CONTRIBUTIONS to SCIENCE, **2** (4): 513–524 (2004)
Institut d'Estudis Catalans, Barcelona

Different forms of the double helix: architecture and function*

Alexander Rich**
Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA

Abstract

The DNA double helix was discovered 50 years ago. This motif was found in several other places, most notably in RNA molecules. It is also found in DNA-RNA hybrids as well as transfer RNA molecules where double helical segments combine to form a more complex structure. A distinct variant of the DNA molecule was discovered in which the helix turns left-handed, which has led to practical consequences in the treatment of certain viral diseases.

Fifty years ago the double helix era began in biology. The model formulated by Watson and Crick (1953) [65] pointed to the solution of two important biological problems. First, it provided a clear method for containing information in the sequence of nucleotides in long DNA molecules. Second, the structure based on complementary pairing between the bases provided a direct explanation for how the molecule could be duplicated through strand separation and synthesis of the complementary strands. In addition, the double helix was aesthetically appealing. A double helical staircase with base pairs in the center and sugar phosphate chains on the outside gently coiling in a right-handed manner. There were two grooves on the molecule, a major groove and a minor groove. The asymmetries in the grooves were related to the fact that the glycosyl bonds connecting the base pairs to the sugar phosphate chains were found on one side of the base pair giving rise to two different types of grooves.

But there was a mystery. The fiber x-ray diffraction data obtained by Wilkins, Franklin, and colleagues had already revealed that there were two forms of the double helix, an A form that was obtained from air-dried DNA fibers, and a B form in which the fiber was maintained in a hydrated state. Because of the evident simplicity of the B form diffraction pattern, the Watson-Crick formulation described the B form. Since the molecule readily converted from A to B form, it clearly suggested that there was another form the double helix could adopt in the A form. Thus, the initial mystery –why are there two such forms?

It was known at the time that there are two major nucleic acids, one DNA and the other RNA, in which the sugar phosphate backbone contained ribose rather than the deoxyribose of DNA. One of the early questions that arose immediately after the Watson-Crick formulation was related to the nature of the RNA molecule-could it form a double helix?

Discovery of the RNA Double Helix

My research work on RNA structure started in the early 1950s when nothing was known of its three-dimensional structure, and the question asked was whether RNA could form a double helix comparable to the one described by Watson and Crick (1953). In their paper, they suggested that the double helix could not form with RNA molecules because of the van der Waals crowding associated with the additional oxygen atom on the ribose ring. The issue was further complicated by the fact that it was not known at that time whether RNA was linear or branched, since the additional hydroxyl group on the ribose ring represented a potential branch point. While we were both postdoctoral fellows at Cal Tech in 1953 and 1954, Jim Watson and I tried to make oriented fibers of RNA to see whether their X-ray diffraction pattern was helical. The results were inconclusive [50, 51]. On moving to the National Institutes of Health (NIH) later in 1954, I continued to work on the problem. No significant progress was made until the polynucleotide phosphorylase enzyme was discovered by Marianne Grunberg-Manago in Severo Ochoa's laboratory [11]. That enzyme made very long strands of polyribonucleotides, and with it, structural investigations became more productive.

David Davies joined me at NIH to work on RNA structure. In 1956, we mixed solutions of polyriboadenylic acid (poly rA) and polyribouridylic acid (poly rU) and discovered a remarkable transformation revealed by X-ray studies. These two molecules actually reacted with each other to produce a double helix! A brief note was sent to the Journal of the

* This article is the transcription of a lecture given by the author at the Institute for Catalan Studies on the 10th of November 2003.

** Address for correspondence: Alexander Rich, Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139 USA

American Chemical Society (JACS) in June of that year [48], describing the work and a preliminary interpretation of the pattern. The diffraction pattern clearly indicated the formation of a helical complex that was not present in either of the two individual polymers. Furthermore, the pattern had significant differences compared with those produced from DNA fiber diffraction patterns. In particular, the first layer line was very strong, whereas in the DNA diffraction patterns, the first layer line was weak.

This result, which seems so obvious today, generated a great deal of skepticism at the time. While walking down a long corridor at NIH, I met Herman Kalckar, an eminent Danish biochemist. During our conversation, I mentioned that we discovered that poly rA and poly rU formed a double helix. Kalckar was incredulous. “You mean without an enzyme?” he asked. His attitude was justified, since the only double helix known at that time was one made with the DNA polymerase enzyme that Arthur Kornberg had purified [29]. Other critics thought it was highly unlikely that polymers containing over a thousand nucleotides would be able to disentangle themselves and form a regular structure. They believed it would be hopelessly entangled. Further, some polymer chemists thought it improbable that two highly negatively charged polymers would combine.

Two weeks after sending off the 1956 JACS note, I wrote a letter to my former postdoctoral mentor, Linus Pauling, describing these results. The letter reveals a sense of incredulity on my part that this reaction could happen and that it was “completely reproducible.” This was the first demonstration that RNA molecules could form a double helix. It was also the first hybridization reaction and, as noted in the JACS letter, we pointed out that “this method of forming a two-stranded helical molecule” utilizing specific interactions could be used for a variety of studies.

An important method for studying the nucleic acids was measuring their absorbance in the ultraviolet. For some time, it had been known that polymerization of nucleotides resulted in a decrease (hypochromism) in absorbance at 260 mm. Further, it had been reported earlier that adding poly rA to poly rU led to hypochromism [64]. Although the mechanism of hypochromicity was not understood at the time, it was a useful tool for analysis. Further insight into the reaction of poly rA and poly rU was obtained by carefully measuring hypochromicity in mixtures of varying composition. In work carried out with Gary Felsenfeld (Fig. 1), the absorbance fell to a very sharp minimum at a 1:1 mole ratio [8]. It implied that the system was dynamic-the molecules came together to form a helical duplex, but they then disassembled and reassembled so that ultimately all of the gaps between adjacent molecules were closed. This was a dynamic picture of nucleic acid molecules rapidly associating and dissociating, a view that was different from the prevailing view. Before this experiment, macromolecular nucleic acids were regarded as somewhat immobile. The only available mental picture of the structure was the double helix formulated by Watson and Crick. Most people thought about the nucleic acids in static terms, whereas the reaction between poly rA and poly rU suggested a more dynamic interpretation of nucleic acid molecules. This led to a significant change in thinking, and it expressed itself in a variety of other systems, such as recombination in which the movements of nucleic acid molecules play a crucial role.

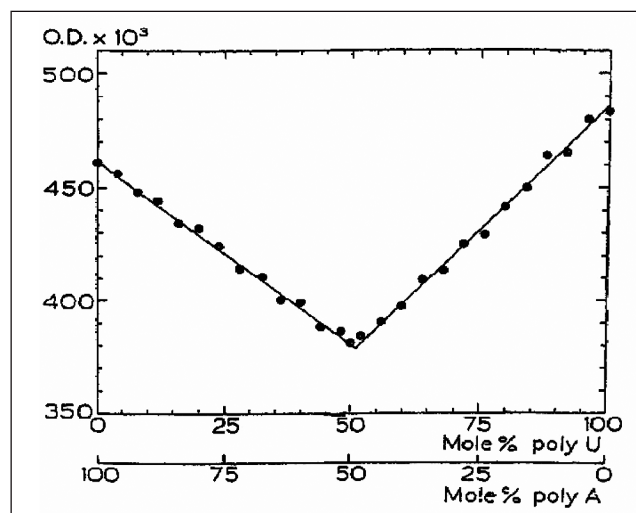

Figure 1. Titration of the optical density at 259 mμ of various mixtures of poly rA and poly rU [8].

