Designing of oligonucleotides: Probes and Primers for diagnosis, epidemiology and research in medicine

Rahul Sharma* and V.M. Katoch*

Detection, identification and amplification of desired gene sequence/fragment are the basics of modern biology. Amplification of the genomic material is required for most of molecular biology and biotechnology related approaches such as molecular diagnosis, DNA fingerprinting, DNA sequencing and gene cloning. Molecular epidemiology has been emerged as a key component of modern epidemiologic research. Genotyping is not only useful for identifying variations among organisms but also host susceptibility markers. Most of these approaches require amplification of the sequence of interest to identify variations among various molecular markers and identification of particular sequence with the help of probes or sequencing.

These advances have helped tremendously in the growth of modern medicine. Development of well defined specific gene probe(s) has been very useful in identification of specific sequences which could be of a pathogen or a mutated/polymorphic host DNA. Among different type of probes, synthetic oligonucleotide targeting specific variable stretches have emerged as most important molecular tool. These tools are not only useful for direct detection by hybridization but are also useful as primers for amplification of desired fragments for restriction analysis, hybridization or sequencing for identifying exact structural differences. All these oligonucleotide act as important keys to open gate for detection, identification and amplification of desired gene targets which can be used for diagnosis of a genetic disorder or infection (its sources and also drug resistance when the target is encoding for drug resistance).

Molecular probes: Molecular probes (Nucleic acid probes) are used to detect sequence of interest in a variety of applications, such as Southern, Northern, dot blots and microarrays. For a sequence to be useful as a probe, it needs to be specific for the target sequence under particular condition. It must not bind to anything else which might be present in the sample being screened. Nucleic acid targeting probes for identification and diagnosis are not very sensitive, DNA targeting probes requires $10^4-10^5$ copies where as rRNA targeting probes are $10-100$ times more sensitive. Designing...
probes is similar to designing primer pairs. In fact, one way to begin is to use some of the same design programmes.

Replication of DNA is naturally occurring process of every living cell. Polymerase Chain Reaction (PCR) is DNA replication artificially done in tube/microtiter plate in which only small region of the DNA replicates many times to generate million folds from a complex mixture of starting material usually termed as “Template DNA” and in many cases require little nucleic acid purification. This is the technique that has replaced in many cases traditional DNA cloning methods for the production of large amount of DNA from limited starting material especially for gene analysis related work.

Core material for various molecular experiments is oligonucleotides to be used as probes and primers for gene amplification. This article gives an overview of technology related to designing and synthesis of oligonucleotide probes. The information needed to produce a PCR primer or gene probe may be explored from many sources but the knowledge from genetic databases like NCBI, TIGR is often the starting material. Softwares developed for these purpose have been listed and discussed so that this review may serve as a ready reference to the readers.

Key to the PCR or gene amplification even by isothermal gene amplification lies in the design of the two oligonucleotide primers. It requires some knowledge of the DNA sequence information which flanks the fragment of

---

**What is PCR?**

- **DNA replication gone crazy in a tube!**
- **Only a small region of the DNA actually gets replicated**

**Fig-1 An over view of PCR**
Designing of oligonucleotides: Probes and Primers

DNA to be amplified. From this sequence two 17-24 mer oligonucleotides may be chemically synthesized which are known as “Primers”. These not only have to be complementary to a stretch of DNA to 3' (Forward) and 5' (Reverse) sides of target sequence, but should not to be complementary to each others to form dimmers, both reactions may prevent DNA amplification.

**Designing of Primers for various application of PCR:**

Differential use of PCR depends upon the selection of the region of DNA template to be amplified as well as designing of PCR primers for the selected sequence.

1. **PCR for Diagnosis**

Microscopy is a widely accepted method for demonstration of various pathogens directly in the clinical specimen (Blood, Sputum, Urin, Stool etc) but it requires about $10^4$ organisms or targets/ml specimen. Use of PCR for diagnosis can improve the sensitivity (fg DNA/ml of clinical specimen can be detected) and specificity (up to species and subspecies level) of clinical diagnosis. A diagnostics PCR ususally amplifies a fragment of smaller size. PCR targeting repetitive sequences (eg. IS-6110 for diagnosis of tuberculosis) may increase the sensitivity, specificity may be cross checked by the Nucleotide-Nucleotide BLAST at [http://www.ncbi.nlm.nih.gov/blast/html](http://www.ncbi.nlm.nih.gov/blast/html). Some specialist software (http://doi.ieeecomputersociety.org/10.1109/BIBE.2004.1317337) are available to design the primers for diagnosis of a disease. Strategy for the purpose is to align nucleotide sequences of common pathogens to discover genus/species specific regions of the particular genomes, so as to design a primer.

