Assembly and Characteristics of Nucleic Acid Double Helices

Patterns of base-base hydrogen bonds-Characteristics of the base pairs

Interactions between like and unlike bases have been observed in crystal structures individual nucleotides. There are 28 different base pairs which can be formed that have at least two H-bonds between them **(FIGURE 26)**, and whose H-bonds are reinforced in strength by virtue of the fact that they are cyclic **(Figure 27)**.
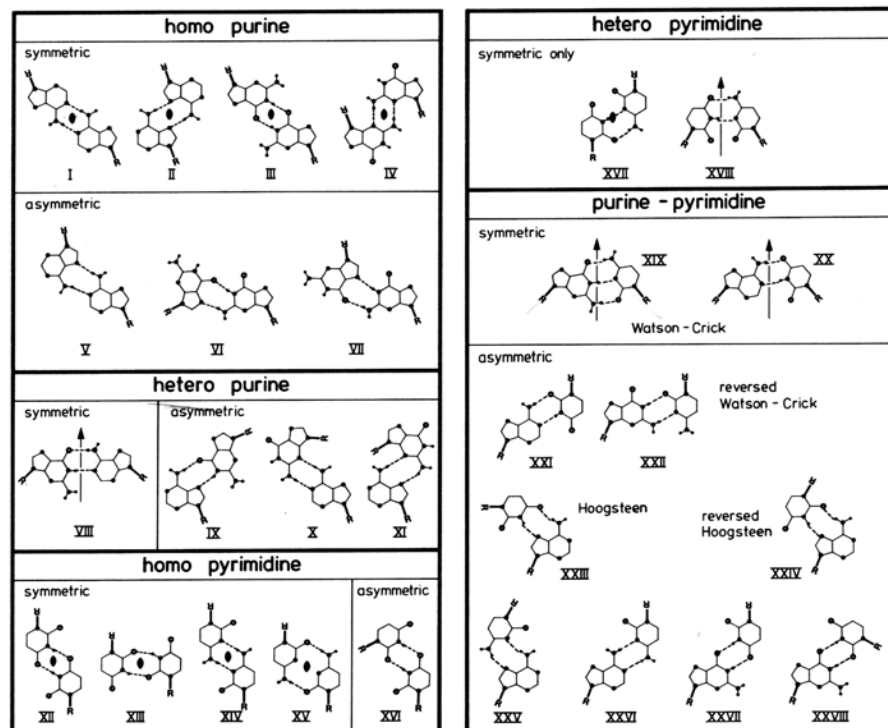

These base pairs are symmetric at least at the level of the C1'-N bond. That is transformations about the central axis between the two base pairs results in the exact placement of one link on top of the other. In some of the cases delineated in the figure, an exact symmetry axis exists; the entire base can be placed on top of the other by rotating thru the dyad axis (SEE FOR EXAMPLE I-IV). Base pairs which have a dyad axis that is perpendicular to the plane of the bases most readily form parallel stranded double-helices,
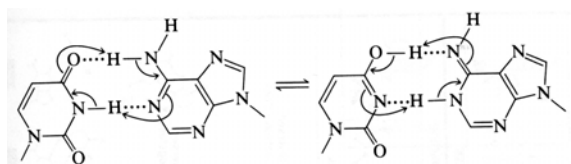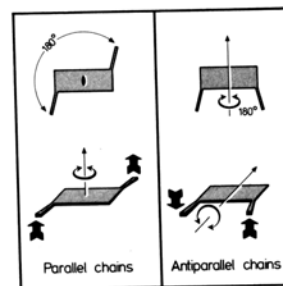
Figure 26

Figure 27

Figure 28 Base-pairs can display two kinds of twofold symmetry, depending on the orientations of the glycosyl $C_{1'}$−N bonds. In this illustration, the latter are indicated as sticks protruding from rectangular plates representing base-pair planes.

whereas base pairs whose dyad axis is parallel to the plane of the bases most readily for anti-parallel double helices **(Figure 28)**.

Not many of these base pairs in Figure 26 are found in naturally occurring double stranded polynucleotides, but more of them exist than we would have expected at first from the Watson-Crick DNA structure.

