*SOME RELATIONS BETWEEN DNA AND RNA**

By Alexander Rich† and J. D. Watson

GATES AND CRELLIN LABORATORIES OF CHEMISTRY AND KERCKHOFF LABORATORIES OF BIOLOGY, CALIFORNIA INSTITUTE OF TECHNOLOGY

Communicated by Linus Pauling, May 7, 1954

There are two nucleic acids: DNA (desoxyribonucleic acid) and RNA (ribonucleic acid). Both are found in all complex organisms, whether plant or animal and, together with the proteins, are thought to be among the most important biological substances in the cell. Like the proteins, they are macromolecular and consequently capable of great specificity. DNA is considered to function as a genetic substance and may, in fact, form the specific element of the chromosome.[1] About the functions of RNA, we possess little definite information. It has been implicated in protein synthesis, but only indirectly. The really interesting thing about both the nucleic acids is that we know very little about how they function chemically in a cell.

This uncertainty arises in part from the complex structure of the nucleic acids. Both are polymeric compounds formed from the linear aggregations of the nucleotides, their fundamental subunit. Four main types of nucleotides exist, and they are believed to be randomly situated along the chain. A polynucleotide chain may contain thousands of nucleotides, and thus an almost infinite number of specific nucleic acids is possible. However, despite the indeterminacy of the nucleotide sequence, a certain degree of regularity exists.

In the first place, the union between successive nucleotides is always the same. The main internucleotide linkage in both DNA and RNA is 3′,5′-phosphate diester linkage.[2, 3] Nucleic acids are thus long-chain molecules in which the sugar-phosphate backbone is highly regular, which derive their individuality from the order in which specific purine and pyrimidine bases are attached to the backbone.

The second level of regularity arises from the spatial configuration of the backbone. We learn from X-ray diffraction studies that all DNA's investigated, no matter what their source, have a similar structure.[4] The same is true of all RNA's.[5]

Up to now we have had success in understanding only one of these two structures. DNA appears to be a two-stranded helical structure in which the two polynucleotide chains are joined together by hydrogen bonds between the purine and pyrimidine bases.[4, 6, 7] The hydrogen-bonding arrangement is both regular and specific and always unites the bases together in pairs. Adenine is specifically paired with thymine and guanine with cytosine. The resulting structure is unique, for the specific pairing produces a complementary relationship between the sequence of bases on the two chains. If we know the sequence on one chain, we will automatically have the sequence on the opposing chain. In addition to producing a structure which seems to fit the X-ray data, specific pairing furthermore provides a simple explanation for the analytical observation that, in all DNA's examined, the molar content of adenine is very nearly equal to that of thymine, and that of guanine to that of cytosine.[8]

The most attractive feature of the two-stranded complementary helix is the fact that it suggests an answer to the question of how DNA can replicate itself exactly, a function it must possess if it is a genetic material.[9] The complementary structure

fits this requirement neatly if we make the assumption that one strand can serve as a template for the formation of its complement. We visualize, then, a mechanism involving initial separation of the two strands, with each of the separated strands serving as a template for its complement—the whole process occurring in zipper-like fashion. This method of replication is likely to be very exact, as the necessity for specific pairing is absolute, and misformed pairs will not fit into the structure.

We should also like to know how DNA influences the physiology of the cell. An obvious possibility is that it controls, either directly or indirectly, the synthesis of specific proteins.

Several objections may be cited against a direct role of DNA in protein synthesis. In the first place, we can see no obvious complementary relationship between the shapes of the individual amino acids and the surface of the DNA molecule. Second, protein synthesis appears to proceed normally for some time in cells which lack a nucleus. Both Actabularia[10] and the reticulocyte stage of red blood cells[11] are examples in which protein synthesis occurs in the absence of nuclear DNA. In addition, the sites of greatest protein synthesis as measured by isotope incorporation are the microsomes and the mitochondria, particles in which DNA is completely absent. It would thus appear that protein synthesis often occurs in the absence of DNA.

