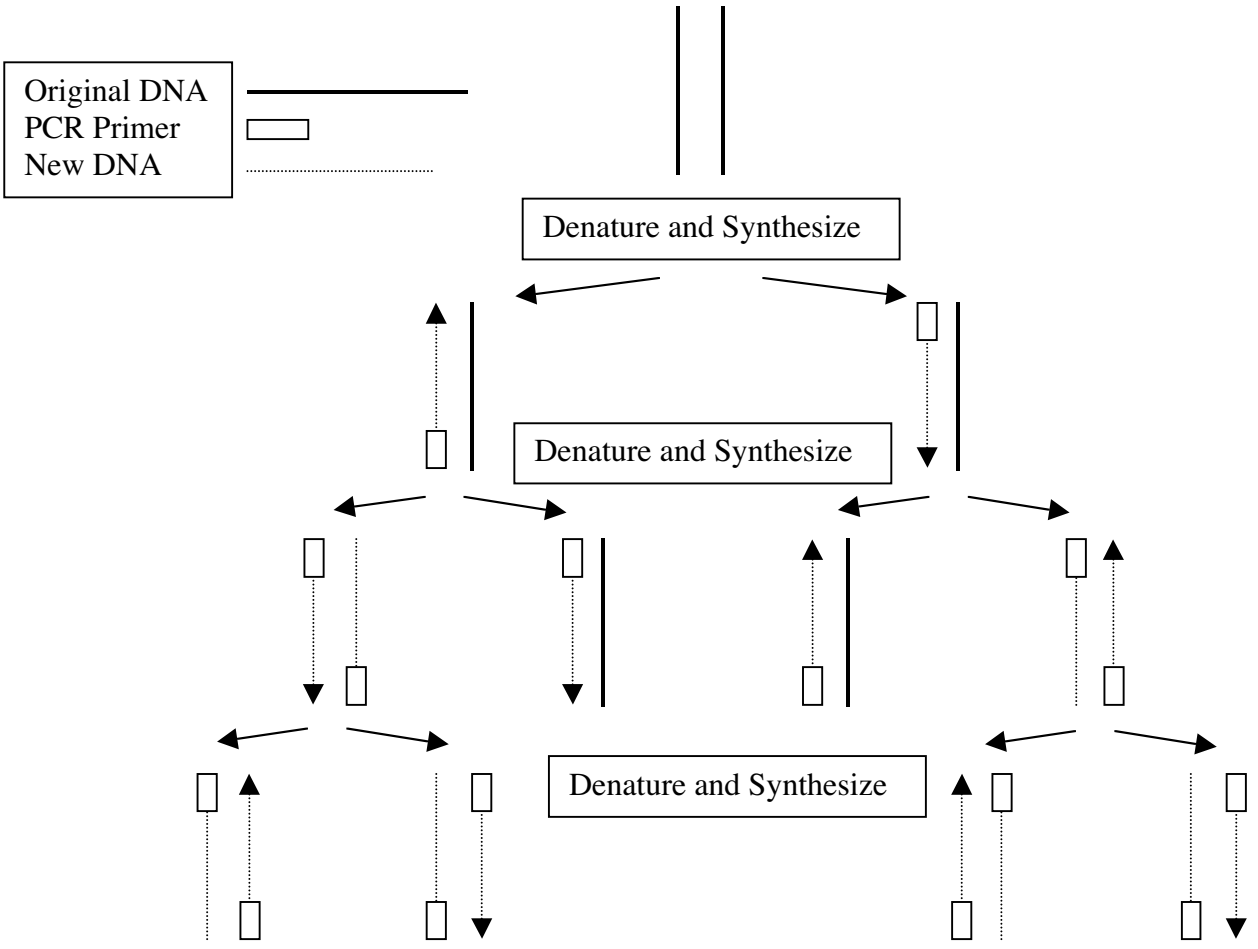
POLYMERASE CHAIN REACTION (PCR)

The polymerase chain reaction (PCR) is a technique used to amplify specific segments of DNA that may range in size from ca. 200 - 2000 or more base pairs. Two recent papers that summarize the development and applications of PCR and discuss its specific applications to Organismal and Population Biology are Mullis. K. Scientific American April 1990 and Arheim et al. BioSciences, 1990, respectively.