Let's examine the most common base pairs present in DNA and RNA, the Watson Crick base pair. This base pair stoichiometry of the W/C base pair satisfies the most striking observation of the composition of DNA and RNA double helices. That is A=T and G=C, indicating that these A+T and G+C occur in pairs. The bonding pattern of the bases gives a partial symmetry, at least at the level of the C1' carbons, a feature that was key to discerning the bonding pattern in DNA.

The Hoogsteen base pairs also can satisfy the stoichiometry requirements of G=C and A=T, but failed to make the symmetry test. In no way can these be rotated to give any dyad symmetry, pseudo or otherwise. Several other type of non-Watson-Crick base pairs exist, most notably, the wobble pairings surmised from the complementarity of the tRNA with mRNA, and accounts for the observation that despite the fact that there are 64 possible codons and all are usable, there are not 64 different tRNA anticodon stem sequences. Thus, mispairings must occur. Examples of these are shown. Note that all are missing a dyad axis, rotation about an axis can not be done to give C1'-C1' superposition.
These wobble base pairings do not only occur in tRNA-codon interactions but have been observed in crystal structures B-DNA containing mispairings.

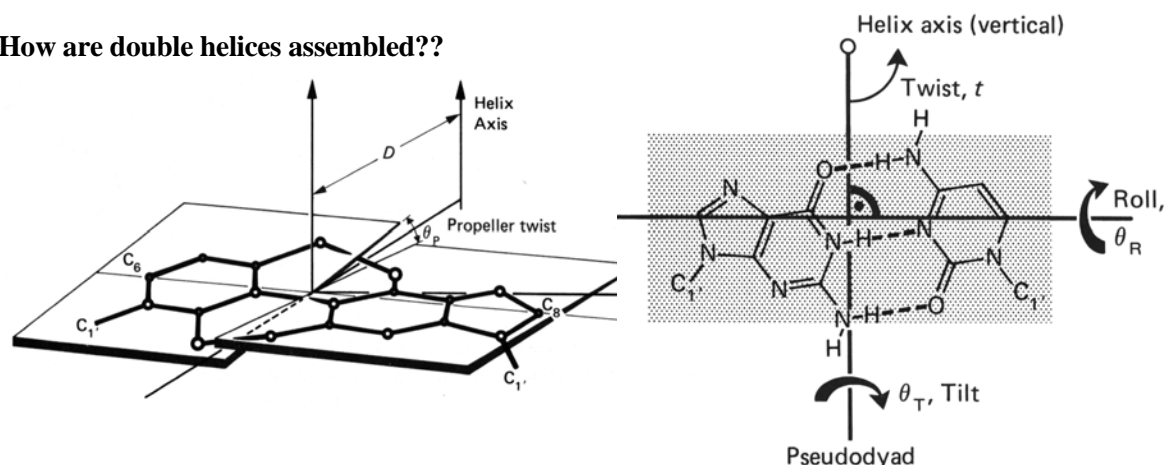**How are double helices assembled??**

Figure 29

Let us first examine the angular characteristics of base pairs. **Figure 29** diagrams a base pair. The bases in a base pair are usually not coplanar; they instead are twisted about the hydrogen bonds that connect them, like the blades of a propeller. The dihedral angle that defines the non-coplanarity is called the propeller twist angle. To view the angle, hold the base pair such that you are looking down its long axis, the angle is defined as positive when the nearer base rotates clockwise.

If the base pair is imbedded in a helix, then there are several more angular attributes of the base pair that we must consider:

1) D-displacement from the helix axis. By virtue of the symmetry axis we discussed above, in a double stranded nucleic acid, there exists a helix axis which is defined by the average symmetry axes of the base pairs. In some cases, the base pair "slips" from this axis, and this displacement from the axis is measure as distance from the helix axis.

2) Base pair tilt. Despite the propeller twist of a base pair, an average mean plane of the base pair is defined (GRAY AREA IN FIGURE 29). The rotation of this plane about the pseudo-dyad axis defined above is the

14

base pair tilt.

3) Base pair roll. This parameter measures the degree of departure of the mean plane of the base pairs from the perpendicular helix axis on the short axis of the base pairs.

4) Helix twist. Defines the orientation of a base pair with respect to the helix axis. That is how big an arc the base pair traces as it measured from one base pair to the next.

