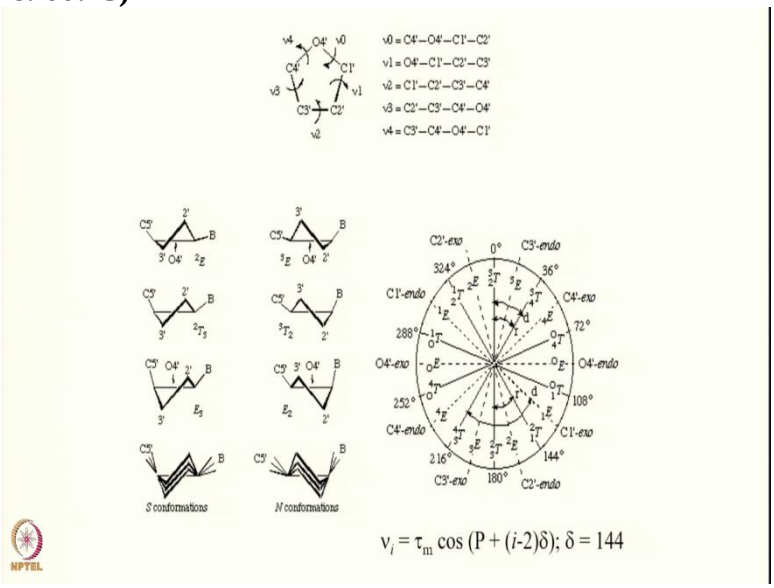
NMR spectroscopy for Structural Biology
Prof. Ashutosh Kumar and Prof. Ramkrishna Hosur
Department of Chemistry
Indian Institute of Technology - Bombay

Lecture: 25
Application of NMR in the Area of Structural Biology: Structure of DNA and RNA 2

(Refer Slide Time: 00:25)



So, we continue with the discussion of the DNA and the RNA structures and you see here as I said the ribose ring or the 5 membered ring is the important element what are the elements of the of the DNA structure with the phosphate backbone is one. And the second thing is the sugar ring and the sugar ring as I said is a 5-membered ring here. Now we have to define the structure of the sugar ring.

So, to and that is an important element we have to define the structure of the phosphate backbone the various angles which are present here the torsion angles which are there we will come to that and then you have this sugar ring. The sugar ring structure is determined by these torsion angles which are indicated here. So, ν_0 , ν_1 , ν_2 , ν_3 , ν_4 not all of them are very independent the torsion angle is defined in this manner.

So, the ν_0 torsion angle is defined as angle around this bond here rotation angle around this bond which defines the relative positions of C4' and C2', C4', O4', C1', C2' and the central bond around the central bond there is distraction rotation angle these are also called as dihedral angles. Because these define the angle between the 2 planes formed by the these groups C4', O4', C1'.

These atoms for one plane and O4', C1', C2' this form another plane and it is the angle between these 2 planes that defines the dihedral angle or it is also the torsion angle around this. So, likewise you have this in ν_1 which is O4', C1', C2', C3' and then you have ν_2 which is C1', C2', C3', C4' and ν_3 C2', C3', C4', O4' and you have so, ν_4 C3', C4', O4', C1'.

But however notice that all these torsional angles are not independent they are dependent on each other because the ring is closed. And then there are also important you can see physical

chemistry calculations one can see how many degrees of freedom are there. So, you have you have the in a typical one you have the 4 bonds and 3 bond angles and you have 2 torsion angles.

Typically 2 torsion angles are sufficient to describe this but then once you have the ring closure then you actually have only one torsion degrees of freedom left and that is what is indicated in this picture here you see. These ones are all related all these torsion angles are related in a particular manner and one degree of freedom which is there, which is called as the p when one defines this p this is called the pseudo rotation angle.

The pseudo rotation angle depending upon the value of the pseudo rotation angle so, you have different kinds of nu i's here. This nu i's the individual torsion angles is given by tau m this defines the maximum distortion from the plane of the ring how much a particular atom is going out. So, this is the kind of quantity number which indicates how much is the a particular atom going out of the plane what is the maximum value.

But this will have a certain constraint because if the bond angles and the bond distances are well fixed and they cannot be changed too much and therefore this actually is a kind of a well defined element. This is typically of the order of 38° and then the cos P is a value which can take values from 0 to 360. And i - 2δ, δ is a fixed number like 144 you put for put $\delta = 2$ and keep varying i through different values then you will see by you get different kinds of sugar geometries.

You depending upon the value of P you get your different geometries and all of these are indicated in the kind of a circle because of the p can go from 0 to 360 or you can say it goes from 0 to – 180 or +180 and so on so forth. And these are typically indicated how what is the kind of a structure that we will get these are all good certain labels it is what is the kind of structure the sugaring has got.