Triplex as well as duplex

A year later, it was discovered that the addition of divalent cations such as magnesium would change the picture dramatically [9], leading to the formation of a three-stranded molecule (Fig. 2). We concluded that a second strand of polyuridylic acid bound in the major groove of the poly rA-poly rU duplex. Addition of this strand did not increase the diameter of the molecule and neatly accounted for the observed 50% increase in sedimentation constant. It was further proposed that the second poly rU strand bound to the N6 and N7 of adenine using two hydrogen bonds to uracil. This proposal was considerably strengthened 2 years later by the X-

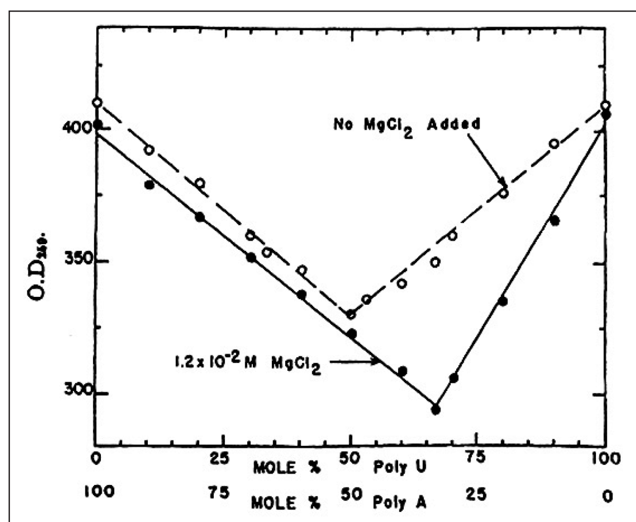

Figure 2. Titration to the optical density at 259 mμ of mixtures of poly rA and poly rU showing the effect of $MgCl_2$ leading to formation of a three stranded molecule [9].

ray analysis of the co-crystal containing 9-methyl adenine and 1-methyl thymine by Karst Hoogsteen (1959) [18], which had the same hydrogen bonding. It was an early indicator of the structural complexity inherent in RNA molecules.

Over the next several years, a variety of polynucleotide interactions were studied [5, 42, 43, 44], leading to the formation of two- and three-stranded molecules. Fiber X-ray diffraction and other studies led to the conclusion that the molecules all had bases on the inside and the sugar phosphate chains on the outside. The bases were generally held in place by Watson-Crick or alternative types of hydrogen bonding involving at least two hydrogen bonds.

Sugar rings pucker in two different forms

The deoxyribose sugar ring contains five atoms; they cannot all lie in one plane, and at least one atom must be out-of-plane (Fig. 3), which is called ring pucker. With continued analysis of the DNA fiber patterns, it became clear that the B form contained a ring pucker in which the C2' atom was out-of-plane on the same side as the base (C2' endo). Because of that pucker, the phosphate groups were nearly 7 Å apart, giving rise to an extended polynucleotide chain. Study of the more complex A form led to the conclusion that the C3' atom was out-of-plane (C3' endo). Although the energy barrier is low for pucker changes in deoxyribose, it is considerably higher in ribose. In the C3' endo conformation the phosphate groups were about 5.8-6 Å apart. Thus, the sugar phosphate backbone was shortened, and this led to a double helix in which the base pairs were slightly displaced from the center of the helix to produce a flatter helix and a somewhat thicker molecule. A relative scarcity of water molecules stabilized that conformation. It became clear that the normal conformation in the hydrated in vivo environment involved the C2' endo sugar pucker of B form DNA. However, diffraction studies of double-stranded RNA molecules yielded patterns very similar to the A form DNA pattern. The RNA diffraction patterns did not change with humidity, as the molecule was held in that form by the C3' endo conformation, frozen by the energy barrier for changes in ribose pucker.

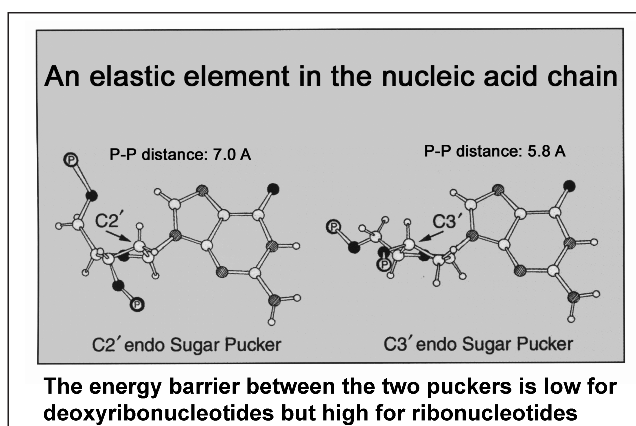

Figure 3. Two major nucleic acid sugar puckers. Deoxyguanosine with attached phosphorous atoms is shown in the two major nucleic acid sugar puckers. The C2' endo pucker (left) is found in B-DNA, whereas the C3' endo pucker (right) is found in A-DNA or in RNA. The distance between successive phosphate groups is close to 7.0 Å in C2' endo and shortens to 5.8-6 Å in the C3' endo pucker. Nucleic acids can convert from one pucker to the other, although it takes greater energy for conversion of ribonucleotides. Carbon, large open circles; hydrogen, small open circles; nitrogen, cross hatched circles; oxygen, black circles; and phosphorous, heavily outlined open circles.

Can DNA "make" RNA?

By 1960, several investigators had shown that crude preparations of an RNA polymerase activity could incorporate ribonucleotides into RNA in the presence of DNA, but the mechanism was not at all clear. It was widely believed that information transfer went from DNA to RNA. But how did that occur [45]? The availability of chemically synthesized oligomers of poly deoxythymidylic acid (poly dT) [58] made it possible to study this experimentally. It was known that the RNA backbone was significantly different from the DNA backbone due to the 2´OH in RNA, so it was not obvious that they could combine. Nonetheless, we showed that these two molecules could accommodate each other to form a hybrid helix containing one strand of poly dT and one strand of poly rA [46], as seen from hypochromism (Fig. 4) and other studies. This was the first experimental demonstration of a DNA-RNA hybrid, and the discovery of messenger RNA was still one year in the future. It immediately provided experimental support for a model of how DNA could "make" RNA, using complementary base-pairing, as in DNA replication. A year later in 1961, experiments by J. Hurwitz with a purified RNA polymerase preparation demonstrated that this was the mechanism underlying information transfer from DNA to RNA [10]. The reaction between poly dT and poly rA was the first experimental demonstration that the two different backbones could adapt to each other in this method of informa-

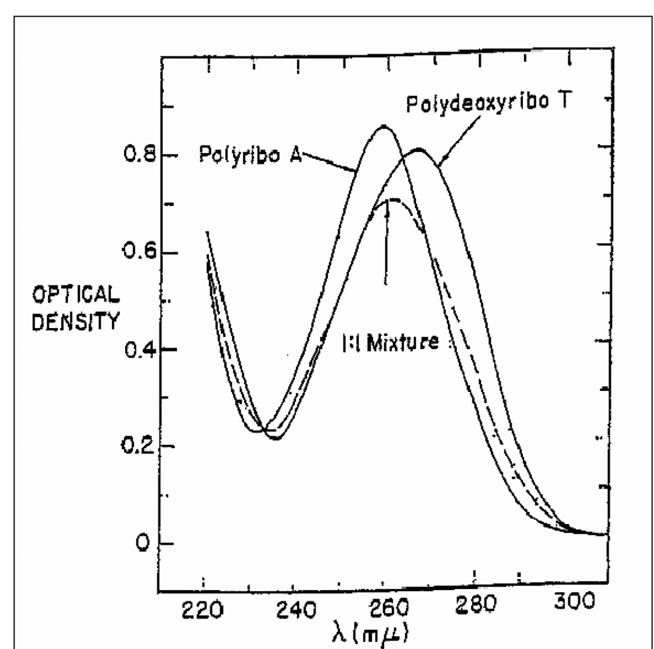

Figure 4. The spectrum of poly rA, poly dT and a 1:1 mixture showing hypochromism.

tion transfer. The reaction was also the first hybridization of a DNA molecule with an RNA molecule. The same hybridization is still widely used today in the purification of eukaryotic mRNA by hybridizing poly dT to their poly rA tails.

A half year later, in 1960, Marmur, Doty, and their colleagues demonstrated that it was possible to renature naturally occurring denatured DNA duplexes by incubating them at an intermediate temperature that would allow the single strands to anneal together with the correct sequence [7, 32]. A year later, this annealing method was also adopted to form DNA-RNA hybrids in viral systems [12].

