Form IV – the "Final Puzzle Piece"

Transcript of narration

By Ken Biegeleisen, M.D., Ph.D.

SLIDE 1
Opening Title

SLIDE 2

Hello, I'm Dr. Ken Biegeleisen, and I'm here today to discuss an extremely important DNA structure called "Form IV"; a structure which was discovered in the 1960s, then promptly forgotten. "Form IV" is the form assumed by native Form I circular duplex DNA, upon being subjected to alkali denaturation. It was discussed at considerable length in "The Double Non-Helix, Part I", on this web site, but we need to look more closely now.

SLIDE 3

The two most distinguishing characteristics of Form IV are, (1) that it is remarkably dense, sedimenting in velocity ultracentrifugation experiments at nearly *three times* the rate of the native chromosome, and (2) although this mysterious structure, depicted here by the question marks, is easy to create by alkali *de*naturation, it is notoriously difficult to *re-nature*. This behavior is markedly different from that of linear DNA, which can be denatured by simple boiling, and readily *re*natured by simply allowing it to cool slowly.

SLIDE 4

Before we proceed, a word about the screen. If you haven't downloaded this PowerPoint presentation to your computer, then you're online, and you're looking at a Flash or HTML5 export of the PowerPoint, accomplished by means of a converter called iSpring.

This is the 2016 version of iSpring, which was current at the time this presentation was created. For reasons which are not at all clear to me, they have removed some of the old user controls over the screen, but they've also added some nifty new ones.

The first thing to note is this slanted double-arrow icon, which converts from this default view, to a full-screen view. Note that the slide strip is still present in full-screen

view. There used to be a button to remove the slide strip, but that button is no longer present in the current version.

Next, please note the volume control down here at the bottom.

If we hit the double arrow again, and return to the default view, you'll see a new set of iSpring controls at the top called "Marker Tools". These tools include a virtual red pen and a yellow highlighter, with which you can actually write directly onto these slides, even when they're playing. I don't know how many people will actually have a use for these tools, but they're very cool, and if you have the time, or the need, check them out.

SLIDE 5
Table of Contents

SLIDE 6
"Introduction"

SLIDE 7

I don't know of a single scientist in the world today, who has the slightest interest in Form IV, or who even knows that it exists. My only interest was in completing the task to which this NotAHelix web site is dedicated, that is, to apply non-helical DNA structural theory to all important forms of DNA.

SLIDE 8

In previous slide presentations, we have done that for DNA when complexed with protamine, and when associated with the histone octamer in nucleosomes. Now we do it for Form IV.

SLIDE 9

I'd had simplistic ideas about Form IV for many decades, and at the outset of this endeavor, I thought the model would be completed in a matter of weeks at the most. Nothing could have been farther from the truth. The Form IV model wound up taking nearly a year of hard labor to solve, and even now the solution is not a single, clear-cut molecular model, but rather a *series* of models for several imprecisely-defined molecular species, none of which are carved in stone.

Although this turned out to be an exhausting and time-consuming task, in the end it yielded a precious and unexpected reward: When the modeling was complete, there unexpectedly appeared a solution to the oldest problem in non-helical DNA science...

SLIDE 10

...to find a simple, inexpensive and readily-repeatable method of non-destructively separating the strands of a plasmid or viral chromosome.

SLIDE 11

This is a feat that was accomplished only once previously in DNA history. That accomplishment took place in 1996, through the work of my colleague, Tai Te Wu, but the experiment was only accepted for publication in the Bulletin of Mathematical Biology, an obscure journal which few scientists read. Moreover, the Wu separation was extremely difficult, time-consuming and expensive, wherefore no one will ever attempt to repeat it.

The finding of a simple and inexpensive way to achieve this non-destructive strand separation is therefore of the utmost importance to DNA science, and well worth the one-year modeling effort.

SLIDE 12
"Historical Background"

SLIDE 13

In 1963, Renato Dulbecco and Marguerite Vogt discovered that the chromosome of the oncogenic virus polyoma was circular. This extended the parallel discovery made in E. coli bacteria earlier that same year, by John Cairns.

SLIDE 14

Here's a drawing showing these two chromosomes together, to give us some idea of the size difference. The E. coli chromosome, having been borrowed from the famous Cairns autoradiograph, is caught here in the act of replicating. The polyoma is not drawn precisely to scale, because if we did that, it would be nearly invisible, being only 1/1000th the length of E. coli. Even at the size shown here it's hard to see, so let's blow it up a bit.

OK, now we can see it a little better, including the supertwisted conformation generally found in native viral chromosomes.

The point of discussing this size discrepancy is to introduce the subject of the severe size limitations which always pertain to the study of DNA structure in intact chromosomes. The fact of the matter is that DNA structural studies *cannot* be done on large chromosomes; certainly not human chromosomes, and not even on the smaller bacterial chromosome, because the purification procedures for DNA will cause even a bacterial chromosome the size of E. coli, to shatter into a thousand pieces:

SLIDE 15

The polyoma chromosome, however, having a length of a mere 5 kb, is small enough to survive, fully-intact, during routine handling in the laboratory. Therefore, most of what we know about circular DNA comes from the study of the chromosomes of viruses, and of small cellular plasmids.

SLIDE 16

Using the small polyoma chromosome, Dulbecco and Vogt made the following discoveries, which have proven applicable to a host of small circular DNA species:

1. The native DNA has two components, which sediment separately in velocity gradient centrifugation.
2. The faster component consists of intact chromosomal DNA. This is invariably supertwisted, which accounts for its higher sedimentation rate.
3. The slower component consists of chromosomes which are nicked in one or both strands. This removes the supertwists, which accounts for its lower sedimentation rate.

SLIDE 17

Further studies with the intact, superhelical chromosome disclosed that, unlike linear or nicked circular DNA, it could not be heat-denatured. But it *could* be denatured by alkali. At a pH around 13, it converted to a novel form; one of astonishingly high density; sedimenting 2-3x *faster* than the native chromosome in velocity gradient centrifugation titrations, such as this one.

We looked at this data extensively in the PowerPoint presentation entitled "The Double *Non*-Helix, Part I", on this web site. In a little while, we're going to have to look at it again, even more closely.

What is the structure of this rapidly-sedimenting denatured form? It was clear, in the 1960s and 70s, when these studies were done, that alkali-denatured DNA remained duplex; that is, the strands did *not* physically separate upon denaturation.

Other than that, however, the molecular biological establishment was, and remains to the present day, 100% *clueless* as to what the structure of this astonishingly high-density form of DNA is. Concerning that structure, there have been no theories to explain it, no molecular models, and, in fact, there has never been any discussion about it at all.

SLIDE 18

This rapidly-sedimenting form is the subject matter of the current PowerPoint presentation. I'm going to show you its probable detailed molecular structure, but first, we need some terminology. So, largely from the laboratory of Jerome Vinograd, who was the world's reigning monarch of circular DNA science in the mid-20th century, here are the 4 forms of DNA which were recognized at that time.

Form I is the form in which small circular DNA is most frequently encountered in nature. As we have seen, it is superhelical, and has many interesting topological properties.

Form II is the form resulting from the introduction of as little as a single nick, into a single strand of Form I. The superhelicity, and all the interesting topological properties, are instantly destroyed, leaving an open-circular form which behaves, in most respects, no differently from common linear duplex DNA.

SLIDE 19

Form III is a rarely-used terminology, appended to Form I which has been subjected to full-duplex cleavage, yielding a linear duplex molecule. Since this is how most people think of DNA in the first place, this form is therefore usually just called…DNA!

Form IV is the structure of alkali-denatured Form I, and is the subject of this presentation.

SLIDE 20

The terminology "Form IV" is the terminology I was taught as a graduate student, and it's the terminology I have employed for over 40 years. There is, however, no uniformity in terminology for this form of DNA. It's been referred to by at least 4 other names that I know of, and probably others besides these:

1. Some authors just call it "**denatured DNA**".
2. Some turn this phrase into a species name, such as "**denDNA**".
3. Some think of it as being a subsidiary of the native chromosome structure, referring to it by terminology such as ("Form I primed") **Form I'**.
4. Robert Warner, whose work we shall review in great detail, called it ("Form I sub-d") "**Form I_d**".

SLIDE 21

This bewildering array of names betrays the fact that scientists generally have regarded alkali-denatured circular DNA as being merely an unimportant variation of the native Form I structure; one which does not merit very much serious consideration, since serious consideration generally begins with a serious name. I would take issue with that view, opining instead that the structure of alkali-denatured circular DNA is of *tremendous* importance to the understanding of *all* DNA structure, and that it is therefore deserving of its own personal form number. Therefore, for the duration of this presentation, we're going to stick to the terminology "Form IV".

SLIDE 22

DNA structural science was 10 years old when circular DNA was discovered, back in 1963. From the outset, it was clear that circularization gives rise to a structure that's full of surprises. For example, when linear DNA is boiled, the two strands separate, a process called "denaturation", and when slowly cooled, the strands rejoin, which is called "renaturation" or "reannealing". When *circular* DNA is boiled, however, nothing happens; the strands do *not* separate. Nevertheless, circular DNA *can* be denatured, by alkali at pH 13, whereupon it turns into the mysterious Form IV. Studies of A_{260}, that is, light absorption at 260 nm, clearly show that at pH 13, Form IV is in fact denatured, meaning that the base pairs are disrupted, but the structure remains duplex. Whether

examined by ultracentrifugation, electron microscopy, gel electrophoresis, or any other method, one thing is certain: Those strands do *not* physically separate.

SLIDE 23
"Renaturation of Form IV"

SLIDE 24

Form IV, which, since the early days of DNA science, has been depicted by silly drawings like this…[PICTURE]…is a denatured structure. Like linear DNA, it can be *re*natured, but getting it to do so, that is, getting it to revert back to the native Form I structure, proved historically to be a major undertaking. The renaturation of Form IV turned out to be a complex affair, requiring precise control of:

1. pH,
2. temperature, and
3. ionic strength.

SLIDE 25

If, for example, the temperature and ionic strength were fixed, the pH at which Form IV would renature, *i.e.*, convert back to Form I, was limited to a surprisingly narrow range.

I'll illustrate with a single example, employing data adapted from Strider & Warner, which we covered in detail in "The Double *Non*-Helix, Part I. If we were to set the temperature to 60º and the ionic strength to 1M NaCl, alkali-denatured DNA from the virus ϕx174, that is, Form IV DNA from the virus ϕx174, would rapidly *re*nature at pH 11, with 100% of the chromosomes resuming the native Form I conformation in the space of 10 minutes or less. But if we were to alter the pH, either up or down, by so little as a few tenths of a pH unit, the renaturation rate would drop drastically, as shown, down to about 35%. If we were to further alter the pH, by about ±0.5 pH units, as indicated by the orange extrapolation arrows, the renaturation rate would drop to absolute zero — there would be *no* renaturation at all, at least not in the time frame of this experiment.

SLIDE 26

The primary researcher who systematically investigated the *re*naturation of Form IV was Robert Warner, who, for the last 30 years of his life, was the chairman of the biochemistry department at the University of California at Irvine. The lab work, however, was actually done at the *New York* University School of Medicine, where I was, at the time, a graduate student in the MD-PhD program. That work was *not* done by me, but by a grad student down the hall a few feet from me, named Bill Strider. Bill's results were first published in a brief 1971 note in Federation Proceedings, then in great detail in his PhD thesis, written the same year. I sincerely believe that I'm the only man on the entire planet earth who has ever read Bill's thesis from cover-to-cover, including perhaps even his own thesis advisors.

That work was not published in a major peer-reviewed scientific journal until ten years later, 1981, at which time Bill Strider, ironically, had quit science, entered and graduated medical school, and had gone into medical practice in the field of Obstetrics and Gynecology. But Bob Warner, his UC Irvine advisor, was still very much in the business of science, and his 1981 publication incorporated everything in Bill's 1971 PhD thesis, and added much more to it. In fact, it was so minutely detailed and complex as to be a major challenge just to read.

SLIDE 27

The most essential details of the Form IV renaturation process, were, however, the same as those written up by Strider ten years earlier. Here now is the heart of the matter; the **Strider-Warner renaturation curves**. Take a good look at them, because we're going to be referring to them repeatedly throughout the remainder of this presentation.

These data are discussed in great length in the first PowerPoint presentation on this web site, entitled **The Double Non-Helix, Part I**. By the way, in case it's not obvious, that presentation can readily be found by looking for the "see no evil, hear no evil, speak no evil" apes, which mark the portal to the presentation:

The complete discussion of the Strider-Warner renaturation curves, in **The Double Non-Helix, Part I**, extended from slides 76-256, 180 slides in all, and I cannot repeat it all here, since that would take as long as the entire current presentation, thus doubling its length. But we can, and shall, take a few moments to discuss their essential features.

SLIDE 28

Each of these curves show the % *re*naturation of a sample of ϕx174 DNA, previously converted to Form IV, that is, previously denatured by addition of NaOH to pH 13, then incubated at the temperatures indicated adjacent to each of the five curves. In general, the time necessary for 100% reannealing was in the range of 2-10 minutes. In each of these experiments, the salt concentration was fixed at 1 M NaCl.

We've already looked at one of these curves, the one for 60º. We saw that at that temperature, and at that salt concentration, reannealing only took place within a narrow range of pH, and dropped to zero if the pH was altered by even a half unit.

Now we can see that this peculiar behavior of Form IV is the same at all temperatures examined, except that at room temperature, 100% reannealing couldn't be attained at any pH.

SLIDE 29

Strider and Warner then repeated this exercise, only this time with the temperature fixed at 40º, but with a systematic alteration of the salt concentration:

For each of the salt concentrations shown, we see the same dependence on precise control of pH. At 1M NaCl, for example, renaturation was optimal at about pH 12, but

moving the pH up as slightly as two-to-three *tenths* of a pH unit totally abolished all renaturation.

SLIDE 30

To illustrate that all three parameters – *i.e.*, pH, ionic strength, *and* temperature – must be carefully controlled to get renaturation of Form IV, I have re-plotted some of the data to illustrate the effect of temperature.

Each of the curves now shows % renaturation in 1M NaCl at the indicated pH, as a function of temperature. Thus, for example, at pH 11.3, the green curve, renaturation is rapid and complete at about 50º, but drops almost to zero at either 40º or 60º.

How do these, or any of the other Form IV authors, explain any of this? They don't! No one has ever made the slightest effort to determine the structure of Form IV. The closest thing to a "molecular model" was a hand-drawn picture published by Vinograd in the 1960's. We looked at that picture a moment ago; here it is again: the "state-of-the-art" with respect to Form IV structure, as of 2015, the date of this presentation.

Maybe this was acceptable as a "molecular model" in its day, but that was 60 years ago, and we're going to have to do a lot better than that now. And there's no one to help us. We're going to have to figure it out for ourselves.

SLIDE 31
" Sedimentation coefficient vs pH titration"

SLIDE 32

In this endeavor, we are hardly clueless. There's actually a rather large body of Form IV data to guide us. We have just looked at one of the main data sources: these Strider-Warner renaturation curves. Whatever structure we propose for Form IV, that structure will have to explain its extreme fastidiousness for rigorous, pinpoint control of pH, temperature and ionic strength during renaturation.

The other major source of information about Form IV, which must be taken into consideration in any molecular model-building effort, emerges from the sedimentation coefficient vs. pH titrations that several groups undertook to do in the 1970s. The most important of these titrations was done by a man whose name we mentioned before, Prof. Robert Warner, who, for many years, served as chairman of the biochemistry department at the University of California, Irvine. The work was actually done by his postdoctoral student Mark Rush. Mark had previously earned his PhD in the lab of the also above-referenced Jerome Vinograd of CalTech, who was the inventor of CsCl density gradient centrifugation, and one of the most prestigious biochemists of the mid-20th century. Mark then himself became a professor of biochemistry, at the New York University School of Medicine, the institution at which I was accepted into the MD-PhD program. When I heard that Mark was working with small circular DNA, I made a beeline for his lab, and he became my research advisor.

SLIDE 33

While still a student himself, Mark Rush had coauthored a paper with Warner, on the sedimentation coefficient vs pH titration of DNA from the virus ϕx174. Here is an annotated adaptation of the result.

Like the Strider-Warner Form IV renaturation curves, this *sedimentation coefficient vs. pH* data was exhaustively covered in "The Double Non-Helix, Part I", on this web site. But we're going to have to talk about some of it again.

The data shown here are sedimentation coefficient values for the ϕx174 chromosome at various pH's. If, perchance, you have no background in velocity gradient sedimentation, then I'll tell you that it involves placing samples into an extremely high-speed ultracentrifuge, spinning at rates up to 100,000 rpm, and measuring the speed with which the samples drop to the bottom of the tube. This speed can then be represented by a parameter, called the "sedimentation coefficient", usually abbreviated ***s***. Comparison of the ***s*** values of different molecules can provide important information about size and shape.

Here, the sedimentation coefficient is shown on the ordinate, labeled "S_{w20}". For brevity, we shall hereafter dispense with the subscript "$_{w20}$" in "S_{w20}", which merely indicates that the solvent was "w" for "water", and that the temperature was 20º Celsius, and henceforth we shall refer to the sedimentation coefficient simply by the single letter "***s***".

In experiments in which polymers or aggregates of the same substance differ in size, the different ***s*** values may reflect mainly those size differences. In this experiment, however, where all the molecules are the same size, the differing ***s*** values reflect mainly differences in *density*, or *compactness* of the molecules, which differences come about because of conformational changes, especially changes in the amounts of *twisting*.

The most prominent features of this curve are the large *dip* in ***s*** at a pH just below 12, and a shoulder at pH 12.3. The shoulder is not very prominent here, but there are a number of these denaturation studies in the literature, and each one of them shows this shoulder, which is therefore established as a true and reproducible attribute of this sort of pH vs. ***s*** titration curve.

The significance of the pH 12 dip and the pH 12.3 shoulder were also discussed at great length in "The Double *NON*-Helix Part I", but we must now repeat the part that pertains to the structure of Form IV.

SLIDE 34

Before we proceed, it is essential that we keep firmly in mind the conditions which prevailed in the centrifuge tubes during the centrifugations whose results are shown in this figure, because, as we now know, any statements made about circular DNA conformation are meaningless unless we know the temperature, ionic strength and pH at which the conformation was observed. Here, the temperature, as the ordinate axis label indicates, was indeed 20º. The ionic strength was variable, depending upon the pH, but was, in all cases, between 0.3 and 0.4 M sodium.

Now let's look at the data. The lower curve, composed of white circles, is a *de*-naturation curve, which, through changes in the sedimentation coefficient ***s***, gives us

clues about conformational changes which the chromosome undergoes, as the pH is increased from neutrality...up to 13...the point of permanent alkali denaturation.

SLIDE 35

The little pictures above the denaturation curve illustrate the changes of state presumed to be taking place in those regions of the curve. We'll discuss some of these shortly.

The upper curve, composed of black circles, is, as we shall see, *not* a *re*-naturation curve, but merely a *neutralization* curve, showing what happens to the denatured chromosome when the pH is lowered back to 7. Clearly, in this experiment, the material does *not* ***re***nature, because, as the top curve indicates, the ***s*** value of denatured ϕx174 DNA, after neutralization, does *not* return to normal, even after the pH is lowered all the way back down to 7.

SLIDE 36

It is very important that we are clear about the difference between neutralization and renaturation. If we really wanted our denatured ϕx174 DNA to literally *re*nature, that is, return to the native conformation, with the native ***s*** value restored, the only way to have done that would have been to apply one of the sets of conditions defined by the Strider-Warner curves we looked at a moment ago.

But these conditions, which are absolutely essential for renaturation, are totally lacking in the Rush-Warner ***s*** vs pH titration experiment. As we have seen, the temperature, in accordance with the ordinate label, was 20º, and, as the figure caption explains, the ionic strength, including the contribution of NaOH for pH adjustment, ranged from 0.3-0.4 M. Are these conditions under which *any* true renaturation would be possible? Absolutely not. Let's see why.

SLIDE 37

Returning to the Strider-Warner curves, we notice that as we look from left-to-right, that is, *up* the pH scale, the temperatures of the reannealing optima *decrease*. There is therefore a reciprocal relationship between temperature and pH in the reannealing process. Clearly, it appears that reannealing requires conditions that are of a ***disruptive*** nature, *i.e.*, *high* temperature, *high* pH, or *high* ionic strength. But if the temperature is lowered, as when we go here from 70º to 60º to 50º to 40º, then the pH required for optimal renaturation must be commensurately *increased*, from about 10.7 to about 11.7. In other words, in the face of decreasing temperature, an increase in the pH is evidently necessary to maintain the required *disruptive force* at a level adequate to maintain the goal of 100% reannealing.

At the *lowest* temperature tested, however, it is important to note that this rule can be seen to be breaking down. Thus, at 25º, *i.e.*, "room temperature", we find that reannealing has become sluggish; 100% reannealing is *not attainable*; not even at pH's of 12 and above. It would thus appear that at room temperature, there is simply not enough *disruptive* energy to promote renaturation, no matter how high we make the pH.

SLIDE 38

Now, trying hard to keep all this in mind, let us reconsider the Rush-Warner **s** vs pH titration study; particularly the upper curve marked by the black circles, which depict the changes in ***s*** associated with the lowering of the pH of Form IV back to neutrality. The question is, *what are the conformational changes represented by those black circles?*

In particular, before we start to propose novel conformations for that DNA, which we *shall* shortly do, I feel strongly that it's well-worth taking a few minutes to rule out the confounding possibility that any significant part of the DNA, represented by the upper curve, *reanneals,* and returns to the native Form I structure.

Let's start by considering the temperature, which was 20º. Is it even possible that some of the Form IV could renature by merely neutralizing it at 20º? I doubt it. Let's bring back the Strider-Warner renaturation optima. We have already seen that at 25º, it's impossible to get 100% reannealing, even at very elevated pHs. So how much reannealing could we expect at 20º, which is 5º *lower* than that? Not much! If Strider had attempted to find a pH of optimal reannealing at 20º, it would have been up here someplace. At such a low temperature, the hypothetical reannealing optimum pH would have to be extremely high, perhaps as high as pH 13 or more, but that might be physically impossible, because at that pH, all the bases would be *deprotonated,* and DNA would only be able to *de*nature, not *re*nature.

It is therefore very unlikely that there could have been any significant reannealing in the neutralization curve of the Rush-Warner ***s*** vs. pH titration experiment.

SLIDE 39

This important conclusion is even *further* established by a re-consideration of the effect of ionic strength. Let's look again at the Strider-Warner figure showing the dependence of Form IV reannealing optima on salt concentration:

Lowering the salt concentration, from 2M to 1M to 0.1 M, *increases* the pH required for optimum reannealing, moving it to the right on the pH scale. It can be estimated, from the available data, that at the ionic strength of the ***s*** vs. pH titration at the top of the slide, namely 0.3→0.4 M salt, Bill Strider's pH optima for renaturation, shown at the bottom of the slide, most of which were determined at 1M salt, would be pushed somewhere between 0.2 and 0.3 pH units *higher*.

Now let's take a last look at the Strider-Warner renaturation optima for 1M salt. They already show poor reannealing at 25º. Now we know that lowering the temperature to 20º is going to move that curve to the right, and lowering the salt concentration to 0.3-0.4 M is going to move it even *further* to the right, into a realm of pH where no reannealing is physically possible, because all the bases would be deprotonated.

SLIDE 40

And yet we see that the upper neutralization curve, formed from the black circles, clearly depicts a change-of-state of the DNA, culminating, at pH 7, in a form that is significantly less compact than the Form IV seen at pH 13, but still *twice* as compact as the native chromosome. We have, I believe, now firmly excluded the possibility that any

part of this DNA is truly reannealed. If that be the case, then *what is the structure of the neutralized DNA at pH 7?* And if we call the product observed at pH 13 "Form IV", what shall we call the form at pH 7, which is only half the compactness? Is it still Form IV, somehow made less compact? Or is it something else?