We consider it plausible to suppose, like many others,[12] a connection between RNA and protein synthesis. Under such a scheme, DNA could control RNA synthesis, with RNA responsible for protein synthesis. Several considerations lead us to favor this hypothesis. First, DNA is lacking entirely in some plant viruses and is replaced by RNA.[13] Second, RNA is found in large amounts in both the microsomes and mitochondria,[14] the particles which are the sites of high protein synthesis.

We shall not be able to check a structural relationship between RNA and protein synthesis or between RNA and DNA until we know the structure of RNA. This compound appears more complex than DNA in several fundamental ways. The chemical formula is not yet established with certainty. While the main internucleotide linkage is similar to DNA (a 3′,5′-phosphate diester linkage), the possibility is still open that the structure is branched.[2] If such branches exist, they most likely arise at the hydroxyl group in the 2-position of a ribose residue in the main chain. Such a branch could not exist in DNA, owing to the absence of this hydroxyl group. Branching of this type has been suggested by many workers,[15, 16] although there is no evidence which is convincing. This problem remains a key one for the chemist to resolve.

The analytical composition of the bases in RNA also seems more complex than that in DNA. The experimental results shown in Table 1 indicate the probable existence of two classes of RNA. The first class is represented by RNA's derived from the plant viruses and appears to indicate a fairly random distribution of the various bases. No evidence for the existence of the specific pairs of bases is provided. On the contrary, in RNA's of the second type the molar amount of adenine is approximately equal to that of uracil (for pairing purposes, uracil can be considered equivalent to thymine), and the amount of guanine tends to equal that of cytosine. This second class includes many RNA's from sources other than plant

viruses. A few exceptions to the 1:1 ratio are found, but it is not certain that they are significant. It is obvious that in many RNA's the specific pairs are found, and this must be significant. One possibility is the existence of two types of RNA structure, in only one of which are the base pairs utilized. This, however, cannot be true, since, as mentioned earlier, RNA of all sources produces the same X-ray pattern. A simple interpretation of the analytical data does not appear possible.

Until recently there have existed no satisfactory X-ray diffraction data on RNA. We are therefore beginning to obtain X-ray data from various sources of RNA. We have found that it is possible to draw fibers from RNA preparations and to obtain X-ray diffraction patterns from them.[5]

These photographs do not show the high degree of fibrous orientation which can be obtained from DNA. We do not know whether this is owing to partial degradation of our specimens or whether it is because the molecule is intrinsically less

TABLE 1

PURINE AND PYRIMIDINE COMPOSITION OF VARIOUS RNA'S

Source	Adenine*	Uracil	Guanine	Cytosine	Analysis by†
Rabbit liver:					
Mitochondria	0.193	0.199	0.308	0.302	Crosbie, *et al.* (1953)[17]
Microsomes	0.195	0.199	0.305	0.300	
Cell sap	0.193	0.197	0.298	0.296	
Whole cytoplasm	0.192	0.213	0.304	0.292	
Calf brain	0.216	0.206	0.318	0.260	Deluca, *et al.* (1953)[18]
Calf thymus	0.162	0.157	0.352	0.330	Cohn and Volkin (1953)[15]
Carp nucleotropomyosin	0.166	0.178	0.348	0.318	Hamoir (1952)[19]
Sea urchin oöcytes	0.200	0.220	0.307	0.298	Vincent (1952)[20]
Carcinoma of liver (human)	0.098	0.071	0.407	0.426	Chargaff, *et al.* (1950)[21]
Fowl carcinoma	0.140	0.120	0.460	0.280	Harris, *et al.* (1950)[22]
Tobacco mosaic virus	0.296	0.264	0.251	0.186	Knight (1952)[23]
Cucumber mosaic virus	0.258	0.301	0.255	0.187	Knight (1952)[23]
Potato virus X	0.342	0.212	0.215	0.235	Knight and Dorner (1953)[24]
Tobacco necrosis virus	0.278	0.257	0.245	0.225	Markham (1953)[25]
Tomato bushy stunt	0.248	0.252	0.280	0.219	Markham and Smith (1951)[26]
Southern bean mosaic	0.258	0.252	0.260	0.232	Knight and Dorner (1953)[24]
Turnip yellow virus	0.227	0.222	0.172	0.382	Markham and Smith (1951)[26]

