SYNTHETIC LIFE SL2

CHAPTER 1

OLIGONUCLEOTIDES

Part 1 – modified canonical nucleobases

Canonical nucleobases and Watson-Crick pairing in DNA

PCR – Polymerase Chain Reaction
*PCR – Polymerase Chain Reaction*

Thermus aquaticus is a thermophilic bacteria from hot springs in Yellowstone Park. 70°C – optimum, living range: 50-80°C. It is a source of thermostable enzymes.

Taq polymerase withstands denaturing conditions (hot temperatures) detrimental for most enzymes. Activity optimum: 75-80°C, half-life at 95°C > 2.5 h.

1990 – Kary Mullis optimized the PCR technique with Taq polymerase (1993 Nobel Prize).

---

**PCR – Polymerase Chain Reaction**

We begin with a single molecule of DNA. We can melt the DNA (break the hydrogen bonds holding the helix together) by heating it to 98 degrees.

Two DNA primers (18-22 bp, T_m: 50-60°C) are designed to anneal to a known sequence. The primers are separated in the sequence that we are targeting by a few hundred base pairs. Cooling the reaction from 98°C to a more moderate temperature allows annealing to take place.

Now we have two primed templates. With dNTPs and DNA polymerase in the reaction mixture, new DNA is synthesized.
PCR – Polymerase Chain Reaction

The DNA is molten for another cycle. Because there is a vast molar excess of primers, when we cool the mixture, we again anneal primers.

PCR – Polymerase Chain Reaction

New DNA is synthesized.

PCR – Polymerase Chain Reaction

In the next cycle, we begin to see DNA molecules whose ends are defined by the primers.

PCR – Polymerase Chain Reaction

After many cycles of melting, annealing, and replication, the overwhelming majority of DNA molecules in the mixture have ends defined by the primers.
**DNA sequencing**

Electropherograms are commonly used to sequence portions of genomes.

An ABI PRISM 3100 Genetic Analyzer. Such capillary sequencers automated the early efforts of sequencing genomes.
Sanger sequencing

What good is dideoxyribose?

**deoxyribose**

**dideoxyribose**

Sanger sequencing

Fluorescent dideoxynucleotides

standard chemically cleavable (reduction)

J. Ju *et al.* PNAS 2005, 102 (17), 5926-5931

Sanger sequencing

**primer**

5’ 3’

TAGCT AGCATAGGCGGCTGCTCTTT

3’ 5’

DNA template + primer
DNA polymerase
dNTPs

Sanger sequencing

ddATP  ddTTP  dddCTP  ddGTP

**double stranded DNA**

**Hydrogen bond**

**Backbone**

**purine Nucleotide**

**pyrimidine**

**Covalent bond**

**STOP**

J. Ju *et al.* PNAS 2005, 102 (17), 5926-5931
Sanger sequencing uses ddNTPs (dideoxynucleotide triphosphates) which do not have a free 3' OH mixed in with dNTPs. Whenever the DNA polymerase incorporates a ddNTP, it won't be able to add any other nucleotides. Then gel electrophoresis is used to separate the DNAvv.

https://www.youtube.com/watch?v=ONGdehkB8jU (from 0:50)

Canonical nucleobases and Watson-Crick pairing in DNA

Purines

Pyrimidines

http://example.com
**Hoogsten base pairing of canonical DNA nucleobases**

Figure 3.24 Hoogsten base pairs. Hoogsten base pairs use a different edge of the purine from a Watson-Crick base pair.

H. E. Moser, P. B. Dervan Science 1987, 238, 645-650

**Modifications of nucleobase structures tolerated by polymerases**

Chemical structures of common modified bases generated by DNA methyltransferases.

