The Probable Structure of the Protamine-DNA Complex

By Ken Biegeleisen, M.D., Ph.D.
(212) 717-4422
kb@NotAHelix.net

Transcript of Soundtrack

Slides 1-4 (TITLES)

Slide 5

TERMINOLOGY NOTE: The term "unit cell" is used in two different ways, which might be confusing. In the larger sense, the term refers to the complete protamine P1-P2 dimer and its associated DNA. In the smaller sense, it refers to a cross-section of the (larger) unit cell containing four arginine residues plus the four adjacent nucleotide residues to whose phosphate groups they electrostatically bind. (This is followed by instructions for the Flash Player controls).

Slide 6 (Introduction)

Hello. I'm Dr. Ken Biegeleisen. I've believed, for 34 years now, that DNA is non-helical in cells. The experiments which prove this are described in Part I of this series. Since almost no scientists pay any attention to these experiments, I got the idea that maybe I could approach the problem in another way; namely by demonstrating that some of the mysteries of nucleoprotein structure could be solved by starting with the correct DNA structure; that is, a NON-helical model.

Slide 7

Here is a preview of this structure. I'm not going to describe anything in detail yet, but I'll show you the final structure, so you can appreciate its beauty and symmetry.

Shown here is a longitudinal view of the unit cell of the protamine-DNA complex, consisting of the 2 protamine chains P1 & P2, which bind electrostatically to a pair of perfectly-charge-aligned DNA chains. The DNA chains are stretched out to their full lengths, with nearly 7 Å base-pair spacing. They have no helical twist.

Slide 8

When adjacent unit cells come together, the DNA base pairs mutually intercalate, giving rise to the familiar Watson-Crick 3.4Å spacing found in artificially-deproteinized crystals of DNA in the laboratory.

Slide 9

This is the same structure viewed from the top. The arrows indicate that the alternating columns of DNA and protein continue indefinitely in either direction.

Slide 10

If we introduce an adjacent row of DNA and protein columns…

Slide 11

…and displace it by a distance corresponding to ½ the width of a unit cell, we wind up with a 2-dimensional array of salt bridges which is so startlingly perfect that it is impossible to attribute it to mere coincidence.

At this point, if you're a good scientist, you should be filled with a deep skepticism about this structure, but that's because you haven't seen how it was developed. It was a step-by-step logical development, and at each step there really were no acceptable alternatives but this one.

But I'm getting ahead of myself. This project did not start with protamine. It started with histones.

Slide 12 (HISTONE STRUCTURE)

Slide 13

It has never been doubted, by any author on the subject, that the basis of the interaction between nuclear proteins and DNA is the preponderance of positively-charged basic amino acid residues in the protein, and the preponderance of negatively-charged phosphate groups in DNA. Therefore, in any nucleoprotein structure, one would hope to see some sort of alignment between these oppositely-charged groups.

Slide 14

Let us, then, do an analysis of the numbers and locations of basic (*i.e.*, positively-charged) amino acid residues in the Nucleosome Core Particle. This is currently believed to be the fundamental building block of nucleoprotein structure, consisting of the Histone Octamer plus its associated DNA.

The data to be shown are from Luger *et al's* high-resolution x-ray study of reconstituted nucleoprotein, located in the Protein Data Bank at the location shown.

Slide 15

Here are the statistics on the locations of the basic amino acids in the Histone Octamer.

Slide 16

It consists of 4 distinct subunits, now highlighted in the left-hand column in purple. They are called H3, H4, H2A and H2B. Each of these subunits occurs twice in the octamer.

Slide 17

Each subunit contains approximately 20 basic residues, mostly arginine and lysine, with a lesser amount of histidine. The total number in all 4 subunits is 82.

Slide 18

Looking at the right-hand column, we see that some of the basic residues are contained in a region called "Helix II". This is the name appended to a part of each subunit which is in the subunit core, relatively inaccessible to solvent, and therefore less likely to be available for electrostatic binding to DNA. The total number of such basic residues is 9.

Slide 19

If we subtract these 9 residues, we're left with 73 basic residues in external locations, available for binding to DNA.

Slide 20

Now we have to multiply by 2, because each of the 4 subunits occurs *twice* in the Histone Octamer. This gives us a grand total of 146 arginine, lysine and histidine residues available for electrostatic binding to DNA.

Slide 21

Now, according to published studies, the number of DNA base pairs associated with the Histone Octamer in the core particle is…….146! A remarkable coincidence, is it not? Although perhaps *not* so remarkable when we take into consideration the essentially-universally-held belief that the binding between DNA and Histone is electrostatic. So perhaps the equivalence of the number of positive charges on Histone and the number of associated negatively-charged base-pairs of DNA should only be expected.

Slide 22

If so, however, then it ought to be possible to demonstrate some alignment between the opposite charges. Let's take a look at the structures to see whether we can do so.

Slide 23

This is the complete nucleosome core particle, with 146 base-pairs of DNA wrapped around it like a scarf. The movie is an AmiraMol export. The protein has been colored entirely white, and the DNA entirely black. You should clearly see the helical twists of the DNA as the complex spins around.

What holds the DNA in this neat spiral arrangement? It is said that there is a "superhelical groove" on the surface of the protein structure, and that the DNA lies in this groove. If so, the ordinary rules of protein chemistry *ought* to apply to the groove; that is, there ought to be some recognizable chemical force cementing this relationship. The mere existence of a structural "bed" in one chemical does *not* automatically cause a second chemical to lie down in it. There has to be a bond of some sort.

Hydrophobic bonding seems rather unlikely, as Watson-Crick DNA is, in effect, a totally symmetrical cylinder coated with negative charges. The nature of the bond therefore defaults to electrostatic. We might therefore be justified in expecting a special concentration, or special arrangement, of positive charges on the surface of the groove, which attract DNA either by their large number, or by their peculiar spatial arrangement.

Slide 24

Let's therefore look at the spatial arrangement of positive charges on the protein surface. Because the octamer is somewhat flattened, I've changed the view to one which shows more of its surface. I've also removed the DNA. The black color now represents the basic amino acid residues Arg, Lys and His. You will note that they are distributed *all over* the surface of this structure. Can you see any concentration of charges in a superhelical groove? If anything, the charges seem to be concentrated on the extreme periphery, well-removed from the place where DNA binds.

Slide 25

When I realized this, I decided to examine, in more detail, the charge relationships between the positive basic residues in histones, and the negative phosphate groups in DNA.

Here's a literal AmiraMol projection of the core particle, with a new coloring. Now the DNA is colored entirely green, except for the negatively-charged phosphate groups, which are colored blue. The protein is now entirely yellow, except for the positively-charged extremities of the Arg, Lys and His residues, which are colored red. There's certainly no obvious alignment of opposite charges evident in this view. Let's move in and take a closer look.

Slide 26

What I'm about to do, I've done in many randomly-selected parts of the superhelical groove. This is a flat projection, so you can't tell how far apart the atoms are in 3-dimensional space. And what I'm going to do is to take measurements, employing AmiraMol's measuring tool. Here are the results:

Slide 27

Like a hydrogen bond, a good electrostatic bond will be about 3Å in length. The charge attraction drops exponentially as the distance increases beyond this. I emphasize that this section of the protein-DNA complex was chosen entirely at random. There are no true ionic bonds here. We have, at the upper left, what appears to be a good bond in this 2-D view, but the distance between the atoms, in 3-dimensional space, is actually over 31Å. The one below it is nearly 11Å, and the one below that 16.7Å. Finally, at the bottom, we have a few apparently coincidentally-close contacts, each, however, *greater than* 3Å. In short, there are no meaningful ionic bonds anywhere in this section.

I reiterate that I have performed this exercise in numerous randomly-selected regions of the superhelical groove, and my conclusion is that there are no significant ionic bonds anywhere in this structure.

Slide 28

If you're a good scientist, you'll be skeptical, and you certainly won't want to take my word for it. But we now have the miracle of internet technology, and if you have good virtual chemistry software installed on your computer, you can click the Protein Data Bank link shown here, and instantly be taken to the original structure file. Then you can do your own measurements, wherein you'll see for yourself that there are no real ionic bonds in this structure. *Why is nuclear protein positively-charged, and DNA*

negatively-charged, if there are no charge-charge interactions in the nucleoprotein structure? It makes no sense whatsoever.

