Organizing committee:

Prof (Dr.) Mohan Joseph Modayil
Director, CMFRI, Kochi

Dr. P.C. Thomas
Winter School Director

Co-ordinators:

Dr. R. Paul Raj, Head, PNPD
Dr. K.C. George, Principal Scientist
Dr. P. Jayasankar, Senior Scientist
Dr. D. Noble, Senior Scientist

WINTER SCHOOL

ON

Recent Advances in Mariculture Genetics and Biotechnology

4th to 24th November 2003

Course Manual

INDIAN COUNCIL OF AGRICULTURAL RESEARCH
CENTRAL MARINE FISHERIES RESEARCH INSTITUTE
P.B. No. 1603, Tatapuram P.O.,
Kochi – 682 014

DESIGNING OF PRIMER FOR PCR

P.C. Thomas
Central Marine Fisheries Research institute, Cochin

Introduction

The polymerase chain reaction (PCR) is the technique for *in vitro* synthesis of multiple copies of a given DNA molecule. Primers are the most important components of PCR and success of PCR largely depends on the primers. A primer is an oligonucleotide. It is utilized as the starting point for the synthesis of the DNA fragment in a PCR. Primers bind to its complementary sequences on the DNA template and extend in length by addition of nucleotide to its 3' end. Amplification of specific regions of the template DNA depends on the design of primer sequences.

In a PCR, primers are needed as pairs. They are designated as "forward" and "reverse". These primers are complimentary to the 5'ends of regions on the opposite DNA strands which has to be amplified, so that they can be extended toward one another with DNA polymerase , forming new DNA molecules (indicated by the arrows in the diagram).

Since the primers used in a PCR will largely determine the success of the reaction proper attention has to be paid in its designing. There are many factors that guide the primer design .The following points are to be considered while designing a primer.

Sequence specificity

The primers should be specific to the regions flanking the DNA segment to be amplified . The most important requirement of a primer is its sequence specificity. The primer sequence determines what nucleotide sequence it can anneal to, how firmly it will anneal and how good it will serve as a starting point for the addition of new nucleotides to the primers for its extension into new DNA strand. Specific primer are generally , designed to target a DNA sequence flanking the regions to be amplified.

Primer Length

Primer length is an important factor. In general, a primer in the length range of 17 to 30 nucleotides is considered ideal. Longer primer sequences can be used for enhanced specificity in conjunction with higher annealing temperature. The prime requirement of a primer is that it should be complex enough so that the likelihood of it annealing to sequences other than the chosen target is very low.

For example, there is a 1/4 chance (4^{-1}) of finding a single A, G, C, or T nucleotide in any given DNA sequence. Similarly, there is a 1/16 chance of finding any dinucleotide sequence (ie., AG/AC/AT/ etc.) and a 1/256 chance of finding a given tetra nucleotide sequence (ie.,AGCT/ATCC/ etc.). Thus, a sixteen base sequence will statistically be present only once in every 4^{16} bases (Once in every 4294967296, or 4 billion) which is about the size of the human or maize genome. Thus, an oligonucleotide primer of 17 bases or more is extremely sequence specific. Generally, primers of 17 to 30 nucleotides are routinely used for amplification of specific regions of genomic DNA of animals and plants. Extra long primers may result in mismatch pairing and nonspecific priming even at high annealing temperatures. The optimum length of a primer depends on its (A+T) content.

Base composition of primers

The optimum G + C base composition is in the range of 35-60%. It is preferable that the GC content difference between the two primers is within 5%. It is also desirable to avoid long runs of Gs and Cs in primers. Minor adjustments in the lengths of the primers may be made to compensate the differences in them. The G=C pairing is much stronger than the A=T as it has three rather than two hydrogen bonds between them. Hence a GC pair require more heat to melt than the AT pair. High GC content sometimes leads to primer dimer association even after heating to 95°C and as a result in poor amplification of the desired products.

Annealing temperature

The two primers designated as "forward" and "reverse" should have similar melting/ annealing temperatures so they can both work under the same thermal regimes optimally leading to the exponential amplification of DNA in PCR.

As a rule of thumb the following formula is used for determining the Tm values.

$$\mathbf{Tm}\ (\text{degrees C}) = 4 \times (G+C) + 2 \times (A+T) - 5$$

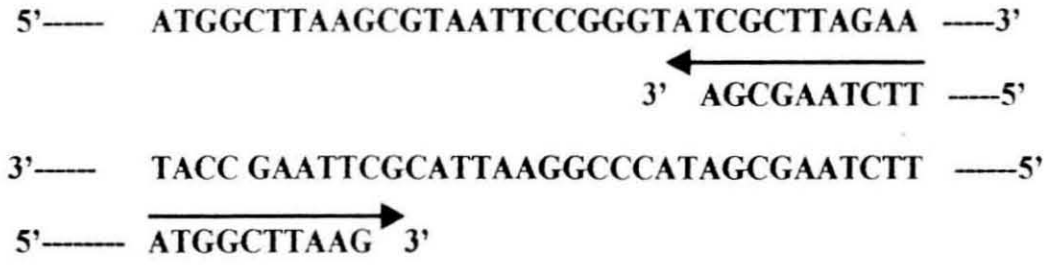
For example the melting temperatures of the following primers work out to be

$$\mathbf{TGGCTTACGAATCGC} \rightarrow 4 \times (9) + 2 \times (7) - 5 = 45°C$$

While designing the primers take care to keep the Tm° of both primers within 1-4°C of each other, and in the 40-65°C range. This formula is applicable only for primers between about 15 and 25 bases length. Other factors like ionic conditions, traces of detergent and solvents, DNA quantity / quality etc.strongly affect actual melting temperature.These factors and time allowed for annealing affect the rate of annealing of primers during PCR.

