The double helix: a personal view

Francis Crick

Medical Research Council Laboratory for Molecular Biology, Hills Road, Cambridge, UK

Francis Crick reviews the papers published 21 years ago on the structure of DNA and the reaction to them.

FOR this anniversary I thought it might be appropriate to look back, in a rather informal way, at the original papers on the structure of DNA to see how they appear today in the light of 21 years of research.

During the spring and summer of 1953 Jim Watson and I wrote four papers on the structure and function of DNA. The first appeared in *Nature* on April 25 accompanied by two papers from King's College London, the first by Wilkins, Stokes and Wilson, the other by Franklin and Gosling. Five weeks later we published a second paper in *Nature*, this time on the genetic implications of the structure. A general discussion was included in the volume that came from that year's Cold Spring Harbor Symposium, the subject of which was viruses. We also published a detailed technical account of the structure, with rough coordinates, in an obscure journal[1] in the middle of 1954.

The first *Nature* paper was both brief and restrained. Apart from the structure itself the only feature of the paper which has excited comment was the short sentence: "It has not escaped our notice that the specific pairing we have postulated immediately suggests a possible copying mechanism for the genetic material". This has been described as 'coy', a word that few would normally associate with either of the authors, at least in their scientific work. In fact it was a compromise, reflecting a difference of opinion. I was keen that the paper should discuss the genetic implications. Watson was against it. He suffered from periodic fears that the structure might be wrong and that he had made an ass of himself. I yielded to his point of view but insisted that something be put in the paper, otherwise someone else would certainly write to make the suggestion, assuming we had been too blind to see it. In short, it was a claim to priority.

Why, then, did we change our minds and, within only a few weeks, write the more speculative paper of May 30? The main reason was that when we sent the first draft of our initial paper to King's College we had not yet seen their own papers. Consequently we had little idea of how strongly their X-ray evidence supported our structure. The famous 'helical' X-ray picture of the B form, reproduced by Franklin and Gosling in their paper, had been shown to Watson, but he certainly had not remembered enough details to construct the arguments about Bessel functions and distances which the experimentalist gave. I myself, at that time, had not seen the picture at all. Consequently we were mildly surprised to discover that they had got so far and delighted to see how well their evidence supported our idea. Thus emboldened, Watson was easily persuaded that we should write a second paper.

The papers in *Nature*

The two experimental papers of April 25 overlap to a considerable extent. Rosalind Franklin's paper mentions the crystalline A structure, but only briefly, except for the claim that the Patterson superposition function (which was in the press at the time) supported two chains rather than three. Both papers stress that there must be more than one chain in the structure. Indeed Maurice Wilkins had personally told Chargaff that a year or so earlier. Both present the argument that the positions of the intensity maxima ruled out two (parallel) chains related by a dyad parallel to the fibre axis. Neither gave the neat argument, due to Watson, that their own density measurement, together with the observed change in length between the two forms, supported two chains rather than three. Franklin noted that if there were several chains they could not be equally spaced and that 'equivalence' favoured two rather than three. It was not explicitly stated, however, that equivalence implies dyad axes perpendicular to the fibre axis and that therefore the two chains must run in opposite directions. Nor did she realise that the monoclinic unit cell of the A form also suggested this, although we had deduced this from her own experimental data.

Both papers correctly concluded from the intensity positions that the phosphate–sugar backbone was on the outside of the structure and that the bases were stacked on the inside. Franklin repeated the argument, which she had made to us verbally a year earlier, that the phosphates would be hydrated (in which she was perfectly right) and therefore that they would probably be on the outside of the molecule. In short, both the groups at King's College had obtained a fairly general idea of the structure but they had done no proper model building. Mainly because of this they had missed the pairing of the bases and they had completely overlooked the significance of Chargaff's rule.

The omissions in the paper by Watson and myself are also striking. The structure is produced like a rabbit out of a hat, with no indication as to how we arrived at it. No dimensions are given (let alone coordinates) except that the base pairs were 3.4 Å apart and that the structure had 10 base pairs in its repeat. The exact nature of the base pairing was not immediately obvious; nor even unambiguous since at that time there were two systems for numbering pyrimidine rings. Most of this information was provided in the subsequent papers. However the general nature of the structure was clear enough, though the tone of the paper ("it must be regarded as unproved until it has been checked against more exact results") was, apart from the short first paragraph, rather muted.

