Biol/Chem 475 Primer Design Spring 2007
Generating a 100 bp ladder using phage lambda DNA as template

FAQS

What are we doing in lab for the next couple of weeks?

- Week 7: design primers for a 100 bp ladder of PCR products
- Week 8: try primers out; assess PCR results and work up strategy to optimize yield and specificity
- Week 9: each student reruns PCR varying one or two parameters of the reaction

What's the Point?

- The goal of this exercise is to introduce you to the the basics of primer design and troubleshooting PCR reactions – where artifacts and suboptimal reactions can kill an experiment or, worse yet, provide misleading restuls

Where should I go & What should I do? *In this handout*
What do I need to hand in? See extra sheet
Appendix A: I thought molecular biologists don't usually care about details like Tm
Appendices B: Jeff Young's general guidelines for primer design
Appendix C: Downloading sequence from NCBI in Fasta Format

Where should I go?

NCBI: Obtain template DNA sequence
http://www.ncbi.nlm.nih.gov/

Primer-3: Design Primers
http://frodo.wi.mit.edu/

Sequence Extractor: Display primers on DNA sequence
http://bioinformatics.org/seqext/
Sequence Extractor features
http://bioinformatics.org/seqext/features.html

GeneWalker: analyze potential for primer artifacts/interactions
http://www.cybergene.se/primerdesign/genewalker/genewalker11.html

GeneWalker Manual
http://www.cybergene.se/primerdesign/help.html

IDT-DNA Another site to analyze primer-primer interactions – gives ΔG's!
http://www.idtdna.com/analyzer/Applications/OligoAnalyzer/Default.aspx

Other sites of interest Online PCR TOOLS
http://molbiol-tools.ca/PCR.htm
Tips on primer design
http://www.premierbiosoft.com/tech_notes/PCR_Primer_Design.html

What's the point?

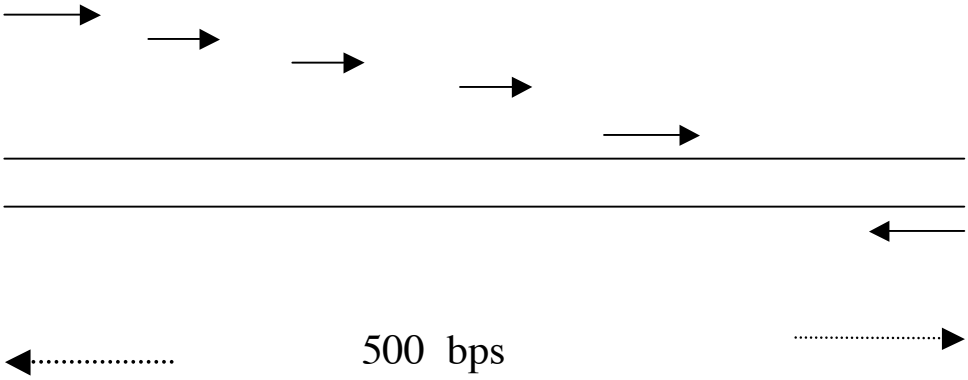
DNA size standards are expensive, so we are going to make our own.

Inexpensive Reagents: TAQ polymerase (buffers, dNTPs),
PCR primers,
DNA template (*lambda bacteriophage*)

Invaluable: Your Time. You will be designing primers to amplify, from a single template, a 1000 bp product & a nested set of products 100-500 bps in length in 100 base pair increments &. In other words we want (using 7 or more primers) products of 100, 200, 300, 400, 500 bp, & 1000 [+/- as few bases as possible]. The primers will be used in six different PCR reactions; the products of the reactions will then be combined to make a 100 base pair ladder mix.

