4 Progress in Molecular and Subcellular Biology

With Contributions by
St. Bram · P. Chandra · U. Ebener · H. Laube
W. Lotz · D. Oesterhelt · R. A. Olsson
R. E. Patterson · L. K. Steel · M. Sussman
G. Will · H. J. Witmer · M. Woltersdorf

With 88 Figures

Springer-Verlag Berlin Heidelberg New York 1976

ISBN 3-540-07487-2 Springer-Verlag Berlin Heidelberg New York
ISBN 0-387-07487-2 Springer-Verlag New York Heidelberg Berlin

Printed in Germany.

Offsetprinting and bookbinding: Konrad Triltsch, Graphischer Betrieb, Würzburg.

Twenty Five Years of Molecular Biology

Fred E. Hahn

Molecular Biology has apparently come of age. The name was introduced into the modern scientific literature a quarter of a century ago (ASTBURY, 1950). The Journal of Molecular Biology began its publication in 1959, followed by Molecular Pharmacology, Molekularnaia Biologiia, Molecular Photochemistry, Journal of Molecular and Cellular Biochemistry, Clinical Science and Molecular Medicine, Molecular and General Genetics, Molecular Etiology and probably other periodicals dedicated to molecular aspects of the life sciences whose titles may have escaped our cursory search. If one adds to this the numerous sets of monographs and review collections, including this series, the source and secondary literatures in molecular biology comprise, by now, a sizeable library.

Molecular biology is, no doubt, a chemical science. However, during the past two decades it has become fashionable to substitute for the traditional adjective, biochemical, the more alluring "molecular". One author (KOSOWER, 1962) even published a treatise, entitled Molecular Biochemistry whose redundancy in title was commented upon by writing, "In the past few years, a new research area has emerged from the application of the physical-organic approach to the problem in chemical transformation found in biochemistry. We choose to call this area molecular biochemistry, and delimit it as the study of the detailed chemical mechanisms of the chemical transformations in biology, usually as they are described in biochemistry".

Hence, underneath the molecular vogue in communicating biochemical research, there lives an expectation that the mechanics of life can be resolved into their component processes and molecular entities with the distant goal of being able to reassemble this resolved totality conceptually or even physically into a working resemblance of the original living object. "Having pulled apart the chemical continuity of the living organism, we are challenged to reintegrate the scattered pieces into a whole" (LIPMANN, 1971).

In one of his satirical writings, CHARGAFF (1963) recognized clearly that the term molecular biology, in superceding biochemistry, has a programmatic connotation and signals the resolutionistic position in the life sciences. However, resolutionism has not halted at the molecular level of biological organization. We point to the existence of quantum biology whose early entries into the literature, after SCHRÖDINGER's prophesy (1945), were SZENT-GYÖRGYI's Introduction to a Submolecular Biology (1960) and PULLMAN and PULLMAN's Quantum Biochemistry (1962). Among the ideas introduced were that (1) biopolymers possess certain

solid-state physical properties which cannot be derived or extrapolated from the physical properties of their component molecules so that the electronic states of the polymers depart from those of the individual subunits; and (2) the consideration of biological functions of molecules must penetrate below the classical aspects of structural organic chemistry to consider "the essential importance of molecular systems with mobile electrons, and therefore of electronic delocalization" (PULLMAN and PULLMAN, 1962).

We recognize that resolutionistic, i.e. molecular and submolecular biology with its built-in ambition of logical and/or physical reassembly of components into a biologically functional whole is related to positivism. Ultimately it strives, therefore, to "explain" the phenomenon of the living state by uncovering its underlying mechanics. LIPMANN (1971) writes about "molecular technology". Molecular biology, hence, is a contemporary form of mechanistic, i.e. deterministic biology. It has derived its impetus and philosophical justification from the successful development of a physical theory of heredity, i.e. from the resolution of classical genetics into molecular genetics: its symbol has become the genetic code.

