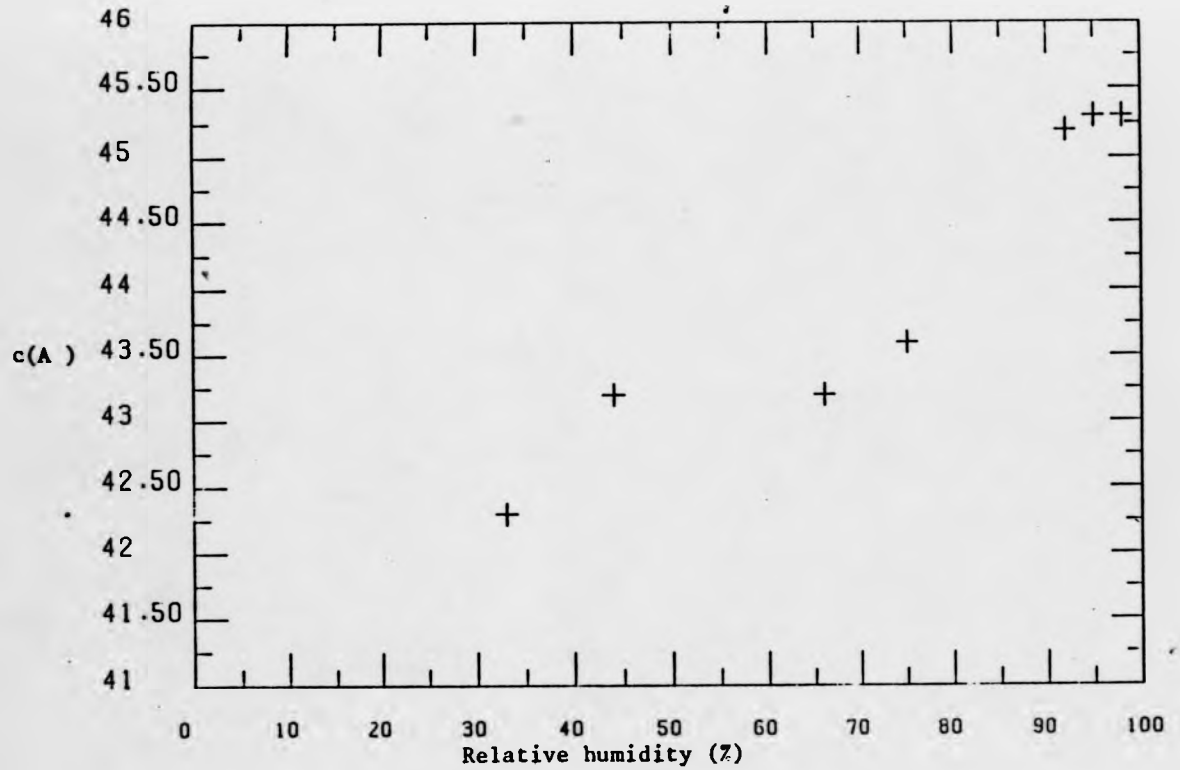



Figure 8a : c(A ) Vs relative humidity (%)

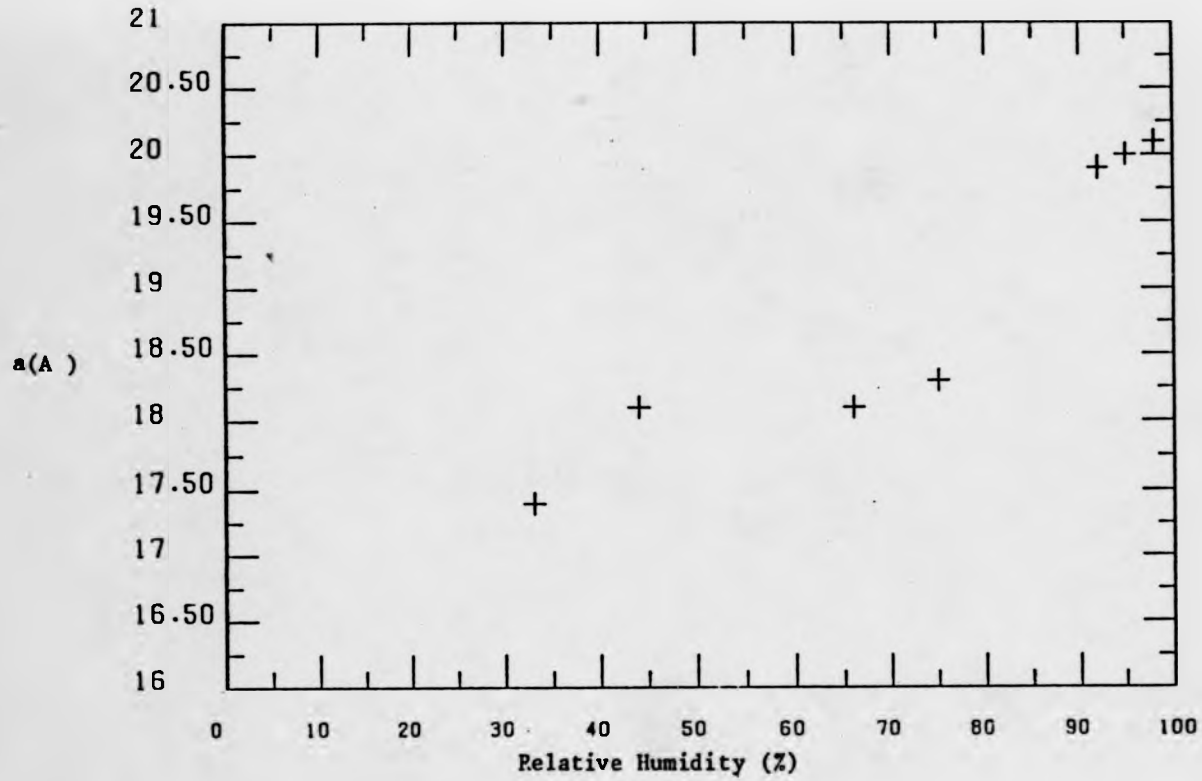



Figure 8b : a(A ) Vs relative humidity (%)

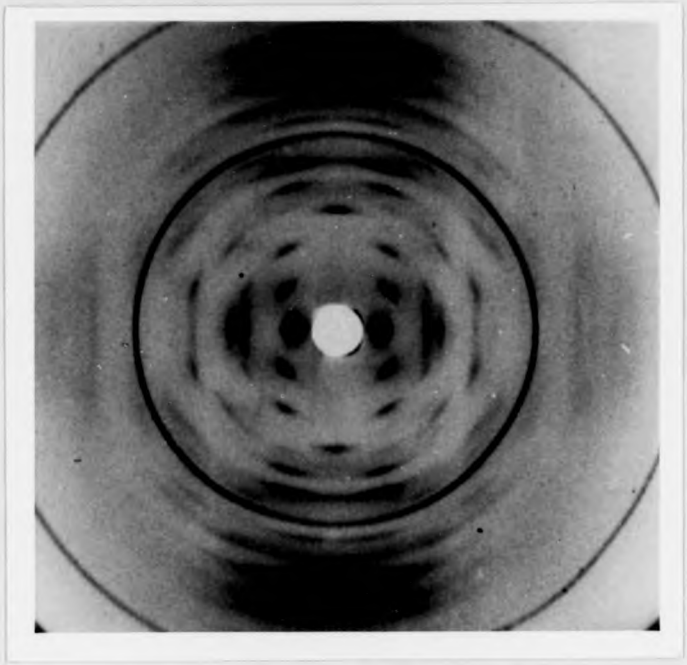


Figure 9a : Crystalline $S_{II}$-type form Diffraction Pattern of K poly d(GC).poly d(GC) at 95% R.H.

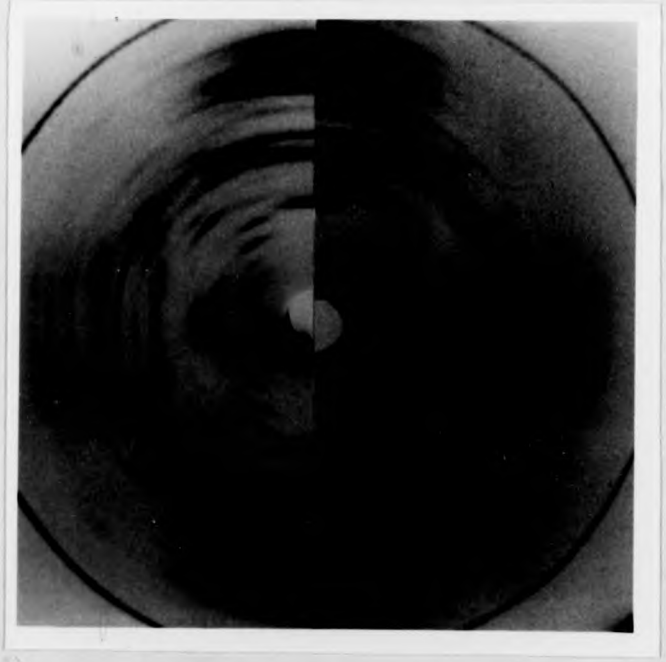


Figure 9b : $S_I$-type (left) and $S_{II}$ (right) forms. Diffraction Pattern of K poly d(GC).poly d(GC)

that the phosphate positions in the relative humidity S conformation differ less in relative position on a regular backbone basis than they do in the sugar-phosphate backbone conformation of the low humidity conformation (i.e. $P\text{-}O_3$, $P\text{-}O_5$ conformations are different in two S conformations).

5. The lattice parameters show a sharp increase at 92% and then remain more or less the same up to 98% (in some fibres they change to semi-crystalline B-form at 95% - conditions are discussed later in this chapter).

The lattice parameters are given below:-

(a) low humidity pattern at 66%

a = 17.85 A c = 42.80 A

(b) high humidity pattern at 95%

a = 20.0 A c = 45.35 A

The $\rho$ values observed were used to refine the lattice parameters from relative humidity 33% to 98%. They are given on table 1 (low humidity patterns) and table 2 (high humidity patterns).

6. A mixture of the two conformations was observed at 92% or 98% relative humidities (see figure 10). From this pattern the lattice parameters of the two components were determined:-

(a) Major component

a = 19.6 A c = 44.4 A

which is similar to the high relative humidity conformation

(b) Minor component

a = 18.0 A c = 43.0 A

which is the low humidity conformation.

7. The observed intensities of these two patterns were compared with the squared transforms calculated from extended $Z_I$ and $Z_{II}$ helices proposed by Wang et al. (1981). The low humidity patterns showed

Analysis of the lattice parameters as a function of relative humidity

Table 1a

Observed and calculated $\rho$ values for the low relative humidity S patterns ($S_I$-type)

| RH = 33% | | | a = 17.4 A | c = 42.3 A |
|---|---|---|---|---|
| h | k | ℓ | $\rho_0(A^{-1})$ | $\rho_c(A^{-1})$ |
| 1 | 1 | 0 | 0.1148 | 0.1149 |
| 2 | 0 | 0 | 0.1325 | 0.1327 |
| 1 | 0 | 2 | 0.0816 | 0.0815 |
| 1 | 0 | 4 | 0.1160 | 0.1155 |
| 1 | 0 | 5 | 0.1354 | 0.1356 |
| 1 | 1 | 5 | 0.1642 | 0.1649 |
| 0 | 0 | 6 | 0.1417 | 0.1418 |

| RH = 44% | | | a = 18.1 A | c = 43.2 A |
|---|---|---|---|---|
| h | k | ℓ | $\rho_0(A^{-1})$ | $\rho_c(A^{-1})$ |
| 1 | 0 | 0 | 0.0625 | 0.0638 |
| 1 | 1 | 0 | 0.1111 | 0.1106 |
| 2 | 0 | 0 | 0.1276 | 0.1277 |
| 1 | 0 | 2 | 0.0787 | 0.0789 |
| 1 | 2 | 2 | 0.1754 | 0.1751 |
| 1 | 2 | 4 | 0.1928 | 0.1926 |
| 1 | 0 | 4 | 0.1124 | 0.1125 |
| 1 | 0 | 5 | 0.1328 | 0.1323 |
| 0 | 0 | 6 | 0.1389 | 0.1390 |
| 1 | 1 | 5 | 0.1597 | 0.1602 |
| 1 | 0 | 8 | 0.1961 | 0.1961 |

Cont.

Table 1a (Cont.)

| RH = 66% | | | a = 18.1 A | c = 43.2 A |
|---|---|---|---|---|
| h | k | ℓ | $\rho_0(A^{-1})$ | $\rho_c(A^{-1})$ |
| 1 | 0 | 0 | 0.0629 | 0.0637 |
| 1 | 1 | 0 | 0.1109 | 0.1103 |
| 2 | 0 | 0 | 0.1273 | 0.1273 |
| 1 | 0 | 2 | 0.0789 | 0.0787 |
| 1 | 2 | 2 | 0.1749 | 0.1747 |
| 1 | 0 | 4 | 0.1128 | 0.1123 |
| 1 | 2 | 4 | 0.1919 | 0.1921 |
| 1 | 0 | 5 | 0.1318 | 0.1320 |
| 1 | 1 | 5 | 0.1595 | 0.1598 |
| 0 | 0 | 6 | 0.1382 | 0.1388 |
| 1 | 0 | 8 | 0.1962 | 0.1956 |

| RH = 75% | | | a = 18.3 A | c = 43.6 A |
|---|---|---|---|---|
| h | k | ℓ | $\rho_0(A^{-1})$ | $\rho_c(A^{-1})$ |
| 1 | 0 | 0 | 0.0620 | 0.0633 |
| 1 | 1 | 0 | 0.1099 | 0.1096 |
| 2 | 0 | 0 | 0.1263 | 0.1266 |
| 1 | 0 | 2 | 0.0780 | 0.0782 |
| 1 | 2 | 2 | 0.1735 | 0.1736 |
| 1 | 0 | 4 | 0.1119 | 0.1115 |
| 1 | 2 | 4 | 0.1915 | 0.1909 |
| 1 | 0 | 5 | 0.1312 | 0.1310 |
| 0 | 0 | 6 | 0.1371 | 0.1376 |

Table 1b

Observed and calculated ρ values for the S pattern ($S_I$) in figure 3. This pattern was used in the structural analysis of $S_I$-conformation

RH = 66% a = 17.9 A c = 42.8 A

| h | k | ℓ | $\rho_o(A^{-1})$ | $\rho_c(A^{-1})$ |
|---|---|---|---|---|
| 1 | 0 | 0 | 0.0636 | 0.0646 |
| 1 | 1 | 0 | 0.1119 | 0.1119 |
| 2 | 0 | 0 | 0.1289 | 0.1292 |
| 2 | 1 | 0 | 0.1719 | 0.1710 |
| 3 | 0 | 0 | 0.1956 | 0.1939 |
| 3 | 2 | 0 | 0.2821 | 0.2817 |
| 1 | 0 | 1 | 0.0685 | 0.0687 |
| 1 | 0 | 2 | 0.0798 | 0.0798 |
| 2 | 1 | 2 | 0.1774 | 0.1773 |
| 1 | 0 | 3 | 0.0951 | 0.0955 |
| 2 | 2 | 3 | 0.2339 | 0.2346 |
| 3 | 1 | 3 | 0.2423 | 0.2433 |
| 1 | 0 | 4 | 0.1138 | 0.1138 |
| 2 | 1 | 4 | 0.1951 | 0.1950 |
| 1 | 0 | 5 | 0.1339 | 0.1337 |
| 1 | 1 | 5 | 0.1619 | 0.1620 |
| 2 | 0 | 5 | 0.1740 | 0.1744 |
| 2 | 1 | 5 | 0.2075 | 0.2072 |
| 3 | 0 | 5 | 0.2262 | 0.2265 |
| 0 | 0 | 6 | 0.1405 | 0.1405 |
| 1 | 0 | 6 | 0.1548 | 0.1547 |
| 1 | 0 | 7 | 0.1758 | 0.1762 |
| 1 | 0 | 8 | 0.1993 | 0.1982 |
| 2 | 0 | 8 | 0.2271 | 0.2276 |

TABLE 2

Observed and calculated ρ values for the high relative humidity S patterns ($S_{II}$-type)

| RH = 92% | | | a = 19.9 A | c = 45.2 A |
|---|---|---|---|---|
| h | k | ℓ | $\rho_o(A^{-1})$ | $\rho_c(A^{-1})$ |
| 1 | 0 | 0 | 0.0559 | 0.0581 |
| 1 | 1 | 0 | 0.1013 | 0.1006 |
| 2 | 0 | 0 | 0.1164 | 0.1161 |
| 1 | 0 | 2 | 0.0725 | 0.0730 |
| 1 | 0 | 4 | 0.1058 | 0.1059 |
| 0 | 0 | 6 | 0.1321 | 0.1328 |
| 1 | 0 | 6 | 0.1488 | 0.1449 |
| 1 | 0 | 7 | 0.1646 | 0.1654 |
| 1 | 0 | 8 | 0.1852 | 0.1863 |
| 0 | 0 | 12 | 0.2655 | 0.2655 |
| 1 | 0 | 12 | 0.2717 | 0.2718 |

| *RH = 95.1 | | | a = 20.0 A | c = 45.3 A |
|---|---|---|---|---|
| h | k | ℓ | $\rho_o(A^{-1})$ | $\rho_c(A^{-1})$ |
| 1 | 0 | 0 | 0.0568 | 0.0578 |
| 1 | 1 | 0 | 0.0998 | 0.1001 |
| 2 | 0 | 0 | 0.1153 | 0.1156 |
| 2 | 1 | 0 | 0.1543 | 0.1530 |
| 1 | 1 | 1 | 0.1022 | 0.1025 |
| 1 | 1 | 2 | 0.1085 | 0.1094 |
| 2 | 1 | 2 | 0.1593 | 0.1592 |
| 1 | 0 | 4 | 0.1060 | 0.1056 |
| 1 | 1 | 4 | 0.1335 | 0.1335 |
| 2 | 0 | 4 | 0.1458 | 0.1455 |
| 0 | 0 | 6 | 0.1324 | 0.1325 |
| 1 | 0 | 6 | 0.1464 | 0.1446 |
| 1 | 0 | 7 | 0.1647 | 0.1650 |
| 1 | 0 | 8 | 0.1854 | 0.1859 |
| 1 | 0 | 10 | 0.2278 | 0.2283 |

*This pattern was used in the structural analysis of $S_{II}$ conformation and shown in figure 9.

| RH = 98% | | | a = 20.1 A | c = 45.3 A |
|---|---|---|---|---|
| h | k | ℓ | $\rho_o(A^{-1})$ | $\rho_c(A^{-1})$ |
| 1 | 0 | 0 | 0.0563 | 0.0575 |
| 1 | 1 | 0 | 0.1002 | 0.0996 |
| 2 | 0 | 0 | 0.1147 | 0.1150 |
| 1 | 0 | 2 | 0.0727 | 0.0725 |
| 1 | 1 | 2 | 0.1082 | 0.1089 |
| 2 | 1 | 2 | 0.1591 | 0.1584 |
| 1 | 0 | 4 | 0.1054 | 0.1053 |
| 2 | 0 | 4 | 0.1449 | 0.1449 |
| 0 | 0 | 6 | 0.1324 | 0.1323 |
| 1 | 0 | 7 | 0.1647 | 0.1647 |
| 1 | 0 | 8 | 0.1855 | 0.1856 |
| 1 | 0 | 10 | 0.2279 | 0.2279 |
| 1 | 0 | 12 | 0.2708 | 0.2708 |

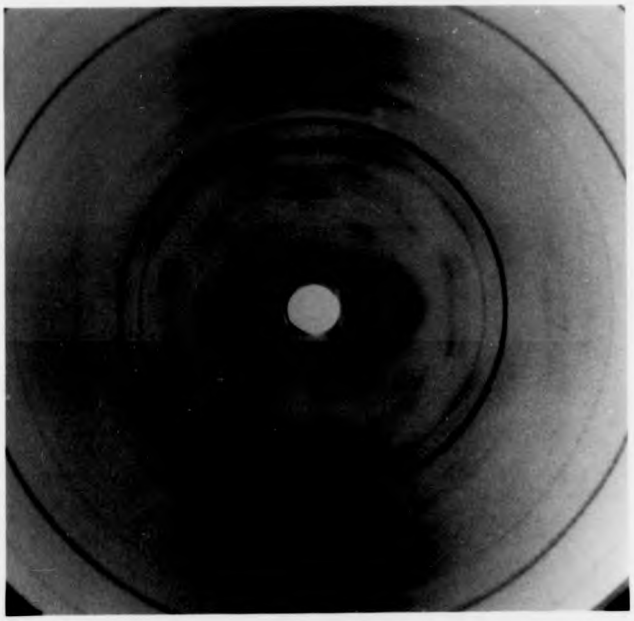


Figure 10 : Mixture of $S_I$ and $S_{II}$ forms. Diffraction Pattern of K poly d(GC).poly d(GC) at 92% R.H.

strikingly good agreement with $Z_I$ helix transform (see figure 11). However, there are few discrepancies which need to be corrected by adjusting the $Z_I$ helix. The discrepancies are:-

(a) The transform predicts a weak 11th layerline. However the observed intensities are higher than predicted by $Z_I$ helix. This could be corrected by changing the tilt of the bases.

(b) The (105) reflection is fairly strong in the diffractin pattern and the model predicts almost zero intensity.

(c) The (100) reflection is fairly weak reflection in the diffraction pattern and the model predicts a very strong intensity.

This could be adjusted by moving the base position relative to the helix axes.

The detailed analysis of the $S_I$ and $S_{II}$ diffraction patterns is discussed in section 8.5.

The semi-crystalline B" pattern (figure 1) is of particular interest. Preliminary measurements indicate that it is 10-fold with a helix pitch of 33.4 A. Since the pattern contains a relatively small number of Bragg reflections the lattice determination must be tentative, but from the initial analysis it appears to be orthorhombic with a = 18.6 A, b = 27.9 A, c = 33.4 A and two double helices passing through each unit cell. The overall intensity distributions is similar to that observed at high humidity from the semi-crystalline B form of DNA. There is however in the B" pattern additional well-defined diffraction on the meridion of the fifth layerline. This diffraction might be from ions or ordered water in the structure, but it is tempting to consider the possibility that it is due to differences in the conformation of the G and C residues analogous to the variation proposed by Klug et al. (1979) in the alternating B form of poly d(AT).poly d(AT). It may be that the B" reported here is closely related if not identical to the modified B pattern observed for poly d(GC).poly d(GC) by Sasisekharan

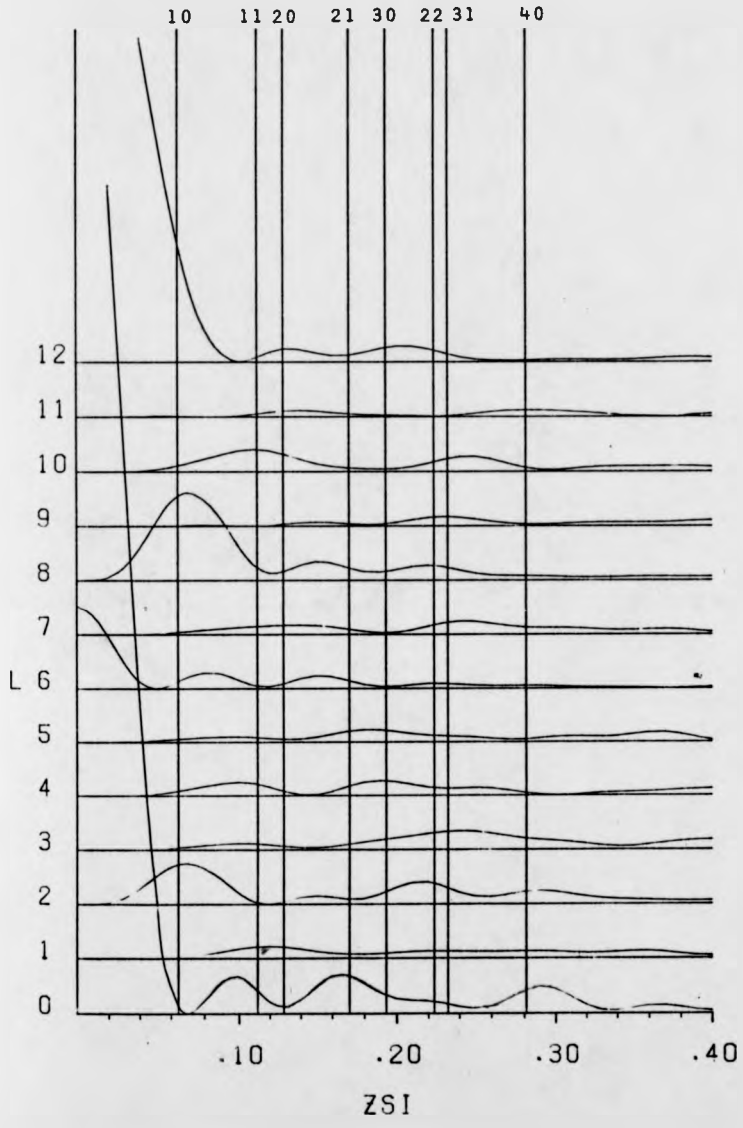



Figure 11 : The cylindircally averaged squared molecular transform for $Z_I$-DNA (Wang et al., 1981)

and Brahmachari (1981). However, the published pattern is not sufficiently well-defined for this to be verified. While the description of this pattern states that it has, in addition to the features of B-DNA, meridional or near-meridional reflections on various layerlines, it makes no particular mention of meridional diffraction on the fifth layerline. It would be of great interest if the KF environment in our studies and the water-ethanol environment in these studies resulted in a similar modification of the B form of the poly d(GC).poly d(GC) double helix.