Slide 29

Here's a final look at the nucleosome core particle with its DNA. This time, each subunit is colored differently. It's really quite lovely to look at, and I'm sure the structure, assembled outside its normal physiological environment, in an artificial laboratory setting, is nevertheless laden with significance. But I believe that it needs adjustment, if there is to be some semblance of alignment between the opposite charges. The amount of adjustment necessary, to bring about a better charge alignment, might be small, or it might be large – there's no way of knowing up front.

Slide 30

I intended to undertake the above-referenced "adjustment", beginning with a revision of the DNA structure, to remove the helical twist, in accordance with the experimental results reported in Part I of this series. The trouble is, the Histone Octamer has over 10,000 atoms of protein, and the associated DNA has at least another 10,000 atoms. Where do you start revising such a structure? What scientific criteria can you invoke to judge whether a revision has any biological validity at all? There are none!

Therefore, I sought a simpler system. I decided to revise the structure of the protamine-DNA complex instead. This proved to be a most fortunate choice, for the mystery of the protamine-DNA structure proved to have only a single solution, and that solution has provided a type of virtually incontrovertible proof that DNA is non-helical inside living cells, as we shall now show.

Slide 31 — The Structure of Human Protamine

Slide 32

My first stop on the road to protamine structure was the Protein Data Bank. I logged into their home page, and typed in "PROTAMINE". Nothing came up! I checked the spelling, and tried again, but still nothing came up. I tried entering the names of the subunits, P1 & P2, but still drew a blank. Hmmmmm…that's odd, I thought.

Slide 33

Next I tried Internet search engines. The search term "protamine" yielded mostly articles on protamine sulfate, a sperm protamine derivative used by physicians to reverse the anticoagulant effects of heparin. The search phrase "protamine structure" yielded mostly articles on primary structure. I won't bore you with the list of alternative search terms I employed.

Slide 34

The long and the short of it is that there were no published molecular models for protamine! It would appear that its structure was a complete mystery!

Slide 35

It was thus the PERFECT SUBJECT protein to test my theory that our knowledge and understanding of nucleoprotein structure would be vastly increased by starting with a non-helical model for DNA.

Slide 36 — Analysis of Protamine P1-P2 amino acid sequences

Slide 37

Here's the amino acid sequence of human protamine. There are two chains, called P1 and P2. Let's look at the residue count, and then remove the number scale. P1 has 51 residues and P2 has 57 residues.

Slide 38

There; that's a little easier on the eye. Now, for those of you who don't remember your protein conventions, the N-terminal end is always on the left, and the C-terminal on the right. This is not indicated in the slide, so just keep it in mind. I won't review the entire 1-letter amino acid alphabet, but you'll need to keep certain letters in mind. Chief among those is Arginine, represented by the letter "R". As you can perhaps see at a glance, there are a lot of "R"s in this structure.

Slide 39

Of lesser importance in Protamine are the basic amino acids Lysine, indicated by the letter "K"… and Histidine, by the letter "H". These latter basic residues play a more important role in Histone than they do here, where they occur far less frequently than Arginine.

Slide 40

The letter "C" stands for Cysteine. Certain of the Cysteine residues seem remarkably well-aligned with respect to the possibility of disulfide bonds, which are suggestively drawn in for emphasis.

Slide 41

Let us now examine the entire amino acid sequence by class. The first and foremost class is that comprising the basic amino acid residues, each of which carries a positive charge. Without a doubt, the teleological purpose of these positive charges is for binding to DNA. The overwhelmingly most frequently-occurring of these is Arginine, now colored red. Arginine accounts for *half* the residues in Protamine.

Let's bring in a little picture of Arginine. It's decidedly the *longest* of all the amino acids, with an R-group side chain of 6 atom's length. Let's also add the positive charge to the guanidinium group at the end of the side-chain.

Slide 42

Looking again at our amino acid sequence, we next bring in Lysine and Histidine, both now colored blue. Here's a picture of Lysine coming in from the left. Note that it's side-chain is one atom *shorter* than that of Arginine, and that there are far fewer Lysines in the structure than Arginines. It would appear that the Protamine design favors the

longest, most positively-charged R-groups. Histidine, shown at the bottom, is shorter still.

Slide 43

Let's remove the pictures, and color the Lysine and Histidine residues red.

Slide 44

With all the basic amino acid residues colored red, it's easy to see that this is an *extremely* positively-charged molecule. There are only a few amino acids left. Let's look at *them* now.

Slide 45

First we'll color the hydrophilic residues blue. There are 33 such residues, each one being either Tyrosine, Serine, Glutamine, Threonine or Cysteine. Except for Cysteine, whose R-group terminates in a sulfhydral group, all the others terminate in hydroxyl groups. In other words, they are all capable of hydrogen-bonding with water, or, more likely, with the phosphate groups of DNA.

Slide 46

OK, let's color *them* red also.

Slide 47

Now we can clearly see the extremely hydrophilic nature of this protein. All that remains are a pair of Methionines, a pair of Alanines, a pair of Prolines, 3 Glycines, and a single occurrence each of Leucine and Isoleucine, these latter 2 being the only seriously hydrophobic side groups in the entire 108-amino-acid structure. Protamine thus has an intensely hydrophilic nature, with a powerful positive charge.

Slide 48 (SUMMARY)

To summarize, protamine is therefore essentially a long string of Arginine residues, interspersed, for the most part, with hydrophilic residues capable of donating a proton for a hydrogen bond.

Concerning the order in which any of these residues occur in the amino acid sequence, no logical pattern can be clearly discerned, either in human Protamine, shown here, or in the Protamine amino acid sequences of many different species which have been studied.

("As a philosophical aside"), it would appear that nature is demonstrating here a sort of tolerance which might be a good role model of tolerance for human beings generally: where additional Arginine residues are not needed (which, if present, would presumably result in a biologically inactive, quasi-crystalline structure), any of the above-referenced hydrophilic side-chains will suffice, in any order, and even an occasional hydrophobic group will be tolerated.

But what is the structure of Protamine?

Slide 49 (3 CLASSES OF PROTEIN STRUCTURE)

In elementary chemistry, we learn that protein structures are generally divided into 3 classes: Globular, alpha helix and beta sheet. Let's start with the first category, "globular proteins".

Slide 50

Is it possible that Protamine has a globular structure?

Slide 51

The term "globular" encompasses a huge variety of different substances, 3 examples of which are pictured in the top of the slide. On the upper left, we have alkaline phosphatase, a liver enzyme. In the middle, we have hemoglobin, the well-known oxygen-carrying protein which imparts to blood its red color. On the right we have interleukin, a protein involved in the immune response.

Although the shapes of these three substances are quite different, they all have one thing in common: Their cores are composed mainly of amino acids having non-polar aliphatic and aromatic side chains, whereas their surface side-chains are mainly charged and hydrophilic.

Can Protamine have this structure?

Slide 52

Not very likely! Fewer than 10% of the residues comprising the chains of Protamine are non-polar, and 3 of those residues are Glycine, which has no R-group at all. In short, Protamine simply doesn't have enough non-polar residues to form a core stabilized by hydrophobic bonds.

In nature, one can find globular protein cores which are stabilized by *salt bridges*, but human protamine has absolutely *zero* acidic, negatively-charged residues, wherewith to neutralize the great preponderance of basic residues in its primary structure. Even in species whose protamines have acidic residues, the most that will be found is one or two per protamine dimer.

I believe, therefore, that I'm standing on very firm ground, when I say that a globular structure for Protamine is ruled out.

Slide 53

Is it possible that Protamine is an a-helix?

Slide 54 (Movie of alpha-helix backbone)

Before proceeding, I must comment on the perspective of these rotating virtual molecular models. They *have* no perspective, that is, the atoms and bonds do not get smaller when they move to the background. Consequently, your eye may perceive them to be rotating to either the right *or* to the left, and the perceived direction of rotation may seem to change from time-to-time. Often, but not always, the molecule will look essentially the same regardless of the perceived direction of rotation. But in the case of a helix, if your eye sees it rotating to the right, it will appear to be a right-handed helix, whereas if your eye tells you it's rotating to the left, likewise the direction of perceived helical twisting will be to the left.

Slide 55

The situation is quite analogous to the old perspective problem with the 2-dimensional representation of a cube. I'm sure you've all seen this picture before — if you stare at it long enough, the letter "A" will seem to move back and forth from the front of the picture to the back. There is no "correct" interpretation — it all depends on your perspective.

Slide 56

As a general rule, with, I think, only a single exception which we've already passed, the correct direction of rotation, in all the movies to come, is to the *right.* If your eye is telling you otherwise, then keep staring, and — hopefully! —the direction will eventually right itself.