Single-crystal X-Ray diffraction

X-ray diffraction studies of nucleic acid fibers were carried out extensively by Wilkins, Franklin, and colleagues in the 1950s and 1960s, even though it was realized that the limitations of such studies were enormous. In fiber X-ray diffraction, a rather small number of reflections are registered. However, the number of variables needed to define the structure (at least 3N, where N is the number of atoms) is so great that it was clear that fiber diffraction could not "prove" a structure. It could only say that a particular conformation was compatible with the limited diffraction data from fibers. During this period, the number of single-crystal X-ray diffraction studies was increasing, many of them involving co-crystals of purines and pyrimidine derivatives such as those initiated by Karst Hoogsteen (1959) [18]. These studies were useful in obtaining information about components of nucleic acid structure. For example, we were able to form a co-crystal of cytosine and guanine derivatives which was the first experimental demonstration that these were held together by three hydrogen bonds [57], not two, as initially suggested by Watson and Crick. Linus Pauling had already emphasized this point based on general structural considerations [37]. Several co-crystals were solved of derivatives of adenine and uracil or adenine and thymine during this period in my laboratory [20, 33, 60], as well as in others (reviewed in [61]). However, a disturbing trend emerged from these studies; namely, all of them were held together by Hoogsteen base-pairing involving N7 and N6 or adenine, and none of them had the Watson-Crick base pairs involving the adenine N6 and N1 atoms. This led some investigators to suggest that the double helix might be held together by Hoogsteen pairing, which would involve protonation of the cytosine residues if two hydrogen bonds were used in connecting cytosine to the O6 and N7 of guanine. The calculated diffraction pattern of such a helix had many similarities to that predicted by a double helix held together by Watson-Crick base pairs, even though the fit was not good [1]. However, the question remained: What is the real structure of the double helix?

The Double Helix at atomic resolution

The first single-crystal structures of a double helix were solved in 1973 in my laboratory. This was before it was possible to synthesize and obtain oligonucleotides in significant quantities suitable for crystallographic experiments. However, we succeeded in crystallizing two dinucleoside phosphates, the RNA oligomers GpC [6] and ApU [53]. The structures are illustrated in Figure 5. The significant point in this analysis was that the resolution of the diffraction pattern was 0.8 Å: the atomic resolution allowed us to visualize not only the sugar phosphate backbone in the form of a double helix, but also the positions of ions and water molecules. It could be shown that extending the structure with all sugars in the C3' endo conformation using the symmetry of the two base pairs made it possible to generate RNA double helices that were quite similar to the structures that had been deduced from studies of double-helical fibers of RNA. The bond angles and distances from these structures provided the library of acceptable angles and distances and, in addition, gave rise to the nomenclature for identifying torsion angles in the sugar phosphate backbone.

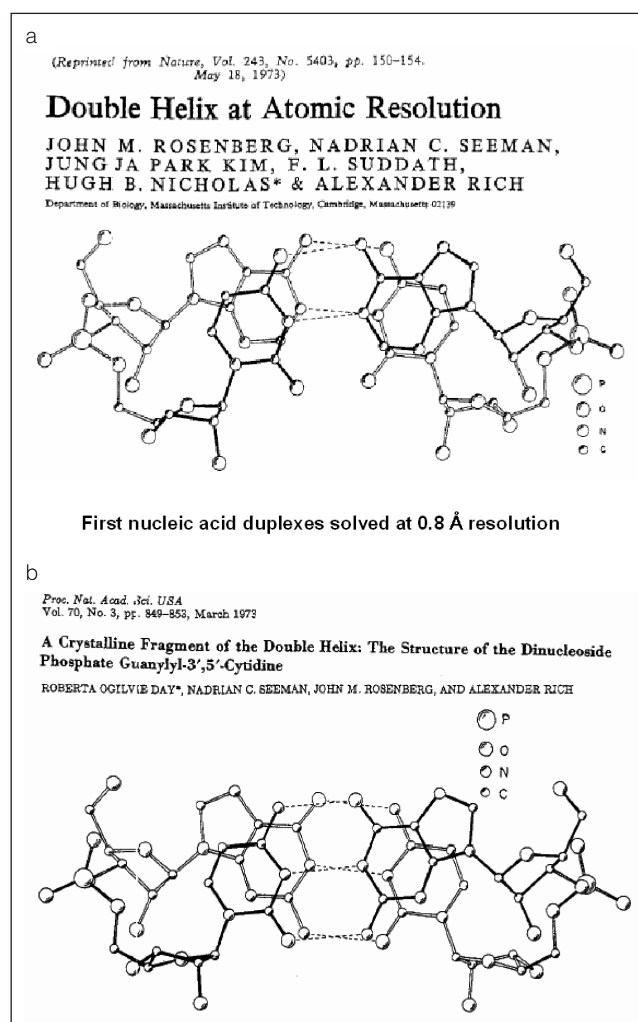

Figure 5. The 0.8 Å resolution double helical crystal structure of the dinucleoside phosphates (a) ApU as displayed on the cover of Nature magazine [53], and (b) GpC [6].

The GpC structure had the anticipated base pairs connected by three hydrogen bonds. However, the ApU structure showed for the first time that Watson-Crick base pairs formed when the molecule was constrained in a double he-

lix, as opposed to the Hoogsteen base pairs that were favored in the single-crystal complexes of adenine with uracil derivatives. I mailed preprints of these to several people, including Jim Watson. He phoned me, saying that after having read the ApU manuscript, he had his first good night's sleep in 20 years! This indicated to him that the uncertainty about the organization of the double helix was resolved. The significance of the double helix at atomic resolution was recognized by the editors of Nature who, in their "News and Views" commentary, called it the "missing link" and recognized that "the many pearls offered" helped resolve one of the big uncertainties in nucleic acid structure ("News and Views" [1973] Nature, 243: 114).

These structures capped the effort that I had started some 20 years earlier which had been initiated in earnest in 1956 with the recognition that poly rA and poly rU would form a double helix. Here, at last, was the demonstration at atomic resolution of the details of that structure. High-resolution crystallographic analysis of larger fragments of the double helix (DNA or RNA) did not emerge until almost a decade later with the availability of chemically synthesized and purified oligonucleotides, available in large enough quantities to permit single-crystal diffraction analysis.

A single-crystal X-ray structure of a hybrid helix did not appear until 1982 when my colleagues and I solved the structure of a DNA-RNA hybrid linked to double-helical DNA [63]. This was 22 years after the hybrid helix was first observed [46]. It showed that the dilemma of two different backbone conformations was resolved by having the DNA strand adopt the RNA duplex conformation. This had been inferred from fiber diffraction studies and has remained a constant feature, reflecting the relative conformational flexibility of the DNA backbone compared with the less flexible RNA strand.

tRNA structure

The ways in which a double helix can be incorporated into a complex folding were first seen in the structure of transfer RNA. Transfer RNA plays a central role in protein synthesis. Inside the ribosome it holds the growing polypeptide chain and interacts with the messenger RNA. There was intense interest in its three-dimensional structure. As methods for purifying tRNA improved during the 1960s, there was an increase in the number of attempts to form single crystals. This effort was very frustrating because it was easy to fail, and most people involved in the effort failed repeatedly. In 1968, working with Sung Hou Kim in my laboratory, we were able to obtain single crystals of E.coli $tRNA^{phe}$ [21]. Three other groups also obtained single crystals of various tRNAs in that same year, and all of these crystals were poor, in that they were somewhat disordered and the resolution was limited. Our earliest crystals diffracted to ~20 Å. By the next year we were able to get crystals that diffracted to 6 or 7 Å resolution, and a study of the three-dimensional Patterson function using the 12 Å data from the crystals of E.coli $tRNA^{Fmet}$ yielded approximate molecular dimensions of 80 x 25 x 35 Å [22]. These crystals represented progress of a sort, but at the same time, the frustration was great because they were not suitable for solving the structure of the molecule.