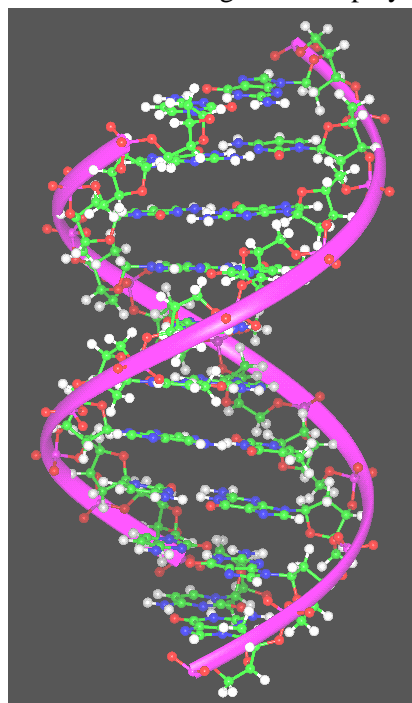What do these things mean to polynucleotide structure?

Figure 30

Propeller twist, base pair roll and displacement are extremely important components in maintaining the stacking interaction of DNA. **Figure 30** shows that in standard Watson Crick B-DNA, the bases are coplanar, no propeller twist and the bases stack readily upon each other.   In this type of DNA, the helical twist is 36°, meaning that there are 10.0 base pair/1 turn of helix. However, if any of these parameters are varied the stacking of the base pairs is changed dramatically.   For example we lower the temperature or bind a protein. If this change in environment induces an overwinding of the DNA to give 9.33 base pairs per 1 turn of helix (helical twist of 38.6^{o}) and there were no change in either propeller twist, tilt and roll occurred, stacking would be disrupted or worse, the base pairs would crash into one another.

More importantly, and as many things are in nature, more subtly, base pair tilt, roll, helical twist and propeller twist are SEQUENCE dependent, both from the influence of stacking interaction energies and the van der Waals constraints imposed by different base pairs.

Table 5 shows the conformational variations in DNA, dependent both on composition and sequence of the DNA as well as the composition of the surroundings.   For example, native DNA can be inter converted between two different families of DNA, B and A, just by changing the

Table5

Principal Crystalline Forms of DNA and RNA in Fibers: Dependence of Form and Relative Humidity (Equivalent to Salt Concentration)^a

Polynucleotide	Counterion	Relative humidity (%)	Form
Native DNA	Na	75	A
	Na	92	B
	Li	57–66	C
	Li	44	C
	Li	66	B
Poly(dA)·poly(dT)	Na	70	β–B'
	Na	92	α–B'
Poly(dG)·poly(dC)	Na	75	A
	Na	92	B
Poly(dA–dT)·poly(dA–dT)	Na	75	D
	Na	Up to 98	A
	Li	66	B
Poly(dA–dC)·poly(dG–dT)	Na	66	A
	Na	66–92	B
	Na	66	Z
Poly(dA–dG)·poly(dC–dT)	Na	66	C"
Poly(dG–dC)·poly(dG–dC)	Na	95	B
	Na	43	Z
	Na	Up to 92	A
Poly(dA–dA–dT)·poly(dA–dT–dT)	Li	81	B
	Na	66	D
Poly(dA–dG–dT)·poly(dA–dC–dT)	Na	92	B
	Na	Up to 98	A
	Li	98	B
Poly(dA–dI–dT)·poly(dA–dC–dT)	Li	66	C
	Na	66	D
	Na	81	C
Poly(dI–dC)·poly(dI–dC)	Na	92	B
	Na	66	B
Native RNA (reovirus)	Na	75	D

15

humidity. The compositional isomers serve to illustrate the sequence dependence; poly (dA-dT)$_2$ exists A, and B DNA, but poly (dA-A-T).dT-T-A) can exist only as a member of the B-family. These are only a few of the sequence dependent things that can occur in DNA, lets now examine A, B and Z-DNA structure in detail.

Information content of DNA.

