Visit: **http://www.chem.umass.edu/people/cmartin/Courses/MCB642**

Elements of Structure

Polyanion polymeric structure

DNA is a phosphate-linked sugar polymer.

Forces that influence DNA structure

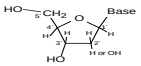
Electrostatics - phosphates, polarizable base and sugar functional groups

Hydrophobic interactions - base functional groups

Hydrogen bonding - Watson-Crick and others... more later...

Torsional constraints - sugar pucker, bond rotations, steric clash

Furanose sugar is primary building block



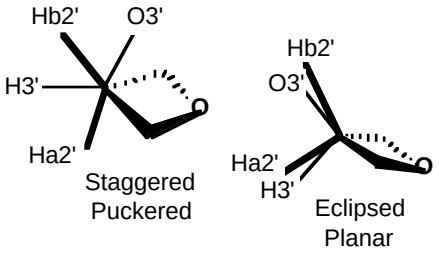
The furanose ring occurs in either the "ribo" form (as in RNA), containing an OH group at the 2'-position, or in the "deoxyribo" form (as in DNA), containing H at the 2' position. In polymeric DNA and RNA the 3' and 5' hydroxyls form phosphodiester linkages. The 3' (and 2') OH extends on the opposite side of the ring (exo) with respect to the base. The 5' CH_2OH occurs on the same side of the ring (endo) as the base.

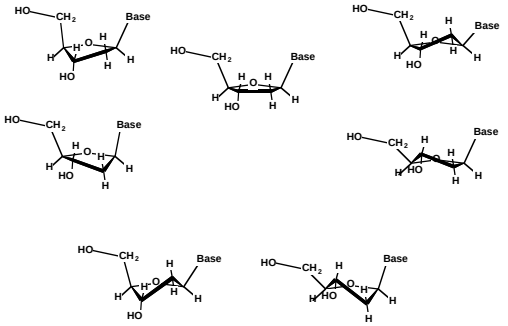
Torsional Constraints

Sugar pucker defines the backbone path (A,B,Z)

Note that the ring is not conjugated and is generally NOT planar. The ring adopts a "puckered" configuration.

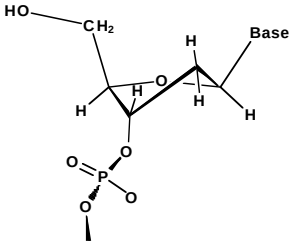
All nucleic acid sugar rings are puckered. The reason for this is that in the all-planar form, there is an eclipsing of substituents on adjacent carbons (see discussions of ethane in standard Organic Chemistry texts) which is energetically unfavorable. The figure at right shows a view down the C2'-C3' bond axis and illustrates how puckering reduces this steric clash. The same effect occurs along the C1'-C2' and C3'-C4' bonds, leading to more complicated puckering (e.g.. C2' endo - C3' exo).



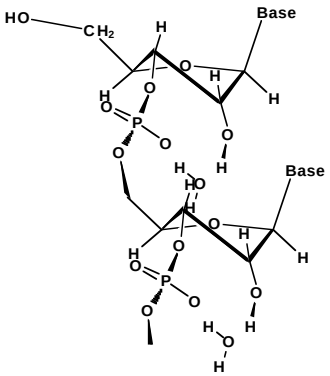


The predominant form in standard B-form DNA is C2' endo (with some C3' exo character). In RNA, however, at least two features combine to favor a C3' endo structure: the 2' hydroxyl proton can form 1) water-mediated hydrogen bond to the 3' phosphate oxygen, and 2) a direct hydrogen bond to the 4' (ring) oxygen on the next sugar (in the 3' direction). As discussed in Saenger (pp64-65), there is some controversy as to whether other factors (such as sterics) play a more important role in RNA's tendency towards a C3' endo configuration.