And here you see if you make a plane of a particular thing put the 3 atoms in a plane C1', C4' if you put them in a particular plane. What is the disposition of the 2' and the 3' atoms with respect to that plane and then you will get different and the C' is always on one side. So, you get different kinds of structures for the sugar ring. So, in this case the 2' is above the plane of the ring and the 3' is below the plane of the ring and the O4' is here.

O4' is always there only and in this structure the 3 prime is above the plane of the ring and the 2' is below the plane of the ring. So, here you see both the things are above and below. So, you have the extent of the things which are above and below. So, you have the 3' above 2' below but the degree is more. So, 3' less 2' more the degree is less and therefore you get different kinds of structures which are indicated by typical N this group of structures which are indicated where the 2' is above the plane of the ring.

These are called as S conformations are called the south and this is the south here this is the south in this in this circle this below this central line this is the south and this is the north. So, all those conformations which are below here these are called as the south conformation. So, the S conformation and these are characterized by the 2' atom going above the plane of the ring and these are called as C2' endo.

If it is above it is called as endo anything which is below is called as exo and therefore you have here in these structures then you have see when it is this calls in the south you have the

C2’ endo the C2’ endo is is this one here. So, all of these are C2’ and O and to the different degree how much is the 3’ below. So, you get what you will say here C2’ endo C3’ exo.

So, but the degree of the 3’ which is below or the degree to which this one is above that determines the different positions along the in this circle here. Therefore you see here this is 0 to 180 on this side similarly what we have the north conformations it is a 3 prime which is above about the center line and this is the 3 prime endo. Therefore this is the 3 prime endo here and if you take the p value how the P value is changing.

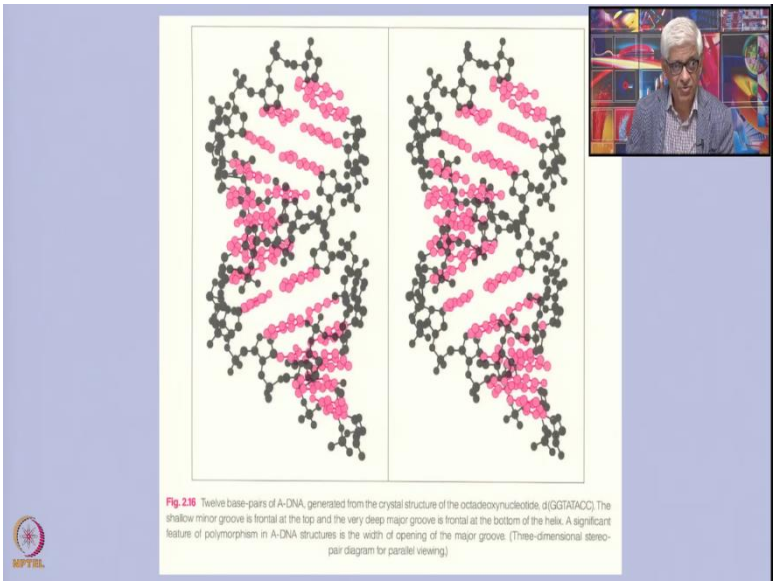
So, this goes from 0, 36, 72 this becomes 90. So, at this 90 the O4’ prime window. Now the other things becomes the O4’ is actually going out of the plane and that is not shown there. So, your 4’ becomes above the plane and put the other three atoms in the in the in the plane of the array in a plane in a particular plane and then you go you have the twisted conformations they are described as endo conformations twisted conformations to twisted means both the things are above and below but to different extents it is not only one.

So, you have in one particular case then only one which is all the 4 atoms are in a plane and only one atom is above the plane you can also think about that and those ones are these are 4’ endo’s or 4 prime endo’s or 4’ exo four atoms in a plane and one of them is above and all other ones are kind of twisted to a certain extent which is closer to this 180 this is called a C2’ endo and then you have C3’ exo here and so on so forth.

So, you see p is an important parameter which determines the nature of the sugar ring and the sugar ring varies from one nucleotide unit to the second nucleotide it can vary. It can vary in the DNA and the RNA why it can vary because the C2’ prime here in the case of DNA has 2 hydrogens here whereas in the RNA it has one oxygen and therefore this can create a kind of a steric differences and interaction differences between the oxygens here and between the one nucleotide unit to the next nucleotide unit.

Therefore this sugar ring can have different conformations. Indeed you will find as we will show later that in the case of DNA this sugar ring adopts the south conformations by and large the C2 endo conformations is called as C2’ prime endo conformations and in the RNA they adopt what is called as the C3’ prime endo conformations and that primarily comes because of the oxygen at this point.

(Refer Slide Time: 09:01)



So, Now you see what it reflects in the case of there are different kinds of structures and and these are classified as A, B, C, D, E, in the DNA. Now the the previous the fibre diffraction data the fiber diffraction data could not give you any more detail about this. Now how do we get this information. So, you can use tools like NMR or crystallography. And people then you actually synthesize small DNA segments to get more details into the structures of the individual nucleotide units or short segments of the DNA segments.