Early work by Bram and co-workers (Bram, 1971; Bram and Tougard, 1972) and by Pilet and Brahms (1972) and more recently the extensive study by Leslie et al. (1980) and have focussed on the dependence of the conformation of the DNA helix on the base sequence of DNA. The two extreme compositions, i.e. exclusively A-T and exclusively G-C are of crucial interest to such an analysis. Synthetic DNAs with both these compositions can assume the classical A and B forms. However, a major difference occurs in that whereas poly d(AT).poly d(AT) readily assumes the D form (Mahendrasingam et al., 1983), poly d(GC).poly d(GC) apparently does not (Leslie et al., 1980), although the very closely related poly d(IC).poly d(IC) does (Mitsui et al., 1970; Ramaswamy et al., 1982). Conversely, the S form is characteristic of poly d(GC).poly d(GC) and has been reported by Arnott et al. (1980) as being observed occasionally for poly d(AC).poly d(GT). From the point of view of their position in the 'pattern' of conformational transitions, the S and D forms show marked similarities. Once assumed both are particularly stable forms of the duplex. The forms are assumed with either increasing relative humidity for a fibre which has previously been in the A form or with decreasing relative humidity for a fibre previously in the B form. Transitions from both forms have so far only been observed when the relative humidity is raised to a very high level (i.e. typically 95%) when the S → B and D → B changes occur. These similarities are of particular interest since, as can readily

be seen from the X-ray fibre diffraction patterns of the S and D forms, the molecular conformations of the two structures are quite distinct.

Excess salt in the precipitated material was removed as described in Chapter 2. A few fibres with very low salt content were prepared. In environments of relative humidity up to 92% these fibres gave a semi-crystalline C pattern, which is similar to the C pattern of Na DNA(see chapter 9). The absence of the salt ring in this C pattern was an indication of the low salt content in those fibres. Further humidification of these fibres at 95% and 98% relative humidity gave a semi-crystalline B pattern. The C → B transition was found to be reversible. A similar conformational transition was observed for low salt ($\sim$ 0.2 $Na^{+}Cl^{-}pO4^{-}$) Na poly d(AT).poly d(AT) fibres (Mahendrasingam et al., 1983a).

Further increases in the salt content of the fibre gave a semi-crystalline C pattern up to 84% relative humidity. However, at 92% relative humidity an A pattern was observed. Further humidification at 95% relative humidity gave an S and B mixture, and a semi-crystalline B pattern was observed at 98% relative humidity. Weak salt diffraction rings were observed in all these patterns, indicating a slight increase in salt content in the fibre. When the relative humidity was reduced from 98% to 92% relative humidity it gave a semi-crystalline B-form as observed earlier. When this fibre was taken down to 33% relative humidity it gave a semi-crystalline C-form. Further humidification of this fibre to 98% relative humidity made it go through the same conformational sequence which had occurred in the first humidification cycle. One very unique aspect of these fibres was the fact that the rate of change of humidity appeared to be important on the 'decreasing' part of the run. If the humidity was reduced slowly, the transition was B → S, and if fast B → A. Due to the salt content in this fibre, it gave a semi-crystalline C-form instead of an A-form when the relative humidity was lowered rapidly.

Fibres prepared with LiF salt in the fibre gave a semi-crystalline C-form up to 75% relative humidity. The diffraction pattern (figure 12) resembles the semi crystalline C-form from Li DNA. However, the molecular packing appears to be rather different. Further humidification of these fibres gave a crystalline B-form (figure 13) at 86% relative humidity. At 92% and 95% relative humidities a semi-crystalline B pattern (figure 14) was observed. A semi-crystalline B-form which is similar to that obtained from LiDNA and K poly d(GC).poly d(GC) was observed at 98% relative humidity. All of these conformational changes were found to be reversible. The crystalline B-form in figure 13 appears to be different in molecular packing from that of Li DNA as reported by Leslie et al. (1980). Further detailed analysis of these diffraction patterns are in progress in our laboratory.

Because it was found that poly d(AT).poly d(AT) (i.e. poly d (purine-prymidine).poly d(purine-prymidine) gave the D form easily in the presence of heavy ions such as $Rb^+$ and $Cs^+$ in fibres, a few fibres of Rb poly d(GC).poly d(GC) were prepared. The polymorphic behaviour of Rb poly d(GC).poly d(GC) was similar to K poly d(GC).poly d(GC). The quality of the fibres were not as good as poly d(GC).poly d(GC). No D pattern was observed. However, the intensity distribution of the low humidity B patterns is slightly different from the B" pattern observed for K poly d(GC). poly d(GC). This may be due to slight structural distortion in B"-DNA.

## 8.4 The ionic content in fibres and its relationship to that in solution

This is the first X-ray fibre diffraction study of poly d(GC). poly d(GC) to be reported in which the polynucleotide was prepared from KF. As reported previously (Rhodes et al., 1982; Mahendrasingam et al., 1983a) estimates of the amount of salt within a polynucleotide fibre have substantial uncertainties. These are particularly acute when, as can be seen from the presence of the sharp diffraction rings in all the patterns in figure 1 - figure 6, part of the salt occurs as a separate phase of

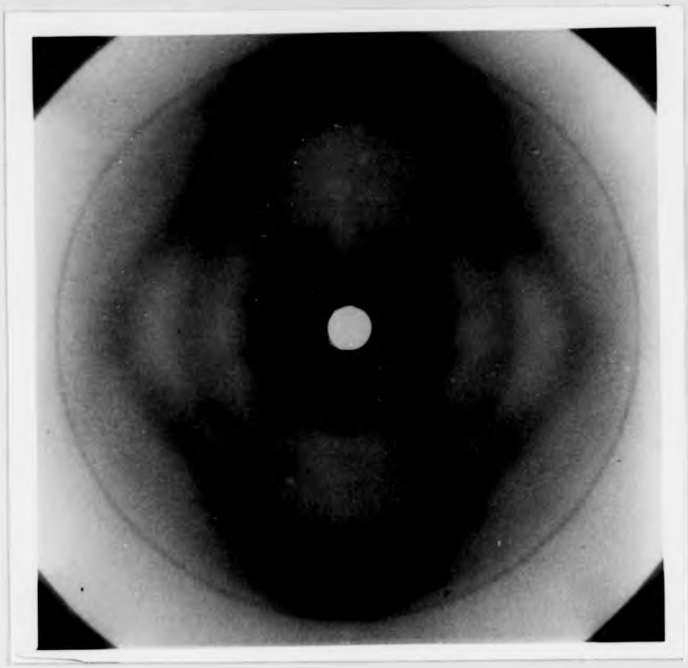


Figure 12 : Semi-crystalline C-form. Diffraction Pattern of Li poly d(GC).poly d(GC) at 75% R.H.

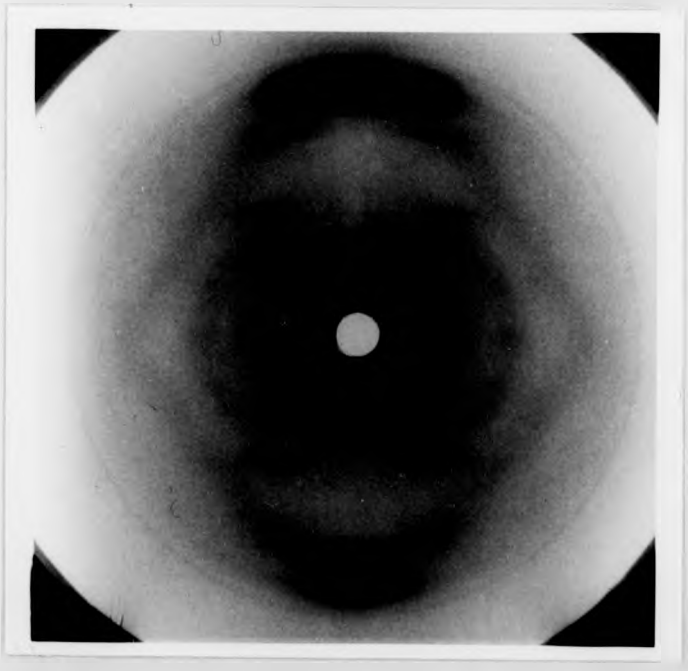


Figure 14 : Semi-crystalline B-form Diffraction Pattern of Li poly d(GC).poly d(GC) at 95% R.H.

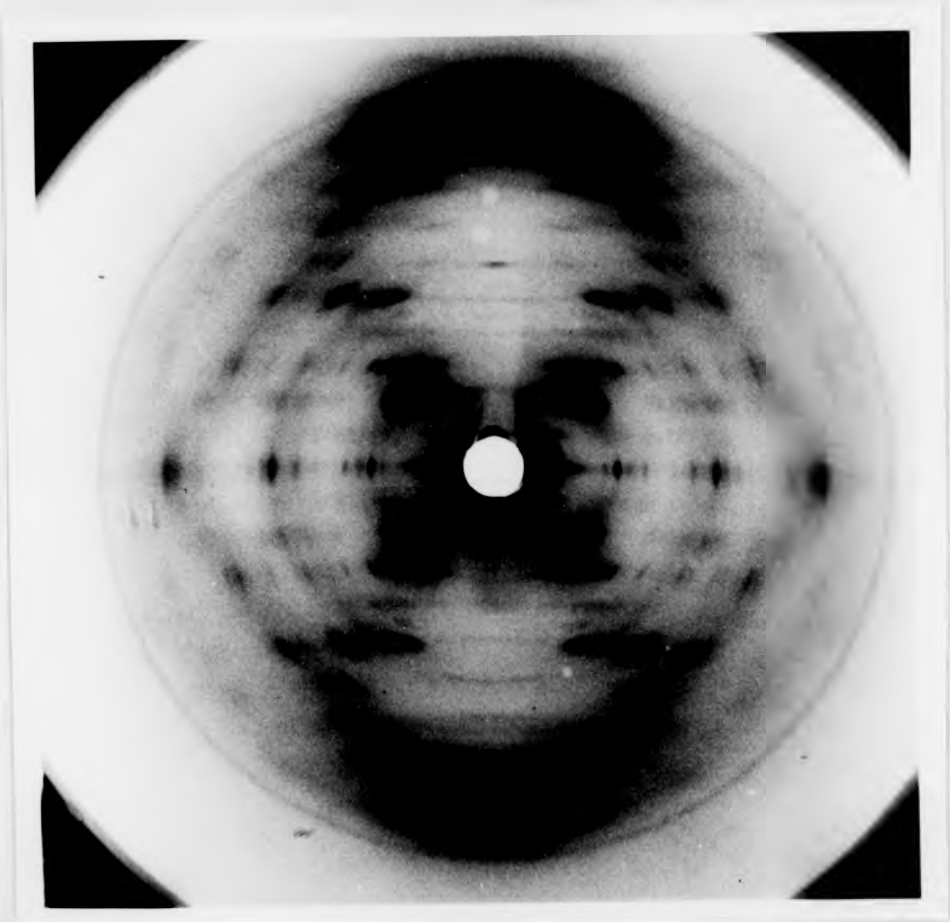


Figure 13 : Crystalline B-form Diffraction pattern of Li poly d(GC).poly d(GC) at 86% R.H.

$KF.2H_2O$ crystallites. The persistence of these crystallites throughout such a wide variation in the relative humidity of the fibre environment (from 33% to 98%) could be taken to indicate that the solvent in the fibre is saturated with KF at all humidities within this range. However, our observation that whether the B form of poly d(GC).poly d(GC) changes to the A form or the S form depends on the rate at which the relative humidity of a wet fibre is reduced, suggests that this simple view of the ionic content of the solvent in the fibre may not be correct. Since the A and S forms can be observed at the same relative humidities and the A form of natural DNAs (Fuller et al., 1965) and of Na poly d(GC).poly d(GC) (Arnott et al., 1980) are generally regarded as low salt forms and the S form as a high salt form (Arnott et al., 1980), it seems somewhat unlikely that in both cases the solvent in the poly d(GC).poly d(GC) unit cell is saturated with KF even though there are crystallites of $KF.2H_2O$ within the fibre. It may be, therefore, that the ionic content of the solvent depends on the rate at which the water content of the fibre is reduced with the more rapid reduction resulting in the lower unit cell ionic strength.

An upper limit for the amount of KF in the polynucleotide unit cell was estimated as less than 0.6 $F^-$ per nucleotide - with a corresponding number of $K^+$ ions in addition to those neutralising DNA $PO_4^-$ groups. In order to relate the observations of fibres of poly d(GC).poly d(GC) to those on the polynucleotides in solution it was necessary to consider what ionic content in fibres might correspond to ionic strength associated with the observation of the Z conformation in solution. Circular dichroism studies typically use polynucleotide concentrations of $10^{-4}$M (i.e. $\sim$ 0.03 mg/ml) whereas for NMR and laser-Raman studies typical concentrations are $10^{-2}$M (i.e. $\sim$ 3 mg/ml) and $10^{-1}$ (i.e. $\sim$ 30 mg/ml) respectively. Despite the fact that these polynucleotide concentrations range over three orders

of magnitude, the salt molarity required to induce the B → Z transition is essentially the same. This can be attributed to the fact that in all these families of experiments the polynucleotide molecules are so far from each other that the difference in polynucleotide concentration has no significant effect on the ionic environment of an individual polynucleotide molecule. Clearly within a fibre the nucleic acid molecules are no longer 'dilute' enough for it to be assumed that they do not interact with each other. However, we can still focus our attention on the contents of the volume of solvent associated on average with each nucleotide. The unit cell per volume per nucleotide in the A, B and S forms is 565 $A^3$, 1038 $A^3$ and 573 $A^3$ respectively and for the purpose of this rather general argument we may take 250 $A^3$ as a typical value for the volume of solvent per nucleotide in the unit cell of a poly d(GC).poly d(GC) fibre when it undergoes the B → Z transition. Therefore, if there is 0.5 $F^-$ per $PO_4^-$ in this volume of solvent, the concentration of the KF surrounding the DNA could be regarded as $\sim$ 4M. Such calculations are no doubt oversimplified. Nevertheless they show that the ionic contents of fibres which are associated with the assumption of "high salt" conformations can be regarded as equivalent to the salt concentrations which induce the corresponding transitions in a wide range of solution studies. In the context of the present work they provide some further support for identifying the Z conformation observed in solution with the S form in fibres.

## 8.5 The Structural Analaysis of $S_I$ and $S_{II}$ Conformations

B-DNA and Z-DNA differ not only in the sense of the helix, but also in the orientation of the base pairs relative to the sugar phosphate backbones (Wang et al., 1979). This is shown schematically in figure 15. Z-DNA has a sequence of alternating deoxyguanosine and deoxycytidine residues. Therefore the asymmetric repeating unit in the

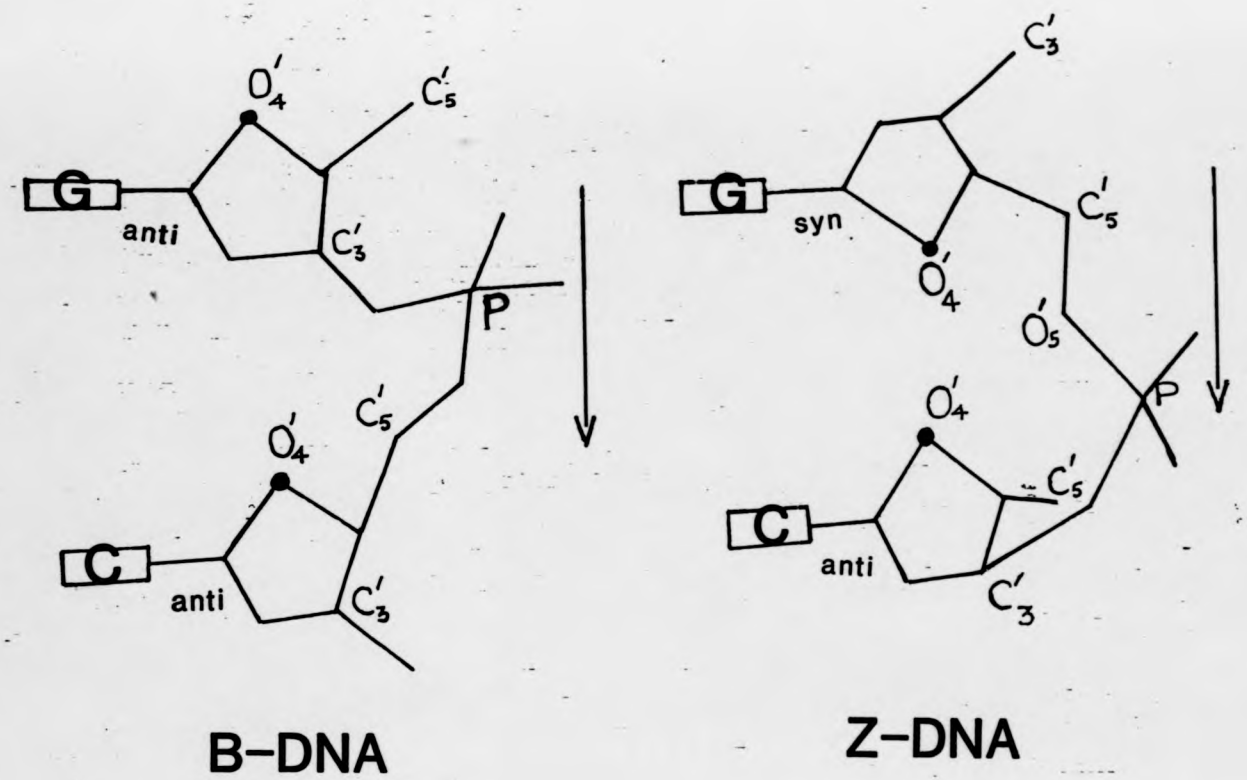



Figure 15 : B-DNA and Z-DNA chains are shown schematically

structure consists of a dinucleotide, in contrast to the single nucleotide which is the repeating unit in right-handed B-DNA. The perpendicular pseudodyad axes is between the base pairs not in the plane of the base pairs. It was reported that in Z-DNA, deoxyguanosine is in a syn conformation with $C_3'$-endo sugar pucker instead of the familiar anti conformation with $C_2'$-endo sugar pucker of B-DNA (Wang et al., 1979).

Wang et al. (1981) reported that the four crystal structures which have been solved have four GpC residues and six CpG residues each, or a total of 16 GpC and 24 CpG linkages. The CpG residues are all in essentially the same conformation and 5 of the 16 GpC linkages are in $g^+t$ conformation (referred to as $Z_{II}$ conformation) and 11 in $g^-t$ conformation (referred to as $Z_I$ conformation). These observations led Wang et al. (1981) to the proposal of the $Z_I$ and $Z_{II}$ structures for poly d(GC).poly d(GC).

In the preliminary analysis it was found that the Fourier transform of $Z_I$ had a good qualitative agreement with the low humidity diffraction pattern obtained for S form ($S_I$). Similar agreement was found to $Z_{II}$ with high relative humidity S pattern $S_{II}$. This prompted the author to undertake a detailed quantitative analysis of the $S_I$ and $S_{II}$ diffraction patterns. The structural details of $S_I$ and $S_{II}$ may provide a plausible explanation for the S $\rightarrow$ B transition with increase in the relative humidity.

As described above the $Z_I$ co-ordinates were used as a starting model for the $S_I$ conformation. A model building program was used to generate different models (as described for D-DNA in Chapter 4) by varying the base displacement form -2.4 A to -3.6 A in 0.2 A steps to find the best fit for the equatorial reflections. The base displacement of $\sim$ -3.2 A was found to give a good fit for the equatorial reflections. Then, by varying the base tilt angle from $-8^o$ to $+8^o$ in steps of $2^o$, a few models were generated. The transforms were calculated and it was found that a positive tilt of the bases gave a weak transform in the 11th layerline. Only a negative tilt

of the bases gave strong calculated diffraction on the 11th layerline. A negative $\sim$ 2 degress base tilt was found to be appropriate for the model. Then the twist of the bases was varied from $-8^{o}$ to $8^{o}$ in steps of 2 degrees. All negative base twists predicted a strong 107 reflection and weak 108 reflection. Only positive base twists gave a strong 108 reflection and a weak 107 reflection. It was found that a positive $\sim$ 6 degrees twist was appropriate for the model. Finally a few models were generated with a small variation of the above parameters. A few structures were then derived for $S_{II}$ using the same method as that described for $S_{I}$. The base tilt, twist and the displacement are $\sim$ 8 degrees, $\sim$ +4 degrees and -3 A respectively. The structure factors were calculated for a few models of $S_{I}$ and $S_{II}$ conformations as a function of molecular orientation in the unit cell. From these calculations the best agreement between observed and calculated structure factors is observed when a dyad axis relating to the two chains of the structure is oriented at an angle of 3 degrees to the a axis for $S_{I}$ and 30 degrees to the a axis for $S_{II}$.

The base tilt, twist and displacement (from the helix axis) are -2 degrees, +6 degrees and -3.2 A respectively for $S_{I}$. For $S_{II}$, the base tilt, twist and the displacement are +8 degrees, +4 degrees and -3.0 A respectively. The observed and calculated structure factors for the structures $S_{I}$ and $S_{II}$ are given in tables 3 and 4 respectively. The cylindrical polar co-ordinates for the structures $S_{I}$ and $S_{II}$ are given in tables 5 and 6 respecitvely. The torsion angles of the models are given in table 7. The cylindrically averaged squared transform for $Z_{I}$ and $S_{I}$ are shown in figure 16. In figure 17 the cylindrically averaged squared transform for $Z_{II}$ and $S_{II}$ are shown. The final R values (residuals) for $S_{I}$ and $S_{II}$ are 35% and 38% respectively.