2. **PCR for DNA fingerprinting**

PCR for amplifying fragments for DNA fingerprinting usually include selective regions (molecular markers) from the genome (Ribosomal genes, Heat Shock Protein (HSP genes), Sattelite regions, Repetitive units). Some supportive programmes such as Meg align (DNA star), Clustal W, Tendam Repeat Finder are useful for selection of suitable region(s).

a. **PCR-RFLP:** This is a very usefull technique for identification and sub grouping of organisms from a diverse population. Primers for this purpose require the ability to amplify a sufficiently large amplification fragment whose restriction profiles will be characteristic. Common regions may be selected by alignment of many sequences. In addition to this restriction analysis is also dependent upon selection of suitable restriction enzymes for RFLP. DNA stare has the complete solution, clestral W and Primer3 are also useful for designing the PCR-RFLP assay.
b. Random Amplified Poly morphic DNA (RAPD): Random single primer (10-15 mer), must anneal in a particular orientation (such that they point towards to each other) and within a reasonable distance to each other. This approach is used for molecular epidemiology of various diseases.

c. Amplified Fragment Length Polymorphism (AFLP): In case of AFLP the template for the PCR reaction is a Restriction Digested genomic DNA. The primers should contain restriction recognition site. The fixed portion gives the primer stability and the random portion allows it to detect many loci. This technique also provides a genome wide scan but needs automatic DNA sequencer for analysis.

3. PCR/ Probes for detection of drug resistance: Molecular routes for detection of drug resistance are similar to diagnosis methods. Approaches include hybridization with probe to detect mutation(s) and eliciting variations in target sequence by PCR-RFLP, PCR-Sequencing and various other molecular methods to detect the mutations.

4. PCR for gene cloning and Sequencing: To carryout the cohesive ended cloning with PCR products, primers design should force on targeting of fragment along with with Restriction Endonuclease site added to the 3' of Primer. Usefull programmes for the purpose are Do Primer (http://doprimer.interactiva.de/) and GeneRunner.

Optimal Primer sequence and appropriate primer concentration are essential for maximal specificity and efficiency of PCR. Besides the primer sequences, it is important to determine several other things such as the length of the product its melting temperature and ultimately the yield, as the ability of an oligonucleotide to serve as primer for PCR is depending on several factors including: (a) The kinetics of association and dissociation of primer-template duplexes at the annealing and extension temperature. (b) The duplex stability of mismatched nucleotides and their location. (c) The efficiency with which the polymerase can recognize and extend a mismatch duplex.

BASIC FACTORS IMPORTANT FOR PRIMER DESIGNING AND USE OF PRIMERS:

1. KINETICS:

   Melting Temperature (Tm): Tm is defined as the temperature at which half the strands are in the double-helical state and half are in the random-coil state.

   Optimal melting temperature range for a suitable primer set should 52-58°C and primer with Tm>65°C should be avoided to avoid secondary annealing.

   Tm of an oligo can be calculated as
   \[
   Tm = 2(A+T) + 4(G+C)
   \]

   and

   \[
   Tm = 4(G+C)
   \]
For a successful PCR reaction ▲G for primer-template DNA < Primer-Primer

**Parameters for all bases stacking DNA-DNA & DNA-RNA duplex:**

- DNA parameters should be symmetric
- Some standard parameters can be computed by analyses of melting curve for oligonucleotide duplex in solution (Santa Lucia 1998).

DNA ▲H~ 7-9 Kcal/mol

---

**Fig-2 Mechanism of amplification during Polymerase Chain Reaction (PCR)**

1. DNA is denatured. Primers attach to each strand. A new DNA strand is synthesized behind primers on each template strand.

2. Another round: DNA is denatured, primers are attached, and the number of DNA strands are doubled.

3. Another round: DNA is denatured, primers are attached, and the number of DNA strands are doubled.

4. Another round: DNA is denatured, primers are attached, and the number of DNA strands are doubled.

5. Continued rounds of amplification swiftly produce large numbers of identical fragments. Each fragment contains the DNA region of interest.