The ascendancy of molecular biology has, therefore, occurred for deeper reasons than the mere opportunism of its adepts, although some of its adversaries, e.g. Erwin CHARGAFF (1963, 1968) keep pointing, not entirely without justification, to opportunistic tendencies at work. It would be surprising if the intellectual hybris with which molecular biologists deal with the objects of their study would not, to some extent, find its counterpart in the manner(s) in which they deal with their fellow scientists and with science as a society.

An Editorial for this Progress series, finally, should at least ask the question if the attainment of the ultimate conceptual aim of molecular biology, *viz.*, resolution and reassembly of a living "system" (the term itself being mechanistic), can explain such a system. Physicists are more cautious and modest. Mathematical formulations which dualistically describe wave and particle functions of light and matter are accepted side by side not as "explanations" but as models. Modern physics does not aspire to explain the physical universe but to develop a unified mathematical imagery which deals with phenomena not in the positivistic sense of what they are but in the functional sense of how they interact and behave. The relationship of such models to physical reality lies in their predictive value. Molecular biology will truly come of age when it begins to examine its underlying philosophical premises.

References

ASTBURY, W.T.: Adventures in molecular biology. Harvey Lect. 46, 3 (1950-51).

CHARGAFF, E.: Amphisbaena, In: Essays on Nucleic Acids. Amsterdam: Elsevier 1963.
CHARGAFF, E.: A quick climb up Mount Olympus. Science 159, 1448 (1968).
KOSOWER, E.M.: Molecular Biochemistry. New York: McGraw Hill 1962.
LIPMANN, F.: Wanderings of a Biochemist. New York: Wiley Interscience 1971.
PULLMAN, B., PULLMAN, A.: Quantum Biochemistry. New York: Interscience 1963.
SCHRÖDINGER, E.: What is Life? New York: Cambridge Univ. Press 1945.
SZENT-GYÖRGYI, A.: Introduction to a Submolecular Biology. New York: Academic Press 1960.

List of Contributors

STANLEY BRAM, Dépertment de Biologie Moléculaire, Institut Pasteur, F-75000 Paris 15

PRAKASH CHANDRA, Gustav-Embden-Zentrum der Biologischen Chemie, 6 Frankfurt am Main 70, Theodor-Stern-Kai 7

U. EBENER, Gustav-Embden-Zentrum der Biologischen Chemie, 6 Frankfurt am Main 70, Theodor-Stern-Kai 7

H. LAUBE, Gustav-Embden-Zentrum der Biologischen Chemie, 6 Frankfurt am Main 70, Theodor-Stern-Kai 7

WOLFGANG LOTZ, Institute for Enzyme Research, University of Wisconsin, Madison, Wisconsin 53706, USA

DIETER OESTERHELT, Institut für Biochemie der Universität Würzburg, 87 Würzburg, Röntgenring 11

RAY A. OLSSON, COL MC, Division of Medicine, Walter Reed Army Instiute of Research, Walter Reed Army Medical Center, Washington D.C. 20012, USA

RANDOLPH E. PATTERSON, COL MC, Division of Medicine, Walter Reed Army Institute of Research, Walter Reed Army Medical Center, Washington D.C. 20012, USA

LINDA K. STEEL, Gustav-Embden-Zentrum der Biologischen Chemie, 6 Frankfurt am Main 70, Theodor-Stern-Kai 7

MAURICE SUSSMAN, Hebrew University, Section of Developmental and Molecular Biology, Migrash Harussin, Jerusalem, Israel

G. WILL, Gustav-Embden-Zentrum der Biologischen Chemie, 6 Frankfurt am Main 70, Theodor-Stern-Kai 7

HEMAN J. WITMER, Department of Biological Sciences, University of Illinois of Chicago Circle, P.O.Box 4348, Chicago, Illinois 60680, USA