* All ratios have been adjusted to total 100 per cent.
† Superior reference numbers refer to notes at end of article.

regular and may not pack well. One of our RNA fiber photographs is illustrated in Figure 1. For comparative purposes, Figure 2 shows X-ray photographs of the crystalline and paracrystalline forms of DNA. Despite the lack of resolution in the RNA photograph, several points of interest emerge: (1) The pattern is dominated by strong reflections on or near the meridian at 3.3 and at 4.0 A. This most likely results from a stacking of purine and pyrimidine bases approximately on top of each other, arranged roughly perpendicular to the fiber axis. This interpretation is supported by the negative birefringence of the fibers.[5] A similar arrangement is found in the DNA structure.[4, 7] (2) The repeat distance along the fiber axis is 25–28 A. This can be seen in some photographs, where the 5.3 A equatorial reflection shows a layer-line periodicity of intensity in a direction parallel to the fiber axis. There is also a subrepeat of half this distance, as is indicated by the strong near-meridional reflections on the second layer line. This feature is also present in the DNA photographs. (3) There is a noticeable absence of any meridional or near-meridional reflection between approximately 13 and 4 A.

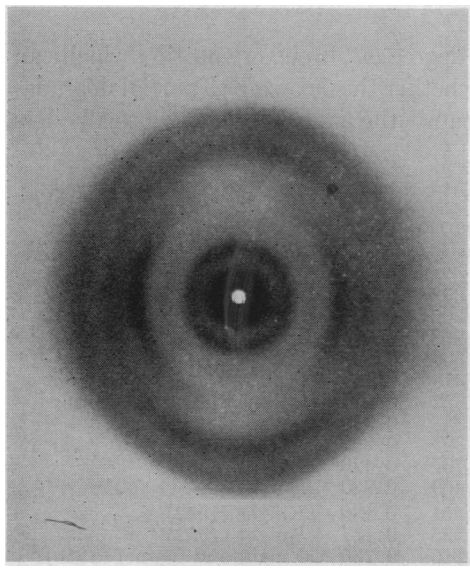

FIG. 1.—X-ray diffraction photograph of RNA fiber. Experimental conditions: fiber axis vertical; Cu $K\alpha$ radiation; 10-cm. cylindrical camera filled with helium gas; relative humidity, 66 per cent.

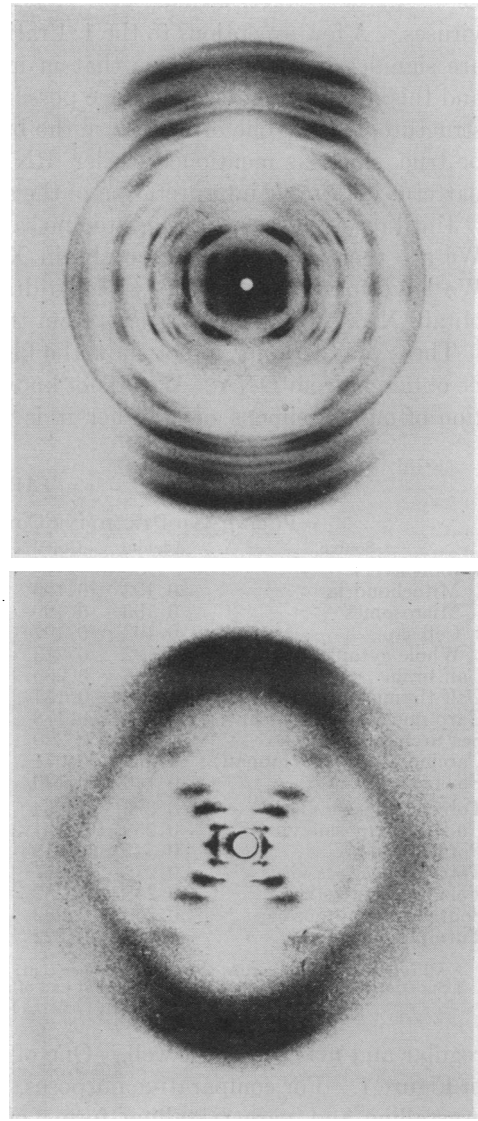

FIG. 2.—X-ray diffraction photographs of DNA fibers. *a*, Crystalline state of DNA (from M. H. F. Wilkins, W. E. Seeds, A. R. Stokes, and H. R. Wilson, *Nature*, **172**, 759–762, 1953). *b*, Paracrystalline state of DNA (from R. E. Franklin and R. G. Gosling *Nature*, **171**, 740–742, 1953).