Tm with salt correction can be computed as: Tm = (▲H/(▲S+R1n(Cr/F)))+ F ([Na⁺]) - 273.15

C_r = Total concentration of strands F = 4, for similar concentration of two strands.

F = 1, if self complementary oligo F = 2, if one strand is in excess

Stability of the chemical reaction can be computed as-

Gibbs free energy ▲G = ▲H-T ▲S
and \[ Tm \text{ (A-T)} = 355.55 + 7.95 \ln [\text{Na}^+] \]
DNA \( \Delta S \approx 19-22 \text{ Kcal/mol} \) and \[ Tm \text{ (G-C)} = 391.55 + 4.89 \ln [\text{Na}^+] \]

- **GC Content:** Between 45-60\%, avoid more than 4 G's or C's in a row\(^9\).
- **3'End Sequence:** Stickier on their end then on their 3' end and a G or C is desirable at 3' because GC clamp reduces spurious secondary bands\(^{14}\).
- **Primer length:** 18-27 nucleotide is considered the best to minimize the chances of a secondary hybridization site\(^{15}\).
- **Dimers and false priming causes misleading results:** Primer should not contain complementary (palindromes) within themselves (self dimer) they should not form hairpins. If this state exits, a primer will fold back on itself and result in an unproductive priming event that decreases an overall signal obtained\(^{16}\). Hairpins that form below 50°C generally do not pose such a problem. Primers should not contain sequences of nucleotides that would allow one primer molecule to anneal to the other primer used in PCR reactions (non self primer dimer formation)
- **Degenerate Primers:** Degenerate primers based on the amino acid sequence of conserved region were also used to search for member of gene family\(^{16}\). Computer programs have also been developed specially for degenerate design\(^{17}\)

2. **DUPLEX STABILITY:**

- **Complementary primer sequence:** Intra or inter primer homology beyond 3 base pair must be avoided.

3. **EFFICIENCY WITH THE POLYMERASE IN PCR REACTION\(^{18}\):**

Factors known to influence the efficiency of polymerase activity are:

- Optimum primer concentration: 0.1-0.2\(\mu\)M
- Optimum Mg\(^{2+}\) ion concentration: 2mM
- Optimum pH of the PCR reaction: 8.5-9.0
- Optimum Taq polymerase concentration: 0.025 U/\(\mu\)l of reaction

4. **HELIX-COIL TRANSITION THERMODYNAMICS:**

The process (PCR Amplification) depends on the adjacent nucleotides on each strand (The Cric's Pair)

The enthalpy and entropy of the duplex can therefore be reduced from the elementary parameters of each Cric's pair.

A correction of the entropy without modification of enthalpy, depends upon temperature of the reaction and the concentration of monovalent salt (Na\(^+\))
The advanced programmes are optimized to reduce the cross hybridization between oligos and targets by filtering the low complexity regions to maintain the specificity of oligos. We have also used OligoPicker for designing the gene specific 70 mer oligos and developed partial DNA Chip for Mycobacterial genome and DNA stare, Primer3 for designing the primers for gene amplification and real-time reverse transcription PCR (RT-PCR).

Following different programs have been developed by various workers to design the primes or probes for different purposes.

1. PCR oligonucleotide resources:

   A. For Calculation of melting point of an oligonucleotide:
      Oligonucleotides pour laPCR (http://www.citi2.fr/bio2/OligoTM.html)

      Oligonucleotide properties calculator (http://www.basic.nwu.edu/biotools/oligocalc.html)

      Oligo Tm Determination: (http://alces.med.umn.edu/rawtm.html)

   B. Programmes to generates Tm, free energy, molecular weight, hairpin and dimer formation structures.
      Oligonucleotide analyzer (http://www.rnature.com/oligonucleotide.html)

      PROLIGO (http://www.biophys.uni-duesseldorf.de/local/POLAND/poland.html)
2. Software available for designing the primers for gene amplification may further divided in following categories:

A. Primer design for individual gene amplification:

Primer Select: Analyzes a template DNA sequence and chooses primer pairs for PCR and primers for DNA sequencing
http://www.dnastar.com

Primer Premier 5: primer design for Windows and Power Macintosh.
http://www.premierbiosoft.com/primerdesign/primerdesign.html

Net Primer: Comprehensive analysis of individual primers and primer pairs.
http://www.premierbiosoft.com/NetPrimer.html

Fast PCR: Software for Microsoft Windows has specific, ready-to-use templates for any PCR and sequencing applications: standard and long PCR, inverse PCR, degenerate PCR directly on amino acid sequence, multiplex PCR.