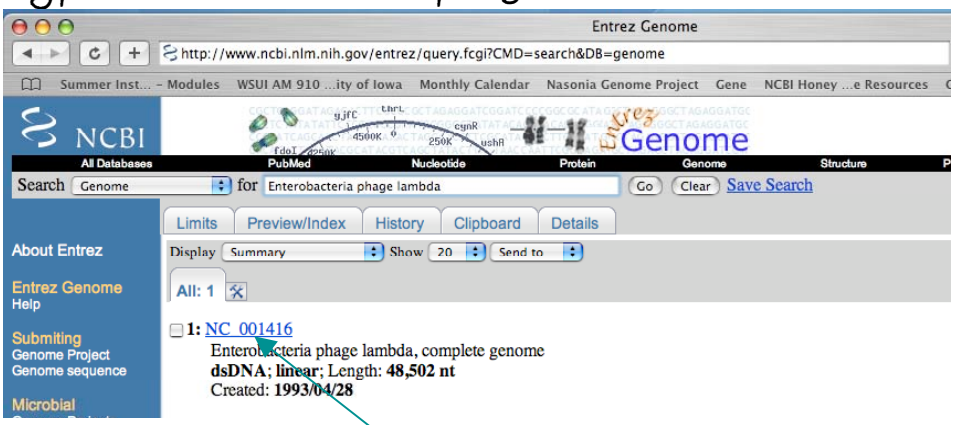
Suggested strategy:

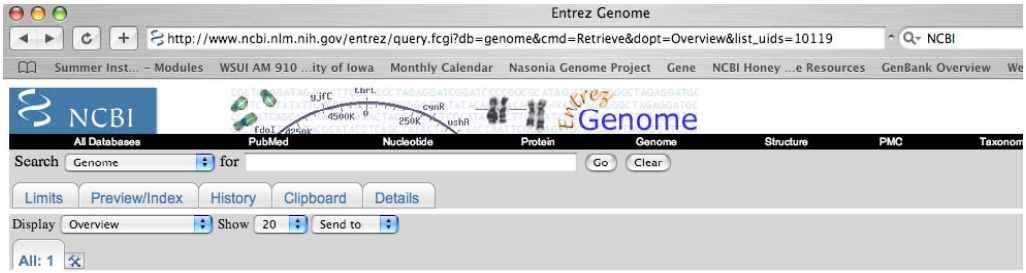
- Note that you can pair one primer with more than one partner to generate different sized PCR products and to reduce the number of primers needed
- Start with largest product and generate primers
- Pick one of these primers and look for nested sets (see drawing below)
- if that primer isnt' working out, pick the other outside primer and look for nested sets



FIRST: GET TEMPLATE SEQUENCE

Go to National Center for Biotechnology Information (NCBI)
http://www.ncbi.nlm.nih.gov/
Choose **Genome** in the pull down menu
Type in Enterobacteria phage lambda





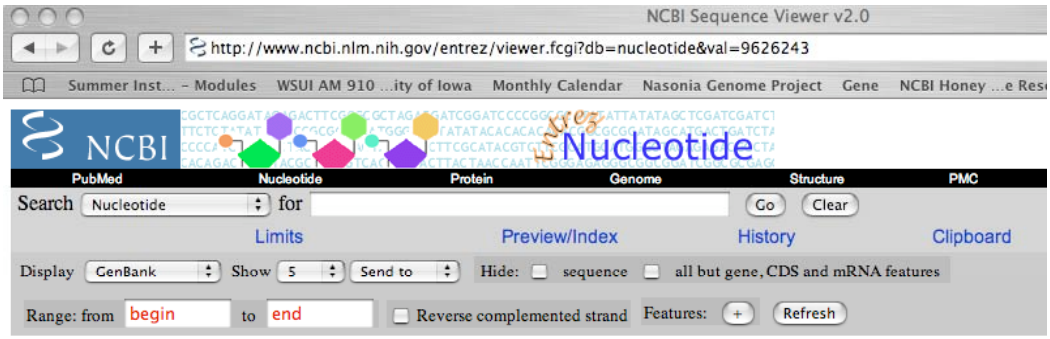
Genome > **Viruses** > ***Enterobacteria phage lambda, complete genome***

Lineage: Viruses ; dsDNA viruses, no RNA stage ; Caudovirales ; Siphoviridae ; Lambda-like viruses ; Enterobacteria phage lambda