Because there is a repetitive in it. So, therefore you can say if you take a short segment of the DNA segment or RNA segment you might be able to think what is the kind of a structure which is propagating through the through the polymer. And that is what was done and you they came out with the different kinds of DNA structures and those ones are indicated here.

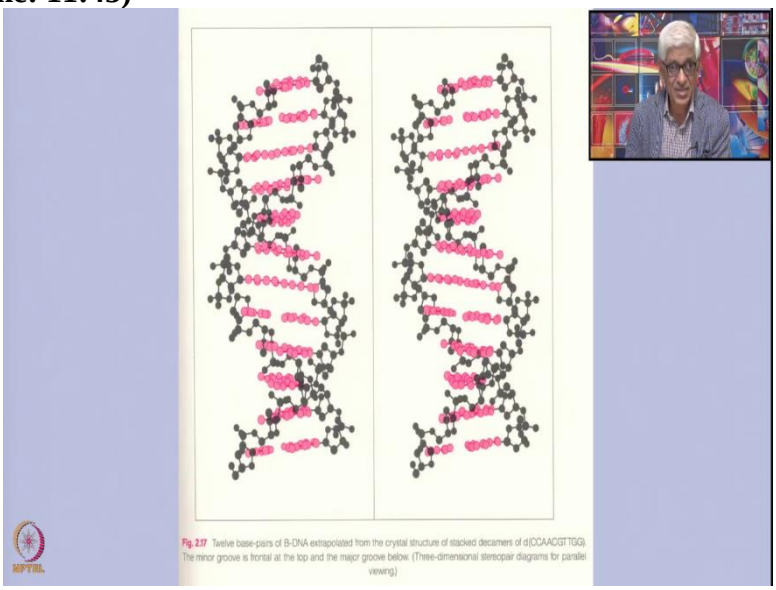
And this is the 12 base pair of A-DNA this is a particular kind of A-DNA structure where it is the A-DNA and this is the structure of this molecule deoxy GGTA TACC and this is the 5' end here and this is the 3' end here the sequence goes as GG TA TAC this is how the DNA sequences are written. And here you see the structure has a certain kind of effect is a fat structure and this the and the base pairs are tilted with respect to the axis of the DNA access of the double helix.

The axis of the double legs may go like this and the base pairs are tilted in this half they are tilted like this in the top half they are tilted like this and that typically the fibre diffraction data had shown that one complete turn of the DNA. So, what is the complete turn. So, you go from here to here like this and then you come to the beginning and that is one complete turn from here to your one complete turn.

And the fibre diffraction data has said that there can be approximately 10 nucleotide 10 base pairs in a particular segment and depending upon the structures of course there can be slightly more and also in the later segments which have been discovered. Therefore you have this a particular half and this particular half and this structure is called as the A-DNA. Because this DNA molecule has structures which are very reminiscent of the 3 prime endo sugar geometry and therefore this is called as A-DNA.

And normally what is present in the larger probability in your solution inside the cell is what is called as the B-DNA.

(Refer Slide Time: 11:45)



And this is the B-DNA and this is the structure obtained by a largest DNA sequence under the different conditions and depends upon the hydration level also inside your structure I mean inside your DNA a crystal how much water molecules are there. So, it also depends upon those ones and this is the structure of a 12-mer DNA segment which has the best sequence C C A A C G T G T T G G notice here this is self complementary what is the meaning of self complementary.

So, if I write here the sequence C C A A C G T T G G. Now this is the 5 prime end and this is a 3' end right. Now what I do, I write the same sequence here back C C A A C G then I have T T G G. So, this is the 3 prime and this is the 5' you see I have written the same sequence twice. So, therefore this is called as the self complementary sequence and therefore one molecule which is there naturally it goes into the form of a double helical structure like this.

And this is easier to work with such molecules otherwise you will have to generate 2 different kinds of sequences and then to make a duplex structure. So, this molecule was chosen though that it automatically forms a duplex structure as the most stable form. So, this is self-complementary. And this molecule the structure showed was like this and it is little bit more elongated you can see compared to the RNA molecule and the A-DNA this structure is called as the B-DNA.

And it has clearly 2 different kinds of grooves one can identify two grooves there is one groove which goes like this which is called as the major groove and there is another group which is the smaller groove which is called as the minor groove. So, there are 2 different kinds of groups described for the DNA structure. So, this amount depending upon the amount of space which is available and you have a groove which is defined.

And the base pairs open in the 2 different grooves one side of the base pair opens in the major groove other side of the base pair opens in the minor groove. So, that is when you explicitly write the base pairs and that is what one sees.

