

## Lecture 1: Nucleic Acids.

### Bases

- Pyrimidines – Uracil/Thymine (RNA/DNA) & Cytosine (U, T, C).
- Purines – Adenine & Guanine (A, G).

#### Chemistry of bases:

- Flat, planar structures because of pi electron structures. This structure means that the bases have a tendency to stack on each other.
- Heterocyclic structures – contains both nitrogen and carbon in the ring.
- Numbering – atom number one is always a nitrogen atom.
- The ring is substituted with keto and/or amino groups.
- Completely uncharged at physiological pH (7), but can be charged at extreme pH because of the basic (alkaline) character of the ring.

#### Watson-Crick base pairing:

- C≡G and T=A or U=A.
- C≡G is stronger than T/U=A because they stack better and have one more hydrogen bond.
- W-C base pairs have the sugar residue of the base in the same relative position, causing GC, CG, AT and TA base pairs to have the sugar residues superimposable on each other. This reserves the regularity of the helix such that the helical backbone is regular and undistorted. Non W-C base pairs can then easily be recognised since these will cause distortions of the backbone.

#### Rare bases:

- Seen much in specific tissues.
- Methylation of C is found in inactive DNA.
- Rare bases are especially found in RNA molecules such as ribosomal RNA and transfer RNA.
- The modification of a normal base to become a rare base always happens after transcription (post transcriptional).
- Examples: 5-methyl-Cytosine ( $m^5C$ ) and 6-methyl-Adenine ( $m^6A$ ).

#### Deamination of bases:

- This refers to the loss of an amino group. All the bases, except Thymine and Uracil have amino groups, but only the deamination of Cytosine gives a base found in DNA/RNA. Deamination of Cytosine yields Uracil, which is found in RNA only. If Uracil is found in DNA it is taken out by Uracil-N-Glycosylase which cleave off the base leaving a sugar with no base on it. This is then removed by AP Nuclease leaving a gap in the DNA strand, which can be sealed by DNA polymerase I. If the Cytosine is  $m^5C$  deamination will yield Thymine which is naturally found in DNA. Methylated Cytosines are therefore mutational hotspots.

### Sugars

- Deoxyribose (DNA) or Ribose (RNA).

#### Chemistry of sugars.

- Form glycosidic bond between C1' of sugar to N1 of pyrimidines or N9 of purines. This bond is stable in alkali but can be broken down by strong acid.
- The 5 membered sugar ring can adopt 4 different conformations – 4 different ways of having four atoms in a plane and one above/below the plane. The two atoms with freedom to move above or below the plane are the C2' and the C3'. These can adopt either endo or exo conformation being on the same or opposite side as the C5' respectively. The endo conformations are the most stable and thus the only found. The C2' endo is found in B-DNA and the C3' endo is found in Z-DNA.

### **Sugar + Base**

- This combination is called a nucleoside
- The binding is a glycosidic bond from 1' of sugar to 1N of pyrimidines or 9N of purines.

### **Sugar + Base + Phosphate**

- This combination is called a nucleotide. It is a phosphate ester of a nucleoside.
- Can be Deoxyribo Nucleotides (DNA) or Ribo Nucleotides (RNA).

#### Chemistry of nucleotides.

- The phosphate group is acidic so nucleotides are acids.
- Negatively charged because of charge carried on P. DNA is a nucleotide polymer with the strong charge is packed away from the bases when DNA folds into a helix – it has the negatively charged P on the outside and the positively charged bases in the middle. In DNA as well as RNA the nucleotides are linked by 5' – 3' phosphodiester linkages. This way of linking gives the nucleotide polymers directionality. A DNA sequence is always read from 5' to 3'.

#### Chemistry of nucleotide polymers.

- The polymers are stable. The phosphodiesterbond can be cleaved on either the 5' or the 3' side of the phosphate. This leads to a 3'-P or a 5'-P respectively. When RNA is broken down with alkali (pH >10) a mixture of 2' and 3' phosphate RNA is generated because of cyclic intermediate (DNA is alkali stable). Most other cleavages of DNA/RNA leaves a 5' phosphate.
- Phosphodiester linkage forms the backbone of the polymers.

#### Nucleases.

- Enzymes that digest/cleave the phosphodiesterbond.
- Can be divided into:
  - Endonucleases – cut internally in chain.
  - Exonucleases – 'eats' away from one end.
- The substrate can be either RNA or DNA (ribonucleases or deoxyribonucleases respectively).
- Examples are:
  - Exoribonuclease – mRNA nucleases – degrades old RNA.
  - Endodeoxyribonuclease – AP endonuclease – finds DNA where base is not present and excises the sugar.