As we discussed in the first lecture, the main function of a nucleic acid is the transfer of genetic information. In the case of DNA, this means not only inter-generational information or passing of the genetic blueprint as well as coding for the general body plan, but it also must contain regulatory information, to help the cell decide when to transcribe a particular gene and when to replicate. These "read-outs" of information are usually made by specific DNA binding proteins. These proteins recognize and bind to DNA sequences that are present in only one or a few copies per genome. Since the DNA of a simple organism like a bacteria contains millions of base pairs, how does the protein recognize a specific sequence--well it "reads" the pattern of H-bond donor and acceptor groups present on a sequence of DNA. The pattern of this information readout must be unambiguous.

The base pairs contain two different surfaces which, when they are contained in double helical DNA are displayed on opposite sides of the molecule. By convention, the sides are defined based on which side is facing where in B-DNA **(Figure 31)**. MAJOR GROOVE-MINOR GROOVE. The information content of the major groove is unambiguous; the minor groove is ambiguous.
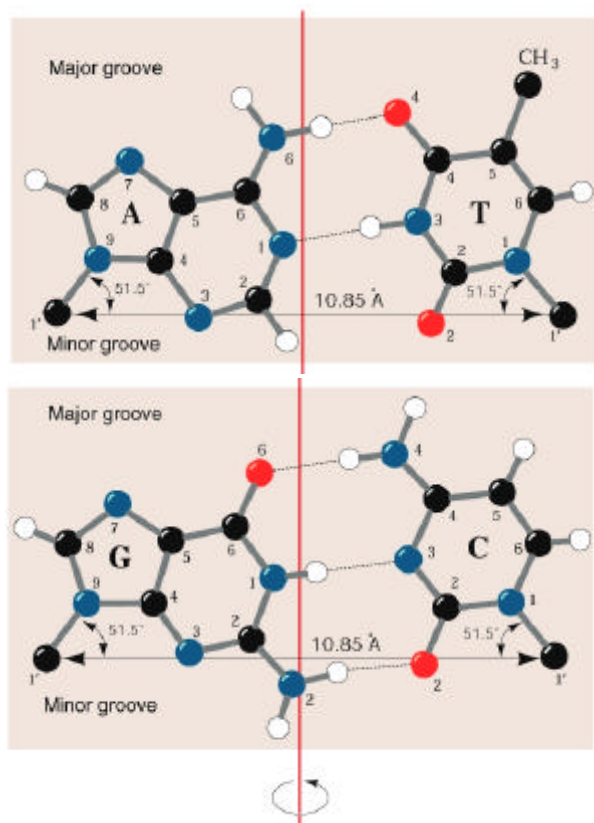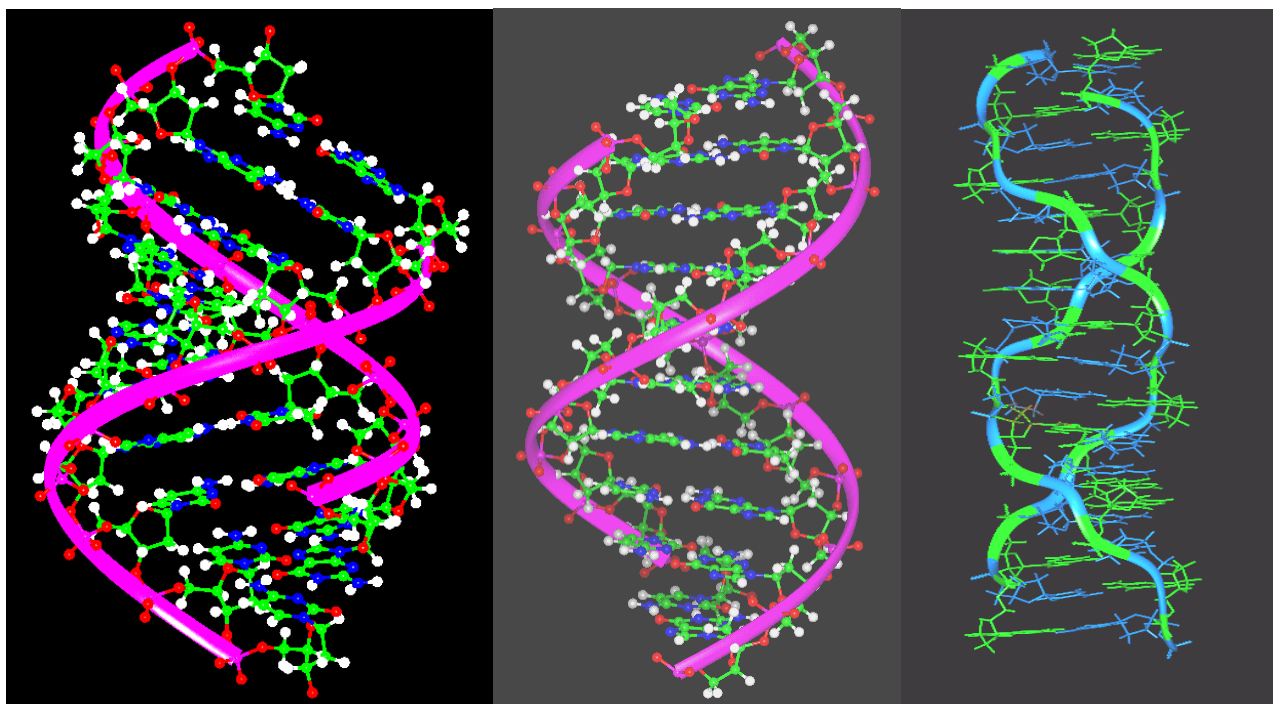

