Z-DNA Moves Toward "Real Biology"

Once considered to be an oddity of no particular significance, this unusual structure is now showing up in DNA segments that control gene expression

Just a few years ago, Alexander Rich of the Massachusetts Institute of Technology (MIT) was nearly alone in his studies of Z-DNA. Most biologists viewed this structure as an oddity with no particular significance in the real world (*Science*, 4 December 1981, p. 1108). This situation is changing now as Z-DNA seems more and more likely to play a role in controlling which genes are turned on and which are not. "It's becoming a much more lively field, and that's good," Rich says. "It makes it much more interesting. Now a number of people are beginning to address the question of how might Z-DNA be working in their system." *

Rich discovered Z-DNA 4 years ago when he was trying to settle the structure of conventional B-DNA. This is the DNA structure proposed 30 years ago by James Watson and Francis Crick. But, despite its widespread acceptance, the B structure was based on x-ray diffraction studies of fibers with limited resolution.

Rich and his associates Andrew Wang, Gary Quigley, Francis Kolpak, James Crawford, and J. H. van Boom decided to look at single crystals of a synthetic DNA molecule consisting of just two of the four DNA bases, guanine and cytosine, in alternating sequence. They got extremely good resolution—0.9 angstrom. "This is atomic resolution. Every atom shows up, as do all of the ions and water molecules," says Rich. But when they did the crystallography, "To our surprise, what emerged was a different molecule, left-handed in its twist." This is the molecule that Rich dubbed Z-DNA.

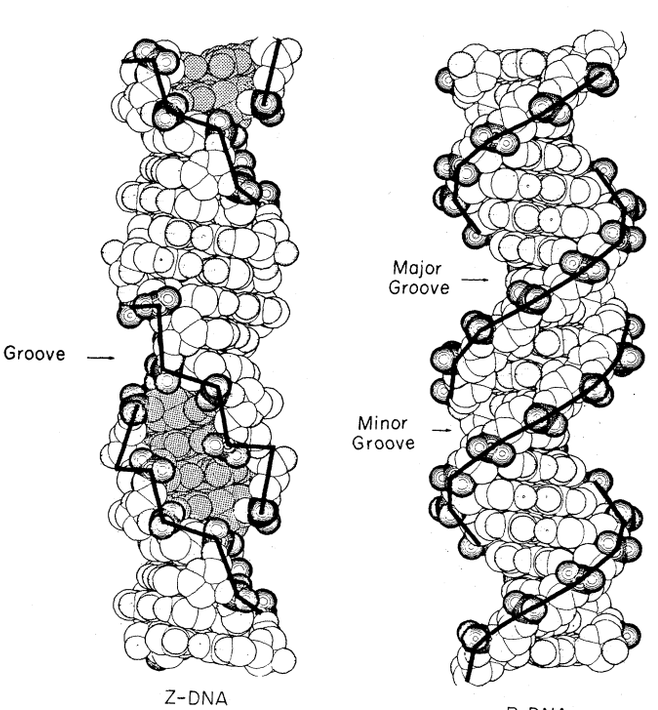
Rich says he chose the name Z-DNA for two reasons. First, there already were A, B, and C forms of the double helix and since Z is so different from them, Rich felt it appropriate to start at the other end of the alphabet. The second reason is that Z can stand for zigzag. The backbone of Z-DNA, unlike that of the B form, zigzags down the molecule. Not only is Z-DNA left-handed, whereas B-DNA is a right-handed helix, but the bases of Z-DNA are also turned around or flipped over. "We didn't know that this kind of change could take place in a polynucleotide and have it still remain a double helix," Rich says.

Among the first questions that occurred to the MIT researchers was, How can the Z-DNA structure be stabilized? They now know, based on their own work and that of investigators in other laboratories, that Z-DNA forms from sequences that largely contain alternating purines and pyrimidines. The reason for this sequence-dependence, Rich and Wang discovered, is that water molecules in the deep groove of the Z-DNA are disordered around sequences with stretches of adenine-thymine base pairs. That disordering contributes to instability of the Z conformation.