Briefly, in a standard PCR reaction, two primers are annealed to the flanking regions of the segment of DNA to be amplified. These primers are oligonucleotides (ssDNA) that are complementary to the two strands and are usually made of ca. 20-30+ nucleotides. The template DNA is first denatured by heating the DNA solution to ca. 94^{o}C. This heating drives the two strands apart. The temperature of the DNA is then dropped to a much lower temperature, such as 55^{o}C (the exact temperature required is determined empirically for each primer pair and template being amplified), to anneal the primers that have been previously added to the DNA solution. After 1-2 minutes, the temperature is elevated to ca. 72^{o}C which is optimal for the polymerization of nucleotides in the presence of Taq DNA polymerase (or some other thermostable polymerase). After this 5' - 3' "extension" step, heating to 94oC again denatures the strands, and the cycle is begun again. Over a number of cycles, this process results in the exponential amplification of the DNA region flanked by the two primers. For example, starting with as little as 10^{-6} mg of DNA, PCR can yield 0.5 - 1 mg of target DNAs, up to 2 kbp in length in 30-35 cycles. After ca. 25-30 cycles, Taq I becomes limited in most reactions.



Problems and Precautions

Since PCR can amplify as little as a single molecule of DNA, problems of contamination become paramount. To minimize the risk of contamination, one must always wear gloves when setting up reactions. Use solutions that are dedicated solely to PCR (it is best to make up solutions and dispense them in small aliquots, for single use only), pipetters and tips dedicated solely to PCR and new glassware and plasticware. To minimize the risk of contamination, many laboratories set up separate rooms to house their PCR machines, or at least, place these in a region of the lab in which no other DNA work is conducted.

Both positive and negative controls should always be run in conjunction with any target reaction. A positive control should include a small amount of the appropriate target sequence and a negative control should include all reaction components except the DNA template.

In addition to potential problems of PCR contamination, variations in sequences obtained from a single PCR reaction occur. This is because Taq DNA polymerase lacks any proofreading activity and therefore occasionally makes an error during template extension[2]. Both transitions and transversions occur, but not large deletions, mosaics or insertions. The error rate is ca. 2×10^{-4} nucleotides per cycle which provides a ~0.25% error frequency. This error rate is ca. 4 times higher than found when the Klenow fragment of *E. coli* DNA polymerase is used. The error rate is known to increase in higher concentrations of dNTP and Mg^{++}.

The problems of PCR errors are not significant if an entire PCR product is used as a hybridization probe or template for direct DNA sequencing. However, if these products are cloned and the cloned DNA is sequenced any sequence obtained should be a consensus sequence from at least 3 different clones.

PCR Components and Parameters

Primers: Primers used in PCR reactions are synthesized oligonucleotides that are at least 16 nucleotides in length and optimally 20-24. When designing primers, one must know enough about the sequence of the region of DNA flanking the segment to be amplified to design complementary primers. Generally these are designed to have a GC content similar to the template and to avoid secondary structures resulting from long stretches of polypurines or polypyrimidines. For some applications, the primer should be exactly complementary to the template. In others, for example, where one is studying a gene that has evolved in various organisms, intentional mismatches are incorporated in the sequence. In all cases, the primer is added in a concentration of ca. 1 μM per reaction. This is enough primer for ca. 30 PCR cycles.

To assist in subcloning PCR products, endonuclease restriction sites are often incorporated into the 5' end of the primer. These sites should not be found in the region of DNA to be amplified.

---

[2] several of the newer thermal-stable polymerases such as some of the vent polymerases have copy correction characteristics.

PCR Buffers: A standard reaction buffer used in PCR consists of

50 mM KCl
10 mM Tris buffer (pH 8.3 at room temperature)
1.5 mM MgC12

We will utilize the buffers supplied with each specific polymerase. Be sure to record the components of each buffer that you utilize.