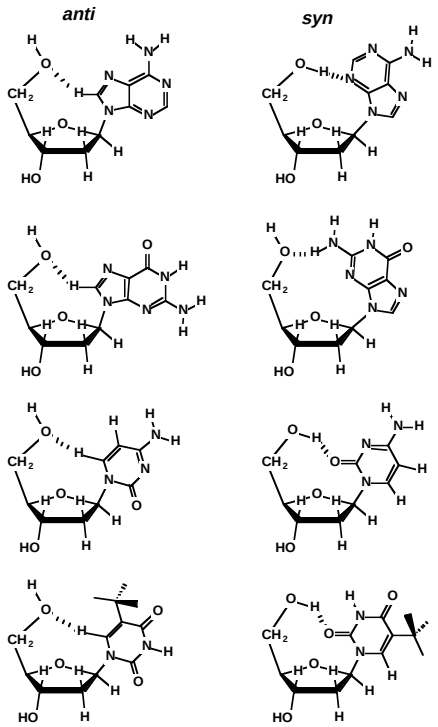
The B-form DNA structure requires a C2' endo sugar pucker in order to place the 5' and 3' linkages at the proper angle with respect to each other. As a consequence, RNA *cannot* form B-form helices!



C2' endo - C3' exo



C2' exo - C3' endo



Note that in the *anti* position, the adenine C8 hydrogen is a H-bond donor. This is unusual for a simple C-H hydrogen, but the ring structure allows electron delocalization, making this hydrogen somewhat acidic (necessary for a good H-bond donor).

Same here for C8 on G in the *anti* position.

Same here for C6 on C in the *anti* position.

Note also that for the *syn* form of G (as for the *syn* form of A above and T below), H-bonding occurs only with a 5' hydroxyl. This is found in *nucleo**s**ides* but not in DNA or RNA (phosphodiester is at 5' end). Therefore, *syn* doesn't occur in nucleo**t**ides, DNA, or RNA.

syn vs. anti

The figures above demonstrate, respectively, *anti* and *syn* configurations of the bases. It also shows the possible stabilizing interactions available. Steric considerations (and those described above) generally tend to favor the *anti* configuration.

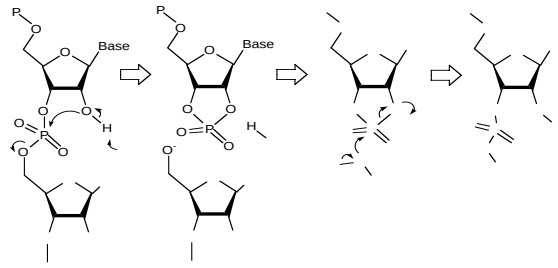
Note that that the *anti* configuration allows the normal Watson-Crick hydrogen bond donors and acceptors to face "out."

Note also that there exists an interplay between sugar pucker (not shown above) and *syn/anti* conformation. It appears that *syn* nucleotides prefer a C2'-endo pucker, to relieve steric clash between the base and the sugar.

Base modifications can sometimes favor the *syn* configuration. Most notably, bulky substitutions at purine C_8 pyrimidine C_6 cause even more steric repulsion for the *anti* configuration and so favor *syn*. This can be an important consideration when employing modified nucleotides such as 8-azido-adenosine.

RNA's extra hydroxyl provides for chemistry - Ribozymes!!

The presence of the OH group at the 2' position in RNA also makes RNA much more chemically reactive than DNA. This has important functional and evolutionary implications.



Less stable to base-catalyzed hydrolysis

Provides for reactivity - ribozymes

(will see importance of this later in the term - evolution)

Also important in that RNA is supposed to exist *transiently* in cells (by contrast, DNA's role in the cell requires that it be very stable).

Summary: All torsional angles interact

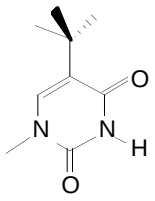
It should be apparent that there is a tight interplay between all of the above torsional angles. Changing one angle, redirects some chemical groups, leading to a need to arrange other angles in order to compensate for new strains introduced.