Although a casual reader could easily have overlooked the significance of the first set of papers, especially as they were full of obscure crystallographic jargon, he could hardly miss the impact of our second one. The biologically important features of the proposed structure were explicitly described. The base pairs were listed with the minimum of hedging about tautomerism and were illustrated in scale diagrams. The proposed duplication mechanism was spelt out in simple terms, unmarred by any trace of algebra. In spite of the discussion of the difficulties of unwinding, the list of unsolved problems and the reservations about the unproved nature of the structure, the final paragraph leaves little doubt that the authors thought they had a good idea.

How do they stand today?

How have these early papers stood the test of time? It can now be taken as firmly established that DNA usually consists of two chains, wound together and running in opposite directions. The evidence for this statement is so extensive that it would take too long to quote it all here. The fact that normally A pairs with T, and G with C, is also well established but the details were less certain until recently. The G:C pair was never in serious doubt. Watson and I drew this with only two hydrogen bonds but mentioned in our technical paper[1] that three was also a possibility. This was made almost certain by the theoretical arguments of Pauling and Corey[2] and was confirmed by X-ray structure determinations of single crystals of base pairs. The same technique showed that the A:T (or A:U) pair in single crystals usually did not have the configuration Watson and I suggested. The matter was only finally resolved about a year ago when Rich and his colleagues published two crystal structures; that of GpC paired with itself[3] and ApU paired with itself[4] (the backbone in each case was ribose), both to about 0.9 Å. They show not only the expected configurations for the base pairs but also make it highly likely that, as we claimed, nucleic acid helices are right handed.

In 1953 it was uncertain whether RNA could form a double helix. Watson and I stated that we thought we could not build our model for the B form of DNA with an RNA backbone. The discovery of double-stranded RNA viruses proved, however, that biological RNA too could form a double helix, though with slightly different parameters. The detailed coordinates we had (tentatively) suggested for DNA were soon shown to be incorrect (we had put the backbone at too big a radius) and much more accurate coordinates were provided by Wilkins and his colleagues, using fairly sophisticated methods of handling their much improved X-ray data. The general correctness of this work has been strongly supported recently by the single-crystal studies, mentioned above, of Rich and his coworkers.

Recently, Bram[5] has put forward evidence that the parameters of a DNA double helix may vary somewhat with base composition, though whether this is a trivial variation or has deep biological implications is at present uncertain. Watson and I were so impressed with the apparent uniformity of the double helix from different biological sources and the regularity of the backbone of our model that we had no hesitation in saying that it "seems likely that the precise sequence of the bases is the code that carries the genetic information", an idea which gave me plenty to think about in the next 10 or 12 years.

Nothing was said about the possibility that the two chains might be melted apart and then annealed together again, correctly lined up. The discovery of this by Marmur and Doty has provided one of the essential tools of molecular biology. I can still remember the excitement I felt when Paul Doty told me about it at breakfast one day in New York in a hotel overlooking Central Park. But in other respects we were almost too far sighted, as witness our remark that recombination would probably depend upon base pairing. We struggled for several years to produce neat models for this, all to no avail, partly because we accepted copy choice too easily but also because we were trying to invent a mechanism which did not need additional enzymes. This showed a gap in our overall grasp of molecular biology, which can also be glimpsed in our tentative suggestion that DNA synthesis might not need an enzyme, a remark I should certainly not make today except perhaps in the context of the origin of life.

As to DNA replication, our earliest description was mainly schematic. We realised that plain nucleotides were not likely to be the immediate precursor but missed the rather obvious idea that they were nucleoside triphosphates, again a lack of insight into biochemistry. We did suggest the so-called Y mechanism (in the Cold Spring Harbor paper) but did not mention the difficulties due to the direction of synthesis of antiparallel chains, though I frequently emphasised it a few years later. Looking back, I think we deserve some credit for not being inhibited by the difficulty of unwinding which we clearly recognised and for our forthright stand against paranemic (as opposed to plectonemic) coiling. In this instance our grasp of X-ray diffraction was invaluable.