### **Lecture 2 The DNA double helix:**

- Exists in two main forms B-form (major) Z-form (minor).
- The base sequences of the two strands are complementary but run in opposite directions. Because new strand synthesis only occurs in the 5' to 3' direction the synthesis of one of the strands must be discontinuous.
- The helix has to be unwound to enable the strands to be separated. This is done by DNA helicases.

#### **Proteins associated with Nucleic Acids.**

Some of the polar groups on the bases are not involved in base pairing. These groups are free to interact with proteins.

DNA helicases. Unwinds the helix ahead of the replication fork. This is an energy requiring process using ATP.

Single Strand Binding protein (SSB). Binds and stabilises ssDNA. Prevent reannealing/ hybridisation of ssDNA.

#### **Supercoiling.**

Also called extra torsional strain. This is the introduction of additional twist in the DNA helix. The DNA double helix is a right-handed helix. If one strand is nicked and wound the opposite direction (anticlockwise) a negative supercoil is introduced. This relieves the torsional strain of the helix and can cause limited regions with distortion of base pairing. The opposite effect is seen if the twist is in the same direction as the intrinsic winding of the double helix. These positive supercoils tighten the structure.

The DNA of most cells is extensively negatively supercoiled. This is essential for the structure of DNA in the cell. This is a high-energy state and it needs to be stabilised. In eukaryotes the DNA is packed into chromatin. The winding of the helix is carried out by topoisomerases. Negative supercoiling is energy demanding and the topoisomerases need ATP to introduce negative supercoils. Since DNA is usually negatively supercoiled the introduction of positive supercoils relaxes the molecule and they are thus free.

The two enzymes that control the state of the DNA are:

Topoisomerase I – This is the omega factor in *E. coli*. They are ATP-independent and remove negative supercoils (introduce positive). They cut a single strand and attach the 5' end temporarily to the enzyme while rotating the nicked strand around the sealed. The strand is resealed after rotation.

Topoisomerase II – This is the enzyme DNA gyrase in *E. coli*. Cleaves both strands of the DNA double helix and passes another part of the helix through the cut. When negative coils are introduced ATP is required but when they are removed ATP is not needed. This enzyme works ahead of the replication fork and transcribing RNA polymerases to remove positive supercoils introduced by the unwinding of the helix.

#### Other forms of torsional strain.

- Planar hydrophobic molecules like ethidium bromide can intercalate between the base pairs of the helix. This increases the torsional strain and base separation. If the DNA molecule is circular there is a limit to how much ethidium bromide dye that can intercalate because of the torsional strain. If the DNA is linear it has free ends and DNA can accommodate much more dye than if circular. Since the density of the DNA decreases as more dye is bound a mixture of linear DNA and circular DNA can be separated on CsCl density gradient. When DNA is isolated from bacterial cells the bacterial chromosome usually breaks into linear pieces but the plasmids are usually intact because of their small size.
- The dyes have mutagenic effect if used in vivo.

#### Lecture 2: Recombinant DNA technologies.

The use of these techniques depends on the use of restriction endonucleases. These are part of the restriction modification system discovered when someone tried to cross infect bacteriophages and found that some phages will not grow in bacteria depending on its previous host.

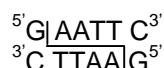
#### Restriction endonucleases.

These are mainly grouped into 3 classes (I, II and III), but more classes exist. Type II is the only of interest to molecular biology since they cut DNA within specific sequences.

Type II restriction modification system.

- The system consist of an endonuclease and a methylase/methyltransferase that methylates the recognition sequence to protect it from restriction.
- Require  $Mg^{2+}$  for restriction. Can be inhibited by chelating agents like EDTA.
- Cut at palindromic/dyad symmetric sequences.
- Recognition sequence is 4 to 8 nucleotides long. An endonuclease recognising 4 bases will generally cut the DNA in more places than one recognising 8 nucleotides. This is due to the statistical occurrence of a short sequence is bigger than that of a long sequence.
- When the DNA is cut it is left with either blunt or a sticky/cohesive ends. The sticky ends can have either 5' or 3' overhang.

Example: EcoRI



Create a 5' overhang.

- Always leaves a 5' phosphate.
- Cleave both strands of the helix.
- Have uncomplicated protein structure.
- The nicks can be sealed with DNA ligase.

Isoschizomers.

- These are restriction enzymes that cleave the same sequence. For example: SmaI (CCCIGGG) and XmaI (CICCGGG).
- The size of the recognition sequence determines how many times the enzyme cut. For example HpaII (CCIGG) will cut more frequent than SmaI and XmaI.