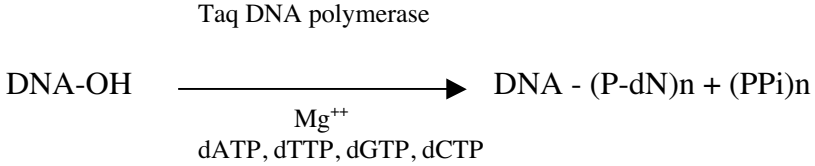
The elevated pH is critical because when the reaction is incubated at 72°C, the pH will drop to 7.2. It is critical that the deoxynucleotides added to the reaction mixture be precisely at pH 7.5.

The concentration of $MgC1_2$ is also critical since this divalent cation is required by Taq polymerase. It is known that neither Mn^{++} or Ca^{++} is able to substitute for this specific cation. Because the concentration of magnesium is so critical, care must be taken to minimize the presence of substances that chelate cations. For example, most DNA is stored in TE that contains 1mM of EDTA, a powerful cation chelator. Template DNA should be stored in TE that has at the most, 0.1 mM of EDTA. In addition, any substance with a negative charge, such as phosphates will effectively remove any available Mg^{++}. Major sources of phosphates are the dNTPs, so the concentration of Mg^{++} must be greater than the concentrations of the dNTPs. Any change in their concentration affects the available Mg^{++} concentration.

Deoxynucleotides: A standard reaction mixture should include ca. 200 μM of each deoxynucleotide (dATP dGTP, dCTP and dTTP). This concentration is enough to synthesize 12.5 mg DNA if only 50% of the nucleotides are incorporated. As mentioned above, the amount of deoxynucleotides affect the available Mg^{++}. Increasing the deoxynucleotides concentrations above 200 μM leads to an increase in the error rate of Taq DNA polymerase.

Target DNA: The DNA to be amplified can be either double- or single-stranded DNA. If very large genomic DNA is to be amplified, it works best to cut the DNA with restriction endonucleases such as Not I that cuts the DNA only rarely. Amplification occurs more efficiently using linear templates rather than closed circular DNA. The concentration of DNA is also critical. Generally one "brackets" the concentrations using 1 ng, 0.1 ng and 0.001 ng of template (target) DNA.

Taq DNA Polymerase: This enzyme catalyzes the following reaction:



Two forms of the enzyme are available. The native enzyme is isolated from the thermophilic bacterium Thermus aquaticus and a genetically engineered form, synthesized by recombinant *E. coli* . Both have 5' - 3' polymerization dependent exonuclease activity and lack 3' - 5" nuclease activity. Generally <2 units of Taq polymerase is used per reaction. An excess of enzyme may result in the amplification of non-target sequences.

As mentioned previously, this enzyme requires Mg^{++}. The melting and annealing temperature required for specific primer template interactions is affected by Mg^{++} concentrations; phosphate buffers and excess EDTA inhibit the activity of Taq polymerase by affecting the Mg^{++} concentration available.

Thermocycler parameters: The exact temperature and time needed at each PCR step must be determined somewhat empirically for each template DNA and primer. Some basic starting points are known.

PCR Denaturation: Generally 90 seconds at 94°C will result in melting apart the strands of dsDNA. It is important that you periodically check the temperature of your thermocycler to assure that it is actually reaching this temperature. Somewhat lower temperatures (i.e. 92°C) may effectively denature short (ca. 200-300 bp) pieces of DNA (the length of the standard Lambda control DNA included in the AmpliTaq kit) but is not hot enough to dissociate the two strands of much longer pieces of DNA. If you do not succeed in amplification, the thermocycler temperature is the first thing that you should check and calibrate.

PCR Primer Annealing: A standard beginning annealing temperature and time to try is 55°C for 1-2 minutes. If your primer has a low GC content (i.e.. <50%), lower temperature may be required. However, lowering the annealing temperature may result in an increase of nonspecific amplification products. If a higher GC primer content is characteristic (i.e., >50%), a higher annealing temperature may be required. There are several formulas that are used to calculate optimum melting temperatures. Often these will give you significantly different predictions. These will be good starting points, but nothing beats the empirical method for optimization of this parameter.