M. WOLTERSDORF, Gustav-Embden-Zentrum der Biologischen Chemie, 6 Frankfurt am Main 70, Theodor-Stern-Kai 7

Contents

The Polymorphism of DNA

Stanley Bram

I. Introduction

The salts of deoxyribonucleic acid (DNA) are in many ways living molecules whose morphological characteristics are dependent on their genetic information and on their past and present environment. However, until very recently, DNA was not considered to possess variable structures – in short, it was not thought to be polymorphic. There were three so-called canonical forms having an almost inorganic regularity. Three invariant forms are not consistent with the diversity of DNA functions or its specific recognition. As recently pointed out by CRICK (1971), double-stranded DNA in the invariant classical states cannot provide enough detail for the very specific recognition of DNA by proteins. Recognition schemes based upon the fixation of mono or divalent ions upon crystals of dinucleotides such as proposed by KIM et al. (1973) are unlikely in solution where ions are not site-bound. Yet we know that a double-stranded stretch of 20 to 50 base pairs of the lac operator is chosen without error in an *E. coli* chromosome of 3 million nucleotide pairs (RIGGS, SUZUKI and BOURGEOIS, 1970).

Nearly all the details of DNA structure have come from X-ray fiber diffraction data on oriented fibers which were obtained ten or twenty years ago. Then and now, structures were defined by their diffraction patterns and not conversely. The fiber diagrams for B-DNA, for example, have remained essentially the same for twenty years but the details of the structure attributed to them have appreciably changed four or five times since the days of WATSON and CRICK (1953). FRANKLIN and GOSLING (1953) with Na DNA, found the so-called A fiber pattern at lower relative humidity and the B pattern at higher relative humidity (rh). Lithium DNA did not yield an A diagram but instead showed either crystalline B or another type or diagram – type C (MARVIN et al., 1961). In 1959 a study was published of the low humidity A diagrams of a variety of DNAs (HAMILTON et al., 1959) and since all of the DNAs gave indistinguishable A patterns, it was concluded that all DNA structures were independent of the base composition. This dogma was challenged only when it was found that the X-ray patterns of various DNAs in solution were indeed dependent upon base content (BRAM, 1971b). It was suggested in this study that the transition to the A form might also depend upon base content and this was then found to be the case (PILET and BRAHMS, 1972). The diffraction experiments in solution lead to fiber diffraction studies (BRAM, 1972a; BRAM and TOUGARD, 1972) and the discovery of four new and distinct X-ray diffraction patterns. Furthermore, it was found that the transitions between the various DNA conformations are far more interesting and com-

plex than had been thought (AZOULAY and BRAM, 1973) (BRAM and BAUDY, 1974).

Other physical chemical techniques, such as infra-red and Raman spectroscopy (PILET and BRAHMS, 1972) (ERFURTH, KISER and PETICOLAS, 1972) are now beginning to play important roles in DNA structural determinations and will certainly reduce the many man-years now required for obtaining good structural coordinates. However, this review will be restricted mainly to recent X-ray studies of DNA. It is worth mentioning that infra-red studies have only recently been interpreted to show that the B-conformations depend on the base composition (PILET, BLICHARSKI and BRAHMS, 1975).

II. A Summary of the Principles of DNA X-Ray Diffraction

Oriented fibers of DNA are obtained by exploiting the high viscosity of DNA gels. When tension is applied to a wet gel the long molecules align parallel to each other in the direction of the tension. This tension can then either be maintained or increased while drying. The rate of drying, which we shall see is a very important parameter, can be controlled by adjusting the temperature and the relative humidity during drying. For example, a gel can be dried to a fiber in half an hour over 44 % rh, at 37°C, while a day may be required to dry it at 4°C over 79 % rh. The fiber is then mounted in a collimated beam of X-rays and the intensity diffracted is recorded on a film behind the fiber. The direction on the film parallel to the long axis of the fiber is called the meridian, and the perpendicular to this direction, the equator. One usually fixes the water content of the fibers by carrying out the experiments in helium which has been equilibrated with standard saturated salt solutions. Specimens for solution or gel X-ray scattering can be obtained by mixing solvent with DNA fibers; they are then irradiated in sealed thin-walled glass capillary tubes (sometimes after dialysis *vs.* the solvent).