We spent the next two years looking at many different purified tRNA preparations and explored many different crystallization procedures. By 1971, we reached an exciting turning point: Yeast $tRNA^{phe}$ could be crystallized in a simple orthorhombic unit cell with a resolution of 2.3 Å [23]! These were the first crystals of tRNA suitable for analysis. The key event in making crystals of this resolution was the incorporation in the crystallization mix of spermine, a naturally occurring polyamine. The spermine bound specifically to yeast $tRNA^{phe}$ and stabilized it so that it made a high-resolution crystal. This was a significant discovery at the time. The stabilizing effect of spermine on yeast $tRNA^{phe}$ made it possible to form good crystals in other lattices as well. However, spermine would not necessarily stabilize all tRNA molecules in a similar way.

Analysis of the crystal diffraction pattern showed that it had a characteristic helical distribution of diffracting intensities when viewed in one direction but did not show a helical distribution when viewed at right angles. This was taken as evidence that short helical segments containing 4-7 base pairs were found in the molecule, a result entirely consistent with the cloverleaf folding of tRNA molecules postulated by Holley and colleagues after sequencing the first tRNA molecule [17]. This discovery opened the door to the ultimate solution of the structure of yeast $tRNA^{phe}$.

Tracing the polynucleotide backbone of yeast tRNAphe

Myoglobin was the first protein whose three-dimensional structure was solved. The structure was revealed at various levels of resolution. An important event was the tracing of the polypeptide chain which showed how the myoglobin molecule is organized as a series of ?-helical and single-stranded regions folded together. This was our first glimpse of how a protein molecule is folded, even though the high-resolution structure had not been completed.

The structure of yeast $tRNA^{phe}$ was revealed in a similar, gradual way. Crystallographic research moved more slowly in the early 1970s than today. Computers were primitive; advanced area detectors, cryo-crystallography, and synchrotron beams were things in the future. However, before work could continue, heavy-atom derivatives had to be discovered that would be useful for phasing the diffraction pattern of a crystalline nucleic acid molecule. This had never been done before, and it took considerable time to discover appropriate derivatives. Three different types of heavy atoms were developed containing platinum, osmium, or samarium ions. The osmium residue was very important since it was known to form complexes with ribonucleotides involving both the 2´ and 3´ hydroxyl groups. Only one pair of cis hydroxyl groups was found at the 3´ CCA end of the tRNA mol-

ecule. In order to gain some appreciation of the geometry, the structure of an osmium-adenosine complex was solved which enabled us to visualize the interaction [4]. The single osmium derivative in the tRNA crystal made it possible to identify the 3´ hydroxyl end of the tRNA chain [24]. The samarium ions were very useful, and they occupied more than one site. The platinum residue was only useful for 5.5 Å data. An interim electron density map at 5.5 Å [27] made it possible to uncover the external shape of portions of the molecule. However, the true shape of the molecule was not revealed until a map at 4 Å was visualized in 1973 [24].

At 4 Å resolution, peaks were seen throughout the electron density map that were due to the electron-dense phosphate groups. We knew a great deal about the distance constraints between adjacent phosphate groups in a polynucleotide chain, and this made it possible to look for peaks separated by 5 and 7 Å. Tracing the chain led to the discovery that the tRNA molecule had an unusual L-shape. The CCA acceptor double helix was colinear with the T pseudo-U helix, and it was almost at right angles to the anticodon double helical stem which is colinear with the dihydro U stem. The molecule had the shape of an "L," with the amino acid acceptor 3´ hydroxyl group at one end of the L and the anticodon loop at the other. At the corner of the L, there was a complex folding of the T pseudo-U and dihydro U loops. Figure 6 shows a perspective diagram of the L-shaped molecule.

The L-shaped folding of the tRNA polynucleotide chain was a dramatic and surprising discovery. Because of the constraints in the cloverleaf folding of the molecule, several models had been proposed for the folding of tRNA molecules. All of them were wrong. No one had anticipated that the molecule would organize in this fashion. Even at 4 Å resolution, this folding was compatible with much experimental data concerning tRNA molecules. For example, it was known that photoactivation of E. coli $tRNA^{val}$ resulted in the formation of a photo dimer involving the 4-thio-U residue in position 8 and the cytosine in position 13. In the 4 Å folding of the polynucleotide chain, these two bases were in close proximity, and the distance between the phosphate groups of these two residues was short enough to allow formation of the photo dimer [24].

The L-shaped tRNA molecule with the folding shown in Figure 6 is now a standard feature of molecular biology, having been found in virtually all tRNA molecules, even when they are complexed to aminoacyl synthetase enzymes. The significance of the folding is twofold. First, it revealed that the 3´ acceptor end is over 70 Å away from the anticodon loop, which has implications for understanding the interaction between tRNA molecules and tRNA aminoacyl synthetases. Second and most important, it suggested that the interaction of the tRNA molecules with the message occurs at one end of the L, whereas the segment responsible for forming the peptide bond is considerably removed from the site. This makes it possible to have great specificity with many interactions at either end of the molecule due to this separation. These features have been incorporated into our present view of protein synthesis in interpreting the three-dimensional structure of ribosomes and the movement of tRNA molecules inside the ribosome. Indeed, a proposal was made regarding the movement of tRNA molecules in the ribosome using the 4 Å folding [47], pointing out the necessity for bending the mRNA between the two adjacent codons binding two tRNA molecules. A 45º kink is seen in the mRNA in the three-dimensional structure of the ribosome [69].

Today, we are accustomed to seeing a variety of complex ribonucleotide molecules in which double-helical segments and single-chain segments are juxtaposed to make complex structures with a variety of functions, especially in ribozymes. However, the beginning of our understanding of the manner in which complex polynucleotide chains can fold started with this first tracing of yeast $tRNA^{phe}$ visualized at 4 Å resolution.

This tracing was seen in more detail a year later in our 3 Å analysis of the folding of yeast $tRNA^{phe}$ in the orthorhombic lattice (Fig. 6) [25, 26]. Simultaneously, Aaron Klug and colleagues published the 3 Å structure of the same spermine-stabilized yeast $tRNA^{phe}$ in a monoclinic lattice [52]. This further confirmed the L-shape folding of the polynucleotide chain, even though the lattice was different. These 3 Å structures were very similar and revealed in great detail the manner in which base-pairing of nucleotides, both in the double-helical regions and in the single-stranded regions, stabilizes the three-dimensional fold of the molecule. The folding was held together by a variety of hydrogen-bonding interactions, including many in the nonhelical regions of the molecule. These hydrogen-bonding interactions included the formation of triplexes and a variety of interactions beyond that seen in Watson-Crick base-pairing. Recognition of the importance of these alternative types of hydrogen bonds explained why the model builders of that period, trying to anticipate the structure of tRNA, were all incorrect in their conclusions. The main reason was that they relied excessively on Watson-Crick base-pair interactions and did not recognize the stabilizing effect of many other types of hydrogen bonds.

The L-shaped folding was predicted to be a general conformation found in all tRNA molecules [26]. Subsequent work has amply verified the relative constancy of the hydrogen-bonding networks [40, 41].

