

## Chapter 8

### Nucleotides and Nucleic acids

Problems: 1, 2, 6, 7, 9, 11, 12, 14

Also start memorizing base structure and how they H-bond to complement  
There is a short organization homework for this chapter

#### 8.0 Intro

Nucleotides have many roles

- energy currency

- essential chemical link in hormone and external stimulus of cell

- structural component of many cofactor

- ad DNA and RNA

#### 8.1 Basics

Every protein and every RNA in cell is specified by a sequence of DNA

a GENE - the segment of DNA required for functional biological product

1000's of protein so DNA is very large

several classes of RNA

- rRNA - structural components of ribosomes

- mRNA - information intermediate between nuclear DNA and protein synthesis

- tRNA - adapter molecules to connect a 3 letter sequence to an AA

#### A. Nucleotides and Nucleic Acid Structure

Nucleotide- (figure 8-1 a) on board

- Nitrogenous base

- Pentose

- Phosphate

Nucleoside

- Same minus the phosphate

Nitrogenous base Figure 8-1-b on board, figure 8-2 on board

- Pyridine - 6 member ring

  - Cytosine

  - Thymine(DNA)

  - Uracil (RNA)

- Purine - 9 member ring (or a 6 fused to a 5)

  - Adenine

  - Guanine

Some modified bases found (figure 8-5a&b)

- When specifying atoms or groups on rings

- Count around ring using # system fig 8-1

  - 5-methylcytidine

- If talking about a groups on a substituent on the ring

- Name atom with ring # as superscript

### N<sup>6</sup>-Methyladenosine

Base modifications made after base has been incorporated into a DNA or RNA polymer

Usually used for regulation or protection of DNA

Or as structural element in RNA

### Sugar

Closed 5 member ring

D-Ribose (RNA)

Deoxyribose (DNA)

Locating substituents done with ' on # to indicate the sugar instead of the base

So DNA is 2' deoxyribose

Sugars not planar but a slight pucker (figure 8-3 b)

### Phosphate

Generally attached 3' or 5' C of sugar

Via a phosphoester linkage

Can be on other positions

Can even be both

Predict structure of a 2',3'-cyclicmonophosphate

Will see in cAMP and cGMP

Sometimes have multiple phosphates on 5' sugar

ADP, ATP, GDP, GMP

Naming an nomenclature of bases -tides and -sides summed in figure 8-4 and table 8-1

### B. Making the polymer

Link 3' of one sugar to 5' of another through a phosphodiester linkage  
phosphate in this linkage has pKa of 0

so is ionized and has a negative charge at pH 7

make a backbone with a regular negative charge

Anything that binds needs to counteract

+ proteins

+2 metal ions

Polyamines

All linkages are the same

Helps to define a linkage orientation

5' end lacks a nucleotide at 5' position (No further bases on 5' end)

3' end lacks a nucleotide at 3' position (No further bases on 3' end)

Diagram left column page 276 on board

Other groups, often 1 or more P's may be on either 3' or 5' end

By convention single strand written with 5' on left and 3' on right

Other representations

pApCpGpTpA

pACGTA

“Short” nucleic acid oligonucleotide <50

> 50 called polynucleotide

### C. Properties that affect 3D structure

free bases are weakly basic

both bases conjugated ring systems

Resonance structures make rings flat planar

Exist in two or more tautomeric forms depending on pH

#### Figure 8-9

Structures that started with, figure 8-2, are dominant tautomer form at pH 7

Strong absorbance in UV near 260 nm

(Where was protein absorbance?)

Hydrophobic and relatively insoluble in cell

More soluble at high or low pH because push into charged form

Hydrophobic interaction tend to make stack on top of each other

Stack also help van der Waals and dipole-dipole interactions

bases have lots of units that like making H bonds

H-bonds between CG and AT that preserves sugar to sugar

distance are Key to double stranded DNA structure

## 8.2 Nucleic Acid Structure

### A. DNA stores genetic Information

DNA first isolated by Miescher in 1868

Suspected has something to do with inheritance

No proof until 1940's

Avery, Macleod, McCarty showed DNA from one bacteria could transform another

### B. Base Composition (Not a book subheading)

covalent structure understood in 1940's

Chargaff – Chargaff's rules

4 bases, occurring different ratios in different organisms

but same ratio in different tissues of same species

base composition does not change with are or environment

A=T

G=C

A+G=T+C (purines = pyrimidines)