Figure 31

16

Comparison of A B and Z type DNA.

Figure 32

Table 6   Comparison of Structural Characteristics of A- and B-Type Polynucle-
otide Double Helices

Family type	A	B
Sugar pucker	$C_{3'}$-endo	$C_{2'}$-endo
Intrastrand phosphate · · · phosphate distance	5.9 Å	7.0 Å
Dislocation of base-pairs from helix axis	4.4 to 4.9 Å in major groove	−0.2 to −1.8 Å in minor groove
Rotation per nucleotide	30° to 32.7°	36° to 45°
Axial rise per nucleotide	2.56 to 3.29 Å	3.03 to 3.37 Å
Base-pair tilt	Positive, 10° to 20.2°	Negative, −5.9° to −16.4°

Figure 33

How different are A, B and Z-DNA? Grossly, these two types of DNA are quite different (**FIGURE 32**). A-DNA is a short stubby helix, while B-helices are rather thin.   A-DNA is underwound with respect to B-DNA, having 11 residues/turn of the helix, while B-form has 10-10.5 **(Table 6)**. The most obvious difference is evident if one looks down the helix axis of the DNA. The pseudo dyad axis of B-form base pairs lies almost exactly on the helix axis, therefore the end-on view of B-DNA shows the center of the "cylinder" of B-DNA filled with the base pairs, and the sugar phosphate backbone meanders around the outside of it (**Figure 33**). By contrast the end on view of A-DNA shows that the helix axis is "hollow". The polynucleotide chains wrap around the axis like a ribbon. Both the base pairs and the sugar-phosphate backbone are driven out towards the periphery of the double helix. Moreover, the groove sizes and widths of A and B DNA are dramatically different.
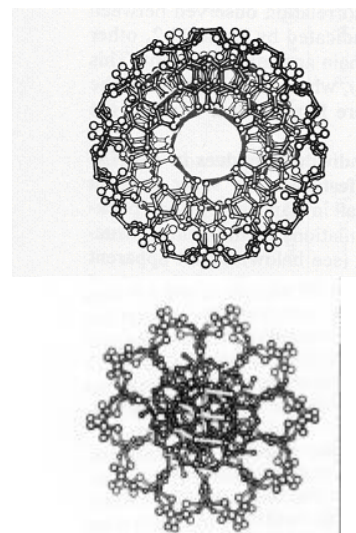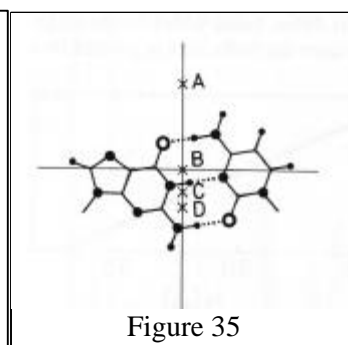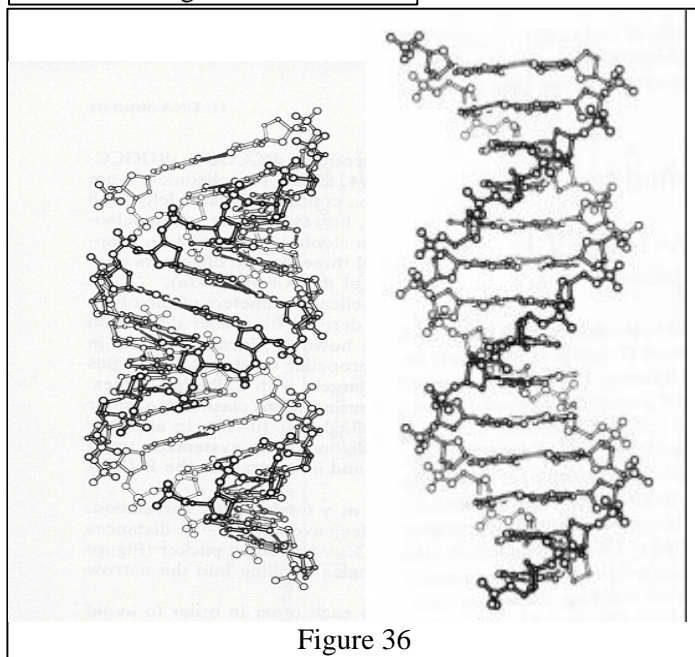
17