PCR Primer Extension: Extensions are usually carried out at 72°C for ca. 2-3 minutes. The optimum temperature for Taq polymerase is 75°C but this temperature can cause the primers to fall off the template. Primer extension actually begins during annealing since Taq polymerase is partially active at 55° C. The length of time required for primer extension is dependent upon the length of the sequence. A general rule of thumb is Taq polymerase will extend about 1 kb per minute. Longer stretches may require up to 15 minutes. It is best to keep the extension time as short as possible to prolong the "life" of Taq polymerase.

Thermocycler Ramp time: Some PCR machines allow you to set the ramp time, or the time between each cycle. Generally this should be as short as possible. Certain types, notably the "air thermocyclers", have virtually no ramping time. This results in faster overall PCR amplifications.

Optimization of PCR

A. Routine PCR

A typical amplification reaction includes the sample of target DNA, a thermostable DNA polymerase, two oligonucleotide primers, deoxynucleotide triphosphates (dNTPs), reaction buffer, magnesium and optional additives. One cycle of amplification is defined by the series of temperature and time adjustments. Each cycle of PCR after the first cycle theoretically doubles the amount of targeted template sequence (amplimer). Therefore, ten cycles theoretically multiplies the amplimer by a factor of about one thousand; twenty cycles, by a factor of more than one million.

Each cycle of PCR amplification consists of a defined number of reaction steps. The steps are designed using temperature and duration time to denature the template, anneal the two oligonucleotide primers and extend the new complementary DNA strands by polymerization. The target DNA is denatured by heating to 95°C or higher for 15-120 seconds. During denaturation, the two intertwined strands of DNA separate from one another, producing the necessary single-stranded DNA (ssDNA) template for primer annealing and polymerization (extension) by a thermostable polymerase. To anneal the oligonucleotide primers, the temperature of the next step in the cycle is reduced to approximately 40-60°C. At this temperature, the oligonucleotide primers can anneal to the ssDNA strands and serve as primers for DNA synthesis. This step requires approximately 30-60 seconds. Finally, to extend from the primer-bound template DNA, the reaction temperature is raised to the optimum for most thermostable DNA polymerases, which is approximately 72-74°C. Extension of the primer by the thermostable polymerase requires approximately 1-2 minutes and is target length dependent. Extension completes one cycle, and the next cycle begins by returning the reaction to 95°C for denaturation. After 20-40 cycles, the amplified nucleic acid may then be analyzed (for size, quantity or sequence) or it may be used in further experimental procedures (cloning).

B. Magnesium Concentration

Magnesium concentration is a crucial factor affecting the performance of Taq DNA Polymerase. Reaction components, including template DNA, chelating agents present in the sample (EDTA or citrate), dNTPs and proteins can affect the amount of free magnesium. In the absence of adequate free magnesium, Taq DNA Polymerase is inactive. Conversely, excess free magnesium reduces enzyme fidelity and may increase the level of nonspecific amplification. For these reasons, it is important to empirically determine the optimal $MgCl_2$ concentration for each reaction. This is accomplished by preparing a series of reaction s containing 1.5-3 mM Mg^{2+}, in increments of 0.5 mM, by adding 3, 4, 5 and 6 µl of a 25 mM $MgCl_2$ stock to 50 µl reactions.

C. Enzyme Concentration

It is recommended that 1.25 units of Taq DNA Polymerase be used per 50 µl reaction. For most applications, enzyme will be in excess; the inclusion of more enzyme will not significantly increase product yield. Increased amounts of enzyme and excessively long extension times

increase the likelihood of generating artifacts due to the intrinsic 5'-3' exonuclease activity of Taq. Artifacts generally can be seen by the smearing of bands in EthBr-stained agarose gels.

The most frequent cause of excessive enzyme levels is pipeting errors. Accurate dispensing of submicroliter volumes of enzyme solutions in 50% glycerol is nearly impossible. It is recommended that reaction master mixes sufficient for the number of reactions being performed be used to obviate this problem. A master mix increases the volumes of pipeted reagents and reduces pipeting errors.