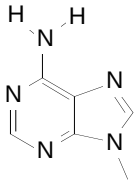
Steric clash is the simplest principle to understand. Be aware that electronic effects also play a major role in determining torsional angles. The substitution of a Br for H at a given position not only increases the steric bulk at that site, but also introduces an electronegative substituent into the system to which it is bonded.

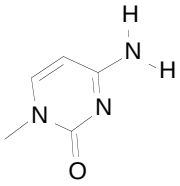
Four base units

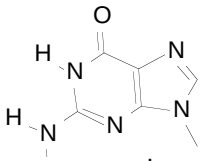
A,C,G,T (U)

modifications arise in nature (and outside of nature)


thymidine


adenine


cytidine


guanine

Base functionalities

Physical Properties

Hydrogen bond <u>donors</u>

careful with angles!!!

Hydrogen bond <u>acceptors</u>

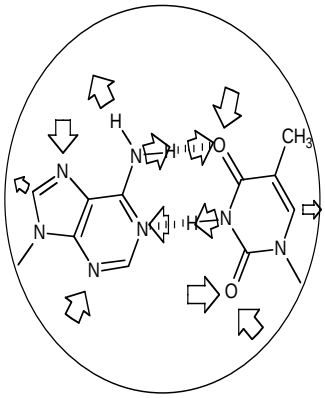
careful with angles!!!

Hydrophobic interactions

Exocyclic amino groups: A,G,C
Protonated ring nitrogens: T,G

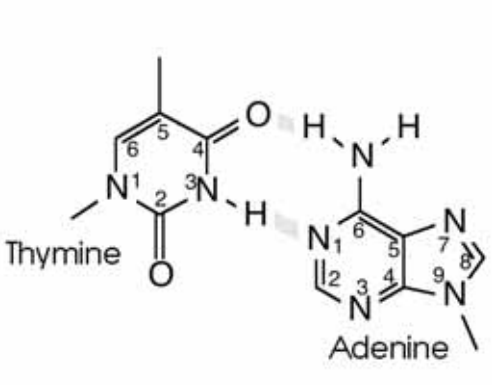
Deprotonated ring nitrogens: A,C,G
Carbonyl oxygens: C,G,T

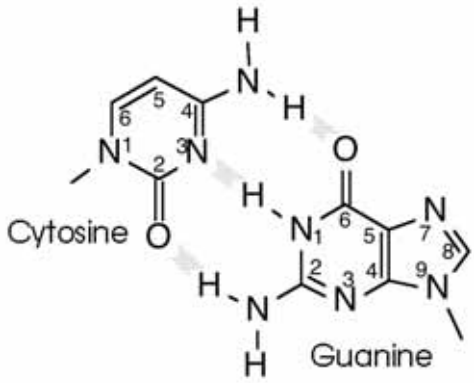
Base aromatic systems
Methyl group: Thymine



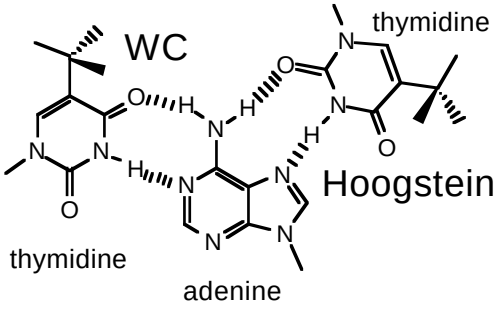
Base pairing

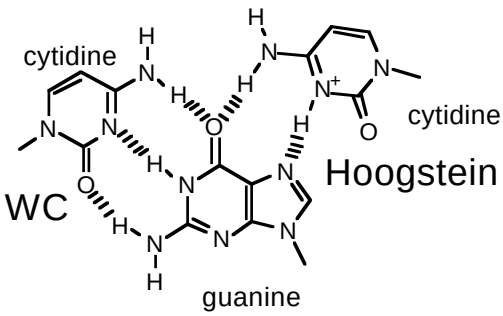
Watson Crick - strong bonds





Hoogstein - triple helices!