The diffraction from a discontinuous helix such as DNA has the following properties: (a) it is confined to layer lines separated by a distance proportional to 1/pitch; (b) on the lower layers the diffraction is described by a Bessel function of the same order as the layer number; (c) the pattern has a cross-shaped form in its center; (d) there is a strong meridional intensity spot at a distance corresponding to 1/inter-unit separation; (e) for the double helix the intensity diffracted is modified by a fringe function corresponding to the relative position of the two strands. The reader can obtain an estimate of the pitch and interbase pair separation by taking the ratio between the distance to the first real meridional reflection on the photographs (which corresponds to 3.4Å in the B and 3.3 Å in the C diffraction) and the distance separating the layer lines. This ratio is 11.3 for a helix of 10 base pairs per turn and is about 9 for a helix of eight bases per repeat. The two B diagrams of Fig. 2a and b show most of the helical diffraction features, but the classical B distribution of weak intensity on the first and third layers and strong on the second is observed only with DNA

relatively rich in GC. This alteration of weak and strong intensity on the inner layer lines was one of the major reasons why WATSON and CRICK (1953) concluded that DNA must have a double helical structure whose chains were separated by a rotation of about 180°.

In the A-DNA fibers, the molecules are always packed into a crystalline lattice; consequently the intensity on the various layer lines is discontinuously sampled, making the helical nature of the diffraction less evident.

III. Generalities of D A Transitions

In vivo, many other molecules are present about DNA, especially large amounts of water, proteins and mono and divalent ions. DNA is in a complex of either nucleoprotein or polyamines. These various proteins, macro-ions and polyamines will exclude water from the DNA surface to a degree dependent on their interaction with DNA. Consequently, it follows that much of the DNA will be more or less dehydrated due to steric exclusion by proteins. Those regions which remain relatively hydrated will exist in a B-like conformation, while those regions of reduced water content will undergo structural transitions in response to the local environment.

Upon drying pure sodium or potassium DNA, a B to A transition occurs at less than 6 - 10 H_2O per nucleotide. In the more than 20 years since the discovery of the A form (RILEY and OSTER, 1951), all X-ray experiments have found it co-existent with a crystalline lattice. Consequently it was suggested that intramolecular lattice interactions are responsible for this configuration of DNA (BRAM and BAUDY, 1974). To test this suggestion, further experiments were carried out on DNA in 80 % ethanol and isopropanol. Here also, only crystalline A patterns were discerned (see Fig. 1). Furthermore, spontanous orientations was sometimes observed (BRAM and BAUDY, 1974) indicating that lattice formation has a preferred direction.

Previously, it seemed possible that DNA could adopt the A form in local regions of a chromosome when provided with the correct environment by bound proteins. However, a requirement for crystallization would eliminate this possibility. Furthermore, it has been known for some time that high concentrations of histones (WILKINS, ZUBAY and WILSON, 1959) and polyaminoacids (ZUBAY, WILKINS and BLOUT, 1961) block a transition to the A form. Probably the strongest evidence against the existence of the presence of A-DNA *in vivo* comes from the recent finding in our laboratory that certain histone proteins will block the A transition in fibers at concentrations of one to two percent, which represents about one protein per thousand base pairs. Such small amounts of protein probably act as an impurity, which prevents crystallization. It is therefore important to consider what is known about the structure and transitions of DNA when the formation of the A state is inhibited.