D. Primer Design

PCR primers generally range in length from 15-30 bases and are designed usually to flank the region of interest. Primers should contain 40-60% (G+C) and care should be taken to avoid sequences that might produce internal secondary structure. The 3'-ends of the primers should not be complementary to avoid the production of primer-dimers. These delete primers from the reaction and result in an unwanted polymerase reaction that competes with the desired reaction. Avoid three G or C nucleotides in a row near the 3'-end of the primer as this may result in nonspecific primer annealing, increasing the synthesis of undesirable reaction products. Ideally, both primers should have nearly identical melting temperatures (Tm); in this manner, the two primers should anneal roughly at the same temperature. In any case, the annealing temperature of the reaction is dependent on the primer with the lowest Tm.

The sequence of the primers can also include regions at the 5'-ends that may prove useful for downstream application. For example, restriction enzyme sites can be designed into the primer pair for ease in downstream manipulations such as cloning. The final concentration of the primer in the reaction must be optimized. It is recommended to add 50 pmole of each primer (1 μM final concentration in a 50 μl reaction) as a starting point for optimization. Generally, nanograms of primer DNA equivalent to 50 pmole are: 16.3 ng time the number of bases in the primer.

E. Template Considerations

Successful amplification of the region of interest is dependent upon the amount and quality of the template DNA. Reagents commonly used to purify nucleic acids (salts, guanidine, proteases, organic solvents and SDS) are potent inhibitors of DNA polymerases. A final ethanol precipitation of the nucleic acid sample will eliminate most of the inhibitory agents.

The amount of template required for successful amplification depends on the complexity of the DNA sample. For example, whereas a 4 kb plasmid containing a 1 kb insert equates to 25% of the DNA being the target of interest, a 1 kb gene in human genomic DNA (genome of 3.3×10^9 bp) represents approximately 0.00003% of the input DNA. Therefore, approximately 1,000,000-fold more human genomic DNA is required to maintain the same number of target copies per reaction.

Two common mistakes when trying to amplify target DNA are using too much plasmid DNA and too little genomic DNA. As a general guide for how much template DNA to use, start with a minimum of 10^4 copies of the target sequence to obtain a signal in 25-30 cycles, but keep the final DNA concentration of the reaction <10 ng/μl.

Conversion of Nucleic Acids from microgram amount to number of molecules:

Nucleic Acid	Amount	# of molecules
1 kb RNA	1 µg	1.8×10^{12}
1 kb dsDNA	1 µg	9.18×10^{11}
pGEM® Vector DNA	1 µg	2.9×10^{11}
Lambda DNA	1 µg	1.9×10^{11}
E. coli genomic DNA	1 µg	2×10^{8}
Human genomic DNA	1 µg	3.0×10^{5}

DNA yields from different human tissue sources:

Source of DNA	Amount of Tissue Used	Typical Yield
Whole blood	30 µl	0.5-1 µg
Cell suspension	5×10^{5} cells	2-5 µg
Hair root	Single root	10-200 ng
Semen	30 µl	5-10 µg

F. Primer Annealing Temperature

The sequences of the primers are a major consideration in determining the optimal temperature of the PCR amplification cycles. For primers with a high Tm, it may be advantageous to increase the annealing temperature. Higher temperatures minimize nonspecific primer annealing, increase the amount of specific product produced and reduce the amount of primer-dimer formation.

Numerous formulas exist to determine the theoretical Tm of nucleic acids and these may serve as a starting point for annealing conditions. However, it is best to optimize the annealing conditions by perfuming the reaction at several temperatures starting approximately 5°C below the calculated Tm. The formula below can be used to estimate the melting temperature for any oligonucleotide:

$$Tm = 81.5 + 16.6 \times (\log_{10}[Na^+]) + 0.41 \times (\%G+C) - 675/n$$

Where $[Na^+]$ is the molar salt concentration ($[K^+] = [Na^+]$) and n = number of bases in the oligonucleotide.