Genome Info:	Features:	BLAST homologs:	Links:	Review Info:
Refseq: NC_001416	Genes: 90	COG	Genome Project	Publications: [13]
GenBank: J02459	Protein coding: 71	3D Structure	Refseq FTP	Refseq Status: **Reviewed**
Length: **48,502 nt**	Structural RNAs: **None**	TaxMap	GenBank FTP	Seq.Status: **Completed**
GC Content: **49%**	Pseudo genes: **None**	TaxPlot	BLAST	Sequencing center:
% Coding: **87%**	Others: **94**	GenePlot	TraceAssembly	Completed: **1993/04/28**
Topology: **linear**	Contigs: 1	gMap	CDD	Organism Group
Molecule: **dsDNA**			Other genomes for species	

Gene Classification based on COG functional categories Search gene, GeneID or locus_tag: Find Gene

Click here for Sequence Viewer presentation (base sequence and aligned amino acids) of selected region



1: NC_001416. Reports Enterobacteria ph...[gi:9626243]

Comment Features Sequence

```
LOCUS       NC_001416              48502 bp    DNA     linear   PHG 30-MAR-2006
DEFINITION  Enterobacteria phage lambda, complete genome.
ACCESSION   NC_001416
VERSION     NC_001416.1  GI:9626243
PROJECT     GenomeProject:14204
KEYWORDS    .
SOURCE      Enterobacteria phage lambda
  ORGANISM  Enterobacteria phage lambda
            Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae;
            Lambda-like viruses.
REFERENCE   1  (sites)
  AUTHORS   Chen,C.Y. and Richardson,J.P.
  TITLE     Sequence elements essential for rho-dependent transcription
            termination at lambda tR1
```

Scroll (way) down to nucleotide sequence
Select and copy your assigned sequence block next page)

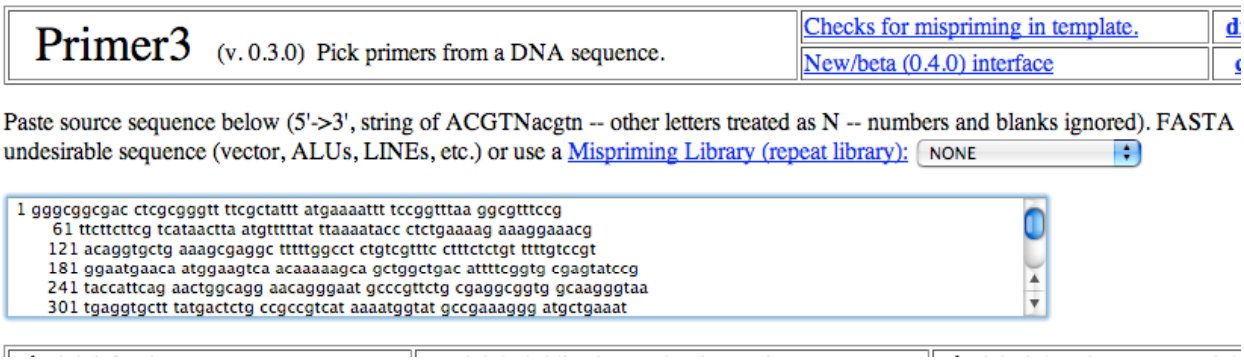
```
ORIGIN
        1 gggcggcgac ctcgcgggtt ttcgctattt atgaaaattt tccggtttaa ggcgtttccg
       61 ttcttcttcg tcataactta atgtttttat ttaaaatacc ctctgaaaag aaaggaaacg
      121 acaggtgctg aaagcgaggc tttttggcct ctgtcgtttc ctttctctgt ttttgtccgt
      181 ggaatgaaca atggaagtca acaaaaagca gctggctgac attttcggtg cgagtatccg
      241 taccattcag aactggcagg aacagggaat gcccgttctg cgaggcggtg gcaagggtaa
      301 tgaggtgctt tatgactctg ccgccgtcat aaaatggtat gccgaaaggg atgctgaaat
      361 tgagaacgaa aagctgcgcc gggaggttga agaactgcgg caggccagcg aggcagatct
      421 ccagccagga actattgagt acgaacgcca tcgacttacg cgtgcgcagg ccgacgcaca
      481 ggaactgaag aatgccagag actccgctga agtggtggaa accgcattct gtactttcgt
      541 gctgtcgcgg atcgcaggtg aaattgccag tattctcgac gggctccccc tgtcggtgca
      601 gcggcgtttt ccggaactgg aaaaccgaca tgttgatttc ctgaaacggg atatcatcaa
      661 agccatgaac aaagcagccg cgctggatga actgataccg gggttgctga gtgaatatat
      721 cgaacagtca ggttaacagg ctgcggcatt ttgtccgcgc cgggcttcgc tcactgttca
```