A quarter-century later: a DNA surprise

It was not until the late 1970s that the development of DNA synthesis made it possible to obtain significant quantities of oligonucleotides and carry out single crystal X-ray diffraction studies which could "prove" the structure. In 1978 I met a Dutch organic chemist Jacques Van Boom, who could synthesize DNA oligomers. He made $d(CG)_3$, and Andy Wang crystallized it and discovered it diffracted to 0.9 Å resolution. Heavy atoms were used to solve the structure which revealed a surprising left-handed double helix with two antiparallel chains held together by Watson-Crick base pairs

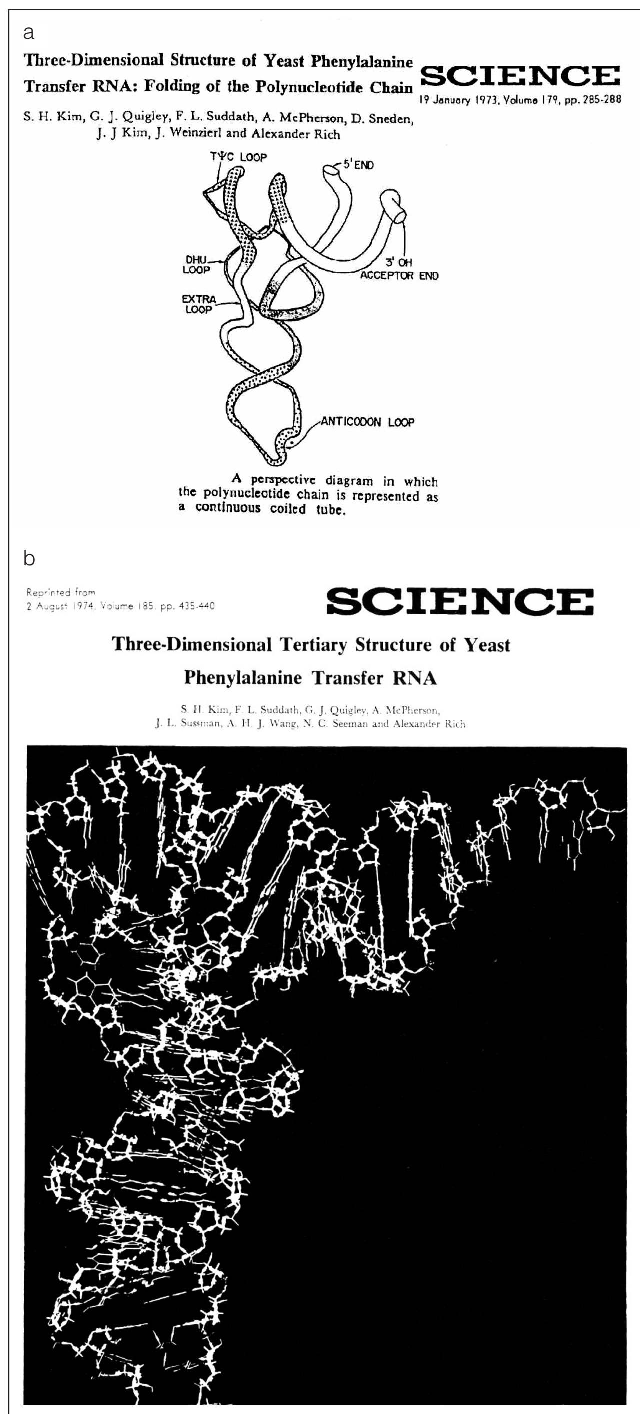

Figure 6. (a) At 4 Å resolution the fold of the tRNA^{phe} chain could be visualized, as shown in this perspective diagram [24]. (b) The 3 Å tRNA^{phe} structure reveals the complete interactions of the L-shaped molecule, as shown on the cover of Science [25, 26].

[62]. Every other base had rotated around the glycosyl bonds so that the bases alternated in anti and syn conformations along the chain. The zig-zag arrangement of the backbone (hence, Z-DNA) was different from the smooth, continuous coil seen in B-DNA (Fig. 7). The general response to this unusual structure was amazement, coupled with skepticism.

The relationship between Z-DNA and the more familiar right-handed B-DNA began to be apparent from the earlier work of Pohl and Jovan (1972) [39] who showed that the ultraviolet circular dichroism of poly (dG-dC) nearly inverted in 4M sodium chloride solution. The suspicion that this was due to a conversion from B-DNA to Z-DNA was confirmed by examining the Raman spectra of these solutions and the Z-DNA crystals [59]. The conversion to left-handed Z-DNA was associated with a "flipping over" of the base pairs so that they were upside down in their orientation relative to what would be found in B-DNA. Sequences that most readily converted had alternations of purines and pyrimidines, especially alternations of C and G [49] and alternations of CA on one strand and TG on the other strand [36].

This discovery stimulated a burst of research from a large number of chemists who were very interested in studying DNA conformational changes. It tended to leave most biologists rather puzzled, since the ionic conditions suitable for stabilizing Z-DNA were very far from those present in a cell. This view changed somewhat with the discovery that negative supercoiling would also stabilize Z-DNA [38]. Supercoiling was known to be a part of biological systems, and it suggested a connection between this alternative conformation and biological phenomena.

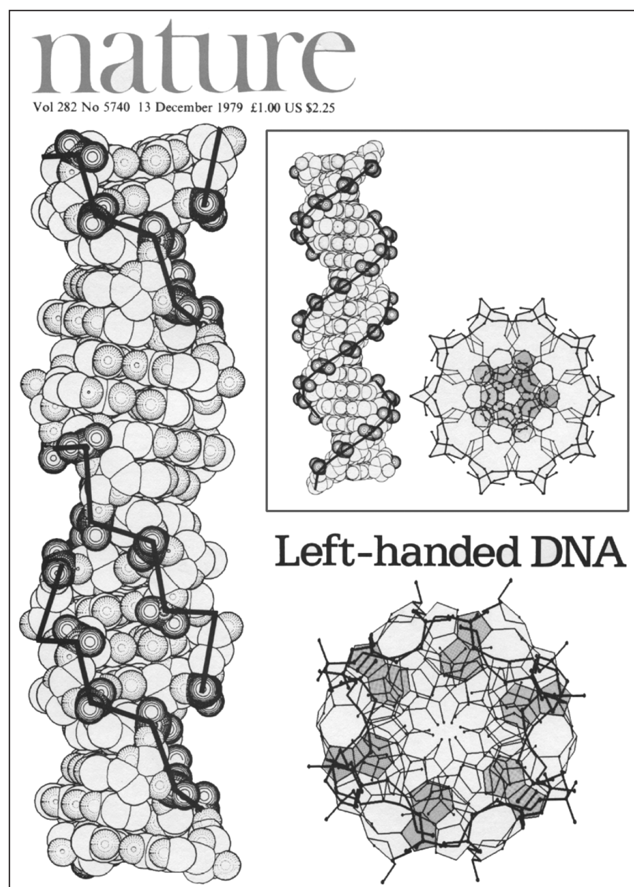

Figure 7. A diagram showing a comparison between B-DNA and Z-DNA with solid lines running from phosphate to phosphate, as shown on the cover of Nature [62].

A Z-DNA biology?

Research work on the biology of Z-DNA progressed very slowly. By the mid-1980s after several years of research in

which nothing definitive emerged about Z-DNA, most biologists were very skeptical about its role. Many felt that Z-DNA was a non-functional conformational phenomenon. My conviction was very simple. Here was an alternative DNA conformation, and I felt it likely that it would be used because evolution is opportunistic. The challenge was to find out how it was used.

The first indications of a biological role for Z-DNA came from immunological work. In collaboration with David Stollar we found that, unlike B-DNA, Z-DNA is highly antigenic, yielding polyclonal and monoclonal antibodies [34]. Antibodies to Z-DNA provided a useful tool for characterizing chromosome organization. They bound specifically to the interband regions of the Drosophila polytene chromosomes; the binding was particularly strong in the puff regions, the sites of enhanced transcriptional activity [35]. Ciliated protozoa have two nuclei: the macronucleus, which is the site of transcription, and the micronucleus, which contains DNA involved in sexual reproduction. AntiZDNA antibodies stained the macronucleus of the ciliated protozoan Stylonychia, but not its micronucleus [30]. These were the first data to suggest a connection between Z-DNA and transcriptional activity.