SLIDE 41
(Transitional slide, no audio or visual content)

SLIDE 42

There's a further detail we must pay attention to in the upper neutralization curve. Please note that at a pH of about 12, there appears to be a *shoulder* in the curve. Note that I have inserted a vertical gray shading bar into the figure, to emphasize the fact that the pH at which this apparent shoulder is seen, coincides with the pH range at which many dramatic events are taking place in the lower *de*-naturation curve. But there's a problem: Compared to the dramatic events in the lower *denaturation* curve, this upper curve shoulder is rather subtle; if we just nudge a few data points, by a few tenths of a Svedberg unit, the shoulder disappears...

OK, let's put the shoulder back. If this shoulder is real, then, as we shall see, it is of considerable importance in the interpretation of this data. But *is* it real? In order to answer that question, we would have to look at other similar experiments.

And the bad news is this: the data you're looking at right now is the only *neutralization* curve that's ever been published! In a moment, I'm going to show you several more *de*-naturation curves which have been published, employing different species of DNA, and they all have the same salient characteristics as this one. But none of those authors ever bothered to do an ***s*** vs. pH titration for the *neutralization portion of the data*, that is, a titration comparable to the upper curve, composed of the black circles. The one you're looking at now is the only one in existence!

So should we take the apparent shoulder in the upper curve seriously? Is it real, or just a random fluctuation of data points? I have concluded that it *is* real, not just for one, but for two good reasons which I'll present shortly.

SLIDE 43

Right now, let's look at some other ***s*** vs. pH titrations, done with the DNA of other organisms. As I just said a moment ago, each of these will consist of a *de*naturation curve only; there will be no neutralization data. Here, for example, is the denaturation curve for the penicillinase plasmid of Staphylococcus aureus, also from the laboratory of Rush & Warner.

Note that the general shape of the denaturation curve is similar to that of the comparable curve for ϕx174 which we have been looking at, except for one thing, which is that the entire curve seems to be shifted about 0.3 pH units to the left. I'll briefly fade-in the ϕx174 curve, so we can see the pH shift to the left when we go back to the penicillinase plasmid.

Whether this is a species difference, or a difference in the experimental protocol, I cannot say. Nevertheless, it is evident that the penicillinase plasmid curve possesses all

the main features we saw for ϕx174, in particular the very prominent dip in ***s*** at about pH 11.7, and an unmistakable *shoulder* in the curve at about pH 12.

SLIDE 44

Here is the comparable curve for the oncogenic virus polyoma, from Vinograd's lab, whose pH parameters are the same as those of the penicillinase plasmid.

Once again, we see the dip and the shoulder.

We thus see that any Form IV model, in addition to having to explain the extreme fastidiousness of the renaturation process for pH, temperature and ionic strength, will also have to explain every nuance of these ***s*** vs. pH titration curves, including the neutralization curve for ϕx174.

SLIDE 45
"Final requirement: EM"

Title 46

As if the situation was not complex enough, we have one more major species of experimental data to explain, namely that arising from electron microscopic studies of Form IV. The pictures shown here were compiled from the 1974 study of Grossman, Watson and Vinograd. The top row shows pictures of Form IV spread using a formamide-based hypophase; the bottom row shows Form IV spread using an aqueous hypophase. In each row, the panels on the right show the effect of adding ethidium bromide to the denatured DNA. That effect is to induce a high degree of superhelicity, a phenomenon we are not, at the moment, interested in. We shall therefore be focusing our attention on the left-hand column only.

And, incredibly, what we see is that in the absence of EtBr, the Form IV is double-stranded, looking indistinguishable from Form II DNA, *i.e.*, native DNA which has been relaxed by introduction of one or more single-stranded nicks! *What the heck is going on here?!* Just a moment ago, we were looking at graphs showing that Form IV was incredibly dense, sedimenting at *three times* the rate of native DNA, prompting artists to render it in accordance with the belief that it's some sort of hopelessly twisted Gordian Knot, such as is portrayed by this drawing we have looked at several times before. Then we put this stuff under the electron microscope, to visualize that Gordian Knot, and what do we see? *Nothing!* Look at these molecules! Look at this one down here – it looks like ordinary, relaxed circular DNA! Where's the Gordian Knot? Where's the collapsed, compacted and twisted structure? There is none! *What the heck is going on here??*

And don't think that there's some sort of error in these electron micrographs, that is, that maybe if this *looks* like Form II, it *is* Form II. It is *not* Form II, it's Form IV, and there's no error about it. In fact, as we shall see shortly, this is undoubtedly the form which was erroneously declared by Stettler *et al*, that is, by Charles Weissmann and his associates, to be a new DNA form which he dubbed "Form V". What we shall see is that the so-called "Form V" is merely Form IV, but in a conformation identical or similar to

that shown here; a conformation which was somehow brought about by the preparatory steps for electron microscopy.

It is virtually inconceivable that this denatured DNA is base-paired, because it arose at the base-pair-destroying pH of 13, and was not subsequently subjected to any of the conditions demonstrated by Strider and Warner to be essential for reannealing. And yet it looks just like normal duplex DNA. How can we reconcile the repeatedly-demonstrated compactness of Form IV, in the face of these extended, open-circular-looking structures?

SLIDE 47

The beginning of understanding this apparent dilemma is to understand that what we're seeing here is the result of not only lowering the pH back to physiological realms, but also lowering the salt concentration. In every curve we've looked at, up to now, the DNA has been in a salt environment in the range of 0.3-1.0 M sodium. To put this in perspective, the salt concentration of the human body, that is, the concentration of NaCl in human body fluids, is generally given to be 0.9% W/V, or 9g/L = 9/58.44 = 0.15 M. So our data, up to this point, has been data generated under conditions which, relative to the human body, would be well-described as "high salt" conditions.

Here, however, we see the results of spreading DNA, in preparation for electron microscopy, in a solution whose salt concentration is not likely to be any higher than 0.1 M, and usually 5-10x *lower* than that. Therefore, in electron microscopy, we are dealing with DNA in settings considerably *more* dilute than those found in living tissues. These may, relatively speaking, be therefore well-described as "low salt" conditions. And the result of removing the salt appears to be that the denatured DNA reverts to a duplex form closely resembling native DNA which has been relaxed, and which cannot likely have the compact structure of the experimental material at pH 13. But what *is* its structure?

SLIDE 48

We are almost at a point where we can provide detailed molecular models for this, and for the other stages of Form IV denaturation, but first, there are 5 more phenomena of DNA structure and/or science which must be thoroughly understood before we can proceed:

Phenomenon #1, which we absolutely must understand, is Z-, or left-handed DNA.

Phenomenon #2 is tetraplex DNA, consisting of a pair of duplexes with mutually-intercalated base pairs.

Phenomenon #3 is alternative-tautomer base-pairing in DNA.

Phenomenon #4 is that the maximum number of supertwists possible in covalently-closed-circular DNA appears to be a number corresponding to approximately 0.5% of the number of base pairs in the chromosome. Thus, in the typical plasmid and viral chromosomes we shall be discussing, most of which have approximately 5000 bp's, the maximum # of supertwists possible is about

25. If you didn't know that, or if you cannot believe it, hold off on your skepticism – we'll be explaining it shortly.

Phenomenon #5 is the supposedly-novel form of DNA mentioned above, known as "Form V", which is almost surely a misnomer, since — as we shall see — it is probably nothing more than Form IV at low salt concentration. But we need to consider Form V, described by Charles Weissmann and his associates in 1979, because their paper on the subject single-handedly poisoned the minds of an entire generation of molecular biologists *against* non-helical DNA structures, and is still doing so after nearly 40 years.

Unless these 5 phenomena are understood, there is no possibility of explaining the structure of Form IV alkali-denatured DNA.

SLIDE 49

"Phenomenon #1: Z-DNA"

SLIDE 50

Left-handed DNA was first proposed by Mitsui *et al* in 1970, based upon x-ray crystallography of the self-complementary copolymer poly d(I-C)·poly d(I-C). These results were extended by Pohl and Jovin (1972), who showed that a similar copolymer, poly(dG-dC)·poly(dG-dC), underwent a cooperative R→L, that is, right-handed to left-handed helical transition at high salt concentrations. Please keep this latter phenomenon in mind, because we are shortly going to be discussing conformational changes that depend heavily upon salt concentration.

SLIDE 51

In 1979, Alexander Rich and his co-workers published a detailed model of left-handed DNA, shown here, which they dubbed "Z" DNA, choosing the letter "Z" because they thought that that letter well-represented what they perceived to be a zig-zag quality in the sugar-phosphate backbone, wherein the backbone left-handed twist alternated regularly between two distinctly different helical conformations. This is quite unlike right-handed W-C DNA, whose backbone twist is largely homogeneous from base-pair to base-pair.

In case the zig-zag nature of the sugar-phosphate backbone is difficult to see in the complete structure, we can isolate one of the strands, enabling us to see it better. This video shows a single strand from a "standard" B-DNA duplex, juxtaposed with a single strand from a Z-DNA duplex. If you watch closely, you should be able to discern that one of the strands has an irregular, non-homogeneous character, whereas the other is a smooth, continuous helix with no irregularities. The irregular one, you will also note, differs from the other in having a left-handed helical twist.

All this is much easier to see if we use cartoons to depict the single strands. Now you can clearly see that one of the two strands has a zig-zag appearance to its sugar-

phosphate backbone; this is the Z-DNA. The other is standard Watson-Crick B-DNA, and has a smooth helical backbone with no irregularities.

SLIDE 52

Here is a comparison between the basic dimensions of Z-DNA and B-DNA, that is, between the dimensions of left-handed Rich DNA and right-handed W-C DNA. Note the large rise-per-residue in Z-DNA, 3.7 Å, compared with the more familiar 3.4 Å spacing in B-DNA. Doesn't this mean that the bases in the left-handed Z form are farther apart? Well, "yes and no". Rich, in his writings, made the point that, whereas Watson-Crick B-DNA was best thought of as being a continuous ladder of base pairs, with uniform spacing between the steps of the ladder, it was not so in Z-DNA. Z-DNA, although also a ladder, was best thought of *not* as a ladder of base-pairs, but of *pairs* of base-pairs.

Note the large rise-per-residue in Z-DNA, 3.7 Å, compared with the more familiar 3.4 Å spacing in B-DNA. Doesn't this mean that the bases in the left-handed Z form are farther apart? Well, "yes and no". Rich, in his writings, made the point that, whereas Watson-Crick B-DNA was best thought of as being a continuous ladder of base pairs, with uniform spacing between the steps of the ladder, it was not so in Z-DNA. Z-DNA, although also a ladder, was best thought of *not* as a ladder of base-pairs, but of *pairs* of base-pairs.

SLIDE 53

That is, in Z-DNA, the fundamental *structural unit* was not a base-pair, but a *pair* of adjacent base-pairs, *i.e.*, a *dinucleotide*. Within that structural unit, the two base-pairs were separated by the usual DNA spacing of 3.4 Å. The dinucleotide structural units themselves, however, were *not* separated by a spacing of 3.4 Å, but by the considerably larger spacing of nearly 4 Å. Here's a closer look, showing only the bases in two adjacent structural units. The stacking of base pairs *within* the structural units are as expected for DNA; a stacking usually referred to as "parallel displaced". But successive structural units, when viewed along the sugar-phosphate axis, had little or no overlap at all, as seen in this axial view of the same 4 bases.

Presumably because they are all synthetic co-polymers, the published Z-DNA structures I have looked at are very regular and repetitive, from dinucleotide-to-dinucleotide structural unit. We may therefore assign inter-base-pair distances to the base pairs within the Z-DNA structural units, and *between* the structural units. Moving from lower-to-higher residue numbers, the distances — representing measurements along a line parallel to the sugar-phosphate backbone axis — are:

G→C:	3.45 Å (*within* structural units)
C→G:	4.00 Å (*between* structural units)
Average:	3.72 Å

SLIDE 54

What is the reason for the 'Z' structure? That is, why the zig-zag? Why can't left-handed DNA have a smooth and continuous winding, such as that of W-C right-handed DNA?

The answer may lie in a statement made by Francis Crick in 1953, over a half-century ago. That statement was made in the context of a discussion about the serious topological problems inherent in unwinding, replicating, and then re-winding a duplex helical structure whose strands were plectonemically intertwined, that is, a duplex having the Watson-Crick twisted structure.

It's really quite remarkable that whereas today, most molecular biologists work themselves into a veritable snit if you make even the slightest suggestion that the Watson-Crick structure is in any way problematical, Crick, in sharp contrast, way back in 1953, spoke with complete impunity about the difficulty of envisioning a way for his two strands to separate for the purpose of replication. He stated:

> **"The difficulty might be more simple to resolve if successive parts of a chromosome coiled in opposite directions. The most obvious way would be to have both right and left handed DNA helices in sequence but this seems unlikely as we have only been able to build our model in the right handed sense..."** (Watson JD & Crick FHC. The structure of DNA. CSHSQB 18:123-132, 1953, p. 129)

When I first read this, which, for me, was the year 1972, my senior year at Cornell University, I couldn't believe it. I had already become convinced, at that early time in my life, that circular chromosomes most likely had no net helical twists, and that therefore DNA from circular chromosomes, at least in purified, protein-free form, absolutely *had* to have the ambidextrous winding that Crick had rejected as "unlikely". That winding is represented here in rather ridiculously-oversimplified fashion. I should emphasize, in telling this story, that at that time, it would be another 7 years before Alexander Rich would publish the first crystal structure of left-handed Z-DNA. In the year 1972, few scientists had any inkling that DNA might possibly have a left-handed form.

Crick's comment suggested that left-handed DNA was impossible. But was that really true? I had to see for myself.

SLIDE 55

Therefore, while still a pre-med college undergraduate, I spent a month making a very precise model of left-hand DNA, from a large number of Framework Molecular Model kits. The model began, not as left-handed DNA, but as an exact W-C *right*-handed helix.

Every bond length was correct to the nearest mm, and every bond and dihedral angle to the nearest degree. By the time the right-handed model was completed, I had become rather emotionally attached to it, sort of like David Niven in the movie "Bridge Over the River Kwai" [ERROR: THE ACTOR WAS ALEC GUINNESS]. So, when the time came, I had to ruthlessly suppress my attachment. I grabbed the model by its extremities, and heartlessly twisted it to the left. After a few minor adjustments, there

stood before me what appeared to be a perfectly acceptable model of *left-handed* DNA. I *very* carefully checked for VDW violations, and there were none. I concluded that left-handed DNA was entirely possible, and when it was actually proven to exist 7 years later, I, for one, was not the least bit surprised.

SLIDE 56

In successfully creating a left-handed model with no steric hindrances, I thought I had disproved Crick's statement about the unlikelihood of left-handed DNA, but that's because I really hadn't understood his statement. It's only recently that I came to understand what Crick was *really* saying. He was *not* saying that left-handed DNA was impossible, he was merely saying that it wouldn't produce the x-ray patterns which had emerged from Wilkins' crystallography lab. That is, he was saying that if one forced left-handed DNA to occupy the same volume, and to have the same pitch as right-handed DNA, *then,* and perhaps *only* then, there would be Van der Waals violations; *i.e.*, that the left-handed structure was only impossible if it was created with those constraints.

SLIDE 57

I think, therefore, that the reason that Alexander Rich's left-handed DNA structure has the peculiar zig-zag sugar-phosphate backbone that it has, is because the peculiar geometry of the left-handed sugar-phosphate backbone requires it, *i.e.*, that it cannot be any other way.

The key to further understanding of this matter undoubtedly lies in a straightforward study of the bond lengths, bond angles and dihedral angles of the left-handed sugar-phosphate backbone, but I am not enough of a chemist to undertake such a study, because it would take me a very long time to do, and there are better chemists with better software than I possess, who could do it in a fraction of the time.

SLIDE 58

There is another property of Z-DNA that must be understood, if the conformational peculiarities of Form IV are to be explained. Here we have measured the phosphate-to-phosphate distances in Z- and B-DNA, along the sugar-phosphate backbone. These are the direct phosphorus-to-phosphorus interatomic distances, not the distances between the rectilinear planes within which the phosphorus atoms lie. These distances are actually *shorter* in Z-DNA than in B-DNA. Considering that the average *base* spacings in Z-DNA are *longer* than those of B-DNA, this must indeed be accounted a very peculiar and paradoxical feature of the Z structure, but it is a fact readily verified by direct measurements. Here are the P→P spacings for two decamers...

As I said a moment ago, the residue spacings in B-DNA are surprisingly variable, giving rise to a range of P→P distances, with an average of 6.72 Å. The P→P distances in this Z-DNA model, however, alternate rigidly between exactly 5.91 Å and 6.01 Å, with an average of 5.96 Å. Remarkably, even though the residue spacing in Z-DNA is 11% *more* than that of B-DNA, the P→P spacing shown here is 9% *less!*

SLIDE 59

These short P→P spacings are believed to account for an important and well-documented behavior of DNA under a variety of extreme conditions, namely a right-handed→left-handed helical transition. We made mention of this a few moments ago. The phenomenon was first described by Pohl and Jovin in 1972, who employed optical rotatory dispersion studies to demonstrate the change in helical sense. Those authors reported that the transition took place progressively, as the salt concentration increased from 2.5→4.4 M. But why? That is, *why* should an increase in salt concentration cause a reversal of the direction of DNA helical winding?

The current understanding is that high positive counterion concentrations neutralize the negative charge repulsions between adjacent phosphate groups, allowing the DNA backbone to assume the closer phosphate-phosphate spacings found in Z-DNA.

SLIDE 60

"Phenomenon #2: tetraplex DNA"

SLIDE 61

I cannot explain the ***s*** vs. pH titration curve for circular DNA without invoking a tetraplex DNA structure at certain points in the titration curve. Since most scientists are not familiar with tetraplex structures, it would be best if we review them up front, at least briefly.

I personally know of only two such structures. The most straightforward of them is the one proposed by my colleague Tai Te Wu, which he described as a "straight ladder" structure, distinguished by a fully-extended sugar-phosphate backbone, having approximately 6.8 Å residue spacing, *i.e.*, twice the normal DNA base-pair spacing, and no helical twist. A pair of such duplexes could then mutually intercalate their base pairs, giving rise to a tetraplex product with normal 3.4 Å base pair stacking.

SLIDE 62

I was able to use the Wu straight-ladder tetraplex to solve one of the longest-standing problems in molecular biology, namely the structure of the complex between protamine and DNA in sperm cells. This is a structure that is greatly ignored by molecular biologists, but only because it is perceived as having no commercial value at the moment. On the other hand, as a theoretical problem, its importance can hardly be exaggerated.

SLIDE 63

It has never been disputed, in any writing I've seen from any quarter, that the basis of the association of nucleoprotein with DNA is charge attraction. And it is beyond doubt that protamine is the very most prototypical of all nucleoproteins, being little more than a long string of positive charges from basic amino acid residues, mostly arginine.

Conversely, DNA is a long string of negative charges, from phosphate groups in the sugar-phosphate backbone. The process of aligning the long string of positive charges to the long string of negative charges, to create a plausible protamine-DNA model, is so conceptually simple, and so fundamentally prototypical, that one would really have expected it to have been completed in the year 1953, within a few months at the most after publication of the W-C model. But it wasn't completed in 1953.

In fact it was *never* completed, until I completed it, in 2006, over a half-century later. The reason the structure couldn't be solved sooner than that, was that everyone who attacked the problem did so with helically-twisted DNA, which simply doesn't work. The reason I succeeded where everyone else failed, is not that I'm "smarter" than anyone, but simply that I was willing to start with a *non-helical* DNA structure, after which the problem virtually solved itself.

SLIDE 64

The solution to the protamine-DNA mystery, which I published in the Journal of Theoretical Biology, is given in minute detail on this web site, in the PowerPoint presentation entitled "The Double Non-Helix, Part II: The Probable Structure of the Protamine-DNA Complex". It would be foolish to repeat that entire presentation here, wherefore I shall present only a view of the final structure, and refer interested readers to "The Double Non-Helix, Part II" for further details.

SLIDE 65

Let's start with a look at a segment of a Wu duplex, with 6.8 Å spacing.

SLIDE 66

Now let's spin it around, to get a 3D perspective on it. This is essentially no more than the Watson-Crick structure, only stretched to its full length, which removes the helical twist. There are no significant steric hindrances in the backbone, and the only thing which would prevent this structure from forming spontaneously is the double-length spacing between base pairs, which we'll solve momentarily.

SLIDE 67

DNA, in this conformation, can be perfectly aligned with protamine. Here's a view of the unit cell of that structure. The protamine P1-P2 dimer, which can be well-nigh-conclusively shown to be a β-sheet, is in the middle, with its many positively-charged arginine residues, which constitute about 50% of the amino acid sequence, regularly spaced at 6.8 Å. Since arginine has the longest R-group of any of the 20 amino acids, you'll have no difficulty discerning the arginine residues, even in this low magnification view.

The DNA is in the conformation we just examined in the previous SLIDE, although you won't necessarily recognize that at a glance, because it is rotated 90º, so that the base pairs, formerly pointing toward us, are now facing outward, that is, pointing to

the right and left edges of the screen. What you will recognize, however, is that the DNA phosphate groups, colored red, are aligned with the arginine side-chains on either side of the protamine dimer. The charge alignment is as perfect as perfect can be, but there's one thing wrong: DNA will not stack at 6.8 Å. To solve this problem, we bring in additional unit cells. Watch as the base pairs mutually intercalate:

SLIDE 68

That the DNA and protein align so perfectly is in itself remarkable, but the fortuitousness of this structure doesn't end there. We're only looking at the stacking of bases in the x-axis direction, from left-to-right on the screen. Let's now consider the relationships between adjacent unit cells in the z-direction, which is normal to the screen.

SLIDE 69

We start by removing the adjacent unit cells, so we can concentrate on one.

SLIDE 70
(No narration; video of protamine-DNA unit cell being rotated 90º)

SLIDE 71

Here's the axial view of the structure. There are too many atoms to comprehend, so let's look at a cross section.

SLIDE 72

This is a magnified view, showing 3 bp's on either side, plus 4 arginine residues in the middle. Let's add some labels.

SLIDE 73

In this view, we can see the excellent alignment of the charged groups, giving rise to perfect 3Å salt bridges, wherever arginine residues are present. Let's shrink this down and add additional unit cells, which will interact with one another by mutual intercalation of base pairs:

SLIDES 74-76
(No audio)

SLIDE 77

Note the alternating columns of DNA and protein, which will continue from one side of the sperm nucleus to the other.

SLIDES 78-81
(No audio)

SLIDE 82

What we're looking at here is merely an axial view of the mutual intercalation of base pairs between adjacent DNA duplexes, which we're already seen in longitudinal view. Now we need to consider what's going on to the right and left of the unit cells shown.

SLIDE 83

To do that, we'll bring in an additional group of unit cells, starting with them aligned to the first group, protein-to-protein and DNA-to-DNA. Now, let's displace the second group by a distance of ½ of the unit cell.

SLIDE 84

When we position the unit cells thusly, we find that they now interact by means of an extraordinarily fortuitous array of salt bridges. Let's take a closer look.

SLIDE 85

The three gray boxes, now highlighted, show the locations of the salt bridges, with positive charges contributed by arginine, and negative charges by DNA phosphate.

SLIDE 86

Here's a close up, with the charges marked.

SLIDE 87

Each side of the box consists of a perfect 3Å salt bridge. And, in case you haven't had to deal with basic trigonometry in the recent past, and you're wondering whether the structure might be weakened by the charge repulsions at the diagonals, ...

SLIDE 88

...please keep in mind that charge attractions and repulsions change with the square of the distance, wherefore the charge attractions at the sides of the square are twice as strong as the charge repulsions at the diagonals.