Slide 57

Now, back to our question: *Is it possible that Protamine is an a-helix?*

Slide 58

I can answer that question up front, with a fairly emphatic "no". Polypeptides with bulky side chains, and lots of electrically-charged side-chains, do *not* readily form alpha-helices. Protamine is full of the most bulky and most-highly-charged side-chains found in nature. *However*, since there are a large number of negatively-charged DNA phosphate groups available to neutralize the positive charges on Protamine, we shall be obsessive and compulsive, and rule no structure out without fully considering it.

The empty alpha-helix shown here was borrowed from another molecule, which accounts for the non-alpha-helical tail at the C-terminal end. We have stripped off the original amino acids, leaving only the polypeptide backbone.

Slide 59

Now we've added Arginine R-groups at each position.

Slide 60

To make things easier to grasp, let's give the Arginine side-chains an accent-color.

Slide 61

Here's the structure from the top. Essentially what we've created is a cylinder, coated on all sides by positive charges. Remember that DNA too is a cylinder, only coated by *negative* charges. If the DNA is to uphold an alpha-helical structure in a polypeptide whose primary amino acid sequence is highly-*unfavorable* for the alpha-helix, I would assert, with a high degree of confidence, that there would *have to be* good alignment between the positive and negative charges. Without that alignment, what's to keep the poly-arginine alpha-helix, an energetically *unfavorable* structure, from instantly unwinding?

Slide 62

This is an issue which cannot be dealt with in an entirely satisfactory way in the setting of a slide show. Nevertheless, here's a scale model of the Watson-Crick structure, dancing about adjacent to a hypothetical poly-arginine alpha-helix. If you've got a good eye for virtual structures, you can perhaps see that there is essentially *no chance* of attaining any meaningful alignment between the oppositely-charged atoms. The pitches don't match, the winding directions are contrary, and the cylindrical nature of each results in only a tangential contact between them, at best.

Slide 63 (Dancing helix joke)

Since no real alignment can be established here, the only way to reconcile the irreconcilable differences in the distributions of positive and negative charges in helical Protamine and DNA, is to simply close your eyes to them, and pretend that these various molecules just coexist, constantly moving about in a random and meaningless manner.

I think you'll agree with me that this is an act of intellectual desperation, and is too high a price to pay, if the only reward is to be able to keep the concept of a helical winding alive, in a setting in which it just doesn't work.

Slide 64

I should add that I have spent a goodly amount of time in futile attempts to anastomose the charges of these two molecules, and I have concluded that it's impossible. Although my efforts can perhaps be dismissed as inconclusive, the best evidence against the alpha-helix as a structure for Protamine is that in 50 years, no one else has succeeded in making it work either.

Slide 65

Before moving on, a general observation is in order: You *cannot* solve the Protamine structure problem by simply wrapping Protamine around DNA, as has been suggested to me by a prominent nucleic acid / protein scientist who should have known better.

Slide 66

Remember that DNA, *if* it has the W-C structure, is essentially a cylinder coated with negative charges, indicated here by the white minus signs.

If you bring two such cylinders too close together, the negative charges will mutually repel each other, and the cylinders will be driven apart. This will never do in a cell nucleus, where DNA is packed in to a degree which staggers the imagination.

Slide 67

If you simply wrap Protamine around DNA, all you get is a larger cylinder, now coated with *positive* charges! This scheme is fraught with difficulty. First of all, these bloated hypothetical structures will *still* mutually repel each other, forcing the cylinders apart. Secondly, there is no possibility of any further higher-order structure, because you can't then go and wrap more DNA around the already-bloated, Protamine-coated cylinders, since the DNA double-helix is a rigid rod.

In short, the desperate theory that you can preserve the W-C structure in sperm heads, by wrapping Protamine around the DNA, is an exercise in futility.

Slide 68

So, what's the structure of Protamine? Although some will stubbornly deny, I think we have all-but-eliminated a globular structure, as well as an alpha-helical structure. What remains?

At this point, it might be instructive to turn to the master of mystery, Mr. Sherlock Holmes himself. On more than one occasion, he said to Dr. Watson words such as:

"How often have I said to you that when you have eliminated the impossible, whatever remains, *however improbable*, must be the truth?"

This particular form of the quote is from "Sign of the Four". So? What remains?

Slide 69

The only standard protein structure remaining is the beta-sheet. And when we begin to consider the beta-sheet, everything turns around. All of a sudden, in the place of the awkwardness and absurdity we've been grappling with thus far, we at last begin to see logic, beauty, and harmony, as the puzzle pieces finally start to fall into place — PROVIDING that we are willing to admit to the possibility that DNA might not have a helical twist in every setting in which it's found.

So, without further ado, let us begin to examine the structure of the Protamine-DNA complex, a structure which must have implications for DNA structure in many other settings.

Slide 70

This is a model of a fully-extended beta-sheet. It's a fairly flat structure, and is not often found in nature in this simple form. In real life, it bends and twists in accordance with the phi and psi angles, which terms you will hopefully remember from Biochemistry 101.

As you will undoubtedly recall, each amino acid residue is somewhat arbitrarily defined as starting with the peptide nitrogen, proceeding to the alpha-carbon, to which the R-group, or "side-group" is attached, and ending with the carboxyl carbon.

The peptide bonds are all approximately flat, because of their double-bonded nature.

Slide 71

Thus, the only bonds which readily rotate are those associated with the alpha carbon. The proximal bond, connecting the alpha-carbon to the peptide nitrogen atom, allows an axle-like rotation, associated with a dihedral angle called "phi", and the distal bond, connecting the alpha-carbon to the carboxyl carbon, also rotates, resulting in a second dihedral angle called "psi". The values of phi and psi largely determine the structure of the beta sheet.

In the fully-extended sheet, as shown here, both phi and psi are 180 degrees, resulting in a flat peptide backbone, the only deviation from flatness being the R-groups, which, as we shall see momentarily, project both above and below the plane of the sheet.

The most important thing to note on this slide is the ***distance between adjacent r-groups on either side of the sheet, which is* 7.3 Å.** In case the significance of this is not immediately evident, let me point out that this is very close to the distance between adjacent phosphate groups in the DNA sugar-phosphate backbone, when measured along that backbone. This means that when we remove the helical twist from the DNA, we are suddenly faced with the possibility of having a perfect alignment between the opposite charges on DNA and protein.

Slide 72

In this movie you will be able to see the flatness of the fully-extended beta-sheet peptide backbone. Only the beta-carbon atoms of the side-groups, represented by the green bonds, project above and below the plane.

Slide 73

As we have said, real beta-sheets are generally not flat. Watch what happens to the length of this segment of polypeptide backbone when the phi and psi angles are simultaneously altered; to (-) and (+) 130.5°, respectively. Keep your eye on the lower, or C-terminal end, marked by the arrow.

Do you see the shortening?

Slide 74

Here's a movie of the new structure. It certainly has interesting symmetries, but it's no longer flat.

Slide 75

Let's look at the new measurements. Most significantly, by changing the phi and psi angles to (–) and (+) 130.5 degrees, we've shortened the distance between R groups on either side of the beta sheet to 6.8 Å. This is exactly the distance between two Watson-Crick DNA residues, when measured along the sugar-phosphate backbone. Can you guess where we're going with this?

Slide 76

In case you're thinking that the phi and psi angles of (-) and (+) 130.5 degrees have been arbitrarily chosen to force our model to anastomose to DNA, think again. Remember the Ramachandran Plot?

The Ramachandran Plot shows, by color coding, the ranges of phi and psi values which result in stable polypeptide structures. The white portions of the plots are the most stable, followed by the darker brown regions.

Our beta-sheet is marked by a blue star.

Slide 77

With a phi value of –130.5 degrees…

Slide 78

…and a psi value of +130.5 degrees, we find ourselves in the very heart of the most-favored region of the Ramachandran plot for beta-sheets.

Slide 79

By way of comparison…

Slide 80

…note the Ramachandran position of our original fully-extended, flat beta-sheet, marked now by the red star. It's in the dark brown region, meaning it's a possible conformation, but not a highly-favored one.

Slide 81

Our 130.5 degree model, however, is about as favorable as a polypeptide chain backbone can possibly be, and it just happens to conform to the dimensions of DNA, when untwisted. Coincidence?

Slide 82

Before bringing in the DNA and completing our model, we should finish perfecting the protein component.