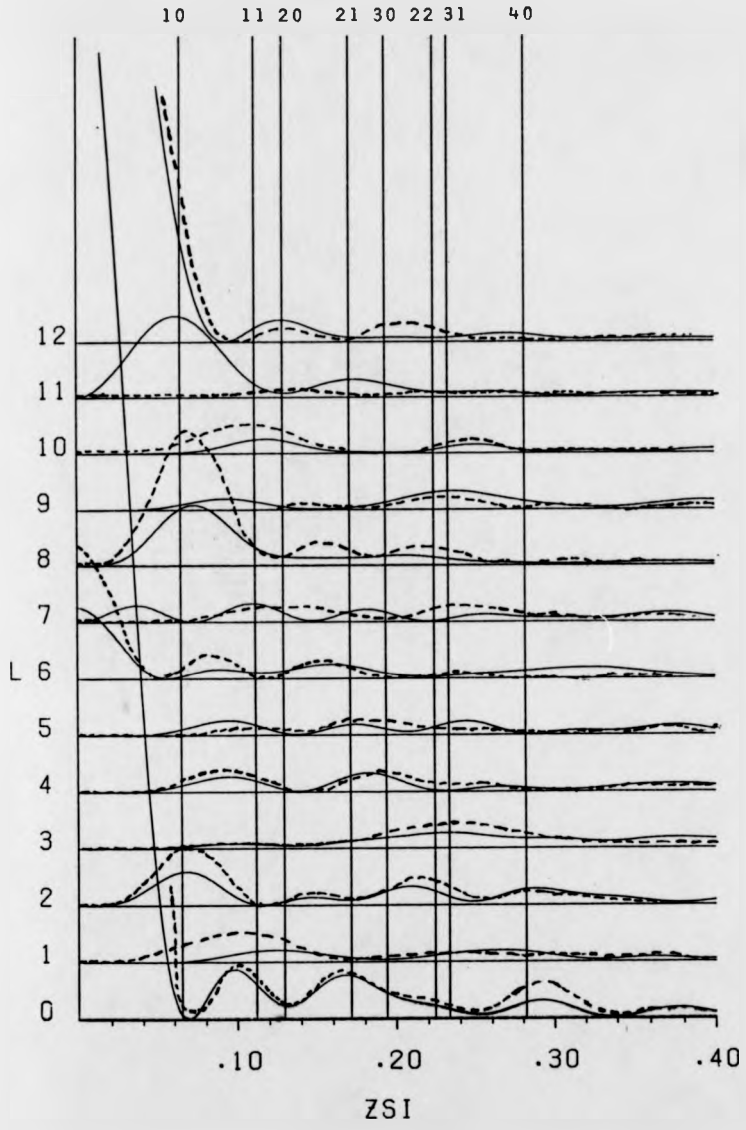



Figure 16 : The cylindrically averaged squared molecular transform for $Z_I$ and $S_I$. $S_I$ : full-line, $Z_I$ : dotted-line

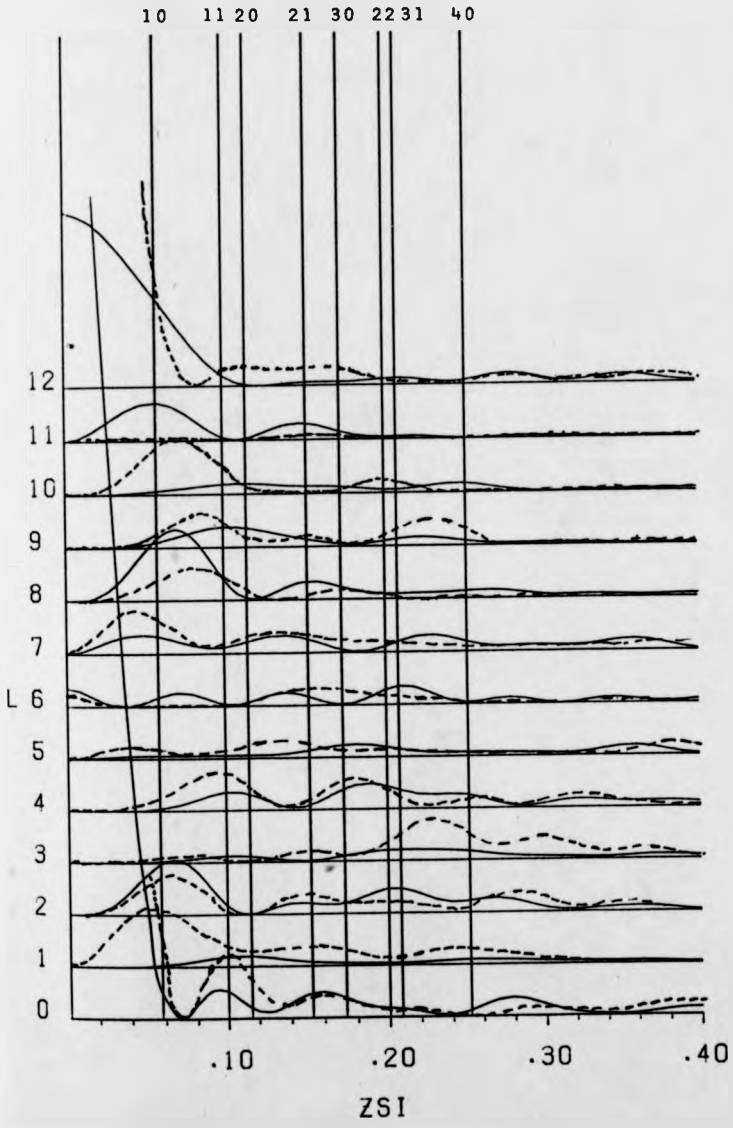



Figure 17 : The cylindrically averaged squared molecular transform for $Z_{II}$ and $S_{II}$

$S_{II}$ : full-line, $Z_{II}$ : dotted-line

TABLE 3

Indices of the Observed Bragg Reflections with the Observed ($F_{obs}$) and calculated ($F_{cal}$) values of the structure amplitudes for $S_I$-DNA model

| h | k | ℓ | $F_{obs}$ | $F_{cal}$ |
|---|---|---|---|---|
| 1 | 0 | 0 | 265 | 512 |
| 1 | 1 | 0 | 1717 | 1663 |
| 2 | 0 | 0 | 1170 | 581 |
| 2 | 1 | 0 | 685 | 628 |
| 3 | 0 | 0 | 562 | 1580 |
| 3 | 2 | 0 | 1384 | 1692 |
| 1 | 0 | 1 | 80 | 161 |
| 1 | 1 | 1 | 830 | 559 |
| 2 | 0 | 1 | 296 | 795 |
| 1 | 0 | 2 | 367 | 1270 |
| 2 | 1 | 2 | 1390 | 810 |
| 1 | 0 | 3 | 107 | 91 |
| 2 | 1 | 3 | 1207 | 747 |
| 3 | 1 | 3 | 1600 | 1434 |
| 1 | 0 | 4 | 308 | 387 |
| 2 | 1 | 4 | 1681 | 1373 |
| 1 | 0 | 5 | 323 | 509 |
| 1 | 1 | 5 | 759 | 636 |
| 2 | 0 | 5 | 393 | 271 |
| 2 | 1 | 5 | 619 | 1058 |
| 3 | 0 | 5 | 767 | 116 |
| 1 | 0 | 6 | 205 | 237 |
| 1 | 0 | 7 | 125 | 269 |
| 1 | 0 | 8 | 1818 | 1794 |
| 2 | 0 | 8 | 645 | 567 |
| 2 | 1 | 8 | 756 | 948 |
| 1 | 0 | 10 | 372 | 259 |
| 1 | 0 | 11 | 1957 | 1863 |

TABLE 4

Indices of the Observed Bragg Reflections with the Observed ($F_{obs}$) and calculated ($F_{cal}$) values of the structure amplitudes for $S_{II}$-DNA model

| h | k | ℓ | $F_{obs}$ | $F_{cal}$ |
|---|---|---|---|---|
| 1 | 0 | 0 | 1099 | 451 |
| 1 | 1 | 0 | 424 | 296 |
| 2 | 0 | 0 | 466 | 404 |
| 2 | 1 | 0 | 349 | 628 |
| 1 | 1 | 1 | 225 | 209 |
| 1 | 0 | 2 | 366 | 550 |
| 1 | 1 | 2 | 225 | 314 |
| 2 | 1 | 2 | 560 | 402 |
| 1 | 0 | 4 | 184 | 178 |
| 2 | 0 | 4 | 238 | 428 |
| 1 | 0 | 7 | 305 | 432 |
| 1 | 0 | 8 | 645 | 601 |
| 1 | 0 | 10 | 205 | 153 |
| 1 | 0 | 11 | 224 | 347 |

TABLE 5

Cylindrical polar coordinates for the left-handed $S_I$-DNA model. The dyad axis relating to the two chains is at (r,0,0)

| Atom | r(A ) | $\phi$(o) | Z(A ) |
|---|---|---|---|
| $G_1O_3'$ | 7.74 | 39.7 | 6.52 |
| $G_1C_5'$ | 6.23 | 43.5 | 3.55 |
| $G_1C_4'$ | 7.30 | 48.5 | 4.44 |
| $G_1O_4'$ | 7.70 | 58.7 | 4.12 |
| $G_1C_2'$ | 7.64 | 57.9 | 6.55 |
| $G_1C_3'$ | 7.00 | 48.2 | 5.97 |
| $G_1C_1'$ | 7.90 | 64.6 | 5.32 |
| $G_1N_7$ | 7.42 | 91.4 | 5.04 |
| $G_1C_5$ | 6.04 | 90.3 | 5.17 |
| $G_1C_4$ | 5.80 | 77.4 | 5.32 |
| $G_1N_3$ | 4.80 | 68.4 | 5.42 |
| $G_1C_2$ | 3.60 | 76.3 | 5.45 |
| $G_1N_2$ | 2.60 | 61.5 | 5.62 |
| $G_1N_1$ | 3.80 | 97.2 | 5.34 |
| $G_1C_6$ | 5.11 | 101.4 | 5.22 |
| $G_1O_6$ | 5.65 | 113.5 | 5.12 |
| $G_1N_9$ | 7.20 | 74.1 | 5.24 |
| $G_1C_8$ | 7.93 | 82.4 | 5.12 |
| $G_1O_5'$ | 6.90 | 38.2 | 2.42 |
| $G_1PO_1'$ | 6.23 | 379.7 | 1.22 |
| $G_1PO_2'$ | 6.60 | 38.7 | -0.07 |
| $G_1P'$ | 6.06 | 33.4 | 1.18 |
| $C_2O_3'$ | 4.62 | 40.3 | 1.47 |
| $C_2C_5'$ | 6.46 | 74.2 | 1.26 |
| $C_2C_4'$ | 5.25 | 67.6 | 1.87 |
| $C_2O_4'$ | 4.42 | 80.2 | 2.42 |
| $C_2C_2'$ | 3.20 | 66.4 | 0.59 |
| $C_2C_3'$ | 4.50 | 56.6 | 0.86 |
| $C_2C_1'$ | 3.07 | 80.2 | 1.94 |
| $C_2C_5$ | 4.90 | 129.7 | 1.60 |

Cont.

Table 5 (Cont.)

| Atom | r(A ) | $\phi$(o) | Z(A ) |
|---|---|---|---|
| $C_2C_4$ | 4.37 | 146.5 | 1.66 |
| $C_2N_4$ | 5.40 | 156.2 | 1.58 |
| $C_2N_3$ | 3.11 | 153.6 | 1.82 |
| $C_2O_2$ | 0.92 | 138.8 | 2.04 |
| $C_2C_2$ | 2.13 | 132.8 | 1.92 |
| $C_2N_1$ | 3.03 | 108.4 | 1.84 |
| $C_2C_6$ | 4.33 | 114.7 | 1.69 |
| $C_2O_5'$ | 6.30 | 85.8 | 0.52 |
| $C_2PO_1'$ | 6.90 | 84.5 | -1.94 |
| $C_2PO_2'$ | 8.70 | 83.5 | -0.18 |
| $C_2P'$ | 7.40 | 87.80 | -0.63 |

TABLE 6

Cylindrical polar coordinates for the left-handed $S_{II}$-DNA model. The dyad axis relating to the two chains is at (r,0,0)

| Atom | r(A ) | $\phi$(o) | Z(A ) |
|---|---|---|---|
| $G_1O_3'$ | 8.25 | 30.8 | 7.57 |
| $G_1C_5'$ | 8.04 | 37.3 | 4.41 |
| $G_1C_4'$ | 8.30 | 41.8 | 5.71 |
| $G_1O_4'$ | 8.10 | 51.8 | 5.65 |
| $G_1C_2'$ | 7.20 | 47.0 | 7.81 |
| $G_1C_2'$ | 7.40 | 37.7 | 6.91 |
| $G_1C_1'$ | 7.74 | 55.9 | 6.92 |
| $G_1N_7$ | 7.20 | 83.1 | 6.47 |
| $G_1C_5$ | 5.80 | 81.3 | 6.41 |
| $G_1C_4$ | 5.63 | 67.8 | 6.51 |
| $G_1N_3$ | 4.66 | 57.8 | 6.51 |
| $G_1C_2$ | 3.44 | 64.3 | 6.31 |
| $G_1N_2$ | 2.54 | 46.5 | 6.25 |
| $G_1N_1$ | 3.50 | 86.6 | 6.14 |
| $G_1C_6$ | 4.80 | 92.4 | 6.21 |
| $G_1O_6$ | 5.32 | 105.3 | 6.01 |
| $G_1N_9$ | 7.00 | 65.2 | 6.71 |
| $G_1C_8$ | 7.71 | 74.0 | 6.67 |
| $G_1O_5'$ | 6.71 | 41.5 | 3.81 |
| $G_1PO_1'$ | 6.35 | 22.4 | 2.56 |
| $G_1PO_2'$ | 6.96 | 40.4 | 1.31 |
| $G_1P'$ | 6.20 | 35.8 | 2.44 |
| $C_2O_3'$ | 4.69 | 42.2 | 2.41 |
| $C_2C_5'$ | 6.36 | 77.8 | 2.78 |
| $C_2C_4'$ | 5.11 | 69.7 | 3.05 |
| $C_2O_4'$ | 4.04 | 81.3 | 3.36 |
| $C_2C_2'$ | 3.56 | 67.6 | 1.21 |
| $C_2C_3'$ | 4.76 | 58.5 | 1.86 |
| $C_2C_1'$ | 2.92 | 78.8 | 2.51 |
| $C_2C_5$ | 4.53 | 133.6 | 2.08 |

(Cont.)

Table 6 (Cont.)

| Atom | r(A ) | $\phi$(o) | Z(A ) |
|---|---|---|---|
| $C_2C_4$ | 3.97 | 151.6 | 1.89 |
| $C_2N_4$ | 5.02 | 162.0 | 1.71 |
| $C_2N_3$ | 2.70 | 159.7 | 1.91 |
| $C_2O_2$ | 0.52 | 137.5 | 2.11 |
| $C_2C_2$ | 1.75 | 135.0 | 2.11 |
| $C_2N_1$ | 2.73 | 108.9 | 2.27 |
| $C_2C_6$ | 4.02 | 117.3 | 2.26 |
| $C_2O_5'$ | 6.50 | 82.5 | 1.41 |
| $C_2PO_1'$ | 7.65 | 73.3 | -0.57 |
| $C_2PO_2'$ | 8.90 | 77.9 | 1.51 |
| $C_2P'$ | 7.83 | 80.5 | 0.56 |

*Table 7* : *The backbone torsion angles of left-handed DNA models*

| Torsion Angle | $S_I$ | | $Z_I$ | | $S_{II}$ | | $Z_{II}$ | |
|---|---|---|---|---|---|---|---|---|
| | C | G | C | G | C | G | C | G |
| $O_3'$-P'$O_5'$-$C_5'$ | -151 (t) | 13 ($g^+$) | -137 (t) | 47 ($g^+$) | -132 (t) | 169 (t) | 146 (t) | 92 ($g^+$) |
| P'-$O_5'$-$C_5'$-$C_4'$ | -158 (t) | -166 (t) | -139 (t) | 179 (t) | -136 (t) | -174 (t) | 164 (t) | -167 (t) |
| $O_5'$-$C_5'$-$C_4'$-$C_3'$ | 77 ($g^+$) | -144 (t) | 56 ($g^+$) | -169 (t) | 27 ($g^+$) | 76 ($g^+$) | 66 ($g^+$) | 157 (t) |
| $C_5'$-$C_4'$-$C_3'$-$O_3'$ | 139 (t) | 100 ($g^+$) | 138 (t) | 99 ($g^+$) | 146 (t) | 93 ($g^+$) | 147 (t) | 94 ($g^+$) |
| $C_4'$-$C_3'$-$O_3'$-P | -100 ($g^-$) | -85 ($g^-$) | -94 ($g^-$) | -104 ($g^-$) | -100 ($g^-$) | -144 (t) | -100 ($g^-$) | -179 (t) |
| $C_3'$-$O_3'$-P'-$O_5'$ | 98 ($g^+$) | -75 ($g^-$) | 80 ($g^+$) | -69 ($g^-$) | 85 ($g^+$) | -50($g^-$) | 74 ($g^+$) | 55 ($g^+$) |
| $C_6$-$N_1$-$C_1$-$O_4'$ | 24 ($g^+$) | - | 21 ($g^+$) | - | 33 ($g^+$) | - | 32 ($g^+$) | - |
| $C_4$-$N_9$-$C_1'$-$O_4'$ | - | 67 ($g^+$) | - | -112 ($g^-$) | - | -121 (t) | - | -118 ($g^-$) |

# CHAPTER 9

## Polymorphism of Naturally Occurring DNAs

### 9.1 Introduction

Deoxyribonucleic acid (DNA) and its synthetic analogues are remarkable in respect of the number of distinct, regular conformations that they can assume, both from the point of view of understanding the biological function of DNA and for an elucidation of the factors important in determining its three-dimensional structure. It is important to establish the conditions under which these various conformations are favoured. Three well-defined conformations have been observed for naturally occurring DNAs in oriented fibres. These have been designated A, B and C and their detailed geometries in a variety of crystalline and semi-crystalline forms have been determined by X-ray fibre diffraction analysis (Fuller et al., 1965, Langridge et al., 1960b, Marvin et al., 1961, Arnott et al., 1972b, 1973).

In the case of naturally occurring DNAs the C conformation is commonly observed as a low humidity form of the lithium salt (Marvin et al., 1961). However, when the counter-ion is sodium rather than lithium previous X-ray fibre diffraction studies have suggested that the C form is at best a poorly favoured conformation. Bram et al. (1974) claimed that fibres of the sodium salt of calf thymus DNA pulled at 37$^{o}$C gave a C pattern at relative humidities of 66% and less. However, the published pattern was not well-defined and a layerline spacing of 31 A was the only parameter quoted as supporting its designation as C. Arnott and Selsing (1975) published a C pattern from a fibre of NaDNA which although better defined than that reported by Bram et al. (1974) was still much inferior,

both as regards orientation and crystallinity, to patterns observed for LiDNA (Marvin et al., 1961). Arnott and Selsing (1975) reported that they observed the C form of NaDNA at low relative humidities from fibres with salt contents intermediate between those appropriate for the A and B conformations of NaDNA. In a more recent discussion of DNA polymorphism Leslie et al. (1980) state that the C form is stable at relative humidities and salt contents which are both intermediate between those which favour the A and B conformations. Zimmerman and Pheiffer (1980) have recently described an extensive X-ray diffraction study of DNA fibres in a variety of concentrated salt solutions and organic solvent/water mixtures. For NaDNA they observed C-like patterns from fibres immersed in mixtures containing 95% or more (v/v) of t-butanol. These authors emphasised that the C-form should be visualised as a family of related structures with a characteristic diffraction pattern. They suggested that continuous smooth transitions may occur between the various members of the C family and also possibly between the B and C forms. Infrared linear dichroism studies have shown that a C-like conformation can occur in oriented films of DNA (Brahms et al., 1973).

Transitions between the A and B forms of NaDNA have been shown to depend on the salt content of the fibre and the relative humidity of its environment, (Cooper and Hamilton, 1966). The remainder of this chapter is devoted to describing the conditions required for routinely observing the C-form of NaDNA and the transitions between the A, B and C forms of NaDNA. These are shown to be applicable to a wide variety of natural DNAs. A systematic study on LiDNA is also described.

## 9.2 Materials and Methods

The following natural DNAs were used in this study: calf thymus clostridium perfringens, herring sperm, pollock roe, salmon sperm, E.Coli

(all from sigma) SP15 (from Professor Mandel), $T_2$phage, micrococus lysodeikticus (both from Miles) $\phi$w-14 (Kropinski et al., 1973), calf thymus (from Professor Warren). The final stage in purification was either by ethanol precipitation from solutions containing 0.1 M or less NaCℓ, or by centrifugation at 50,000 r.p.m. from solutions for which the NaCℓ concentration was as low as 0.001 M. Fibres were drawn from either the precipitated material or from concentrated gels. X-ray diffraction patterns were obtained at a range of relative humidity of the fibre environment from 33% to 98%. For specimens prepared from centrifuged material, ion concentration estimates were made using measurements of $OD_{260}$ for the original solution and supernatant. Hence it was possible to calculate the number of $Na^+$ and $Cℓ^-$ added per DNA phosphate (for details see Chapter 2).

9.3 Results

(a) All the DNAs studied could be induced to assume the C form and examples of their diffraction patterns are shown in figure 1 to figure 4. Where patterns were sufficiently well defined (see figure 1) for analysis, they were found to be hexagonal rather than orthorhombic (Marvin, et al., 1961). The pattern on figure 1 was measured and indexed as a hexagonal lattice with lattice parameters, a = 30.8 A and c = 31.6 A.

(b) Conditions were found for which all the DNAs studied assumed the B conformation and also, with the exception of DNA from $T_2$ and SP15, the A conformation.

(c) C patterns from the various NaDNAs were only observed for the fibres which contained very little excess NaCℓ, and which were placed in an environemnt of low relative humidity.

(d) For all the DNAs which could assume the A form, fibres which gave the C pattern at low humidity exhibited structural transitions first from

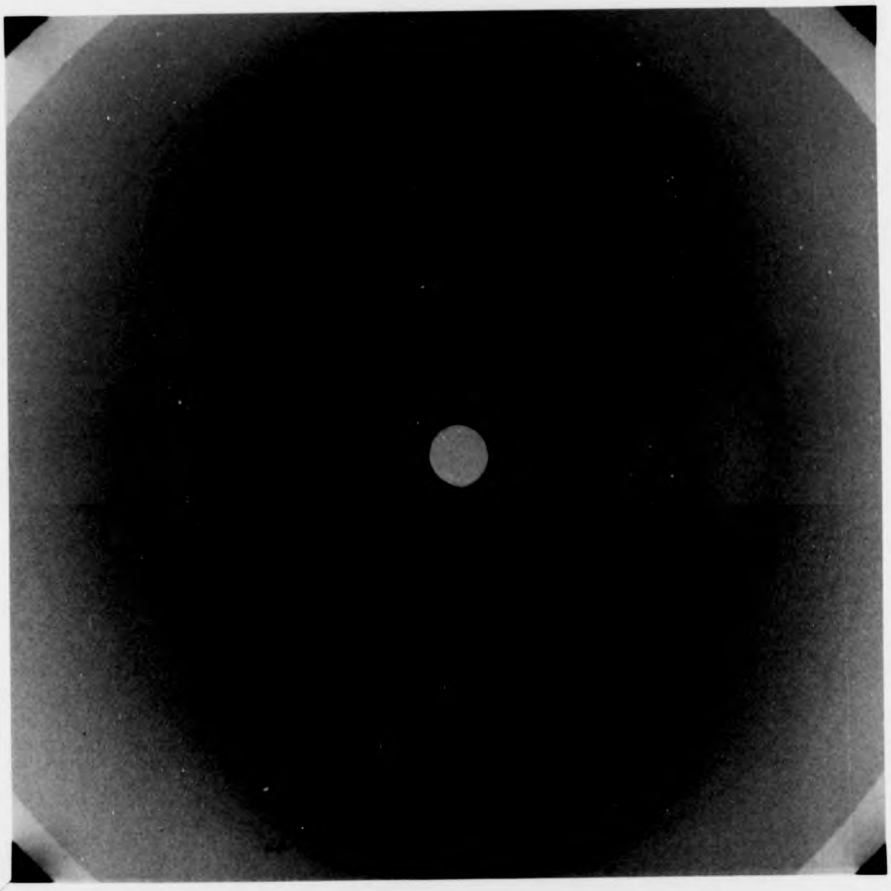


Figure 1 : Semi-crystalline C-form. Diffraction Pattern of Na DNA (Calf thymus) at 57% R.H.

Figure 2 : Semi-crystalline C-form. Diffraction Pattern of Na $T_2$ DNA at 33% R.H.

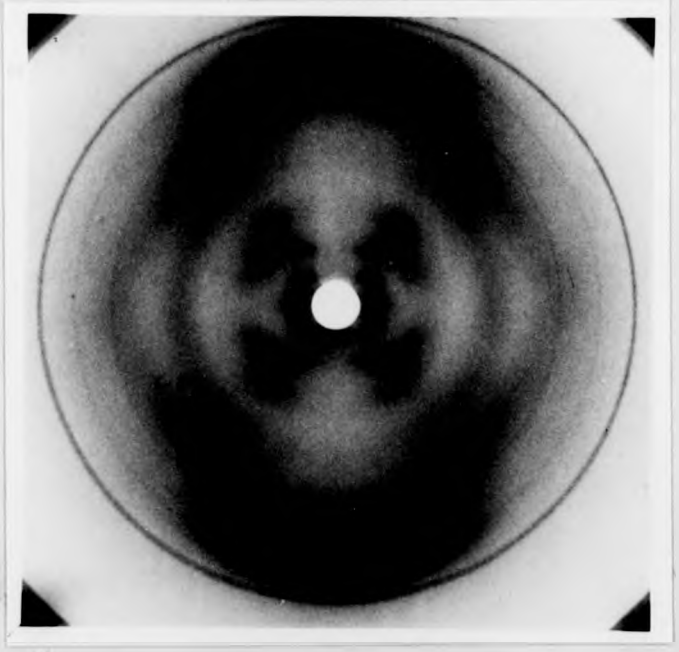


Figure 3 : Semi-crystalline C-form. Diffraction Pattern of Na DNA (E.Coli) at 66% R.H.

the C form to the fully crystalline A form and then from the A form to the semi-crystalline B form as the relative humidity was increased. In general, the greater the salt content of the fibre the lower the relative humidity at which transition to the A pattern occurred. For fibres containing the lowest amounts of excess salt, the C form was usually observed at relative humidities in the range of 32% to 75% but as the salt content of the fibre was increased the upper limit in the relative humidity at which the C form occurred was reduced until it reached a level when the C pattern was no longer observed, (i.e. typically 0.5 $C\ell^-$ per $PO_4^-$). Below 32% relative humidity the X-ray diffraction patterns were badly defined, indicating a progressive collapse of the structure. As reported previously (Cooper and Hamilton, 1966), the relative humidity at which the A to B transition occurred decreased with increasing salt content of the fibre.

(e) The transitions C → A → B were fully reversible as the relative humidity was reduced.

(f) For the fibres which underwent the C → A → B transitions, patterns were observed which indicated mixtures of conformations. C/A and A/B mixtures (figure 5) were noted. (For detailed discussion see Chapter 6).

(g) For fibres of DNA from $T_2$ and SP15 there was a transition from the C form to a semi-crystalline B form with increasing relative humidity (figure 6 and figure 7).

(h) For LiDNA, a semi-crystalline C form and B form (both crystalline and semi-crystalline) were observed as previously reported. (Marvin et al., 1961). No A form was observed.