The entirety of the protamine-DNA complex is therefore stabilized, from one end of the sperm cell to the other, by a massive array of salt bridges which is so unexpectedly perfect, that, to quote my cover letter to the Journal of Theoretical Biology:

> **"In fact, if this structure is not correct, at least in its essential details, then an evil demon in a parallel world must be playing a trick on**

humankind. Nothing that works this well can possibly be just 'coincidence'. "

The editor, in his acceptance letter, referenced this statement, so I suspect he agreed with me.

SLIDE 89

Here's a dodecanucleotide of the final intercalated structure. It actually strongly resembles the W-C structure, except that it looks like there are too many sugar-phosphate backbones. The stoichiometry, however, *is* correct, because even though there are 4 sugar-phosphate backbones, each one is stretched to twice the "traditional" length, so that the combined mass per unit length, of all 4 together, is exactly the same as that of the 2 sugar-phosphate backbones of W-C DNA.

SLIDE 90

Here's an axial view of the Wu straight-ladder structure, showing two stacked base-pairs. The base-pairs are here stacked in a perhaps-*too*-perfect "sandwich" conformation. In 2005, when I was actively working on protamine-DNA structure, I naively presumed that "sandwich" stacking was the energetically most-favorable arrangement. I had very little background in aromatic ring chemistry back then, and I have only just recently learned that there is an energetically-*more*-favorable mode of stacking, called "parallel displaced", which can be brought about by simply pulling the duplexes apart a few angstroms, like so:

SLIDE 91

This small displacement would presumably improve the energetics of stacking, without significantly affecting the packing of protamine-DNA into the sperm head.

SLIDE 92

This is best appreciated in this longitudinal view that we looked at earlier. All we're talking about is taking my original, close-packed model for stacking, and pulling the unit cells apart a few angstroms. Watch closely, because the move is very small...

There it is. This slight movement converts the less-favorable sandwich stacking, which I originally proposed in 2006, to the more favorable "parallel-displaced" stacking. One could question whether this move might impair the packing of protamine and DNA into the sperm cell nucleus, but there's no chance of that. Since the protamine-DNA complex, in my original model, required no more than about 10% of the available volume of the sperm head anyway, this minor adjustment would in no significant way compromise the ability of the model to readily account for the packing of the nucleoprotein into the available space.

SLIDE 93

"Another form of tetraplex DNA: the Gehring tetramer"

SLIDE 94

The second of the two tetraplex structures we must consider is the Gehring tetramer. This, in many respects, is a prototype for the Wu tetraplex. But not in every respect.

The Gehring tetramer is covered in some detail in "The Double Non-Helix, Part II, The Probable Structure of the Protamine-DNA Complex", on this web site. But there are certain key aspects of it that were *not* covered, and which we must look at now.

Here's how the tetramer is constructed from its 4 component strands. Shown here is the 1st strand...

SLIDE 95

... now the 2nd...

SLIDE 96

... now the 3rd...

SLIDE 97

... and finally the 4th. Put them all together, and you get this:

SLIDE 98

Like the Wu tetraplex, the Gehring tetraplex, except for its 4 sugar-phosphate backbones, looks a lot like ordinary DNA, even more so because it's helical.

SLIDE 99

The tetramer is created from the DNA hexanucleotide sequence "T-C-C-C-C-C", at the acid pH of 4.3. At that pH, cytosine forms base pairs with itself, which are known as "C-dot-C+", because a single proton is shared between the two imino N3 nitrogen atoms.

The crystal structure of the Gehring tetramer is in the Protein Data Bank under the accession number 225D.

SLIDE 100

When one examines the Gehring tetraplex structure, one sees that it is composed of two identical DNA duplexes, running antiparallel to one another. Shown here is the first of the two. These are the "A" and "C" strands of the tetramer. This duplex has some peculiar properties. If you have a good eye for nucleic acid virtual models, you'll notice

immediately that the base pairs are very far apart. There's another peculiar feature which will be evident if you look at the ribose O4' atoms, at the apices of the pentagonal ribose rings. You'll notice that they're pointing up. Whichever strand you're looking at, if you now look at the other strand, you'll see that *those* O4' atoms are also pointing up. In other words, these strands are *parallel*, unlike the strands of ordinary DNA, which, as even elementary biology students all know these days, are *anti*parallel. So this is a structure you're not going to see every day of the week.

SLIDE 101

Here's the second duplex, comprising the "B" and "D" strands. If you look at it out of the corner of your eye, it looks just like the first. But when you look more closely, you see that these stands also are parallel, except that here, the O4' oxygen atoms of the ribose rings are pointing *down*.

SLIDE 102

In other words, the complete tetramer, shown here, is composed of a pair of *anti*parallel duplexes, each duplex, however, consisting of a pair of strands which coil about one another in *parallel.*

SLIDE 103

The H-bond lengths in the base pairs are very short, averaging 2.13 Å. These are much shorter than the 3 Å base spacings of commonplace W-C DNA. As we shall have need to discuss again later, hydrogen bond strength is related to size. At 2.13 Å, these are very strong hydrogen bonds, approaching the strength of covalent bonds.

SLIDE 104

This tetraplex has a helical twist, which I find fascinating. The helix is right-handed, and rather narrow in diameter, due to the shortness of the 2.13 Å hydrogen bonds. Within each of the component duplexes, the stacking distance between base-pairs is also small, namely 6.14 Å. That means that after intercalation, the base-stacking distance will be 3.07 Å, considerably more compact than the usual 3.4 Å found in B-DNA.

SLIDE 105

While this small stacking distance may seem unfavorable, it must be considered in the light of the relative orientations of successive layers of base-pairs, which are regularly offset by a full 90º! Here's an axial view; the faint base-pair is the next level down. Look at the offset!

Of the three types of aromatic ring stacking of which I'm aware, namely sandwich, parallel-displaced, and T-shaped, I'm not even sure how to classify Gehring. I guess you'd have to call it "parallel-*extremely*-displaced", wherefore the measured

distances between *comparable* atoms in successive levels of the structure are not only far greater than 3.07 Å, they are, in fact, a lot greater than comparable spacings in B-DNA. In the Watson-Crick structure, at its canonical base-pair spacing of 3.4 Å, direct, "as the crow flies" interatomic distances between comparable atoms in adjacent base pairs, average out to about 4.3 Å; here comparable atoms are spaced at an average of about *5 Å.*

SLIDE 106

The Gehring tetramer forms by association of four short single strands of DNA. The question of whether it would be physically possible for the same structure to form from a much longer supertwisted circular chromosome is difficult to answer, but I think it's at least possible.

In order to visualize this, consider this wildly-distorted picture of a chromosome, showing a G-C base pair in the posterior half of the two strands, intercalated with an A-T base pair in the front. This manner of intercalation is the basis for a 4-stranded structure according to the rectilinear model of Tai Te Wu. In order to convert this into a Gehring tetramer type of structure, we'd have to perform the gymnastics I'll now attempt to depict in the drawing at the bottom of the slide, which is a top view, that is, a view looking down upon the chromosome from the top of the slide. To employ the terminology of an American football game, what the drawing will attempt to show is the entire posterior half of the circular strand on the right, represented by the cytosine residue, separating from its base pairs with the strand on the left, and then doing an "end run" around its own anterior half, that is, the half of the strand bearing the thymine residue, to re-position itself in front of the A-T base pair. The H-bonds between the strands then re-form, in this new position.

I cannot imagine this happening in a single motion, but I can imagine that under denaturing conditions, small segments from the rear of the strand on the right might thusly relocate themselves, forming a nidus for a cooperative conversion of the entire chromosome to a Gehring-type of conformation.

Such a thing, if possible at all, would only be imaginable under the most extenuating of circumstances, but in considering the origin of Form IV, we're going to be considering DNA at pHs just short of 13, which is about as extenuating as things can get without total hydrolysis of the DNA into a multitude of fragments and individual nucleotides.

SLIDE 107

In any event, the features of the Gehring tetramer are, I think, very informative with respect to the structure of tetraplex DNA generally. Like W-C DNA, Gehring DNA strands each have a right-handed helical twist, which is smooth and continuous. Now, the pitch of each of the Gehring pair of duplexes is markedly greater than that of the W-C single duplex, and yet the basic character of the Gehring sugar-phosphate backbone, namely smooth and continuous, is the same as that of B-DNA, that is, it does not display any of the peculiar "zig-zag" characteristics of left-handed Z-DNA.

There's no reason to doubt that the Gehring tetramer could, in theory, be modeled as a "straight ladder", non-helical structure, but I am not enough of a virtual chemist to

explain why it is not so, but rather helical. The answer, of course, lies in the concurrence of conditions resulting from the short and powerful hydrogen bonds, the peculiar form of 90º-offset base stacking, and the Van der Waals's interactions of the atoms comprising the sugar-phosphate backbone; but which of these are most important in determining the helical twist, and exactly how they bring the twist about, I cannot say.

SLIDE 108

Nevertheless, taking all the data into consideration, I am rather inclined to believe, at this point, that close spacings — of both base-pairing and base stacking — favor the right-handed sugar-phosphate backbone, whereas larger spacings favor the left-handed, or Z backbone. The closest I can come to providing a molecular explanation for this is to cite the stretch of sugar-phosphate backbone between what I referred to as the "fundamental structural units" of Z-DNA, which we now re-visit. Since these "functional units" are stacked at nearly 4 Å, which is not known to be a favorable benzene-ring-type stacking distance, we must tentatively conclude that it is a mechanical limitation of the left-handed sugar-phosphate backbone that mandates this spacing.

SLIDE 109
"Phenomenon #3: DNA base tautomers"

SLIDE 110

We need to get ourselves grounded in basic DNA tautomer chemistry, because it's possible that Form IV, under certain conditions, may exist as a duplex which is stabilized by tautomeric base-pairing.

Tautomers are alternative DNA base structures, which pair differently than the standard or "canonical" forms (*i.e.*, the familiar base pairs A-T and G-C). When free in solution, the forms exist in equilibrium, but the equilibria heavily favor the canonical forms.

DNA base tautomers have been of interest for a long time, because it is believed that one of the mechanisms of mutation is the chance unfortunate happening of a base converting transiently to a tautomeric alternative form at the exact moment of DNA replication.

SLIDE 111

Base tautomers form because of the translocation of hydrogen ions.

Here are some of the known base tautomers. This chart is adapted from the 1976 Nature manuscript by Topal and Fresco; an oft-referenced publication in this field. I'll run through it quickly, then show you how the tautomers come about.

First of all, in the upper left, we see the so-called 'canonical' base pairs; that is, A-T and G-C; the ones we learn about in school. But look below them: here we see not just one, but two examples each of A-C and G-T, which are not supposed to exist in nature. On the right side of the figure it gets even more exotic; here we see four

examples of a phenomenon few of us ever have cause to even think about: purine-purine base-pairing, namely A-A, G-G, G-A and A-G.

Let's see how this comes about. Consider the first 'canonical', or standard W-C base pair on the upper left, A-T. Let's look a little closer.

SLIDE 112

This base pair is stabilized by two hydrogen bonds, one connecting the amino group of adenine to the upper keto oxygen of thymine; the second connecting the two imino ring nitrogen atoms.

SLIDE 113

But look what happens if we try to substitute cytosine for thymine: now we not only do *not* have bonding, but actually *repulsion.* The two amino groups are not going to like being forced together. And the imino ring nitrogens have no proton at all by which to bond.

SLIDE 114

But we can remedy this situation by moving an amino proton in cytosine down to the imino ring nitrogen, like so...

After adjusting the double-bond structure accordingly, *voila!* We now have a new form of cytosine, called "imino cytosine", which hydrogen-bonds perfectly with adenine! Let's bring in the H-bonds, to make this easier to see.

There it is: a "C-A" base pair; something most students will never learn about in biology school.

SLIDE 115

We can achieve the same outcome by doing the same amino→imino proton shift on the adenine. Let's backtrack for a moment, and bring back the original 'canonical' cytosine tautomer. As we have previously seen, cytosine and adenine cannot cohabit, because there are only repulsions and no attractions. But if we move one of the adenine amino hydrogens down to the imino ring nitrogen, and adjust the double-bond structure accordingly — *voila!*— we now have a new form of adenine which hydrogen-bonds perfectly with cytosine!

SLIDE 116

Now let's look at a 'canonical' Watson-Crick G-C base pair. Here we see three good H-bonds, but if we replace C with T, we get this awkward thing...

Although there's a possibility of a single H-bond at the bottom, the imino hydrogens in the rings will offset that by repulsion, and the two oxygens at the top are too far apart to interact at all, which is actually good, because oxygen atoms get *really* temperamental if you try to get them closer than about 3 Å.

This natural incompatibility can be solved by doing the opposite of what we did for the A-C tautomeric base-pair, namely by moving the thymine ring imino hydrogen up to the keto position, to give the enol tautomer of thymine, resulting in a T_{enol}-G base pair!

And once again, we can achieve the same result by enolizing the guanine instead. I'm not going to go through all the steps; I'll just show you the outcome. And here it is: a T-G_{enol} base pair, something you don't learn about in school.

SLIDE 117

Now I'll show you the even more exotic-looking purine-purine base pairs.

Let's start with adenine. Here's standard Watson-Crick adenine. DNA bases can potentially exist in two different rotational forms, called "anti" and "syn". In ordinary DNA, the bases are believed to be in the 'anti-' conformation, which is what you see here. Watch what happens when we rotate about the glycosidic bond; that is, the bond which connects the base to the DNA sugar-phosphate backbone: The adenine is now in the "syn" conformation; that's "s-y-n", not "s-i-n"! Human "sin" has nothing to do with this!

SLIDE 118

Now we'll bring in *another* adenine, to pair with our syn-adenine. Well, this obviously isn't going to work. What we need is our old friend imino adenine. *Voila!* We adjust the double bonds to the new conformation , bring in some H-bonds, and we now have something that many scientists, including myself, never knew possible: an adenine-adenine base pair!

SLIDE 119

Here are 3 more purine-purine base pairs, each of which arise in the same manner as the tautomeric "alternative" base-pairs we have already seen, that is, from some combination of anti→syn, keto→enol, and amino→imino transitions.

SLIDE 120

First we have G=A; that is, $G_{\text{-SYN}}$=$A_{\text{-IMINO}}$.

SLIDE 121

Then G=G; that is, $G_{\text{-SYN}}$=$G_{\text{-ENOL, IMINO}}$.

SLIDE 122

And finally, A=G; that is, $A_{\text{-SYN}}$=$G_{\text{-ENOL, IMINO}}$.

Having now briefly reviewed the chemistry of DNA base tautomerization, we can see that virtually any base can form a base-pair with any other. But we must now ask how *likely* this is.

SLIDE 123

Based upon solution studies of purified nucleotides, the fraction of bases which are spontaneously in tautomeric forms, at any given moment in time, is said to be in the neighborhood of 10^{-4} to 10^{-5}. Assuming these numbers would be approximately the same for bases covalently bound within DNA molecules, it becomes clear that "alternative" base-pairings are *not* energetically favorable, and are therefore *not* going to be seen with any great frequency under normal circumstances.

SLIDE 124

But consider the case of a circular DNA molecule which has been converted into Form IV, represented here by two circular strands whose bases are totally out-of-register. We know that the two strands of Form IV remain associated, and that, under the right conditions of pH, temperature and ionic strength, they can be renatured. As explained in "The Double Non-Helix, Part I", it can be deduced that at some point in the *re*naturation, conformational changes which can be represented as a relative circular motion of the two strands *must* take place, if the proper base-pairing is to be re-established, as shown here.

That being the case, we may therefore represent the *denaturation* process as being the reverse, namely a relative circular motion which *dis*-establishes proper base-pairing, as shown here...

As I shall try to show shortly, Form IV is almost surely a structure where the two strands remain well-aligned, base-for-base, although the bases are not properly paired, due to circular drift. The question shall therefore arise, as to whether, at any position in the structure, the presence of an apposed base might *induce* an otherwise unfavorable tautomeric shift, because of the free energy decrease associated with tautomeric base pairing.

SLIDE 125

Calculating the energetics of this has been difficult, since tautomer stability studies are typically done with isolated bases in solution, and their results must therefore be extrapolated back to DNA. According to these studies, the relative ΔG necessary, for the conversion of the normal tautomers to the "alternative" forms, seems to fall generally in the range of $5 \rightarrow 10$ kcal/mol. For isolated bases in solution, these energetics are not very favorable. But when the bases are polymerized in the form of DNA, these unfavorable numbers may be offset by the free energy benefit of base-pair formation, which is in the $-5 \rightarrow -10$ kcal/mol range. These numbers are obviously too vague to constitute proof that non-complementary DNA duplexes can actually exist in the real world; only that they are possible in theory.

SLIDE 126

"Phenomenon #4: the maximum number of supertwists (in a 5 kb chromosome) is 25"

SLIDE 127

In order to understand this section, you must be familiar with the topology equation for circular DNA ($L_K = T + W$).

L_K, the first term in this expression, is called the "linking number", which may be defined as the total number of secondary helical twists in a circular chromosome, when it is forced to lie in a plane, without tertiary twists. Tertiary twists are usually referred to as "supertwists". Most circular chromosomes, as isolated in nature, are supertwisted, wherefore, in the picture on the left side of the slide, for demonstration purposes, we have imagined removing the supertwists by means of imaginary pushpins.

The value of the parameter L_K is totally independent of experimental conditions, that is, it has nothing to do with pH, temperature or ionic strength, but is, rather, an immutable, built-in topological characteristic, which is permanently established at the moment the strands of the chromosome are closed into a circle.

SLIDE 128

But in real life, circular DNA, of course, is *not* compelled to lie in a plane, but rather, it takes on tertiary windings, *i.e., supertwists,* as depicted in the drawing on the right. Tertiary supertwists can either increase or decrease the *secondary* helical winding.

The second term in the topology equation, "T", is the number of secondary helical twists we can still count *after* the chromosome takes on its inevitable tertiary supertwists. Note that the teleological purpose of supertwisting is to correct the secondary winding when a chromosome is either underwound or overwound. Thus, "T" will generally have the standard Watson-Crick winding number of {1 helical twist} per {34Å of length}. Unlike L_K, however, "T" *is* dependent upon ambient conditions of pH, temperature and ionic strength.

"T" and " L_K " are always *about* the same, but not exactly the same. If, for example, the chromosome is "underwound", like so, then L_K will be *less* than the standard Watson-Crick winding number, but that standard winding number will be restored by supertwisting, so that "T" will be a little greater than L_K.

Conversely, if the chromosome is "overwound", like so, then "T" will be a little *less* than L_K.

SLIDE 129

The third term, "W", or "writhe", is the number of supertwists. For the purposes of this equation, molecular topologists have seemingly-arbitrarily defined right-handed supertwists as being numerically *negative* for computational purposes, and left-handed supertwists as being positive.

There's something a bit cognitively-dissonant about this, since right-handed *secondary* helical twists are counted as being *positive,* and left-handed *secondary* helical twists negative. To re-state this succinctly, a right-handed *secondary* helical twist is positive, but a right-handed *tertiary super*helical twist is negative. Conversely, a left-handed *secondary* helical twist is negative, but a left-handed *tertiary super*helical twist is positive.

If you have good powers of mental visualization of things, or if you decided to take the trouble to work with string or rubber-band models, you'd immediately grasp the functionality of these definitions. You'd see that a right-handed tertiary supertwist *removes* a left-handed secondary helical twist, whereas a left-handed supertwist removes a right-handed secondary twist. So the mathematical definitions are *not* arbitrary after all.

Moreover, these arithmetic definitions are what enables the topology equation to work. Accordingly, the equation shows that for a superhelically-twisted chromosome, the sum of the *secondary* helical twists, plus the tertiary supertwists, is a constant, equal in value to L_K. Thus, if the chromosome is "underwound", "W" must be negative, that is, the supertwists will be right-handed, as shown. If the chromosome is "overwound", "W" must be positive, that is, the supertwists must be left-handed, as shown now.

If any of this seems confusing, then you need to watch at least SLIDEs 263-269 of "The Double Non-Helix, Part I", on this web site. At the present time, all you need to know is that for any superhelical chromosome, as isolated from nature, the sum of {the number of observed secondary helical twists + tertiary *super*helical supertwists} is a constant, equal to the linking number, L_K.

The moral of the story is that changes in the secondary winding cause exact but opposite changes in the tertiary winding, and *vice versa*.

SLIDE 130

I argued, in "The Double Non-Helix, Part I", that the reason that native DNA is supertwisted is *not* because it's mysteriously "underwound", as our current textbooks all teach, but rather because its topological net *non*-helicity causes it to be, when forced to lie in a plane, 50% right-handed and 50% left-handed, that is, it has a linking number, L_K, of exactly zero.

This can be envisioned, for simplicity, as a structure such as "A", in which there are exactly 3 right-handed turns on the top of the drawing, and 3 left-handed turns on the bottom. This is a topological statement; the distribution of right-handed and left-handed secondary turns need not be quite this simple. In certain non-helical models that have been proposed, right-handed and left-handed helical regions alternate regularly, but are only ½ twist in length, giving rise to a structure such as that shown in "B", where there are no complete twists at all, but merely a meandering ribbon-like structure which twists a bit to the right, then a bit to the left, without ever completing a full helical turn.

In most of the arguments that follow, it will often seem that I'm assuming model "A" to be the correct one, but that's merely for convenience in drawing and explaining things. In fact, models "A" and "B" are entirely equivalent, *topologically*-speaking, and both would be expected to behave identically in any setting in which ambient conditions impacted on chromosome configuration according to the topology equation we discussed above.

SLIDE 131

It is perfectly well-known that *linear* DNA, which has no topological constraints, "prefers" to be right-handed under physiological conditions, whence we may deduce that

a *circular* chromosome will also do whatever it must to maximize its right-handedness. According to our *non*-helical point of view, a chromosome the size of ϕx174, about 5000 bp in length, *if* constrained to lie in a plane, would take on a secondary helical twist for about every 10 base pairs of length, or about 500 twists in total. Two-hundred-fifty of these would have to be right-handed, which are energetically-favorable, and 250 would have to be left-handed, which are energetically *unfavorable.* Since, when constrained to lie flat in a plane, there are no supertwists, "W=", the 'writhe', *i.e.*, the number of supertwists, is zero. Therefore L_K and T will be exactly the same number; that number being the sum of the right-handed and left-handed secondary twists, which is {250-250}, or *zero*.

SLIDE 132

Because the left-handed secondary twists are energetically *less* favorable, the chromosome will therefore take on about 25 right-handed, or *negative* supertwists, each of which removes one of the "unwanted" left-handed secondary twists. The Topology Equation, in this case, has a rather trivial appearance, but it still states the facts succinctly: "0 = 25-25"! On the next line, we see a more descriptive way to state the case: The left side of the equation shows that the L_K of zero is the sum of 250 right-handed and 250 left-handed secondary twists. The right side of the equation shows that 25 of the unwanted left-handed twists are removed by the addition of 25 right-handed, or negative *supertwists.*

That number, 25, does not arise from any theory presented here, but rather from years of experimental observations on native chromosomes in many laboratories. But, one might ask: *If,* as we allege here, native DNA has no net helical twists, *i.e.*, has a linking number L_K of 0, *why* would there be *only* 25 supertwists?

SLIDE 133

To answer this question, let us return to the maximally-simplistic view introduced a moment above, where the chromosome has been forced to lie flat on a table, with no supertwists.

This simplistic picture shows only 6 twists in total. Let us use our imagination to add more. For ϕx174, the number would be *500* twists; 250 right-handed, and 250 left-handed. We know that the left-handed turns, under normal physiological conditions, are *all* energetically *unfavorable.* If, as the topology equation teaches, right-handed *tertiary* supertwisting can remove unwanted *left-handed secondary* twisting, why then doesn't the ϕx174 chromosome, at neutral pH, take on *more* than 25 supertwists? Why not 50, or 100, or even 250 right-handed supertwists? If the chromosome took upon itself 250 right-handed supertwists, then every single unwanted left-handed secondary twist would be unwound, leaving only 250 right-handed secondary twists, and 250 right-handed supertwists. The chromosome would then be 100% right-handed, and everyone would live happily ever after. No?

