

This protocol will outline the basic steps to take when designing a set of primers for PCR amplification. In addition to designing the primers using online bioinformatics programs, this document will help to ensure certain criteria are met to minimize the occurrence of undesirable secondary structures (i.e., hairpins, self dimers, cross dimers) and non-target amplifications (i.e., cross amplification, cross homology).

STEP 1: Start with the literature. It is possible that primers have already been designed and tested, which means you can simply order the primer set and begin your research. Just because other researchers have used the primers, does not preclude the novelty of your research. Besides, many excellent research ideas and projects start with replicating what others scientists have done. Always start with the literature.

STEP 2: You will need to start with sequence data for the targeted gene. This information may be found in the literature or through bioinformatics websites like BLAST and GeneCards (human only). For example, if you are using eDNA to barcode an animal species, you will probably start with the CO1 sequence and design a species-specific primer to amplify a unique region within that gene. More on that subject later.

STEP 3: Now that you have your target sequence, you may use an online bioinformatics program like Primer3Plus or Primer-BLAST to design several sets of possible primers. Let's look at an example using Primer3Plus to design primers to amplify the human SV2A gene, which is a synaptic vesicle protein found in the brain. The SV2A sequence was found using GeneCard and is 14,521 base pairs long. Navigate to the Primer3Plus website and simply paste the sequence (or FASTA code) as shown below:

The image shows a screenshot of the Primer3Plus web application. The interface includes a navigation menu with links for 'Primer3Manager', 'Help', 'About', and 'Source Code'. The main content area is titled 'Primer3Plus' with the subtitle 'pick primers from a DNA sequence'. A 'Task' dropdown menu is set to 'Detection', with a description: 'Select primer pairs to detect the given template sequence. Optionally targets and included/excluded regions can be specified.' There are 'Pick Primers' and 'Reset Form' buttons. Below this is a tabbed interface with tabs for 'Main', 'General Settings', 'Advanced Settings', 'Internal Oligo', 'Penalty Weights', and 'Sequence Quality'. The 'Main' tab is active, showing a 'Sequence Id:' field and a 'Paste source sequence below' section. The sequence is a multi-line FASTA-style format with line numbers on the left. Below the sequence is a 'Mark selected region:' section with buttons for '<>', '[]', '{}', and 'Clear', and a 'Save Sequence' button. At the bottom, there are fields for 'Excluded Regions:', 'Targets:', and 'Included Region:'.

To start, you can ignore the other fields on the website, unless your target gene is in a plasmid vector. Press the [Pick Primers] button and the program will automatically generate a list of possible primers.

Below you will find a set of primers for the SV2A gene. Primer3Plus will typically generate numerous primers for you to choose from. In this case five primers (not shown) were generated and the program chooses the most suitable primer as primer #1 (circled below). However, you need to consider all of the primer options before simply going with primer set #1. For example, smaller product sizes are desired for eDNA barcoding, so a researcher would opt for ~150 bp product and choose a different primer and not choose primer #1.

Primer3Plus

pick primers from a DNA sequence

[Primer3Manager](#)
[About](#)

[Help](#)
[Source Code](#)

WARNING: Numbers in input sequence were deleted.

< Back

Pair 1:

Left Primer 1:

Sequence: →

Start: 9751 Length: 20 bp Tm: 60.0 °C GC: 55.0 % ANY: 4.0 SELF: 0.0

Right Primer 1:

Sequence: →

Start: 9989 Length: 20 bp Tm: 60.0 °C GC: 55.0 % ANY: 3.0 SELF: 0.0

Product Size: 239 bp Pair Any: 4.0 Pair End: 1.0

```

1      ttttgtattt  tgaggatttg  tttccgtttt  atttttogtc  aatatttttc
51     actgcatttt  tcacatcaac  agagttgggt  ttagggttcc  ccaggagagc
101    aagggagatg  gatgggtccc  tcccaggggt  gcgaggaaga  gggctacgtc
151    cagcatgtgg  ggagtgaggg  acaagggcac  agggcacaca  ctcatacact
201    gggcactgga  gaagcagtc  ttttgagcac  agacaccaa  gagttaaca
251    gtcaccacct  ggtacagtg  taaaaatgg  ggagagtgc  cacagggttg
301    gggcactgga  cgcagcatga  cgaagtgaa  ttttgatttt  cttttttgtt
351    attttcaact  ttttctttt  aacattttt  taaacatttt  acaaacatct
401    -----  -----  -----  -----  -----
          
```