## 9.4 Discussion

This study establishes the C form as a major conformational possibility for the DNA double-helix in biological environments. Previous

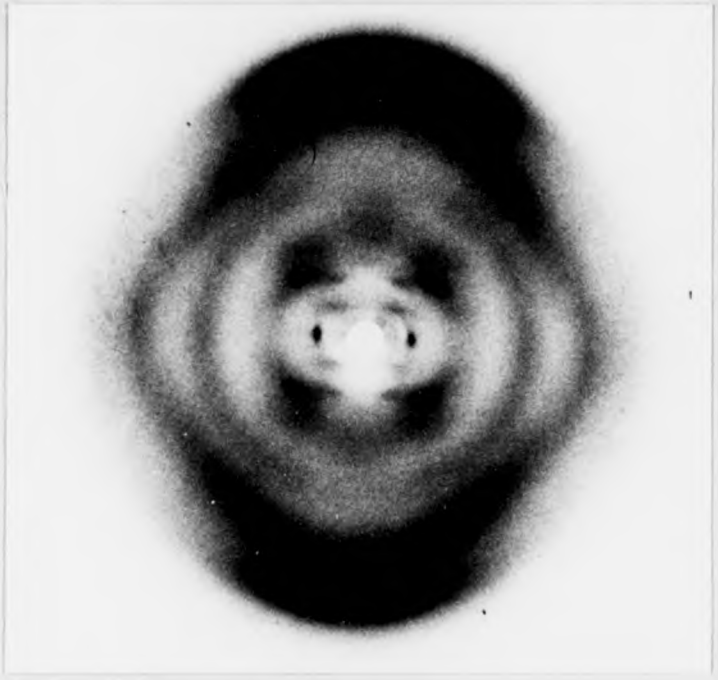


Figure 4 : Semi-crystalline C-form, Diffraction Pattern of Na $\phi$w-14 DNA at 57% R.H.

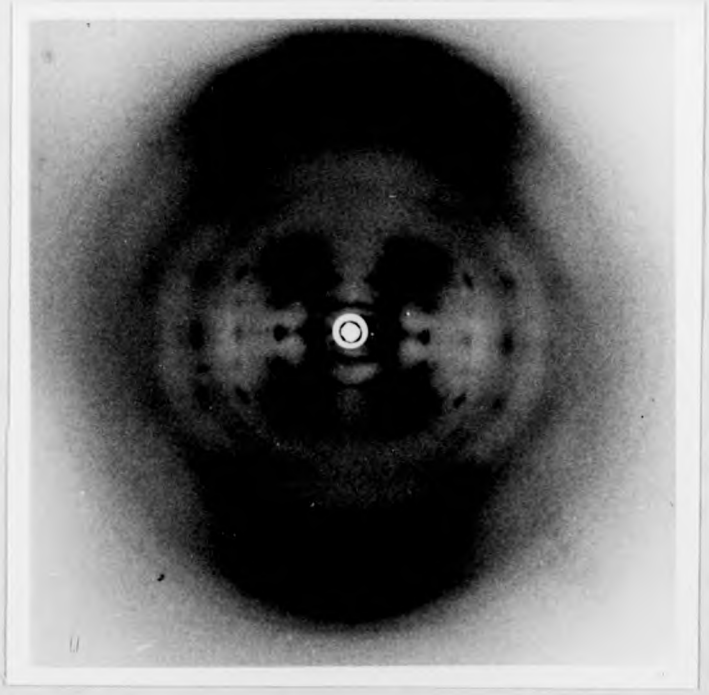


Figure 5 : Mixture of A and B forms, Diffraction Pattern of Na DNA (Calf thymus) at 86% R.H.

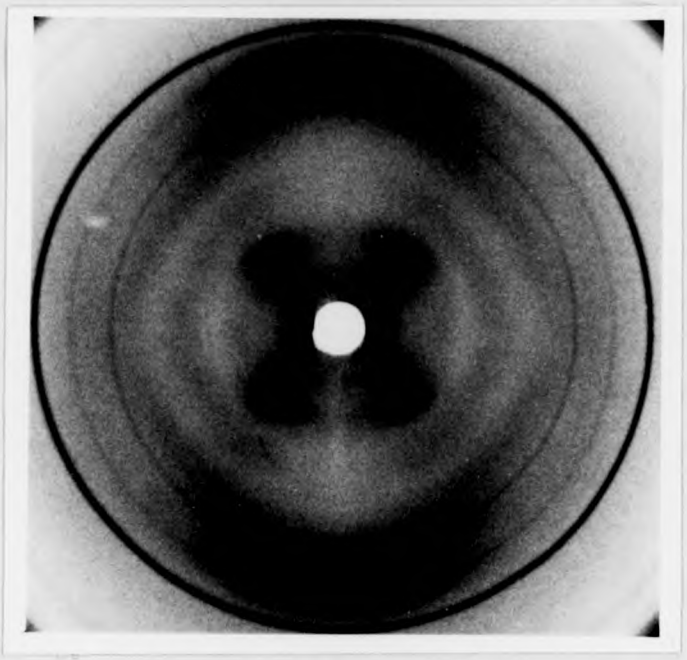


Figure 6 : Semi-crystalline B-form, Diffraction Pattern of Na SP-15 DNA at 98% R.H.

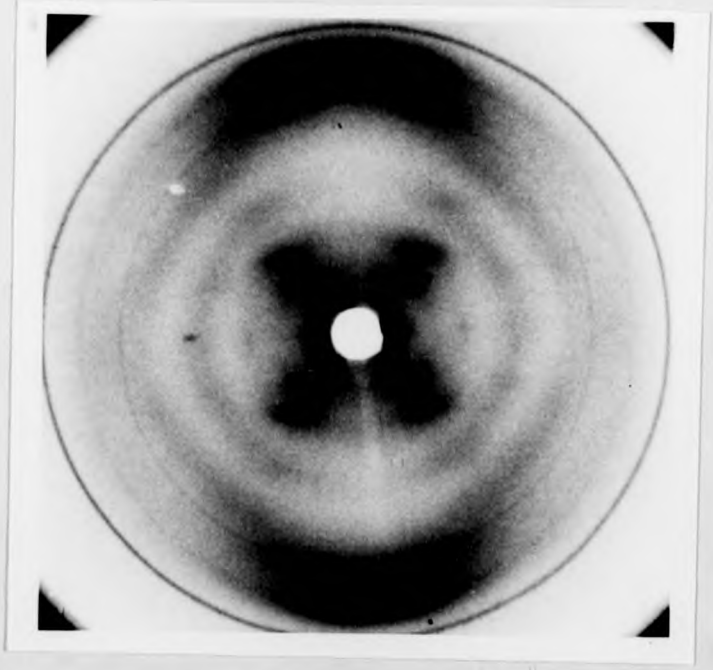


Figure 7 : Semi-crystalline B-form, Diffraction Pattern of Na $T_2$ DNA at 98% R.H.

work indicated that while the C form was readily assumed by the lithium salt of DNA, this form was at best a very poorly favoured conformation when the counterion was sodium. From the present work it is clear that there is no difficulty in inducing NaDNA to assume the C conformation if both the water content and the level of excess salt in the fibre are sufficiently low.

In the biological situation DNA can be expected to occur predominantly either as the sodium salt or in a state where the sodium ions are displaced by basic groups such as lysine and arginine side chains in proteins. It is therefore possible to envisage situations in which the conformation of a functionally active region of DNA is controlled by processes which change the degree of hydration of the DNA and the concentration of $Na^+$ and $C\ell^-$ ions in the vicinity.

Earlier reports have suggested that the C form of DNA might be seen as a variant of the B conformation with ready transition between two forms. While such a transition has been observed for the lithium salt of DNA, in this study the sodium slat, the transition from C to B is observed as passing through an intermediate A conformation except for those DNAs that do not assume the A form. In contrast to the earlier claim that at low hydration the salt content of fibres of NaDNA which gave the C form was intermediate between those appropriate for observing the A and B forms, it was found that the C form is favoured by salt contents which are lower than those which favour the A form. It should be emphasised that in view of the present study the presumption that the A form is an intermediate between the C and B forms of NaDNA is based on two distinct sets of observations. First, the C → A → B transitions occur as a function of the relative humidity for fibres containing small amounts of salt. Second, the conformational transitions C → A → B occur at a fixed relative humidity (say 57%) as the salt content of a fibre is raised.

One of the most striking features of this study is the wide range of natural DNAs over which the C form was observed. In particular it was observed for DNAs for which the G+C content ranged from 31% to 72%. Perhaps of even greater significance is that the C form was observed for three DNAs ($\phi$w-14, SP15, $T_2$) in which a substantial fraction of the nucleotides were extensively modified. In one case the chemical modification involves approximately 1 in 8 nucleotides carrying a positive charge. The fact that the C form is observed for this very wide range of 'natural' DNAs suggests that it is a structure whose occurrence is sufficiently favoured stereochemically for it not to be destabilised by substantial changes in either its charge distribution of the addition of bulky groups. The A form was not observed for $T_2$ DNA. This is in accord with earlier studies which also suggested that this was a consequence of glucosylation of a large fraction of the cytosine residues (Mokulskii et al., 1972). However, the A form was observed for $\phi$w-14 DNA and is identical to A forms from native DNA (see figure 8 and figure 9). Mokulskii et al., (1972) also suggested that at low humidity $T_2$ DNA adopts a novel conformation, the T form. In view of the observation of the C form as a low humidity conformation of $T_2$ DNA and the fact that the helix pitch and the overall intensity distribution reported for the T form are broadly similar to those which characterise C patterns, it is possible that the T form is at least a close relative if not a member of the C family of structures.

The semi-crystalline B patterns of all natural DNA stuied were very similar (figures 10 and 11). However, the intensities of the 1st and 3rd layerlines are enhanced in G-C rich DNAs. The intensity enhancement may be due to minor structural variation in B-DNA as discussed in Chapter 5.

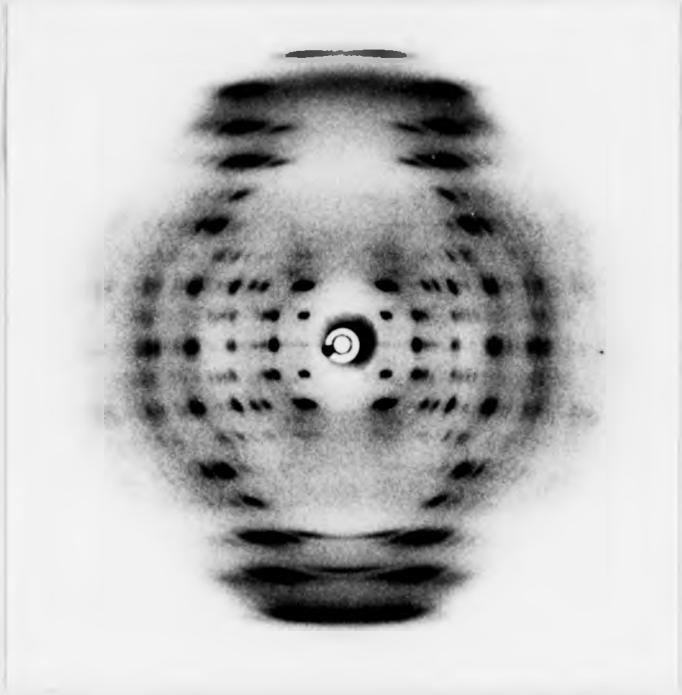


Figure 8 : A-form Diffraction Pattern of Na $\phi$w-14 DNA at 92% R.H,

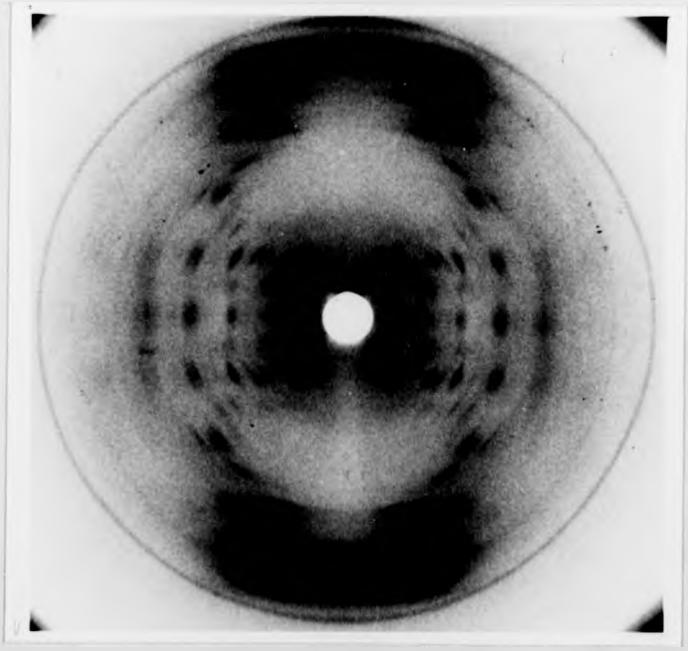


Figure 9 : A-form Diffraction Pattern of Na DNA (Calf thymus) at 75% R.H.

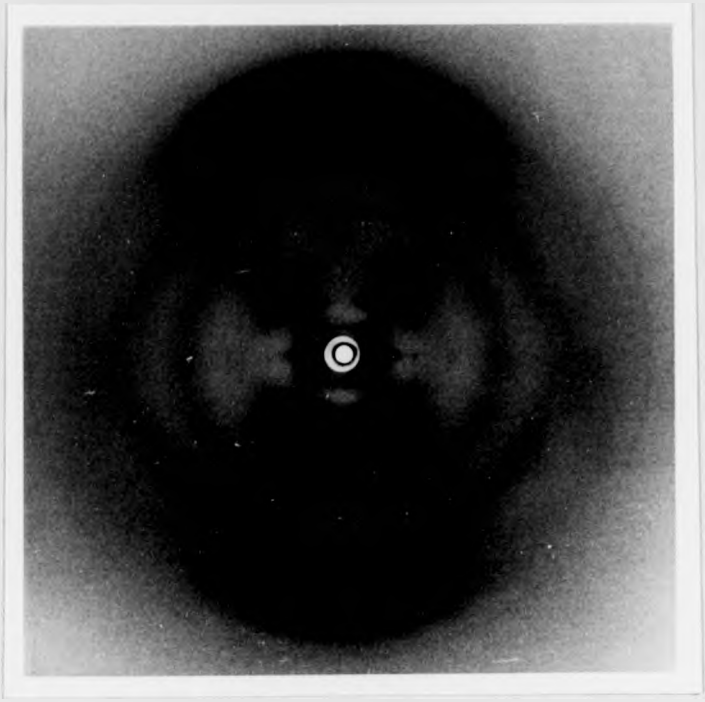


Figure 10 : Semi-crystalline B-form Diffraction Pattern of Na DNA (Calf thymus) at 98% R.H.

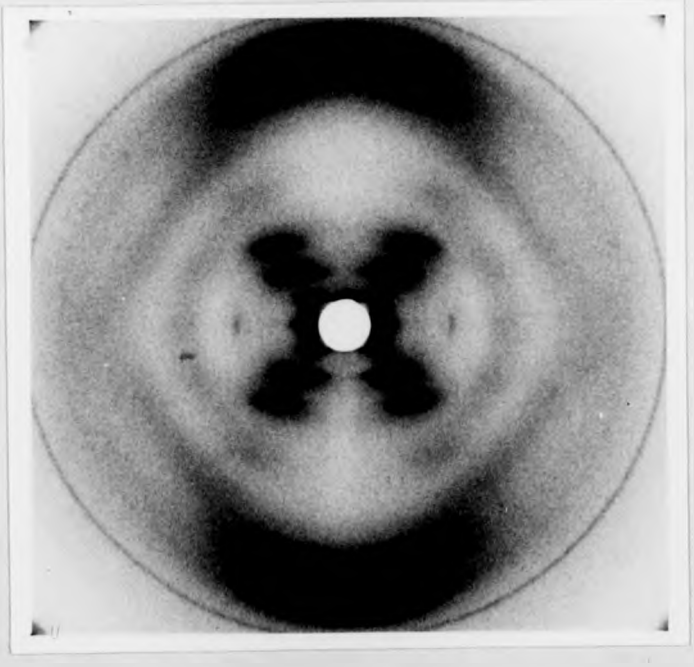


Figure 11 : Semi-crystalline B-form Diffraction Pattern of Na DNA (Cl.perfringens) at 98% R.H.

# CHAPTER 10

## Does the Heteronomous Structure Exist in Homopolymers?

### 10.1 Introduction

A homopolymer has identical purines in one chain and identical pyrimidines in the other (i.e. poly d(purine).poly d(pyrimidine)). Early fibre diffraction studies led Langridge, Bollum and Chamberlain (Langridge, 1969) to conclude that poly d(A).poly d(T) and poly d(G).poly d(C) have molecular structures rather different from those of the native DNA forms, and from one another. Arnott and Selsing (1974a) reported that one conformation of poly d(A).poly d(T) was in fact very similar to B-DNA (termed B'-DNA). They also reported that the existence of a triple stranded structure in fibres of poly d(T).poly d(A).poly d(T). B'-DNA was reported to have two crystal forms depending upon the prevailing relative humidity. At low humidities (up to 84%) the B' structure was found to pack into an orthorhombic lattice with lattice parameters a = 17.8 A, b = 20.0 A, c = 32.4 A, and at relative humidities of 92% and above it was found to be in a hexagonal form with lattice parameters a = 22.8 A, c = 32.9 A. The orthorhombic crystal form is called β-B'-DNA. Arnott et al. (1974a) proposed a $10_1$ helix with identical anti-parallel chains based on α-B'-DNA diffraction data. Further they claimed that the α and β forms have the same molecular conformation (which is a slightly modified B-DNA structure, Arnott et al., 1972b) and are distinct only in their packing arrangement.

More recently, however, Arnott and his co-workers proposed a heteronomous DNA structure (Arnott et al., 1983) for poly d(A).poly d(T) based on β-B'-DNA diffraction data (which has more Bragg reflections than

the α-B'-DNA). In the new structure both chains are $10_1$ helices, mutually hydrogen-bonded in the standard (Watson-Crick) fashion, but have quite different conformations. It is proposed that the poly d(A) chain has $C_3$'-endo sugar puckering, a characteristic of the A-family, while poly d(T) has $C_2$'-endo puckered rings, a characteristic of the B-family. Arnott et al. (1983) further suggests that the heteronomous duplexes may apply also to the α-form. Similar patterns to these have been obtained from fibres of poly d(I).poly d(C), and fibres of poly d(AI).poly d(CT) (Leslie et al., 1980). It has also been reported that the double-stranded form of poly d(I).poly d(C) was unstable in fibres, and transformed irreversibly to the triple-stranded structure poly d(C).poly d(I).poly d($C^+$) after a few days. Arnott et al. (1974c) reported that the diffraction pattern obtained from poly d(G).poly d(C) fibres was similar to the A-DNA pattern of calf thymus DNA, both in terms of intensity distribution and unit cell dimensions. They did not observe the A → B transition which is normally observed in calf thymus DNA.

The remainder of this chapter is devoted to a description of data obtained from poly d(A).poly d(T) and poly d(G).poly d(C).

## 10.2 Materials and Method

The polynucleotides, poly d(A).poly d(T) and poly d(G).poly d(C) were obtained from Boehringer. Fibres of different salt content were prepared as described in Chapter 2. X-ray diffraction patterns were recorded from 33% to 98% relative humidities.

## 10.3 Results and Discussion

Low salt fibres prepared from Na poly d(A).poly d(T) gave a crystalline β-B'-DNA (orthorhombic) pattern for relative humidities up to 84%. This type of pattern was very similar to the pattern reported

by Arnott et al. (1974a). Further humidification of this fibre at 92% and 95% relative humidities gave an α-B'-DNA pattern. At 98% relative humidity a semi-crystalline B-pattern was observed, similar to the semi-crystalline B-form observed for A-T rich DNAs. When the relative humidity was reduced from 98% to 92%, a transition from the semi-crystalline B pattern to α-B'-DNA was observed. Further reduction of the relative humidity gave a "C-type" pattern (figure 1) instead of β-B'-DNA. The pitch and the base separation are 90.9 A and 3.24 A respectively. Hence, the helical symmetry is $28_3$. The molecules are packed in a hexagonal lattice with lattice parameters a = 19.2 A, c = 30.3 A (see table 1). Fibres prepared with slightly increased salt content gave a good crystalline β-'B-DNA (figure 2) up to 84% relative humidity. Further humidification at 92% and 95% gave a α-B'-DNA pattern (figure 3). At 98% relative humidity a semi-crystalline B pattern was observed (figure 4). The helical parameters and the lattice parameters of α-B'-DNA and β-B'-DNA are similar to those reported by Arnott et al. (1983). High salt fibres gave a triple-stranded DNA pattern (figure 5). No structural transition was observed for relative humidities between 33% and 98%.

Recently Arnott et al., (1983) proposed a heteronomous structure for β-B'-DNA. Further they suggested that both α-B'-DNA and β-B'-DNA have the same molecular structure. Since the α-B'-DNA pattern was observed for poly d(AT).poly d(AT) (see figure 12, Chapter 5) it is unlikely that α-B'-DNA can have a heteronomous structure as proposed by Arnott et al. (1983). The α-B'-DNA observed for poly d(AT).poly d(AT) and poly d(A).poly d(T) are shown in figure 6. The lattice parameters of α-B'-DNA from poly d(AT).poly d(AT) are a = 24.4 A, c = 32.8 A. The observed and calculated ρ values are given in table 2. The pitch and the base separation are 32.8 A and 3.28 A respectively. The two antiparallel chains in poly d(AT).poly d(AT) are chemically identical (as shown on the next page) and alternating adenine and thymine bases are present in both chains.

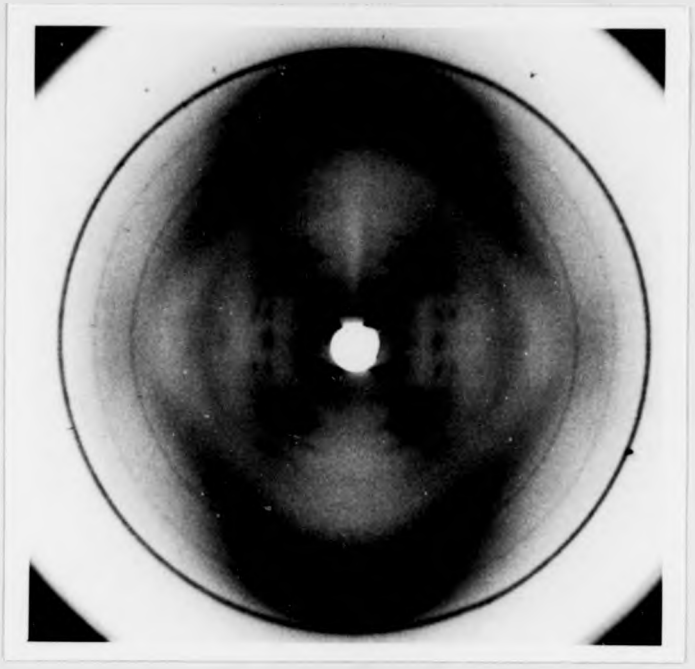


Figure 1 : Semi-crystalline C-type form Diffraction Pattern of Na poly d(A).poly d(T) at 57% R.H.

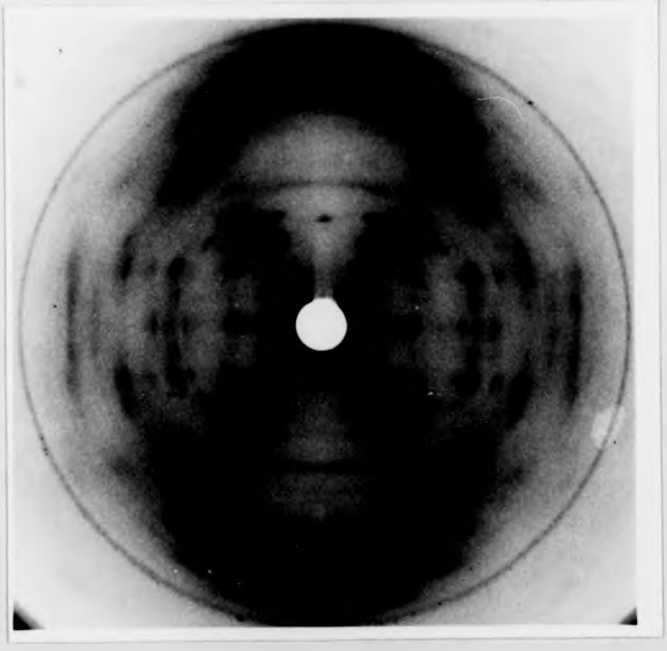


Figure 2 : β-B'-form Diffraction Pattern of Na poly d(A).poly d(T) at 44% R.H.