Protamine has two chains, P1 and P2. There are two orientation questions which arise. The first relates to longitudinal chain direction, and the second to axial rotation.

Slide 83

The first question is this: Are the chains parallel, meaning the N-terminal-to-C-terminal direction is the same for both chains? Or are they anti-parallel?

The second question relates to rotation. Are they "cis", by which I mean that they have the same axial rotation, so that the R-groups from both chains always point the same way, or is one chain rotated 180 degrees with respect to the other?

The answers to these questions become evident after a consideration of the cysteine residues of the two chains.

Slide 84

Any Protamine structure must take into account the obvious existence of disulfide bonds. When cysteine residues line up with the arithmetic perfection seen in human protamine, it would be foolhardy to *not* presume, as a first assumption at least, that they are there for the purpose of forming disulfide bonds.

The near-perfect alignment of cysteine residues is not a coincidence of human Protamine primary structure. It is confirmed by examination of the primary sequences of other protamines. They are all about half Arginine, and they all display a striking alignment of the cysteine residues.

Slide 85 (Interspecies comparison of Protamine P1-P2 amino acid sequences)

Slide 86

Here are the protamines of 4 mammalian species.

Slide 87

At the top of the slide is human protamine, which we have been examining since the beginning of this presentation. Note the near-perfect alignment in the distribution of

cysteine residues between the two strands. There are 23 and 9 amino acid residues, respectively, separating the first 3 cysteine residues in the P1 strand, and exactly the same cysteine spacing in P2. This can't *possibly* be mere coincidence, can it?

Slide 88

This is the protamine of the rat. The primary amino acid sequence is markedly different from that of the human, and yet the general character of the protein is the same. It's about half arginine, and the cysteine residues are lined up with striking precision.

Slide 89

Likewise for the mouse.

Slide 90

The cysteine residues of the stallion are a little less-perfectly aligned than those of the other species, but still clearly suggestive of the likelihood of disulfide bond formation.

Slide 91

The P1-P2 structure must be parallel, because an anti-parallel structure interferes with the formation of disulfide bonds.

In order to really understand this, you've got to work with virtual modeling software. There you learn that in an *antiparallel* arrangement, the process of approximating the cysteine residues of the P1 & P2 strands sufficiently to form a sulfhydral bond is extremely disruptive of the secondary protein structure; so much so that nearly half the beta-sheet architecture is destroyed.

This would be difficult to demonstrate in a slide show.

Slide 92

An easier way to illustrate the problems which arise when the strands are antiparallel is to simply turn one of the amino acid sequences upside-down. Here we've done that for human protamine. On top of the slide is the normal arrangement. On the bottom, the P1 strand has been turned upside-down. If you look at the C-terminal Histidine, you see that it's now on the left, so that the P1 strand has indeed been rendered antiparallel by this simple maneuver.

With respect to cysteine residues, there are a grand total of 3 possible alignments in the antiparallel orientation. That is, if we slide the inverted P1 sequence back-and-forth, we find only 3 positions where more than one disulfide bond is possible. This is the first. Note that there are only 2 possible disulfide bonds in this position.

Slide 93

This is the second alignment. Again there are only 2 possible disulfide bonds.

Slide 94

And here's the third and last. Again only 2 disulfide bonds can be formed. Since, as demonstrated on the top of the slide, we can get at least 3, and maybe 4 good disulfide bonds in the parallel orientation, I would opine that for all the reasons shown, plus those I

can't show you without virtual modeling software, the P1 & P2 chains are in a *parallel* orientation.

Slide 95

Here's what the P1-P2 dimer looks like in the vicinity of a disulfide bond, depicted in yellow. These bonds have a mildly disruptive effect on the local beta-sheet structure, but I think you'll agree, it's hardly noticeable on visual inspection.

Slide 96

Let's add some Arginine residues. Although the P1 and P2 chains are parallel, we have arranged the Arginines to be maximally-extended, in opposite directions. What you're now looking at may be regarded as sort of a "unit cell" of protamine structure. It turns out that the spacing of the guanidinium groups of fully-extended P1 and P2 Arginine residues, when separated by the length of a disulfide bond, is the same as the cross-duplex spacing of DNA phosphate groups. Another coincidence? Or is it design?

I'll show you the exact measurements in a moment. But first, let's assemble all the data we've accumulated thus far, and present an overview of the probable structure of the Protamine-DNA complex.

Slide 97 (Basic Features of Protamine-DNA Complex)

Slide 98

This is a highly-schematic representation of the unit cell of the protamine-DNA complex. In the center are the polypeptide chains, P1 & P2, in a parallel orientation, with arginine residues projecting in both directions. The P1-P2 dimer is held together by disulfide bonds, which, however, are not indicated in this simplified drawing.

Flanking the protamine on either side is a DNA duplex, stretched out to remove the helical twist. For simplicity's sake, the bases are not shown here. Note the good alignment between positively-charged Arginine guanidinium groups in protamine, and negatively-charged phosphate groups in DNA.

If, perchance, the protamine beta-sheets are fully-extended, with the phi and psi angles both 180°, then the DNA backbone phosphate spacing will be about 7.3 Å. As I shall show you shortly, this will, in the final structure, result in base-pair stacking at half that distance, or about 3.7 Å, which is the base-pair spacing found in left-handed Z-DNA. Neither of these forms – that is, neither the fully-extended protein beta-sheet or the left-handed form of DNA – are favored under physiological conditions, but both the protein and DNA backbones can be reduced in length by altering backbone dihedral angles.

We have already seen this for the beta-sheet, where altering the psi and phi angles to (+) and (-) 130.5° respectively, reduced the R-group spacing to 6.8Å on either side of the beta-sheet. This, as we shall shortly see, results ultimately in base-pair stacking at 3.4 Å, the spacing found in "traditional" B-DNA, as shown in this highly schematic movie:

Slide 99

(Movie - no audio)

Slide 100

Now here's a more realistic representation of the protamine-DNA unit cell. This is, in fact, a literal AmiraMol projection, and this time the base-pairs *have* been included. We see the P1-P2 protamine dimer on the right side, now including a disulfide bond, but this time, for graphic clarity, only one of the two associated DNA chains is depicted.

Two of the four ionic bonds between protamine Arginine and DNA phosphate are shown; I'll flash the two others for a second – there they are – then remove them to avoid any more cluttering in the drawing than is necessary to get the point across.

Now, if you've perchance been marveling about the fortuitousness of the alignment between the charged groups in protamine and DNA, as I certainly did when I first realized it, you may still be wondering what we're going to do about the excessive spacing between base-pairs, which is certainly unnatural, and is never seen in any known DNA structure. The answer is, we're going to bring in an adjacent unit cell, and let the base pairs mutually intercalate:

Slide 101

This restores the base-pair spacing to that of known DNA models. If you're not familiar with this sort of mutually intercalated structure, you may be thinking that I made it up. I didn't. These sorts of intercalated structures are actually quite well-known, and I'll be showing you a very real, and very closely-studied example shortly.

Slide 102

If, perchance, the psi and phi angles are 180 degrees each, resulting in a fully-extended beta-sheet, then the DNA base spacing will be 3.65 Å, almost identical to the base spacing in Alexander Rich's Z-DNA structure.

Slide 103

If, however, we adjust the psi and phi angles in the beta-sheet to (+) and (-) 130.5°, and adjust the dihedral angles of the DNA backbone accordingly…

Slide 104

…then we wind up with the "traditional" DNA base-spacing of 3.4Å, the spacing found in B-DNA.

Slide 105

Throughout this presentation so far, I've been hinting that the dimensions of the beta-sheet align naturally with those of DNA. Let's go back, for a moment, to the "classic" Watson-Crick structure, and see where those spacings lie.

Slide 106

As far as the world is concerned, this is the structure of DNA. People have been mesmerized by it for over 50 years. The trouble is, it just doesn't work with Protamine. But the critical dimensions we need are contained within it.

Slide 107

The residue spacing which is ingrained into the minds of young scientists, from high school if not earlier, is the base-pair spacing of 3.4 Å. This, however, is not the true distance between residues, but merely the apparent distance when measured in helical axial space, which is twisted, and therefore distorted.

Slide 108

The true distance between residues is the distance between phosphate groups, when measured along the sugar-phosphate backbone. That distance is always about 7 Å, plus or minus a few tenths of an angstrom. I've looked at numerous structures, including the classic "B" structure, the left-handed "Z" structure, the Rodley-Sasisekharan Side-By-Side structures, the Delmonte Paranaemic structure, and the Wu "Straight Ladder", and this distance is about 7 Å in each one.