This is again similar to the DNA photographs, where it is interpreted in terms of a helical structure. (4) The equatorial reflections are not well resolved but seem to be compatible with a fiber of diameter approximately 21–25 A. (5) The exact details of the patterns are strongly influenced by the water content of the specimen.[5] The innermost reflection at 22 A (probably equatorial) disappears at very low

relative humidities but appears strongly when the humidity is high. The exact distance of this spacing varies somewhat with water content. However, the general appearance of the photograph remains approximately the same; hence the structure cannot have undergone any serious modifications while taking up water.

The X-ray pattern therefore suggests a DNA-like structure for RNA. However, since the DNA model is based upon complementary base ratios which are not found in many RNA's, this suggestion has many difficulties. It is possible that non-complementary side chains may arise from a complementary main structure, but proof of this awaits more direct chemical evidence of branches in RNA.

We have been able to construct single-chain helical models for RNA in which the free ribose hydroxyl group is satisfactorily hydrogen-bonded to a negatively charged phosphate group. However, we have not been able to form satisfactory intramolecular hydrogen bonds between the bases, which, in this model, remain free to form external hydrogen bonds. Experimentally, it appears essential to determine whether the bases in RNA form significant numbers of intramolecular hydrogen bonds, as in DNA. A re-examination of the electrometric titration curve on undegraded specimens may settle this point.

In the above, we have discussed RNA structure from the viewpoint of its possible connection with protein synthesis. We might also briefly consider the possibility of a genetic role for RNA. This arises from the complete absence of DNA in the plant viruses and its replacement by RNA. In these viruses the genetic material must be the RNA component or the protein component, or possibly both. We might hope to obtain a clue by examining the base ratios. It is conceivable that the RNA's which may have a genetic role would show complementarity. The results, however, seem just the opposite. Plant virus RNA's show great departure from the 1:1 ratio, while RNA's from sources to which we need not necessarily postulate a genetic role (e.g., microsomes, mitochondria) often provide beautiful examples of complementarity. We have no explanation for this finding.

Summary.—We have discussed possible functions and relations between DNA and RNA. Current structural studies on RNA have been described. Further chemical and crystallographic work is necessary before we can discover the relationship between the structure of RNA and the origin of protein specificity.

Acknowledgment.—We have profited from discussions with Dr. L. E. Orgel.

* Contribution No. 1917 from the Gates and Crellin Laboratories of Chemistry. This investigation was supported in part by a grant from the National Foundation for Infantile Paralysis.

† United States Public Health Service.

[1] O. T. Avery, C. M. MacLeod, and M. McCarty, *J. Exptl. Med.*, **79**, 137–157, 1944; A. D. Hershey and M. Chase, *J. Gen. Physiol.*, **36**, 39–56, 1952.

[2] D. M. Brown and A. R. Todd, *J. Chem. Soc.*, pp. 52–58, 1952.

[3] D. M. Brown, G. D. Fassman, D. I. Magrath, A. R. Todd, W. Cochran, and M. M. Soolfson, *Nature*, **172**, 1184–1185, 1953.

[4] M. H. F. Wilkins, A. E. Stokes, and H. R. Wilson, *Nature*, **171**, 738–740, 1953.

[5] A. Rich and J. D. Watson, *Nature* **173**, 995–996, 1954.

[6] J. D. Watson and F. H. C. Crick, *Nature*, **171**, 737–738, 1953.

[7] R. E. Franklin and R. G. Gosling, *Nature*, **171**, 740–742, 1953.