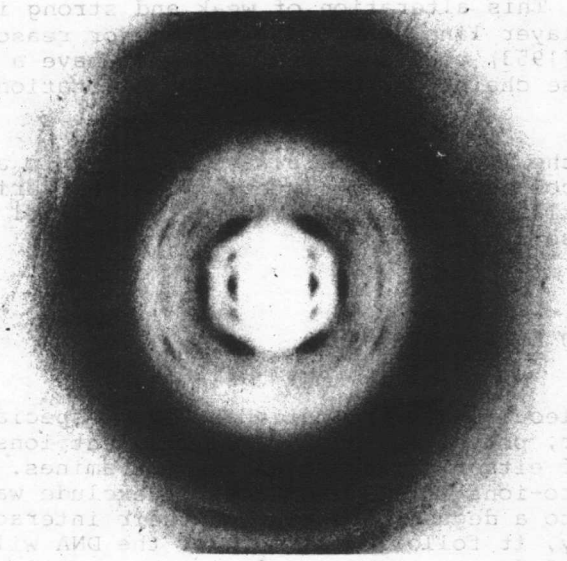

Fig. 1. An X-ray fiber diagram of a sodium calf thymus DNA fiber in 90 % isopropanol by volume. The fiber was photographed with a toorridal mirror diffractometer after equilibration with 98 % relative humidity and subsequent immersion in a large excess of 90 % isopropanol containing 5 mM NaCl. The fiber diagram is identical to those of A fibers in air except for the ring at 4 Å from isopropanol (compare to the A diagram of Fig. 2). The diffraction patterns of DNA in ethanol, isopropanol or in air equilibrated with water have always been crystalline

IV. B-like Structures

The B configuration is probably the most prevalent structure *in vivo* . Yet there exist for DNAs of various base content a large family of distinct B-like structures which are related, in that they all have about ten base pairs per turn of 34 A.

Although it had been inferred for some time, the first conclusive evidence that DNA in solution had a B-like configuration came from small- and wide-angle X-ray scattering experiments (BRAM and BEEMAN, 1971). The mass per unit length obtained was exactly that expected from B-DNA, and the wide-angle scattering pattern was much more similar to that calculated for a B structure than an A. Nevertheless, the wide-angle scattering patterns were not the same as those calculated with the atomic coordinates based upon LiDNA fiber diffraction diagrams at 66 % rh (BRAM, 1971a). These results have more recently been repeated and confirmed by MATHAIS et al. (1973). The spacing of the first scattering maxima was observed to be at 13.6 A instead of at 12.6 A as calculated with the B coordinates. Therefore, it had to be concluded

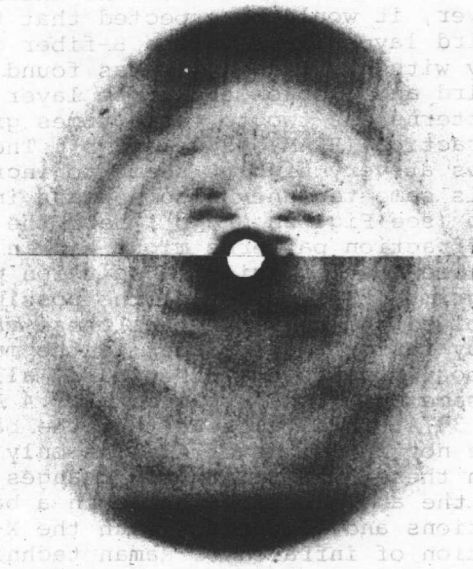

Fig. 2 a and b. Two sodium B diffraction diagrams at 44 % rh, after drying under tension at 37°C, 44 % rh. (a) *C. perfringens* DNA (66 % AT). Note the strong intensity on the first three layer lines. (b) *S. lutea* DNA (29 % AT). Note the very weak intensity on the first and third layer lines. The NaCl content in the fiber was 3 % of the weight of the anhydrous DNA

that either the pitch or the radius of the DNA was larger than that given by the atomic coordinates. Since I did not want to dispute the crystallography at that time, I accepted the published coordinates for a base pair, and suggested that the best explanation involved a variation in the pitch. But now after examining many very high-humidity B fiber patterns and finding that all have the same pitch of 34 A (BRAM, 1973a and b), it seems much more likely that the coordinates were incorrect. Recently rather different coordinates have been published (ARNOTT and HUKINS, 1974) and the scattering calculated with the new values agrees better with experiment.

In 2 M Li or NaCl, where the hydration about the DNA is lowered, the pitch or the effective radius of the DNA is about 5 % smaller than in dilute salt solution (BRAM, 1971 b). Likewise, in chromatin, the wide-angle scattering patterns showed that the structure is more compact than in pure DNA solutions, but here a decrease in both the effective radius and the pitch would best agree with the data. In any case, it is unwise to attribute the changes in an experimental measurement, whether it be X-ray scattering or especially circular dichroism, to a detailed structural transition (for instance to a C form) expecially when the structural parameters may not be directly related to the measured values.