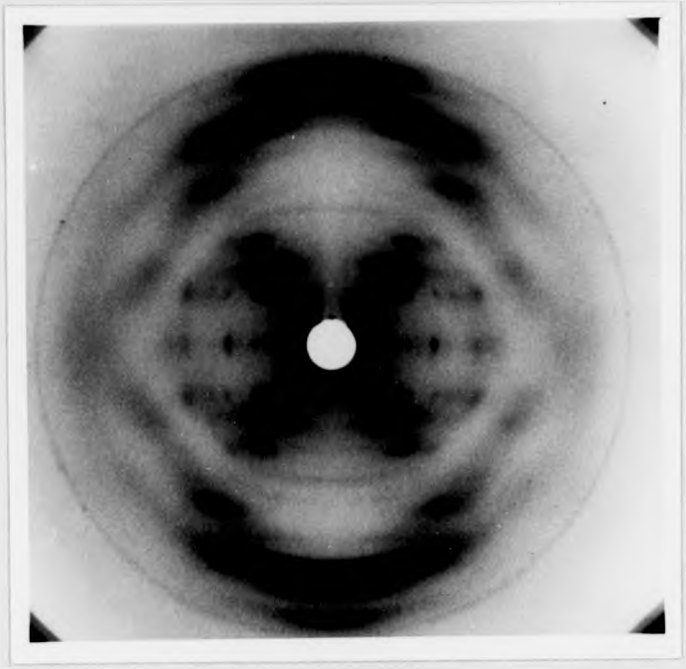


Figure 3 : α-B'-form Diffraction Pattern of Na poly d(A).poly d(T) at 92% R.H.

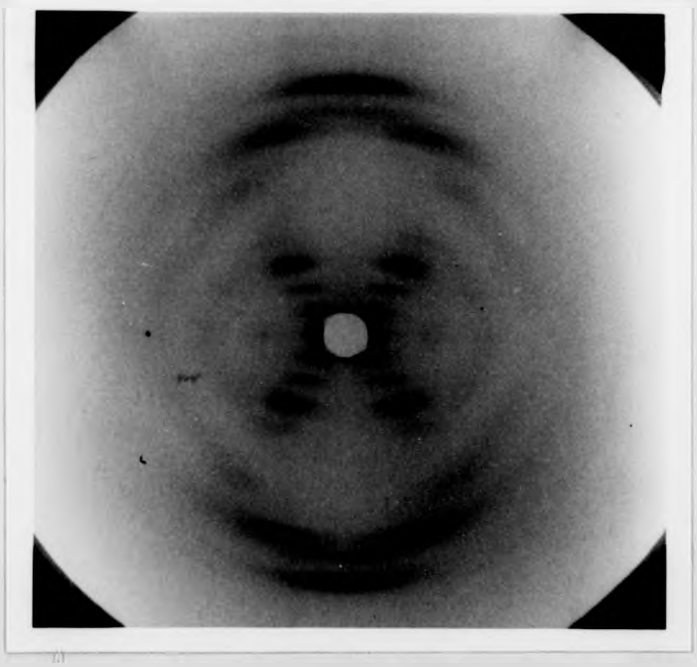


Figure 4 ; Semi-crystalline B-form Diffraction Pattern of Na poly d(A).poly d(T) at 98% R.H.

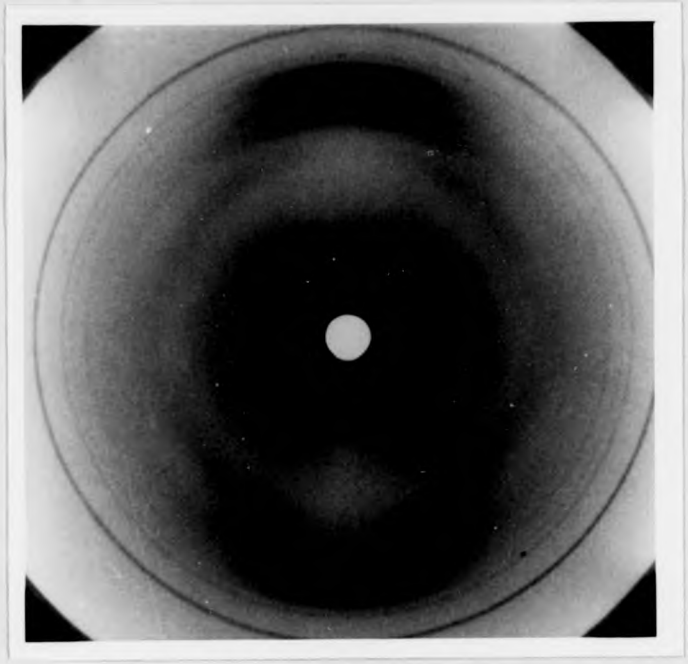


Figure 5 : Diffraction Pattern of Na poly d(T).poly d(A).poly d(T) at 75% R.H.

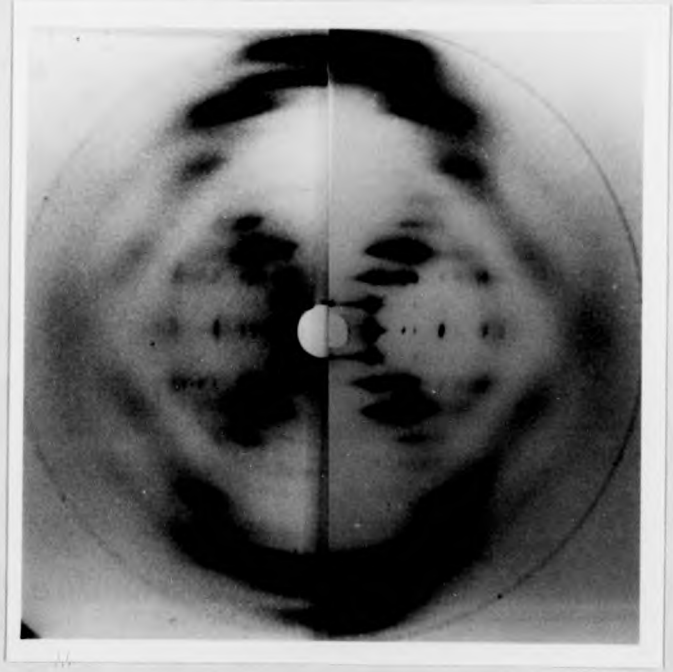


Figure 6 : α-B'-form Diffraction Pattern of Na poly d(AT).poly d(AT) (right) and Na poly d(A).poly d(T) (left)

A ---------- T

T ---------- A

A ---------- T

T ---------- A

However, in poly d(A).poly d(T) one chain is poly d(A) and the other chain is poly d(T) as shown below:

A ---------- T

A ---------- T

A ---------- T

A ---------- T

Hence, whilst the heteronomous model for α-B'-DNA seems unlikely, it is perhaps an acceptable structure for β-B'-DNA, as seen in poly d(A).poly d(T) at low humidities. When the relative humidity is increased the poly d(A) chain may change to a B-like conformation reversibly (which is similar to A ⇌ B transition in ordinary NaDNA fibres). Therefore, it is possible that the α-B'-DNA of poly d(A).poly d(T) has a 'homogeneous' (or mild heteronomous) structure in a high relative humidity environment as observed for poly d(AT).poly d(AT). Further structural studies are needed to establish the structural similarities between α-B'-DNA as observed for poly d(A).poly d(T), and poly d(AT).poly d(AT).

Since LiDNA fibres have never been observed in the A conformation it was decided to study the conformational polymorphism of Li poly d(A).poly d(T) fibres in the hope that it would case some light on the structural nature of α-B'-DNA and β-B'-DNA.

The polymorphic behaviour of Li poly d(A).poly d(T) fibres was distinct from Na poly d(A).poly d(T). At 44% relative humidity Li poly d(A).poly d(T) fibres gave a C-pattern (figure 7). The helical parameters, pitch and base separation are 29.0 A and 3.24 A respectively. Hence, the helical symmetry of this C-pattern is approximately $9_1$. When

the relative humidity of the fibre environment was increased to 66% and 75%, the pitch was increased to 30.4 A. The base separation was 3.28 A. Hence, the helical symmetry then became approximately $28_3$, and this was very similar to the C-pattern obtained from LiDNA. Only a few Bragg reflections were observed in these C-patterns, and it was thus difficult to determine the lattice parameters accurately. However, the molecular arrangement is hexagonal with a unit cell size of approximately 18 A. Hence, the number of molecules per unit cell can only be one (the C-form is normally observed to have three molecules per unit cell). At 92% (and above) relative humidities, the α-B'-DNA pattern was observed (figure 8, similar to that from Na poly d(A).poly d(T)). The C-patterns were observed upon reduction of the relative humidity to 75% and below. The crystallinity of the α-B'-DNA pattern was improved with a slight increase in the salt concentration in the fibre. It is interesting to note that Li poly d(A). poly d(T) was not observed in the β-B'-DNA conformation. This observation, coupled with other earlier X-ray fibre diffraction results of LiDNA fibres (natural and synthetic) strongly implies that neither chain in α-B'-DNA is in an A-type conformation (i.e. it is unlikely to be a heteronomous structure as proposed by Arnott et al., 1983). However, it is possible that β-B'-DNA is either a heteronomous structure as proposed by Arnott and his co-workers, or a homogeneous structure with both chains have the $C_3'$-endo sugar puckering (i.e. characteristic of the A-family).

The polymorphic behaviour of Li poly d(A).poly d(T), coupled with earlier X-ray fibre diffraction results of Li DNA fibres (natural and synthetic) strongly implies that both chains in α-B'-DNA are in a 'B-type' conformation (i.e. not a heteronomous structure as proposed by Arnott et al., 1983). The structure of β-B'-DNA, could, however, be heteronomous in nature (as proposed by Arnott et al., 1983), but there is no reason to reject outright the plausibility of a 'homogeneous' double helix with the

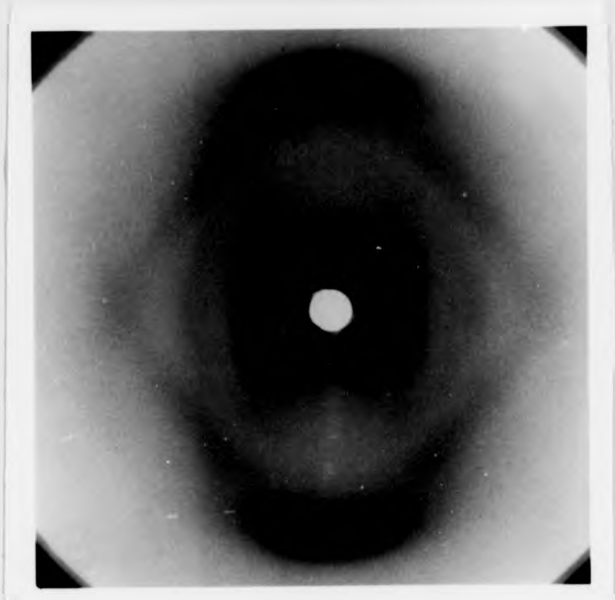


Figure 7 : Semi-crystalline C-form Diffraction Pattern of Li poly d(A).poly d(T) at 44% R.H.

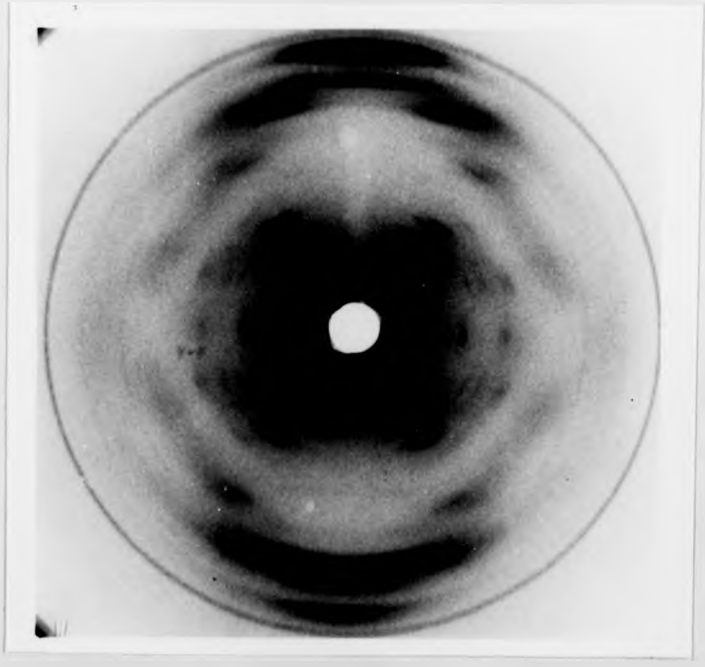


Figure 8 : α-B'-form Diffraction Pattern of Li poly d(A).poly d(T) at 98% R.H.

$C_3$'-endo sugar puckering characteristic of the A-family. Support for these two structural variants is given by the fact that whilst lithium is assumed never to favour the A conformation, the β-B'-DNA diffraction pattern is not obtained from the lithium salt of poly d(A).poly d(T). Hence, the fact that differing low humidity structures exist in the two salts (sodium and lithium) tends to suggest the involvement of an 'A-type' conformation in β-B'-DNA. This involvement could either take the form of a heteronomous structure, or a homogeneous structure.

The polymorphic behaviour of Rb poly d(A).poly d(T) was found to be similar to that of Na poly d(A).poly d(T). No D-form was observed as in Rb poly d(AT).poly d(AT).

Fibres prepared from poly d(G).poly d(C) gave a "B-type" pattern (figure 9) for humidities ranging from 33% to 92%. No conformational changes occurred with increase in relative humidity. However, at 98% relative humidity a semi-crystalline 'B-type' diffraction pattern was obtained. The 'B-type' (seen in figure 9) structure is a 10-fold helix with a pitch of 33.4 A. The molecular packing is hexagonal, with lattice parameters a = 24.8 A and c = 33.4 A. The calculated and observed ρ values are given in table 3. No A-form was observed as reported by Arnott et al., (1974c). Therefore, the homopolymers seem more prone to addopt 'B-like' structures rather than 'A-like' structures.

$C_3'$-endo sugar puckering characteristic of the A-family. Support for these two structural variants is given by the fact that whilst lithium is assumed never to favour the A conformation, the β-B'-DNA diffraction pattern is not obtained from the lithium salt of poly d(A).poly d(T). Hence, the fact that differing low humidity structures exist in the two salts (sodium and lithium) tends to suggest the involvement of an 'A-type' conformation in β-B'-DNA. This involvement could either take the form of a heteronomous structure, or a homogeneous structure.

The polymorphic behaviour of Rb poly d(A).poly d(T) was found to be similar to that of Na poly d(A).poly d(T). No D-form was observed as in Rb poly d(AT).poly d(AT).

Fibres prepared from poly d(G).poly d(C) gave a "B-type" pattern (figure 9) for humidities ranging from 33% to 92%. No conformational changes occurred with increase in relative humidity. However, at 98% relative humidity a semi-crystalline 'B-type' diffraction pattern was obtained. The 'B-type' (seen in figure 9) structure is a 10-fold helix with a pitch of 33.4 A. The molecular packing is hexagonal, with lattice parameters a = 24.8 A and c = 33.4 A. The calculated and observed ρ values are given in table 3. No A-form was observed as reported by Arnott et al., (1974c). Therefore, the homopolymers seem more prone to addopt 'B-like' structures rather than 'A-like' structures.

Table 1

The observed and calculated ρ-values for the "C-type" pattern in Figure 1

RH = 57% a = 19.2 A c = 30.3 A

| h | k | ℓ | $\rho_o(A^{-1})$ | $\rho_c(A^{-1})$ |
|---|---|---|---|---|
| 1 | 0 | 0 | 0.0593 | 0.0600 |
| 1 | 1 | 0 | 0.1040 | 0.1040 |
| 1 | 0 | 1 | 0.0682 | 0.0685 |
| 1 | 1 | 1 | 0.1090 | 0.1091 |
| 2 | 0 | 1 | 0.1252 | 0.1245 |
| 1 | 0 | 2 | 0.0885 | 0.0892 |
| 1 | 1 | 2 | 0.1227 | 0.1232 |
| 2 | 0 | 2 | 0.1374 | 0.1370 |
| 1 | 0 | 3 | 0.1162 | 0.1158 |

Table 2

The observed and calculated ρ-values for the α-B'-DNA pattern (from poly d(AT).poly d(AT)) in figure 6.

RH = 95% a = 24.4 A c = 32.8 A

| h | k | ℓ | $\rho_o(A^{-1})$ | $\rho_c(A^{-1})$ |
|---|---|---|---|---|
| 1 | 0 | 0 | 0.0470 | 0.0473 |
| 1 | 1 | 0 | 0.0823 | 0.0819 |
| 2 | 0 | 0 | 0.0950 | 0.0946 |
| 2 | 1 | 0 | 0.1254 | 0.1251 |
| 3 | 0 | 0 | 0.1416 | 0.1419 |
| 3 | 1 | 0 | 0.1710 | 0.1705 |
| 4 | 0 | 0 | 0.1886 | 0.1892 |
| 1 | 0 | 1 | 0.0550 | 0.0563 |
| 1 | 0 | 2 | 0.0775 | 0.0772 |
| 1 | 1 | 2 | 0.1022 | 0.1021 |
| 1 | 0 | 3 | 0.1031 | 0.1030 |

Table 3

The observed and calculated ρ-values for the 'B-type' pattern shown in figure 9

RH = 66% a = 24.8 A c = 33.4 A

| h | k | ℓ | $\rho_o(A^{-1})$ | $\rho_c(A^{-1})$ |
|---|---|---|---|---|
| 1 | 0 | 0 | 0.0468 | 0.0466 |
| 1 | 1 | 0 | 0.0801 | 0.0808 |
| 3 | 1 | 0 | 0.1682 | 0.1682 |
| 1 | 0 | 1 | 0.0549 | 0.0554 |
| 1 | 1 | 1 | 0.0866 | 0.0862 |
| 2 | 0 | 1 | 0.0984 | 0.0980 |

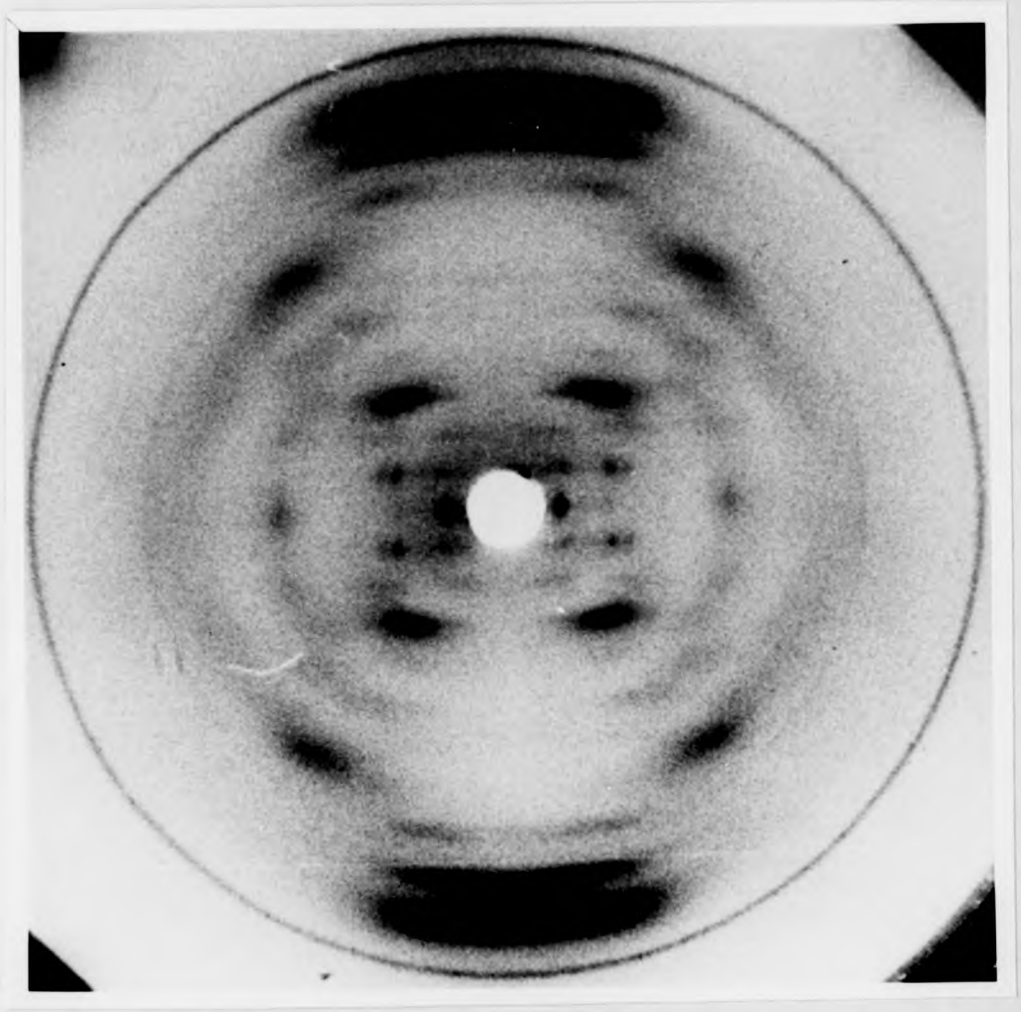


Figure 9 : "B-type" form Diffraction Pattern of Na poly d(G).poly d(C) at 66% R.H.

# CHAPTER 11

## Polymorphism of poly d(AC).poly d(GT)

### 11.1 Introduction

A greater variety of double helical conformations have been observed for synthetic than for naturally occurring deoxypolynucleotides. While this structural variety appears to be related to the nucleotide sequence it has not so far been possible to deduce any very strong correlation between it and the conformations assumed by a particular polynucleotide double-helix. Synthetic polynucleotides are not normally available in sufficient quantities to allow the techniques described by Cooper and Hamilton (1966) to be used systematically to investigate the effect of ionic content of the fibre on the polynucleotide conformation. Fibres of different ionic contents were prepared as discussed in Chapter 2. Poly d(AC).poly d(GT) is of particular interest because it contains equal molar fractions of the four nucleotides commonly found in naturally occurring DNAs and its study might therefore be expected to allow effects due to a regular nucleotide sequence to be distinguished from those due to overall nucleotide composition. Furthermore, poly d(AC).poly d(GT) shares a characteristic property of poly d(GC).poly d(GC) and poly d(AT).poly d(AT), namely, an alternating purine-pyrimidine sequence of bases. It contains A-T and G-C base pairs, and as such provides a useful comparison with the above mentioned co-polymers. Vorlickova et al., (1982) reported from circular dichroism studies that poly d(AC).poly d(GT) possibly gives a B → Z-like transition in solutions. Zimmer, et al., (1982) reported similar observations for circular dichroism studies in solutions. In a previous report of an X-ray fibre diffraction study of poly d(AC).poly d(GT)

(Langridge, 1969), it was stated that good x-ray diffraction patterns were obtained which were identical to those given by native DNA. However, apart from a brief discussion of the intensity of the eleventh layerline in the A pattern, no details were given of the patterns obtained nor of the conditions required for observing them. Leslie et al. (1980) reported a fully crystalline B-form in addition to a quite conventional A-form, and claimed that the A-pattern was indistinguishable from those obtained from other DNAs. Arnott et al. (1980) reported that the S form was observed occasionally for this polynucleotide, but supplied no qualifying data.

The remainder of this chapter is devoted to a description of the polymorphism of poly d(AC).poly d(GT) in the presence of NaCℓ, LiCℓ and KF salts.

## 11.2 Materials and Method

Poly d(AC).poly d(GT) was obtained from Boehringer and Dr. J. Brahms of the University of Paris. The final stage of purification of the material was by centrifugation at 50,000 r.p.m. from solutions for which the salt concentration was as low as 0.001 M. Fibres were drawn from either precipitated material or from a concentrated gel. X-ray diffraction patterns were obtained at a range of relative humidities of the fibre environment from 33% to 98%.

