

Designing PCR Primers

INTRODUCTION

The polymerase chain reaction (PCR) is a technique that has revolutionized many aspects of modern genetics. The technique was developed in 1986, earning its developers the Nobel Prize in 1993. It is a method of amplifying segments of DNA that generates millions of copies of a specified sequence in a period of hours. Since its development, PCR has replaced many of the traditional methods of molecular biology as a method of choice for various applications. These include aspects of DNA cloning (i.e., sequence amplification), DNA sequencing, and even the generation of specified mutations.

In this lab activity, you will use the World Wide Web to search a database for a gene sequence, and you will then use a Web-based program to design PCR primers for that sequence (an overview of the PCR process is presented shortly). As a follow-up exercise, your instructor may have you design primers for a gene of your choice. Such primers could actually be synthesized and tested if time and other resources permit (see Lab Activity 7, entitled “Using PCR to Find New Gene Sequences”).

POLYMERASE CHAIN REACTION: BACKGROUND INFORMATION

In PCR, regions of DNA that are flanked by short, known sequences can be easily amplified. This provides enough copies of the amplified DNA so that it can be visualized on an electrophoresis gel and manipulated in various ways. Sometimes, the flanking sequences are known, but what lies between them is unknown. In these cases, PCR can be used to amplify a sequence that can be further studied to characterize it completely (this approach is illustrated in the last section of this lab activity entitled “Discovering New Gene Sequences Using PCR”). In other instances, PCR is used to manipulate DNA sequences that are already characterized. For example, PCR can be used to incorporate specific new mutations into genes carried on plasmids (PCR-based, site-directed mutagenesis).

In all cases, PCR amplifies DNA using a DNA polymerase that requires a primer. Primers are short, synthetic, oligonucleotides that are complimentary to DNA sequences that flank the region to be amplified. In most cases, the sequences of these primers are based on known sequences; however, one technique, called randomly amplified polymorphic DNA PCR, uses random primers. In addition to primers and DNA polymerase, PCR reactions must contain template DNA (the DNA to be amplified) and the DNA “building blocks” deoxynucleotide triphosphates (dNTPs, which include dATP, dTTP, dGTP, and dCTP). Furthermore, since the PCR process includes a high-temperature step, the DNA polymerase must be a thermostable polymerase.

Amplification of DNA by PCR is accomplished by a process known as *thermal cycling*. The PCR reaction (template DNA, primers, dNTPs, thermostable DNA polymerase, etc.) is combined in a tube

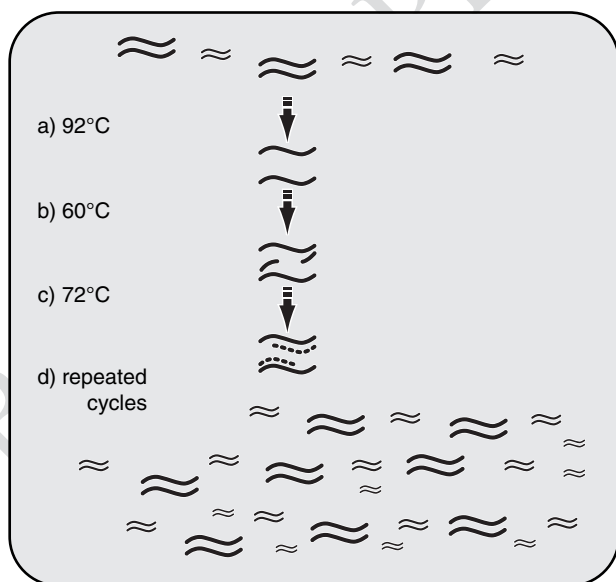


Figure 6.1 – A diagrammatic representation of the PCR process: (a) template DNA is denatured by high temperatures; (b) primers anneal to template DNA; (c) thermostable DNA polymerase synthesizes new DNA; (d) multiple cycles of the three temperatures result in an exponential increase in the sequence associated with the PCR process.

these three temperature stages, essentially doubling the sequence flanked by the primers with each cycle. This process results in an exponential increase in the number of copies of the sequence, so that millions of copies of the PCR product are present at the end of the process. This amplified DNA is so concentrated that it can be seen as a discrete band when the DNA is visualized on an electrophoresis gel.

It is interesting to note that practical PCR was not possible until a thermostable DNA polymerase was isolated (one that would not be denatured by the 92° to 94°C DNA denaturation step). Such an enzyme (now called *Taq* polymerase) was first isolated from the bacterium *Thermus aquaticus* (= *Taq*) that thrives in hot springs, where most living things cannot survive.