Slide 109

Another important dimension in DNA, but one which is *not* frequently mentioned, is the cross-duplex phosphate-to-phosphate distance, which is also fairly constant from structure-to-structure, and this distance is always about 20 Å. These two measurements, namely 7 Å and 20 Å, provide a perfect match for the spacings in the Protamine beta-sheet. Let's see how.

Slide 110

Here's the beta-sheet again.

Slide 111

I've already pointed out that the distance between residues on either side of the sheet is about 7Å. Obviously, this 7 Å spacing in the peptide backbone is an excellent fit for the 7 Å spacing in the DNA sugar-phosphate backbone, *provided that the helical twist is removed.* We thus learn that the failure of the scientific establishment to come up with a protamine-DNA structure in over a half-century is surely because of its stubborn insistence on a helical structure for DNA, which simply doesn't work here.

We also learn that each protamine strand will, in all probability, interact with *two* DNA strands, since the 7 Å spacing is present on *both sides* of the structure.

What about the 20 Å cross-helix dimension of DNA? How does that work with protamine?

Slide 112

Let's return to this picture, which I showed previously. It's a short segment of the beta-sheet containing a disulfide bond and a pair of Arginine residues, maximally extended, but in opposite directions.

Slide 113

The lengths of two Arginine residues, when separated by the length of a standard disulfide bond, is 14 Å. Since we're looking for a spacing of 20Å, you may think we're a bit short, but actually we're right on target. First of all, let's tilt the structure forward, and look down upon it from the top:

Slide 114

Here's the same picture, viewed in the axial direction, from the top down.

Slide 115

Now let's add a DNA duplex. The guanidinium-to-guanidinium cross-sheet distance in Protamine is 14Å, and the phosphate-to-phosphate cross-duplex distance in DNA is 20 Å.

That leaves a space of something like 3 Å, which is perfect for a salt bridge, that is, an ionic bond. Like a hydrogen bond, the best spacing for an ionic bond is 3 Å. Unlike a hydrogen bond, it is non-directional, dependent only upon the distance between the atomic nuclei. If this distance is shorter than 3 Å, the Van der Waals radii start to get violated, and if it is greater, the electrostatic attraction between the members of the bond starts to decrease exponentially.

Since Nature has left us a convenient 3 Å space for an ionic bond, we may conclude that the 14 & 20 angstrom spacings, in the protein and DNA respectively, are *indeed* ideal spacings for our protamine-DNA complex.

Slide 116

Remember that in the beta-sheet, Arginine residues project from both sides, so to complete the picture, we need to add a second DNA duplex. Here's the axial view of the complete unit cell. By the way, if the DNA bases look improperly paired, that's because they're *not* paired in this particular view. Because the strands of the DNA duplex are antiparallel, the phosphate groups at any given level are *not* associated with the same base pair. One is associated with a base above the plane, and the other is associated with a base *below* the plane.

This is a literal AmiraMol projection of our structure. If you look carefully at the bases on top of the drawing, you'll see that the base on the right is smaller, because it is several angstroms farther away from the observer than the one on the left. Their associated phosphate groups, however, are at about the same level.

Slide 117

Here again is our greatly-oversimplified drawing showing the essential features of this unit cell when viewed from the front. Again, for graphic clarity, only the DNA sugar-phosphate backbone is shown; the bases are excluded from the drawing.

Slide 118

Although, being afflicted with human frailty, I cannot help being proud of my solution to the 50-year-old protamine-DNA structure mystery, the truth of the matter is that in addition to the fact that I'm hardly the discoverer of the beta-sheet, neither did I deduce the DNA structure in the complex. The structure was worked out years ago by the brilliant Northwestern University immunologist, Tai Te Wu.

Wu's biographical information, and the details of his experiments which prove that the strands of plasmid chromosomes are not topologically intertwined — that is, that plasmid chromosomes do NOT have the Watson-Crick structure — are in Part I of this series, entitled "The Science and History of Topologically Non-Linked DNA".

In the current presentation, however, we shall focus only on the structural details of the Wu model, since it provided the inspiration for the protamine-DNA complex structure I shall propose.

Slide 119

Now I'm going to read through some sections of Wu's paper on 4-stranded DNA structure, from PNAS. If you've seen part I of this series, you can use the space bar to skip through these slides.

Slide 120

The paper is called "Secondary Structures of DNA", and it was published long ago, in 1969.

Slide 121

In this paper, Tai Te Wu discussed the pitch**:**diameter ratio for the Watson-Crick double helix at 66% and 92% humidity, as revealed by x-ray crystallography. He stated:

Slide 122

"At 92 percent, the surrounding of the DNA fiber resembles that inside a cell, while at 66 percent, the state of the fiber becomes completely artificial. The differences…should then provide the necessary clue for us to resolve the intricate secondary structure of DNA *in vivo.*"

Slide 123

After a thorough analysis of the x-ray diffraction patterns of DNA at the two humidities, Wu concluded that "if the structure of the DNA fiber at 66 percent consists of a double helix, than at 92 per cent [it] must consist of two double helices".

Slide 124

In the Wu structure, the distance between base pairs in either double helix was twice as large (*i.e.*, about 6.8 angstroms) as in the Watson-Crick structure, but the stacking of bases at 3.4 angstroms was preserved by mutual intercalation of the base pairs of the two duplexes.

Slide 125

Finally, he predicted that at 100% humidity, *i.e.*, the condition prevailing in living cells, the two mutually-intercalated duplexes would lose all vestiges of helical twist and exist as a pair of mutually-intercalated "straight ladders", each one having its base pairs spaced at about 7 angstroms.

Slide 126

This is Wu's own highly schematic diagram of his proposed structure for circular chromosomes. I found it confusing when I first looked at it, so I'll try to walk you through it quickly. The single-stranded, circular, antiparallel sugar-phosphate backbones for the entire chromosome are represented by the elongated circles on the sides.

Slide 127

I should note that although Wu presumes a superhelical twist will be found in purified, protein-free circular DNA in the laboratory, the structure shown here is Wu's conception of DNA at 100% humidity, *i.e.*, in the natural intracellular environment, where he believes DNA to have a "straight ladder" structure with no helical twist at all.

The rectangular shapes attached to the backbones are not individual bases, but base-*pairs*, the small number shown obviously schematically representing a much larger number than can be drawn in a picture like this. In order to clarify the drawing, the base-pairs in the back have been arbitrarily shaded, and the base-pairs in the front have been left white. This hopefully makes the mutual intercalation of base-pairs easier to see.

The longitudinal backbone distance between adjacent base pairs is 6.8 Å, which, after the intercalation, gives the more familiar 3.4 Å spacing expected for DNA.

Slide 128

Many scientists, who are not familiar with the rather large variety of DNA structures already known, may find the idea of adjacent duplexes with mutually-intercalated bases to be arbitrary, and perhaps even unbelievable, but it's actually *very* believable, and is in fact known to exist in well-characterized structures which have been reported. The best-known of these is the Gehring tetramer.

Slide 129

Here's a top view. It's pretty to look at.

Slide 130

And here are the details. The DNA sequence was peculiar, and the pH was low, but the point is that under the right circumstances, DNA will in fact form a tetramer consisting of two intercalated duplexes. In the case of the Gehring tetramer, the "right circumstances" are found in the base sequence, and in the low pH. In the case of the protamine-DNA complex, the predisposing circumstances are a multitude of extremely favorable ionic interactions and hydrogen bonds.

Slide 131

And this is what it actually looks like. Here's the first strand.

Slide 132

Here's the second strand...

Slide 133

…and the 3rd strand — it's starting to get a bit cluttered-looking…

Slide 134

…and finally, the 4th strand.

Slide 135

Let's spin it around and see what it really looks like. If you're familiar with the Watson-Crick double helix, you'll immediately notice that the Gehring structure, in spite of its 4 strands, is really quite similar.

Being that it is linear, and protein-free, it forms a helix, as all DNA does in the absence of topological constraints. Even though it has twice as many strands as normal DNA, the amount of DNA per unit length remains about the same as that of the classic double-helix, because here the strands are stretched out to about twice their normal length. The general layout of the structure is also very similar in *appearance* to the Watson-Crick structure, and if you looked at it out of the corner of your eye, you might not even notice the extra 2 strands.

