CH 4. CLONING BY PCR – PCR APPLICATIONS

I. SIMPLE PCR

1. Principle

PCR (polymerase chain reaction) is a powerful method of *in vitro* DNA synthesis developed in 1983. PCR is a rapid, sensitive and inexpensive procedure for cloning and amplifying DNA (by selective replication) of specific interest without preparing libraries. The cloned DNA fragment has defined length and sequence. An important feature is that PCR produces large amount of specific DNA molecule from a small amount of template of relatively impure DNA (drop of blood, a hair root or a chunk of bone,…). Consequently, the technique has revolutionized molecular biology and is used in virtually every area of natural sciences and medicine thereby cutting across the boundaries separating basic and applied research, commercial technology and medicine.

The principle of PCR is rather simple and involves enzymatic amplification of a DNA fragment flanked by two oligonucleotide primers (of known sequences) hybridized to opposite strands of the template with the 3'ends facing each other. DNA polymerase synthesizes new DNA starting from the 3'end of each primer. Repeated cycles of heat denaturation of the template, annealing of the primers and extension of the annealed primers by DNA polymerase results in amplification of the DNA fragment. The extension product of each primer can serve as template for the other primer resulting in essentially doubling the amount of the DNA fragment in each cycle. The result is an exponential increase in the amount of specific DNA fragment defined by the 5'ends of the couple of primers.