PROCEDURE

The following steps will direct you through the process of generating two pairs of primers for the CRHB2 gene from the model plant system *Ceratopteris richardii*. *Ceratopteris* is a fern that is a relatively “new” model organism with great promise for studies in genetics (see the C-Fern page at <http://www.c-fern.org> for more information). In this exercise, each pair of primers you design will be specific for a particular region of the gene (designated as either region 1 or 2). Regions 1 and 2 were arbitrarily chosen to include the sequences from base number 50 to 150 (region 1) and from 500 and 600 (region 2). You will be given step-by-step instructions (to ensure success) for designing a primer pair for region 1 and less-specific instructions for designing primers for region 2.

Step 1: Go to the “Entrez Nucleotide” search page of the National Center for Biotechnology Information (NCBI) website as described below. The NCBI homepage provides links to many useful tools and databases. You will only be using a small portion of the resources available there

and placed into a DNA thermal cycler. The thermal cycler rapidly heats up and cools down, taking the reaction mixture through a series of temperature changes that occur as cycles. These cycles consist of three temperature stages that each accomplish different aspects of DNA amplification (see Figure 6.1). The first stage uses a high temperature (usually 92° or 94°C) that denatures the double-stranded DNA into single strands. This step is necessary before new DNA can be synthesized on the single strands. The second stage uses a lower temperature at which the synthetic primers recognize and form hydrogen bonds with the single-stranded DNA (this process is called *annealing*). The specific temperature of this stage depends on the base composition of the two primers (Figure 6.1 shows an annealing temperature of 60°C). Specific annealing temperatures can be calculated for any given PCR primer. The third stage uses a temperature that is optimum for the function of the DNA polymerase that synthesizes the new DNA (usually 72°C). The thermal cycler takes the reaction mixture through many cycles (often as many as 35) of

for this activity. However, you may wish to surf the NCBI homepage a bit when you visit it to see what's available there. To obtain the sequence for the *Ceratopteris* CRHB2 gene, you will be using the "Entrez Nucleotide" search page. To get to this page, go to the NCBI homepage (www.ncbi.nlm.nih.gov/).

After arriving at the NCBI homepage, click the words "Entrez Home" in the column on the right. This will take you to a page where you can choose from several links, including one labeled "Core Nucleotides"; select this link. This will take you to a page entitled "Entrez Nucleotide."

Step 2: Searching for *Ceratopteris* gene sequences — In the search box, type in the name "Ceratopteris" (without the quotes) and click the Go button. When the search engine has completed the database search, you will see the results of the search, including an indication of the number of documents that were found that are related to *Ceratopteris*. Since *Ceratopteris* is a relatively "new" model organism, the number of sequenced genes is somewhat lower than would be the case for an organism like *Arabidopsis* (try the same search using "Arabidopsis" as the search term, and see how many records are listed).

Q1 How many sequences did you find when you searched using *Ceratopteris* as a keyword?

Step 3: Viewing information about the CRHB2 sequence — Now go back to the "Entrez Nucleotide" search box, type in the term CRHB2, and click the Go button. When the search engine has completed its search, click the link for the citation it finds, and you will see a page with information about this gene and the individuals who have worked on its sequence.

Q2 What is (are) the name(s) of the author(s) who published an article describing this sequence?

Q3 According to the article title in the GenBank report, what kind of gene is CRHB2?

Step 4: Copying the CRHB2 gene sequence — At the top of the page, you will see a pull-down menu labeled “display.” Use the pull-down menu on that box to select “FASTA.” When “FASTA” is selected this will take you to a page containing the nucleotide sequence for the CRHB2 gene in a form that can be copied and manipulated. Using the mouse or the keyboard, highlight the complete gene sequence (only the nucleotides), and hit the “copy” command in the browser menu. Once the sequence is copied to the computer’s clipboard, you will access another Web page to work with the sequence to generate PCR primers.

Step 5: Generating PCR primers for the CRHB2 gene — In this step, you will be using a Web-based program that designs primers based on the parameters you designate and other factors.

After you have supplied some specifications, the program uses various algorithms to generate the best possible primers that fit your specification. Various factors are involved in generating good primers. Ideally, they should only bind to the template DNA in the regions for which they were designed. They should also be composed of sequences that will not form internal hydrogen bonds or bonds between the two primers.