(Refer Slide Time: 14:42)

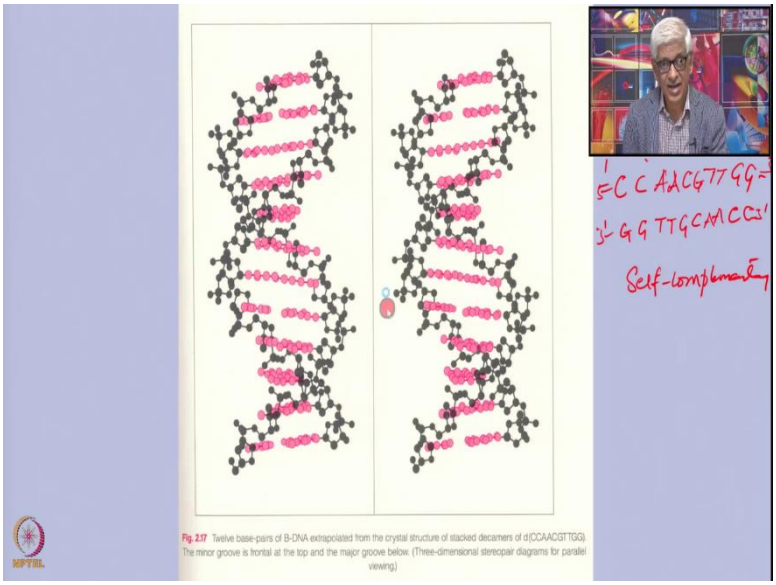


Fig. 2.17 Twelve base-pairs of B-DNA extrapolated from the crystal structure of stacked decamers of d(CCAACGTTGG). The minor groove is frontal at the top and the major groove below. (Three-dimensional stereopair diagrams for parallel viewing.)

Now this is the third type of DNA segment and in naturally. Now let us go behind here. Now you see these ones are right handed helices here you see this is one right-handed helix the double helix goes like this, this is the right-handed helix and the other one is coming in the opposite direction. So, that is also right-handed deluxe handed helices are coming like this and then they are held together by this hydrogen bonds in the base pairs.

So, in this B-DNA as well as in A-DNA we have 2 right handed helices which are going which are intertwined at 2 for base pairs and that is but there are differences in the structural elements how fast is the how much is the distance between one base pair to another base pair typically it is about 3.64 angstroms here. So, therefore if I take 10 base pairs you will have the one end of once complete this turn of the DNA is about 34 angstrom.

This is what I used earlier for the calculation as what is the total length of the DNA segment. And therefore you have this; the unit rise is 3.4 angstroms and then you have a 34 angstroms in the B-DNA. In the case of RNA is slightly smaller and the number of units also is is given the same the height is slightly smaller.

(Refer Slide Time: 15:58)

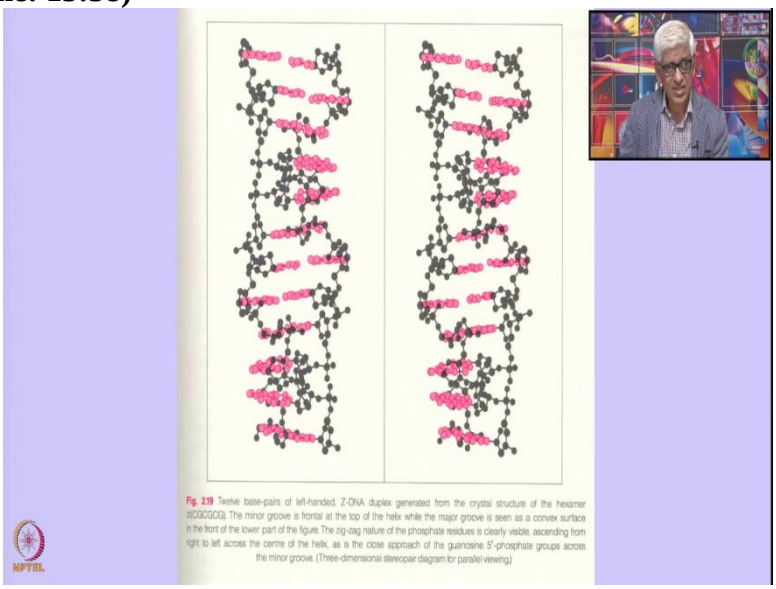


Fig. 2.19 Twelve base-pairs of left-handed, Z-DNA duplex generated from the crystal structure of the hexamer d(CGCGCG). The minor groove is frontal at the top of the helix while the major groove is seen as a convex surface in the front of the lower part of the figure. The zig-zag nature of the phosphate residues is clearly visible, ascending from right to left across the centre of the helix, as is the close approach of the guanosine 5'-phosphate groups across the minor groove. (Three-dimensional stereopair diagram for parallel viewing.)