➔ You can directly copy the Genbank sequence (numbers and all) and paste it into the Primer3 sequence window. The program will ignore the numbers.
[also see appendix C for alternative way to download sequence]

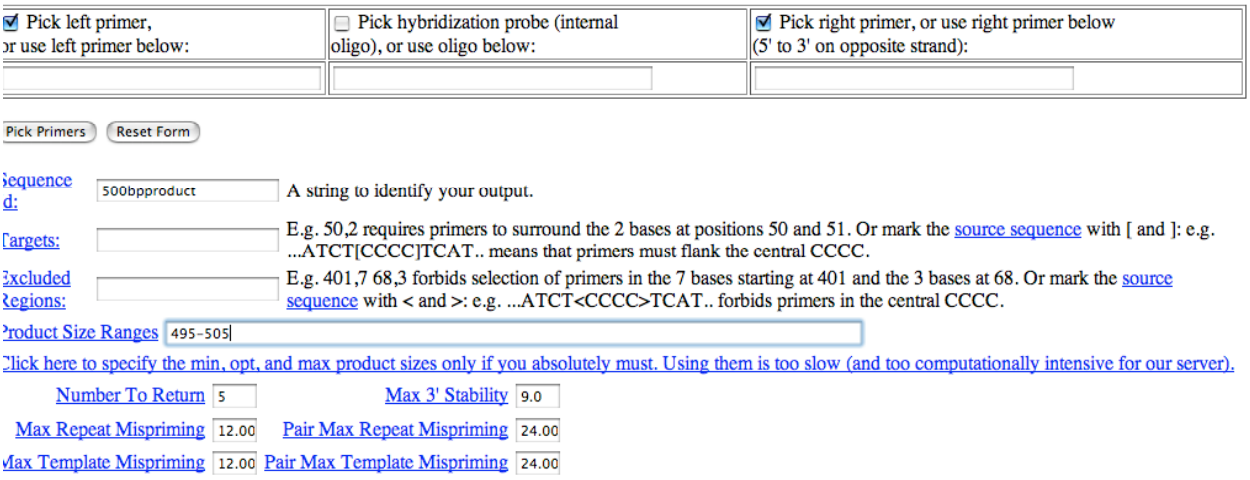
OPEN PRIMER 3

http://frodo.wi.mit.edu/

PASTE LAMBDA SEQUENCE in to source sequence box



- Tell program to pick right and left primers and indicate product size range
- Keep the size range narrow at first –open only if the program can't find primers in the range that you have set
- Leave defaults setting for other parameters



General Primer Picking Conditions:

The specificity of amplification of target sequences in a PCR reaction is controlled by the length and sequence of the primer (which affecs Tm) and the annealing temperature used in the PCR run. The sequence of the 3' end of the primer is also important.