No! The answer to the question is that anything above 25 supertwists will introduce a prohibitive amount of strain into the chromosome, and will therefore *never be seen.* Whether that strain would be due to phosphate-phosphate charge repulsions, or

Van der Waals repulsions, or backbone strain, or all of the above, cannot at the present time be precisely specified. But, be that as it may, beyond 25, no further superhelical twists are possible.

SLIDE 134

What gives us the right to say this? Well, first of all, there's the evidence from the Rush-Warner ϕx174 sedimentation vs. pH curve we have been looking at:

The dip whose center is at pH 11.8 is quite symmetrical on either side of the low point. At the beginning of the dip, at approximately pH 11.6, it is known from various biochemical analyses, and from electron microscopic observations, that the superhelix twist count for ϕx174 is indeed about 25. This supertwist count pertains to the chromosome at all pHs from 7 up to the beginning of the pH 11.6 dip, and is not affected by changes in pH in that lower range.

I should point out that the exact supertwist count in the little picture above this portion of the data is apocryphal, but the direction of superhelical winding shown, namely right-handed, or negative, is correct.

SLIDE 135

It is generally assumed by all authors that at pH 12.0-12.1, the other side of the dip, the conformation is very similar, except that it's opposite in winding sense. At that higher pH, we note that the sedimentation coefficient returns to its pre-dip value, which, for ϕx174, under the conditions of this experiment, is 21***s***. There is therefore no reason to doubt, *a priori,* that at pH 12.0-12.1 there are 25 *left-handed* superhelical turns. And that is very likely the upper limit of superhelicity for this chromosome. While it's theoretically possible that raising the pH further might bring about a further increase in superhelicity, I doubt very much that that would ever be seen.

Why? Because any further increase in pH brings us into the shoulder at pH 12.3, labeled χ, where it is clear that some sort of conformation change is taking place. As to what that conformational change is, we shall be addressing that shortly.

If the chi shoulder was the only evidence that 25 supertwists is the maximum number possible, then the argument would be weak. But this is only the beginning. Let's look now at other data which support the reality of the 25 supertwist limit.

SLIDE 136

Here's an agarose gel electrophoresis study of a typical topoisomerase experiment, where DNA is incubated to equilibrium, to yield the complete set of topoisomers possible. This data is from the Stanford University Biochemistry 201 course; the plasmid used was about the same size as ϕx174:

Topoisomerase randomly nicks and re-seals DNA, giving rise to a set of so-called 'topoisomers', each of which differ from their gel electrophoresis neighbors by ±1 superhelical twist. If DNA is allowed to remain in the presence of topoisomerase for a while, an equilibrium "family", so-to-speak, of topoisomers is seen. Within this "family", the supertwist count can assume any value from zero (which will co-migrate with the

nicked, or 'relaxed' chromosome) to the maximum possible (which will co-migrate with the fully-intact native chromosome).

SLIDE 137

Each distinct band in the agarose gels shown here represents a single topoisomer. The native chromosome is at the bottom, where it co-migrates with the highly-supertwisted topoisomers.

(Note that the supertwist count in the little picture shown is again apocryphal; for ϕx174 the count would be 25, but, for visual clarity, we're only showing about half of those supertwists).

As the topoisomerase reaction proceeds, topoisomers of lesser supertwisting appear, culminating in the zero-supertwist topoisomer, which co-migrates with the nicked, or "relaxed" chromosome.

On the left, the gel has been treated to maximize separation of the topoisomers with low supertwist counts, which would otherwise be hopelessly clumped together (as they are on the right). Conversely, on the right, the gel has been treated to maximize separation of the topoisomers with *high* supertwist counts, which also would be otherwise hopelessly clumped together (as they are on the left). Between the two gels, you can see the complete "family" of topoisomers for this particular species of DNA.

SLIDE 138

Here's the point: *Note that the native chromosome has the same electrophoretic mobility as the most-highly-supertwisted topoisomers,* that is, *no topoisomer has a higher electrophoretic mobility than that of the native chromosome.* That is *always* the case, regardless of the DNA species.

This needs to be explained, but how can we explain it? In "classical" Watson-Crick theory, DNA is an all-right-handed double-helix, which, when circularized, is sealed shut with a 5-7% deficiency in Watson-Crick secondary right-handed turns. To employ the terminology of the helicists, the DNA is supposedly "underwound". This arbitrary and capricious theory is alleged to account for the observed right-handed superhelicity of native DNA, which in turn is alleged to be somehow necessary for the packaging of DNA into viral capsids and cellular organelles. If so, then the fact that no topoisomer *ever* has a higher electrophoretic mobility than the native chromosome, would then have to be regarded as one hell of a "coincidence". That is, the maximum superhelicity observed in a topoisomerase experiment is merely the result of nicking and re-sealing of a chromosome being buffeted by random thermal motion in the laboratory. It has nothing at all to do with biology. *Why* should the electrophoretic mobility of the native chromosome, supposedly reflecting some *biologically-mandated* DNA-packaging requirement, invariably be exactly the same?

In Watson-Crick theory, it makes no sense, and must be regarded as merely an extraordinary coincidence.

On the other hand, if we accept the fact that DNA has, in reality, *no* net helical twists, then the native superhelicity merely reflects the chromosome's thermodynamic need to maximize the right-handedness of the secondary helical winding, by converting

some of the energetically-unfavorable left-handed secondary twists into right-handed tertiary supertwists. That process will proceed until the upper limit of permissible superhelicity is reached. Since that very same upper limit is exactly what would be predicted in a topoisomerase experiment, it's no longer a mystery at all why the electrophoretic mobility of the most highly-supertwisted topoisomers never exceeds that of the native chromosome; both electrophoretic mobilities result from precisely the same structural restraints.

SLIDE 139

In a simple topoisomerase experiment, each band consists of a *pair* of topoisomers, one having a unique integral number of right-handed, or "negative" supertwists because of underwinding, and the other having the same number of supertwists, only in the opposite or left-handed sense, because of *over*winding. This is now indicated in the slide, where you'll see, if you look closely, that all the topoisomers on the left-hand side of the slide are drawn as left-handed superhelices, whereas all the topoisomers on the right-hand side of the slide are drawn right-handed. For any given supertwist count, the left-handed and right-handed superhelical forms have essentially the same electrophoretic mobility, wherefore they form a single band.

Since the maximum number of superhelical turns is the number seen in the native chromosome, it follows that the maximum number of right-handed, or "negative" supertwists will always be exactly the same — or at least very nearly the same — as the maximum number of left-handed, or "positive" supertwists. For ϕx174, that number is about 25.

SLIDE 140

Further evidence for the existence of an upper limit to superhelicity comes from experiments with the intercalating agent, EtBr.

Here's an old ***s*** vs [EtBr] concentration titration by Bauer and Vinograd, employing the DNA of the oncogenic virus SV-40, which is about the same size as ϕx174, and which, like ϕx174, has about 25 right-handed supertwists, as isolated in nature. The analysis shown here is, once again, based upon the sedimentation coefficient, determined by velocity ultracentrifugation.

EtBr intercalates itself between base pairs, stretching out the DNA, and thereby causing it to unwind. If the DNA is linear, the EtBr can intercalate between each base pair, because there's no strain induced in so doing:

If, however, the DNA is closed into a circle, the chromosome supertwists upon binding EtBr, which eventually causes steric hindrances. As we shall see in a moment, at saturation binding, less than half the base pairs of a circular chromosome are able to accommodate an EtBr molecule.

SLIDE 141

The data for ***s*** vs. [EtBr] are similar to the ***s*** vs. pH titration data we have been looking at up to now. The native chromosome, which is known to be a right-handed

superhelix, has a sedimentation coefficient of about 21 ***s***. As the [EtBr] concentration is increased, supertwists are progressively unwound, and ***s*** decreases, exactly as it did in the alkali denaturation curve described above. Let's bring in some pictures, to keep the superhelical windings in mind.

Eventually ***s*** reaches a minimum, at which point all the supertwists are unwound, and the chromosome is an open circle. If further EtBr is added, however, left-handed supertwists appear, and the sedimentation coefficient begins to increase again, eventually returning to 21***s***, at which point it levels off. The leveling off is specifically *not* because each inter-base-pair position is occupied by an EtBr molecule -- on the contrary, as the ordinate shows, at the maximum ***s*** value attainable in this experiment, the molar ratio of EtBr to nucleotide residue was only about 0.1. The SV40 chromosome has plenty of room for more EtBr, so why can't the superhelicity be pushed higher? The obvious answer is that approximately 25 supertwists is the upper limit of what's possible, whether in the native state, where the supertwists are right-handed, or in the state of saturation binding of EtBr, where the supertwists are left-handed.

In case you may be thinking "What if we just added a lot more EtBr? Mightn't we then overcome the 25-supertwist limit?" Charles Weissmann did just that.

SLIDE 142

If you looked at my "The Double Non-Helix, Part I" slide show, then you know that I'm no fan of the work of Charles Weissmann. But it's only his conclusions I disagree with. His laboratory technique and experimental design, however, are beyond reproach. In this figure he shows the effect of EtBr on the electrophoretic mobility of DNA from the virus PM2, which has about 10,000 bps, and therefore about *50* right-handed superhelical twists in the native state. PM2 is a little larger than ϕx174, and the other viral chromosomes we have discussed up to now.

The apparatus shown in the figure holds some 30 agarose lanes, and each one has a different concentration of EtBr, with the [EtBr] concentration increasing toward the right. The DNA in each gel is a mixture of Form I, Form II, and the so-called "Form V" (the last of which forms we shall be discussing presently). These three forms are labeled on the left-hand side of the figure, which also shows the direction of electrophoretic migration.

SLIDE 143

We can ignore Forms II & V, which are irrelevant to the subject matter at hand. Look only at the Form I, that is, the native DNA bands. As you move your eye from left-to-right, in the direction of increasing EtBr concentration, the electrophoretic mobility of Form I *decreases*, reflecting an unwinding of superhelical turns. Because the Form I, when fully unwound, co-migrates with Form II, I shall add a guide line to help us visualize the location of the Form I bands:

Now we can clearly see that increasing the EtBr concentration first unwinds the right-handed superhelical turns, culminating in an open circle which co-migrates with Form II, the latter of which is also an open circle. Let's bring back our little superhelix drawings, to illustrate this.

SLIDE 144

Further EtBr causes the appearance of left-handed superhelical turns in the Form I, culminating in a symmetrical re-establishment of the original electrophoretic mobility. Note that after the electrophoretic mobility returns to the native level, about halfway across the figure, there is *no further change brought about by additional EtBr.* That's because the DNA is fully-wound, superhelically, in the left-handed sense, and no further supertwists can be added. And the number of supertwists at that point, as I've been saying all along, *is the same as the number in the native state*, which, in this case, is not 25, but 50, reflecting the larger size of this particular chromosome.

SLIDE 145
"Phenomenon #5: 'Form V' DNA is really Form IV"

In 1979, in the paper whose electrophoresis results we have just examined, Charles Weissmann and his associates described what they referred to as a supposedly "novel" form of circular DNA, which they dubbed "Form V".

SLIDE 146

This was supposedly a base-paired duplex created by reannealing the previously-isolated circular single strands of the parent chromosome. That paper was thoroughly reviewed in "The Double Non-Helix, Part I", on this web site, but we're going to have to re-review some of it here.

The Weissmann experiment was done in response to a public suggestion by Francis Crick, who was himself responding to the 1976 publication of the non-helical DNA structure of New Zealand engineer Gordon Rodley:

SLIDE 147

Rodley *et al* had proposed that the true structure of DNA was what they called the "side-by-side" structure, which was similar to the W-C structure, except that instead of being all right-handed, the direction of helical winding was variable. Specifically, every 5 base pairs the direction of helical winding in the Rodley structure was *reversed*, so that the net number of helical twists was exactly zero.

SLIDE 148

This is better seen in this ribbon model, which has a length of 30 bp. A W-C double helix of this length would have 3 complete right-handed helical twists, wherefore you would see the strands crossing one another three times. But the side-by-side model shown here has only 3 right-handed ***half***-twists, combined with 3 ***left-handed*** half twists, for a grand total of *zero* twists, or, to use our current terminology, an L_K of zero when circularized. If you watch the model as it partially rotates, a bit to the right, then to the

left, you'll see that neither strand ever actually crosses the other entirely, wherefore there are no net helical twists.

SLIDE 149

The strands of a circular chromosome having this structure would be possible to separate without strand breakage, which, as of 1976, the time of the Rodley publication, had never been done. That is, as we have previously noted, when naturally-occurring circular DNA is subjected to conditions which are denaturing for *linear* DNA, the strands of the *circular* chromosome do not separate. But Rodley and others suggested that perhaps there was some undiscovered reason for that, and that if we knew how to do it, we *could* separate the strands. The logical argument goes like this: Whereas separating the strands proves that they can be separated, *not* separating the strands does *not* prove that they can't be separated.

To lay the matter to rest, Crick suggested deliberately creating an artificial circular chromosome with $L_K = 0$, and proving that it was not normal DNA. To do this, it would be necessary to *destructively* separate the strands of a plasmid or viral chromosome, by introducing small numbers of nicks, denaturing the nicked chromosomes, and purifying the intact circular single strands that still remained.

SLIDE 150

Then, Crick suggested, those separated single strands should be reannealed. His reasoning was that the expected product of such a reannealing, namely a side-by-side duplex with $L_K = 0$, couldn't possibly have the same physical properties as naturally-occurring DNA, because it couldn't possibly have a net helical twist.

Note that the Watson-Crick structure, with its numerous right-handed helical twists, *cannot* arise in this experiment, because there's no way to introduce a net helical twist to a duplex created by the reannealing of separate strands that are not topologically intertwined.

So it was simply a matter of creating the side-by-side duplex, and showing that its physical properties were abnormal, and not those of real DNA.

SLIDE 151

Weissmann and his associates performed the Crick experiment, creating a duplex with $L_K=0$, by reannealing single-stranded circular DNA at 60º, pH 8.5. They called their new structure "Form V", and readily demonstrated that its physical properties were indeed abnormal, supposedly laying to rest forever the proposition that the strands of circular DNA were not topologically intertwined.

But there was a fatal flaw in the Weissmann reannealing protocol. As I pointed out in "The Double Non-Helix, Part I", Slides 236-244 (partially repeated here in Slides 23-30), circular DNA cannot possibly be reannealed at pH 8.5. And yet the structure Weissmann created was undoubtedly circular DNA with $L_K=0$. But was it a properly base-paired structure? Or was it some sort of anomalous duplex, lacking complementary base-pairing?

SLIDE 152

Weissmann's answer to this question was "properly base-paired", but he presented no direct evidence for that. Rather, quite to the contrary, he presented the melting curve shown here, which is totally lacking in cooperativity. Note that the Form II control experiment curve displays the typical shape of a cooperative melting curve, but the Form V melting curve is better described as *linear*, and in any respect in which it differs from linearity, it's certainly not in the direction of a cooperative acceleration of the melting as the temperature increases, but rather the opposite. These results essentially rule out true complementary base-pairing, wherefore it can be said that Weissmann's own data seem to argue against his own stated conclusions.

If Form V was really what Weissmann said it was, namely a *properly-base-paired* duplex with $L_K = 0$, then this entire NotAHelix web site, all of the presentations and other data on it, and my life in general have all been for naught, because the properties of Form V, as Crick had predicted at the outset, were indeed *abnormal*, that is, *different* from those of the native DNA from which its strands were derived. Therefore, either I'm wrong, or Weissmann was wrong – we cannot both be right.

SLIDE 153

Why dwell on the old Weissmann paper now, nearly 40 years later? Because ever since the day it was first published, it has been almost universally presumed to have provided *conclusive* evidence that native DNA has a net helical twist. This implies that non-helical DNA structures, such as those presented here at NotAHelix.net, and other similar structures presented by other authors at other places, all of which have $L_K=0$, are all works of fiction that have no biological relevance.

That single paper therefore poisoned the minds of an entire generation of scientists, and is now doing the same with the current generation. So, you see, we *cannot* ignore that paper, because a whole universe of false DNA structural theory rests upon its very slim pedestal.

Our defense, which we can assert with considerable certainty, is that the term "Form V" is almost surely a misnomer, since — as we shall see — it is probably nothing more than Form IV at low salt concentration. But we still need to consider Weissmann's *data*, because, while we cannot accept his *interpretation* thereof, there's no reason to otherwise doubt its veracity.

SLIDE 154

The key finding that Weissmann reported in 1979 was that when the previously-separated single strands of a circular plasmid or viral chromosome were reannealed at pH 8.5, 60º, in 0.5 M NaCl, a new duplex DNA form appeared, the above-referenced "Form V", which was demonstrably *different* from native DNA. To demonstrate this, the various forms were examined by gel electrophoresis.

Under the conditions of that experiment, the mysterious "Form V" had a higher electrophoretic mobility than Form I (*i.e.*, native DNA), but lower than that of single-stranded DNA from the same species.

Since the day the Weissmann study was first published, I have been asserting, with considerable certainty, that in all probability, "Form V" was nothing more than Form IV, arising because of the reannealing of complementary single strands under conditions at which the native structure cannot possibly have formed. To understand this, it is necessary to bring clearly to mind everything you've read here about the *reannealing* of denatured circular DNA, *i.e.*, "Form IV", and about how *particular* Form IV is with respect to precise control of pH, temperature and ionic strength, if any Form I native DNA is to result from its reannealing.

SLIDE 155

We must return, for a moment, to the data painstakingly obtained by Strider and Warner, who carefully determined a number of sets of conditions under which denatured Form I DNA, *i.e.*, Form IV, could be successfully reannealed. By the way, if this data looks a bit different than it did earlier in this presentation, that's because this slide was created at a time when I was working with data from Bill Strider's 1971 PhD thesis, whereas the earlier slides in this presentation were created using data from Strider *et al's* final J Biol Chem publication on the subject, about 10 years later.

Now let us compare these reannealing parameters with those employed in Weissmann's 1979 paper: he reannealed single strands of circular DNA at 60º, pH 8.5. Could this possibly have worked?

SLIDE 156

Please look again at the isolated Strider-Warner curve for 60º, and note that at that temperature, the pH of optimal reannealing was about 11.

How much Form I would you expect at Weissmann's reannealing pH of 8.5? According to this extrapolation of the Strider-Warner 60º curve, even at the relatively high pH of 10.4, the percent reannealing would have dropped to zero! At Weissmann's pH 8.5, two entire pH units *below* the zero-reannealing point, no normal DNA can possibly have appeared.

SLIDE 157

Conversely, however, the Strider-Warner data show that *Form IV* is stable over the entire pH range. That is, excepting for the few data points which lie precisely on these 5 curves, every other data point in the entire two-dimensional space of this graph defines a set of conditions of pH and temperature under which Form IV DNA *can* exist, *will* be stable, and will show *no tendency* to renature back to Form I.

It follows that Weissmann created Form IV, not a novel new DNA structure, and that the so-called "Form V" is merely Form IV, created by performing a reannealing under severely sub-optimal conditions of pH, temperature and ionic strength. But why didn't *he* know that?

SLIDE 158

The answer is probably that he was expecting the 2-3 fold increase in density known to exist for Form IV in velocity sedimentation experiments, such as we have been looking at all along.

Recall that ***s*** for circular native DNA increases 200-300% upon alkali denaturation, remaining nearly 200% increased even after complete neutralization. The data shown here is from a velocity ultracentrifugation experiment. Shouldn't we expect to see a comparable increase in the *electrophoretic mobility* of these same species, when examined by agarose gel electrophoresis, the method employed by Weissmann for his analyses? The intuitive fast answer would be "Yes", but that's *not the correct answer;* the correct answer is "No". That correct answer, however, could never have been predicted by logical deduction.

SLIDE 159

The answer came rather from laboratory observations by Pouwels and associates in the Netherlands, who examined the ***s*** values of Form IV as a function of ionic strength. Like much of the data in our current presentation, this was discussed extensively in "The Double Non-Helix Part I", on this web site, but we need to re-examine it now.

This study, one of the most important ever done on Form IV, shows that the Form IV sedimentation coefficient lies on a straight line, when plotted against the negative logarithm of the salt concentration. Thus, we see that at most common laboratory salt concentrations, ranging from 1.0 M down to about 0.01 M, Form IV does indeed demonstrate an elevated ***s*** value, although we also note that that value drops continuously, as the salt concentration drops. Incredibly, as the salt concentration approaches the salt concentration indicated by the double-headed arrow, which is the salt concentration of the typical agarose gel, the ***s*** value of Form IV has dropped so low, that it is only slightly higher than that of the native chromosome! And if we drop the salt concentration even lower, to 0.001 M, the ***s*** value of Form IV actually drops *below* that of Form I! What logical process of mind would have predicted that?

This was an important discovery; a discovery which one either knew about, or did not. I knew about it, Weissmann presumably did not. Therefore, seeing that his so-called "Form V" had a sedimentation coefficient only slightly greater than that of native Form I, and not the 200-300% increase in ***s*** which would have been expected for Form IV in velocity sedimentation experiments, the possibility that his DNA, reannealed at 60º, pH 8.5, was merely Form IV, most likely never occurred to him.

But it has occurred to me, and I believe that we can conclusively reject the 1979 Weissmann study as being a largely-irrelevant collection of experimental errors and mis-interpretations. But, all that notwithstanding, he did present a set of carefully-derived experimental data, which must be considered and explained. And if we reject *his* explanations, then it is incumbent upon us to come up with better ones. This we shall do presently.

SLIDE 160

To put this back into perspective, let us return now to Weissmann's original agarose gel. He proclaimed that that which is labelled here "Form V" was a novel and previously unknown DNA form. This conclusion was based upon the electrophoretic mobility, which was greater than that of native DNA, but only *slightly* greater. Since no previously-known form of DNA had a sedimentation coefficient which was greater than that of the native chromosome, except for Form IV, whose sedimentation coefficient, however, was not merely *slightly* greater than that of the native chromosome, but *massively* greater, it therefore seemed reasonable to declare this to be a new DNA form.

But now we have the benefit of the ionic strength studies of the Pouwels group in the Netherlands. These reveal that Form IV, when examined *not* by velocity ultracentrifugation, where the salt concentration is typically very high, but rather by gel electrophoresis, where the salt concentration is very low, would in fact be expected to behave exactly as we see here, displaying an electrophoretic mobility greater than the native chromosome, but only slightly so.

Therefore, in the absence of data to the contrary, common sense mandates that we regard this allegedly "new form of DNA" as being merely Form IV, displaying the behavior expected for Form IV at low ionic strength, and not a new form of DNA at all.

SLIDE 161

Well, this completes our discussion of the 5 phenomena of DNA science which must be thoroughly understood before we can proceed to directly address the question of Form IV structure. We've done our homework; let us now begin to address the main question.

SLIDE 162

"Elucidation of the structure of Form IV.
Beginning of the denaturation curve."

We now have the background information necessary to propose a structure, or structures for Form IV.

SLIDE 163

What we are going to do is to take a journey through the entire denaturation-neutralization curve for ϕx174. We'll start with the lower curve, composed of the white circles, starting at pH 7, where the sedimentation coefficient is about 21 ***s***, and follow it up to the point of irreversible denaturation at pH 13:

Then we'll follow the neutralization curve back to pH 7, at which point the sedimentation coefficient is still quite elevated, at about 36 ***s***.

SLIDE 164

That's not the end, however. In order to fully-understand Form IV we must then follow the Pouwels {sedimentation coefficient vs [salt]} titration curve, down to the end, which is the realm of salt concentrations typical of those used in electron microscopy:

This is a long journey, with a lot of stops to make. Let's get started.