How are these difference generated by the microscopic parameters we discussed earlier?

The primary distinction that can be made between A and B type DNA is their differences in preferred sugar pucker and the degrees variability of the backbone torsion angles allowed in each type of DNA. A-DNA contains exclusively C3'-endo type sugar puckers, while B-DNA tolerates the C2'endo family of puckers which includes the lower right quadrant of the pseudorotation cycle, (**FIGURE 6**) from c3'exo to O4' endo. This difference in sugar puckerings causes a variation in the distance between the adjacent phosphates in the same polynucleotide chain; ranging from 5.9 A in C3'endo, to 7.0 A for C2'-endo configurations (**FIGURE 34**). The decrease P-P distance causes the helical rotation of A-DNA to be less than that of B-DNA. This difference has a consequence in determining the helical arrangements of B and A- DNA, such that these helices are MACROSCOPICALLY different.

Figure 34
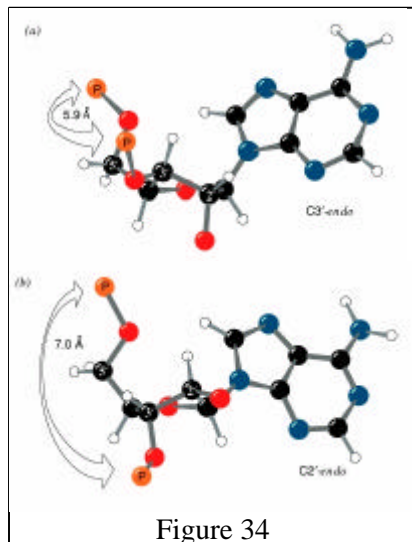
The sugar puckers characteristic of each type of DNA cause the base pairs in the helices to assume different tilt angles with respect to the helix axis. In A-DNA, this tilt, defined as the angle formed between normals to the base pairs and the helix axis is positive in A-DNA, but negative in B-DNA. Looking from the minor groove sides of the bases, tilt in the clockwise direction is positive, and counterclockwise is negative. In B-DNA, the tilt parameter is variable, but is usually quite small, average of about -6°. In A-DNA, tilt is much larger and less variable, $+20^{o}$.

Outside of the sugar puckering modes and its affect on base pair tilt sense and base stacking, the most important feature distinguishing A from B-DNA is the dislocation of the base pairs (D), from the helix axis. As we have already described, the base pair in B-DNA is astride the helix axis, displacement of ~0.2A. In A-DNA, however, the helix axis is pushed far out into the major groove side of the base pairs with D amounting to 4.4-4.9A (**FIGURE 35**).

Figure 35

This gives A-DNA its hollow appearance from the end-on. In A-DNA the displacement of the helix axis into the major groove gives rise to a very deep and narrow major groove and a reduced and relatively shallow minor groove (**Figure 36**). The shallow major groove of A-DNA is only accessible to small molecules, water and metals. In B-DNA the major groove is freely accessible. Since the unambiguous information of a double helix is present in the major groove face of the bases, this limits the utility of A-DNA in storage of regulatory information.

Figure 36