### Separation of DNA.

DNA can be separated according to their size by their mobility through a matrix. The matrix can be made of either agarose which have a 10 – 15 kbase resolution or polyacrylamide which can separate DNA with only one base difference in length.

Large DNA molecules (like bacterial chromosomes) can be separated on pulse field gel electrophoresis.

### Restriction mapping.

A piece of DNA can be cut with different restriction nucleases and then run on a gel to separate the DNA pieces according to their size. This can give their relative lengths, which can be compared to the lengths obtained using another nuclease and a combination of the two nucleases. This can give information about where the restriction sites are on the genome if the nucleases do not cut to many times. If the restriction enzyme cuts more than once each piece obtained should then be cut with the next restriction enzyme and run on the gel again.

An easier way of dealing with enzymes that cut more than twice is to radiolabel one end of the DNA and then use a low concentration of restriction enzyme or a low activity nuclease to give a partial digestion. This will give DNA pieces of differing lengths but only the ones with a radiolabel on will be visible on the gel.

### Lecture 3: DNA manipulating.

Creation of recombinant DNA requires the joining of DNA from different sources. One of these is usually a vector such as a plasmid or a bacteriophage and the other is the sample of interest. The enzyme DNA ligase is used to join two bits of DNA together.

### Recombinant DNA formation.

DNA ligase requirements:

- Compatible DNA termini. Either two blunt ends or two complementary sticky ends.
- A 5' mono phosphate and a 3' hydroxyl group, preferably on both the ends of the two DNA molecules that are to be joined.
- High concentration of both types of DNA favours intermolecular ligation, whereas low concentration of DNA will favour intramolecular ligation (religation).
- Energy source. In T4 and eukaryotes DNA ligase uses ATP as energy source, bacterial DNA ligases use NAD<sup>+</sup>.

To avoid religation of the vector the vector can be phosphatase treated. The phosphatase enzyme cleaves off the two 5' phosphates leaving nothing for the DNA ligase reaction. Religation of the vector is thus avoided but if another piece of DNA comes along with two 5' phosphates it can be inserted. The molecule thus created will have nicks in each strand where the vector 5'-OH meet the other DNA's 3'-OH but these will be sealed when the vector plus DNA is reconstituted in vivo.

Modification of incompatible ends.

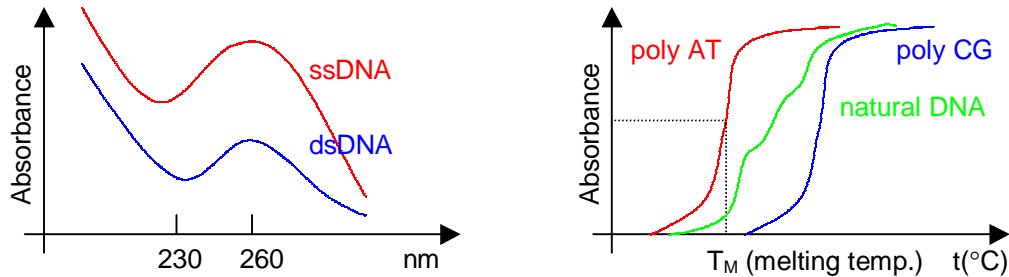
- Cohesive ends to blunt ends. For this DNA polymerase I large sub fragment (the Klenow fragment) is used. This have 5' to 3' polymerase activity and 3' to 5' exonuclease activity.
  - If there is a 5' overhang a new strand is synthesised using pol activity.
  - If there is a 3' overhang this is trimmed using exonuclease activity.

- Blunt ends to cohesive ends. This is done using oligonucleotide linkers with one end sticky and one end blunt. These are ligated onto the blunt ends using high concentration of DNA ligase.

### Melting/denaturing of DNA.

The major stabilising factor on the DNA helix is the stacking of the base pairs. The hydrogen bonds between the bases are of less significance.

The melting of DNA can be followed by the absorption at 260 nm. At this wavelength ssDNA absorb better than dsDNA.



AT rich sequences melt easier than CG rich sequences. This was found using synthetic poly AT (red) and poly CG (blue) DNA. Natural DNA has a much more complex melting pattern (green). In the electron microscope DNA that is partly melted have bubbles of non-base paired DNA – this is the AT rich regions. The reason why CG melt last is that is that CG base pairs stack better and have more hydrogen bonds.