SLIDE 165

From pH 7 to a pH of about 11.6, the sedimentation coefficient of ϕx174 is approximately constant, around 21 ***s***. Then, from pH 11.6-11.8, ***s*** drops to about 16. This drop, and the subsequent rise up to pH 13, are extensively discussed in "The Double Non-Helix, Part I", on this web site, and I will only present a general summary here.

The little pictures above the curve tell the story. It is universally understood, by all commentators, that the drop in ***s*** reflects an unwinding of supertwists, which gives rise to an open-circular form. The open-circular form presents a larger surface area than the more-compact supertwisted form, with a corresponding increase in viscosity-induced drag. This explains the slowing down of the chromosome during this phase of the ***s*** vs. pH titration.

SLIDE 166

Why exactly does increasing the pH cause an unwinding of supertwists? This question is not discussed much, in fact I'm not sure I've ever heard it discussed at all. It cannot likely be due to a change in the sugar-phosphate backbone, because the pK_a of DNA phosphate is variously stated as being anywhere from 0-2. In that range of pH, the phosphate groups are protonated.

But by the time the pH has increased to the 11-12 range, the phosphate groups have long since become deprotonated. Therefore, to explain the unwinding of supertwists, we must look at the bases, all of which have pK_a's in the pH 11-12 range.

SLIDE 167

By the way, the data shown here may look unfamiliar because of the inclusion of the low pH portion of the pH scale. These data are from an early paper by Vinograd and Lebowitz, showing the effect of pH on the sedimentation coefficient of the chromosome of the oncogenic virus polyoma. This study, to the best of my knowledge, is absolutely unique in that it looked at both the basic *and* acidic pH domains, the acidic being a realm of pH that has been historically neglected.

You'll note that there's a dip in ***s*** at low pH which in most respects mimics that seen at high pH, except that there's no plateau at the end. All writers on the behavior of circular DNA at low pH state that at low pH, DNA strands are cleaved, destroying all the distinctive topological properties of the native form. That is, we're to believe that the left-hand end of this curve is also the end of the line for this species of inquiry, because fully-intact duplex circular DNA supposedly cannot exist below it.

When one encounters any of the rather rare mentions of the behavior of circular DNA at low pH, one gets the distinct impression that every writer on the subject merely quoted from previous writers, all eventually leading back to the single study whose data you see here.

The implication is that no DNA topological studies are possible at pHs lower than those shown, because of strand cleavage. I've always wondered whether that was entirely true, especially if the process of strand cleavage was slow. But if, like myself, you're curious to know what *really* happens to circular DNA at low pH, you're going to have to repeat this work yourself, because this is the only low-pH study I've ever seen, and it's entirely possible that it's the only one ever done.

But I digress. The point of this slide is simply to dramatize that all important pH changes in the phosphate groups take place at *low* pH, and that therefore the unwinding of supertwists at *high* pH cannot be due to pH-related changes in the DNA sugar-phosphate backbone.

SLIDE 168

The reason for the unwinding must be rooted in the fact that each of the "standard" base pairs depends upon an amine or imine group capable of donating a proton, and that these groups become progressively deprotonated as the pH rises. This explains the unwinding as an acid-base phenomenon, but really visualizing what goes on at the molecular level is very difficult. The beginning of understanding is to consider that, although much has been written lately about the relatively-greater importance of base-stacking than base-pairing to DNA stability, the fact remains that base-pairing is essential to maintenance of the DNA duplex. As the pH increases, base-pairing is progressively weakened, therefore, in principle at least, secondary turns must unwind.

SLIDE 169

But how can we envision this unwinding? If we could shrink ourselves down to the size of a DNA molecule, and watch it denature, what exactly would we see?

The first question that comes to mind is this: Do the hydrogen bonds in the DNA base pairs literally stretch? At the outset of my Form IV inquiry, I actually didn't know the answer to that question. So I had to look it up.

SLIDE 170

I found the answer in a Proteopedia article, which states that hydrogen bonds can have lengths anywhere from 2.2 to 4 Å in length. At 2.2 Å, the strength begins to approach that of a covalent bond. Conversely, at 4 Å, it's hardly a bond at all.

SLIDE 171

It is therefore easy to imagine a DNA helix gradually and continually unwinding, as dramatized by this rather crude rope model, with a concomitant widening of the helix,

that is, an increase in the hydrogen bond length, accompanied by an increased spacing between the two members of each base pair.

Since left-handed DNA is clearly favored under conditions promoting denaturation, that might suggest that the H-bond lengths in Z-DNA would be greater than those of B-DNA. Are they?

SLIDE 172

To answer this question, I took some measurements, employing, for B-DNA, the 1981 crystal structure 1-BNA, whose Protein Data Bank accession # is the same, *i.e.*, 1-BNA. The Z-DNA virtual model, which I have been using for many years, was from the U. of Liverpool.

These figures may seem to show some potentially-interesting differences in H-bond lengths, but they're unlikely to be of any real significance with respect to the question at hand, because they're all in the 2.8-2.9Å range, and a 1/10 Å change in H-bond length is not going to be very helpful in explaining the unwinding of a DNA helix.

We therefore must look for changes in base stacking.

SLIDE 173

Here too, however, there is not much support for our rope model of continuous unwinding. DNA cannot, to the best of my knowledge, have a continuously-varying base-stacking distance. Speaking anthropomorphically, DNA bases "like" to stack at about 3.4 Å, and, if what we see in various published DNA structures is any indication, there's not much leeway in the stacking distance.

We examined Z-DNA in detail in the current presentation, slides 49-59, specifically so that we could address questions such as this. What we found was that, even though Alexander Rich's left-handed DNA structure has a much larger pitch than Watson-Crick DNA, the primary stacking distance is still 3.4 Å.

We also saw that Z-DNA has a secondary 4.0 Å stacking distance, but I, for one, have not heard of any intermediate stacking distances. What about shorter stacking distances? It's true that the bases of the Gehring tetramer appear to be stacked at 3.08 Å, but, as we saw above, in slides 93-108, that apparently short distance is misleading in that they are extremely laterally-displaced. Moreover, this spacing only occurs in the setting of low pH and unrealistically short and powerful hydrogen bonds, wherefore Gehring cannot be of much use as a model for ordinary DNA denaturation.

SLIDE 174

I am therefore drawn to the conclusion, pending information to the contrary, that the unwinding of DNA, in an alkali titration experiment, is not smooth and continuous, as in our rope model, but, rather, *incremental.* According to this view, each incremental step in unwinding would consist of a single right-handed W-C secondary turn unwinding, then re-winding into a left-handed Z-DNA turn, accompanied, as it must, by the removal of a single right-handed tertiary superhelical twist.

Again, if you don't know what I mean when I say that introduction of a left-handed secondary twist must be accompanied by the removal of a right-handed tertiary twist, then you must go back to "The Double Non-Helix, Pt. I", and review the section on the topology equation, L_k=T+W, which is explained in slides 263-269.

SLIDE 175

With each supertwist removed, ***s*** decreases, as the conformation moves closer and closer to an open circle, as depicted by the small and somewhat apocryphal drawings above the curve.

At a pH of about 11.8, all 25 right-handed supertwists have been removed, the circle is fully-open, and the sedimentation coefficient reaches its minimum. At this point, assuming, as we do throughout this web site, that DNA, in reality, has no net helical twists, then at pH 11.8 there must be an approximately equal number of right-handed secondary twists with base-pair spacing of 3.4 Å, and left-handed secondary twists with average base spacing of 3.7 Å, and *no* superhelical twists at all.

SLIDE 176

If the pH continues to increase, the only option for the chromosome is to convert *more* right-handed secondary "B" turns, to left-handed secondary "Z" turns. Each such conversion was previously accompanied by the removal of a right-handed tertiary supertwist, but there are no more to remove. Therefore, left-handed supertwists must be wound in, as suggested by the apocryphal 3-turn superhelix picture above the curve. This causes ***s*** to begin to increase again.

At a pH of about 12.3, the shape of the curve suggests that ~25 left-handed supertwists have been wound in. As I demonstrated earlier, in slides 126-144 above, 25 supertwists is the upper limit of superhelicity that is possible for this species of DNA; above this point in the pH scale, no further supertwists can be added.

There is a shoulder in the ϕx174 curve at this point, labelled "χ". The same shoulder is found at comparable points in all the other published pH vs. ***s*** curves that we looked at earlier. This shoulder suggests a conformational change. What is this change?

SLIDE 177

"What is happening at the 'χ' shoulder at pH 12.3?"

SLIDE 178

We have seen, in four completely different physical settings, that circular DNA cannot supertwist beyond a certain limit, which, for a 5 kb chromosome such as ϕx174, the organism whose DNA is represented here, is 25 superhelical turns. If no further superhelical turns can be added at increasing pH, what, we must ask, is happening at the "χ" shoulder? Is there anything at all we can say about this?

SLIDE 179

We are not utterly without clues, and in fact we do have some good ones. First of all, I think we can be confident that this is not merely a continuation of left-handed superhelical twisting of any structure which continues to be based upon the W-C right-handed helix, and/or the Rich left-handed helix.

After all, we have just seen that in 4 totally different biochemical settings, a 5kb chromosome *cannot* superhelically wind itself beyond 25 tertiary turns. So the first thing we know is what this is *not:* It is *not* merely the addition of more superhelical twists.

SLIDE 180

Secondly, we know, from the work of Rush and Warner, who were the original source of the curves shown here, that whatever the change of state is, at pH 12.3, it is *reversible*, at least up to, or nearly up to the top of the entire denaturation curve. There *is* a point in the denaturation curve at which the denaturation becomes *irreversible,* and that point was shown, by Rush and Warner, to be considerably above the pH 12.3 shoulder, namely at about pH 12.7-12.8.

SLIDE 181

We may now ask this question: At points between the pH 12.3 shoulder, and the pH of irreversible denaturation at approximately pH 12.7-12.8; a range within which we know that the DNA, although partially denatured, can still be *re*-natured; what exactly is the *mechanism* of renaturation?

While the question cannot be answered with absolute certainty, there is one thing that can be said with *considerable* certainty: If the DNA can be renatured, there must be substantial residual base-pairing, sufficient to form a nidus for the resumption of the W-C structure, or the left-handed Z structure, when the pH is lowered. Of the need for residual base-pairing we can be *reasonably* confident, because once denaturation is complete, at pH 13, where *all* the base-pairs are definitely disrupted, *re*naturation becomes a very difficult task.

SLIDE 182

This means that to deduce the structure of DNA at the pH 12.3 shoulder, we must look first for structures that are base-paired, but which are not subject to the sorts of steric clashes which would result from excessive superhelical windings.

If we assume furthermore that the denaturation curve, between the pH 12.3 shoulder, and the point of irreversible denaturation as one approaches pH 13, represents the evolution of a single structure, then that structure, at the higher pH, must have the extraordinary property of being 2-3 times as compact as native DNA, since the sedimentation coefficient of the native chromosome is only 21 ***s***, whereas at pH 13 the sedimentation coefficient of the new structure has leaped up to nearly 50 ***s***; a two-to-threefold increase!

In the entire history of DNA research, no one has ever described a *duplex* structure which could take on such extreme compactness, regardless of the degree or direction of secondary or tertiary twisting. I know of only one *type* of structure which can be invoked here: the DNA tetraplex, which consists of two stretched-out duplexes, with their base-pairs mutually intercalated.

Among the known types of DNA structure which have been discovered to date, this is the only type that can account for such a massive increase in density.

SLIDE 183

Two such structures were mentioned earlier. First, on the left, we have the tetraplex structure proposed by Tai Te Wu, which was briefly described here in slides 60-92, and described in much greater detail in the PowerPoint slide presentation entitled "The Double *Non*-Helix Part II: The Probable Structure of the Protamine-DNA Complex", on this web site. An axial view of two adjacent base pairs from this structure is shown on the left. This structure, in the form in which I published it in 2006, has what aromatic ring chemists refer to as "sandwich" stacking of bases, although it is not necessary that the stacking be so perfect. If the base pairs were offset a few angstroms, to give a parallel-displaced spacing, the general features of the structure would be unaffected.

Secondly, on the right, we have the Gehring tetramer, also described above, as well as in "The Double *Non*-Helix Part I", on this web site. An axial view of two adjacent base pairs of this structure is shown on the right, with 90º "parallel-displaced" stacking of bases.

Because the Gehring structure only forms at low pH, and because the orientation of the strands is unnatural, I'm going to assume that the Wu structure, on the left, is the correct one.

SLIDE 184

In order to form a clear picture in our minds, of what happens at the pH 12.3 shoulder, we're going to have to decide upon a tentative working model for the non-helical DNA duplex, for demonstration purposes. I'm referring only to a model for DNA extracted from cells, stripped of all nucleoproteins, and floating freely in solution; that is, a structure that might not have much relevance to the nucleic acid chemistry of real intracellular life, but only to the sorts of artificial laboratory settings that pertain to data such as those we are considering here.

There are several such non-helical structures that have been proposed, largely similar to each other, and, insofar as the differences between them relate to the matter immediately at hand, it's not so important, at this juncture, which one we choose. Therefore, I am arbitrarily going to use the 1976 Gordon Rodley side-by-side structure, the first and most famous of all the non-helical structures, for demonstration purposes. This is the structure we looked at previously, in Slides 147-148 above. It's 50% right-handed and 50% left-handed, consisting of a perfectly regular alternating pattern of half twists, *i.e.*, 5 base-pairs, or ½ twist to the right, then ½ twist to the left. There are no complete twists, and therefore neither strand ever crosses the other:

Now, in accordance with my determination to leave no stone unturned, I must point out that we've been talking all along about *excesses* of right-handed and/or left-handed twists at various pH's, whereas the 30 bp segment of Rodley-structured DNA shown in this slide has exactly and precisely *equal* numbers of each, so one might protest that the slide does *not* provide an accurate working model of the DNA at the pH 12.3 shoulder, where we've invoked a *preponderance* of left-handed windings. But I'm going to assert that this is, in fact, accurate enough for our current purposes. Remember that the taking on of the maximum-possible number of *tertiary* superhelical twists only changes the *secondary* winding of circular DNA by about 5%, *i.e.*, for a 5kb chromosome such as that of ϕx174, a change of only 25 secondary helical twists out of a total of 500. In a segment the size of the one shown here, having a length of only 30 bp's, the 5% inequality between right-handed and left-handed twist count would hardly be noticed. Therefore this illustration, which is exactly 50% right-handed and 50% left-handed, is sufficiently accurate for our current purposes.

SLIDE 185

Here we see an artist's representation of a 5 kb circular chromosome at the χ shoulder at pH 12.3, maximally-supertwisted in the left-handed sense. Unfortunately, the so-called "artist" is me, and I ran out of energy at 12 supertwists, so instead of a 5000 kb chromosome with 25 left-handed supertwists, I humbly beseech you to accept this hypothetical substitute chromosome, which — if it was possible for me to draw them...*which it isn't!*..would have only 2500 secondary helical twists, but whose 12 left-handed tertiary supertwists *can* be, and *are* clearly shown.

What I envision starting to happen, at the χ shoulder at pH 12.3, is that at the 12-supertwist limit for this particular chromosome, the phosphate groups of the sugar phosphate backbones are already squeezed as close together as is possible. The limit of closeness is undoubtedly established, generally, by the charge repulsions between the negatively-charged phosphate groups, and more specifically by the van der Waals repulsions of the phosphate oxygen atoms. I'm going to take the sum of the van der Waals radii of two adjacent oxygen atoms, which is about 3 Å, as a fairly-reasonable estimate for the minimum comfortable distance between apposing strands of the superhelix. Any attempt to further supertwist will violate the van der Waals radii of the oxygen atoms, and exponentially increase the phosphate group charge repulsions, leading the DNA to seek an alternative conformation.

SLIDE 186

If, in this hypothetical chromosome with 12 left-handed supertwists, a 13^{th} supertwist tries to form, the phosphate group charge repulsions, and phosphate oxygen atom van der Waals repulsions, will pass the limits of tolerance, and, if we may speak anthropomorphically, the chromosome will "seek an alternative" to a 13^{th} supertwist.

By default, since I, for one, can imagine no other option, that alternative will be the Wu structure. At some point in the chromosome, which we may arbitrarily designate as the region now highlighted in the center, the existence of conditions which might, in other circumstances, bring about additional supertwisting, will, here at the supertwist

limit, bring about instead the merger of the two duplex strands, which constitute the highlighted supertwist, into a single mutually-intercalated structure, like so...

SLIDE 187

The picture on the left shows a small section of only *one* of the two duplexes which constituted the highlighted supertwist in the previous slide. We'll see the second strand in a moment. The process depicted here can be thought of as the molecular biological equivalent of a little genie grabbing the ends of the DNA duplex on the left, and stretching it out to its full length, as shown on the right, which is identical to the structure on the left, except that it has twice the base-pair spacing.

The other duplex strand of the superhelical structure will then mutually intercalate with this strand, like so...

This results in a chromosome with a hybrid structure...

SLIDE 188

...most of it still in the duplex left-handed superhelical conformation, but a small segment in the middle now being in the Wu *tetraplex* 4-stranded mutually-intercalated conformation.

The final result is that at the pH 12.3 shoulder, instead of an impossibly-overwound left-handed superhelix, we instead begin to see the thermodynamically sound, orderly hybrid structure shown here.

Note that the small tetraplex segment is being portrayed as a perfectly straight, non-twisted structure, but this by no means must *necessarily* be the case. Like the Gehring Tetramer, the tetraplex structure I'm proposing here, at pH 12.3, can surely have its own helical twist, and I would presume, in fact, that it does. That is, any structure of this sort could, in principle, have either a right-handed twist, a left-handed twist or no twist. Pending information to the contrary, I would naively guess that this proposed new tetraplex segment would have the same left-handed twist that the parental superhelical regions have. But, whereas further supertwisting of the parental structure would give rise to steric conflicts, the appearance of *tetraplex* structure brings no such destabilization, as is evident if we flip the structure on end and look at it from the axial perspective.

At this angle we can see, as we have seen previously, that the phosphate groups are now neatly arrayed on the *outside* of the tetraplex column of DNA, where there are no charge repulsions or Van der Waals violations:

SLIDE 189

Let us now assume that the pH 12.3 mutual intercalation process begins precisely at points 180º opposite one another in the circular chromosome, as depicted here, and proceeds like the closing of a zipper, both upward and downward, ultimately giving rise to a single, perfectly intercalated structure at a pH just below 13. There's good reason to think that it may not be quite that simple, but let's assume, for the moment, that it *is* that simple.

SLIDE 190

Here's the DNA at the beginning and the end of this zippering process. The drawing is designed to emphasize the differences in size and structure at the two extremes of this pH range. At pH 12.3 the DNA is a supertwisted duplex, with a probable diameter of about 43 Å, which I'll explain momentarily. At pH 12.8, just below the point of irreversible denaturation, the structure is a closely-packed tetraplex, with a much smaller diameter, but at least twice the length. The reason the length may *more than double* is that (a) the DNA had to stretch to twice the normal length to allow for intercalation, and (b) I'm allowing for the possibility that some of the supertwists may unwind as the apposing sides of the circular chromosome mutually intercalate, which could add a bit more to the length.

These conformational changes, as we shall see, are entirely sufficient to explain the huge increase seen in the sedimentation coefficient, ***s***, in the pH vs. ***s*** titrations whose data we need to explain.

Let's look at the diameters of these structures.

SLIDE 191

For the supertwisted form, for the reasons given earlier, we're going to assume that the spacing between the duplex strands is in the neighborhood of 3 Å. The diameter of each of the two duplex strands will be presumed to be the standard W-C 20 Å cross-duplex phosphate-to-phosphate distance. Therefore, the total superhelix diameter will be about 43 Å.

SLIDE 192

For the tetraplex, we saw, in the presentation called "The Probable Structure of the Protamine-DNA Complex", on this web site, that the cross-sectional dimensions of the Wu intercalated structure may be represented as a rectangle with sides 10 x 20 Å. We can now begin to do some calculations. The question is: Do these structures, as presented here, provide us with a satisfactory explanation for the 200-300% greater sedimentation coefficient of the Wu tetraplex at pH 12.8, compared to the sedimentation coefficient of the 25-twist superhelix at pH 12.3?

SLIDE 193

Let us compare the cross-sectional areas. The area of a cylindrical cross section is, of course, πr^2. For the superhelix, we just saw that the diameter is 43 Å; therefore the radius is half that, which, when squared and multiplied by π, comes out to 1451 square angstroms.

For the Wu tetraplex, the cross sectional area is merely the product of the length and width of the rectangular cross-section, which is 200 square angstroms.

The ratio of these cross-sectional areas is 7.25, or 725%. That is, the sheer mass of solvent interface through which the DNA must fight, in order to move forward in the centrifuge tube, is clearly much larger for the superhelix on the left, even

notwithstanding it's smaller length. That means it's going to move more slowly, as illustrated by this sky-diving analogy, which we invoked previously, for similar reasons.

SLIDE 194

The structures of the DNA forms represented in this slide are considered, by the entire molecular biological establishment, to be "unknown", and the structures shown are therefore hypothetical. We must therefore ask: "Would DNA forms, having the structures shown here, also have the sedimentation coefficients shown here, that is, would they be predicted to have the sedimentation coefficients which have actually been measured for DNA at the pHs at which these structures appear?

This question may be re-phrased, to ask the fundamental underlying question: Can the structures shown here explain the 200-300% *increase* in ***s***, which is observed at high pH for all circular chromosomes that have been studied?

In order to attempt to do a more precise estimate of the relative rates of sedimentation of our 2 structures during ultracentrifugation, the only mathematical formula I know is that of Theodor Svedberg, the chemist who won the 1926 Nobel Prize for his invention of the ultracentrifuge, and whose name, Svedberg, is now the name for the unit of terminal velocity during ultracentrifugation, which we have been representing all along by his initial, the letter ***s***. The use of this mathematical formula, however, is problematical, because the Svedberg equation only strictly applies to ideal spheres, which is a far cry from what we're dealing with here. Nevertheless, let's do the math, and see what we can come up with.

SLIDE 195

"Does the Wu tetraplex explain the 200-300% increase in ***s*** at high pH?"

SLIDE 196

The Svedberg equation, for calculating the value of the sedimentation coefficient ***s***, states that ***s*** equals ***m*** divided by "6-π-η-r". The numerator ***m*** is the mass, which is obviously the same for both forms of DNA, since we're merely looking at conformational changes of what is otherwise the same molecule.

Eta, the letter that looks like an elongated letter "n", is the viscosity of the medium, which can be regarded as a constant for these experiments. Therefore, the only variable on the right side of the equation is ***r***, the radius of the particle. I reiterate that this formula was originally written for ideally-spherical particles, and cannot be literally applied to complex shapes such as those of our two DNA forms. But perhaps, if we make the effort, we can accommodate the formula to our DNA forms.

SLIDE 197

Here we see an imaginary metal rod, representing a length of DNA, sedimenting down an ultracentrifuge tube, in the direction of the large arrows. The purple dots

represent the bulk solvent, which is essentially stationary; that is, the motion being portrayed here is the motion of the *DNA* relative to a stationary solvent.

In the position shown, the DNA is presenting the maximum possible surface area to the solvent, creating the maximum possible viscous drag, and therefore moving at the slowest possible sedimentation rate. But if we rotate the DNA 90º, as if it was a spear being hurled down the centrifuge tube, then it would present the *minimum* surface area to the solvent, as we can readily see if we move ourselves to the bottom of the tube and look up at the descending DNA.

SLIDE 198

Now let's return to our original view, only this time, let us *replace* the rod with a set of perfect spheres, attached together to form a rod-like structure comparable to the original in size and overall shape. It is intuitively-obvious that, to a first approximation at least, each spherical unit of the new rod would sediment at approximately the same rate as it would if it was sedimenting all by itself.