It has been shown that solutions of DNAs with different base composition have different wide-angle X-ray scattering patterns and as a direct consequence must have different structures (BRAM, 1971b). The ratio of the X-ray intensity maximum near 10 A to that observed near 13 A increased almost two-fold over the range of 31 to 100 % AT. Since this ratio of intensity (10/13 A) is

directly related to the ratio of intensity between the third layer line and the second layer, it would be expected that the relative intensity of the third layer lines in the B-fiber diagrams would increase strongly with AT. Indeed, it was found that the intensity on the third and also on the first layer line of the very AT-rich DNA B patterns was two to three times greater than in the GC-rich DNA diffraction (BRAM, 1973a and b). The intensity on the odd layer lines at very high rh seems to increase semicontinuously with AT. This same tendancy also appears in the B diagrams observed at low rh (see Fig. 2a and b); here the crystalline nature of the diffraction patterns might aid in structural determinations. Since the intensity in solution parallels that of the fiber diagrams, this rules out the possibility that lattice effects are involved and requires that these variations in layer-line intensity reflect differences in the molecular structure with base composition. On the other hand, all of the fiber diagrams have the same meridional spacing of 3.4 A and layer-line separation of 34 A. This tells us that the base-pair separation and pitch are not functions of AT. The only metamorphoses consistent with these constraints are changes in either the base-pair tilt or the angle between bases in a base-pair. Detailed model calculations and comparison with the X-ray data (or perhaps the utilization of infrared or Raman techniques) will be necessary to decide between these two possibilities. In any case, the differences in the X-ray patterns result from either (1) a structural variation over large regions or clusters of base-pairs, or (2) a difference in the structure of each individual GC or AT pair or grouping of a few nucleotides. The apparently progressive rise of the odd layer-line intensity with AT, which is probably related to the monotonic decrease in density of DNA with AT on both CsCl (SCHILDKRAUT, MARMUR, and DOTY, 1962) and NaI gradients (ANET and STRAGER, 1969), would be more consistent with (2). The second mechanism also finds support in recent gel electrophoresis studies on various DNAs which show that the mobility increases continuously with AT; these studies were also interpreted to imply that the cross-section morphology of DNA is a function of the base content (ZEIGER et al., 1972). This latter mechanism provides a direct answer to the question of how DNA is recognized by proteins; each base-pair has a distinct and recognizable configuration. This would imply that DNA is almost infinitely polymorphic in the B state.

B- X-ray diffraction diagrams are also observed in fibers at lower water content (see Fig. 2). With lithium DNA it is a stable state (LANGRIDGE et al., 1956) but with sodium DNA, where it has only recently been shown to exist (BRAM, 1972a), it is a metastable state (AZOULAY and BRAM, 1973) (BRAM and BAUDY, 1974). Low-humidity sodium B states can be obtained by quickly dehydrating a gel containing 3 % or more NaCl to 66 % or less relative humidity. Depending on the fiber-pulling conditions, the fibers contain variable amounts of A-DNA. It is still an open question as to whether or not the low humidity conformations are equal to those at higher water contents.

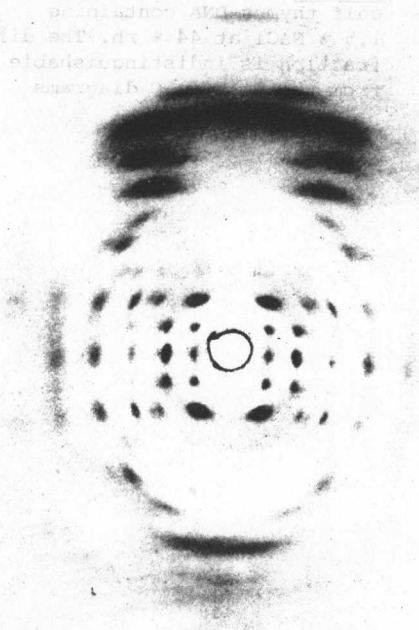

Fig. 3. An A-diffraction pattern from calf thymus DNA at 79 % rh. This fiber was first quickly dried from a wet gel under tension and exhibited a C-diffraction pattern at 44 % rh. The relative humidity was then raised to 79 % and photographed after two hours. A C→A transformation has occurred. The fiber was tilted by about 20° from the vertical