### Reannealing.

Reannealing must be thermodynamically favourable to take place. The creation of order has to be paid for by the interactions between the two strands. Reannealing is performed at high concentration of DNA and neutral pH. For reannealing to take place complementary sequences must find each other, this has a statistical importance since DNA sequences that are repetitive will have more partners and thus reanneal first. This can be used to find the **complexity** of a given DNA sample the more complex the DNA the longer time it will take for it to reanneal. It can also be used to separate the highly repetitive DNA from moderately repetitive and single copy DNA by stopping the reannealing at different time intervals.

When performing a reannealing experiment the temperature and the ionic strength of the reannealing solution determines the **stringency** of the reannealing. The stringency is high if only strictly complementary sequences are allowed to hybridise. If matching but not complementary sequences are allowed to reanneal the stringency is said to be low.

Reannealing can be followed by:  
(physical properties)

- Changes in UV absorbance – see figures above.
- Chromatography on hydroxylapatite.
- Increased viscosity of the solution.

(enzymatic susceptibility)

- Resistance to single-strand specific nuclease.

## Lecture 4: Kinetics and uses of reannealing.

### Reassociation kinetics.

Reannealing obey second order kinetics. This is because the rate-limiting step is the coming together of the two strands. After the two complementary strands have met and formed the first base pairs the rest of the hybridisation is fast. The reaction is as follows:



The backward reaction does not take place. By integration of the rate equation for this reaction the following is found:

$$C_0/C = 1 + k C_0 t$$

In this equation the  $C_0 t$  it takes for the DNA to reanneal can be used as a measure of the complexity of the DNA – it takes longer time for single copy DNA to find its complementary strand and this is also dependent on the total concentration of DNA.

The rate of reannealing depends on:

- The concentration of DNA.
- The complexity of the DNA. This is defined as the number of base pairs in non-repeating sequences. The complexity is inversely proportional to  $k$  (the rate constant).
- The single stranded length.
- The cation concentration.

From the above equation it can be found that when half the DNA have hybridised ( $C/C_0 = 1/2$ ) the factor  $C_0 t_{1/2}$  is equal to  $1/k$ . By comparing  $C_0 t_{1/2}$  values of new genomes with that of genomes of known complexity it is possible to find the complexity of the new genome. (Genes VI pp 648).

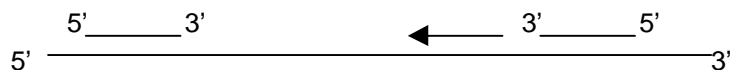
Uses of reassociation kinetics.

- Repeated sequence analysis. This is used to plot  **$C_0 t$  curves** from which  $C_0 t_{1/2}$  can be found.
- Can determine [DNA] if the complexity of the DNA is known.
- Or the complexity can be determined if the [DNA] is known.
  - These are both under defined conditions of strand length and cation concentration.

### Uses of reannealing.

(same numbering system as in lecture note book, see p23)

1. **Detection and isolation of repeated sequences** by separating ssDNA from dsDNA at intermediate stages of the reannealing process.
2. **Quantitation of a particular sequence** using radio labelled probe specific for this sequence. This involves DNA/DNA reassociation.
3. **Measurement of transcript abundance** using DNA/RNA annealing. It is used that RNA will hybridise specifically to its DNA template. DNA/DNA hybridisation will compete with RNA/DNA hybridisation so the RNA is run on gel and then immobilised on a membrane (Northern blot). Or the hybridisation is done under conditions where RNA/DNA hybridise but DNA/DNA does not. DNA This method can measure down to one transcript per cell.
4. **In situ hybridisation** can be used to detect specific sequences on chromosomes, in phage plaques, bacterial colonies, tissue cross sections or intact embryos. Can be used to screen for specific recombinant DNAs.
5. **Hybridisation to blots** can be used to build restriction maps of sequences present as one part per million and for analysis of RNA transcripts. A sample of either digested DNA or cellular RNA is separated by agarose gel electrophoresis and then blotted onto a filter membrane. In Southern blots the sample is DNA and in Northern blots the sample is RNA.
6. **Polymerase Chain Reaction (PCR)** can be used for selective amplification of a gene or part of a gene. Two probes are used both have their 3' hydroxyl ends pointing towards each other. This means that when DNA polymerase is added it will extend the 3' end of



the strand that is in the right direction (opposite of the template strand) until it meets the 3' of the other probe where it will stop. The 3 stages of PCR are melting, annealing (of probes) and elongation. The process can give up to  $10^6$  molecules of DNA in 20 cycles, but the process have a high mutation rate since thermostable bacterial polymerase is not perfect and no DNA repair system is added.

7. **Blocking of gene expression** by engineering cells to express antisense RNA that will hybridise to naturally occurring RNA to prevent translation. The development of the tomato that does not go soft.