Slide 136

We are now ready to view the Wu structure, which I'm quite certain is the structure of DNA in sperm cells, and undoubtedly has important applications elsewhere. Wu did not have the benefit of virtual modeling software at the time that his work was done, so this is the first time it has ever been presented visually in 3-dimensional format.

Slide 137

Here's the top view. It's more rectilinear than the Gehring tetramer, but it has a nice symmetry.

Slide 138

Here's the first of the two DNA duplexes. It's just normal Watson-Crick DNA, untwisted and stretched to twice the length seen in synthetic laboratory crystals.

Slide 139

This is a movie of the structure, showing its features.

Slide 140

Now watch carefully. I'm going to remove the first duplex, and replace it with the second duplex in its correct position relative to the first.

Slide 141

Here's the second duplex. If you have a good eye for subtleties like 3' → 5' polarity, you'll notice that it's rotated 180°. You'll see why shortly.

Slide 142

Here are all 4 strands, with the base-pairs mutually intercalated. It's very similar to the Gehring tetramer, except that the helical twist is removed.

Slide 143

Here's what it looks like in 3D. I consider this to be the most significant slide in this show, because I believe that what you're looking at right now is an important structure for DNA in life; perhaps *the most* important such structure. In accordance with this, I've inserted curved arrows to remind you that if you're eye is playing that

perspective trick on you, making the structure appear to be rotating to the left, then you're seeing it backwards, and not as it really is.

However your eye perceives it, however, this structure has order, symmetry and even aesthetic beauty. If, perchance, it is deemed less mesmerizing than the Watson-Crick helix, so be it. If you need to see spinning helices, take a trip to the local barber shop, but leave helical structures out of replicating circular chromosomes. Please.

Slide 144

Let's look at the exact dimensions of the protamine-DNA structure, then we'll assemble the whole thing and show it to you.

Slide 145

By way of an introduction, let's take a look at "classic" double-helical DNA, because our structure is not really so different. Here's a dinucleotide in the W-C conformation.

Slide 146

When you rotate it, it looks like this.

Slide 147

Now here's *our* structure. In general, our bond lengths and angles do not differ significantly from those of the Watson-Crick structure, and the changes mainly involve alterations of the backbone dihedral angles.

Slide 148

Here's our structure in 3-D.

Slide 149

This is a literal AmiraMol projection, showing our sugar-phosphate backbone dihedral angles. There is no hint of steric hindrance in this backbone, and therefore no energy barrier to its formation.

The backbone is completely straight, and the bases are at perfect right angles. Neither of these things needs to be quite as rigidly rectilinear as we have made them — after all, the sperm cell, within which this substance is found, is the very antithesis of rectilinearity. But we did it this way because the methodology involved in modeling the virtual structure mandates a very regular, repeating pattern from residue-to-residue.

Slide 150

The chi angle, which determines the positions of the bases, is about -100°, which places it in a fairly energetically-favorable portion of the *anti* region of the energy vs. chi curve.

Slide 151

How were these dihedral angles determined? The answer is "through laborious trial and error". Going into this task, I naively imagined that it would be easy. It was not. It was a nightmare.

These days, professional virtual modelers, which I am *not*, employ computer programs which analyze millions of structures, and report back which ones are the best.

Slide 152

It's a lot like computer chess. The computer just looks at the current position, then stupidly goes through a few million alternative moves until it finds one which matches a winning game stored in its memory.

Slide 153

With DNA backbones, the process is not very different. The computer, armed with certain criteria, will stupidly check a few million dihedral angle combinations, until it finds one it "likes". Or, perhaps I should say, until it finds one that *you* like.

To give you an idea of the magnitude of this problem, consider the fact that there are six dihedral angles, alpha-to-zeta, and even though one of them, delta, is relatively fixed by the fairly rigid constraints imposed by the ribose structure, the other five, even if we assume only 10 possible values for each of their dihedral angles, give rise to 10^5, or 100,000 structures.

Slide 154

But, as this slide shows, limiting the dihedral angle to only 10 possible values is extremely crude. Here, we've represented a tetrahedral carbon atom as a steering wheel, and show just how pathetically low the resolution is when we limit dihedral angle changes to only 10 values, *i.e.*, changes of 36° at a time. In actuality, there are an infinite number of possible values for any dihedral angle, and changes of far, far less than 36° can have profound effects on structure.

So, in actuality, there are a lot more than 100,000 possible structures. There are millions. The only way to easily solve a structure problem of this sort is to use one of those chess-like computer programs, which will calmly and quietly look at a million structures and tell you which one is best.

Unfortunately, I have no such program, so I had to do it "by hand", so-to-speak.

Slide 155

What you want to do is find a "unit cell" which, when cloned up the Z-axis, will generate a polymer whose bond lengths and bond angles are well-within normal limits, and whose dihedral angles are such that there will be no steric hindrances.

It's not difficult to come up with an initial model for the unit cell. But when it's cloned, to make 2 residues, you almost invariably find bizarre bond lengths and bond angles at the junction between them. To correct them, you have to revise the entire unit cell structure, which you do by the process of "educated guess".

Once you get it right, you can just clone the structure up the Z-axis indefinitely.

The particular structure shown is the final unit cell used to generate the published protamine-DNA structure. But it wasn't my first unit cell. I can't count how many I went through.

Slide 156

Here, for your amusement and entertainment, are a few stages in the development of a DNA backbone which has no net helical twist, and which progresses vertically up the Z-axis. The first one has obvious problems, which are partly solved in #2, and perfected even more in #3. But I was never able to use #3 either**;** when it was cloned, and bases added, it just didn't work.

Modifying these structures *always* has unpredictable effects. Here's an example of a bright idea which didn't work out too well:

Slide 157

Oops! Back to the drawing board.

Slide 158

The point is that this structure did not come about easily. Those of you who are inflexible "helicists" will surely find my methods arbitrary. Who am I to decide that DNA has a particular structure, then spend a month "making it up"? Ah, but I didn't make it up. If you start this slide show over, and watch more carefully this time, you'll see that there is a logical imperative at each step. Protamine *must* have some structure, and no structure has yet been published. The only structure which has even a prayer of being correct is the beta-sheet, and only a linear, untwisted, stretched-out DNA will anastomose with protamine's multiplicity of positive charges. So our only remaining task is to assemble a DNA backbone with all bond lengths and bond angles standard, and all dihedral angles such that there are no steric hindrances. This is one such structure, and I reiterate that since there is no evident barrier to this structure forming, it certainly *will* form if DNA and protamine are brought close together.

On the other hand, though our solution is a good one, it is not likely to be the best one possible. Only a computer can perfect it, by playing that molecular biological "chess" game I alluded to above. This notwithstanding, the backbone, as we present it here, is entirely reasonable, and, if indeed not perfect, is certainly not in need of very much modification.

Slide 159

Now, having patted myself on the back for creating a structure which obeys all the rules of chemistry, I must now confess that the structure does have a single serious imperfection. At the end of this entire modeling task, I added all the protons, as a precaution, not expecting to find any serious problems. The protonated structures, by the way, are available, or *should be* available on my web site. If you want them, and can't find them online, email me (kb@NotAHelix.com) and I'll send them.

When I examined the protonated structure, I found a single significant steric hindrance. It was not in the backbone, but in the excessive proximity of ribose hydrogen atom 2H2' and pyrimidine hydrogen atom H6. These atoms were only 1.6 Å apart, a 33% violation of the Van der Waals radii.

Slide 160

As I tried to illustrate earlier, you can't just fix this with a few strokes of the pen. Correcting this would have entailed a total revision of the entire structure from top to

bottom, and I confess that I got lazy, and refused to face up to the days or even weeks of additional work necessary.

That the error *can* be corrected is, I think, abundantly self-evident in view of the excessively large number of degrees of freedom in this structure, and so I forgive myself; partially, at least. Thus, I decided to leave this task to the scientists, mentioned above, who own the million dollar computer programs that play the equivalent of chess with DNA and protein structures, who can do in milliseconds what will take me countless hours.

I further excuse myself on the grounds that the Protein Data Bank, which has filters for errors in bond lengths, bond angles, and dihedral angles, passed this structure without comment.

Slide 161

Here are the detailed specifications of the model. The only points worth specifically noting are that (1) our O3'–P bond, 1.64 Å in length, is a few hundredths of an angstrom longer than most published structures, and (2) our 106.44º O3'–P–O5' angle is a bit large. The values for this latter angle, however, vary from 86º-108º in structures I've examined in the Protein Data Bank, so it's well within the range of published values.