To access a Web page with a primer design program, go to the following Web address:

Primer3 (a Primer Design page):
<http://frodo.wi.mit.edu/>

When you arrive at this page, paste the gene sequence into the box at the top of the page. Next you will set some parameters that allow you to modify the primer selection process. As you set these parameters, click the associated links so that you can learn what types of specifications you’re setting. If no setting is given in the following list for a particular parameter, leave the default setting as it is.

Sequence Id: Supply a name of your choice.

Target: 100,1

Product size range: delete the numbers given as possible default settings and enter 60-90

Primer size: minimum, 18; optimum, 18; maximum, 18

After entering the information, scroll down the page and click the “pick primers” icon, and the sequences for an optimum primer pair will be generated (a number of alternative primer pairs will also be listed following these). You will also see a representation of the gene with arrows indicating where the primers will associate with the gene sequence. Notice from this image and also from the information regarding the starting point for each primer that the PCR product should represent a portion of the gene located in region 1 (as described earlier). This result occurred because you entered appropriate parameters (e.g., target and product size) to obtain primers that should produce a product in region 1.

Q4 How big will the PCR product be if the optimum primers are used?

Q5 What are the sequences of the two optimum primers that were generated?

Now go back to the Primer3 input page and try modifying some of the specified parameters to see how it affects the size and placement of the PCR product that would be generated. After you are familiar with how the program works, generate a primer pair according to the following specifications. The primer pair you generate next should produce a PCR product that will be less than 80 base pairs long and located within region 2, as described earlier. The PCR primers themselves may be any size from 18 to 27 bases long.

Print the portion of the Primer3 output page that shows the optimum primer pair that you generate according to the specifications given earlier, and attach the printout to your lab assignment when you hand it in.

Q6 The sequence for CRHB2 in the GenBank database was generated from mRNA (i.e., the sequence is equivalent to a cDNA sequence). If you wanted to use the primers you generated in this exercise to amplify sequences from genomic DNA (instead of cDNA), what other information would be beneficial to know about the gene sequence for CRHB2 before using these primers in an experiment? *Hint:* Sequences in eukaryotic mRNAs are generally quite different from genomic sequences because of a phenomenon known as posttranscriptional modification. In answering question 6, think about what these differences are and which of them might cause unexpected changes in the predicted size of your PCR product.

DISCOVERING NEW GENE SEQUENCES USING PCR

Your instructor may ask you to repeat the preceding process using a gene of your choice (or one assigned to you) as a follow-up activity. If time and resources permit, your primers (or those designed by another classmate) may actually be synthesized and tested, as described in a separate lab activity. If the gene and the primers are chosen carefully, you may be able to discover a new gene sequence (in a different organism) using these primers. The following are some tips that will help you design your primers for this purpose.

Choosing an Organism to Study

The organism for which you design new primers should be one for which relatively few genes have been sequenced. For example, as you saw earlier, far fewer genes from *Ceratopteris* have been sequenced than from *Arabidopsis*. Similarly, while the entire genome of *Drosophila melanogaster* has now been sequenced, very few genes from ant species have been described. To see what kinds of genes have been sequenced for a given organism, just run an “Entrez Nucleotides” search for that organism in the NCBI database. In a separate lab exercise, instructions are given for extracting DNA from plants and several small insect species, so that PCR primers can be tested using the DNA of these organisms.

Choosing a Gene to Study

If you hope to discover a new sequence, obviously you must focus on a gene that hasn't been sequenced in your organism (check GenBank). To design primers, you'll need to work from a known gene that has been sequenced in a related organism. For example, if you want to find a new sequence in *Ceratopteris*, start with a sequence from another plant and maybe even another fern. You'll have a better chance of picking primers that will work if you select a gene that is known to be highly conserved. Many genes that serve essential roles are highly conserved, meaning that they are structurally more similar between species than are other genes. For example, ribosomal genes, calmodulins, tubulins, and parts of most homeotic genes are known to be highly conserved. Finally, consider the source of the sequence information that you use to design your primers — is it from genomic DNA or cDNA, and what are the implications (i.e., see question 6)?

Testing Your Primers

If you are able to have your primers synthesized, you can isolate the DNA from your organism of choice and conduct PCR using these primers (see Lab Activity 7, entitled “Using PCR to Find New Gene Sequences”). If your primers amplify a PCR product of the predicted size, this will indicate that you have successfully generated functional primers for the gene of interest. The next step to discovering a new sequence will be to isolate the PCR product from the gel and have it sequenced. If you can do this, you will be making a contribution to our current knowledge of genetics!