- **Click on each parameter to view a description – just browse through the info here -- we'll come back and explore a couple of these parameters more closely**
- Start out using the default values for parameters in this section
- See appendix D for a definition of Tm
- Ignore **Other Per-Sequence Inputs** and what follows…

General Primer Picking Conditions

Primer Size	Min: 18	Opt: 20	Max: 27				
Primer Tm	Min: 57.0	Opt: 60.0	Max: 63.0	Max Tm Difference: 100.0			
Product Tm	Min:	Opt:	Max:				
Primer GC%	Min: 20.0	Opt:	Max: 80.0				
Max Self Complementarity: 8.00			Max 3' Self Complementarity: 3.00				
Max #N's: 0			Max Poly-X: 5				
Inside Target Penalty:			Outside Target Penalty: 0	Note: you can set Inside Target Penalty to allow primers inside a target.			
First Base Index: 1			CG Clamp: 0				
Salt Concentration: 50.0			Annealing Oligo Concentration: 50.0	(Not the concentration of oligos in the reaction mix but of those annealing to template.)			

☑ Liberal Base ☐ Show Debuging Info ☑ Do not treat ambiguity codes in libraries as consensus

Pick Primers | Reset Form

➔ Tell program to pick primers and then carefully inspect the primer-3 output – be sure to scroll down to the bottom of the page

Primer3 Output

WARNING: Numbers in input sequence were deleted.

```
No mispriming library specified
Using 1-based sequence positions
OLIGO            start  len      tm      gc%   any   3'  seq
LEFT PRIMER       2260   20   59.73    55.00  4.00  0.00 gtacggataccgcgaaagag
RIGHT PRIMER      2457   20   60.30    50.00  5.00  3.00 gctttttgctgtcccacagt
SEQUENCE SIZE: 4080
INCLUDED REGION SIZE: 4080

PRODUCT SIZE: 198, PAIR ANY COMPL: 5.00, PAIR 3' COMPL: 3.00
```

NOTE info a bottom of output page

```
ADDITIONAL OLIGOS
                    start  len      tm      gc%    any    3'  seq

 1 LEFT PRIMER       1556   20   60.53   55.00   6.00   2.00 ccagcaggagctggacttta
   RIGHT PRIMER      1760   20   59.93   45.00   2.00   1.00 cgtatccccttcgttttca
   PRODUCT SIZE: 205, PAIR ANY COMPL: 3.00, PAIR 3' COMPL: 1.00

 2 LEFT PRIMER        625   20   59.50   45.00   6.00   2.00 ccgacatgttgatttcctga
   RIGHT PRIMER       823   20   59.89   45.00   6.00   2.00 taattagcatccgcccattc
   PRODUCT SIZE: 199, PAIR ANY COMPL: 5.00, PAIR 3' COMPL: 2.00

 3 LEFT PRIMER       1230   20   59.82   50.00   5.00   3.00 gtggcgggttatgatgaact
   RIGHT PRIMER      1427   20   59.57   40.00   4.00   1.00 ggcaacatgaaaacgcataa
   PRODUCT SIZE: 198, PAIR ANY COMPL: 5.00, PAIR 3' COMPL: 1.00

 4 LEFT PRIMER       1032   20   60.08   50.00   3.00   3.00 gatggtgatgccgagaactt
   RIGHT PRIMER      1227   20   60.54   55.00   4.00   1.00 ccaccgacttttcacggtag
   PRODUCT SIZE: 196, PAIR ANY COMPL: 5.00, PAIR 3' COMPL: 1.00

Statistics
         con   too    in    in          no    tm    tm  high  high        high
         sid  many   tar  excl   bad    GC   too   too   any    3'  poly   end
        ered    Ns   get   reg   GC% clamp   low  high compl compl     X  stab    ok
Left   31766     0     0     0   319     0  3614 22272     0    16   105   760  4680
Right  32185     0     0     0   188     0  3243 23274     0     4   103   721  4652
Pair Stats:
considered 20667, unacceptable product size 20643, high end compl 2, ok 22
primer3 release 1.1.0
```

- What do any , 3' , pair any compl and pair 3' comple refer to?
- How do the alternate primer picks (additonal oligos) differ from the "chosen" primer?
- Note the statistics which indicates the number of primer candidates examined and why the primer was rejected

➔ If you are OK with the primer pick, copy the info shown on previous page and paste it into a word file – keep Courier 10 font so the collums will stay aligned.