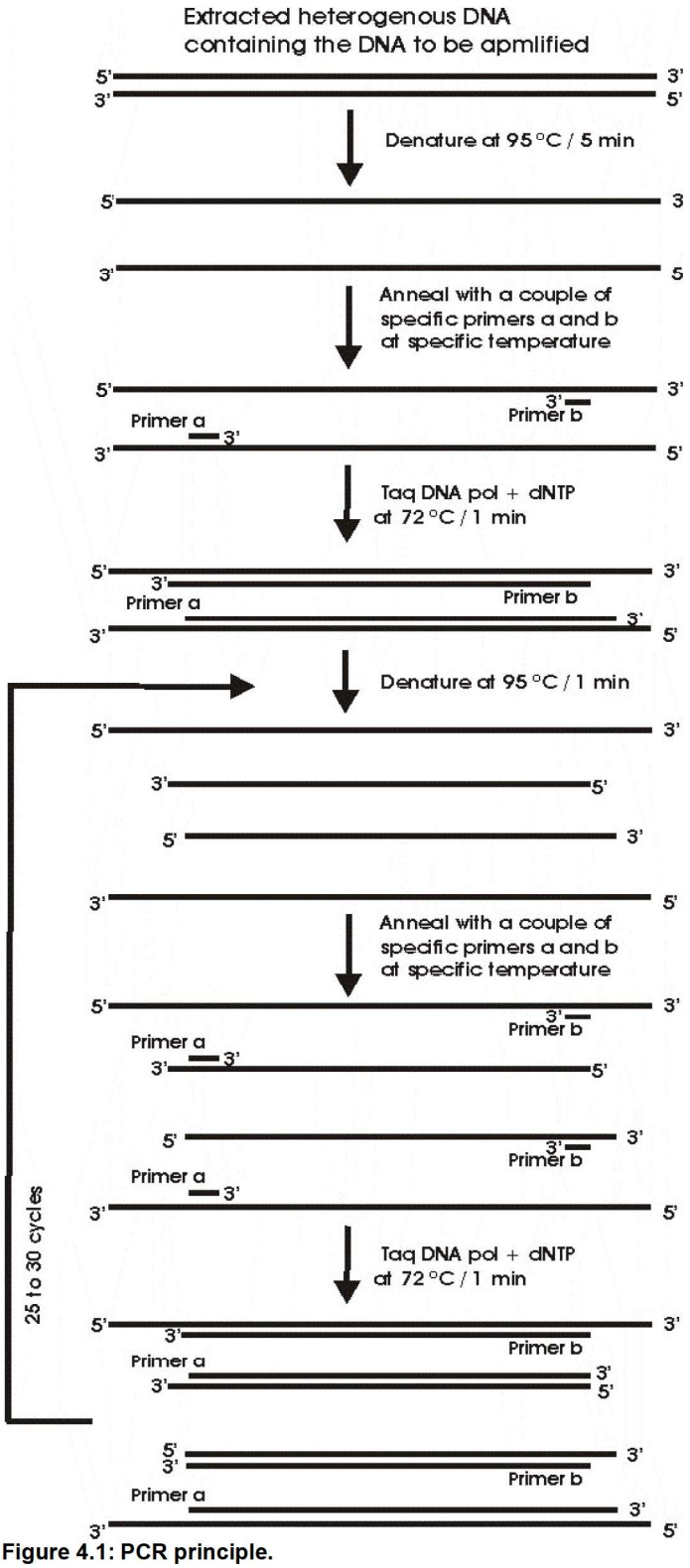

Figure 4.1: PCR principle.

2. Primers

PCR cloning requires a couple of specific primers complementary to the two extremities of the desired DNA molecule which is being amplified. The primers are usually named “forward” and

"reverse" for upstream and downstream, respectively. The sequences at both extremities of the fragment to be cloned must be known in order to design the complementary primers. The sequence of primers may be deduced from partial peptide sequence or from orthologous cloned genes or cDNA (in other species)[32]. Primers should be at least 20 to 25 nucleotides[33] in length to ensure high specificity. However, for some applications requiring multiple random initiation (see RAPD analysis in Ch 7.II. 4.), primers are usually 8 to 10 nucleotides in length.

The distance between primers can vary considerably. However, the efficiency of amplification decreases considerably for distances greater than 3 kb. In general, short fragments amplify with higher efficiency than long ones since the enzyme has more chance to reach the template end.

3. Technique

After DNA extraction, several nanograms of the sample (mixture of DNA) are added by the primers, a thermostable DNA polymerase (Taq DNA polymerase), and the dNTP contained in a $MgCl_2$ buffer. Mg is a cofactor for DNA polymerase. The mixture undergoes a first denaturation step at 95ºC for 5 minutes (Figure 4.1). Temperature is then decreased to 5ºC below primers Tm, and the sample is incubated at that temperature for about a minute in order to allow primers annealing on the denatured DNA. Temperature is raised to reach optimal temperature of the DNA polymerase and the sample is incubated during 30 seconds to one minute according to the length of the amplified fragment. In this step, the enzyme uses the primers and all the necessary components to synthesize the complementary strands. This is the first cycle which is repeated 25 to 30 other times. The following cycles are identical to the first one, except for the denaturation step which lasts about one minute.

After 30 cycles, one gets 2^{30} copies (~10^9 copies) from each DNA copy in the initial mixture. The amplified DNA may be visualized after electrophoresis on an agarose gel and staining with ETB. Size of the amplified fragment is equal to the distance separating primers 5'ends on the template.

These thermal cycles were performed manually at the early beginning of PCR but rapidly programmable thermocyclers were developed and improved. The device is programmed and the cycles are performed automatically and tubes are collected at the end. However, Klenow fragment of E. coli DNA polymerase I used initially is not suitable for automation since it must be added fresh during each cycle because the enzyme is inactivated by each denaturation step. The problem was solved by the introduction of thermostable DNA polymerase from the thermophilic bacterium Thermus aquaticus (Taq), which substantially reduced the work needed to perform PCR.

To verify whether DNA of interest is obtained, PCR products should be characterized by Southern blotting using a specific probe and by DNA sequencing. The sequence can be compared with known or expected DNA sequences in databases.

4. Unwanted Mutagenesis in PCR

Taq polymerase has very high error rates because it lacks 3'-5' exonuclease activity. It incorporates one incorrect nucleotide for every 9000 nucleotides and causes a frameshift once every 41000 nucleotides. In addition, the enzyme adds nucleotides to 3'ends in a template-independent manner, making the amplification product difficult to clone.

If the mutagenized PCR product is used directly for sequencing or as a template for an *in vitro* transcription system then the generated mutations should not have a significant effect because most of the molecules have an unaltered nucleotide at any position. However, if the PCR product is later cloned into a vector, then the probability is greater that the cloned molecule will contain a sequence alteration at other than the desired position.

High fidelity of PCR amplification can be achieved by using thermostable enzymes with proofreading exonuclease activities.