## 11.3 Results obtained from Na poly d(AC).poly d(GT) and K poly d(AC).poly d(GT)

The fibres prepared with low salt content gave a C-form (figure 1) at low relative humidities (typically from 33% to 57%). An increase in relative humidity gave a mixture of the A and C-forms (at 66% relative humidity) (figure 2). Further humidification (typically from 75% to 92% relative humidity) of the fibre environment gave a crystalline A-form (figure 3). Occasionally, mixtures of the A and B forms were observed at

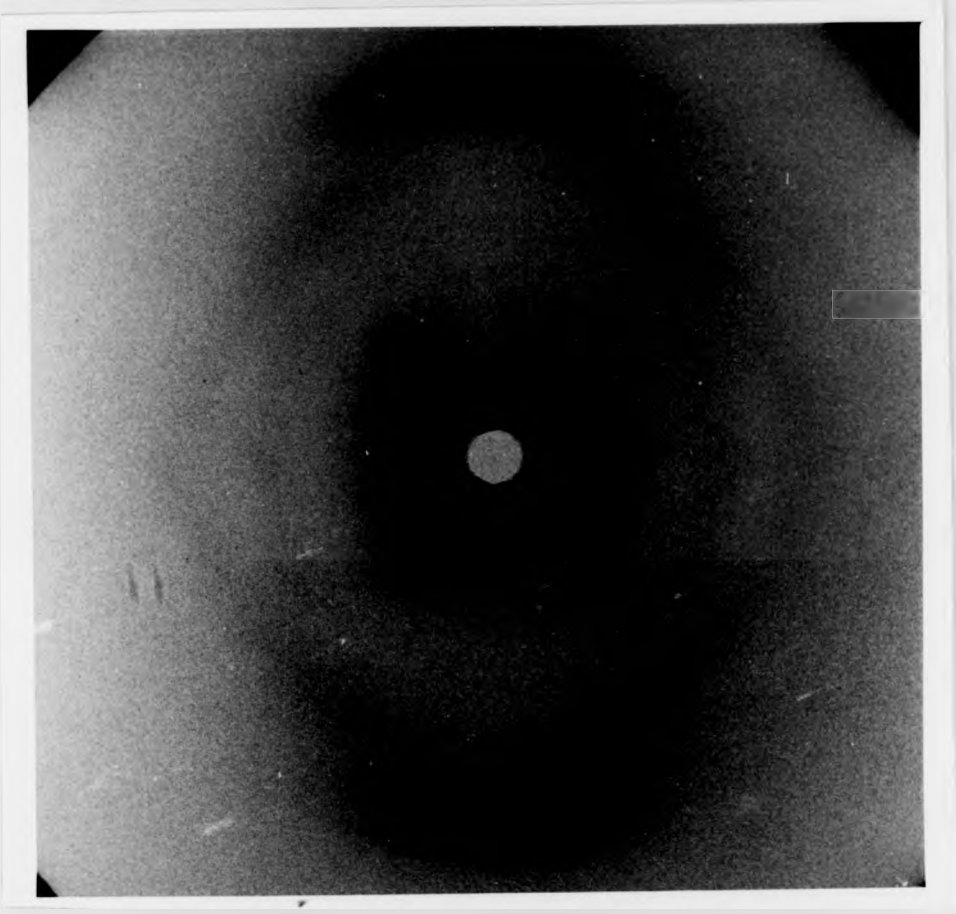


Figure 1 : Semi-crystalline C-form Diffraction Pattern of Na poly d(AC).poly d(GT) at 44% R.H.

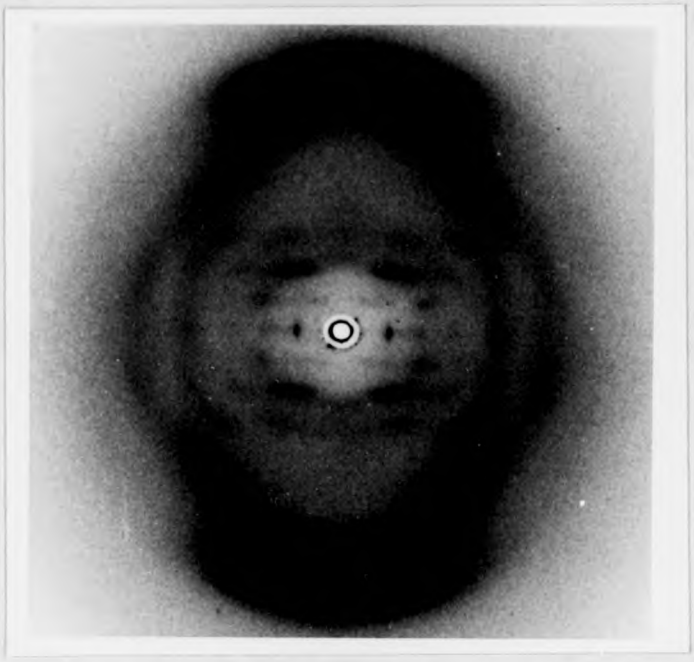


Figure 2 : Mixture of A and C forms. Diffraction Pattern of Na poly d(AC).poly d(GT) at 66% R.H.

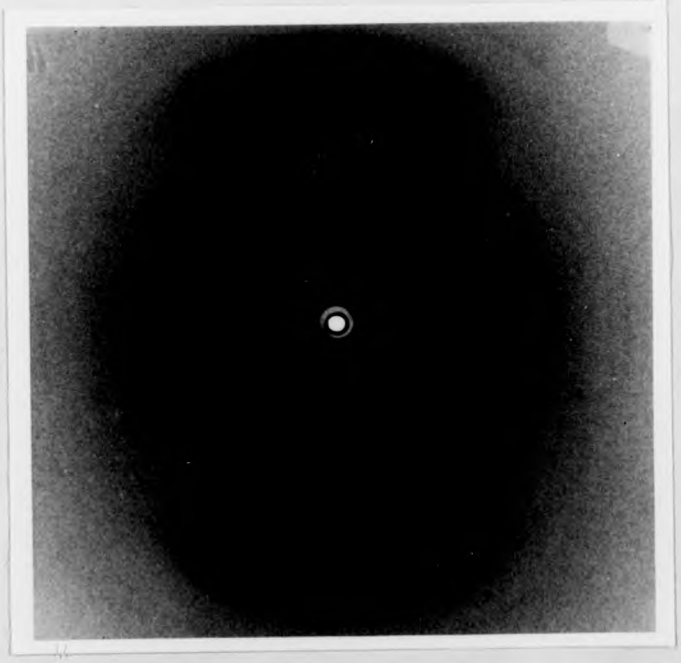


Figure 3 : A-form Diffraction Pattern of Na poly d(AC).poly d(GT) at 75% R.H.

92% or 95% relative humidities (figure 4). At 98% relative humidity, a semi-crystalline B-form was observed (figure 5).

Reducing the relative humidity of the fibre environment from 98% to 44% does not result in a simple reversal (as observed for natural DNAs, see Chapter 9), of the C → A → B conformational change as described above. While the B to A transition does occur when the relative humidity is reduced from 98% to 75%, an A-type diffraction pattern persists even when the relative humidity has been reduced to 33%. However, when fibres (which had been through the above described humidity sequence) were re-wet and re-drawn, they behaved in exactly the same way as the original in all respects of the humidity cycle.

Further increase in the salt content of the fibre gave an A and C mixture at low humidities (typically from 33% to 44%). An A-pattern was obtained at intermediate relative humidities (typically from 57% to 86%). At 98% relative humidity, a semi-crystalline B-form was observed. Reducing the humidity resulted in the appearance of the A-form from 86% to 33% relative humidity. High salt fibres gave poorly crystalline and generally unoriented B-type patterns, dominated by salt diffraction at lower humidities.

The polymorphic behaviour of low and medium salt K poly d(AC).poly d(GT) is very similar to that of Na poly d(AC).poly d(GT) fibres. However, the fibres prepared with slightly higher salt content gave an A- and "D-type" mixture pattern (figure 6) up to 92% relative humidity. A semi-crystalline B pattern was obtained at 98% relative humidity (figure 8). Reduction of the relative humidity of the fibre environment down to 33% gave a mixture of the A and D forms.

Further increase in the salt content of the fibre produced semi-crystalline B patterns, even at low humidities. The S form was not observed as reported by Arnott et al. (1980) neither for Na poly d(AC).poly d(GT) nor for K poly d(AC).poly d(GT) fibres.

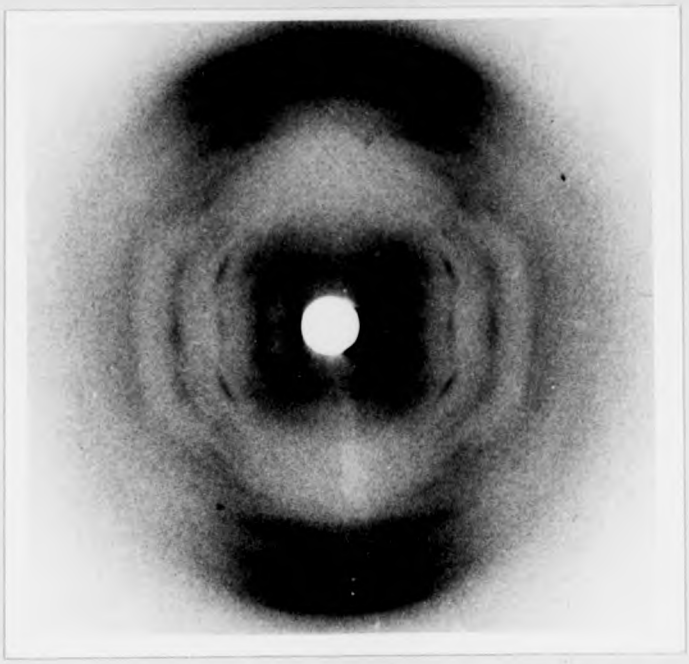


Figure 4 : Mixture of A and B forms, Diffraction Pattern of Na poly d(AC).poly d(GT) at 92% R.H.

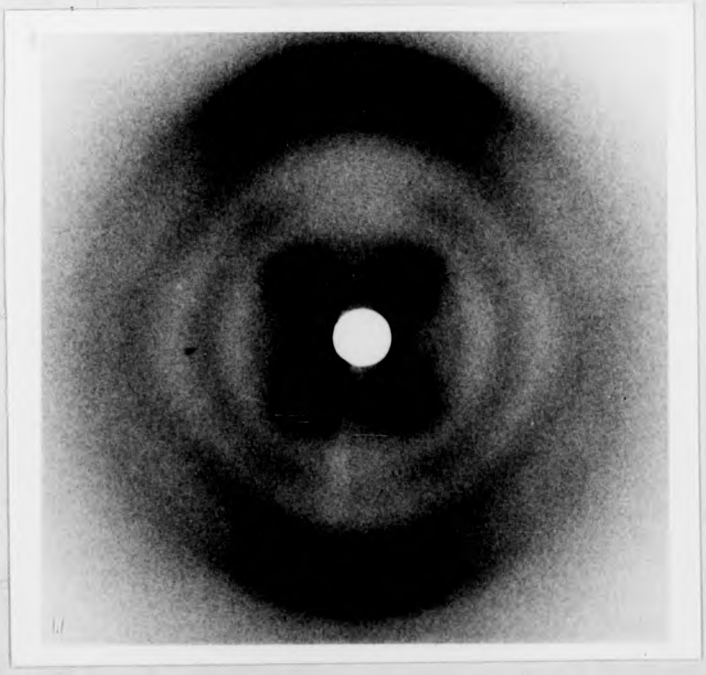


Figure 5 : Semi-crystalline B-form Diffraction Pattern of Na poly d(AC).poly d(GT) at 98% R.H.

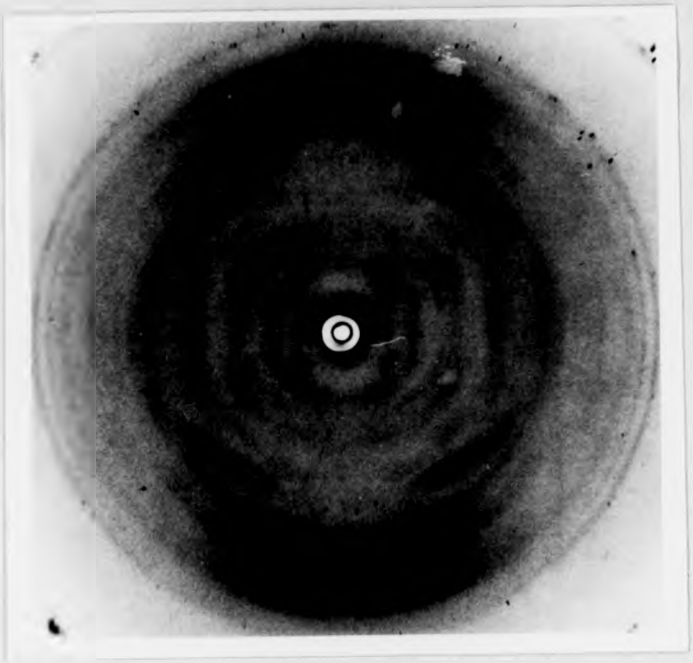


Figure 6 : Mixture of A and "D-type" forms. Diffraction Pattern of K poly d(AC).poly d(GT) at 92% R.H.

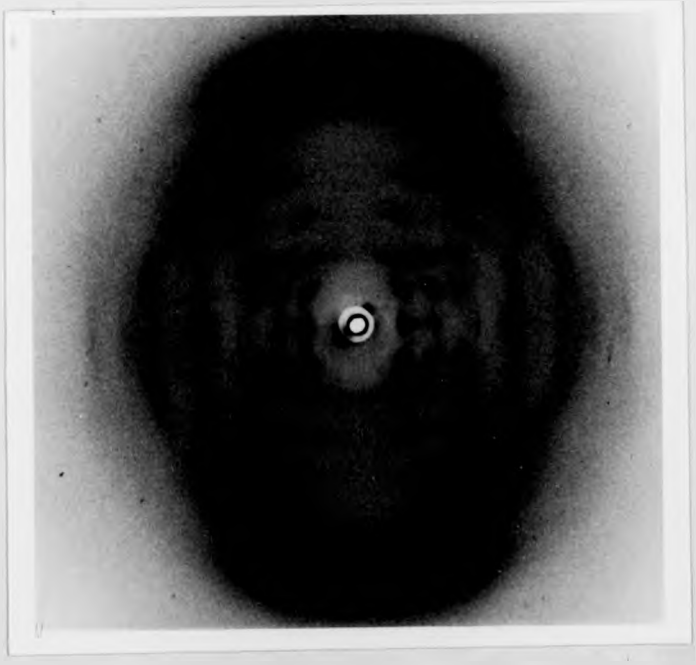


Figure 7 : Mixture of A and "D-type" forms. Diffraction Pattern of Na poly d(AT).poly d(AT) at 33% R.H.

11.3.1 Discussion

The synthetic polynucleotide poly d(AC).poly d(GT) is an example of DNA with an "average" nucleotide composition, but a highly regular base sequence. Whilst the C-form is also observed for fibres of this DNA containing low amounts of excess NaCℓ at low relative humidities, these fibres differ from those of natural DNAs in the reversibility of the transitions. The behaviour of poly d(AC).poly d(GT) is reproducible in this respect, as is evidenced by the results obtained from many fibres (both re-wetted and new). It may be that during drying of the gel in fibre preparation, excess salt tends to crystallise as a separate phase leaving the salt content in the polynucleotide phase so low that the C-form is favoured at low relative humidity. Gradual increase in the relative humidity of the fibre environment results in water entering the fibre which would be expected to dissolve the salt crystallites and lead to a more uniform distribution of salt. Subsequent reduction of the water content of the fibre may not result in the recurrence of a separate salt phase because the salt ions aggregate less readily in the wet fibre than in the gel, and as a consequence, the concentration of excess salt in the vicinity of the polynucleotide may then be above the level which even at a low humidity favours the C-form. At relative humidities in the range of 44% to 66% these fibres give C-patterns with a well-defined salt ring characteristic of NaCℓ micro-crystallites (figure 9, taken by Dr. N.J. Rhodes, in this laboratory). Increasing the relative humidity about 15% results in the disappearance of the salt ring. Neither this ring nor the C-pattern reappears when the relative humidity is reduced to 44%.

It should be emphasised that no indication of irreversibility in the C → A transition has been observed for any of the natural DNAs studied in the present work (see Chapter 9). It may be that this unusual behaviour of poly d(AC).poly d(GT) is a consequence of its particular nucleotide sequence resulting in a significantly different base-stacking contribution

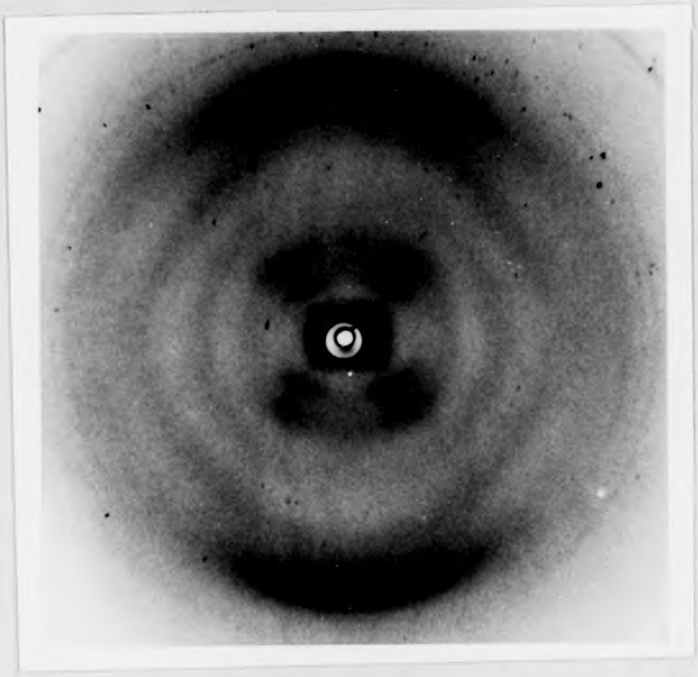


Figure 8 : Semi-crystalline B-form Diffraction Pattern of K poly d(AC).poly d(GT) at 98% R.H.

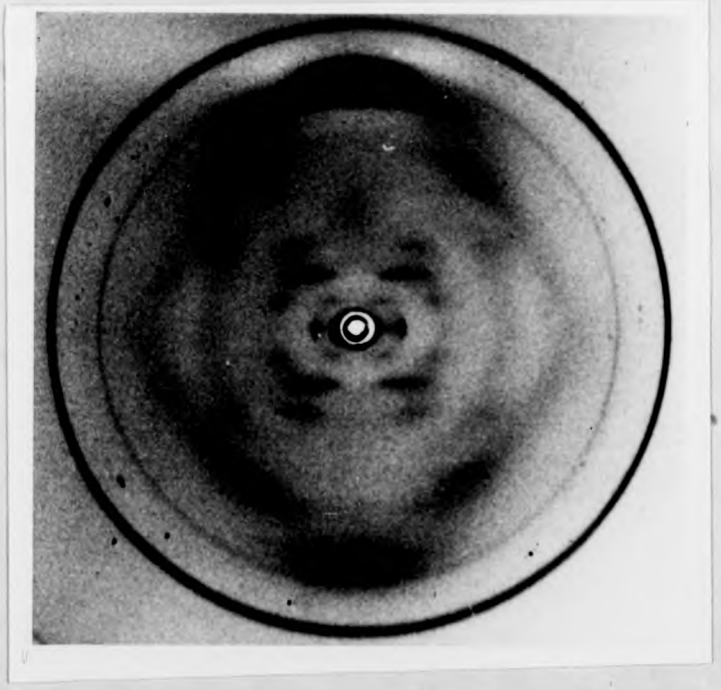


Figure 9 : Semi-crystalline C-form Diffraction Pattern of Na poly d(AC).poly d(GT) at 57% R.H.

to the conformational energy, or to favouring a different pattern of solvent interaction around the nucleic acid.

Arnott et al. (1980) reported that they have observed occasionally the S pattern for Na poly d(AC).poly d(GT) which is normally observed for poly d(GC).poly d(GC) under 'high salt' conditions (Arnott et al., 1980 and Mahendrasingam, et al., 1983). In the present work no S form was observed. However, a pattern suggesting an A/D mixture was observed for K poly d(AC). poly d(GT). This sort of mixture is normally observed for Na poly d(AT). poly d(AT), (Figure 7), (see Chapter 5). This is the first report to suggest that the D-conformation may be adopted in fibres of poly d(AC).poly d(GT). The polymorphic behaviour of poly d(AC).poly d(GT) thus appears to be very similar to either poly d(GC).poly d(GC) or poly d(AT).poly d(AT) depending on salt conditions in the fibres.

## 11.4 Results and Discussion for Li poly d(AC).poly d(GT) and $NH_4$ poly d(AC).poly d(GT) fibres

The material used in this study was precipitated with ethanol from 0.1 M salt solution. Fibres prepared from this material by redissolving in deionized water (without adding salt solution) gave well defined C patterns (figure 10) at low humidity (typically 33%). This persisted at increasing relative humidities until a transition occurred (typically at 95%) to a semi-crystalline B-form (figure 13) of the type observed at high humidity for both LiDNA and NaDNA (Langridge et al., 1960). The C-patterns were well defined throughout the humidity range over which they were observed, with the highest degree of crystallinity obtained at 66% (figure 12). The helical parameters of this C-pattern are quite different from those observed for LiDNAs (Marvin et al., 1961). The semi-crystalline C-forms from the lithium salt of naturally occurring DNAs are non-integral with approximately 9.3 nucleotide-pairs per continuous helix pitch of 31 A and are classified as $28_3$ helices and designated C (Marvin et al., 1961). The unit cell

parameters in this work for Li poly d(GGT).poly d(ACT) by Leslie et al., (1980) and quite distinct from the hexagonal lattice with a = 22.1 A and c = 58.2 A which these authors observed for the C" form of Na poly d(AG). poly d(CT). This C"-pattern has a $9_2$ helical symmetry. Hence, hereafter the C pattern observed for Li poly d(AC).poly d(GT) is denoted as a C'-form. The best crystalline pattern was obtained at 66% relative humidity and has been indexed in the present study. The observed and calculated $\rho$ values are given in table 1. The detailed structural analysis of this pattern has been carried out by Mr. V.T. Forsyth in our laboratory. Changes were observed in the detailed intensity distribution of the C-patterns as a function of the relative humidity (figure 15). Some of these changes were typical of a sampling effect due to variation in the lattice parameters as a function of the water content of the fibre, but others such as the variation in the relative intensity of the eighth and ninth layerlines indicate 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%) to a C-form which gave a pattern similar to that in figure 11 except that it was semi-crystalline with streaks rather than Bragg reflections on lower layerlines of odd order (figure 14). C-patterns of this general form continued to be observed as the relative humidity was reduced to 32%. This form is referred as C'*. The presence in the C'* pattern in figure 14 of streaks on $\ell$=1 and 3 and Bragg reflection of $\ell$=2 indicates that the molecules in the unit cell are randomly displaced from the relative position determined above for the C' structure by $^{c}/_{2}$.

Fibres prepared with slightly increased salt content gave crystalline and well defined C-patterns even at low relative humidities (typically from 33% to 57%) (figure 11, taken by Dr. N.J. Rhodes). In this pattern an additional meridional reflection half way between $\ell$= 4 and $\ell$=5

Table 1

The observed and calculated ρ-values of the C'-pattern at 66% relative humidity

RH = 66% a = 32.4 A c = 29.0 A

| h | k | ℓ | $\rho_o(A^{-1})$ | $\rho_c(A^{-1})$ |
|---|---|---|---|---|
| 1 | 1 | 0 | 0.0613 | 0.0616 |
| 3 | 0 | 0 | 0.1068 | 0.1068 |
| 1 | 0 | 1 | 0.0489 | 0.0496 |
| 2 | 1 | 1 | 0.1009 | 0.1003 |
| 3 | 0 | 1 | 0.1123 | 0.1122 |
| 1 | 0 | 2 | 0.0776 | 0.0776 |
| 2 | 0 | 2 | 0.0990 | 0.0991 |
| 2 | 1 | 2 | 0.1165 | 0.1167 |
| 2 | 1 | 3 | 0.1397 | 0.1399 |
| 2 | 2 | 6 | 0.2411 | 0.2409 |

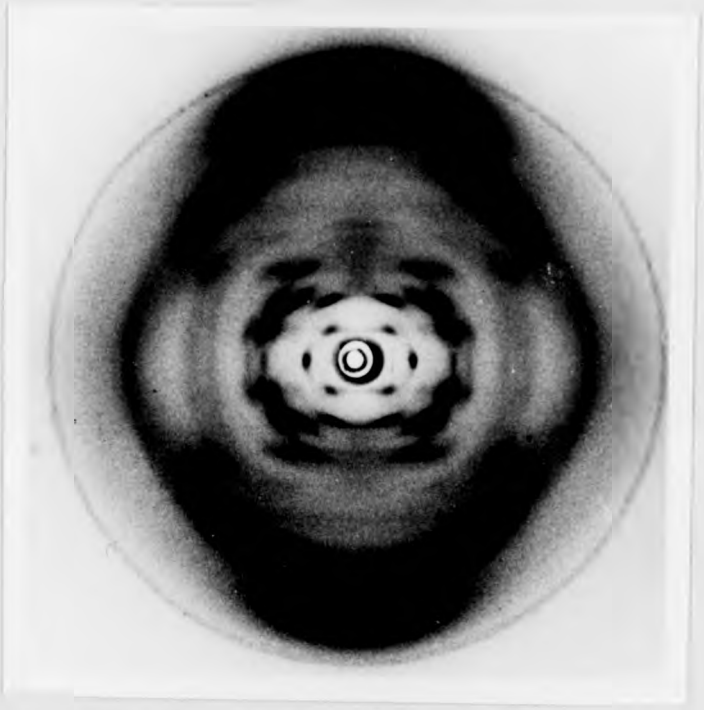


Figure 10 : Semi-crystalline C'-form Diffraction Pattern of Li poly d(AC).poly d(GT) at 57% R.H.