This seems to imply that the sedimentation coefficient of a cylinder can indeed be at least *estimated* by the Svedberg equation, treating the radius of the cylinder as if it was the radius of each of a set of identical spheres of the same diameter, fused together into a rod-like structure about the same length as the original cylinder.

In suggesting this, I'm ignoring the question of how the act of *linking together* all the spheres might affect the sedimentation coefficient of the resulting rod-like structure as a whole, because I have no exact mathematical formula for answering the question *quantitatively*.

SLIDE 199

Considering the question *qualitatively*, however, is possible. If we mentally reduce ourselves to the size of a sphere, to observe what goes on as our rod-of-spheres passes through the solvent, we can see that at the junction of each pair of adjacent spheres, there will be a small contiguous zone which a water molecule cannot pass through, necessitating that it "leap" over the junction. This will increase the viscous drag relative to that which would pertain if the spheres were truly independent of each other, but probably not by much, because the proportion of the solvent-accessible free surface area which is lost at each sphere-to-sphere junction is small, each such junction being small in surface area, sort of like a tangent to a circle.

This by no means conclusively addresses the question of how altering the length of a rod of fixed mass will affect its sedimentation coefficient, and we'll be returning to this question in a moment, but first let's look at the rod in the orientation we previously referred to as "spear-throwing position", and replace *that* with perfect spheres.

SLIDE 200

Here, we can speak with more certainty of the sedimentation coefficient, because, if we mentally go back down to the bottom of the centrifuge tube, looking up at the rod-of-spheres as it descends, we see that the solvent only interacts with a surface which is

the size of a single one of the spheres, so that, as long as the DNA is constrained to assume this orientation, it really doesn't matter much what the length of the rod is.

SLIDE 201

In this instance we can apply the Svedberg Equation literally, for the comparison of the sedimentation coefficients of our two DNA forms, because, if and when both forms of DNA are in this "spear-throwing" orientation, then the only variable in the equation is ***r***, the radius of each of the DNA forms. And the ratio between the radii is, of course, the same as the ratio of the diameters, which we know. Let us therefore look at the diameters of our two DNA forms, and thereby determine the ratios of the sedimentation coefficients, at least for this "spear-throwing" orientation.

SLIDE 202

We previously saw that the superhelix had a cross-sectional diameter of about 43 Å. The tetraplex cross-section, however, was approximately a rectangle, one side having a length of 20 Å and the other 10 Å. Well, a rectangle isn't a cylinder, so we're going to have to improvise. As a first approximation, I'm going to represent the tetraplex as being fairly equivalent to a cylinder with a diameter of 15 Å, the average of the two sides of the rectangle. Then, the ratio of the diameters of the two conformations of this DNA, which is, of course, the same as the ratio of their radii, is...

SLIDE 203

... 43 Å divided by 15 Å, or just a shade below 3.

Since the structure with the *smaller* diameter will have the *higher* sedimentation coefficient, this predicts that the Wu tetraplex, fully-formed at pH 12.8, will sediment about 3x *faster* than the 25-twist superhelix which is seen at pH 12.3. And that's not far from what is actually seen for ϕx174, where the ratio is 50 ***s*** divided by 21 ***s***, or 2.4.

SLIDE 204

We've now quantitatively grappled with the question of the predicted ratio of the sedimentation coefficients for our two DNA forms, but only for one orientation, namely the "spear-throwing" orientation which presents the minimum possible surface area to the solvent. What can we predict about the ratio of the sedimentation coefficients when the DNA is in the *other* orientation we made mention of, namely the orientation which presents the *maximum* possible surface area of DNA to the solvent, which occurs when the DNA rods are oriented at right angles to the direction of sedimentation?

It is important that we try to estimate this ratio, because whatever this ratio proves to be, we may be confident that all other possible orientations of the DNA rods will result in sedimentation coefficient ratios which will be *between* those seen at these two extremes.

SLIDE 205

Although the sphere analogy we've been looking at suggests, thus far at least, that in a situation where the mass of a DNA rod is constant, the length does not have a large effect on the sedimentation coefficient, that still does not mean that it will have *no effect* at all on ***s***. But can we hope to get a little more ***quantitative*** in our answer to this question?

There are, in other settings, ways of predicting the effect of DNA length on mobility, as in the instance shown here, which gives *electrophoretic mobility* values for differently-sized strands of *linear* DNA. But in these cases, the *molecular weight* of the DNA increases with the length. In our example, however, the molecular weights of our two DNA forms are exactly the same, and the length difference is due merely to a conformational change.

This much can be said with certainty: the greater length of the Wu tetraplex *will* slow it down relative to the sedimentation rate of a shorter rod-shaped molecule. Exactly *how much* slower it will sediment, however, we cannot say with mathematical precision. Nevertheless, in a comparison with the shorter superhelix, I'm strongly inclined to believe that the amount of slowing down of the tetraplex, as a result of its greater length, is not very great.

SLIDE 206

Why do I say this? To answer this question, let us return to the centrifuge tube, only this time dispensing with the metal rods and spheres, and looking instead at the real DNA. I've added another detail here, namely an attempt to graphically illustrate the disordering of the ordered structure of the solvent, in a zone surrounding each DNA form.

The zone of turbulence surrounding the superhelix, which has the much larger diameter, is represented here as being correspondingly wider than that which surrounds the tetraplex.

SLIDE 207

I'm going to attempt to execute a simple Einsteinian thought experiment, by cutting the tetraplex in half. It is intuitively obvious that the half-sized tetraplex, now approximately the same length as the superhelix, will sediment faster than the superhelix, in accordance with its vastly-smaller cross-sectional area and correspondingly-smaller diameter. But there's a "catch": by cutting it in half, we've halved the molecular weight, so that we have now introduced a new variable to the Svedberg equation. Whereas previously, we only had to deal with changes in ***r***, the radii of the cylinders, the "bad news" here is that we now must also consider ***m***, the mass. The "good news", however, is that by cutting the tetraplex in half we have rendered the lengths of the two DNA forms approximately equal, so that the effect of length on ***s***, that troublesome relationship for which we had no mathematical formula, can now be essentially ignored! That is indeed good news.

SLIDE 208

Moreover, if we arbitrarily assign the letter ***m*** to represent the mass of the *half*-tetraplex, then we can see immediately that the mass of the superhelix, having *not* been cut in half, will simply be ***2m***. After doing a trivial rearrangement of the Svedberg equation, dispensing with the constants, and taking the liberty of substituting the diameters of the DNA forms for their radii, we arrive at the following annoying-looking, but actually quite trivial algebraic expression, which simply reveals that by cutting the tetraplex in half, we have also halved the ratio of the sedimentation coefficients of the two forms.

Whereas with the full tetraplex and superhelix we were dealing with the ratio of the diameters of the two forms, now we are dealing with one-half of that ratio. Thus, the ratio of the sedimentation coefficients of the *half*-tetraplex to the superhelix is simply one-half the inverse ratio of the diameters, or 1.45.

SLIDE 209

To complete this tedious exercise, we now bring back the top half of the tetraplex, but as a free-floating additional molecule, sedimenting independently of the bottom half. It is intuitively obvious that in this setting, the top and bottom will sediment at exactly the same speed, because they are exactly the same thing.

Suppose we now insert a bridge between the top and the bottom, causing them to once again be a single structure. How could this possibly cause any important change in the sedimentation coefficient? The reconstituted tetraplex would, after all, experience nearly the same total viscous drag as the two halves did, when they were sedimenting separately.

SLIDE 210

I have therefore concluded that the sedimentation coefficient ratio *in this orientation*, where the DNA presents the maximum possible surface area to the solvent, is very likely to be approximately the number we just derived, namely 1.45, even though that number was derived for the tetraplex after it was cut in half. What I'm saying is that, in this position, the length of the tetraplex is unimportant enough to ignore, for our current purposes.

SLIDE 211

In our quest to determine whether or not the appearance of the Wu tetraplex can account for the 200-300% increase in *s* at high pH, we have one task remaining. Thus far, all our calculations have been for DNA rods in extremes of orientation with respect to the degree of exposure of their surface areas to the viscous drag of the solvent. What about the infinite number of other orientations?

A better mathematician than I could probably come up with an exact formula for this, but I'm content to note that all other orientations will give rise to sedimentation coefficient ratios somewhere between the two extremes we have calculated. That is, if, in

the orientation of maximum solvent drag the ***s*** ratio is 1.45, and if, in the orientation of minimum solvent drag that ***s*** ratio is 2.9, then in the average middling position, the typical ***s*** ratio must be approximately the average of the ratios at the extremes, *i.e.*, about 2.2. And this compares very favorably with the observed value of 2.4.

Now, admittedly, this thought experiment falls far short of Einstein's flashlights on speeding trains, and in elevators flying through space. Oh-oh; it looks like Albert has been watching us the whole time. Oh, well, be that as it may, I still think that this illustrates, at least in a crude qualitative way, that the accelerating effect of the diminished *diameter* of the Wu tetraplex on its ***s*** value is very likely more important than the *decelerating* effect of its greater length.

SLIDE 212

In any event, and lack of precise mathematics notwithstanding, I think we can see that the conformational change proposed here, between pHs 12.3 and 12.8, is almost surely *entirely* adequate to explain the extraordinary increase in the sedimentation coefficient, from 21 ***s*** at pH 12.3, to 50 ***s*** at pH 12.8.

SLIDE 213

In case it is not perfectly obvious, we have not yet proposed a structure for Form IV, but only a structure for the Wu tetraplex, which is the direct *precursor* of Form IV. The Wu tetraplex can be readily renatured, but Form IV is extremely resistant to renaturation, and therefore must have a different structure.

The sedimentation coefficient calculations we have gone through have in no way been done in vain, because, as I shall show you shortly, the structure of Form IV will very likely have about the same density and compactness as the Wu precursor.

SLIDE 214

Before proposing a structure for Form IV, there is one last important detail we must discuss, concerning the reversible portion of the denaturation curve. This is a detail which was touched upon earlier, relating to what I referred to as the "unlikelihood" that the Wu tetraplex literally began with the intercalation of points precisely 180º opposite one another in the circular chromosome, and then proceeded in both directions simultaneously to give a perfect, zipper-like closure. I suggested that the situation might not be quite so perfect. Here's why.

SLIDE 215

In Robert Warner's final publication on the subject of Form IV, he made the interesting observation that the kinetics of renaturation of Form IV were *heterogeneous*. This was demonstrated in the figure shown here; a renaturation reaction in 1 M salt, at pH 12.1, 25º. At that elevated pH, renaturation proceeded slowly enough to study the kinetics.

At the start of the renaturation reaction, the kinetics were first order, that is, on the natural logarithmic scale he used, the data was pretty much a straight line up to about the 50% renaturation point, but by 60%, the line was flattening out, giving rise to a long tail within which renaturation was progressively slower over time. This suggests that preparations of Form IV are predominantly one species, but also include a mixture of other species even more mysterious than the predominant one. Warner and his associates made little attempt to account for this heterogeneity in terms of any particular atomic model, so we're going to have to at least *try* to do that here.

SLIDE 216

It is well to keep in mind that the mutual intercalation we are proposing, between pHs 12.3 to 12.8, has nothing to do with specific base-pairing, but is rather a *stacking* of already-paired bases. It is therefore not necessary to presume that the intercalation begins at any particular point in the chromosome, or proceeds in any particular way. Rather, it is entirely possible, and perhaps probable, that many of the regions of *un*-zipped duplex DNA, *between* the newly-formed tetraplex regions, may be mismatched in length, so that when they eventually coalesce into the tetraplex conformation, they give rise to a variety of looped structures, such as are illustrated here.

In this drawing I've taken the liberty of representing the duplex chromosome as a single line, which is merely an artistic expedient. Thus, drawing "A", although drawn with a single line, represents a typical circular *double*-stranded chromosome whose sides are about to be drawn together by supertwisting. For graphic clarity, I've also taken the liberty of omitting the supertwists, which are superfluous to the point I'm going to make here.

Drawing "B" shows the entire chromosome closed into a zipper-perfect tetraplex, except for the terminal loops. I must presume that the Form IV molecules which will eventually arise from structure "B" constitute the predominant species noted by Warner; the species that renatures with first-order kinetics. What about the tail in the Warner renaturation curve?

Drawing "C", although in a melodramatic and perhaps wildly-exaggerated way, depicts a possible explanation for Form IV heterogeneity. This drawing shows a chromosome at some pH in the vicinity of 12.3, which has converted to the tetraplex structure at three different sites, but with the unmerged regions between them being mismatched for length. After the intercalation process is completed, at pH 13, these mismatched loops may self-intercalate, giving rise to peculiar branched structures such as are depicted in drawing "D".

Because all publications on Form IV show it to sediment as a single, well-defined band, I suspect that the most realistic of these drawings is "B", and that the lengths of the branches in the branched structure shown in "D", if indeed such branches form at all, are likely much smaller than is suggested by the latter intentionally-melodramatic drawing.

SLIDE 217

Well, this completes our study of the steps leading up to Form IV. At some pH between 12.7 and 12.8, ϕx174 DNA has totally converted into the Wu tetraplex form, from which Form I native DNA can be recovered by merely neutralizing the solution. But once the pH hits 13, the denaturation becomes irreversible. That new structure, which appears at pH 13, is what was dubbed "Form IV" by early researchers. We are now well-equipped to suggest a likely structure for it.

SLIDE 218

Structure of Form IV

1. Forces available at pH 13 to stabilize a new structure

We now arrive at the heart of this presentation, namely a description of the final product of alkali denaturation; the irreversibly-denatured form which appears as the pH approaches 13. This is the form known as "Form IV".

We start by asking the question: What forces are available, at pH 13, to stabilize a nucleic acid structure? First of all, let's look at base-pairing, starting with the two canonical base-pairs found in typical DNA.

SLIDE 219

If we look all around these two base pairs, seeking protons that might still be available at pH 13 for hydrogen bonding, we find only those *already* known to be so involved. That is, other than the protons and moieties involved in standard W-C base-pairing, which we all learned about in school, there are no other donatable hydrogen atoms in any of the 4 bases, that is, no additional hydrogen atoms which might contribute to the formation of atypical base pairs through tautomeric transformation.

SLIDE 220

Based upon the pK_a's of the amino and imino nitrogen atoms which are responsible for ordinary, everyday A-T and G-C base pairing, it is clear that all four bases would be overwhelmingly deprotonated at pH 13. In other words, there can be *no* significant base-pairing at that pH. Therefore, whatever structure we propose for Form IV, it cannot be a structure that depends upon base-pairing for its stability. What about base stacking?

SLIDE 221

Base stacking, of course, is alive and well at pH 13, and, in view of the higher salt concentration at the higher pH, probably even enhanced. Is it therefore possible that Form IV is merely the Wu 4-stranded structure, only without base-pairing? That is, can a viable structure form from 2 or more DNA strands, stabilized solely by base stacking?

SLIDE 222

Rodley, the author of the non-helical "side-by-side" structure, apparently thought so. The idea was that at pH 13, all the atoms and moieties that constitute DNA would just stay put, only without hydrogen bonds — which, in an axial view of the Wu structure might look something like this.

The slightly-increased distance between members of the base pairs would presumably come about because of charge repulsions, since all the bases would be deprotonated, and therefore all negatively-charged.

But merely proposing that everything just "stands still", only without hydrogen bonds binding the bases together, is insufficient, because if that was the case, then one might perhaps presume that the structure would instantly reanneal when the pH was neutralized, like so. After all, if everything's just "standing still", then reannealing wouldn't require the significant relocation of a single atom, and all that would be necessary would be to lower the pH to the point that protons began to re-appear at the sites of the former hydrogen bonds.

Well, we know that that most assuredly does *not* happen. In fact, the great resistance of Form IV to renaturation is one of its most characteristic and defining properties. It follows that *if* this is presumed to be the structure of Form IV, then we would have to assume additionally that the strands undergo a circular "drift" at pH 13, as indicated in this very crude video:

SLIDE 223

In the beginning, the base pairs are properly aligned, but at pH 13, where there is no base pairing, there is therefore no longer any force to compel the stands to maintain that alignment, wherefore they might drift. Then, when the pH was restored to neutrality, the bases, now out of alignment, would have no possibility for base-pairing, *other than* the relatively weak tautomeric sorts of base-pairing we looked at previously in Slides 109-125. Might it be possible that the bases, having no alternative, would simply *remain* as they are, *i.e.*, in the Wu conformation, but without the native complementary base pairing, just comfortably stacked together in a hydrophobic core, with the hydrophilic phosphate groups arrayed on the outside?

That is, is it possible that Form IV is nothing more than the Wu tetraplex structure – only *without* true base pairing, stabilized mainly by just base stacking? In the previous slide I suggested that Form IV's resistance to renaturation might rule that out, but now we see that it isn't ruled out, because the resistance to renaturation might simply be the result of the strands having undergone this sort of circular drift, so that proper restoration of true complementary base-pairing thereby becomes impossible.

Nevertheless, I know of no precedent for a stable nucleic acid structure whose stabilization is solely through base stacking. The fact remains, however, that there has never been a published proposal for any Form IV structure at all, and since we are therefore dealing with what has been, to date, a complete mystery, it would be prudent to entertain all possibilities, including this one.

I do not propose, however, to dwell excessively on this particular possibility, because there is another model for Form IV which is considerably more plausible, since it

is stabilized not merely by base stacking, but also by salt bridges, the latter of which are known to exist in a multitude of nucleic acid and protein structures.

SLIDE 224

Structure of Form IV

2. Form IV models based upon salt bridges between phosphate groups

Let us continue our consideration of the forces available to stabilize a DNA structure at pH 13. Base pairing is excluded at that pH. Base stacking, however, persists, and may even be stronger. There is one more force available, which we have not discussed previously: salt bridges.

SLIDE 225

Every residue in DNA has a negatively-charged phosphate group, ready and willing to bond ionically to any positively-charged atom or molecule which is available. In any solution of pH 13, there is a cation concentration of at least 0.1 M from the added base alone, and in both the Strider-Warner and the Rush-Warner data at which we have been looking, the salt concentration was a lot higher than that. In Strider-Warner, except in the experiment where it was the salt concentration itself that was varied, the reannealing salt concentration was otherwise always 1 M. In the Rush-Warner pH vs. ***s*** titration, the salt concentration was between 0.3-0.4 M (corresponding to the pH range 7-13). These salt concentrations are significantly higher than the physiological salt concentration of the human body, which is approximately 0.15 M, and may therefore be regarded as "high salt".

SLIDE 226

That is why, since 1980, I have been writing and saying that Form IV most likely has a structure that is stabilized by both base stacking and salt bridges. In my 2002 publication on the topology of non-helical circular DNA, I proposed the primitive conceptual model shown here. In this drawing, I attempted to indicate little more other than that the phosphate groups, at pH 13, turned inward to the axial position, bound together by salt bridges with the ambient cation, while the bases turned to the outside.

The drawing, and the concepts underlying it are, in retrospect, problematical. First of all, Panel C suggests the occurrence of the very sort of overwinding which, in this *current* 2016 slide show, I have labored to rule out. The reason I invoked this overwinding in the 2002 publication was that I believed, at that time, that the entire explanation for the increase in ***s***, of the ϕx174 RF chromosome at pH 13, was relentlessly-increasing degrees of left-handed superhelicity. I used to employ an analogy to the wringing of water out of a wet towel, implying that as the DNA supertwisted further and further, water would be squeezed out, eventually culminating in a dense, largely-anhydrous nucleic acid core.

It wasn't until some years later that I realized there is a supertwist limit, and that there had to be an entirely different intermediate structure between Panels B and D, namely the Wu intercalated tetraplex we have been looking at. Here's that structure

again, at the moment when it first begins to intercalate the apposing sides of the superhelix.

But the problems don't end there. My idea, that one could arrange the phosphate groups into a quasi-crystalline core built around salt bridges, was originally based upon the 1953 Linus Pauling structure for DNA. That turned out to be the wrong model to follow.

SLIDE 227

In 1953, one month before the publication of the Watson-Crick "double helix" structure in the journal Nature, Linus Pauling had pre-empted them with his own proposed *3*-stranded model for DNA, which was published in PNAS. Here's a virtual model of the Pauling structure, which I created from his original polar coordinates. It's very nice-looking from the top.

SLIDE 228

Now here's a video. If you look carefully, you'll note that there are 3 strands, each with a left-handed helical twist. The bases, colored blue and gray, project outward, while the phosphate groups, colored orange and red, face inward.

The Pauling model had two embarrassing errors: First of all, his sugar-phosphate backbone was RNA, not DNA, that is, the ribose was *oxy*ribose instead of *deoxy*ribose. If you take the trouble to look, you'll have no difficulty discerning the red hydroxyl groups at the 2' position of each ribose moiety. This was a minor error, with, at that time at least, little structural significance for the model overall. The second error, however, was far, far worse: his structure was stabilized by axial ***hydrogen bonds*** (!) between the phosphate groups, which would be impossible, because at any pH much above pH 2, DNA phosphate is ***deprotonated!*** We discussed the pK_a of DNA phosphate in Slide 166 above.

Pauling, who was considered by many to be the world's leading authority on the chemical bond, was so demoralized by having made these amateurish blunders that he subsequently and forever removed himself from all DNA research!

But I came to realize that Pauling, although having provided us with what quickly proved to be an *incorrect* structure for naturally-occurring DNA, may have inadvertently provided an important model for the structure of *denatured* DNA, *i.e.*, Form IV. All that was necessary, or so I thought, was to change two things. The first was to increase the strand count from 3 to 4. That was not a problem, because a strand count of 4 was actually *more* consistent with Pauling's underlying hypothesis than his published 3-stranded model, as he himself had already suggested in 1953, at which time he stated that he actually preferred a 4-stranded model, and had only chosen a 3-stranded model because it was more consistent with the Wilkins x-ray crystallography data.

My second change was to replace Pauling's erroneous hydrogen bonds with sodium salt bridges. My thinking on the subject was that a good hydrogen bond and a good salt bridge were both 3 Å, so the change, or so I thought, would be trivial.

SLIDE 229

Although I didn't realize it until years later, the idea that I could simply replace Pauling's hydrogen bonds with salt bridges was naïve. In fact, in thinking that, I was following in Pauling's footsteps, by making a foolish blunder. I had neglected to keep in mind that in a 3 Å hydrogen bond, the distance, from each member of the bond, to the proton in the middle, is only half that, *i.e.*, 1.5 Å, which, curiously, is much shorter than the length of an analogous covalent bond. Thus, in Pauling's theoretical hydrogen bonding between DNA phosphate groups, the distance between the phosphate groups would have indeed been about 3 Å, but in a sodium salt bridge between a pair of DNA phosphate groups, the distance between the phosphate groups would be *twice that*, namely 6 Å. So I was starting off on the wrong foot. But I didn't realize that until years later.

SLIDE 230

By the year 2005 I had finally acquired virtual modeling software, and I created a sort of "throwaway" first model for Form IV. Here's the model in axial view. The base pairs are disrupted, and the bases are no longer axial, as in ordinary DNA, but rather all stacked peripherally, in sandwich formation. The phosphate groups, whose oxygen atoms are colored red, are pointing neatly toward an axial core which would presumably contain sodium ions, each one chelated to an unspecified number of DNA phosphate groups, giving rise to a multitude of 3 Å salt bridges.

SLIDE 231

This model was created in a very naïve way, by simply isolating 4 straight-ladder DNA chains from my protamine-DNA model, and rotating them approximately 180º. Here's what that process looks like, taking a 1-bp-thick section through the tetraplex structure...

When we add the rest of the 4 chains, you get the 2005 version of my proposed Form IV structure.

SLIDE 232

Here's the structure, rotated from the axial to the upright position.

