Lecture 36: Reverse Transcription and PCR

Generation of dsDNA from HIV viral RNA (Step 1):
Reverse transcriptase has the following activities (p66/p51)
- RNA dependent DNA polymerase activity
- RNase activity, degrades RNA template.
- DNA dependent DNA polymerase activity.

HIV virus uses tRNAs for primers, we will use synthetic DNA primers, complimentary to the known sequence of the HIV genome

1A. Generate 1st DNA strand

DNA primer I

RNA template

(10 kb)

partially digested vRNA

cDNA (copy DNA)

1B. Generate 2nd DNA strand.

DNA primer II

~600 bp

HIV protease coding sequence

dsDNA (HIV genome)

PCR = Polymerase Chain Reaction (Step 2):
- PCR is a method of amplifying a user-selected segment of DNA within a larger DNA molecule.
- The location of the amplified segment is defined by two primers, they anneal to their templates according to W-C basepairing rules and initiate polymerization from those sites. The primers are DNA and synthesized chemically.
- The final product contains the DNA between the two primers and the primers themselves.
- Any non-complementary bases at the 5' end of the primers are also incorporated into the PCR product.

The actual reaction mix contains:
- Double stranded DNA template
- Excess "right" and "left" primers
- dNTPs
- Thermostable DNA polymerase

The last item on the above list is quite important, since each cycle of PCR requires a heating step to denature the newly synthesized DNA and to allow the primers to bind. Therefore the polymerase must be able to withstand high temperature.

The amplification cycle [DNA amplification = DAP] consists of the following three steps, repeated 15-25 times. After the third cycle, each cycle doubles the amount of PCR product.
- Denaturation of double stranded DNA (T=95 C) [Denature]
- Annealing of primers on each strand (T<Tm) [Anneal]
- Primer extension with DNA Polymerase + dNTPs (T ~ 78 C) [Polymerize]
Primer Design:
1. Sufficiently long that it will hybridize (bind via Watson-Crick H-bonds) uniquely to the desired location.
   - Left primer — sequence of the top strand, at the boundary of the region to be amplified. This primer will anneal to the bottom strand.
   - Right primer — sequence of the bottom strand, at the boundary of the region to be amplified. This primer will anneal to the top strand.

2. Sufficiently long that it will give annealing temperatures in the range of 50-60 C, e.g. if $T_m = 60$ C then anneal at 55 C.
   \[ T_m = 81.5 + 0.41(\%GC) - 675/N \]
   (675/N accounts for the dependence of melting temperature on length):

3. May contain additional sequences at the 5’ ends of the primers that will result in the incorporation of restriction sites, often including additional bases to facilitate binding of the res. endonuclease to the DNA.

Example — Simple amplification with no additional 5’ sequences.

Left (upstream) primer:

Right (downstream) primer:

Cycle 1:

Cycle 2 (after denaturation step) (Only considering the new DNA strands that were made in cycle 1).
Amplification of HIV protease.

The desired PCR product encoding the HIV protease gene is:

5’ CCGTATACGGCGGCTGATCTGTTGGGAA........TTTAATTTCCGATCTCCG 3’
3’ GCCTTAGGCTGATCTGTTGGGAA........TTTAATTTCCGATCTCCG 5’

EcoR1 ProGlnIleThrLeuTrpGln----------LeuAsnPhe BamH1

The sequences at each end, containing GAATTC and GGATCC are not part of the HIV genetic material. These DNA sequences were added during the PCR reaction and contain restriction endonuclease sites that will be used to precisely insert the HIV gene into the expression plasmid.

C-G-A-T-T-C-HIV--G-G-A-T-C-C-C-G

We would design primers that are the sequence of the start (top strand) and end (bottom strand) of the HIV protease DNA coding region, with an EcoR1 site (GAATTC) on the left primer and a BamH1 site (GGATCC) on the right primer. The EcoR1-containing left primer would prime DNA synthesis using the lower strand as the template and the BamH1-containing right primer would prime synthesis using the upper strand as the template.
Left Primer Sequence: 5' CGGAATTC CCTCAGATCAC 3'
T_M = 81.5 + 0.41(675/N) = 81.5 + 0.41(100*10/19) = 67.6

Right Primer Sequence: 5' CGGGATCC GAAATTAAAGTA 3'
T_M = 81.5 + 0.41(675/N) = 81.5 + 0.41(100*8/21) = 65.0

What annealing temperature would you use with these primers?
(These T_Ms apply to the entire sequence – the T_M values for the regions that are complimentary to the HIV coding region are lower. Hence you may want to use a lower T_M for the first few cycles.)

The final product, after PCR.

PCR – Showing how the 5' sequences on the end of the primers are incorporated into the PCR product (note that only selected templates are followed in this diagram, as indicated by the curved arrows in the margins)