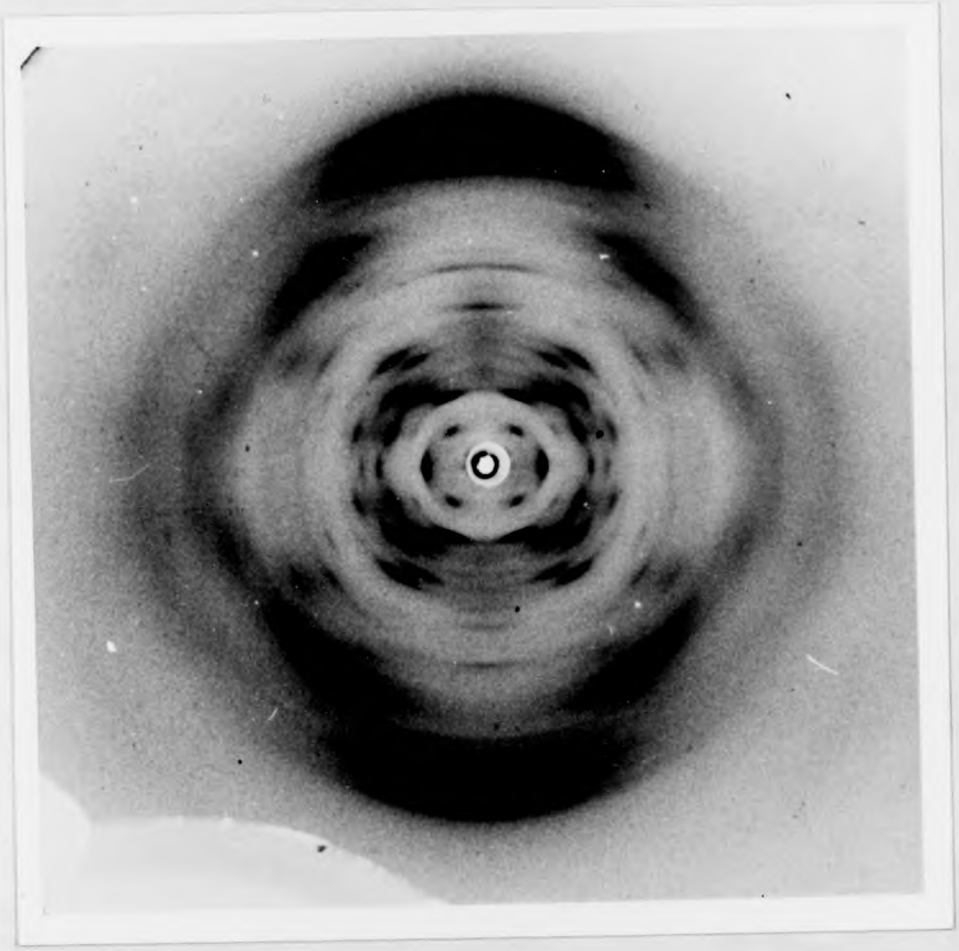


Figure 11 : Crystalline C'-form Diffraction Pattern of Li poly d(AC).poly d(GT) at 44% R.H.

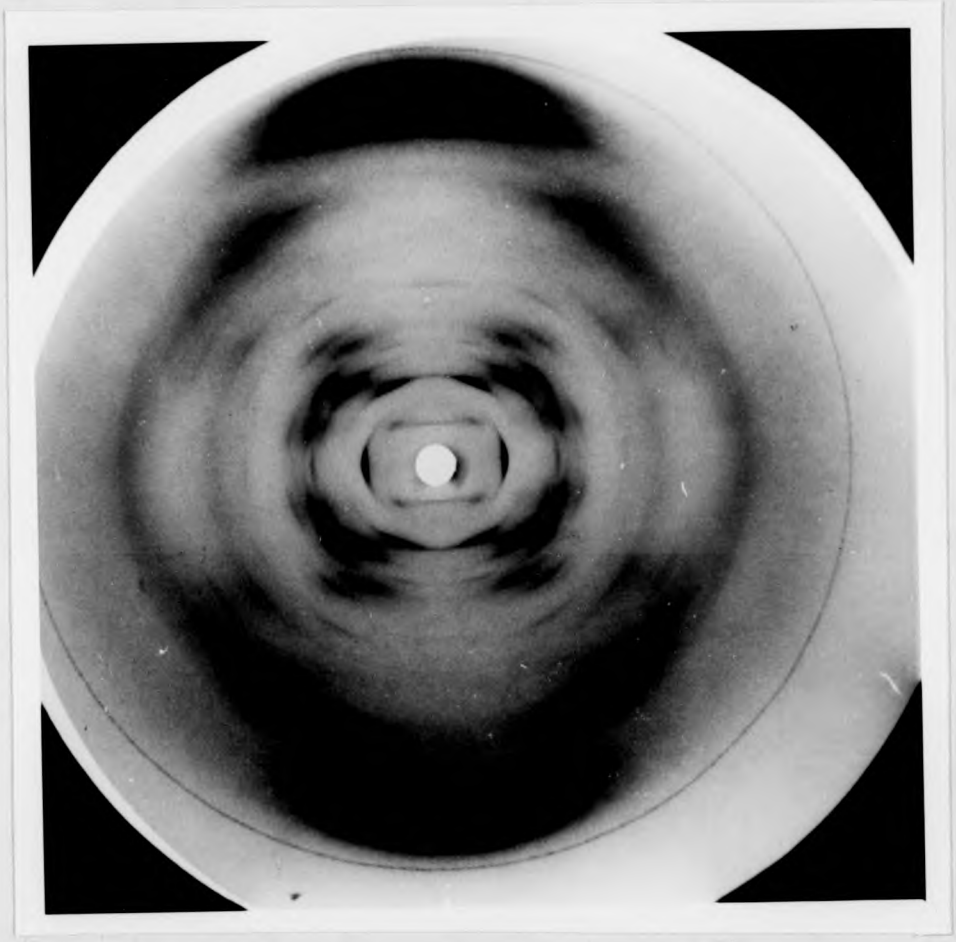


Figure 12 : Crystalline C'-form Diffraction Pattern of Li poly d(AC).poly d(GT) at 66% R.H.

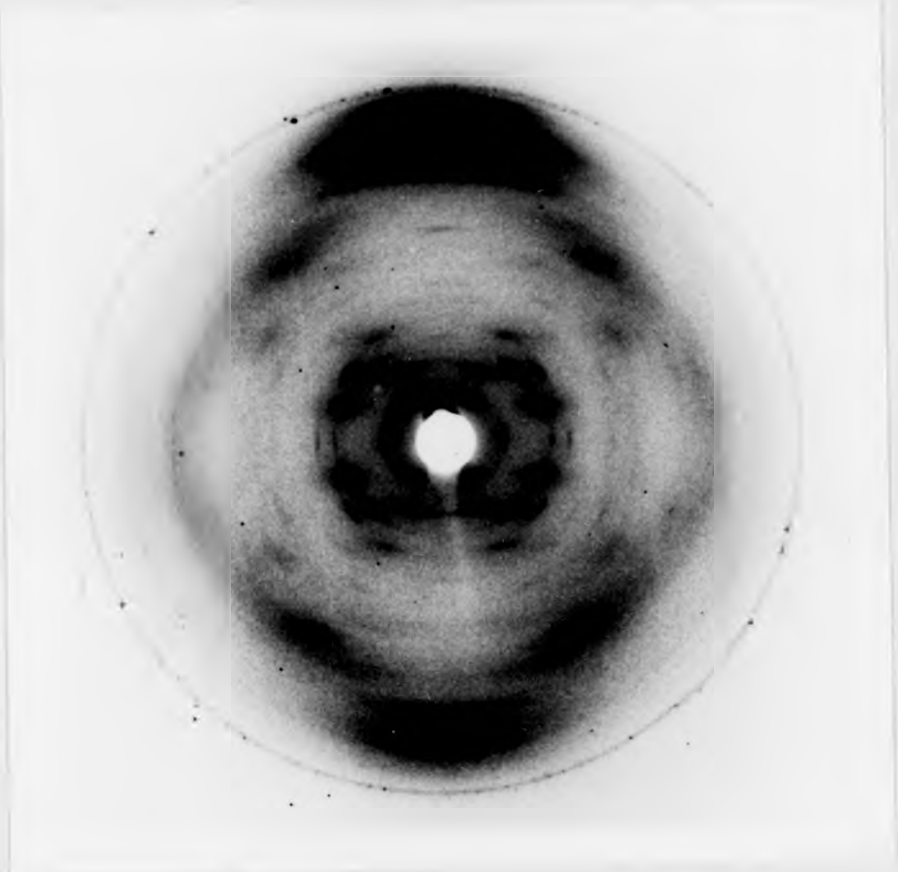


Figure 14 : Semi-crystalline C-form (C'*) Diffraction Pattern of Li poly d(AC).poly d(GT) at 33% R.H.

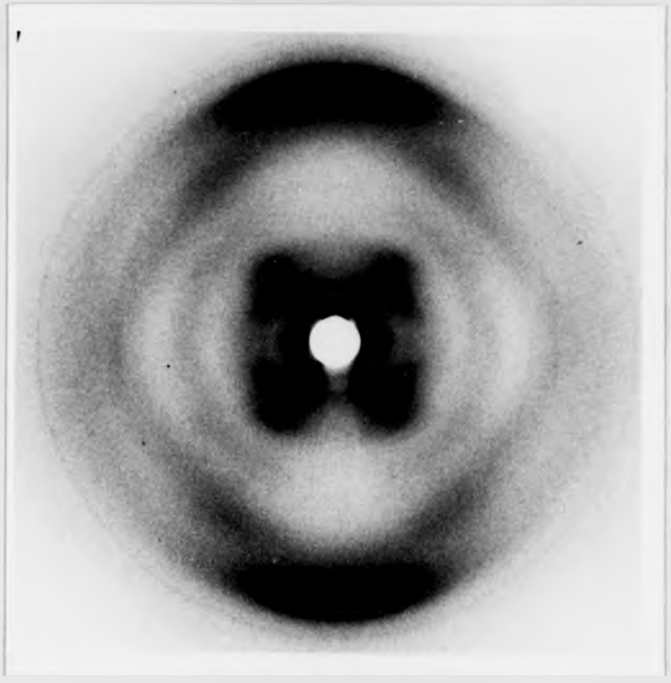


Figure 13 : Semi-crystalline B-form Diffraction Pattern of Li poly d(AC).poly d(GT) at 98% R.H.

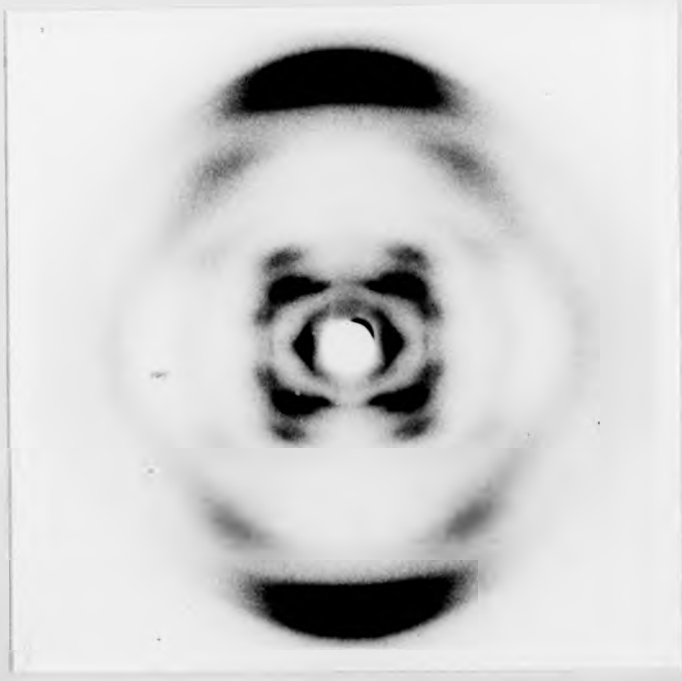


Figure 15 : Semi-crystalline C'-form Diffraction Pattern of Li poly d(AC).poly d(GT) at 75% R.H.

was observed. This can be attributed to either small differences between successive residues along the double-helix (as discussed in Chapter 7) or to ions or water ordered with a periodicity equal to twice the translation per nucleotide. The C'-form persisted up to 95% relative humidity. A semi-crystalline B-pattern was observed at 98% relative humidity. Reduction of the relative humidity of the environment resulted in a transition to the C-form and gave a pattern similar to figure 14 with streaks rather than Bragg reflections on lower layerlines of odd order (figure 16 taken by Dr. N.J. Rhodes).

The polymorphic behaviour of $NH_4$ poly d(AC).poly d(GT) is very similar to Li poly d(AC).poly d(GT). However, the C-patterns are not crystalline like in Li poly d(AC).poly d(GT) C' patterns. $NH_4$ poly d(AC).poly d(GT) C patterns (figure 17) are semi-crystalline and very similar to those obtained from Na poly d(AC).poly d(GT).

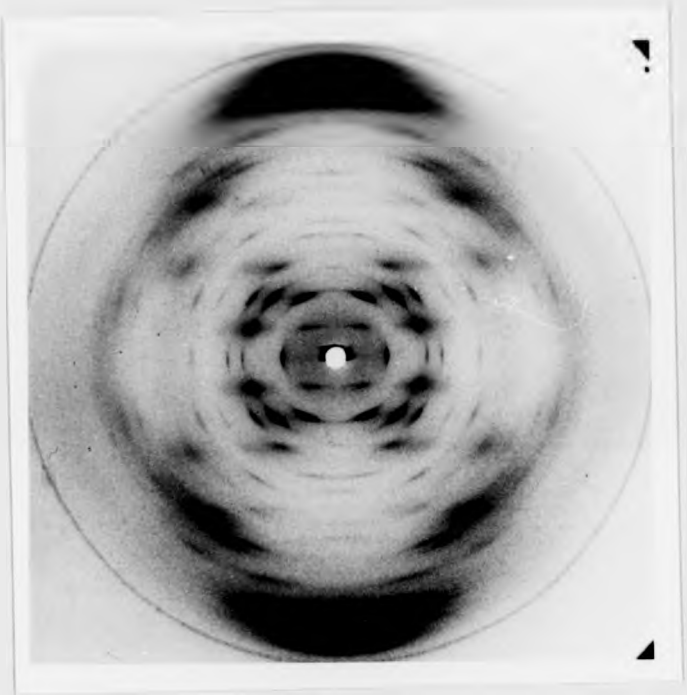


Figure 16 : Semi-crystalline C'-form (C'*) Diffraction Pattern of Li poly d(AC).poly d(GT) at 44% R.H.

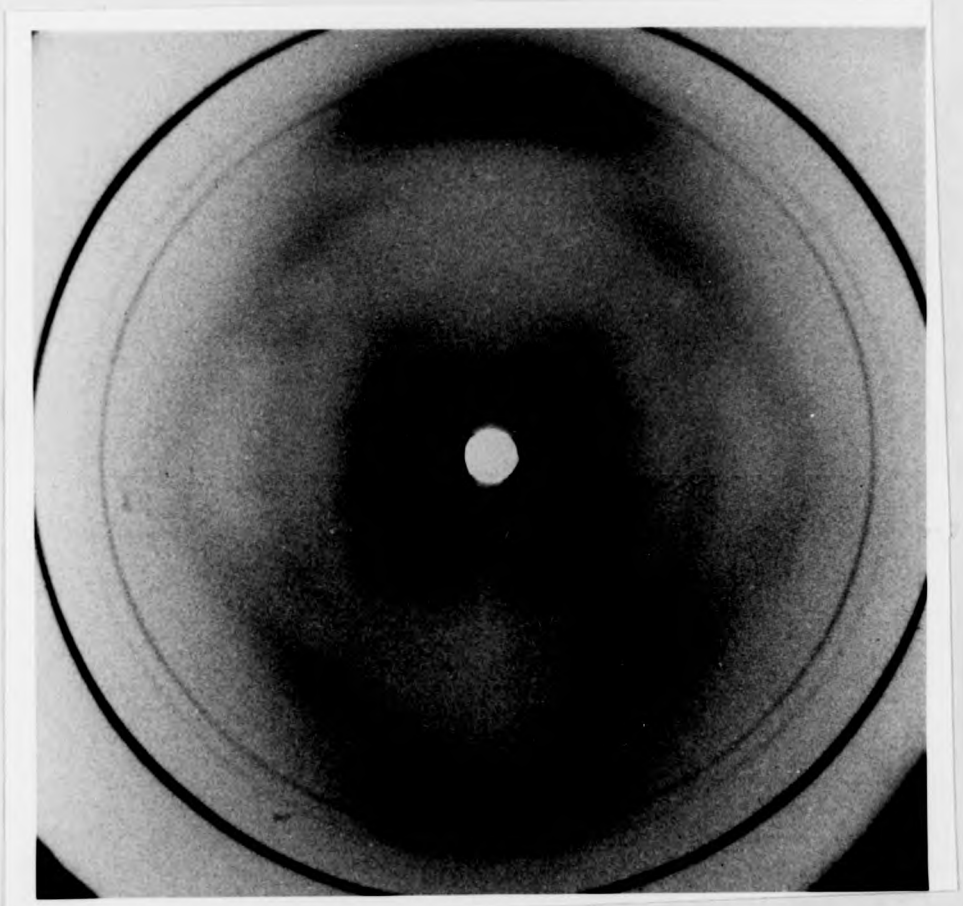


Figure 17 : Semi-crystalline C-form Diffraction Pattern of $NH_4$ poly d(AC).poly d(GT) at 44% R.H.

# CHAPTER 12

## Conclusion and Suggestions for Further Work

### 12.1 Introduction

A number of different techniques have been used in the study of synthetic and native polynucleotides since the overall structure of DNA was first elucidated in 1953 by Watson and Crick. During the course of these investigations, it has become abundantly clear that DNA is capable of a high degree of conformational flexibility and perusal of this polymorphism is marred by the fact that fibre diffraction only enables resolution down to $\sim$ 3 A. This limitation in our knowledge of the DNA molecule and the way in which it changes has to some extent been alleviated by the relatively recent availability of high quality single crystals which enable data collection and subsequent analysis to atomic resolution. The one proviso that must be attached to these data and the solutions obtained therefrom is an awareness of the fact that in all likelihood there are going to be differences between the polymer structure as determined from a crystal containing (for example) oligomers, and that determined from a 'continuous' polymer as envisaged in a fibre. Hence, although single crystal studies must be considered with caution they can be used constructively to corroborate evidence already obtained from fibres to stimulate new ideas. Examples of such corroboration have been provided by single crystal studies of the duodecamer (d(CGCGAATTCGCG)) which has been shown to be similar to the classical B structure (Wing, et al., 1980). Other workers have reported 'A-like' structures from work done on oligomers (Conner et al., 1981, Shakked, et al., 1981, 1983 and Wang et al., 1982a, 1982b). The availability of such data has further enabled the evolution of new theories relating

to sequence dependent local variation and hydration (Dickerson and Drew, 1981, Dickerson, 1983 and Drew and Dickerson, 1981).

12.2 Is D-DNA Left-Handed?

It has been found that there are striking similarities in the scheme of conformational transitions seen in poly d(AT).poly d(AT) and poly d(GC).poly d(GC) fibres.

(a) In Na poly d(AT).poly d(AT) fibres the sequence of transitions can be represented as follows

$$C \underset{\leftarrow}{\overset{\rightarrow}{}} A \rightarrow D \underset{\leftarrow}{\overset{\rightarrow}{}} B$$

$\rightarrow$ increasing relative humidity

(b) In K poly d(GC).poly d(GC) fibres the sequence is:-

$$B'' \underset{\leftarrow}{\overset{\rightarrow}{}} A \rightarrow S \underset{\leftarrow}{\overset{\rightarrow}{}} B$$

$\rightarrow$ increasing relative humidity

Once the D-form is assumed it is particularly stable even at high relative humidity, a feature which is reminiscent of the behaviour of S-DNA for poly d(GC).poly d(GC). Furthermore, another interesting similarity between the D and S forms is the observation that in K poly d(AT).poly d(AT), once the D form is assumed in the fibre, the following conformational transitions are observed:-

$$\alpha\text{-D-DNA} \underset{\leftarrow}{\overset{\rightarrow}{}} \beta\text{-D-DNA} \underset{\leftarrow}{\overset{\rightarrow}{}} B$$

The transition from $\alpha$-D-DNA to $\beta$-D-DNA was observed at 92% R.H. coupled with an increase in pitch from 24.0 A to 25.7 A, i.e. a 7% increase. In poly d(GC).poly d(GC), analogous behaviour is seen: once the S-form is assumed, the following sequence of transitions was observed:

$$S_I \underset{\leftarrow}{\overset{\rightarrow}{}} S_{II} \underset{\leftarrow}{\overset{\rightarrow}{}} B$$

The transition from $S_I$ to $S_{II}$ was observed at 92% or 95% relative humidity, coupled with an increase in pitch from 42.8 A to 45.3 A, i.e. a 7% increase.

This similarity in the behaviour of poly d(AT).poly d(AT) and

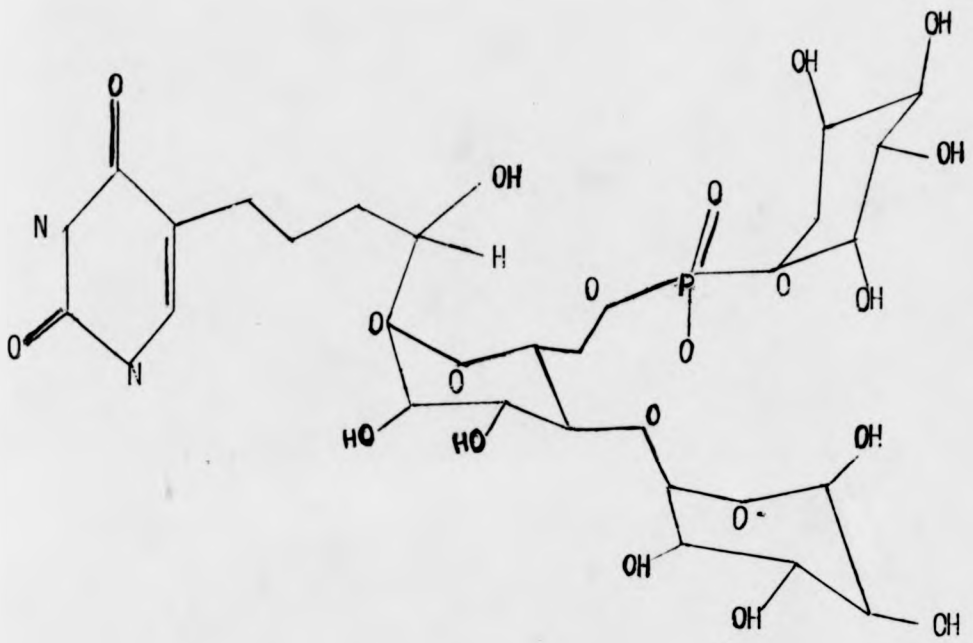
poly d(GC).poly d(GC), together with the fact that good left-handed models can be generated to explain the D-form diffraction pattern, give fairly strong support to the suggestion that D-DNA is a left-handed helix (see Chapter 4).

Another interesting feature about the D-DNA patterns obtained from poly d(AT).poly d(AT) is the first layerline meridional reflection (see chapter 4). This could be due to an ordered ion-water assembly in the major or minor grooves. The same reflection was not observed for A-DNA patterns (see Chapter 5) obtained from the same fibre. However, B-DNA patterns obtained from this fibre did contain the first layerline meridional reflection. The specific binding of ions in D-DNA may be one of the reasons for its great stability.

## 12.3 The B → A → C Transition

The crystal structure of B-DNA revealed a well ordered water structure in the minor groove, which is believed to be the stabilizing feature of the structure (Drew, et al., 1981). Conner et al., (1982) reported that in A-DNA, there is no trace of this spine of hydration. However, the hydrated phosphate groups participate in a network of solvation that laces together the two edges (phosphate group) of the major groove. The first layerline meridional reflection seen in B-DNA and B'-DNA patterns (see Chapter 5) may be due to an ordered water structure similar to the one observed in the crystal structure. Hence, the B → A transition could be caused by a disruption of the ordered water structure in the minor groove. Added support for this idea is furnished by studies on different native DNAs (Chapter 9). It was found that all natural DNAs (including phage, which has modified bases) could attain the B conformation. However, not all the DNAs could attain the A-form. In studies of phage DNA ($T_2$ DNA and SP15 DNA) fibres, the A conformation was never found to occur. These phage DNAs have modified bases. In $T_2$ DNA, all the cytosines

in it are replaced by hydroxymethyleytosines, 70% of which are glucosylated and 5% of which are diglucosylated (Wyatt et al., 1953 and Lehman et al., 1960). In SP-15 DNA, the fifth position of the methyl group is modified by a bulky group as shown below (private communication, Professor M. Mandel).




These modified groups lie in the major groove, and as such, may disturb the hydration layout necessary for the acquisition of the A-form. However, it is interesting to note that another phage DNA ($\phi$w-14) which has a small linear group attached to thymine methyl group (putrescene; $-CH_2$ NH $(CH_2)$ $NH_2$) does give an A-DNA pattern, but the transition was significantly difficult compared with that seen in calf thymus DNA. Therefore it is possible that the $-NH_2$ group at the end of the putrescene group forms

an electrostatic interaction with the phosphate group (i.e. across the major groove) oxygen in the next chain. Hence, the stability of the A-DNA structure found in $\phi$w-14 DNA could be higher than that from ordinary DNAs.

