Lecture 12 - Polymerase Chain Reaction (AMG text pp. 143-151)
September 27, 2001

Biochemistry of the Polymerase Chain Reaction

The logic of the Polymerase Chain Reaction (PCR) protocol follows directly from well-understood principles of nucleic acid biochemistry. The basic components of a PCR reaction are:

- one or more molecules of target DNA
- two oligonucleotide primers
- thermostable DNA polymerase
- dNTPs

This reaction mix is repeatedly heated and cooled a total of 25-35 times to produce a greater than million-fold amplification of the target DNA.

Each PCR cycle requires three temperature steps to complete one round of DNA synthesis. Before cycle 1 can be initiated, the double-stranded DNA target needs to be heat-denatured to provide single-stranded regions for primer annealing.

In practice, amplifications of a billion-fold amplification are not usually found, why do you think that is?

How would events in early cycles effect the outcome, i.e., primer specificity and enzyme activity?

PCR is extremely sensitive, being able to detect just a few molecules of target DNA in a forensic sample. What type of quality control issues need to be considered in light of this exquisite sensitivity?

Who thought of PCR, it seems pretty simple.
The idea for PCR is credited to Kary Mullis who was a research scientist in the 1980s at a California biotechnology company called Cetus (bought out by Chiron). Mullis, and five other researchers in the Human Genetics Department at Cetus, demonstrated that oligonucleotide primers could be used to specifically amplify defined segments of genomic DNA (or cDNA). Mullis was co-winner of 1993 Nobel Prize in Chemistry. He is an unconventional scientist to say the least!

Michael Smith was the other winner of the 1993 Nobel Prize in Chemistry. He was honored for his discovery of a method to create site-specific mutations in cloned DNA. Michael Smith acquired a small fortune when he sold his interest in a biotech company he started in the early 1980s called Zymogenetics.
The predicted yield of PCR product can be calculated by a simple "investment" equation that incorporates terms for the number of target molecules at the onset, the efficiency of each cycle, and the number of cycles. This PCR investment formula can be expressed as:

\[
\text{PCR product yield} = (\text{input target amount}) \times (1 + \% \text{ efficiency})^{\text{cycle number}}
\]

This equation can be used to calculate that ~26 cycles are required to produce 1 mg of PCR product from 1 pg of a target sequence (million-fold amplification) using an amplification efficiency value of 70% [1 mg PCR product = (1 pg target) \times (1 + 0.7)^{26}].

The temperature profile of a PCR cycle is controlled by the thermal cycler program which results in a near exponential increase in PCR product accumulation for about the first 30 cycles.

PCR target molecules accumulate as a function of cycle number. The exponential phase lasts for about 30 cycles under standard reactions conditions. The plateau phase results from limiting amounts of enzyme and reduced enzyme activity.

**PCR Requires a Thermostable DNA Polymerase**

The first commercially-available thermostable DNA polymerase for PCR came from the thermophilic eubacterium Thermus aquaticus. Several other thermostable DNA polymerase have been isolated and characterized which offer several advantages for specialized PCR assays.

In the 1960s, microbiologist Dr. Thomas Brock (then at Indiana University) traveled to Yellowstone National Park (Wyoming) to research heat-loving microorganisms called thermophiles. He discovered a bacterial species that thrived in high temperatures. Twenty years later, scientists at Cetus Corporation realized that the DNA polymerase from Thermus aquaticus would solve their problem of denatured enzyme after each cycle (Mullis added fresh E. coli DNA polymerase enzyme after each cycle). A number of other thermophilic organisms have since been discovered and exploited for PCR applications.

Taq DNA polymerase has the same overall structure as most DNA polymerases, but their must be some differences, right?

*How would you determine the biochemical basis for thermal stability of Taq polymerase?*

**Design of PCR Primers**

The amplification product of a PCR reaction is defined by the sequence of the PCR primers. The primers anneal to complementary sequences on the DNA template and thereby determine the boundaries of the amplified product.

Once a DNA target has been chosen, there are several rules of thumb for primer design that are important to consider. These general rules are:
Primer length: Choose primers that will anneal to complementary sequences that are 18-24 nucleotides long.

Duplex stability: Both primers in a PCR reaction should have similar melting temperatures (Tm) to ensure that they will have the same hybridization kinetics during the template annealing phase.

Non-complementary primer pairs: The two primers cannot share complementarity at the 3’ ends or else they will give rise to primer dimer products.

No hairpin loops: Each primer needs to be devoid of palindromic sequences which can give rise to stable intrastrand structures that limit primer annealing to the template DNA.

Optimal distance between primers: This is very application specific, but for most diagnostic PCR assays, it is best when the opposing primers are spaced 150-500 bp apart.

Computer programs are often used to predict suitable PCR primer pairs for a given DNA target. The coding sequence of the hypothetical AMG gene was analyzed for optimal PCR primer pairs using the MacVector analysis program (Oxford Molecular Group). A composite map of representative primer pairs for one region of the AMG sequence is shown below.

Primer sequences of individual oligonucleotides, predicted melting temperatures (Tm) of primers and PCR products, and predicted annealing temperature for optimal PCR amplification. The computed information for the F1-B1 primer pair is shown below.

On-line primer design algorithms are available - let's test them! AMG target sequence to be analyzed:

Copy and paste into query windows of the following two applications:

Stanford University’s Web Primer primer analysis application
MIT Genome Center’s Primer3 primer analysis application

Did any of the same forward or reverse primers come up with both applications?

Are there any major differences between the default settings of Web Primer and Primer3?

When might you need to design primers to target a specific sequence even if these types of primer analyses did not reveal any “useful” primers?

What type of sequence might make it difficult for the default settings to find a primer pair?
Optimizing a PCR Assay
A critical step in developing a new PCR assay is to optimize the reaction conditions to obtain maximum specificity and sensitivity. The primary reason for these optimization steps is determine what deviations from the "standard" reaction conditions are necessary to promote functional primer annealing and extension.

Agarose gel electrophoresis is used to characterize PCR products from a temperature optimization experiment. This photograph of an ethidium bromide stained gel below illustrates the effect of temperature on product yield and reaction specificity. Primer pairs 1 and 2 generate higher levels of the correct PCR product at 55 oC, whereas, primer pair 3 functions best at 50 oC.

Name 3 examples of applied molecular genetic methods in which PCR has changed the strategy. For each example, describe the "old" method (before PCR), and the new strategy with PCR.

Based on what you know about the Nobel Prize and its implications, do you think that "discovery" of the PCR method deserved the prize? Explain.

What parameter would you change first if your PCR reaction gave too many products? What would you do if the PCR reaction gave very little if any of the correct product (or any other product for that matter)?

What would you change if all the parameters listed above in the table had no effect on your ability to generate the correct product?