Once you have an appropriate sequence for forming Z-DNA, it still will prefer to form B-DNA because this structure requires slightly less energy. Rich and his associates found, however, that there are several ways to hold a DNA sequence in the Z conformation. The first is with ions such as polyamines, or with methyl groups on the 5 position of cytosine-guanine sequences. This observation about the methyl groups, first made by Michael Behe and Gary Felsenfeld at the National Institutes of Health, is of great interest because when certain cytosine-guanine sequences of eukaryotic DNA are methylated, genes nearby are inactive and when they are not methylated, the genes are active. The hint is that these sequences might be using changes between the Z and B forms as a switch for gene activity.

Also of great interest is Rich's discovery that certain proteins not only bind specifically to Z-DNA and not to B-DNA but can flip the B structure into Z and hold it there. Working with Alfred Nordheim, Paul Tesser, Fernando Azorin, Young Kwon, and Barry Rosen, Rich finds these proteins in a diverse group of cells, including cells from the fruit fly *Drosophila*, human cancer cells, wheat germ cells, and cells from the bacterium *Escherichia coli*. The Z-DNA–binding proteins vary in their abundance in these different cells. The wheat germ cells, for example, have high concentrations of the proteins, whereas the concentration is low in *E. coli*.

These Z-DNA–binding proteins include large proteins of molecular weight 70,000 to 150,000. Rich says he is particularly interested in these because large proteins may help control gene activity. The DNA in eukaryotic cells is coated with proteins, but most of these are



Z-DNA and B-DNA

The left-handed helix of Z-DNA is not simply a left-handed version of the well-known B-DNA but has an entirely different conformation. [A. Rich]

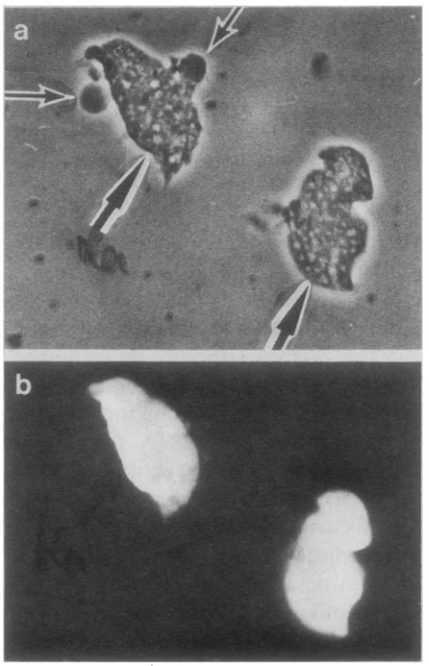
*Rich reviewed progress in Z-DNA research in the National Institutes of Health 2nd annual DeWitt Stetten, Jr., lecture, 5 October 1983.

histones—small proteins of molecular weight 10,000 to 15,000 that do not seem to recognize specific DNA sequences. Groups of histones cluster together, forming structures like spools and the DNA wraps around them. But, says Rich, "there are also associated with chromatin a number of proteins that are not well understood and play an important role in transcription. It would not surprise me to find that the Z-binding proteins are of this type."

Another way to stabilize the Z conformation is by negative supercoiling. Rich explains: "DNA usually is underwound. If you have a circle of DNA, the number of turns in it is less than what you would have if the molecule were nicked and relaxed. Suppose you have a 4000 base pair plasmid. Roughly speaking, you should have one turn of the helix for every 10 base pairs, so that you would then expect 400 turns of the helix. What you find is that, typically, the plasmid will have 380 turns—it is underwound by 20 turns. There are enzymes that produce this negative supercoiling. The important thing is that if you have that kind of negative supercoiling strain energy, you can relax the molecule if you flip a portion of it into Z-DNA. Negative supercoiling tends to stabilize Z-DNA." (This has been shown by scientists in several laboratories as well as those at MIT, including James Wang at Harvard, Robert Wells at the University of Alabama, David Pulleyblank at the University of Toronto, and Fritz Pohl at Konstanz University in Germany.)