Another observation is the fact that low humidity C-DNA patterns obtained from poly d(AT).poly d(AT) have a meridional reflection on the first layerline that is similar to the one seen in B and B'-DNA. Given the fact that the C-conformation is supposedly a minor variant of the B-form it is conceivable that a similar hydration process is occurring, although perhaps not to the same extent (the relative intensity of the reflection is lower than that seen in B-DNAs). When the relative humidity of the environment of such fibres is increased, the C-pattern disappears and is replaced by an A-pattern in which there is no sign of the aforementioned reflection.

It is difficult to formulate structural theories of hydration on the basis of fibre diffraction alone. However, single crystal studies (Drew et al., 1981 and Conner et al., 1982) have, however, enabled resolution of water molecules and it is on the evidence of these studies (combined with that from fibre studies) that the following hypothesis has been put together.

The transition from C → A involves 'a winding up' of the helix such that the rotation per residue changes from 38.6 to 32.7, and it occurs at relatively low humidities. It is reasonable to assume that the most electronegative parts of the double helix (namely the phosphate groups) will become hydrated first, leaving a relative absence around the exposed polar base atoms of the minor groove. Such partial hydration could (in the scheme of the results obtained by Conner et al., 1982) be seen to result in a 'lacing together' of the two strands which deliniate the major groove, and which may then cause changes in sugar puckering, conformation angles and ultimately a

'winding' of the helix to form the observed A-structure.

In the context of this theory, further humidification is seen as a process whereby the polar base atoms exposed in the minor groove are more actively hydrated. This means that the sugar rings are likely to be influenced in a different way from that outlined above for the C → A transition. If such hydration induces a change in puckering back to the original $C_2$-endo, it could be envisaged that the A-conformation would then undergo a change (involving unwinding) back to a structure similar to the original C type. In practice, a 'B' structure is usually observed, but it is believed that B-DNA and C-DNA are closely related. At high humidities, the B conformation is assumed to be stabilized by a strong water network in the minor groove, implying that the minor groove is narrow and the major wide enough to prevent the 'lacing' which possibly stabilizes the A-form.

## 12.4 B" → A → S → B Transition

The observation of the left-handed Z-form of DNA has created a great interest in transitions between the various conformations of different DNAs. As discussed in Chapter 8, K poly d(GC).poly d(GC) fibres have been observed to undergo the following conformational transition: B" $\underset{\leftarrow}{\overset{\rightarrow}{}}$ A → S $\underset{\leftarrow}{\overset{\rightarrow}{}}$ B by simply changing the relative humidity.

### 12.4.1 B" → A transition

The mechanism of the B" → A transition is possibly analogous to the C → A transition in Na poly d(AT).poly d(AT) and other NaDNAs discussed in this thesis. During humidification of the fibre, the phosphate group atoms are more quickly hydrated than the groups at the edges of the bases. This then might produce winding of the helix and cause a change in the sugar pucker from $C_2$-endo to $C_3$-endo, the former a feature of the B-form, and the latter a feature of the A-form. Furthermore, as suggested for the

C → A transition, the hydration around the phosphate groups possibly laces the edges of the major groove together, resulting in the eventual appearance of a stabilized A-conformation.

12.4.2 A → S → B transition

It is reasonably well established that the A and B forms are right-handed and that the S-form is left-handed. Hence a transition involving a change in helix sense is of considerable importance. Observations that tend to support this idea are:-

(a) The A → S and S → B transitions are observed through intermediate mixtures of (A/S) and (B/S) respectively. These intermediate mixtures do contain diffuse scattering in the centre of the diffraction pattern, which possibly indicates the presence of non-crystalline (or amorphous material). This could be construed as a consequence of the change of the handedness involving some sort of amorphous intermediate state. A similar state might explain diffuse scattering on (B/S) mixture patterns. This type of diffuse scattering in the centre of the pattern has not been observed in either (C/A) mixture patterns of (A/B) mixture patterns. In the C → A → B transitions, no change in helix sense is involved, and changes possibly just involve simple winding and unwinding. On this basis one would not expect any great reorganisation involving an amorphous state.

(b) It was found that a period of annealing at about 92% is required for B → S transition to occur. Once the S-form is assumed, it appears to be locked in this structure and can only be induced to change from it by raising the relative humidity to 98% when it changes to the B-form. If the relative humidity is then dropped from 98% to 75%, a B → A transition was observed. These observations again imply to S-DNA is left-handed, since fast changes in humidity do not appear to permit its reacquisition.

12.5 Polymorphism of poly d(AT).poly d(AT)

It was shown that fibres prepared from poly d(AT).poly d(AT) could attain a variety of conformations (C, A, α-D, β-D, D', D-type, B' and B) depending on the 'type' and 'concentration' of ions in the fibre, and also the relative humidity of the fibre environment.

Also it has been shown that a single fibre could acquire the C, A, D, B' and B forms by simply changing the relative humidity of the fibre environment.

A crystalline B-form has also been observed for Na poly d(AT). poly d(AT). This form is normally associated uniquely with the lithium salt.

12.6 Polymorphism of poly d(GC).poly d(GC)

It has been shown that poly d(GC).poly d(GC) can adopt a variety of conformations (C, B", A, $S_I$, $S_{II}$ and B) depending on the 'type' and 'concentration' of ions in the fibre, and also the relative humidity of the fibre environment.

It has also been shown that a single fibre can adopt the B", A, $S_I$, $S_{II}$ and B forms by simply changing the relative humidity of the fibre environment.

Left-handed structures have been proposed to explain the $S_I$ and $S_{II}$ diffraction patterns from K poly d(GC).poly d(GC).

12.7 Polymorphism of poly d(AC).poly d(GT)

This polynucleotide is one in which the base composition is quantitatively what one would expect (on average) from most natural DNAs. It is, however, distinct in that whereas natural DNAs (from a structural point of view) are, 'random' in sequence, poly d(AC).poly d(GT) is obviously very regular. These features make it a particularly interesting polymer to study in relation to the work that has already been described on

alternating A-T base pairs and alternating G-C base pairs. It also makes it interesting in that any difference between native DNA and poly d(AC). poly d(GT) must presumably be due to the repetitive regularity of the synthetic polymer.

It was found that Na poly d(AC).poly d(GT) fibres prepared under low salt conditions gave a C → A → B sequence of transitions, which is similar to that observed for the sodium salt of native DNAs. However, the C → A transition was irreversible for Na poly d(AC).poly d(GT) and reversible for Na DNAs. Arnott et al. (1980) did observe the S-form for Na poly d(AC). poly d(GT). The S-form was not observed for Na poly d(AC).poly d(GT) in the present work. However, in the present work 'D-type' patterns were observed from K poly d(AC).poly d(GT) fibres. At high relative humidity this pattern changed into semi-crystalline B-pattern. This is somewhat similar to the polymorphic behaviour of Na poly d(AT).poly d(AT). Vorlickova et al., (1982) reported from circular dichroism studies that the salt-induced transition of poly d(AC).poly d(GT) is by low degree of cooperativity, the ionic specificity and kinetics being much more similar to that of poly d(AT).poly d(AT) (Vorlickova et al., 1980) than to poly d(GC).poly d(GC) (Pohl and Jovin, 1972). The polymorphic behaviour of poly d(AC).poly d(GT) thus seems to be similar either to poly d(AT).poly d(AT) or to poly d(GC).poly d(GC) depending on the conditions (e.g. 'type' and concentration of ions, relative humidity).

For Li poly d(AC).poly d(GT) fibres, a crystalline C'-form was observed at 66% relative humidity. It changed to a semi-crystalline B-form at 98% relative humidity. When the relative humidity was reduced the C'-pattern was observed, but it was screw disordered.

## 12.8 Mixtures of (A/C) and (A/B)

In A-DNA patterns, the first equatorial reflection (at $\sim$ 0.055 A) and the meridional reflection at $\sim$ 3.3 A were normally regarded as being a

B-form impurity. Since the observation of the C-form at low relative humidity (Rhodes et al., 1982) the above reflection observed in the A-DNA diffraction pattern could possibly be from the C-form rather than the B-form. The above possibility has been discussed in Chapter 6.

## 12.9 Does the heteronomous structure exist in homopolymers?

Arnott et al., (1983) has proposed a heteronomous structure for α-B'-DNA and β-B'-DNA for Na poly d(A).poly d(T). In Chapter 10, the possibility of heteronomous and homogeneous structures has been considered. It was concluded from the evidence available that β-B'-DNA is possibly a heteronomous structure and α-B'-DNA a homogeneous structure.

## 12.10 Alternating Structures

The possibility of the existence of the alternating structures in the co-polymers has been discussed in Chapter 7. There is reasonable evidence available to suggest the existence of alternating structures in co-polymers.

## 12.11 A-DNA

A-DNA has been observed for all natural DNAs with the exception of SP-15 and $T_2$-DNAs and DNA's containing the lithium ion. The A-form was observed in poly d(AT).poly d(AT), poly d(GC).poly d(GC) and poly d(AC). poly d(GT). However, in the presence of $Rb^+$ or $Cs^+$ ions poly d(AT). poly d(AT) was not observed to adopt the A-conformation.

## 12.12 B-DNA

The B-form was observed for all DNAs (natural and synthetic) irrespective of base composition, sequence, and the type of ions present. However, the intensity distribution on these pattterns varies considerably.

In the B-form of A-T rich DNAs, intensities on the first and third layerlines are stronger in relation to those of the second layerline. However, in G-C rich DNAs this observation is reversed: the second layerline reflections are stronger than those of the first and third layerlines. One explanation for this is that the minor groove of the B-form in A-T rich DNAs is narrower than that in G-C rich DNAs. Hence, these two regions of DNA could possibly have a distinct biological role.

## 12.13 C-DNA

The C-form was observed for LiDNAs at low relative humidities (Marvin et al., 1960). For the sodium salt, it was regarded as an unstable conformation. However, it has been established from this work and Rhodes et al., (1982) that the C-form is a low salt and low humidity form. The C-form is also observed with the lithium and sodium salts of all natural DNAs. Most of the synthetic DNAs gave a C-form at low humidity and low salt conditions.

## 12.14 Suggestions for Further Work

It is abundantly clear that DNAs (natural and synthetic) are capable of a stunning degree of polymorphism. However, very little is understood about the mechanism of these conformational transitions. Hence, further work in this direction is very important and highly necessary for a fuller knowledge of the molecular basis of genetics. The conformations observed in DNA fibres depends mainly on the following variables:-

(a) 'Type' and 'concentration' of the ions present in the fibre.

(b) The relative humidity of the fibre environment.

Condition (a) is very difficult to change during an experiment (i.e. once the fibre is made, it is constant). Therefore, the only variable accessible during the experiment is the relative humidity of the fibre environment. It has been shown throughout this work that fibres made in precisely the right

ionic conditions can be made to undergo a great variety of transitions simply by changing the relative humidity of the fibre environment (Rhodes et al., 1982, and Mahendrasingam, et al., 1983a, b). Hence the ability to make fibres in a highly controlled way is of paramount importance. Another extremely important factor is the X-ray source in use. A typical exposure time using a conventional X-ray source (GX6; rotating anode) is fifteen hours, and time resolution is obviously limited. However, data can be collected at Daresbury Laboratory, England, such that an average exposure on film can be made in $\sim$ 5 minutes and this offers considerably increased time resolution. The use of the synchrotron radiation source (SRS) has been further supplemented by the availability of a multi-wire proportional chamber (MWPC) which has been used in place of film. Using this detector exposures can be made so that different conformations can be well identified within 30 seconds. In its present state of commission, the SRS system has already enabled quantitative examination of the C $\rightarrow$ A transition. A project to study conformational transitions in poly d(AT).poly d(AT) has recently been supported by SERC and is currently in progress.

During the past few years the handedness of the DNA double-helix has emerged as a central question in investigation of the relationship between its molecular structure and biological function. For most of the thirty years since Watson and Crick proposed their two-stranded model for DNA, the major emphasis in physical and chemical studies of DNA has been on the analysis of data in terms of right-handed helical models. However, recent studies of the synthetic polynucleotide poly d(GC).poly d(GC) and related oligomers has focussed attention on the detailed structure of left-handed conformations, and the possible mechanisms of transition between left- and right-handed conformations. A project to study the transitions between left- and right-handed conformations of poly d(GC).poly d(GC) has just been funded by SERC and is currently in progress.

Another important area for the future work is in the interaction of drug and protein molecules with DNA. How do these molecules interact with DNA? What effect do they have on DNA structures? How do they effect conformational transitions? All these problems are not only of direct interest to the biophysical community but are also of importance in the search for a sensible biostructural approach to drug design.

## BIBLIOGRAPHY


Altona, C. and Sundaralingam, M., (1972), J.Am.Chem.Soc., 94, 8205.

Arnott, S., (1970), Prog.Biophys.Mol.Biol., 21, 265.

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

Arnott, S. and Hukins, D.W.L., (1972b), Biochem.Biophys.Res.Commun., 48, 1392.

Arnott, S., (1973), Trans.Am.Crystallog.Assoc., 9, 31.

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

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

Arnott, S., Chandrasekharan, R., Hukins, D.W.L., Smith, P.J.C. and Watts, L., (1974b), J.Mol.Biol., 88, 523.

Arnott, S. and Selsing, E., (1974c), J.Mol.Biol., 88, 551.

Arnott, S., (1974), Biochem.J., 141, 537.

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

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

Arnott, S., Chandrasekharan, R., Hall, I.H. and Puigjaner, L.C., (1983), Nucl.Acid.Res., 11

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

Behe, M., Zimmerman, S. and Felsenfeld, G., (1981), Nature, 293, 233.

Blakeley, P.J., (1976), Ph.D. Thesis, University of Keele.

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

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

Bram, S., (1971), Nature New Biol., 232, 174.

Bram, S. and Tougard, P., (1972), Nature New Biol., 239, 123.

Bram, S., (1973), Proc.Nat.Acad.Sci. U.S.A., 70, 2167.

Bram, S. and Baudy, P., (1974), Nature, 250, 414.

Buerger, M.J., (1945), X-ray Crystallography, John Wiley and Sons, New York.

Cantor, C.R. and Schimmel, P.R., (1980), Biophysical Chemistry, Part II. W.H. Freeman and Company.

Chesley, G.F., (1947), Rev.Sci.Instr., 18, 422.

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

Conner, B.N., Takano, T., Tanaka, S., Itakura, K. and Dickerson, R.E., (1982), Nature, 295, 294.

Cooper, P.J. and Hamilton, L.D., (1966), J.Mol.Biol., 16, 562.

Davies, D.R. and Baldwin, R.L. (1963), J.Mol.Biol., 6, 251.

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

Dickerson, R.E., (1983), J.Mol.Biol., 166, 419.

Drew, H.R., Takano, T., Tanaka, S., Itakura, K. and Dickerson, R.E., (1980), Nature, 286, 567.

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

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

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

Finch, J.T. Lutter, L.C., Rhodes, D., Brown, R.S., Rushton, B., Levitt, M. and Klug, A., (1977), Nature, 269, 29.

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

Franklin, R.E. and Gosling, R.G., (1953a), Nature, 171, 740.

Franklin, R.E. and Gosling, R.G., (1953b), Acta.Cryst., 6, 673.

Franklin, R.E. and Gosling, R.G. (1953c), Acta. Cryst., 6, 678.

Fuller, W., (1961), Ph.D. Thesis, University of London.

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

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

Goodwin, D.C., (1977), Ph.D.Thesis, University of Keele.

Greenall, R.J., Pigram, W.J. and Fuller, W., (1979), Nature, 282, 880.

Greenall, R.J. (1982), Ph.D. Thesis, University of Keele.

Gupta, G., Bansal, M. and Sasisekharan, V., (1980), Int.J.Biol.Macromolecules, 2, 368.

Hamilton, L.D., Barclay, R.K., Wilkins, M.H.F., Brown, G.L., Wilson, H.R., Marvin, D.A., Ephrussi-Taylor, H. and Simmons, N.S., (1959), J.Biophys. Biochem.Cytol., 5, 367.

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

Ivanov, V.T. and Minyat, E.E., (1981), Nucl.Acid.Res., 9, 4783.

Inman, R.B. and Baldwin, R.L., (1962), J.Mol.Biol., 5, 172.

Klug, A., Crick, F.H.C. and Wyckoff, H.W., (1958), Acta.Cryst., 11, 199.

Klug, A., Jack, A., Viswamitra, M.A., Kennard, O., Shakked, Z. and Steitz, T.A., (1979), J.Mol.Biol., 131, 669.

Kropinski, A.M.B., Bose, R.J. and Warren, R.A.J., (1973), Biochemistry, 12, 151.

Kypr, J., Vorlickova, M., Budeskinsky, M. and Sklenar, V., (1981), 99, 1257.

Langridge, R., Seeds, W.E., Wilson, H.R., Hooper, C.W., Wilkins, M.H.F., and Hamilton, L.D., (1957), J.Biophys.Biochem.Cytol., 3, 767.

Langridge, R., Wilson, H.R., Hooper, C.W., Wilkins, M.H.F. and Hamilton, L.D., (1960a), J.Mol.Biol., 2, 19.

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.

Langridge, R., (1969), J.Cell.Physiol., 74, 1.

Laskowski, M.Sr. (1971), The Enzymes (Boyer, P.D. ed.), Academic Press, London.

Leslie, A.G.W., Arnott, S., Chandrasekharan, R. and Ratliff, R.L., (1980), J.Mol.Biol., 143, 49.

Lin, S.-Y. and Riggs, A.D., (1971), Biochem.Biophys.Res.Commun., 45, 1542.

Lin, S.-Y. and Ri-gs, A.D., (1972), Proc.Nat.Acad.Sci., U.S.A., 69, 2574.

Lehman, I.R. and Pratt, E.A., (1960), J.Biol.Chem., 235, 3254.

Lutter, L.C., (1977), J.Mol.Biol., 117, 53.

Mahendrasingam, A., Rhodes, N.J., Goodwin, D.C., Nave, C., Pigram, W.J., Fuller, W., Brahms, J. and Vergne, J., (1983a), Nature, 301, 535.

Mahendrasingam, A., Pigram, W.J., Fuller, W., Brahms, J. and Vergne, J., (1983b), J.Mol.Biol., 168, 897.

Marvin, D.A., Spencer, M., Wilkins, M.H.F. and Hamilton, L.D., (1961), J.Mol.Biol., 3, 547.

Massie, H.R. and Zimm, B.H., (1965), Proc.Nat.Acad.Sci., U.S.A., 54, 1641.

Mitsui, Y., Langridge, R., Shortle, B.E., Cantor, C.R., Grant, R.C., Kodama, M. and Wells, R.D., (1970), Nature, 288, 1166.

Mokulskii, M.A., Kapitonova, K.A. and Mokulskaya, T.D., (1972), Molec. Biol., 6, 716.

O'Brien, F.E.M., (1948), J.Sci.Inst., 25, 73.

Patel, D.J. and Tonelli, A.E., (1974), Biopolymers, 13, 1943.

Patel, D.J., Canuel, L.D., Pohl, F.M., (1979), Proc.Nat.Acad.Sci., 76, 2508.

Pigram, W.J., (1968), Ph.D. Thesis, University of London.

Pilet, J. and Brahms, J., (1972), Nature New Biol., 236, 99.

Pohl, F.M. and Jovin, T.M., (1972), J.Mol.Biol., 67, 375.

Pohl, F.M., Jovin, T.M., Baehr, W. and Holbrook, J.J., (1972), Proc.Nat. Acad.Sci., 66, 3805.

Pohl., F.M., Ranade, A. and Stockburger, M., (1973), Biochem.Biophys. Acta., 335, 85.

Pohl. F.M., (1976), Nature, 260, 365.

Premilat, S. and Albiser, G., (1975), J.Mol.Biol., 99, 27.

Ramachandran, G.N., (1960), Proc.Indian.Acad.Sci., A52, 240.

Ramaswamy, N., Bansal, M., Gupta, G. and Saskisekharan, V., (1982), Proc.Nat.Acad.Sci. U.S.A., 79, 6109.

Rhodes, N.J., Mahendrasingam, A., Pigram, W.J., Fuller, W., Brahms, J., Vergne, J. and Warren, A.J., (1982), Nature, 296, 267.

Rhodes, N.J., (1983), Ph.D. Thesis, University of Keele.

Richmond, T. and Steitz, T.A., (1976), J.Mol.Biol., 103, 25.

Riggs, A.D., Lin, S.-Y. and Wells, R.D., (1972), Proc.Nat.Acad.Sci., U.S.A., 69, 761.

Podley, G.A., Scobie, R.S., Bates, R.H.T. and Levitt, R.M., (1976), Proc.Nat.Acad.Sci. U.S.A., 73, 2959.

Sasisekharan, V., Pattabiraman, N. and Gupta, G., (1978), Proc.Nat.Acad.Sci., U.S.A., 75, 4092.

Sasisekharan, V. and Brahmachari, S.K. (1981), Curr.Sci., 50, 10.

Scheffler, I.E., Elson, E.L. and Baldwin, R.L., (1968), J.Mol.Biol., 36, 291.

Selsing, E., Arnott, S. and Ratliff, R.L., (1975), J.Mol.Biol., 98, 243.

Shakked, Z., Rabinovich, D., Cruse, W.B.T, Egert, E., Kennard, O., Sala, G., Salisbury, S.A. and Viswamitra, M.A., (1981), Proc.Roc.Soc.Sev.B., 213, 479.

Shakked, Z., Rabinovich, D., Kennard, O., Cruse, W.B.T., Salisbury, S.A. and Viswamitra, M.A., (1983), J.Mol.Biol., 166, 183.

Sherwood, D., (1976), Crystals, X-rays and Proteins, Longman Group Ltd.

Shindo, H., Simpson, R. and Cohen, J.S., (1979), J.Biol.Chem., 254, 8125.

Shindo, H. and Zimmerman, S.B., (1980), Nature, 283, 690.

Shindo, H., Wooten, Jan B. and Zimmerman, S.B., (1981), Biochemistry, 20, 745.

Spencer, M., (1959), Acta Cryst., 12, 59.

Sundaralingam,M. and Jenson, L.H., (1965), J.Mol.Biol., 13, 94.

Sundaralingam,M., (1969), Biopolymers, 7, 821.

Thamann, T.J., Lord, R.C., Wang, A.H.J. and Rich, A., (1981), Nucl.Acid.Res., 9, 5443.

Vand, V., Eiland, P.F. and Pepinsky, R., (1957), Acta.Cryst., 10, 303.

Viswamitra, M.A., Kennard, O., Shakked, Z., Jones, P.G., Sheldrick, G.M., Salisbury, S. and Falvello, L., (1978), Curr.Sci. 47, 289.

Viswamitra, M.A., Kennard, O., Jones, P.G., Sheldrick, G.M., Salisbury, S., Falvello, L. and Shakked, Z., (1978), Nature, 273, 687.

Vorlickova, M., Kypr, J., Kelinwachter, V. and Palecek, E., (1980), Nucl.Acid.Res., 8, 3965.

Vorlickova, M., Sedlacek, P., Kypr, J. and Sponar, J., (1982), NuclAcid.Res., 10, 6969.

Vorlickova, M., Kypr, J., Stokrova, S. and Sponar, J., (1982), Nucl.Acid.Res., 10, 1071.

Vorlickova, M., Skelnar, V. and Kypr, J., (1983), J.Mol.Biol., 166, 85.

Wang, A.H.J., Quigley, G.L., Kolpak, F.J., Crawford, J., van Boom, J.H., van der Marel, G. and Rich, A., (1979), Nature, 282, 680.

Wang, A.H.J., Quigley, G.L., Kolpak, R.J., van der Marel, G., van Boom, J.H. and Rich. A., (1981), Science, 211, 171.

Wang, A.H.J., Fujii, S., van Boom, J.H. and Rich, A., (1982a), Proc.Nat. Acad.Sci., U.S.A., 79, 3968.

Wang, A.H.J., Fujii, S., van Boom, J.H. and van der Marel, G., van Boeckel, S.A.A. and Rich., A., (1982b), Nature, 299, 601.

Watson, J.D. and Crick, F.H.C., (1953), Nature, 171, 737.

Wilkins, M.H.F., Stokes, A.R. and Wilson, H.R., (1953), Nature, 171, 738.

Wing, R., Drew, H., Takano, T., Broka, C., Tanaka, S., Itakura, K. and Dickerson, R.E., (1980), Nature, 287, 755.

Wrinch, D., (1950), Acta.Cryst., 3, 475.

Wyatt, G.R. and Cohen, S.S., (1953), Biochem.J., 55, 774.

Zimmerman, S.B. and Pheiffer, B.H., (1980), J.Mol.Biol., 142, 315.

Zimmer, C., Tymen, S., Marck, S. and Guschlbauer, W., (1982), Nucl.Acid.Res., 10, 1081.

TITLE X-RAY FIBRE DIFFRACTION STUDIES ON THE POLYMORPHISM OF DNA AND ITS SYNTHETIC ANALOGUES.

AUTHOR Arumugam MAHENDRASINGAM

DEGREE Ph.D

AWARDING BODY Keele University

DATE 1983

THESIS NUMBER DX187422