The functions of DNA

It is, of course, somewhat a matter for surprise that DNA synthesis is not fully understood even today. It would take too much space to discuss the complex and rapidly moving field here. Semiconservative replication in many instances is firmly established. The process certainly occurs as if base pairing were taking place, but I have often asked myself what evidence would make it certain that base pairing really occurs rather than some elaborate allosteric mechanism, even though the latter seems unlikely. Perhaps only an X-ray determination of the structure of the polymerase will finally answer the question. Meanwhile the topics of Okazaki fragments, rolling circle models, RNA primers and the exact roles of the various polymerases will keep many people busy. Even at that early period we did at least ask whether the DNA of a chromosome was in one long molecule, though the idea of circular DNA never occurred to us. Nor did we suggest that a virus might have single-stranded DNA. There is however one remark which may turn out to be perspicacious ". . . we suspect that the most reasonable way to avoid tangling is to have the DNA fold up into a compact bundle as it is formed". As we struggle with the structure of the *E. coli* chromosome and the even more formidable problem of the structure of the chromosomes of higher organisms—probably the major unsolved problem of molecular biology today—it might be worth remembering this tentative suggestion from the distant past.

The other topic we touched on was mutation. This was of the base-substitution type—there is no hint of frameshift mutants. We totally missed the possible role of enzymes in repair although, due to Claud Rupert's early very elegant work on photoreactivation, I later came to realise that DNA is so precious that probably many distinct repair mechanisms would exist. Nowadays one could hardly discuss mutation without considering repair at the same time.

There is no hint in these early papers that nucleic acid might form a complex three-dimensional structure such as we now find in transfer RNA nor even the idea of the hypothetical Gierer loops. Our message was that DNA was simple and alone carried the genetic information. We saw no reason to complicate it till we had to. For the same reason although we must have drawn a G:U pair we attached no importance to it. "Wobble" was still far in the future, but these, it seems to me, are forgivable oversights.

Reactions to the structure

It is really for the historian of science to decide how our structure was received. This is not an easy question to

answer because there was naturally a spectrum of opinion which changed with time. There is no doubt, however, that it had a considerable and immediate impact on an influential group of active scientists. Mainly due to Max Delbrück, copies of the initial three papers were distributed to all those attending the 1953 Cold Spring Harbor Symposium and Watson's talk was added to the programme. A little later I gave a lecture at the Rockefeller which I am told produced considerable interest, partly I think because I mixed an enthusiastic presentation of our ideas with a fairly cool assessment of the experimental evidence, roughly on the lines of the article which appeared in *Scientific American* in October, 1954. Sydney Brenner, who had just finished his PhD, at Oxford under Hinshelwood, appointed himself, in the summer of 1954, as Our Representative at Cold Spring Harbor and took some pains to get the ideas over to Demerec. It was about this time that Matt Meselson, just moving into biology from physical chemistry, grasped the importance of inventing a new method to tackle the problem of semiconservative replication, a theoretical analysis which led to density gradient centrifugation. But not everyone was convinced. Barry Commoner insisted, with some force, that physicists oversimplified biology, in which he was not completely wrong. Chargaff, when I visited him in the winter of 1953–54, told me (with his customary insight) that while our first paper in *Nature* was interesting, our second paper on the genetic implications was no good at all. I was mildly surprised to find, when, some years later, in 1959, I talked with Fritz Lipmann who had arranged that I should give a series of lectures at the Rockefeller, that he had not really grasped our scheme of DNA replication. (It emerged that he had been talking to Chargaff.) By the end of the lectures, however, when he summed up, he gave a remarkably clear outline of our ideas. Arthur Kornberg has told me that when he began work on DNA replication he did not believe in our mechanism, but his own brilliant experiments soon made him a convert, though always a careful and critical one. It was his work which produced the first good evidence that the two chains run in opposite directions. All in all it seems to me that we got a very fair hearing, better than Avery and certainly a lot better than Mendel.

Not that it was all plain sailing. We were naturally delighted with the work of Meselson and Stahl, and of Herbert Taylor, on semiconservative replication, though I have never thought this the essence of our ideas which lies rather in the base pairing. Seymour Benzer's genetic analysis of the r_{II} locus of phage T4 encouraged us greatly. But we had to live through the claims of Marshak that there was no DNA in *Arbacia* eggs and of a Canadian group that the amount of DNA synthesis in one cell cycle was twice the expected amount. At a later stage Cavalieri claimed that the basic DNA structure had four chains, rather than two, an idea which cropped up again more recently. On the crystallographic side Donohue, whose advice had been crucial to our understanding of base pairing, was a persistent critic of the validity of the later X-ray work, but in recent years he carried it too far, refusing, for example, to admit as evidence the great accumulation of data showing that the two chains are antiparallel. (In 1956, he had rashly published, with Stent, a quite erroneous structure having like-with-like pairing.) I hope the recent papers by Rich, referred to above, have to some extent reduced his doubts, which at times had some justification.