V. The A Form

If one considers form to mean a particular unique structure, then the A configuration is the only real form of DNA. The A-diffraction diagram is the only one that has been found to be independent of base composition (HAMILTON et al., 1959) and salt content (COOPER and HAMILTON, 1966). It is the equilibrium form of all pure sodium DNA and also probably of potassium DNA at lower water contents (FRANKLIN and GOSLING, 1953). However, some synthetic (LANGRIDGE, 1969) and some very AT-rich DNAs (BRAM, 1972) have not adopted the A form under the limited range of conditions so far studied. The B to A transition is favored by higher GC content (PILET and BRAHMS, 1972) and lower excess salt contents (COOPER and HAMILTON, 1966). Although it is generally held that low water content is the most important factor, fibers of *M. lysodeikticus* DNA which is very rich in GC have remained in the A form at relative humidities above 98 %, and very high NaCl contents (above about 8 % NaCl) seem to block the A transition of calf-thymus DNA at all humidities. Both high AT and salt contents and also high water contents probably inhibit the A-transition by making crystallization less favorable. As previously mentioned, the A form of DNA has only been found in a crystalline lattice. The A-fiber diagram is characterized by a layer-line separation of 28 A and an intensity distribution of weak intensity on the inner part of the first layer line, a very strong second layer

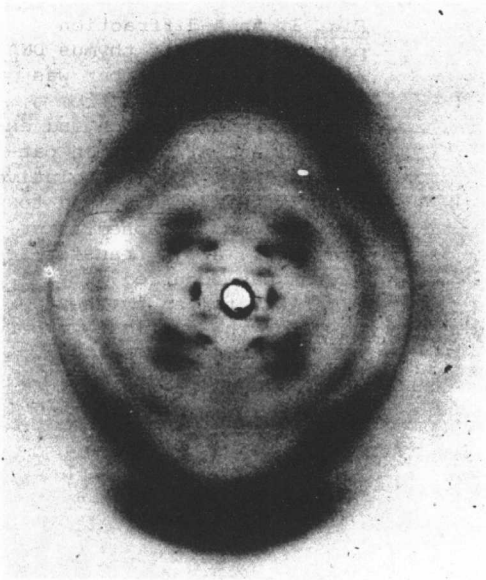

Fig. 4. A sodium-C diagram from calf thymus DNA containing 1,5 % NaCl at 44 % rh. The diffraction is indistinguishable from the lithium-C diagrams

line, and almost zero intensity on the central portion of the third layer (Fig. 3 is an example of an A-X-ray diffraction diagram). The A conformation features 11 base-pairs per turn, each of which is tilted 20° from the perpendicular to the helix axis and is separated from its nearest neighbors by 2.6 A. The base-pairs are also displaced off the helix axis by 4 A and the sugar ring is puckered in the so-called (C2-endo) conformation (ARNOTT et al., 1968).

Crystallographers have shown that the structure of the hybrid of RNA and DNA is very similar to that of A-DNA (MILMAN, CHAMBERLIN, and LANGRIDGE, 1967) and have suggested that this is consequently the form of DNA active in transcription (ARNOTT et al., 1968) where hybridization of RNA to DNA is the most important step. Also, the lack of polymorphism in the A form would make it a more readily used substrate for polymerase enzymes. Since the transition to the A state may be blocked in the nucleus by divalent ions or histone-like proteins, a catalyst may be required to drive the DNA into the A state. A likely candidate for this catalyst is RNA itself, which forms only an A-like state with DNA.

VI. C-like Structures

Of the three classical states, the C state is the most polymorphic and least well-defined. Calf thymus lithium DNA adopts structures