Now this is another kind of a DNA structure which was discovered this actually came quite late this came somewhere in the 19 close to the 1980s 1979 and this is called as the z DNA or

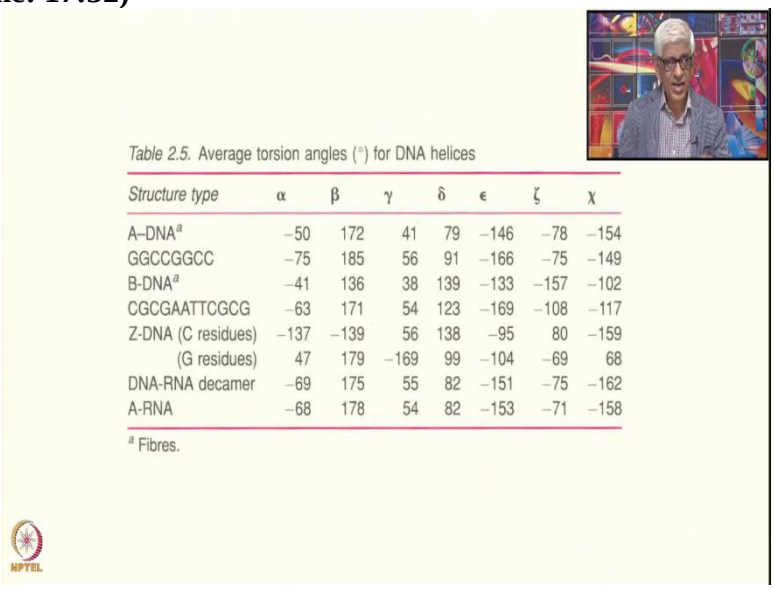
the Z-DNA and this actually is a left-handed helix this is not right-handed helix this is the left-handed helix and this very characteristic feature of this is that this is it is not a monomer which is repeating here but it is a dimer.

So, the dimer structure is the basic unit of this repeating unit and this is a left handed DNA structure and it occurs only in such point of a sequences which are rich in CG segments. CG segments is initially it was not clear whether actually biologically present or not this is called as the zigzag nature of the of the phosphorus phosphate residues that you can see. So, this is also that is why it is called Z-DNA it is zigzag nature.

It is not a proper helix as you can see in the in the case of B-DNA or the A-DNA but this is the zigzag nature and the dinucleotide is the repeating unit unlike the mononucleotide which is the repeating unit in the A and the B-DNA and in the dinucleotide the 2 units are different structural features and that is an important factor how they become different. And these are typically observed in very high concentration of whenever there is high concentration of salt.

And why did they do it of course just curiosity and they discovered this and then of course it became an important point.

(Refer Slide Time: 17:32)



So, here is the summary of all of these DNA segments the structures of the different DNA segment. So, here you have the A-DNA then you have the B-DNA and this is the crystal structure which these are these are from the model A-DNA in the small a which is here this indicates from whatever the build model they built the Watson and Creek these are from fibres, fibre diffraction data.

And these are from single crystal structure GGC GGGCC this is also you know it is a self-complementary sequence and this forms an A-DNA. The B-DNA this form which is there CGCGAATTCGCG this is also a self complementary sequence because if you write in the similar reverse way you will find it is the duplex. And the Z-DNA is as I said the CCG that was a structure which was there and it is a dinucleotide which is repeating unit here.

You have the C residues and the G residues they have different structures the individual units are different structures. Now what are the structures here what are these α, β, γ and δ, ε these ones these are the torsion angles along the backbone of the DNA. So, you have this backbone

phosphodiester linkages all over there and that is the various six torsion angles which are present along the phosphodiester backbone.

And this χ is the angle which connects the sugar ring to the base sugaring to the base that is the C1' nitrogen that bond the torsion angle around that bond that actually defines what is the orientation of the base with respect to the sugar ring and that angle is called as chi torsion angle. And you notice here these have certain ranges of values typical ranges of values by and large.

So, these ones the alpha is around between minus 50 to -75 and B-DNA it has -41 this is slightly lower compared to these ones and the B and the β angle this is 172 to 175 and here in the B-DNA it is smaller 136 and the γ, δ these are along the backbone. So, along the polypeptide backbone you have this different torsion angles there. So, if I want to draw here this is the CH_2 this is the 5' this goes to the oxygen here.

This is the 5' then I have the C4' the C3' then the O3' then you have the CH_2 again this is the phi of the other end and then you have the P here then you have the O here again and your sugar ring is here this goes to the O here C1', C2' and this is connected to this. And this angle is the δ torsion angle there and this is the γ and this is the beta and this is the α.

And this angle is the epsilon and this is your X_i and the one which connects to the C1' to this on there that is the χ that one angle which is not there is no space here and this goes this is the this is the χ here. This is the 5' again CH_2. So, therefore you have this α, β, γ and δ, ε and X_i. So, these are the six torsion angles what you have in this here and how these values are changing.