S-Adenosylmethionine (SAM)


K. Ishikawa et al. DOI: 10.1093/dnares/dsq027

**Modifications of nucleobases**

**Restriction modification system**

“Immune system” of bacteria and archaea against attacking viruses
EcoRI – a typical restriction enzyme

Products of restriction enzymes

Structure of the homodimeric restriction enzyme EcoRI (cyan and green cartoon diagram) bound to double stranded DNA (brown tubes). Two catalytic magnesium ions (one from each monomer) are shown as magenta spheres and are adjacent to the cleaved sites in the DNA made by the enzyme (depicted as gaps in the DNA backbone).

Non-canonical fluorescent nucleobases

a. Nucleobases composed of aromatic hydrocarbons. b. Nucleobases composed of planar heterocyclic fluorophores. The lack of hydrogen bonding and weaker π-stacking are compensated by versatile energy states brought from the heteroatoms. These fluorophores contribute to a broader spectrum of emission wavelengths. Functional groups can be added to expand functionality, such as metal binding.


Non-canonical fluorescent nucleobases

c. Examples of nucleobases based on hydrocarbons. d. Photoreaction of adjacent phenethynylpyrene nucleobases yields a colour change in emission. The left image shows phenylalkynylpyrene excimer emission whereas the right image shows pyrene monomer emission, both excited at 360 nm. e. C-glycosidic nucleobases based on known fluorophores. f. Simple heterocyclic nucleobases used in the detection of DNA repair activity. g. Nucleobase pairs based on shape complementarity. Although they lack hydrogen bonding, the conformation of these bases counterpart each other, thus forming unnatural base pairing.

Strategies for incorporating fluorescent nucleobases

<table>
<thead>
<tr>
<th>Labelling methods</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Direct chemical synthesis</td>
<td>Site-specific incorporation at any position</td>
<td>High cost on preparative synthesis scales</td>
</tr>
<tr>
<td></td>
<td>Little or no constraint on fluorescent structure</td>
<td>Requires access to DNA synthesizer</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-100 nt or less in length</td>
</tr>
<tr>
<td>Post-synthesis modification</td>
<td>Site-specific incorporation at any position</td>
<td>Limited structural diversity available</td>
</tr>
<tr>
<td></td>
<td>Less expensive than direct synthesis</td>
<td>May require challenging purification</td>
</tr>
<tr>
<td>Enzymatic incorporation</td>
<td>Low cost</td>
<td>Some constraints on positional labelling</td>
</tr>
<tr>
<td></td>
<td>Access to labelled DNAs/RNAs - 100-1000 nt in length</td>
<td>Fluorescence structure limited by enzyme controls</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Base-pair choices limited</td>
</tr>
</tbody>
</table>

Fluorescent nucleobases for studying DNA and RNA

a, Fluorogenic sensing of a demethylation enzyme, ALKBH3. The emission of pyrene is initially quenched by the positive charge of 1-methylated adenine (m1A). When ALKBH3 demethylates m1A, the quenching effect is removed and a signal is generated.

b, Fluorogenic analysis of adenine-to-inosine RNA editing enzyme. The emission maxima of the thiolated adenine (thA) and inosine (thI) are different. Hence by measuring the intensity of thA and thI at their respective maximal wavelengths, the activity of the A-to-I enzyme can also be measured.

c, Kinetic and thermodynamic investigation of the effects of mercury on DNA metabolism. The fluorescent thymine can chelate mercury with another thymine ring and link DNA strands. This can be used to probe mercury metabolism in vivo and to study the effects of mercury on DNA status.

d, Visualization and analysis of human concentrative nucleoside transporters (hCNTs). The fluorescent nucleoside can enter the plasma membrane through the transporters, thus allowing the measurement of the transport activity. TGF-β1, transforming growth factor β1.
DNA Polymerase is evolutionarily optimized by mutagenesis to accept fluorescent Cy3 or Cy5 nucleotides instead of simple dCTP. Here, the polymerase replicates a short segment of its own encoding gene. Water/oil compartmentalization allows testing many independent mutations in parallel.