Primer orientation

If both the forward and reverse primer sequences are designed from the sequence data of coding strand of the DNA molecule, the reverse primer has to match the sequence of the noncoding strand. The reverse primer must be the reverse compliment of the target site of the coding strand from which primer is designed.

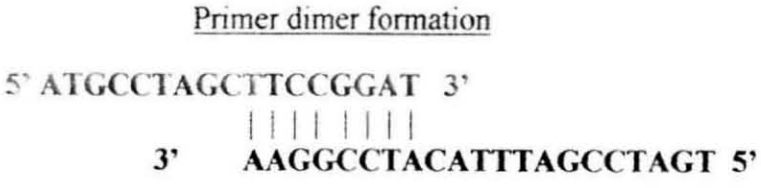


During PCR, the primer pairs bind to opposite strands, and elongation of both primers occur. The forward primer is complimentary to and binds to the non-coding strand, while the reverse primer binds to the coding strand. Not only is the base sequence important, but the 5'to 3' orientation is as well.

Complementarity between primers

The primers should be designed in such a way that there is no complementarity between them. Partial or full complementarity with each other result in the formation of primer dimers (primers binding to each other), thus leaving little of the primers for the priming of the target DNA. Very little primers will be available for target DNA and no/ less amplification product will be formed.

Similarly, care should be taken to avoid complementarity between the 3' ends of the primers. Complementarity between 3' end of primers will also lead to " primer dimer" formation, which inhibits the synthesis of desired PCR fragment.



Self complementarity

In primers of sufficient length reverse-complimentarity between both the ends has also to be avoided especially in long primers as it can loop around and bind to themselves. Like primer dimerisation, this self-annealing will compete with priming of the desired DNA target leading poor amplification.

GC-rich 3′ ends ("GC clamp")

G and C bases have 3 hydrogen bonds and thus bind more strongly than A and T, which share only two hydrogen bonds. Having several G or C at the 3′ end (elongation end) of the primer will make that end more stable and can increase PCR yield.

A/T bases at last codon

Primers should not have T at its 3'end. As T is the least discriminating nucleotide, primers with 3′ T have greater chance of mismatch**.** Further, it is advisable for each primer to have at least one A or T within the last triplet at its 3'to discourage mismatch tolerance of primers with consecutive G's or C's.

5′- End Modification

The PCR primer can work with a poorly matching 5′ end, because the 5′ primer end is not elongated. Therefore, it can accommodate incorrectly matched bases. This is often used for while designing primers for introducing a particular restriction site to facilitate cloning the PCR product.

Primer Degeneracy

When primers are designed from the sequence information of heterologous sources, one cannot often find primer targets that are sufficiently well conserved over a wide enough group of organisms. Thus, we can order a mix of the two (or more) combinations of primer that will cover all sequences. However, as degenerate combinations will greatly increase the chance of priming of an undesirable sequence unrelated to your sequence of interest.

Primer designing guidelines in brief:

1. Primers should be 17 to 30 nucleotides in length.
2. Base composition should be 35-60% (G+C)
3. The 3′ end of primers should have preferably a G or C, or GC or CG.
4. Tm values of primers between 55-65° C are preferable.
5. Runs of three or more Cs or Gs at the 3′ ends of primers should be avoided.
6. Primer orientation should be correct while designing.
7. The 3′- ends of primers should not be complementary to each other.
8. Self-complementarity of primers should be avoided.

Assistance for Primer designing

The primer sequence should be specific to the regions flanking the DNA segment to be amplified. Sequence specificity can be worked out and checked using software programmes. Database which provide sequence information for regions of interest (NCBI: ENTREZ) can be utilized for this. Comparison of sequences can be made with large number of sequences stored in database and similarities can be checked by NCBI: BLAST programme. Many soft wares are available on Internet for designing of primers of a desired sequence.

Internet site(s)

http://biobase.dk/index.html

http://bcf.drl.arizona.edu/gcg.html

http://www.blocks.fhcrc.org/

http://www.biodisk.com/

http://www.oligo.net

http://www.hgmp.mrc.ac.uk/

http://bioinformatics.weizman.ac.il/blocks/index.html

http://alces.med.umn.edu/webprimers.html

http://www.willamstone.com/

http://doprimer.interactiva.de/

http://www.genome.wi.mit.edu/cgi-bin/primer/primer3_www.cgi

http://dot.imgen.bcm.tmc.edu:9331/seq_util/seq_util.html

http://www.med.jhu.edu/medcenter/primer/primer.cgi

Procedure

1. Search National center for Biotechnology Information (NCBI= 'gene bank') for selected /assigned type of sequence / gene family, using the 'Entrez Browser' program. Try various keywords and ' wildcard' search combinations to try to find all known entries of selected sequences/gene family. Also use the "taxonomy" browser to find closely related species.
2. Download the sequence files and clearly label these files

3. Prepare a text file containing only the sequences

4. Remove the nucleotide numbers from the sequence.

5. Submit the edited sequences to the primer 3 server, by pasting the edited sequences into the submission window.

5. Make entries in the parameters boxes such as sequence ID, target sequence and length of amplified product.

7. For initial stages of primer designing, leave other parameters at default settings

8. Run the programme. The output will give several combinations of primer pairs for a particular target sequence. Primer 3 gives the sequences of primers- forward and reverse in their proper orientation.

9.The primers are then tested for the secondary structure formations in software in which this facility is available e.g. DNASIS.