---

[32] To obtain the promoter sequence by PCR, a reverse primer can be designed according to the 5'end sequence of a known cDNA or based on the N-terminal amino acid sequence. This primer will be employed to anneal with denatured genomic DNA. The primer serves for synthesis of a new strand, which is experimentally lengthened by an oligo-dC (TdT activity). Then, PCR is continued with the reverse primer and oligo-dG as forward primer. As a result, major PCR products including the promoter sequence will be obtained. Thus any part of a gene can be selectively amplified even though the gene is not yet cloned.

[33] Caution must be taken during design of primer sequences in order to avoid any pairing of primers together or intramolecular pairing. Moreover, Tm of the primers should be in the same range.

II. RT-PCR

Since its introduction, numerous modifications and applications of PCR technique have been developed such as RT-PCR and real time PCR. RT stands for reverse transcription which is the synthesis of DNA using RNA template by means of a reverse transcriptase. RT-PCR aims to amplify a particular DNA which is complementary to a specific RNA in order to determine its relative amount or for cloning purpose simply. PCR is preceded by a reverse transcription on an RNA population using either an oligo-dT primer (to reverse transcribe mRNA population in eukaryotes) or random decaprimers (to reverse transcribe nearly all RNA classes). It is also possible to use specific primer(s) to reverse transcribe a specific RNA species (type) among a heterogeneous population.

After reverse transcription, specific couple of primers are used during the PCR reaction carried out on the reverse transcription product. The primers can be designed from published sequences (from the same species or from another) of the 3'UTR and 5'UTR, or they can be designed based on the very N-terminal and C-terminal amino acid sequences in case the cDNA sequence is not available. Whenever partial cDNA[34] are desired, the forward or the reverse primer may correspond to the poly-A tail or to a sequence added by TdT on RT products.

RT-PCR products are sorted on a gel in order to determine the relative amount of a particular RNA in the cell (whose cDNA is being amplified). Indeed, it is possible to compare amount of the amplified cDNA between different cell types (or different experimental conditions) since it reflects the initial amount of RNA (Figure 4.2). There are diverse methods to compare the amount of the amplified products; the simplest one is the comparison of bands intensities on the agarose gel but this approach is semi-quantitative, that is it does not enable accurate quantification of the initial target amount in the analyzed sample.

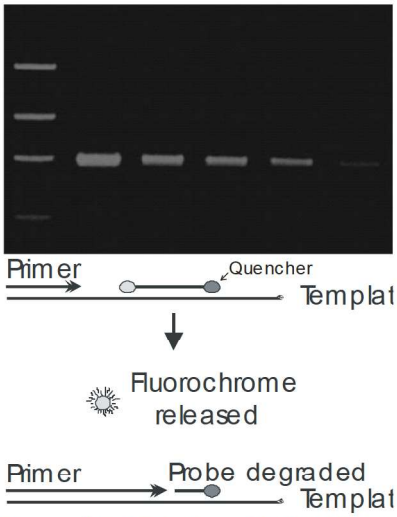

Figure 4.2: Gel of an RT-PCR (top) and principle of real-time PCR (bottom).

III. REAL-TIME PCR

Real-time PCR combines amplification of target DNA with simultaneous detection of amplification product in the same closed reaction vessel. The apparatus that enables real-time PCR is different from a thermocycler of the traditional PCR. For example, it is provided with a fluorescent detection system to follow amplification kinetic. Real Time PCR requires, in addition to primers and all typical PCR reagents, a probe labeled with a fluorochrome coupled with a quencher (e.g. TaqMan probe) and complementary to the amplified fragment near its middle. The fluorochrome is inactive when coupled with the quencher (the probe annealed with its target). The probe annealed to its target DNA is degraded by the 5'-exonuclease activity of Taq DNA polymerase during amplification process which releases a fluorescent molecule (Figure 4.2). Assessment of fluorescence simultaneously to amplification determines the amount of the amplified product. This method is quantitative and much more sensitive than staining with ETB after electrophoresis. It shows the crude result in form of curves of fluorescence intensity as a function of time (or PCR cycles). Fluorescence increase reflects amplification kinetic, so that an earlier appearance of fluorescence means a greater target amount in the analyzed sample. Curves are used to calculate the Ct (cycle threshold) values that help performing accurate quantification of the target amount.