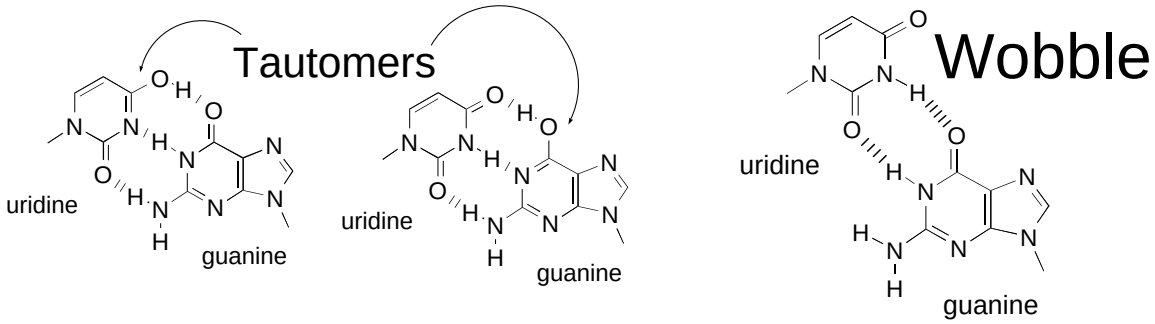




It is not that Hoogstein base pairs don't form strong H-bonds - they do! - but WC pairs fit more readily into the *repeating* structure of the helix. This allows a **cooperativity** of interactions

Hoogstein-type interactions do occur frequently in nature. *In fact, in yeast tRNA-Phe, of the 9 extra-helical base pair interactions which contribute to its structure, only 1 is WC! The rest are Hoogstein.*

Other base pairs occur frequently

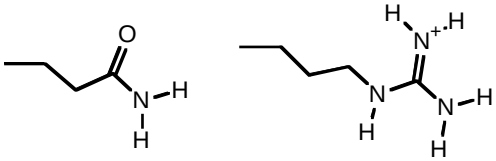


The GU base pair occurs frequently in RNA structures. This requires formation of either a tautomeric form of one of the nucleic acid bases (two structures at left above), or the formation of a "wobble" base pair (structure at right above). In the tautomers, look closely at the bonds...

Amino acid - nucleoside interactions

Hydrogen bond donors & acceptors - Bidentate possibilities

Hydrophobic interactions - Methyl groups on T?



Review of fundamental chemical principles (§ 6)

Bond type	Bond Length (Å)	Bond energy (kcal/mole)	Energy to lengthen by 0.1 Å	Fraction broken at Equilib (298K)
"RT'	-	"0.6"	-	-
C-C	1.54	83.1	3.25	10^{-60}
C-H	1.09	98.8	3.60	10^{-75}
O-H•••O	2.75	3-6	0.1	0.00005-0.005

Hydrogen bonds are typically 20-30 times weaker than covalent bonds.

The **chemistry of water** is dominated by hydrogen bonding

	Melting Point	Boiling Point
H_2O	273 K	373 K
H_2S	190 K	211 K

Water is a very small molecule and on that basis *should be a gas* at room temperature (compare CO_2). But as we know, it has relatively high melting and boiling points. This is due to its ability to form extensive intermolecular hydrogen bonds. By comparison, the sulfur in hydrogen sulfide does not function as well as oxygen in accepting a hydrogen bond; consequently both melting and boiling points are depressed dramatically.

The **hydrophobic interaction** and **hydrogen bonding**

Hydrophobic interactions play a very major role in protein and DNA structure. Unfortunately, these "interactions" are more difficult to "inventory" than are hydrogen bond and covalent interactions.

The major force driving hydrophobic groups inside DNA or proteins is the entropy of the water molecules. In "pure" liquid water, the individual water molecules are free to move around, making and breaking hydrogen bonds as they go. If we stick a benzene molecule into the water phase, those water molecules near the benzene will adopt a very specific orientation around that molecule - thus the entropy of the water has decreased - an energetically unfavorable result.