### C. The double helix (Not a book subheading)

Watson-Crick 3D structure figure out 1953

Early 1950's Rosalind Franklin + Maurice Wilkins did X-ray diffraction

Indicated 2 repeating units 3.4 Å and 34 Å Couldn't interpret

1950 Watson & Crick played with known base structures, chemical reasoning, and

Put X-ray data and Chargaff's rule together into the double helix models

#### Figure 8-11 & 8-13

Two strands running opposite directions

Complementary bases

C=G

A+T

Stack hydrophobic bases on inside

Put charged polar on outside

Distance between bases 3.4 Å

Distance from one strand to next 36 Å

Structure allows replication because self complementary

Also note minor and major groove

### D. Other forms

The standard Watson-Crick is called the B form of DNA

See specs Figure 8-17

A form - twist it tighter

This is what you see in DNA-RNA hybrids

A and B are right handed helices

Z form

Left-handed helix

Structure more slender and elongated

Takes special solvent conditions or special sequences

GC or 5methyl GC

Some evidence for short stretches of Z in prokaryotes and

Eukaryotes, but role in cell not known

### E. Unusual structures

Bend in helix when more than 4 A's on one strand

(6 A's make 18 degree bend)

May be important in protein binding

Palindromes

A primary/secondary structure change

Palindrome a word or sentence that is spelled the same frontwards

or backwards

In DNA this means a sequence with 2-fold symmetry over both stands (see figure 8-18)

Note is self complementary

This allows to form cruciforms (cross structures)(double strand or hairpins (single strand) see figure 8-19

Again in vivo implications of cruciforms structures not known

However large amount of DNA is seen in palindrome structure

Often protein dimer binding site same protein binds to both sides

Mirror repeats

Mirror of DNA on same strand

Not self complementary so no hairpin or cruciform

3 or 4 strand structures

Can appear in sites of initiation or regulation of DNA replication recombination or strand separation

But again in vivo implication unclear

Depend on additional base pairing

See figure 8-20 both 3-ple and 4-ple

So can fuse 3 or 4 strands together

These non-Watson-Crick base pairing called

**Hoogsteen pairing**

after discover Karst Hoogsteen (1963)

All of these structure (triple, 4ple, cruciform, Z form) have been see in vitro in DNA sequences involved in regulation of gene expression. It is not known if these structures are actually part of the control mechanism, or whether the DNA is simply bound by protein for control, and they just happen to form these structure in artificial conditions. Watch this space for future developments

### F. Messenger RNA's code for polypeptide chains

DNA largely confined to nucleus  
use RNA copy to transfer information to cytoplasm where make proteins

Three kinds of RNA in cell  
mRNA carries the message  
copying DNA to RNA then correctly processing that RNA into mature mRNA called **transcription**

Prokaryote - single message may code for one or many proteins  
1 protein called monocistronic  
Many proteins called polycistronic

In Eukaryotes mostly monocistronic

Minimum length of mRNA set by protein  
3 bases/amino acid  
usually longer other signals, control processing messages included

### G. RNA's have complex structure

t RNA, different from mRNA from r RNA  
Will look at details in chapter 26  
focus on mRNA from DNA - transcription  
Single stranded, tends to form a right handed helix (figure 8-22)  
Base stacking is dominant force  
Purine-purine base stack stronger than all others Why?  
(Double ring, more surface area)  
Purines will pop pyrimidines out just to do this  
If any self complementarity - will try to form double helical secondary structure  
When helical adopts A form structure  
Have an alternate way to H-bond G and U as well (figure 8-24)  
  
Makes for complex secondary structure with helices and bulges and turns  
  
Going to 3D is difficult Figure 8-25  
TRNA has a 'cloverleaf' secondary structure  
3D structure is an L  
Other RNA structures equally complex

### 8.3 Nucleic acid Chemistry

To have DNA be a stable genetic material, need it not to react  
will examine the chemistry of DNA to see how this might have biological implications

#### A. Double helical DNA and RNA can be denatured

Native DNA highly viscous glob at RT  
heat up or change pH goes through transition where loses viscosity sharply  
just like denaturation of protein, are denaturing DNA  
disrupting H bonds between bases and base stacking interaction  
two strand unwind and can separate  
NO BONDS BROKEN

If a strand haven't complex separated, can anneal quickly & zip back together

if strand separates then more difficult, 2 step process

Step 1 strand find each other

Step 2 strand zip together

In double helical form UV absorbance of DNA at 260 nm is LOWER in double stranded than in free nucleic acids

This is due to base stacking interactions changing electronic properties

Effect called **hypochromism**

Guess what happens when DNA denatures?