The software I used to create this primitive model was AmiraMol, a product of Mercury Computer Systems. I never owned the software, which cost over $100,000 to buy; it was merely loaned to me, by means of a license that had to be renewed each month with a new activation key. Shortly after I made the old video you see on this slide, Mercury, apparently having made no money on AmiraMol, sold the software! The new owner had no interest in my work, and stubbornly refused to renew my license. I was therefore unable to perfect the model until I could somehow acquire the use of another $100,000 virtual modeling software program. That took nearly 10 years! During that 10 year interval, the project was at a standstill.

Then, in Feb. 2015, another company called Schrödinger, G-d bless them, had just decided to make their Maestro virtual modeling software freely available to academicians. I downloaded it immediately, and set out to perfect the Form IV model, which, as the title of this PowerPoint presentation proclaims, was the "Final Puzzle Piece" in basic DNA chemistry.

As soon as I set out in earnest to perfect the model, however, I ran into trouble. Thus, although I had thought my model to be exceedingly clever when I first proposed it, I must tell you now that it's *wrong*. The phosphate core is possible, but the bases are stacked at 6.8 Å, which is impossible -- or, at least, very unfavorable. The reason the bases are thusly stacked, at exactly twice the normal 3.4 Å stacking distance found in all current DNA models, is because of the way in which the model was created, which I showed you in the previous slide. All I did was to naively perform a 180º rotation of four strands of DNA copied literally from my 2006 protamine-DNA model. In that earlier structure, however, the mutual intercalation of adjacent *pairs* of 6.8 Å-spaced duplexes resulted in a final base-spacing of 3.4 Å. The problem in the structure shown here is that the strands are isolated, and not intercalated with one another, so that the excessively-large 6.8 Å base-spacing is not corrected.

SLIDE 233

But when I tried to correct the model, by adjusting the base spacing to 3.4 Å, I found that adjacent chains clashed terribly in the axial core, with Van der Waals violations so severe that some pairs of atoms were occupying the same space. Here's one of my models, viewed from the axial perspective. It looks a little like the Pauling model from this angle, but when we rotate it to a longitudinal position, and spin it about, the terrible steric conflicts become apparent.

The model shown here was actually one of my better models, but it was a failure. I made eight different models altogether, thinking, each time, that I would learn from the mistakes of the previous one, and solve the spacing problem in the axial core, but no model I was able to come up with was free of severe Van der Waals violations. The best I could do was to reduce the Van der Waals violations to interatomic distances of about 1 Å, which is far too close, even for pairs of hydrogen atoms.

I finally realized that the problem was not merely one of perfecting the model more skillfully, but that, in fact, the concept was just plain wrong. The addition of a 4th strand to the Pauling structure meant that a large number of phosphate groups had to be packed into a defined cylindrical volume. That volume was going to be determined by the base spacing on the exterior of the cylinder. If the base spacing was left at 6.8 Å, as in my naïve 2002 prototype, then the interior could be plausibly modeled. But that base spacing would be energetically unfavorable -- DNA bases don't "like" to stack at 6.8 Å. If, conversely, the base spacing was set at 3.4 Å, then the stacking of bases would be thermodynamically-satisfactory, but the interior would be unavoidably compressed, giving rise to an illogical structure having an impossibly-overstuffed jungle of atoms in the core.

SLIDE 234

The key to solving the Form IV structure was letting go of the pretty and symmetrical Pauling-based design, and keeping in mind that Form IV was not known to have any physiological significance, but was, insofar as any man knows at this point, merely a laboratory artifact. It therefore did not have to align with any other biomolecules, whether charged or not; nor did it have to be elegant, beautiful or symmetrical.

In creating 8 failed structures by rotating the strands of the Wu tetraplex 180º, I was aware that I had *passed by* a position which could be the basis for a plausible Form IV structure, but which seemed an unlikely choice for that structure. But when I finally faced up to the fact that my Pauling-based models were categorically impossible, I found myself, for the second time in my life, being forced to accept the Sherlock Holmesian logical imperative, "When you have eliminated the impossible, whatever remains, *however improbable*, must be the truth". The logical imperative to which I refer was to begin to seriously consider rotating the Wu tetraplex strands, *not* the full 180º, but rather stopping midway, at 90º.

Now, this is a structure with promise. It's a two-part structure, each part consisting of a pair of mutually-intercalated, stacked bases, with the phosphate arms of each part pointing toward the analogous phosphate arms of the other part. Let's add some more base pairs, so we can get a more realistic picture of the structure.

In this structure, the bases are perfectly stacked at 3.4 Å, but the phosphate arms are spread very far apart, so that there's a large hole in the middle. This large angle between the phosphate arms is necessary, because if we were to attempt to decrease this angle by moving the phosphate groups closer to the axial position, Van der Waals conflicts would quickly develop in the space between them. In the form shown here, there are no Van der Waals violations at the apex of the angle between the phosphate arms. But what about that large empty space in the middle? Aristotle said that "nature abhors a vacuum", which is still true today.

In order to turn this into a meaningful model, we must offset its two parts, like so...

Now we have a model that answers to the requirements of Form IV, based on the known and/or likely chemistry of that form:

1. There are no steric clashes or Van der Waals violations.
2. The bases are perfectly stacked – perhaps *too* perfectly stacked.
3. There is no dependence upon base pairing.
4. The phosphate groups face each other, and are properly positioned to form multiple ligands with sodium ions.

Now we must address the issue of optimal base stacking. It is known that this sort of "perfect" stacking, called "sandwich" stacking, is energetically less-favorable than a similar stacking which is less perfect, called "parallel displaced".

SLIDE 235

We can see parallel-displaced base stacking in the ordinary B-DNA W-C double helix, and here's what the bases look like if we remove the sugar phosphate backbone...

Now let's tip it over, to get an axial view, and let's move in to see better.

This is a dodecanucleotide, and here's the bottom base. What we're going to do is to move up the ladder, base-by-base, to see just how much axial overlap there is between successive rungs of the ladder...

As you can see, the answer is "not much". In several instances, there's no overlap at all.

SLIDE 236

One might deduce, from the example of the B-DNA model, that my 2006 protamine-DNA model is flawed, in that the perfect "sandwich" stacking of base-pairs, as seen in this axial view, is energetically suboptimal. This may be true, but the two duplexes, which constitute the structure's DNA tetraplex, can be moved apart a bit without affecting the basic features of the model at all, other than to slightly increase the axial cross-sectional area. Since the total sperm DNA, in the published form of the model, only occupies about 10% of the available volume in a sperm head anyway, a small increase in the cross-sectional area will introduce no difficulty at all, with respect to the DNA fitting into the available space.

SLIDE 237

How will parallel displaced base stacking affect our new Form IV model? Not very much, actually. Here's the model in its primordial form, with that large empty space in the middle. To get away from this sandwich structure, all we need do is to rotate the 4 strands, best seen in this narrow cross-section; to slightly displace the base-stacking; like so. Then, when we offset the two sides, to get rid of that vacuum in the center, we'll see something like this.

By the way, that little purple sphere is a sodium ion. What is the actual size of a sodium salt bridge in a setting such as this?

SLIDE 238

In researching that question, I came across a better data source than I ever thought possible, namely the U. of Edinburgh's massive compilation entitled "Metal Coordination Sites in Proteins". Insofar as the site indicates, the compiler was one Marjorie M. Harding. Thank you Prof. Harding. This is a collection of linked spreadsheets for every likely metal ion found in conjunction with a protein or nucleic acid. The numbers are staggering; for sodium alone the table contains 17,342 entries, with substantial information about each one, including, in many cases, Protein Data Bank references for virtual model files.

The mini-table shown here is my own set of summary statistics, assembled by loading the original table into Microsoft Excel. Thus we learn that the most commonly-

observed sodium coordination #'s are 5 and 6, accounting for more than half the structures which have been studied. The average coordination # for all 17,342 structures was just under 5, and the statistic we most urgently require, the average size of a sodium salt bridge, was 2.57 Å.

If we accept that as a working estimate of salt bridge lengths in Form IV, then the distance between ionically-bonded phosphate oxygen atoms in our model, separated by a sodium ion in their midst, would be twice that, or 5.14 Å.

SLIDE 239

This is my final proposal for the structure of Form IV. This is, of course, an axial view. Seen here are parallel-displaced base stacking, and multiple sodium salt bridges linking the phosphate group oxygen atoms. The sodium ions are not shown , only the oxygen-to-oxygen distances.

This is a Schrödinger Maestro graphic export; some of the measurements are obscured because they are lying directly on top of each other, but I'll tell you that the range of salt bridge lengths, in the current model, is 4.69-5.20 Å, with an average of about 4.9 Å. This is not far from the 5.1 Å value calculated as the average of 17,342 sodium salt bridges from the Edinburgh database.

SLIDE 240

We need to consider the cross-sectional area of the new structure. In Slides 195-212 above, we went to great lengths to demonstrate that the Wu tetraplex structure, shown here again, can account for the 200-300% increase in sedimentation coefficient of ϕx174 between pHs 12.3 and 12.8. If the pH is further increased, we see Form IV, which, at high pH, has the same high ***s*** value as the Wu tetraplex from which it arises. That suggests that it also ought to have about the same cross-sectional area. Does it?

SLIDE 241

Let's look again at the steps in the proposed transition from the Wu tetraplex to my proposed Form IV structure. Here's the tetraplex. Let's put a highlight box around it. Now, please look carefully at this box. I'll bring in the PowerPoint-measured dimensions of this box, because they're going to be important. Now I'm going to bring in a slightly larger box, which I'll explain momentarily, along with its own dimensions. This new box has the same length as the smaller box I just removed, but a slightly-greater height. We're going to use this larger box to examine the size of the new Form IV structure.

Let's now repeat the 90º rotation we looked at earlier -- the rotation which was the basis for the new Form IV structure -- and highlight the rotated structure with the larger box. Note that this larger box fairly-perfectly encloses this primordial sandwich-stacked structure. Now let's repeat the offset, to give the final structure, and bring back the same highlight box, which, once again, fairly-perfectly encloses the final Form IV structure.

We could thus say that the Wu tetraplex structure, at pH 12.8, has *almost* the same cross-sectional area as the Form IV structure shown here, but not exactly the same.

Let's look at the Wu structure again. This smaller highlight box ,which fairly-perfectly encloses the tetraplex, has about 85% the surface area of the larger box which we used to measure the cross-sectional area of my proposed Form IV structure. One might protest that the Wu tetraplex, being a little more compact, might perhaps be expected to sediment a little faster — which it *doesn't*. But we have to recall one additional fact.

SLIDE 242

Let's bring back the larger bounding box for this. As I have been saying, my 2006 protamine-DNA model, which was the source of the Wu tetraplex structure we're using here, had the bases in the perfect sandwich mode of stacking. If we were to offset the two duplexes, as I did above in Slide 236, then we would increase the cross-sectional area a bit. It would be impossible for me to provide exact numerical parameters to precisely define or describe that increase, because I'm not a benzene ring chemist, but mere visual inspection of this graphic model certainly suggests that the increase would be entirely adequate to account for the apparent small discrepancy between the cross-sectional area of the Wu tetraplex and my new model for Form IV.

SLIDE 243

In conclusion, I'm going to propose that the Form IV model presented here is close enough to the Wu tetraplex, in volume per unit length, that we do not need to repeat the laborious comparative sedimentation coefficient calculations we did above, but can fairly-safely presume that this Form IV structure would indeed sediment at 200-300% the rate of the left-handed superhelical form of ϕx174 DNA which is seen at the χ shoulder at pH 12.3.

SLIDE 244
(FORM IV STRUCTURE, AXIAL VIEW)

We close this section by showing a rotating model of the new Form IV structure. Here's the structure in axial perspective…

SLIDE 245
(FORM IV STRUCTURE, LONGITUDINAL VIEW)

…and here's the structure in longitudinal perspective. A Jmol model of the structure may be viewed on this website, and a pdb virtual structure file may be downloaded.

SLIDE 246

Structure of Form IV
3. What happens to Form IV during neutralization (pH 13 back to pH 7)?

SLIDE 247

We have now traced the denaturation curve for ϕx174 chromosomal DNA up to pH 13, proposing that in the final stages before complete denaturation, between pH 12.3→12.8, the DNA assumes the Wu intercalated tetraplex structure; the same one I proposed in 2006 for DNA in the protamine DNA complex. We have also considered, *and, for the most part, rejected* the proposition that at the final pH of 13, where base-pairing is essentially ruled out, the DNA simply remains approximately in the Wu conformation, only without base-pairing.

In place of that latter unwieldy hypothesis, we have instead put forth a novel structure for Form IV; one stabilized by salt bridges between the phosphate groups. This structure satisfies all the requirements laid upon it by the available data. It has no dependence upon base pairing, but relies only on forces that are still very much available at pH 13, namely base stacking and electrostatic interactions. The model is devoid of steric hindrances, and there is therefore no reason to doubt that this structure, or one very much like it, could appear under the circumstances.

We must now embark upon the neutralization leg of this journey, the trip from pH 13 back to pH 7, and show that our new Form IV structure can explain these data as well.

SLIDE 248

Actually, we have not one, but *four* tasks still remaining:

1. To explain why the denatured species becomes *less compact* as the pH is dropped back toward neutrality.

2. To explain the apparent shoulder in the upper neutralization curve at approximately pH 12.

3. To suggest a specific structure at pH 7.

4. To suggest a specific structure for the ***duplex*** form which appears under the low ionic strength conditions employed for electron microscopy.

Explaining the drop in ***s***, as the pH is dropped from 13 back down to 7, has been the most difficult challenge of this entire Form IV project. I have two potentially different explanations, although it's also possible that they both pertain. The first is immensely complex, and it's difficult to form a clear mental picture of it. Nevertheless, we shall have to invoke its key features, for without them, there will be no way to explain most of what goes on in the upper curve.

Conversely, the second explanation is immensely simple, although it proved to be a bit *too* simple, and, by "the end of the day", it proved insufficient to completely explain the data. I shall include it anyway, because the complete explanation may require invoking both mechanisms.

SLIDE 249

"Explanation #1: neutralization reverses all the denaturation steps"

SLIDE 250

The first, and more complex explanation may be succinctly described as a reversal of all the steps in the denaturation curve. In other words, we might first see a return to the Wu tetraplex structure, with the bases once again axial and the phosphate groups once again peripheral.

SLIDE 251

As the pH drops, we might see some or all of the chromosome returning to a duplex form, as I've tried to suggest with this greatly simplified drawing, where two representative base pairs depict a Watson-Crick sort of double-stranded structure.

SLIDE 252

The first of these steps, the return to the Wu tetraplex, can be thought of as being the result of two changes. First of all, lowering the pH also lowers the ionic strength, and that will weaken the salt bridges, destabilizing the Form IV structure. Probably more significant, however, is the self-evident fact that in dropping the pH, as the pH falls below the pK_a's of the bases, the possibility of base-pairing begins to re-emerge. Since, however, the proper base-pairs are no longer juxtaposed, this would have to be *faux* base-pairing, resulting mainly from tautomeric base forms. This is why I insisted that we review tautomers *before* proceeding to this discussion of Form IV.

I said "mainly" tautomeric; we need to keep in mind that there are ordinarily only 4 bases in DNA, so that in any random juxtaposition of two DNA strands it must be the case that, on the average, every 4^{th} base will find itself adjacent to its proper complementary base by mere chance. That's 25% complementarity, which is a good start. Can the other 75% of bases attain to base-pairing by means of tautomeric conversions? As we said in the discussion of tautomers in slides 109-125 above, the energetics are vague. The free energy *decrease* associated with formation of a base pair is in the same general range as the free energy *increase* associated with a change to an unfavorable tautomer. We therefore cannot say with certainty that this scenario, invoking tautomeric base-pairing, will occur, but only that it seems quite possible.

SLIDE 253

Continuing with this conceptual approach, that is, with the view that the neutralization curve proceeds by the reversal of all the steps of the denaturation curve, we would expect next, as the pH/ionic strength drops still further, that we would, at some point, begin to see the reappearance of duplex DNA. This must be the case, because electron microscopy of Form IV reveals duplex DNA, showing that somewhere between {pH 13, high salt} and {pH 7, low salt} the duplex form does indeed return.

SLIDE 254

The problem with this scenario is that even at pH 7, the sedimentation coefficient of neutralized ϕx174 DNA is still at 36 ***s***, much higher than the native sedimentation coefficient of 21 ***s***. I believe that this *necessarily requires us to suppose* that some portion of the chromosome *must* still retain a tetraplex structure, even at pH 7, because no known *duplex* DNA conformation can account for such a high ***s*** value. In other words, the energetics of our hypothetical tautomerically-base-paired duplex structure are sufficiently poor, compared with those of a truly complementary-base-paired structure, that we are all-but compelled to believe that portions of the chromosome remain in either our 4-stranded, salt-bridge-based Form IV structure, or the Wu 4-stranded intercalated structure, or perhaps some of each, all the way down to the pH 7 end of the curve.

SLIDE 255

Forming a picture of these events, in the "mind's eye", is rather difficult. For me, at any rate. In general we're *presuming* the complex process previously illustrated by this picture, where the conversion from the duplex to the tetraplex structure, taking place between pH 12.3 and pH 13, was portrayed as occurring in a stepwise fashion, as opposed to an all-or-none cooperative transition. In bringing back this picture, I must remind you that the size of the mismatched regions is, in all probability, wildly exaggerated here, in order to make the point.

The reverse journey, however, is different. Assuming that during the neutralization titration, all the denaturation steps are gradually reversed, we would then have to conclude -- based upon the persistence of a high sedimentation rate at pH 7 -- that even at pH 7, the conversion back to duplex is incomplete, at least at the relatively high ionic strengths of the Rush and Warner experiments. In order to complete the reversion back to fully-duplex structure, as seen under electron microscopy, we must, additionally, *lower the ionic strength.*

Why should lowering the salt concentration bring about a tetraplex-back-to-duplex transition? Alas, we shall not be able to answer this with pinpoint precision, but two explanatory mechanisms come to mind. It may perhaps be the case that both mechanisms co-exist during the neutralization process.

One possible explanation might be that the mutual intercalation of base pairs creates a significantly hydrophobic core within the tetraplex. Increasing the pH and/or salt concentration, which increases the hydrophilicity of the solvent, might then favor the sequestration of stacked base pairs in the hydrophobic tetraplex core. Conversely, lowering the pH and/or salt concentration would *decrease* the hydrophilicity of the solvent, *reducing* the tendency of the bases to congregate in a hydrophobic core. Under the latter conditions, with the bases tolerating a more intimate contact with the solvent, the favorability of the more open duplex structure might thereby increase.

Another possibility is that the intermediate structure which partially persists all the way down to pH 7 is not the Wu intercalated structure at all, but rather the salt-bridge based structure we have proposed for Form IV. In that case, the effect of lowering the salt concentration would be easy to understand, since lowering the salt concentration

would simply *decrease* the number of salt bridges upon which the fundamental stability of that tetraplex structure depends.

With this in mind, let us take a short detour, and consider an alternative way to explain the entire neutralization curve, which invokes only a single, easy-to-understand principle.

SLIDE 256

"Explanation #2: Theory that the primary effect of neutralization is merely reduction of the ionic strength, with consequential weakening of salt bridges."

SLIDE 257

There is a very simple way to explain the gradual drop in ***s*** in the neutralization curve, which begins by considering the DNA to be persistently and exclusively in the most-compact, 50 ***s*** conformation we have proposed, *i.e.*, in the salt-bridge-based axial phosphate tetraplex form, *throughout* the entire neutralization curve.

SLIDE 258

We then consider the possibility that the primary effect of lowering the pH is *not* the effect brought about by the change in $[H^+]$ concentration, but the rather the effect brought about by the change in $[Na^+]$ concentration, *i.e.*, ionic strength. We must keep in mind here that the changes in pH, in the Rush-Warner titration we are discussing, are brought about by addition of NaOH. Thus, at pH 13, the ionic strength is about 0.4 M, whereas at pH 7 it has dropped to 0.2 M. This latter concentration, while only half the former concentration, is still relatively high in comparison to the physiological salt concentration of human body fluids, which, as we have previously noted, is 0.9% W/V, corresponding to 0.15 M.

Such relatively high salt concentrations are probably necessary to support the 50 ***s*** Form IV model we have proposed, which is dependent upon a crystalloid axial structure. We can say this with some confidence, because the structure that appears at pH 7 in the neutralization curve, a point where the salt concentration remains at the relatively high value of 0.2 M, has the persistently-high sedimentation coefficient value of 37 ***s***, which cannot be accounted for by any known duplex structure, whereas at the much lower ionic strengths used in the preparation of Form IV for electron microscopy, the structure has clearly become *duplex*, suggesting that the lowering of ionic strength, at some point, rules out tetraplex structures based upon salt bridges.

SLIDE 259

We may therefore ask: What might be the predicted effect upon our Form IV structure, when we lower the ionic strength from 0.4 M to 0.2 M? Might it be the case that as the salt concentration drops, the number of salt bridges holding the structure together *diminishes*, resulting in a progressive *weakening* of the collective residual force holding the two halves of the structure together in the tetraplex form? It seems to me that

if there is a weakening of the force holding the two halves of the structure together, then the *lengths* of the salt bridges remaining would *increase*, like so...

If this point of view is correct, then the conversion to a duplex form, at the very low ionic strengths used in preparing DNA for EM, makes perfect sense. We merely surmise that when the salt concentration is low enough, the salt bridge length exceeds that which is necessary to maintain the ionic bond at all, and the 4-stranded structure disintegrates into a 2-stranded form.

SLIDE 260

With this in mind, I revisited the previously-mentioned 17,342-entry sodium ligand table; this time to extract a summary of bond length data. Recall that the average salt bridge length in that table was 2.57 Å. Here we see the distribution of ionic bond lengths, revealing that the great majority of them, about 15,500 out of the 17,342 total, were in the 2-3 Å range. Very few were either *less than* 2 Å, or *greater than* 3 Å.

The main point of the table is to document that the total range of ionic bond lengths was 1.627-3.408 Å. Only about 5% of them were >3Å, and, in the total group of 17,342 structures in the Edinburgh database, none can be found having Na+ ligand bond lengths longer than 3.5 Å. We may surmise, therefore, that under hypothetical conditions where a putative salt bridge, *if* it were to form, would have to be greater than 3.5 Å, such a salt bridge will probably never be seen, because 3.5 Å is very likely the upper limit of what is possible.

SLIDE 261

Let us therefore presume that during neutralization of Form IV, lowering the ambient salt concentration weakens the attractive force between the two halves of the Form IV structure, resulting in a lengthening of the remaining salt bridges. From their starting length of 2.57 Å, they can increase up to about 3.5 Å, which is the distance from the sodium ions to the phosphate group oxygen atoms. Since there are no known sodium ligand lengths greater than 3.5 Å in length, we may presume that when the ionic strength is lowered still further, the salt bridges will become too weak to maintain the 4-stranded structure, and the two halves of the tetraplex structure will not be able to remain together, whereupon the structure will revert to a 2-stranded form, as depicted by this *extremely* whimsical graphic representation.

Note that the two strands of the reconstituted duplex chromosome, as presented in this slide, are being held together not by base pairing, but solely by base stacking. This could very well be the final structure of Form IV, after it is both neutralized, and then prepared for EM studies by greatly lowering the ionic strength. This structure has a single base stack which extends around the entire periphery of the chromosome.

I know of no precedent for such a structure, but nothing about it which is either energetically or sterically problematical, so I can advance this as a candidate for the ultimate Form IV structure – at neutral pH and low ionic strength – with considerable conviction.

SLIDE 262

The alternative possibility for the low-salt structure is for the bases to unstack themselves, and to resume Watson-Crick type base-pairing, as in this equally-whimsical graphic representation. I'm not going to make any attempt to portray this more accurately, since this would merely be the Watson-Crick structure itself, drawings of which are already found in countless textbooks of molecular biology. The only difference would be that the base-pairing here would have to be tautomeric, which, in principle at least, wouldn't change the visual appearance of the chromosome, as compared with the visual appearance of classic Watson-Crick DNA.