An important advance came with the work of Liu and Wang (1987) [31] on the interaction of RNA polymerase with DNA during transcription. They pointed out that the moving complex does not rotate around helical DNA, but instead plows straight through. Because the ends of the DNA molecule are fixed, the DNA behind the moving polymerase was unwound and subjected to negative torsional strain, while positive torsional strain developed in front. Further evidence came later from the work of P.Shing Ho and colleagues, who found a high concentration of sequences favoring Z-DNA formation near the transcription start site [54]. To study the association with transcription more directly, I collaborated with Burkhardt Wittig and colleagues using a technique developed by Peter Cook at Oxford. Mammalian cells were encapsulated in agarose microbeads; mild detergent treatment lysed the cytoplasmic membrane, permeabilizing the nuclear membrane but left the nucleus otherwise intact. The resulting “entrapped” nuclei replicated DNA at nearly the in vivo rate, and they were able to carry out transcription [19]. Using biotinylated monoclonal antibodies against Z-DNA, the level of Z-DNA was shown to be regulated by torsional strain [66]. An increase in transcriptional activity of the embedded nuclei resulted in a parallel increase in the amount of ZDNA [67]. Using a UV laser pulse for protein-DNA cross linking, the biotinylated anti-Z-DNA antibodies were linked to DNA. This made it possible to isolate DNA restriction fragments bound to the antibody. With cultured human cells, three regions upstream of the c-myc gene formed ZDNA when c-myc was expressed. However, these regions quickly reverted to B-DNA upon switching off c-myc transcription [68]. Nonetheless, the actin gene control retained its Z-DNA at all times.

The picture which then emerged was that the negative torsional strain induced by the movement of RNA polymerase stabilized Z-DNA formation near the transcription start site. Even though topoisomerases tried to relax the DNA, the continued movement of RNA polymerases generated more negative torsional strain than the topoisomerases could relax. However, upon cessation of transcription, topoisomerases rapidly converted it back to the right-handed B conformation. Thus, Z-DNA was seen as a metastable conformation, forming and disappearing depending upon physiological activities.

Binding proteins for a left-handed helix

If Z-DNA were to have biological functions, it seemed highly likely that a class of proteins should bind to it specifically. The challenge was to isolate such proteins that bound selectively to Z-DNA with high affinity. The first successful method used a gel shift assay with radioactive-labeled, chemically stabilized Z-DNA in the presence of a ~20,000-fold excess of B-DNA and single-stranded DNA [13]. A Z-DNA-binding protein was found to be a nuclear RNA editing enzyme [14] called double-stranded RNA adenosine deaminase (or ADAR1). This enzyme acts on double-stranded segments formed in pre-mRNA, binding to the duplex and selectively deaminating adenosine, yielding inosine. Ribosomes interpret inosine as guanine. Thus, ADAR1 can alter the amino acid sequence of a DNA-encoded protein. The functional properties of the edited protein (with the amino acid alteration) are often different from those of the unedited protein. The editing enzyme is found in all metazoa; it acts to increase the functional diversity of proteins transcribed from a given locus [2].

Proteolytic dissection of the editing enzyme ADAR1 led to a domain from the N-terminus called $Z?_{ADAR1}$ [15]. $Z?_{ADAR1}$ was found to contain all of the Z-DNA binding properties associated with the editing enzyme, and it bound Z-DNA tightly with a low nanomolar K_d.

The purified $Z?_{ADAR1}$ domain was co-crystallized with $d(CG)_3$ and the structure solved at 2.1Å resolution [55]. It revealed that the DNA was in a form virtually identical to that seen in the first Z-DNA crystal [62]. The 70amino acid binding domain was found to adopt a helix-turn-helix ?-sheet motif (winged helix) in which the recognition helix and the β-sheet bound to five successive phosphate groups in the zigzag backbone of Z-DNA, and it recognized the syn conformation of guanine (Fig. 8).

It is possible that the Z-DNA binding domain of ADAR1 targets Z-DNA forming regions of some transcriptionally active genes, as only they have Z-DNA. $Z\alpha_{ADAR1}$ appears to be active in vivo in the editing of certain transcripts where it may target the gene [16]; however, its role in RNA editing is not resolved.

The co-crystal structure of $Z\alpha_{ADAR1}$ and Z-DNA made it possible to identify those amino acids important for Z-DNA recognition. A computer search rapidly revealed other proteins with similar sequence motifs. One is the protein DLM1 which is up regulated in tissues in contact with tumors and is

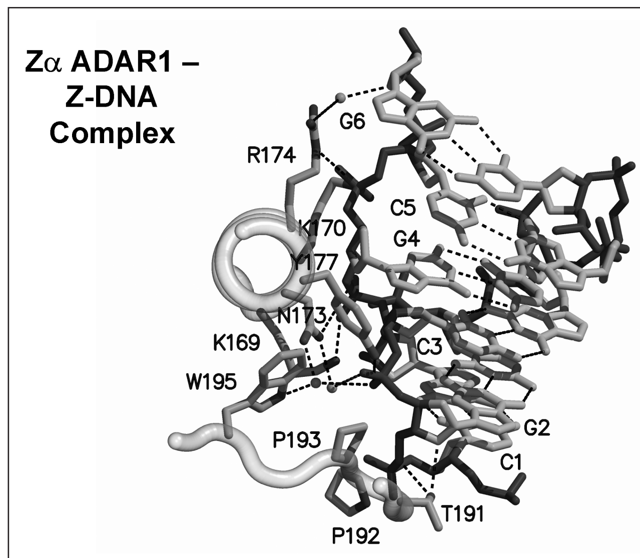

Figure 8. A portion of the crystal structure of Z?$_{ADAR1}$ complexed to Z-DNA. The view is down the recognition helix on the left with a portion of the ?- sheet below it. Z-DNA is on the right with hydrogen bonds shown as dashed lines [55].

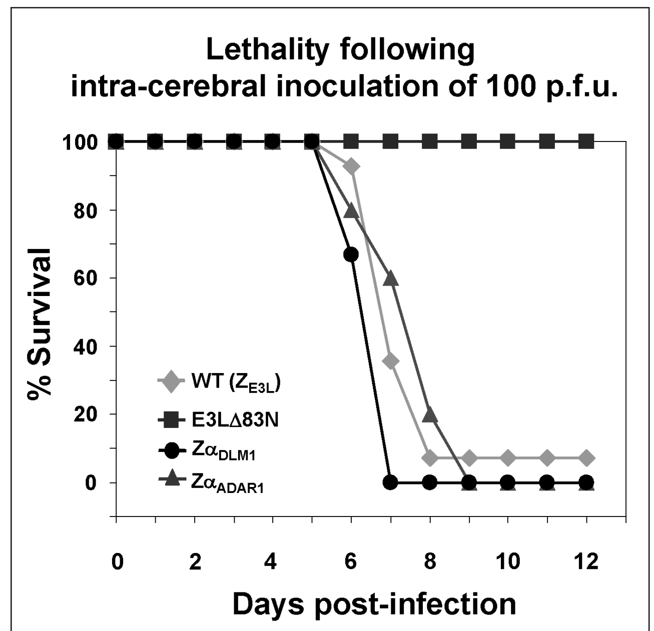

Figure 9. Lethality in mice following injection with vaccinia wild type virus (WT), an N-terminal deletion of 87 amino acids (?1-83) without Z_{E3L} or chimeric virus with Z?$_{ADAR1}$ or Z?$_{DLM1}$ replacing Z_{E3L} (Yang et al. 2003).

also interferon-induced. The co-crystal structure of a domain of DLM1 (Z?$_{DLM1}$) and d(CG)$_3$ was solved at a resolution of 1.85 Å, and it showed that this second protein domain recognizes Z-DNA in a manner very similar to that found with $Z\alpha_{ADAR1}$, but with a few variations [56]. This second structure clearly indicated that a family of such proteins exists.