Absorbance 260 increases

Called **hyperchromic** effect

Makes an easy way to follow Denaturation-annealing

See figure 8-27a

Midpoint of curve is called  $t_m$  or melting point

Melting point correlates with GC content of DNA

More GC higher melting point 8-27b

If Have DNA in early part of melting curve

Can use EM to see AT loops opening up

Figure 8-28

Can do same for DNA-RNA hybrids

More stable than DNA

DNA-RNA is 20° more stable than DNA-DNA

Don't know why

#### B. Can get hybrids between nucleic acid of different species

Logically this can only occur if sequence are similar

can be used to see if two species are related

can be used to probe for similarities with a given piece of DNA

### C. Nonenzymatic transformation of nucleotides and nucleic acids

a number of very slow spontaneous reaction

But over the course of a lifetime, changes in bases may be tied to mutations, aging and carcinogenesis

All bases undergo deamination (figure 8-30a)

C to U 1 in every  $10^7$ / 24 hours

About 100 events/day/cell

A and G about 1 event/day/cell

Thought to me - why DNA uses T instead of U, can recognize U is an error and correct (mech next semester - book also does a paragraph)

Hydrolysis of base-sugar bond (figure 8-30 b)

Purines faster than pyrimidines

10,000 purines lost/animal cell /day!

UV light dimerizes adjacent pyrimidines (usually T's)

Figure 8-31

Another special repair mech

X-ray and gamma ray

Break open rings

Break covalent backbone

Net UV and environmental ionizing cause 10% of damage due to environment

Reaction that occur with chemicals

Sometimes chemical innocuous by themselves but turned into reactive agent when metabolized by body

Nitrous acid ( $\text{HNO}_2$ ) comes from nitrosamines, nitrite and nitrate salts

Accelerated deamination of bases (Fig 8-32a)

Bisulfite the same ( $\text{HSO}_3^-$ )

Both are used as food preservatives

When ingested this way risk seems minimal

In fact risk from spoiled food causing illness > risk from mutagenesis

### Alkylating agents

Dimethylnitrosamine, dimethylsulfate, nitrogen mustard

(figure 8-32b)

If methyl G to O<sup>6</sup> G can't base pair with C

Similar reaction with SAM already in cell!

### Oxidative damage

From H<sub>2</sub>O<sub>2</sub>, hydroxyl radicals, superoxide radicals

From irradiation or aerobic metabolism

Cell has host of defenses, but some damage still done

### D. Some bases are methylated (details chapter 25)

just aids how some bases methylated unintentionally, actual some bases methylated deliberately

In E coli 2 systems

Restriction modification system

E coli me it DNA at specific sequences using SAM

If finds unmethylated DNA it destroys assuming that is from foreign virus

Dam system

DNA adenine methylation system

GATC methylates A

Part of DNA mismatch repair system

Eukaryotic cells

5% C are methylated usually CpG

Suppresses movement of transposons

May have structural significance Z forms more easily

### E. Sequence Determination

sequencing difficult until 1977

Gilbert & Maxam And Sanger

**Sanger method shown figure 8-33**

Need short primers, labels fluorescently or radioactively

Have small amount of ddNTP

Terminates occasionally

Run out on gel length and reaction mix gives terminal base

Automatic sequencer method shown in **figure 8-34**

Use dideoxy

But have different fluorescent label on each base

Use capillary electrophoresis

## F. Chemical synthesis also automated

### Figure 8-35

Method of Khorana 1970's

Similar to Merrifield synthesis in that add protected base then deprotect

Easily get up to 70-80 nucleotides

This is how get primers for sequencing

## 8.4 Other Functions of Nucleotides

### A. Nucleotides as carriers of E

may add 2 or 3 P's at 5' hydroxyl end of RIBOSE (not deoxyribose)

NTP's of deoxy are seen, but not used as E sources, just as intermediates in DNA synthesis

mono, di tri phosphates

$\alpha\beta\gamma$  position

hydrolysis provides E for other reactions

NOT just ATP, but U, G, and C as well for specific reactions

cleaving ester linkage give about 14 kJ of E

cleaving anhydride linkage gives about 30 kJ/mole

### B. Adenine used a component of many cofactors

Some examples figure 8-38 Co A, NAD<sup>+</sup>, FAD

A not taking part in reaction

seems to be handle to pull cofactor into active site and hold it there

Doesn't seem to be anything special about A, probably just easy for cell since was already making lots of A for ATP

Common protein domain often see nucleotide-binding fold in these protein for binding ATP

### C. Some nucleotides are regulatory molecules

hormone in blood - first messenger

hormone binds to protein on cell surface, protein inside cell starts making a second messenger

Often cAMP Figure 8-39

cGMP also used

ppGpp produced by bacteria during AA starvation

Used as signal to inhibit protein synthesis by inhibiting synthesis of rRNA and tRNA