For those who object to these slight irregularities, I should point out that I had little difficulty modifying this structure to create an alternative DNA backbone for our model; one in which *every bond length and angle was identical to that found in the "classic" Watson-Crick B-DNA structure*. When I noted that protamine-DNA unit cells, constructed with this latter model, did not pack together quite as well in 3-dimensional space, I stopped working with it.

The packing together of unit cells of the protamine-DNA complex is dealt with below.

Slide 162

Here are the detailed specifications of the protein component of the protamine-DNA complex. The bond lengths and angles are right out of the basic biochemistry textbook, and will be recognizable by any elementary student of this science.

Slide 163

As pointed out previously, the phi and psi angles we have employed, (-) and (+) 130.5° respectively, place this structure in the most favorable portion of the Ramachandran plot for beta-sheets. The slight discrepancies between the phi angles…

Slide 164

…and the psi angles, for adjacent residues, are merely random fluctuations which arose in the process of preparing a dipeptide unit cell for cloning along the Z-axis.

Slide 165

There are a multitude of hydrogen bonds which stabilize this structure. Each is perfectly formed, having a length of exactly 3 Å.

Slide 166

The closest thing to a steric hindrance is the slightly close distance shown here, which is an insignificant 5% less than the sum of the Van der Waals radii for adjacent hydrogen and oxygen atoms.

Slide 167

In other words, as is the case with the DNA component of this structure, I would unhesitatingly opine that the protein component is so energetically-favorable, that it *must* form when in the presence of the DNA component, unless some external force *prevents* it from doing so.

Slide 168 (COMPLETE STRUCTURE)

Slide 169

This is the unit cell, so-to-speak, of protamine-DNA structure. It consists of the two polypeptide strands, P1 and P2, binding electrostatically to about 25 base-pairs of DNA on either side. The disulfide bonds holding the protein strands together are not visible in this view.

Look how perfect and symmetrical the structure is! Everything is right, and everything fits. The Arginine-phosphate ionic bonds, 4 examples of which are indicated now by arrows, are perfectly aligned, and each ionic bond is very close to 3 Å, an ideal length. The amino acid residues which are not Arginine are mostly hydrophilic residues which can donate a proton for a hydrogen bond with a DNA phosphate group

Slide 170

Let's take a closer look.

Slide 171

In this view, you can see the perfect stacking of bases, and the near-perfect placement of ionic bonds. When viewed from this angle, the Arginine residues are easier to see on the left, but the excellent alignment is also clearly evident on the right side of the structure.

Slide 172

This movie is a bit choppy, but it gives you some idea of what the 3-D structure of this complex is.

Slide 173

Now let's just look for a moment at the N-terminal and C-terminal ends.

Slide 174

Here's the N-terminal end. I should perhaps point out, although I suppose it's superfluous, that the DNA, although indicated as terminating here, obviously does not terminate, but continues without interruption to the next protamine P1-P2 dimer, which is not shown.

Slide 175

Not surprisingly, the structure starts with an Arginine residue. There's also a Histidine among the N-terminal amino acids. Note that the P2 strand of protamine is unaccompanied for the first 7 residues of its length, below which it joins P1.

Slide 176

Here's the C-terminal end.

Slide 177

A disulfide bond can be seen here, as well as both of the two proline residues which are found in human protamine. With respect to proline, I had to decide whether it was there to force a kink into the polypeptide backbone, or whether, alternatively, the proline should itself be forced to conform to the secondary structure of the rest of the chain.

I experimented with a number of kinked structures, but I was unable to come up with anything that made sense.

Because proline is not necessarily found in the protamines of other species, and because, when it is found, its position is totally variable, I somewhat arbitrarily decided to simply treat it as a programming nuisance, rather than as a vehicle of kinking, and I forced ***it*** to accommodate to the structure of the chain in general.

Proline is extremely disruptive to secondary structure, but with a little effort, I was able to remove the proline kink by allowable alterations of bond angles and dihedral angles, resulting in a structure with very little in the way of steric hindrance. I was also able to preserve, almost entirely, the alignment between the DNA phosphates and the C-terminal basic residues of Protamine.

Slide 178

Now, I've been carrying on quite a bit about how "perfect" this is, and how well-aligned that is, but the DNA bases, thus far, are still 6.8 Å apart, which is unnatural. Let us see how adjacent unit cells of protamine intercalate to solve this problem:

Slide 179

Amazing, no? Adjacent unit cells fit together like a hand-in-a-glove, with almost enzyme-substrate specificity.

Furthermore, now the bases are perfectly stacked at 3.4 Å.

Slide 180

Now let's look at the structure from an axial point of view. First, we remove the adjacent cells…

Slide 181

…rotate the structure 90°…

Slide 182

…and tilt it forward, to get a top, or axial view.

Wow. That's a lot of atoms.

Slide 183

Let's take a representative cross section, at a point where there are arginines present, to see how the DNA and protein components interact from the axial perspective.

Slide 184

We've looked at this sort of projection before. Here are P1 and P2, with Arginine residues projecting in all 4 directions. Since all 4 R-groups are Arginine, you can't see the disulfide bonds connecting P1 and P2.

Slide 185

At other levels, however, you would see the connection.

Slide 186

Back to our chosen cross-sectional level, you can see again the near-perfect 3 Å ionic bonds to DNA on all 4 corners of the structure. This is a top view of the protamine-DNA unit cell. Note that in this projection, the structure is quite elongated in one direction. At the beginning of this virtual modeling project, I had thought through most of the basic features of the structure in advance. This included the packing of adjacent unit cells through mutual intercalation of base pairs, shown here:

Slide 187-190

(Shows stepwise addition of parts of adjacent structures)

Slide 191

This gives rise to a pattern of alternating columns of DNA and protein…

Slide 192-194 (Shows stepwise addition of labels)

Slide 195

…which continues throughout the length of the sperm cell. This much I had anticipated. But how, as portrayed in this view, would these long rows of alternating DNA and protein columns interact with neighboring rows? I had given no thought whatsoever to this problem

Slide 196

In the most extraordinarily fortuitous development in this entire project, it turned out that I didn't have to give it any thought! The problem literally solved itself! Here's a second row of DNA-protein columns.

Slide 197

Look what happens when you displace them by a distance of 1/2 unit cell! The resulting alignment gives rise to a regular array of salt bridges which is so fortuitously perfect that it's almost more than one could have wished for! Let's look at them more closely.

Slide 198

Here's the figure from my Journal of Theoretical Biology publication on the protamine-DNA structure. First of all, please note that the unit cell is outlined in a faint gray line. (Pink box moves in) This makes it a little easier to see. The protamine-DNA structure consists of this unit repeated millions of times, with adjacent columns binding together either through intercalation of base pairs, or, as we shall now see, through electrostatic bonds.

Slide 199

Please take note of the smaller gray boxes, which I'll highlight now. These are graphic markers to draw attention to the rather extraordinary pattern of square-array groupings of opposite charges on DNA and protein, wherever Arginine residues are present.

Slide 200

Here's a close-up view. Each DNA phosphate group has a strong negative charge…and each Arginine guanidinium group has a strong positive charge.

Slide 201

Thus, there is a square array of ionic bonds, each of which is very close to 3 Å, making this a very, very favorable disposition of charge interactions which runs the length and breadth of the sperm cell.

Slide 202

Now, this is a somewhat unusual charge array, and, if one has not thought the matter through, one might wonder if the *like* charges in the square array would repel one another, destroying the structure. But you've got to remember that the repulsive like charges lie on the *diagonals* of the squares. As this little diagram will remind you, the diagonals of a square are 1.4x *longer* than the sides. Since the strength of a charge interaction decreases exponentially with the distance, the like charge *repulsions* will therefore be far weaker than the opposite charge *attractions.* In fact, the charge attractions will, in this case, be fully *twice* as great as the charge repulsions. We may conclude, therefore, that this square array is *very* favorable energetically.

Slide 203 (Calculation of volume...)

Although it is self-evident that our protamine-DNA model is quite compact, it's interesting to calculate its volume, if for no other reason, as a check on the validity of the structure. Will it fit into the sperm head?

Slide 204

Here are some statistics on the size of the sperm head. If the picture of the sperm is shocking, I apologize. It's easy to forget what we're actually dealing with when we speak of protamine.

The reproductive, or *germ* cells of humans are, of course, haploid, having only one each of the 23 human chromosomes. This gives a final base-pair count of half the somatic count, or 3 billion base pairs. The typical length and width ranges for the sperm

head are shown, and, after a calculation taking into account the annoying tear-drop shape of the cell, we find a volume range of 40-50 cubic microns.