Other probe detection procedures have been adapted to real-time instruments include FRET (fluorescence resonance energy transfer) and molecular beacons. These are not considered in this chapter.

[34] for 5'-RACE, rapid amplification of 5'-cDNA ends, and the 3'-RACE, rapid amplification of 3'-cDNA ends.

In addition to the qualitative detection of target sequences (detection of infectious agents), real-time PCR instruments can quantitate target nucleic acid and differentiate alleles (determine sequence variation or point mutations). For allelic variation detection, probe-target duplexes in the patient's specimen are compared to wild-type probe-target duplexes by melting curve analysis. Probe-target duplexes will melt (separate from each other or denature) at different temperatures for wild-type target versus mutated (allelic) target. To achieve genotyping by melting curve analysis, denaturation of the PCR product (resulting from amplification of the tested locus) is assessed at high resolution (e.g. 0.2 degrees increase) and the resulting curve is compared to reference samples (mutant and normal, Figure 4.3).

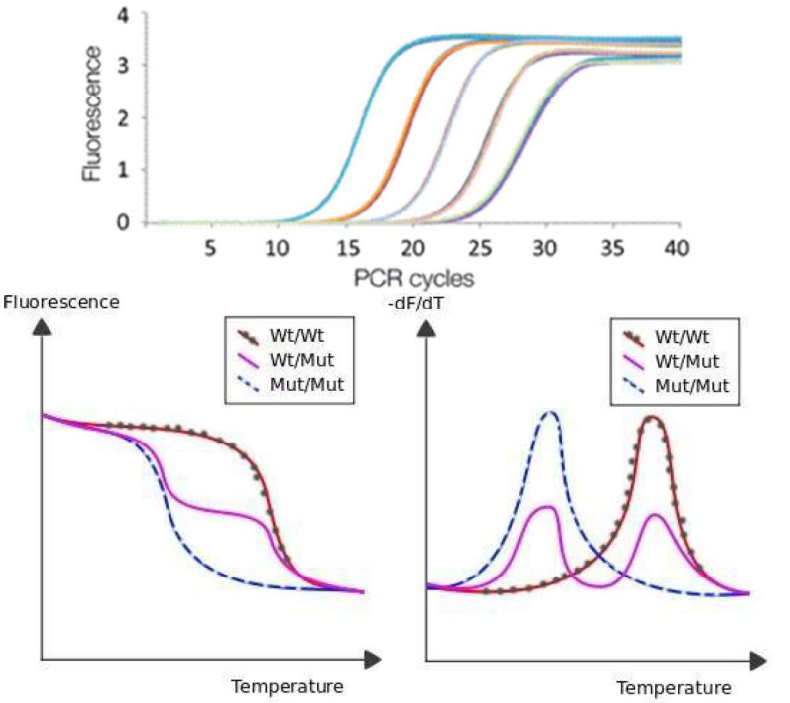

Figure 4.3: Example of real-time PCR result of 5 samples containing different amounts of the target DNA (top). Genotyping by melt curve analysis after PCR (bottom).

IV. APPLICATIONS

In addition to rapidity and specificity of PCR, it has an important advantage which is the high sensitivity and therefore the possibility of amplification and cloning starting from DNA of several cells. This advantage is important since it opens the door for use of PCR in diagnosis and especially in criminology where the available samples are usually very small (a drop of blood, few skin cells, a hair root, …). In another chapter the PCR applications in criminology are considered (see Ch 7.II. 4. b.), here we mention only its applications in terms of basic research and diagnosis. PCR is used for determination of DNA finger print and PCR is of particular use for cloning of ancient DNA starting from dead cells, bone remaining and fossils.

Usually amplification of DNA is not the final aim. In fact, one amplifies DNA fragment in order to determine its sequence and to study it. This could be a study of the encoded amino acids, of the presence of putative binding sites for particular transcription factors, of the regulatory role of the sequences or determination of the presence of specific mutation (diagnosis and genotyping). In addition to diagnosis of genetic disorders, PCR is used in diagnosis of pathogen-caused diseases at an early stage before symptoms appear. For instance, a viral infection could be detected by PCR during the incubation period before symptoms appear. A single copy of the virus present in the collected sample, such as a blood specimen, is theoretically enough to be detected by PCR through the use of specific primers to the viral DNA.