Viruses that use Z-DNA binding proteins

Another member of this family of proteins is E3L, found in poxviruses such as vaccinia. These large DNA viruses reside in the cytoplasm of cells and produce a number of proteins that help to overcome the interferon response of the host cell. E3L is a 25 Kd protein that is necessary for pathogenicity and its N-terminal domain has homologies with Z-DNA binding domains. When vaccinia virus is given to a mouse, the mouse dies in about a week. However, in a virus that has a mutated or missing E3L, it is no longer pathogenic for the mouse, even though the virus can still reproduce in cell culture [3]. To demonstrate the pathogenicity of the vaccinia virus in the mouse and its relationship to possible Z-DNA binding activities of E3L, a collaboration was set up with Bertram Jacobs. Chimeric viruses were created in which the N-terminal domain of vaccinia E3L (Z_{E3L}) was removed, and either the domains Z?$_{ADAR1}$ or Z?$_{DLM1}$ were inserted. In carrying out these domain swaps, a little more than a dozen amino acids in the domain remained unchanged, but over 50 other amino acids were changed. Nonetheless, the chimeric viruses were as pathogenic for mice as the wild type [28] (Fig. 9). Other experiments were carried out in which mutations in the chimeric virus that weakened Z-DNA binding were also shown to weaken pathogenicity. Similar mutations in the wild type weakened mortality. Loss of Z-DNA binding led to loss of pathogenicity. It is likely that the Z_{E3L} domain binds to Z-DNA formed near the transcription start site of certain genes, thereby impairing the anti-viral response of the host cell. This is a new example of the way that viruses seek to exploit features of the host cell in order to overcome the host defense mechanisms.

A small molecule or drug can probably be made that will bind to the Z-DNA binding pocket of the E3L molecule. This drug should prevent mice from dying when infected with vaccinia virus. It may also be active in humans to prevent untoward effects due to vaccination. More significant is the fact that the E3L protein of the closely related variola virus, the agent of smallpox, is virtually identical to the vaccinia E3L [28]. Hence, such a drug binding to E3L may develop into a treatment of smallpox.

To my great surprise, work on Z-DNA and its binding proteins has led us to the possibility of developing a therapy for certain viral diseases, including smallpox. This is a striking example of serendipity in scientific research. However, it also illustrates how fundamental knowledge of nucleic acid molecular structure can lead to practical therapies for human diseases.

References

[1] Arnott S., Wilkins M.H.F., Hamilton L.D., and Langridge R. 1965. Fourier synthesis studies of lithium DNA. III. Hoogsteen models. J. Mol. Biol. 11: 391.

[2] Bass B.L. 2002. RNA editing by adenosine deaminases that act on RNA. Annu. Rev. Biochem., 71: 817.

[3] Brandt T.A. and Jacobs B.L. 2001. Both carboxy-and amino-terminal domains of the vaccinia virus interferon resistance gene, E3L are required for pathogenesis in a mouse model. J. Virol., 75:850.

[4] Conn J.F., Kim J.J., Suddath F.L., Blattman P., and

Rich A. 1974. Crystal and molecular structure of an osmium bispyridine ester of adenosine. J. Am. Chem. Soc. 96: 7152.

[5] Davies D.R. and Rich A. 1958. The formation of a helical complex between polyinosinic acid and polycytidylic acid. J. Am. Chem. Soc. 80: 1003.

[6] Day R.O., Seeman N.C., Rosenberg J.M., and Rich A. 1973. A crystalline fragment of the double helix: The structure of the dinucleoside phosphate guanylyl-3´,5´-cytidine. Proc. Natl. Acad. Sci. 70: 849.

[7] Doty P., Marmur J., Eigner J., and Schildkraut C. 1960. Strand separation and specific recombination in deoxyribonucleic acids: Physical chemical studies. Proc. Natl. Acad. Sci. 46: 461.

[8] Felsenfeld G. and Rich A. 1957. Studies on the formation of two- and three-stranded poly ribonucleotides. Biochim. Biophys. Acta 26: 457.

[9] Felsenfeld G., Davies D.R., and Rich A. 1957. Formation of a three-stranded polynucleotide molecule. J. Am. Chem. Soc. 79: 2023.

[10] Furth J.J., Hurwitz J., and Goldmann M. 1961. The directing role of DNA in RNA synthesis. Biochem. Biophys. Res. Commun. 4: 362.

[11] Grunberg-Manago M., Ortiz P.J., and Ochoa S. 1955. Enzymatic synthesis of nucleic acidlike polynucleotides. Science 122: 907.

[12] Hall B.D. and Spiegelman S. 1961. Sequence complementarity of T2-DNA and T2-specific RNA. Proc. Natl. Acad. Sci. 47: 137.

[13] Herbert A.G. and Rich A. 1993. A method to identify and characterize Z-DNA binding proteins using a linear oligodeoxynucleotide. Nucl. Acids Res., 21:2669.

[14] Herbert A., Lowenhaupt K., Spitzner J., and Rich A. 1995. Chicken double-stranded RNA adenosine deaminase has apparent specificity for Z-DNA. Proc. Nat. Acad. Sci. USA, 92:7550.

[15] Herbert A., Alfken J., Kim Y.-G., Saira Mian I., Nishikura K., and Rich A. 1997. A Z-DNA binding domain present in the human editing enzyme, double-stranded RNA adenosine deaminase. Proc. Natl. Acad. Sci. USA, 94: 8421.

[16] Herbert A. and Rich A. 2001. Role of binding domains for dsRNA and Z-DNA in the invivo editing of minimal substrates by ADAR1. Proc. Natl. Acad. of Sci., USA, 98:12132.

[17] Holley R.W., Apgar J., Everett G.A., Madison J.T., Marguisse M., Merrill S.H., Penwick J.R., and Zamir A. 1965. Structure of a ribonucleic acid. Science 147: 1462.

[18] Hoogsteen K. 1959. The crystal and molecular structure of a hydrogen-bonded complex between 1-methylthymine and 9-methyladenine. Acta Crystallogr. 12: 822.

[19] Jackson D.A., Yuan J., and Cook P.R. 1988. A gentle method for preparing cyto- and nucleo-skeletons and associated chromatin. J. of Cell Sci., 90:365.

[20] Katz L., Tomita K., and Rich A. 1965. The molecular structure of the crystalline complex ethyladenine: Methyl-bromouracil. J. Mol. Biol. 13: 340.

[21] Kim S.-H. and Rich A. 1968. Single crystals of transfer RNA: An X-ray diffraction study. Science 162: 1381.

[22] Kim S.-H. and Rich A. 1969. Crystalline transfer RNA: The three-dimensional Patterson function at 12-angstrom resolution. Science 166: 1621.

[23] Kim S.-H., Quigley G., Suddath F.L., and Rich A. 1971. High resolution X-ray diffraction patterns of tRNA crystals showing helical regions of the molecule. Proc. Natl. Acad. Sci. 68: 841.

[24] Kim S.-H., Quigley G.J., Suddath F.L., McPherson A., Sneden D., Kim J.J., Weinzierl J., and Rich A. 1973. Three-dimensional structure of yeast phenylalanine transfer RNA: Folding of the polynucleotide chain. Science 179: 285.

[25] Kim S.-H., Suddath F.L., Quigley G.J., McPherson A., Sussman J.L., Wang A.H.-J., Seeman N.C., and Rich A. 1974a. Three-dimensional tertiary structure of yeast phenylalanine transfer RNA. Science 185: 435.

[26] Kim S.-H., Sussman J.L., Suddath F.L., Quigley G.J., McPherson A., Wang A.H.-J., Seeman N.C., and Rich A. 1974b. The general structure of transfer RNA molecules. Proc. Natl. Acad. Sci. 71: 4970.

[27] Kim S.-H., Quigley G., Suddath F.L., McPherson A., Sneden D., Kim J.J., Weinzierl J., Blattmann P., and Rich A. 1972. Three-dimensional structure of yeast phenylalanine transfer RNA: Shape of the molecule at 5.5 Å resolution. Proc. Natl. Acad. Sci. 69: 3746.

[28] Kim Y.-G., Muralinath M., Brandt T., Pearcy M., Hauns K., Lowenhaupt K., Jacobs B.L., and Rich A. 2003. A role for Z-DNA binding in vaccinia virus pathogenesis. Proc. Nat'l. Acad. Sci. USA, 100:6974.

[29] Kornberg A., Lehman I.R., Bessman M.J., and Simms E.S. 1956. Enzymic synthesis of deoxyribonucleic acid. Biochim. Biophys. Acta 21: 197.