And of course there is also people did a DNA RNA decamer and of course this is a mixture of the DNA RNA and this has different kinds of torsion angles here the values are indicated there and A-RNA is the sugar ring is more C 3 prime window here there is a sharp contrast and what it reflects which angle reflects the sugaring here this by enlarge it is in the delta because the delta is in the sugar ring C 4 prime C 3 prime that is part of the sugar ring.

And so, it is reflected in the delta torsion angle how these ones are different. You see here the A-DNA has 7991 and the B-DNA has 130 between 120 to 140. So, therefore this is a quite characteristic and the z DNA has alternating this and this. So, you have the C residues has this C2' prime endo and the G residues has the so, called C3' endo. So, if the δe is around the 79, 90 that is the C 3 prime in the structure and if it is another 120 to 140 that is the C2' parameter structure.

So, therefore in the Z-DNA you see the C residues and the g residues have different sugar geometries. So, one of them has the C2' endo kind of a thing other one has a C3' parameter kind of a string. And the beta torsion angle is mostly around the trans value and except for this particular one here it is 136 otherwise it is mostly around the 180 value roughly. And similarly this is the gamma is roughly around the 40 between so, called the Gauss conformation except for here the G residues in the Z-DNA and this goes into the trans area.

So, this is minus 169 is like the terms area all of these ones are. So, called Gauss conformations here close to less or more and things like that. So, this is how the various structures. Now all this information has come from the single crystal structures of short DNA segments. Now there is also one should describe what about the base pairs base pairs are they all parallel to each other or they are different from one another.

(Refer Slide Time: 23:30)

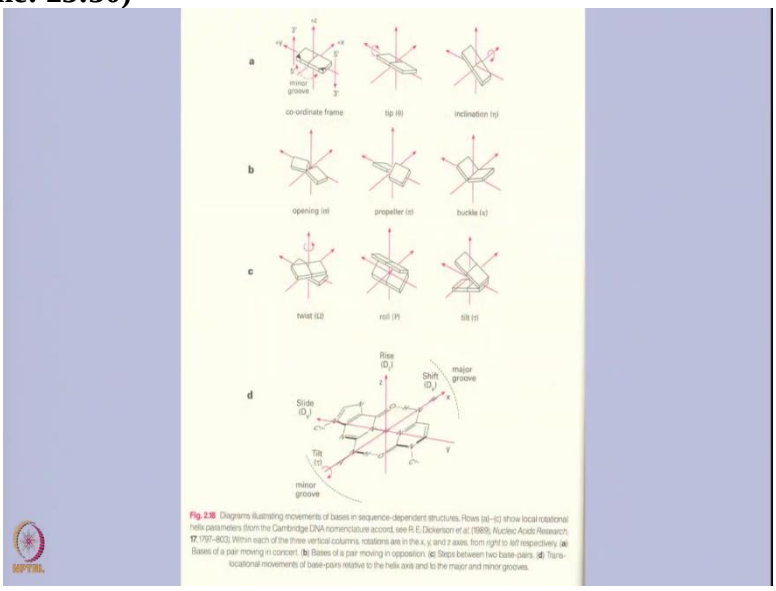


Fig. 2.18 Diagrams illustrating movements of bases in sequence-dependent structures. Rows (a)–(c) show local rotational helix parameters (from the Cambridge DNA nomenclature accord, see R.E. Dickerson et al. (1989). *Nucleic Acids Research*, 17, 1797–803). Within each of the three vertical columns, rotations are in the *x*, *y*, and *z* axes, from *right to left* respectively. (a) Bases of a pair moving in concert. (b) Bases of a pair moving in opposition. (c) Steps between two base-pairs. (d) Translocational movements of base-pairs relative to the helix axis and to the major and minor grooves.

The base pairs can differ to and you have to describe what is the relative orientations of the bases in the base space. So, you can define that in with different parameters here. So, these are the various parameters which are used to describe the base pairs. I told you that the base pair which is there on one side of that is the minor groove other side is the major groove. So, typically the base pairs open on the 2 sides.

Because these are in the interior of the double helix and they open on the 2 sides of the double helix. So, one side is the major group other side is a minor group. So, and now you see here it is the 5 prime to 3 prime it is going here this is 5 prime to 3 prime and the base players are held together like this. And now are these in the same plane or they can be 2 are they parallel to the orientation with respect to the double helical axis?

Or are they tilted with respect to each other various kinds of parameters can be described here what is the orientation with respect to the axis of this? Here it is perpendicular, perpendicular to the helical axis here it is slightly tilted in one dimension one direction this tilted with respect to what. So, this particular axis along this. Now here it is tilted with respect to the other axis. So, there is a rotation here this is indicated by the rotation here.

You indicate a rotation consider this as a planar one and in the whole base pair is the plane and you take a rotation around this axis therefore this becomes like this. Now you take a rotation around this axis it becomes like this. So, therefore you can describe the orientation of this base pair using this one particular parameter that is when these are all parallel if they are parallel. Now but they whether they are all exactly in the same pair like this or there is a small opening here.