As I've pointed out twice previously, the energetics of a tautomerically-base-paired duplex DNA structure are vague. I suspect that they are possible, as long as there is a process for bringing them about. That process is provided here by the denaturation-neutralization sequence. Which of these two candidates for final Form IV structure, the single base stack, or the tautomerically-base-paired form, is energetically most favorable, I cannot adjudicate.

I must say, however, that I favor the tautomerically-base-paired structure. This is not a scientific conclusion, because the energy calculations to confirm the feasibility of tautomeric base-pairing are beyond my meager skills in physical chemistry. Nevertheless, my instincts tell me that the tautomeric base-pairs, while being energetically inferior to proper base pairs, will nevertheless confer some small decrease in the free energy. Then, since 1-out-of-4 base pairs will be coincidentally correct, and since the Watson-Crick species of structure also has the benefit of base-stacking, this form might be energetically "preferred" to the "base-stacking-only" structure, with which it completes for our attention.

In other words, the final structure of Form IV, at neutral pH and low ionic strength, may be this:

SLIDE 263

This is actually a video of the classic Watson-Crick structure. If you look closely you'll see that it's properly base-paired, but if this is, in fact, the ultimate structure of Form IV, at neutral pH and low ionic strength, then the only differences would be, first of all, 3/4 of the base pairs would be tautomeric, and, secondly, some of the structure would have to be left-handed, which I've made no effort to portray here.

SLIDE 264

Now, this explanation -- for the sedimentation coefficient changes in the neutralization curve for Form IV -- which we've enumerated "Explanation #2" -- is all very neat and tidy, with a single mechanism being invoked to account for the entire curve down to pH 7, and beyond that into the realm of low-ionic-strength seen in specimens prepared for electron microscopy. I like the neatness and the tidiness of it, but will it explain the data?

That is, will the lengthening of salt bridges, by itself, explain the entire decrease in sedimentation coefficient, from about 50***s*** at pH 13, down to about 36***s*** at pH 7?

In order to know the answer to that question , we're going to have to take some measurements.

SLIDE 265

Here's an axial view of our presumed Form IV structure at pH 13, with salt bridges averaging 5 Å between phosphate groups. The sodium ions have been excluded from this picture, wherefore the 5 Å spacings must be understood to be the sums of the lengths of pairs of 2.5 Å distances between the phosphate oxygen atoms and the sodium ions.

I used Adobe Photoshop as a tool to facilitate a comparison of the cross-sectional area of *this* structure, which I have invoked at pH 13, with the structure Form IV might assume at pH 7, the latter of which structures I'll show you momentarily. I chose 6 atoms as corners of a polygon which best encompasses the structure, then I used Photoshop to "triangulate" the polygon. The Photoshop measurements are in mm, not Å, which is adequate for our current purposes, because the area is merely going to be used for comparison purposes; to compare this area to another area which will also be given in mm, so the comparison will be valid, even though the units of length are arbitrary.

The area of this axial cross-section can be approximated as the total of the areas of the 4 triangles, which were calculated using the ancient method of Hero of Alexandria, known as Heron's formula, which today, mercifully, we can place in a Microsoft Excel formula cell. This I did, whereby I quickly determined that the Photoshop area of this structure comes out to 16,096 mm^2.

Now let's mentall decrease the salt concentration, which will *increase* the salt bridge length. According to the Edinburgh database we've looked at several times, these sodium salt bridges can expand up to, but not beyond approximately 3.5 Å, giving a phosphate-to-phosphate average distance of 7 Å, as shown here. When we extend the encompassing polygon to the new expanded structure, we find a new and enlarged area of 17,785 mm^2.

Bearing in mind that we're looking at axial sections through long DNA rods, we can refer back to Slides 195-212 above, where we adapted the Svedberg equation to the same sort of analysis. In that earlier discussion we compared the sedimentation coefficients of the pH 12.3 superhelix with the pH 12.8 tetraplex. We learned there that by treating the long DNA rods as if they were cylinders, the ratio of the sedimentation coefficients of the two forms was, to a first approximation at least, inversely proportional to the ratio of their radii.

Here we can do the same sort of thing. We can consider the cross-sectional areas of these structures as being roughly equivalent to the cross-sectional areas of hypothetical rods of perfect cylindrical shape. The ratio of the radii will then be the same as the ratio of the square roots of the areas, so that the predicted ratio of the sedimentation coefficients, which will be the inverse, will be...

SLIDE 266

...1.05. Well, this is not going to work. The actual experimentally-observed ratio between the ***s*** values of the two forms is ... 1.39, which is much higher.

So this hypothesis, that the entire upper neutralization curve can be explained by a lengthening of salt bridges, as the ionic strength decreases, is incorrect. But I have presented it anyway, because it may still *contribute* to the observed decrease of ***s***, even though it cannot explain the entire decrease.

SLIDE 267

There seems to be no way out of the conclusion that, in the Rush/Warner experiment, the structure of the neutralized product at pH 7 *must* contain a mixture of persistently-tetraplex and newly-duplex regions, as depicted by our exaggerated and melodramatic picture.

I simply cannot, as of the time of this writing, find a fully-*tetraplex* structure by means of which I can explain an ***s*** value as low as 36, or, conversely, a fully-*duplex* structure with an ***s*** value that high. Pending new information, then, I must conclude that the structure of neutralized Form IV, at pH 7, 0.2 M salt, is a hybrid structure which is still in the process of converting from the 4-stranded to the 2-stranded form. The completion of that process will not be seen until the ionic strength is lowered to that range which is used in the preparation of DNA for electron microscopy, as depicted by the open circle which is completely outside the neutralization graph.

SLIDE 268
"Tasks Remaining"

SLIDE 269

Earlier, we mentioned 4 tasks still remaining. In a somewhat rambling manner, we have now addressed three of them.

- We have extensively discussed the likely set of explanations for the decrease in compactness as Form IV is neutralized.
- While we are unable to define a single, precise structure for Form IV at pH 7, we have been able to demonstrate that it is very likely a heterogeneous structure within which are found both duplex and tetraplex regions.
- We have shown that there are two plausible competing structures for the low-ionic-strength species of Form IV, at the ionic strengths used in preparation of samples for electron microscopy; one of which structures is a Watson-Crick-like duplex in which the bases are only stacked, with no base pairing, and the other in which the bases are paired, looking very much like normal duplex DNA, only with 75% of the base pairs being in tautomeric conformations.

This leaves one task remaining, namely to explain the apparent shoulder in the neutralization curve, at approximately pH 12.

SLIDE 270

In Slide 42, we considered the possibility that this shoulder may be an artifact, and since there's only one neutralization curve that's ever been published, namely this one, we must draw our own conclusions about this. My conclusion is

that the shoulder is real. When we add back the lower denaturation curve, we are reminded that both the upper and lower curves have perturbations in the same realm of pH, namely pH 12, and I'm inclined to think that this is not mere coincidence.

What's going on in the lower denaturation curve is clear enough. It is not in dispute, in any quarter as far as I know, that during the denaturation sequence of events, as the DNA passes through the pH 12 region, the direction of superhelicity changes from right-to-left.

Although we cannot speak with absolute certainty of the helicity in the final section of the denaturation curve, there are two things we can say with some degree of certainty about it. First of all, with respect to twisting, any and every rod-like polymer has only one conformation where there is no twist, but infinite numbers of conformations in which it may, in theory at least, be twisted in either the right-handed or left-handed sense. It is therefore a virtual mathematical certainty that the DNA, in this portion of the denaturation curve, will have some sort of twist. Secondly, from what we know about the relationship of helical handedness to pH, there is no reason, *a priori,* to doubt that the left-handed superhelical sense, established unequivocally at pH 12.3, will continue as the pH increases up to the top of the curve at pH 13.

Conversely, we know, that under conditions of neutrality and low ionic strength under which DNA samples are prepared for electron microscopy, that the DNA has resumed a duplex form, and that at neutral pH, duplex DNA always "prefers", so-to-speak, a right-handed secondary winding, wherever possible. When not possible, then any deficiency of right-handed secondary twists causes the appearance of compensatory right-handed tertiary supertwists, according to the topology equation we have mentioned numerous times previously.

This means that at some point, between here and here, the backbone helical and/or superhelical winding preference will revert back to right-handed. All things considered, it would appear most likely that that change in winding preference would occur at the same pH in the upper curve, as it does in the lower.

SLIDE 271

Comparison of pH vs ***s*** data (Rush-Warner)
and
[Na^+] vs ***s*** data (Pouwels *et al)*

SLIDE 272

Because most of the data -- from which we've drawn so many conclusions -- comes from the studies of Mark Rush and Robert Warner, it's useful and interesting to try to correlate it with work of a completely different sort, from a completely different lab; namely that of Pouwels *et al* in the Netherlands, whose work we examined previously.

The Pouwels study, done exclusively at pH 7, looked at the sedimentation coefficient of Form IV as a function of salt concentration, starting at 1 M, and decreasing down to 0.001 M. To the data we have previously examined, I have here added the locations of the salt concentrations at the beginning and end of the Rush-Warner

neutralization curve, *i.e.,* from 0.4 M, down to 0.2 M sodium, and the ***s*** values are, in part at least, consistent with those reported by Rush and Warner.

That is, if we return momentarily to Rush-Warner, we can remind ourselves that the calculated salt concentration, largely the contribution of added sodium hydroxide, as the pH was lowered from 13 down to 7, was in the range I just mentioned, namely from 0.4 M down to 0.2 M, and that the sedimentation coefficient, in that range, dropped from about 50 ***s*** to about 36 ***s***. At least one of those two sedimentation coefficients corroborates the Pouwels data, namely the final sedimentation coefficient of 36 ***s***, which agrees closely with the Pouwels sedimentation coefficient at 0.2 M salt, of just under 35 ***s***.

At 0.4 M, however, there is a discrepancy. The Pouwels result there is qualitatively higher, namely just *above* 35 ***s***, but nowhere near the 50 ***s*** seen in the Rush-Warner pH vs. ***s*** titration. To try to explain this in terms of an exact molecular model is something I will not attempt to do. I will note, however, that this discrepancy shows that the 50 ***s*** species of Form IV may not be attainable with salt alone, because even at the Pouwels maximum salt concentration of 1 M, the sedimentation coefficient only goes up to 40 ***s***, which is pretty high, but not quite the 50 ***s*** figure seen in the Rush-Warner experiment.

SLIDE 273

This suggests that the highest observed ***s*** value in the Rush-Warner ***s*** vs pH titration curve, namely 50 ***s***, requires not merely high salt concentration, but also high pH; that is, the total abolition of all base-pairing. Conversely, in the Pouwels ***s*** vs [salt] titration, the pH is always 7, therefore, even at 1M NaCl, the highest ionic strength tested, base pairing is not at all impaired.

SLIDE 274

This suggests to me that at 1M salt, but neutral pH, Form IV is probably heterogeneous, with some regions having reverted to the duplex state, while others have retained the Wu-tetraplex-type of structure, consisting of two strands of tautomerically-base-paired duplex DNA, whose base pairs are mutually intercalated. This I say because the 40 ***s*** sedimentation coefficient corresponds to that seen in the Rush curve at ***this*** point, where I believe the rapidly-increasing compactness comes about because of a rapidly-increasing proportion of the DNA being in that same Wu tetraplex conformation, which does not fully morph into the 50 ***s*** form until the pH is increased all the way to 13.

SLIDE 275

Before moving on, there are two remaining topics worth discussing, with respect to a comparison between the Pouwels data, and the data of the Rush - Warner study, and of the other studies we have been looking at. *Why,* in the first place, would the ***s*** value of Form IV, at neutral pH, *increase* at high salt concentration at all? We began to address this question in Slide 255 above, where I suggested that the mutual intercalation of base pairs would create a hydrophobic core within the tetraplex, which might be energetically-

favored at high salt concentration. Thus, the fraction of Form IV's length which would be in the Wu tetraplex conformation might therefore increase *pari passu* with increasing salt concentration, since "increasing salt concentration" also means increasing *hydrophilicity* of the solvent, which might promote an increase in the *hydrophobicity* of the core.

The other possibility is that, when the salt concentration gets very high, even at pH 7 we might begin to see regions of the chromosome assuming the 50 s type of salt-bridge-based Form IV structure, that is, our maximally-compact Form IV tetraplex structure with the phosphate groups in the axial position, bound by sodium salt bridges.

If we were to further consider this latter possibility, we might also ask whether perhaps even *higher* salt concentrations might push even *more* of the Form IV molecule into the 50 ***s*** conformation, but Pouwels doesn't go beyond 1 M, so we cannot answer the question. Until such additional salt concentration data becomes available, we can only argue, from the evidence, for the complete conversion to the 50 ***s***, salt-bridge-based Form IV structure, at a pH of 13, where all base-pairing is abolished.

SLIDE 276

The second topic worth discussing, with respect to a comparison between the Pouwels experiment and the others we have looked at, is the behavior of Form IV at the low-ionic strength region of the Pouwels data. At those ionic strengths, Form IV has decreased in compactness *so much* that it actually starts to be *less* compact than Form I, clearly indicating a 2-stranded structure with low degrees of superhelical winding.

This is very much in agreement with the duplex structures we have seen in the electron micrographs of Form IV at low ionic strength.

SLIDE 277

SUMMARY and CONCLUSIONS

This slide summarizes and reiterates all we have proposed concerning what can be deduced about the structure of Form IV from the available data, most of which is based upon studies of the RF, or replicative form of the 5 kb chromosome of the virus ϕx174.

At pH 7, the chromosome, as isolated from nature, is known to be a right-handed, or "negatively" supertwisted superhelix, as illustrated by the purple-highlighted drawing. The drawing only shows 3 right-handed supertwists, in actuality there are about 25. There is very little in the way of conformational change between pH 7 and pH 11.6.

But between the pH's of 11.6 and 12.3, the direction of supertwisting reverses, culminating in a new superhelical structure with 25 ***left-handed*** superhelical twists.

At pH 12.3, with 25 left-handed superhelical twists, a variety of different types of published data demonstrate, redundantly and persuasively, that the maximum possible degree of superhelicity has now been reached, wherefore further increases in pH force the appearance of a totally new structure, the Wu mutually-intercalated tetraplex. The shoulder in the curve marks the beginning of this conformational change. The transition

is not cooperative, but occurs gradually, as the pH is increased, culminating, at about pH 12.8, in an extremely compact, totally tetraplex structure.

But the conformational change, even at pH 12.8, is still reversible. *Ir*reversible conversion to Form IV requires a H of 13, at which pH base pairing is totally abolished, requiring that the tetraplex be maintained by the remaining forces which are available at that high pH. These are base-stacking and salt-bridging.

SLIDE 278

Thus, the 4 strands of the tetraplex rotate, so that the phosphate groups now face inward, giving rise to a new axial structure, stabilized by sodium salt bridges between the negatively-charged phosphate oxygen atoms. At the same time, the bases rotate to the periphery, where they stack, without any base-pairing at all. This is the irreversibly-denatured Form IV structure.

SLIDE 279

If this structure is merely neutralized, without any special attention to temperature and ionic strength, the upper curve is seen. The events suggested by the neutralization data in the upper curve are complex, and not easy to illustrate with little drawings such as these.

It is probable that the neutralization brings about a stepwise reversal of all the changes that led to Form IV during denaturation; the only difference being that the native complimentary base pairing is irrevocably damaged at pH 13, so that upon neutralization, only *tautomeric* types of base-pairing are possible. We did not conclusively rule out the possibility that the neutralized forms have *no* base pairing at all, but only base-stacking, but that theory was not developed, because I suspect that tautomeric base-pairing is indeed possible, and that the chromosome will prefer a conformation where there is base-pairing *and* stacking, and not just the latter.

Since tautomeric base pairing is relatively weak, the chromosome will not — if we may speak anthropomorphically — cheerfully relinquish its tetraplex base stacking. Thus, the conformational changes, all the way down to pH 7, are non-cooperative and stepwise, that is, some regions of the chromosome will retain a tetraplex structure, for the base-stacking benefit therein, while others are gradually — not cooperatively — converting to a duplex form.

SLIDE 280

I must presume that, at the beginning of the neutralization process, there is first a reversion of the salt-bridge-based tetraplex structure we have proposed at pH 13, back to the Wu mutually-intercalated tetraplex structure as the pH drops, with the phosphate groups moving back to the more familiar peripheral position which is seen in ordinary DNA.

With further decreases of pH, we must furthermore begin to see portions of the chromosome reverting from the compact tetraplex conformation, back to the much less compact duplex conformation.

Why do we say this? Because at pH 7, the sedimentation coefficient has dropped to about 36 ***s***, a number far too low to explain in terms of a fully-tetraplex structure. At the same time, however, we must also note that 36 ***s*** is far too *high* to explain in terms of a fully-duplex structure, so it must be a *hybrid* structure, with some regions of the chromosome having fully-reverted to a duplex conformation, while other regions still remain in the Wu mutually-intercalated tetraplex conformation.

SLIDE 281

In order to achieve the full return to duplex structure, we must lower the ionic strength to the realm of ionic strengths employed in the preparation of DNA samples for electron microscopy, because it is only under those circumstances that we see, in those electron micrographs, a fully-duplex structure for neutralized Form IV, looking very much like native double-stranded DNA.

True complementary base-pairing, however, cannot exist in Form IV, except under the narrow and precisely-defined conditions of pH, temperature and ionic strength demonstrated by Strider and Warner to be absolutely necessary for true renaturation. Therefore, either the new duplexes are tautomerically-base-paired, or else in a conformation similar to that of Watson-Crick DNA, only with all the bases piled up in a single stack, without base pairing at all, as I attempted to illustrate above, in Slide 261.

I'm inclined to believe that the tautomerically-base-paired form would be energetically preferred, but proving that from energy calculations is quite beyond my meager abilities as a physical chemist.

SLIDE 282

EPILOGUE

What is the interpretation of the Strider-Warner curves?

SLIDE 283

I have now explained, to the best of my meager abilities, the various and complex conformational changes that DNA undergoes, as it is taken through a cycle of alkali denaturation, culminating at pH 13 with a novel tetraplex form known as "Form IV", then back to neutral pH, whereupon it ultimately resumes a form similar to that of native duplex DNA, only without true complementary base pairing.

There is little I can add to the subject of Form IV structure. But there is one question which remains unanswered, and it is a question of such fundamental importance that it can be said that, to a large extent at least, it prompted the entire Form IV inquiry in the first place: What is the meaning of the Strider-Warner renaturation optima? What exactly is happening at these points?

These curves show that the structure of Form IV is stable over a large range of physical conditions, showing little or no tendency to renature, but that under certain narrowly-defined sets of conditions, and *only* under these conditions, Form IV *will* renature, changing back to the native Form I. If the parameters of pH, temperature or ionic strength are made either *more*, or *less* extreme, the renaturation rate decreases sharply, quickly dropping to zero.

What can we say about this? First of all, it is clear that Form I, under physiological conditions, and throughout most of the pH scale, is more stable, *i.e.*, exists at a lower free energy state than Form IV. Of this we can be certain, because, short of complete denaturation at pH 13, there are no conditions known, between pH 7 and 13, where Form I will convert to Form IV, but many sets of conditions, such as those shown here, where Form IV will convert to Form I. There is also unpublished evidence, from the laboratory of Robert W. Chambers, which suggests that Form IV will revert slowly to Form I upon prolonged storage in the refrigerator, whereas Form I does *not* convert to Form IV under those circumstances. This was discussed in "The Double *Non*-Helix, Part I", on this web site.

SLIDE 284

We can therefore portray the energetic relationship between the two forms roughly as per this simple diagram...

This diagram, like the data it represents, suggests that the processes for interconversion between Forms I & IV involve an "activation energy" barrier. Disruptive energy must be added to the system, so that there appears an inherently unstable, transient structure *intermediate* between the two.

SLIDE 285

As I explained, in "The Double Non-Helix, Part I", this intermediate state must, logically speaking, be a state in which the two circular single strands are not tightly locked into a duplex conformation, but rather are capable of rotating with respect to one another, until the original complementary base pairs have "found" each other again, and the original Form I structure can thereby be reconstituted.

SLIDE 286

The explanation for the Strider-Warner reannealing optima then becomes fairly straightforward. We can discuss it more easily if we isolate one illustrative curve, so we'll use the 60º curve. At that temperature, in 1M salt, a pH of 10.4 fails to provide the necessary disruptive energy to bring about renaturation. Energetically speaking, the Form IV can approach the free energy barrier to renaturation, but it cannot get over, so it remains Form IV.

Conversely, at pH 11.4, *too much* disruptive energy is being provided. The only thing that the chromosome can do under extreme conditions of pH, temperature and ionic strength, is to convert into Form IV. But it's *already* Form IV, and there's nowhere else for it to go so, once again it simply remains Form IV.

It's only at conditions of optimum pH, temperature and ionic strength, in this case ~pH 11, that Form IV is finally pushed to the top of the activation energy barrier, but not beyond.

Had the pH been either higher or lower than 11, the chromosome would have remained in the Form IV conformation. But at pH 11, there is a structural uncertainty.

Under these circumstances, neither Form I nor Form IV is favored, but rather the chromosome is exactly between them in free energy. At this point only, if we may speak anthropomorphically, the chromosome cannot "decide" whether to couple together its two strands by means of its bases, to form base pairs, or rather to rotate its strands by 90º, in order to couple by means of its phosphate groups, via the formation of salt bridges. It therefore exists as an unstable intermediate, between the two stable structures.

It's in this intermediate state that the strands are free to rotate with respect to one another. When the native complementary base pairs "find" each other, then and only then can the chromosome drop into the Form I ΔG valley, out of which it will not thereafter move, unless and until it is once again subjected to extreme conditions of pH, temperature and/or ionic strength.

SLIDE 287

There are a number of other topics which I'd like to discuss, such as the possibility that the Rush-Warner point of permanent denaturation of ϕx174 at pH ~12.8 might be a type of Strider-Warner renaturation optimum, only one where the unstable species lies between the Wu tetraplex, and our salt-bridge-based Form IV tetraplex.

However, if denatured Form IV was incubated in 1M salt, 20º, and pH 12.8, renaturation, if it occurred, would only yield the Wu tetraplex! This would not be noticed, because the Wu tetraplex has about the same sedimentation coefficient as Form IV. No Form I could appear, because Form I cannot exist under these extreme conditions.

But I sincerely doubt that anyone in the world, with the exception of myself, would be very interested in that subject. So I won't mention it further. Besides, this presentation is already very long – much longer than I anticipated at the outset. Therefore it's time to end.

I said at the outset of this PowerPoint presentation that, in the course of the studies that had been necessary to elucidate the structure of Form IV, there had unexpectedly been revealed to me a probable way to easily separate the strands of Form I. This is important. The failure of the strands of Form I to separate, under the typical sorts of conditions used to denature linear DNA, has, for over a half-century, been the chief reason that the molecular biological establishment has been fiercely resistant to non-helical DNA structures.

Tai Te Wu is the only scientist who has ever succeeded in non-destructively separating the strands of circular duplex DNA, but the Wu experiment was very difficult, time-consuming and expensive, and moreover, it was only accepted for publication in an obscure journal which no one reads.

My method is startlingly inexpensive, preposterously easy to perform, and, moreover, can be completed in about 5 minutes. The only catch is that the products must be verified by electron microscopy.

I have decided to place the detailed protocol for this experiment in another area of the NotAHelix web page, because it really isn't logically a part of the Form IV subject. You will have no difficulty finding it, because nothing on this web site is at all hidden.

We are therefore now finished with the subject of Form IV. I hope you have found the presentation interesting, and I thank you for watching.

SLIDE 288

THE END