Polymerases (Pol1 (left compartment)) that are capable of utilizing Cy5-dCTP are able to replicate, i.e., produce "offspring", while polymerases like Pol2 (right compartment) that are unable to utilize it disappear from the gene pool.

380 selected variants from round 1 were screened by Polymerase-ELISA and ranked for their ability to incorporate 4 consecutive Cy5-dCTPs or Cy3-dCTPs. Polymerase-ELISA identified 4 mutant polymerases with significantly enhanced ability to incorporate either Cy3-dCTPs or Cy5-dCTPs compared with wild-type Pfuexo-

- A23 (N400D, I401L, R407I)
- 55 (N400G, R407I), and in particular

The polymerase fidelity in the selected mutant „E10“ was not significantly compromised as compared to the starting Pfu DNA polymerase (Pyrococcus furiosus)

Pyrococcus furiosus – an extremophilic Archaeon from marine sediments Optimal life temperature 100°C
CyDNA – synthesis and replication of highly fluorescently-labelled DNA

(a) Organic phase partitioning of CyDNA is shown for Cy3-DNA (left) and Cy5-DNA (right). Essentially 100% partitioning occurs in the presence of 150 mM NaCl (the yellow color of the phenol phase is due to addition of 8-hydroxyquinoline to prevent oxidation).

(b) Agarose gel electrophoresis of CyDNA restriction digests. Restriction endonucleases are sensitive probes of noncanonical DNA conformations such as those which occur under torsional strain. Both Cy3- and Cy5-DNA are resistant to cleavage by the restriction endonuclease DdeI (C'TNAG) but are cut by MseI (T'TAA). This indicates that at least the local regions of AT-sequence in Cy-DNA adopt a canonical B-form conformation.

Why are A, C, G and T the letters of genetic alphabet.

Nucleotides - components

RNA – most likely evolutionarily older ("RNA World") than DNA → prebiotic origin of ribose + A, C, G, and U nucleobases
Prebiotic synthesis of nucleobases

**Purines**
- Adenine
- Guanine

**Pyrimidines**
- Cytosine
- Uracil

Prebiotic synthesis of purines

1960 - Oró’s synthesis of adenine 2 from hydrogen cyanide 1 and ammonia (general acid–base catalysis, presumed to operate in most steps, is only shown once).

- Heating ammonium cyanide at 70°C for a few days → 0.5% adenine
- Heating HCN with liquid ammonia in a sealed tube → 20% adenine

The photochemical shortcut discovered by Ferris and Orgel is shown by the red arrow. Optimized yields – up to 20% for adenine, 3% for guanine.

Eutectic freezing (-20°C) increases the yield of DAMN formation by concentrating HCN between pure ice crystals.


Prebiotic synthesis of pyrimidines

Cyanoacetylene is a major product of electric discharges in the mixture of nitrogen and methane.

Cyanoacetylene incubated with saturated solution of urea yields up to 50% cytosine. Other methods typically yield up to 5% cytosine. It is further converted to uracil by hydrolysis.
Formose reaction in presence of borates

The aldol chemistry of sugars and cyanide chemistry of nucleobases can be combined at earlier stages than glycosylation.

Cyanosulfidic chemistry

Photochemistry of beta-ribocytidine-2',3'-cyclic phosphate 1. Under conditions of irradiation that destroy most other pyrimidine nucleosides and nucleotides, 1 undergoes partial hydrolysis and slight nucleobase loss. Ura, N1-linked uracil; Cyt-H, cytosine; Ura-H, uracil.


Cyanosulfidic chemistry

Photochemistry of beta-ribocytidine-2',3'-cyclic phosphate 1. Under conditions of irradiation that destroy most other pyrimidine nucleosides and nucleotides, 1 undergoes partial hydrolysis and slight nucleobase loss. Ura, N1-linked uracil; Cyt-H, cytosine; Ura-H, uracil.