So, you can again describe these ones here there is an opening with respect to this axis there are 2 bases are like this or the 2 bases are like this all the 2 bases are like this. So, this is the different torsion this is the parameter which is called as the κ here and you have different these are labelled by different parameters and you can describe this as within the same base pair how they are oriented with respect to each other the 2 bases that.

Now you can also describe the relative orientation of 2 base pairs consecutive base pairs. If you look at this two consecutive base pair with a rotation around this particular bond at this

particular axis will displace the 2 base pairs in this manner. So, they are not exactly talking or sitting on top of each other but slightly displaced with respect to each other their planes.

Now if they are angled like this then difference is called as the roll this is called as the twist this angle is called as the twist this angle is called as the roll and this angle is called as the tilt. So, you differentiate these different parameters with the by different names. So, you one of them you call as a twist this call is the roll this is called as the tilt and similarly here you call them as opening and this is the propeller and this is the buckle.

So, here also we inclinate these ones are called as tip this is inclination and this is the coordinate frame in the same indicating the coordinate frame. And all of those ones are described here in a little bit more explicit manner. So, you have summarize. So, you have the minor groove on this side and the major groove on this side. So, if the base pair is base pair is like this you draw an axis like this to through this through the center of the base pair.

And then you have a vertical one and you have an axis like this you define x y z these 3 axis with respect to the base pair and you can define the positions of all of these groups with this as rotations with respect to these 3 these 3 axis there and that is called as the slide and then you have the shift. So, how they are shifted with respect to the center. So, whether it is exactly going through the center or it is shifted with respect to that.

So, this one is shifted so, this is a kind of a slide and this is the shift. So, these are the different parameters which are used to describe the orientations of the base pairs in the same in the plane and. So, you have so, many parameters to describe the structure of the DNA segment. So, you have the sugar geometry on one hand then you have the torsion angles along the polypeptide backbone.

And then you have the base pair orientations with respect to each other and whether they are in the same plane or they are slightly tilted with respect to each other and how are they stacking whether the base pairs are stacking exactly on top of each other or they are slightly different from one another. So, these are the various parameters all this information has come because one could obtain high resolution structures of small DNA segment. And this is taken from this the paper of R E Dickerson this was this published in 1989.

(Refer Slide Time: 28:58)

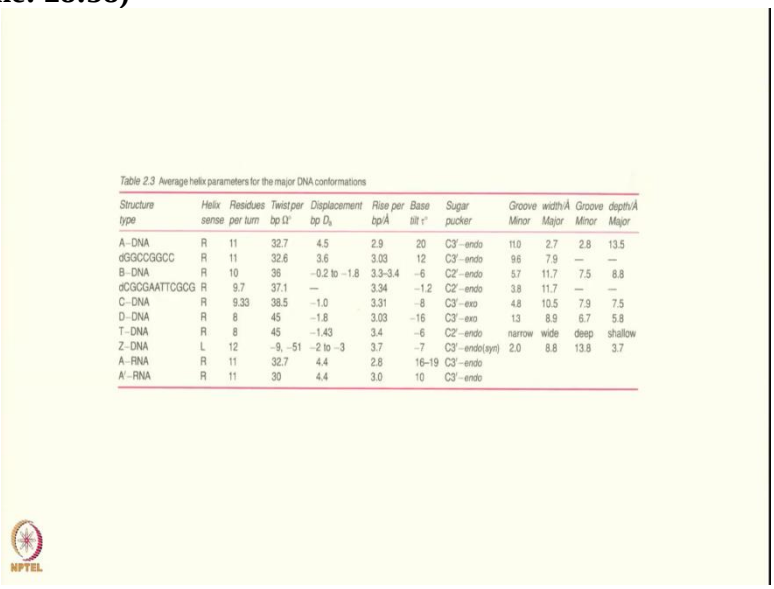


Table 2.3 Average helix parameters for the major DNA conformations

Structure type	Helix sense	Residues per turn	Twist per bp Ω°	Displacement bp D_h	Rise per bp/Å	Base tilt τ°	Sugar pucker	Groove Minor	width/Å Major	Groove Minor	depth/Å Major
A–DNA	R	11	32.7	4.5	2.9	20	C3′–endo	11.0	2.7	2.8	13.5
dGGCCGGCC	R	11	32.6	3.6	3.03	12	C3′–endo	9.6	7.9	—	—
B–DNA	R	10	36	−0.2 to −1.8	3.3–3.4	−6	C2′–endo	5.7	11.7	7.5	8.8
dCGCGAATTCGCG	R	9.7	37.1	—	3.34	−1.2	C2′–endo	3.8	11.7	—	—
C–DNA	R	9.33	38.5	−1.0	3.31	−8	C3′–exo	4.8	10.5	7.9	7.5
D–DNA	R	8	45	−1.8	3.03	−16	C3′–exo	1.3	8.9	6.7	5.8
T–DNA	R	8	45	−1.43	3.4	−6	C2′–endo	narrow	wide	deep	shallow
Z–DNA	L	12	−9, −51	−2 to −3	3.7	−7	C3′–endo(syn)	2.0	8.8	13.8	3.7
A–RNA	R	11	32.7	4.4	2.8	16–19	C3′–endo				
A′–RNA	R	11	30	4.4	3.0	10	C3′–endo				

So, this is the description of the individual base space here average helix parameters for the major DNA conformations we need not go into the details of these numbers here. But this is a kind of an indicator as to what sort of values are there for the different parameters the characterizing the DNA structure in the different DNAs so, depending upon the small variations which are present.