Who might have discovered it?

Then there is the question, what would have happened if Watson and I had not put forward the DNA structure? This is 'iffy' history which I am told is not in good repute with historians, though if a historian cannot give plausible answers to such questions I do not see what historical analysis is about. If Watson had been killed by a tennis ball I am reasonably sure I would not have solved the structure alone, but who would? Olby[6] has recently addressed himself to this question. Watson and I always thought that Linus Pauling would be bound to to have another shot at the structure once he had seen the King's College X-ray data, but he has recently stated that even though he immediately liked our structure it took him a little time to decide finally that his own was wrong. Without our model he might never have done so. Rosalind Franklin was only two steps away from the solution. She needed to realise that the two chains must run in opposite directions and that the bases, in their correct tautomeric forms, were paired together. She was, however, on the point of leaving King's College and DNA, to work instead on TMV with Bernal. Maurice Wilkins had announced to us, just before he knew of our structure, that he was going to work full time on the problem. Our persistent propaganda for model building had also had its effect (we had previously lent them our jigs to build models but they had not used them) and he was proposing to give it a try. I doubt myself whether the discovery of the structure could have been delayed for more than two or three years.

There is a more general argument, however, recently proposed by Gunther Stent and supported by such a sophisticated thinker as Medawar. This is that if Watson and I had not discovered the structure, instead of being revealed with a flourish it would have trickled out and that its impact would have been far less. For this sort of reason Stent had argued that a scientific discovery is more akin to a work of art than is generally admitted. Style, he argues, is as important as content.

I am not completely convinced by this argument, at least in this case. Rather than believe that Watson and Crick made the DNA structure, I would rather stress that the structure made Watson and Crick. After all, I was almost totally unknown at the time and Watson was regarded, in most circles, as too bright to be really sound. But what I think is overlooked in such arguments is the intrinsic beauty of the DNA double helix. It is the molecule which has style, quite as much as the scientists. The genetic code was not revealed all in one go but it did not lack for impact once it had been pieced together. I doubt if it made all that difference that it was Columbus who discovered America. What mattered much more was that people and money were available to exploit the discovery when it was made. It is this aspect of the history of the DNA structure which I think demands attention, rather than the personal elements in the act of discovery, however interesting they may be as an object lesson (good or bad) to other workers.

My own reactions

I have sometimes been asked whether I had ever contemplated writing my own account of the discovery. In the 1950s I did give a lecture on this subject to a group of historians of science at Cambridge and to a similar group at Oxford. I was able to be rather more scholarly than Watson could allow himself in *The Double Helix*, which is better regarded as a rather vivid fragment of his autobiography, written for a lay audience. As to a book I confess I did get as far as composing a title (*The Loose Screw*) and what I hoped was a catchy opening ("Jim was always clumsy with his hands. One had only to see him peel an orange. . .") but I found I had no stomach to go on. Recently we made a film together about it for undergraduates. Much had to be left out when the film came to be cut but it does to some extent supplement Jim's book. Since Olby's detailed and scholarly account[6] will soon be available I doubt if there is now much more I can usefully add.

Finally one should perhaps ask the personal question—am I glad that it happend as it did? I can only answer that

I enjoyed every moment of it, the downs as well as the ups. It certainly helped me in my subsequent propaganda for the genetic code. But to convey my own feelings, I cannot do better than quote from a brilliant and perceptive lecture I heard years ago in Cambridge by the painter John Minton (he later committed suicide) in which he said of his own artistic creations "the important thing is to be there when the picture is painted". And this, it seems to me, is partly a matter of luck and partly good judgement, inspiration and persistent application.

1 Crick, F. H. C., and Watson, J. D., *Proc. R. Soc.*, **A223**, 80–96 (1954).
2 Pauling, L., and Corey, R. B., *Archs Biochem. Biophys*, **65**, 164–181 (1956).
3 Day, R. D., Seeman, N., Rosenberg, J., and Rich, A., *Proc. natn. Acad. Sci. U.S.A.*, **70**, 849–853 (1973).
4 Rosenberg, J., Seeman, N., Kim, J. J., Suddath, F., Nicholas, H., and Rich, A., *Nature*, **243**, 150–154 (1973).
5 Bram, S., and Tougard, P., *Nature new Biol.*, **239**, 128–131, (1972).
6 Olby, R. C., *The Path to the Double Helix* (Macmillan, London, 1974).