"It is easy to show that Z-DNA is stabilized by negative supercoiling," says Rich. "What we wanted to do is to find which sequences of DNA are stabilized in the Z form—to select the little piece that's forming Z." Working with David Stollar and Eileen Lafer at Tufts University, Rich and his associates first obtained antibodies that bind specifically to Z-DNA. Then they developed a method of cross-linking these antibodies on Z-DNA and using restriction enzymes and nitrocellulose filters to separate the Z-DNA sequences from the others. With this technique, they could answer the question of which sequences form Z-DNA. "The main thing is that what we have is a methodology that is completely general. It allows us to take any piece of DNA and look for Z forms."

Together with Nordheim, Rich decided to look at SV40, a monkey tumor virus whose DNA is circular, underwound, and completely sequenced—all of its genes and control regions are known. After introducing negative supercoils into the DNA, they found that



Z-DNA in Protozoa

Fluorescent antibodies to Z-DNA (shown in b) bind to the macronucleus (large arrows in a) of this Stylonychia mytilus *protozoan but not the micronucleus (small arrows). Genes are transcriptionally active only in the macronucleus. [Source: H. L. Lipps, A. Nordheim, D. Ammermann, B. D. Stollar, A. Rich,* Cell **32**, *435 (1983).]*

Z-DNA is formed only in the control region of the viral DNA and specifically in a transcriptional enhancer segment, a sequence of particular interest to molecular biologists. Rich notes that alternating purine-pyrimidine sequences are underrepresented in the SV40 genome. "Thus it is remarkable that they are clustered at the control regions."

Transcriptional enhancers are sequences that turn on gene transcription. If an enhancer sequence is put near a gene, up to 100 times more copies of the gene can be made. But enhancer sequences are peculiar. They act independently of orientation, meaning that it does not matter if you flip them over, and independently of position, meaning that it doesn't matter if they are in front of a gene or in back of it.

"What many people think is going on is that transcriptional enhancers are entrance sites for RNA polymerases—they are a way for the polymerases to get on DNA. Once the polymerase gets on, it carries out a one-dimensional diffusion. It's as though the DNA is a railroad track and the polymerase zips along it until it finds a promotor [a place where the polymerase starts to copy the DNA]. It's a way of increasing access to that promotor," Rich says.

Interested in the possible correlation of Z-DNA sites and gene control regions, Rich and Nordheim made a survey. They looked at other viral sequences that have alternating stretches of purines and pyrimidines, making them candidates to form Z-DNA. They found pairs of such sequences, all in control or regulatory regions of the viruses. For example, the "long terminal repeat" regions of retroviruses are known to have enhancers in them and they also have pairs of potential Z-DNA forming regions. "We find that Z-DNA forming sequences tend to be in enhancer regions. This suggests to us that the formation of Z-DNA may have a significant role in activating genes."

Rich and his colleagues are now isolating and purifying the proteins that bind to Z-DNA to see if they can use them to recreate transcriptional activation in vitro. They are also collaborating with Winship Herr and Yakov Gluzman of Cold Spring Harbor Laboratories in selectively mutating Z-DNA sites in SV40 to see what would happen if a virus could not form Z-DNA. Their first results in this line of research are intriguing.

In one case, they made transversion mutants, changing the thymines to cytosines in the Z-DNA regions. These mutations do not change the regions' potential to form Z-DNA because they maintain alternations of purines and pyrimidines. In a second kind of mutation, called transitions, they changed thymines to adenines, which means that the regions are much less likely to form Z-DNA. The viruses with the transversions grow well. The viruses with transitions grow very very slowly, if at all.

Rich believes that the research on the possible biological function of Z-DNA has at last convinced most molecular biologist that this strange DNA structure is not an artifact. "People used to be skeptical, but they are less and less so now," he says. "It's become very exciting as the field is moving toward what you could call 'real biology.' If I had to guess what the watershed was, I would say it was the isolation of the proteins that bind Z-DNA. Up to then you could always say, 'Well, it could be an artifact.' But the proteins are there and they are found in fairly reasonable amounts."

Rich, of course, believed from the first that Z-DNA had biological meaning. "It was quite clear to me it was real. When you find a molecular conformation in a crystal lattice, it represents a low energy form of the molecule—that's why it's there. The challenge to me was to find out what it does. It seemed highly unlikely to me that it would not be used."

—GINA KOLATA