[8] E. Chargaff, *Experientia*, **6**, 201–209, 1950; G. R. Wyatt, *J. Gen. Physiol.*, **36**, 201–205, 1952.

[9] J. D. Watson and F. H. C. Crick, *Nature*, **171**, 964–967, 1953.

[10] J. Brachet and H. Chantrenne, *Nature*, **168**, 950, 1951.

[11] H. Borsook, C. L. Deasy, A. J. Haagen-Smit, G. Keighley, and P. H. Lowy, *J. Biol. Chem.*, **196**, 669–694, 1952.
[12] T. Caspersson, *Symposia Soc. Exptl. Biol.*, **1**, 127–151, 1947; J. Brachet, *Symposia Soc. Exptl. Biol.*, **1**, 207–224, 1947; A. L. Bounce, *Nature*, **172**, 541, 1953.
[13] R. Markham, *Advances in Virus Research*, **1**, 315–332, 1953.
[14] A. Claude, *Science*, **97**, 451–458, 1943.
[15] W. E. Cohn and E. Volkin, *J. Biol. Chem.*, **203**, 319–332, 1953.
[16] D. O. Jordan, *Progr. Biophys. and Biophys. Chem.*, **2**, 51–89, 1951.
[17] G. W. Crosbie, R. M. S. Smellie, and J. N. Davidson, *Biochem. J.*, **54**, 287–292, 1953.
[18] H. A. Deluca, R. J. Rossita, and K. P. Strickland, *Biochem. J.*, **55**, 193–201, 1953.
[19] G. Hamoir, *Biochem. J.*, **50**, 140–144, 1952.
[20] W. C. Vincent, these PROCEEDINGS, **38**, 139–145, 1952.
[21] E. Chargaff, B. Magasanik, E. Vischer, C. Green, R. Doniger, and D. Elson, *J. Biol. Chem.*, **186**, 51–67, 1950.
[22] R. J. C. Harris, R. N. Beale, and E. M. F. Roe, *J. Chem. Soc.*, pp. 1397–1407, 1952.
[23] C. A. Knight, *J. Biol. Chem.*, **197**, 241–249, 1952.
[24] C. A. Knight and R. W. Dorner, *J. Biol. Chem.*, **205**, 959–967, 1953.
[25] R. Markham, in P. Fildes and W. E. Van Heyningen, *The Nature of Virus Multiplication* (Cambridge: At the University Press, 1953, pp. 85–95).
[26] R. Markham and J. D. Smith, *Biochem. J.*, **49**, 401–406, 1951.

BIOSYNTHESIS OF NUCLEIC ACID IN ESCHERICHIA COLI

BY ELLIS BOLTON

DEPARTMENT OF TERRESTRIAL MAGNETISM, CARNEGIE INSTITUTION OF WASHINGTON

Communicated by M. A. Tuve, May 26, 1954

The nucleic acids of different organisms are similarly constructed and are apparently endowed with common functions. An underlying unity appears also to characterize their biological synthesis. Biochemical studies have shown that microörganisms,[1] birds,[2] and mammals[3] utilize carbon dioxide for the synthesis of nucleic acid purines and pyrimidines; that a number of features of pentose synthesis are common to bacteria[4] and yeast;[5] and that nucleotide-synthesizing enzymes occur in many kinds of cells.[6] However, knowledge of nucleic acid synthesis is not complete for any species. It is not known what kinds of subunits enter nucleic acid macromolecules or what kinds of processes link the units together into strands. Investigations which couple the techniques of microbiology with tracer isotopes and chromatography promise to supply answers to these questions.

This paper reports the results of experimental studies which have surveyed nucleic acid biosynthesis in growing *Escherichia coli*. The results reveal some characteristics of the materials and processes which contribute to the formation of nucleic acid macromolecules.

Escherichia coli is a heterotrophic bacterium which grows readily in simple, chemically defined culture media. At 37° C. its generation time is about one hour. While growing exponentially, it will utilize relatively simple carbon sources such as glucose and carbon dioxide to supply its carbon requirements. It will also utilize a host of supplemental carbon compounds, and it ordinarily makes use of each in a specific fashion. Nutritional characteristics such as these can be ex-