Slide 205

The volume of our unit cell may readily be calculated from the PDB file. The length is 26.5 Å, the width 18 Å, and the height 205.7 Å. The volume is therefore very close to 100 cubic millimicrons.

Slide 206

Once again, we note that the sperm head has 3 billion base pairs. Our unit cell has 59 base pairs. Therefore, there are {3 billion} divided by 59, or very nearly 50 *million* unit cells in the sperm head. Each of these unit cells has a volume of 100 cubic millimicrons, so all 50 million together have a total volume of 5000 cubic millimicrons, or 5 cubic microns.

Since the sperm head has 40-50 cubic microns of space available, it is clear that a protamine-DNA complex having our structure will fit very nicely into the available space.

Slide 207 (SUMMARY AND CONCLUSIONS)

Let's now summarize what we've seen.

Slide 208

We started with the nucleosome core particle.

Slide 209

I've shown you that there is no alignment between the positive and negative charges in the currently-accepted histone-DNA structure, indicating, to me at least, that the structure is not perfected.

But the prospect of perfecting a structure with over 20,000 atoms is a nightmare to even contemplate. Thus I turned to the much simpler protamine-DNA complex.

Slide 210

We showed that Protamine is about 50% Arginine, and considerably *more* than 50% basic, when we add in the lysine and histidine residues. Most of what remains are hydrophilic residues which are hydrogen bond donors: Tyrosine, Serine, Glutamine, Threonine, and Cysteine. The tiny number of remaining hydrophobic residues are woefully insufficient to form the core of a globular protein structure. Since every part of this protein is laden with positively-charged basic residues, there would be no way to avoid placing many of them into any hypothetical core. Once there, the positive charges would all repel each other, unless neutralized by negative charges. Human Protamine, however, does not have a single *negatively* charged residue.

A globular protein structure for Protamine is therefore almost entirely ruled out by mere examination of the primary amino acid sequence.

Slide 211

We showed that the alpha-helix can be almost conclusively excluded as well.

Slide 212

Even assuming that a polypeptide such as Protamine, which is little more than a poly-arginine chain, could somehow form an alpha helix with the help of the negative charges from DNA, the problem still remains that the positive charges on Protamine cannot possibly be aligned with the negative charges on DNA, if the Protamine has the alpha helix structure, and the DNA has the Watson-Crick structure.

Slide 213

Unless, that is, one is prepared to assume that the Protamine-DNA complex *has no structure at all*, and that the sperm nucleus is no more than an unordered collection of poorly-defined, oppositely-charged helices, just dancing about mindlessly in the dark.

Slide 214

I showed you that the beta-sheet is a two-sided structure, with the R-groups on either side having the same spacing as residues in the DNA sugar-phosphate backbone. This raises the possibility of a perfect alignment between positively-charged guanidinium groups in Protamine and negatively-charged phosphate groups in DNA. Is this mere coincidence? You'll have to decided for yourself.

Slide 215

I showed you that the beta-sheet structure most favorable for aligning with DNA occurred at psi and phi angles of ±130.5°, placing it in the most-energetically-favorable portion of the Ramachandran Plot. Another coincidence? Or a clue to the true structure of protamine? You'll have to decide.

Slide 216

I showed you that the rather startling alignment of Cysteine residues in Protamine chains P1 and P2, virtually mandating a parallel beta sheet arrangement, with the chains linked by disulfide bonds.

Slide 217

I showed you what is almost surely the unit cell of the Protamine-DNA complex, consisting of a Protamine P1-P2 dimer ionically binding to two DNA duplexes, one on either side of the beta-sheet, each one un-twisted and extended to a residue spacing of 6.8 Å.

Slide 218

I showed you that the normal DNA base-pair spacing of 3.4 Å was restored when adjacent unit cells mutually intercalated their base pairs. This structure was originally deduced by Tai Te Wu nearly 20 years ago, from Maurice Wilkins' own x-ray data.

Slide 219

I showed you the Gehring Tetramer, a well-known and well-characterized form of DNA proven to have essentially the Wu 4-stranded intercalated structure.

Slide 220

I showed you that in the Protamine P1-P2 beta sheet, the distance between extended Arginine residues was 14 Å.

Slide 221

I showed you that this spacing was closely matched to the DNA cross-duplex distance between phosphate groups, allowing just enough room for perfect 3Å ionic bonds between the guanidinium groups of Protamine and the phosphate groups of DNA…

Slide 222

…a spacing which would be found on both sides of the Protamine beta sheet.

Slide 223

I showed you what I believe to be the true structure of DNA in the Protamine-DNA complex. I also strongly suspect that this structure, or something like it, will be found in many other settings in addition to the sperm cell nucleus.

Slide 224

I showed you how easy it is to fit adjacent unit cells of the Potamine-DNA complex together, by mutually intercalating their base pairs.

Slide 225-226

I showed you a top view of the structure…

Slide 227

…revealing the essential feature of rows of regularly alternating columns of protein and DNA…

Slide 228

…and the manner of association of adjacent rows through an amazingly fortuitous pattern of charge interactions.

Slide 229 (NO AUDIO) (Close-up of a single square-array of charge interactions).

Slide 230

What relevance does the Protamine-DNA structure have for somatic cells?

Slide 231

I showed you earlier that in the nucleosome core particle, the total number of basic amino acid residues in all 8 histone subunits was 146, which is precisely the same as the number of DNA base pairs associated with the particle. Quite a coincidence, is it not?

Slide 232

Now we see that the total number of basic amino acid residues in our Protamine P1-P2 dimer is also very nearly identical to the number of DNA base pairs associated with the unit cell. The base pair count is 59. The basic residue count is a little less, namely 51, if we count Arginine only, or a little more, namely 63, if we also count Lysine and Histidine. I, for one, cannot dismiss this as mere coincidence.

Slide 233

I therefore would be strongly inclined to presume that the structure of the nucleosome, when it is finally perfected, will prove to be very similar to the structure shown here for the protamine-DNA complex, at least in the parts of the structure where DNA binds to protein.

Slide 234

In conclusion, then, I have given you a choice. Concerning an important chemical structure which has apparently eluded scientists for over 50 years, namely the structure of the complex between DNA, the genetic material, and Protamine, the simplest of the nuclear proteins, I have proposed a solution. The solution embodies the elements of logic, order, symmetry, and even beauty.

Yet many will object, even though they cannot propose any alternative whatsoever, simply because they have an emotional attachment to the DNA double helix which overwhelms all logic.

Slide 235

In that case, I guess they'll just have to live with this. Holding on to the past can be comforting, but isn't this too high a price to pay?

REFERENCE SLIDES

(Note references are printed in full below the last slide transcript)

Slide 236

Author's references
Structure files (Protein Data Bank)

Slide 237

More references:

- Histone structure
- Tai Te Wu DNA structure
- Gehring Tetramer structure
- Volume of sperm head

Slide 238 (THE END)

Author's references:

Biegeleisen, K. The probable structure of the protamine-DNA complex. J Theor Biol, 241:533-540, 2006.

Biegeleisen, K. Topologically non-linked circular duplex DNA. Bull Math Biol, 64:589-609, 2002.

Structure files

Protein Data Bank, http://www.rcsb.org, Accession numbers 2AWR (Protamine-DNA complex 1), 2AWS (Protamine DNA complex 2).

These files may also be downloaded directly from https://www.NotAHelix.net.

Histone structure

Luger, K., Mader, A.W., Richmond, R.K., Sargent, D.F., and Richmond, T.J., 1997. X-ray structure of the nucleosome core particle at 2.8 Å resolution. Protein Data Bank, accession #1AOI.
http://www.rcsb.org/pdb/explore/explore.do?structureId=1AOI

Tai Te Wu DNA structure

Wu, T.T. Secondary structures of DNA. PNAS 63:400-405, 1969

Gehring Tetramer structure

Gehring, K., Leroy, J.L., Gueron, M. A tetrameric DNA structure with protonated cytosine.cytosine base pairs. Nature 363:561-565 (1993).
Protein Data Bank Accession #: 225D
http://www.rcsb.org/pdb/explore/explore.do?structureId=225D

Volume of sperm head

Sheynkin, Y., 1998. Understanding semen analysis.
http://www.uhmc.sunysb.edu/urology/male_infertility/SEMEN_ANALYSIS.html