Hydrogen bond "Inventory" - the ***net*** free energy is the important number.

Watson-Crick Hydrogen Bonds

The *net* binding energies for single H-bonds in DNA duplexes range from 0.8 to 1.6 kcal/mole.

REVIEW / OBJECTIVES

Four bases:

1) Can approach to form multiple H-bonds
2) Donor/acceptor patterns define interactions
3) Watson-Crick pairing allows uniform helix repeat (A,B)
4) Hoogstein allows extra or alternate structural contacts
   - major role in RNA tertiary structure
   - major role in "triplex" gene therapies (small aside)

Sugar backbone

1) Need to stagger along C-C bonds forces puckering
2) Two basic pucker types:
   - C2' exo, C3' endo - RNA / A-form - allows hydroxyl to make more H-bonds
   - C2' endo, C3' exo - DNA / B-form - no 2' hydroxyl
3) Base can be either syn or anti - forces either Hoogstein or WC side of bases out

You should

Be able to draw reasonable Watson-Crick, Hoogstein, and related interactions involving 2 or more H-bonds. Similarly (exact same principles), you should be able to make *multi-dentate* contacts with H-bonding amino acids. Please pay attention to ***angles*** and ***distances*** - a bond 5Å long is pretty useless! Appreciate the interplay between the torsional angles and the central importance of sugar pucker!

Helical forms

	A - Form DNA	B- Form DNA	Z-Form DNA
Diameter	26 Å	20 Å	18 Å
Pitch (rise/turn)	28 Å	34 Å	45 Å
Residues/turn	11	10	12
Axial rise/residue	2.6 Å	3.4Å	3.7Å
Turn angle/residue	33°	36°	30° (average)
Base tilt rel. to axis	20°	6°	7°
Major groove	narrow / deep	wide / deep	"flat"
Minor groove	wide / shallow	narrow / deep	narrow / deep
Sugar pucker	$C_{3'}$ endo	$C_{2'}$ endo	$C_{2'}$ endo / $C_{3'}$ endo
glycosidic bond	*anti*	*anti*	*anti* - $C_{2'}$ endo pyr *syn* - $C_{3'}$ endo pur

Uniform nature of the A and B form helices important to replication, transcription fidelity.

"Real" DNA deviates from ideal structure, result is important to more subtle structural aspects. Some sequence dependence to structural features - very important.

B-DNA ($C_{2'}$ ENDO)

Standard form for most duplex DNA. deoxy sugar favors $C_{2'}$-endo.

A-DNA ($C_{2'}$ EXO)

Standard form for most duplex RNA and for DNA-RNA heteroduplexes. Oxy sugar favors $C_{3'}$-endo ($C_{2'}$-exo).

Z-DNA (alternating pucker!)

pucker alternates, therefore backbone zig-zags
monomer units alternate between *anti* - $C_{2'}$ endo / *syn* - $C_{3'}$ endo.
Leads to a zig-zag path for the phosphodiester backbone.

A and B-Form DNA's present a very uniform charged surface.

Example, the phosphate backbone of B-form DNA presents a very uniform path or surface for proteins which may want to bind the DNA. Great for binding - bad for specificity (mostly). This is probably important in insuring fidelity during transcription and in replication. - Period...

DNA can be *very* long!

	length in kb	length if stretched
bacteriophage λ	49	17 μm
E. coli	4,000	1.4 mm
yeast	13,500	4.6 mm
drosophila	165,000	56 mm
human	2,900,000	0.99 meters!

long pieces of DNA when unwound from their intracellular packing act like long polymer threads - high viscosity. Consequently, when we try to harvest proteins from even *E. coli,* we have to add something to cut the DNA, just to reduce the viscosity to manageable levels.

Forces Stabilizing Nucleic Acid Structure

Duplex melting - a *cooperative* process

Heating of DNA results in disruption of the non-covalent association between DNA strands. The helical structure above disappear and DNA adopts a much more disordered structure. The bonds which hold the duplex together act in a cooperative fashion (forming one makes the next easier to make). Consequently, melting is also cooperative (melting one base pair makes it easier to melt the next). As a result, melting of an entire helix can occur over a very narrow temperature range.