Molecular basis of biological specificity

Linus Pauling

Institute of Orthomolecular Medicine, 2700 Sand Hill Road, Menlo Park, California 94025

Linus Pauling reviews his work on the molecular basis of biological specificity and remembers his erroneous conception of a three-chain helix structure for DNA in 1952.

During the decade 1930–40 I formulated a general theory of the molecular basis of biological specificity, involving the idea that biological specificity results from the interaction of complementary molecular structures, with hydrogen bonds among the most important of the weak intermolecular forces between the interacting molecules. The most striking example of specific biological interactions of this sort is the interaction between the two complementary strands of the DNA molecule in the double helix discovered by Watson and Crick 21 years ago.

Early work

My early work was on the determination of the structure of crystals by the X-ray diffraction technique, the determination of the structure of gas molecules by electron diffraction, and the application of quantum mechanics to physical and chemical problems, especially the structure of molecules and the nature of the chemical bond. In 1929, when Thomas Hunt Morgan came to the California Institute of Technology, bringing with him a number of very able younger biologists, I began to become familiar with biological problems, and to think about possible ways in which biological specificity could be explained in terms of interactions between molecules. I worked on several problems of biological specificity, from the molecular point of view, without success; one of them was the problem of explaining the self sterility of the marine organism *Ciona* (the sea squirt), which was being studied by Morgan. In 1934 the problem of the shape of the oxygen equilibrium curve of haemoglobin attracted my attention. Consideration of the structure of haemoglobin led to the idea that investigation of the magnetic properties of this substance and its derivatives would provide valuable information, and work along these lines, in collaboration with C. D. Coryell and a number of students, was initiated. Alfred E. Mirsky of the Rockefeller Institute for Medical Research, who had been studying haemoglobin for several years, came to Pasadena for a year, and he and I formulated a theory of the structure of native, denatured, and coagulated proteins, based upon the concept that a native protein molecule consists of one polypeptide chain (or of two or more such chains) folded into a uniquely defined configuration, in which it is held by hydrogen bonds between the peptide nitrogen and oxygen atoms, as well as by other weak forces, with denaturation involving a loss of this well-defined structure[1].

Antigens and antibodies

In 1936, while I was on a short visit to the Rockefeller Institute for Medical Research, Karl Landsteiner asked me how I would explain the observed properties of antibodies and antigens by means of their molecular structure. I thought about this problem during the following years, and consulted Landsteiner about the interpretation of sometimes conflicting experimental results. By 1940 I had formulated a theory of the structure and process of formation of antibodies[2]. This theory was based upon the concept that the specific combining region of an antibody molecule is complementary in structure to a portion of the surface of the antigen, with the antigen–antibody bond resulting from the cooperation of weak forces (electronic Van der Waals forces, electrostatic interaction of charged groups, and hydrogen bonding) between the complementary structures, over an area sufficiently large that the total binding energy could resist the disrupting influence of thermal agitation. Precipitating and agglutinating antibodies were assumed to be bivalent, consisting of a central part, with structure common to all or almost all antibodies produced by the animal, and two end parts, the combining regions, with structure complementary to that of the antigen. (The idea of complementary structures for antibody and antigen was suggested by Breinl and Haurowitz[3], Alexander[4], and Mudd[5]. There is some intimation of it in the early work of Ehrlich and Bordet.) The complementary combining regions were assumed to be formed by the folding of polypeptide chains in the presence of the antigen, in such a way that the forces of attraction would mould the folding chain into a structure complementary to that of a portion of the antigen, with the folded chain then being held in this configuration by hydrogen bonds and other interactions, even after the antibody has dissociated from the antigen on which the combining group was moulded. Dan Campbell, David Pressman, and a number of other workers in our laboratory carried out experimental studies that verified the valence 2 for precipitating and agglutinating antibodies[6,7] and that left no doubt that the combining regions of antibodies are complementary in structure to the homologous haptenic groups of the antigen[8]. The fit of the combining region of the antigen to the hapten was shown to be close, better than 20 pm in some cases, and the effects of Van der Waals attraction, electrostatic forces, and hydrogen-bond formation were