STEP4: The next step is to analyze the design statistics of the primers. It is imperative that you familiarize yourself with some of the biochemistry that is crucial for primer specificity and successful amplification of the target sequence using PCR. While the computer programs will provide analytics (see arrow above), the researcher must be able to interpret and discuss these results. Below is a summary of the basic terminology:

Oligonucleotide: This is a short length of DNA that is ~20 base pairs, which is the ideal size for a primer.

Melting Temperature (i.e., T_m): This is the temperature at which the DNA template/primer duplex will dissociate and become single stranded. This is an important consideration when determining the annealing temperature for PCR. The optimal T_m range is around 52-62 °C, but is dependent on variables such as GC content and primer length. It is a good idea to keep the T_m of the two primers within 3-5 °C of each other.

Primer	Tm (DegC)	Salt (mM)	Primer (µM)	%GC	Length
Primer #1: <input type="text" value="AGGGAAGAGTGTGCAGAGGA"/>	58.40	50	.2	55	20
Primer #2: <input type="text" value="GACGGTGTGGAGGTCTTTGT"/>	58.43	50	.2	55	20
Tm Delta = 0.03					

The above image was generated from an online T_m calculator (see link below) for the primers generated in our example for the SV2A gene. Notice that the programs report T_m data that are ~ 1.5 °C different. Please do not stress over this minor discrepancy, just reference the website in your research notebook to refer back to if needed. Different programs use different thermodynamic equations to calculate the T_m . For now, the values are close enough, so let's call it 59 °C. Here is the T_m calculator website:

<http://www6.appliedbiosystems.com/support/techtools/calc/>

Guanine-Cytosine Content (i.e., GC, GC content): Expressed as a percentage of the total bases, you want the GC content to be 40-60%. Recall that GC bonds contain three hydrogen, whereas AT only form two hydrogen bonds. This means that GC bonds are stronger, and thus, they will increase the melting temperature (T_m) of the primer/template pair. Because the function of the enzyme DNA polymerase is temperature specific, the researcher must not have a T_m that is out of this optimal range 52-62 °C. Primers should not have regions of four or more consecutive bases (e.g., ATTGGGGCA).

Annealing Temperature (i.e., T_a): This is the temperature at which the primer will bind and anneal to the template strand of DNA. In general, the annealing temperature is about 5 °C below the T_m value and should usually fall in the range of 50-60 °C. There are always exceptions, so please consider these ranges as only suggestions for a typical reaction.

GC Clamp: The presence of 2-3 Gs or Cs at the 3' end of the primer is referred to as a GC clamp and will help ensure bonding specificity due to the strength of the G-C bonds, as discussed earlier. The GC clamp should be occur within the last five bases of the 3' end and there should be no more than three total Gs/Cs. For example, the SV2A forward primer is agggaagagtgtgcagagga and it has a GC clamp composed of three Gs.

Primer Secondary Structures (see figure on last page): We usually refer to these as primer dimers, which can result from a self-complementary regions within or inter-complementary regions between primers. Depending on which program you are using (red = primer BLAST; blue = Primer3Plus), you will get different scores on the primer's tendency to bond to itself or to the other primer. The “self-3'complementarity” or “SELF” and the “self-complementarity” or “ANY” indicate the likelihood that the primer will bond to itself forming dimers and hairpins. I have used a blue box on our example above to show two additional scores given on Primer3Plus. These two scores, Pair Any and Pair End, indicate the likelihood that the two primers will bond with each other. Please view the figure on the last page of this document. After all of that esoteric jargon, we will shoot for scores of 4 or less on our primer designs.