Primer Designer 4: Will find optimal primers in target regions of DNA or protein molecules, amplify features in a molecule, or create products of a specified length.
http://www.scied.com/ses_pd5.htm

Primer Designer: The Primer Designer features a powerful, yet extremely simple, real-time interface to allow the rapid identification of theoretical ideal primers for your PCR reactions.
http://genamics.com/expression/primer.htm

Primer Premier: Automatic design tools for PCR, sequencing or hybridization probes, degenerate primer design, Nested/Multiplex primer design, restriction enzyme analysis and more.

Online primer design web resources:

CODEHOP: http://blocks.fhcrc.org/codehop.html

Gene Fisher: http://bibiserv.techfak.unibielefeld.de/genefisher

Do Primer: http://doprimer.interactiva.de/

Primer3: Widely accepted, freely available, online probe and primer designing software
http://www-genome.wi.mit.edu/cgi-bin/primer/primer3 www.cgi
http://www.basic.nwu.edu/biotools/Primer3.html

Primer Selection: http://alces.med.umn.edu/rawprimer.html

Web Primer: http://genome-www2.stanford.edu/cgi-bin/SGD/web-primer
PCR Designer: http://cedar.genetics.soton.ac.uk/public_html/ primer.html

Primo Pro 3.4: http://www.changbioscience.com/primo/primo.html

Primo Degenerate 3.4: http://www.changbioscience.com/primo/primo.html

PCR Primer Design: http://pga.mgh.harvard.edu/servlet/org.mgh.proteome.Primer

The Primer Generator: http://www.med.jhu.edu/medcenter/primer/primer.cgi

PRIMO3: http://bioweb.pasteur.fr/seqanal/interfaces/primo.html3
http://atlas.swmed.edu/primo/primo_form.html

Primer Quest: http://www.idtdna.com/biotools/primer_quet/primer_quest.asp

Met Primer: http://itsa.ucsf.edu/urolab/methprimer/index1.html

Raw primer: http://alces.med.umn.edu/rawprimer.html

MEDUSA: http://www.cgr.ki.se/cgr/MEDUSA/

Oligonucleotides for the PCR: http://www.citi2.fr/bio2/Oligo2lib.html

B. Genome scale primer designing:
Also useful for designing oligos for the DNA Chip/microarray.

Array Designer 2: For fast, effective design of specific oligos or PCR primer pairs for microarrays

http://www.premierbiosoft.com/dnamicroarray/dnamicroarray.html

GenomePRIDE 1.0: Primer design for DNA-arrays/chips.
http://pride.molgen.mpg.de/genomepride.html

Sarani Gold: Genome Oligo Designer is software for automatic large-scale design of optimal oligonucleotide probes for microarray experiments.
http://mail.strandgenomics.com/products/sarani/

Genorama chip Design Software: Genorama Chip Design Software is complete set of programs required for genotyping chip design. The programs can also be bought separately.
http://www.asperbio.com/Chip_desin_soft.htm

GAP: Genome-wide Automated Primer finder servers.
http://promoter.ics.uci.edu/Primers/

Array Designer: Design hundreds of primer for DNA or oligonucleotide microarrays.

C. Genome scale probe designing:
useful for designing oligo based high density array.

Oligo Picker: Design specific oligos by skipping regions with contiguous bases common in other sequences. In addition, oligo specificity is double checked by NCBI-BLAST.
Oligo Array2: A free software that computes gene specific oligonucleotides for genome-scale oligonucleotide microarray construction.

Oligodb: a web-based system for interactive design of oligo DNA for transcription profiling (hybridization) of human genes. The oligodb system uses the human DNA-transcripts of ENSEMBL.

D. designing of probes (Molecular Beacons) for mutation detection by real time PCR

Beacon Designer 2.1: Design molecular beacons and TaqMan probes for robust amplification and fluorescence in real time PCR.


REFERENCES


11. Santa Lucia JJ. A unified view of polymer, dumbbell and oligonucleotide DNA nearest-