Melting temperatures can depend on the composition of the DNA: high GC content requires higher temperatures to melt the DNA. Of practical importance in PCR and in hybridizations.

Hydrogen bonding

book says it's not important - wrong!

Base stacking and hydrophobic interactions

Ionic interactions - hydrate those phosphates

more important than the book implies

Charge - charge repulsion between adjacent phosphates can contribute to a lessening of duplex stability. Consequently, addition of Na^+ can increase the melting temperature.

However, very high concentrations of salt, occupy all of the water in forming H-bonds to the salt, leaving few waters available to stabilize the phosphates. This can aid precipitation (when water is further removed by replacing it with ethanol, for example).

Modified structures

thiophosphates & H-phosphonates - important to anti-sense strategies

modified bases extend the genetic code - make RNA more like proteins

DNA Substructure - not as uniform as it looks!

DNA "Steps"

Substructure / Sub-"energetics" determined by neighbor-neighbor interactions

Arises from stacking (base-base) interactions. Therefore, an AT base pair can be very different depending on what base pairs are on either side of it.

Breslauer (and others since) have measured energetics of "dinucleotide steps"

Roll, Tilt, Twist, Propeller, Translocation

see overhead...

DNA Bending / Flexure

Substructure can result in inherent bends in the duplex DNA and in *bendability* (which has a directionality to it).

Very important element of recognition for many known protein-DNA complexes...

Classic example: CAP (also known as CRP) transcription regulator.

Protein-DNA Interactions (return to...)

Direct read-out?

Phosphate interactions - non-specific binding or direct read-out?

Many small proteins recognize *palindromes* - why?

Supercoiled DNA

Linking number, twist, writhing L = T + W

Linking number - # of times one strand winds about the other

the only way to change L is to actually cut a strand

Naturally occurring DNA's are underwound (W<0).

Twist - # of complete revolutions one strand makes about the duplex axis

when there is no supercoiling (W=0), then L = T

Writhing number - # of turns that the duplex axis makes about the superhelical axis. Measures the superhelicity.

One can play T & W off of one another ($\Delta W = -\Delta T$). Breaking covalent bonds not necessary.

Supercoils may be Toroidal or Interwound

see overhead

tend to make the DNA more compact than it would have otherwise been.

Intercalators effect

Intercalators force the stacked DNA bases apart. Typically force the duplex to unwind a bit to accommodate, decreasing T (and therefore increasing W). Since W typically starts out negative, increasing intercalator makes W less negative, passing through 0, and then making it again more supercoiled, but in the opposite sense. So EtBr labeled DNA can be more or less compact than the unstained DNA...

Function? The phone cord analogy

Nicking relaxes

Cut one strand of the DNA and it will swivel to release the supercoiled stress.

Topoisomerases

Create ***transient*** breaks in the DNA, allowing L to change.

Type I - create transient single stranded breaks
evidence shows that the enzyme "hangs on" to the 5' end of the DNA through a covalent bond between DNA and the enzyme. Allow the DNA to relax.

Type II (aka gyrases) - create transient double stranded breaks
Create superhelical stress (increase negative supercoiling), using ATP as the driving force.

Very interesting topological mechanisms required.

Sequencing

Restriction endonucleases

recognize palindromes - why a palindrome?

structure of EcoR I

RFLP's (?)

Chemical sequencing - Maxam & Gilbert

Footprinting

Oligonucleotide synthesis

Phosphoramidite method - the most commonly in use. Synthesizer available in the MCB program. Organic-based chemistry is completely different from DNA and RNA polymerases (although in the end, its a phosphotransfer reaction in which the phosphorous is activated to attack by a hydroxyl) - synthesis is in the opposite direction!

Can easily make 50-60mers. Ideally up to 150mer, but yield and purity drop.