Primer Analyzer Tools:

<http://www.idtdna.com/calc/analyzer>

http://www.bioinformatics.org/sms2/pcr_primer_stats.html

STEP 5: You are almost done. The next step you will check your primer set for cross amplification. In other words, we need to ensure your primers will not amplify other genes that are in the mixture. This will confirm, *in silico*, that you have chosen a good set of specific primers that should amplify you intended target sequence. For this step I like to use Primer BLAST. On Primer BLAST scroll down to “Primer Parameters” and enter the sequences for your primers. Leave the other fields as there are, unless your PCR product exceeds 1000 base pairs. Next, scroll down to “Primer Pair Specificity Checking Parameters” and complete all fields to best of your knowledge. Then click [Get Primers]. In this example, I knew I was using the human SV2A sequence from the human genome. Please refer to the next two images below for clarification.

Primer Parameters

Use my own forward primer (5'→3' on plus strand) [?](#) [Clear](#)

Use my own reverse primer (5'→3' on minus strand) [?](#) [Clear](#)

PCR product size

of primers to return

Primer melting temperatures (T_m) [?](#)

Primer Pair Specificity Checking Parameters

Enable search for primer pairs specific to the intended PCR template

Search mode

Database [?](#)

Exclusion Exclude predicted Refseq transcripts (accession with XM, XR prefix) Exclude uncultured/environmental sample sequences [?](#)

Organism [?](#)
Enter an organism name (or organism group name such as enterobacteriaceae, rodents), taxonomy id or select from the suggestion list as you type.

Entrez query (optional)

Primer specificity stringency total mismatches to unintended targets, including at least mismatches within the last bps at the 3' end. [?](#)
 Ignore targets that have or more mismatches to the primer. [?](#)

Max target size [?](#)

Splice variant handling Allow primer to amplify mRNA splice variants (requires refseq mRNA sequence as PCR template input) [?](#)

[Get Primers](#) Show results in a new window Use new graphic view [?](#)

The image below shows the results of my Primer BLAST. Notice that only one match came back for the human SV2A glycoprotein. At this point you have chosen a good primer set *in silico*, so it is time to order the primers. Please note that if the results come back for other species and other genes, you need to reason through the likelihood that this genetic material would be found in the same sample or by species that coexist. If my result came back and indicated that the SV2A gene, CYPC19 gene, and CYTB gene were all possible results, I would need to start over to find a better primer set.

Primer pair 1

	Sequence (5'→3')	Length	T _m	GC%	Self complementarity	Self 3' complementarity
Forward primer	AGGGAAGAGTGTGCAGAGGA	20	60.18	55.00	4.00	0.00
Reverse primer	GACGGTGTGGAGGTCTTTGT	20	59.89	55.00	3.00	0.00

Products on target templates
 >[NC_000001.11](#) Homo sapiens chromosome 1, GRCh38.p7 Primary Assembly

product length = 239
 Features associated with this product:
[synaptic vesicle glycoprotein 2A isoform X1](#)
[synaptic vesicle glycoprotein 2A isoform 1](#)

Forward primer	1	AGGGAAGAGTGTGCAGAGGA	20
Template	149913068	149913087
Reverse primer	1	GACGGTGTGGAGGTCTTTGT	20
Template	149913306	149913287

Primer Design Tools:

<http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi>

<http://www.idtdna.com/Primerquest/Home/Index>

<https://www.ncbi.nlm.nih.gov/tools/primer-blast/>

Primer Analyzer Tools:

<http://www.idtdna.com/calc/analyzer>

http://www.bioinformatics.org/sms2/pcr_primer_stats.html

Primer Secondary Structures:

Self-complementarity and primer dimers

• 5' ATGCATCGGAATCCGATGC 3'

The primer above has sequences on either side of the vertical line that can base pair with each other, forming a "hairpin" structure. If a primer folds on itself and forms a hairpin, it can't bind the target sequence.

• 5' TGACGTGGATCTACGGT 3'

• 5' AGCTTACTAGCCACGTT 3'



The two primers shown above have sequences at the 3' ends that can base pair with each other (shown in red). Such primers will bind to each other instead of to the target because of the very high concentration of the primers relative to the concentration of template molecules.