So, different kinds of labels have been given for the different DNAs we talked about the A-DNA B-DNA but there are also some variations which occurred and they led to what are called as the C-DNA D-DNA T-DNA then you also have the Z-DNA which you also described in major detail. And then you have the A-RNA and A'- RNA these ones are the RNA segments all those are the DNA segments.

These are only minor variations with respect to the major ones the ones which are by and large available in the in the natural sequences are the B-DNA and the A-DNA and even there the RNAs are mostly in the a form and the DNA is in the b form. And this depends on also on the extent of hydration how much is the water content depending upon that you get slightly different structures.

(Refer Slide Time: 30:20)

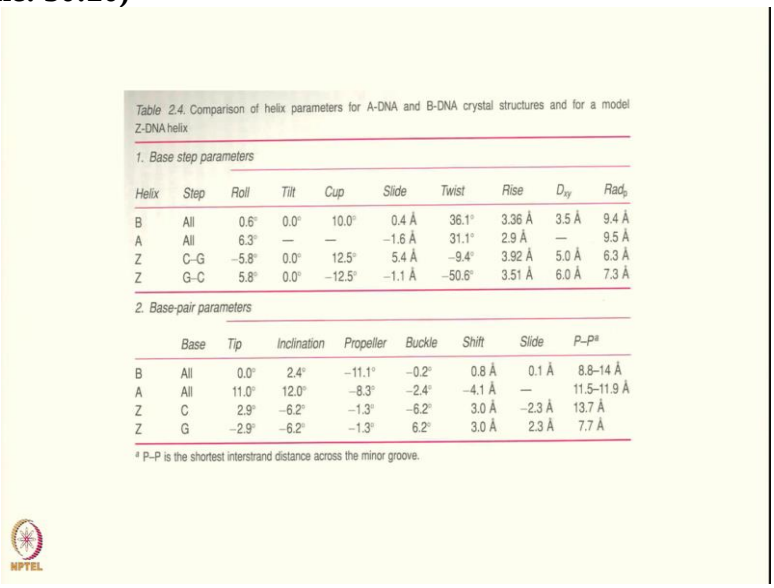


Table 2.4. Comparison of helix parameters for A-DNA and B-DNA crystal structures and for a model Z-DNA helix

1. Base step parameters

Helix	Step	Roll	Tilt	Cup	Slide	Twist	Rise	D_{xy}	Rad_p
B	All	0.6°	0.0°	10.0°	0.4 Å	36.1°	3.36 Å	3.5 Å	9.4 Å
A	All	6.3°	—	—	−1.6 Å	31.1°	2.9 Å	—	9.5 Å
Z	C–G	−5.8°	0.0°	12.5°	5.4 Å	−9.4°	3.92 Å	5.0 Å	6.3 Å
Z	G–C	5.8°	0.0°	−12.5°	−1.1 Å	−50.6°	3.51 Å	6.0 Å	7.3 Å

2. Base-pair parameters

	Base	Tip	Inclination	Propeller	Buckle	Shift	Slide	P–P[a]
B	All	0.0°	2.4°	−11.1°	−0.2°	0.8 Å	0.1 Å	8.8–14 Å
A	All	11.0°	12.0°	−8.3°	−2.4°	−4.1 Å	—	11.5–11.9 Å
Z	C	2.9°	−6.2°	−1.3°	−6.2°	3.0 Å	−2.3 Å	13.7 Å
Z	G	−2.9°	−6.2°	−1.3°	6.2°	3.0 Å	2.3 Å	7.7 Å

[a] P–P is the shortest interstrand distance across the minor groove.

And this is the continuation of the same. So, what sort of parameters are there for the base pairs orientations of the base pairs in the and this is for the z DNA and those ones were for the A B C and D etc and since in the Z-DNA you have different things for the 2 different steps the CG steps are different. therefore you have different parameters here for the 2 things. So, in other words we need to characterize this kind of details for the structure when we want to calculate the structures of the molecules.

And crystallography has been used and NMR also has been used and we will see more what kind of structures have been determined by the NMR data and these happen in the solution. So, this has become possible with various kinds of NMR experiments and certain algorithms which have been developed. So, I think we can stop here.