[30] Lipps H.J., Nordheim A., Lafer E.M., Ammermann D., Stollar B.D., and Rich A. 1983. Antibodies against Z-DNA react with the macronucleus but not the micronucleus of the hypotrichous ciliate Stylonychia mytilus. Cell 32:435.

[31] Liu L.F. and Wang J.C. 1987. Supercoiling of the DNA template during transcription. Proc. Natl. Acad. Sci. USA 84: 7024.

[32] Marmur J. and Lane D. 1960. Strand separation and specific recombination in deoxyribonucleic acids: Biological studies. Proc. Natl. Acad. Sci. 46: 453.

[33] Mathews F.S. and Rich A. 1964. The molecular structure of a hydrogen bonded complex of N-ethyl adenine and N-methyl uracil. J. Mol. Biol. 8: 89.

[34] Moller A., Gabriels J.E., Lafer E.M., Nordheim A., Rich A., and Stollar B.D. 1982. Monoclonal antibodies recognize different parts of Z-DNA. J. Biol. Chem. 257:12081.

[35] Nordheim A., Pardue M.L., Lafer E.M., Moller A., Stollar B.D., and Rich A. 1981. Antibodies to left-handed Z-DNA bind to interband regions of Drosophila polytene chromosomes. Nature 294:417.

[36] Nordheim A. and Rich A. 1983. The sequence (dC-

$dA)_n$?$(dG-dT)_n$ forms left-handed Z-DNA in negatively supercoiled plasmids. Proc. Natl. Acad. Sci. USA 80:l82l.

[37] Pauling L. and Corey R.B. 1956. Specific hydrogen-bond formation between pyrimidines and purines in deoxyribonucleic acids. Arch. Biochem. Biophys. 65: 164.

[38] Peck L.J., Nordheim A., Rich A., and Wang J.C. 1982.Flipping of cloned d(pGpG)n?d(pCpG)n DNA sequences from right to lift-handed helical structure by salt, Co(III), or negative supercoiling.Proc. Natl. Acad. Sci. USA. 79:4560.

[39] Pohl F.M. and Jovin T.M. 1972.Salt-induced co-operative conformational change of a synthetic DNA: equilibrium and kinetic studies with poly (dG-dC). J. Mol. B iol. 67:375.

[40] Quigley G.J. and Rich A. 1976. Structural domains of a transfer RNA molecule. Science 194: 796.

[41] Quigley G.J., Wang A.H.-J., Seeman N.C., Suddath F.L., Rich A., Sussman J.L., and Kim S.H. 1975. Hydrogen bonding in yeast phenylalanine transfer RNA. Proc. Natl. Acad. Sci. 72: 4866.

[42] Rich A. 1957. The structure of synthetic polyribonucleotides and the spontaneous formation of a new two-stranded helical molecule. In The chemical basis of heredity (ed. W.D. McElroy and B. Glass), p. 557. Johns Hopkins University Press, Baltimore, Maryland.

[43] Rich A. 1958a. Formation of two- and three-stranded helical molecules by polyinosinic acid and polyadenylic acid. Nature 181: 521.

[44] Rich A. 1958b. The molecular structure of polyinosinic acid. Biochim. Biophys. Acta 29: 502.

[45] Rich A. 1959. An analysis of the relation between DNA and RNA. Ann. N.Y. Acad. Sci. 81: 709.

[46] Rich A. 1960. A hybrid helix containing both deoxyribose and ribose polynucleotides and its relation to the transfer of information between the nucleic acids. Proc. Natl. Acad. Sci. 46: 1044.

[47] Rich A. 1974. How transfer RNA may move inside the ribosome. In Ribosomes (ed. M. Nomura et al.), p. 871. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

[48] Rich A. and Davies D.R. 1956a. A new two-stranded helical structure: Polyadenylic acid and polyuridylic acid. J. Am. Chem. Soc. 78: 3548.

[49] Rich A., Nordheim A., and Wang A.H.-J.1984.The chemistry and biology of left-handed Z-DNA. Annu. Rev.Biochem. 53:791.

[50] Rich A. and Watson J.D. 1954a. Some relations between DNA and RNA. Proc. Natl. Acad. Sci. 40: 759.

[51] Rich A. and Watson J.D. 1954b. Physical studies on ribonucleic acid. Nature 173: 995.

[52] Robertus J.D., Ladner J.E., Finch J.T., Rhodes D., Brown R.S., Clark B.F.C., and Klug A. 1974. Structure of yeast phenylalanine tRNA at 3 Å resolution. Nature 250: 546.

[53] Rosenberg J.M., Seeman N.C., Kim J.J.P., Suddath F.L., Nicholas H.B., and Rich A. 1973. Double helix at atomic resolution. Nature 243: 150.

[54] Schroth G.P., Chou P.-J., and Ho P.S.1992.Mapping Z-DNA in the human genome: computer aided mapping reveals a non-random distribution of potential Z-DNA forming sequences in human genes.J. Biol. Chem., 267:11846.

[55] Schwartz T, Rould MA, Lowenhaupt K, Herbert A, Rich A. 1999. Crystal structure of the Z? domain of the human editing enzyme ADAR1 bound to left-handed Z-DNA.Science, 284:1841.

[56] Schwartz T., Behlke J., Lowenhaupt K., Heinemann U., and Rich A. 2001.Structure of the DLM-1-Z-DNA complex reveals a conserved family of Z-DNA-binding proteins.Nature Struct.Biol., 8:761.

[57] Sobell H.M., Tomita K., and Rich A. 1963. The crystal structure of an intermolecular complex containing a guanine and cytosine derivative. Proc. Natl. Acad. Sci. 49: 885.

[58] Tener G.M., Khorana H.G., Markham R., and Pol E.H. 1958. Studies on polynucleotides. II. The synthesis and characterization of linear and cyclic thymidine oligonucleotides. J. Am. Chem. Soc. 80: 6223.

[59] Thamann T.J., Lord R.C., Wang A.H.-J., and Rich A. 1981.High salt form of poly (dG-dC)?Poly(dG-dC) is left handed Z-DNA: Raman spectra of crystals and solutions. Nuc. Acids Res. 9:5443.

[60] Tomita K., Katz L., and Rich A. 1967. Crystal structure of the intermolecular complex 9-ethyladenine: 1-methyl-5-fluorouracil. J. Mol. Biol. 30: 545.

[61] Voet D. and Rich A. 1970. The crystal structures of purine, pyrimidines and their intermolecular complexes. Prog. Nucleic Acid Res. Mol. Biol. 10: 183.

[62] Wang A.H.-J., Quigley G.J., Kolpak F.J., Crawford J.L., van Boom J.H., van der Marel G., and Rich A. 1979. Molecular structure of a left-handed double helical DNA fragment at atomic resolution. Nature 282:680.

[63] Wang A.H.-J., Fujii S., van Boom J.H., van der Marel G.A., van Boeckel S.A.A., and Rich A. 1982. Molecular structure of r(GCG)d(TATACGC): A DNA-RNA hybrid helix joined to double helical DNA. Nature 299: 601.

[64] Warner R.C. 1956. Ultraviolet spectra of enzymatically synthesized polynucleotides. Fed. Proc. 15: 379.

[65] Watson J.D. and Crick F.H.C. 1953. A structure for deoxyribose nucleic acid. Nature 171: 738.

[66] Wittig B., Dorbic T., and Rich A.1989.The level of Z-DNA in metabolically active, permeabilized mammalian cell nuclei is regulated by torsional strain. J. Cell. Biol. 108:755.

[67] Wittig B., Dorbic T., and Rich A. 1991.Transcription is associated with Z-DNA formation in metabolically active permeabilized mammalian cell nuclei. Proc. Natl. Acad. Sci. USA. 88: 2259.

[68] Wittig B., Wolfl T., Dorbic T., Vahrson W., and Rich A.1992. Transcription of human c-myc in permeabilized nuclei is associated with formation of Z-DNA in three discrete regions of the gene. EMBO J. 11:4653.

[69] Yusupova G.Z., Yusupov M.M., Cate J.H.D., and Noller H.F. 2001. The path of messenger RNA through the ribosome. Cell 106: 233.