➔ Pick your second primer pair

This time enter your left or right primer from above, and tell program to find the right or left primer

➔ Don't forget to enter a new product size

➔ Repeat instructions from above –don't forget to copy the primer 3 output info for each primer pair.

What if Primer 3 can't find acceptable an acceptable primer or primer pair?

- First examine the statistics to get clues as to what the limitations are
- Look at general primer picking parameters and make some conservative adjustments
- You can open the size range limits, but too big a size window will make the 100bp ladder less useful
- Try the other outside primer
- NOTE: exert a little effort here to generate nested PCR products -- the fewer primers you need for your ladder, the better

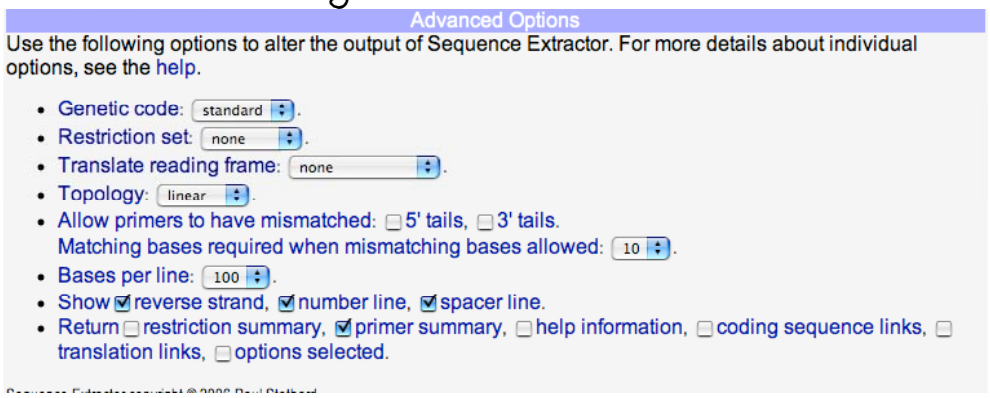
➔ REpeat until you have primer pairs that will generate 100,200,300,400,500 and 1000bp products

➔ In your word document, name your primers and make a list of each unique primer . In Appendix B, there are primer design rules put together by Jeff young. Do your primer sequences meet the standards suggested in this appendix? Briefly summarize your assesment

Go to Sequence Extractor and align primers on sequence:

Sequence Extractor: Display primers on DNA sequence
http://bioinformatics.org/seqext/
Sequence Extractor features
http://bioinformatics.org/seqext/features.html

- First clear primer and sequence box
- Paste in your sequence and your primer list as instructed
- Tell program that you don't want the sequence translated
- Don't allow any mismatches



Carefully examine output:

Inspect alignments

- Capture desktop picture of part of alignment (as above) and paste into word file
- Capture primer summary (below) and past into word document

Primer Summary
100L ACGATTTGCTGAACACACCA: 5712.
200L CAGCACCACAGAGTGCACA: 5607.
300L TTATCGCGCATCCTCAACTGT: 5499.
400R ATCAGATGCCGGGTTACCT: -5705.
500L GATGCACGTAAATCCCGTCT: 5309.
500R GTATGAGCCGGGTCACTGTT: -5809.

An ideal set of primer should anneal efficiently with the target template sequences and not with other template sequences or with each other. The Primer 3 program assesses the ability of primer pairs to hydrogen bond with each other at the 3' end (serving potentially as templates for DNA pol) or generally with each other at any position.

Max Complementarity

The maximum allowable local alignment score when testing a single primer for (local) self-complementarity and the maximum allowable local alignment score when testing for complementarity between left and right primers. Local self-complementarity is taken to predict the tendency of primers to anneal to each other without necessarily causing self-priming in the PCR. The scoring system gives 1.00 for complementary bases, -0.25 for a match of any base (or N) with an N, -1.00 for a mismatch, and -2.00 for a gap. Only single-base-pair gaps are allowed. For example, the alignment

```
5' ATCGNA 3'
   || | |
3' TA-CGT 5'
```

is allowed (and yields a score of 1.75), but the alignment

```
5' ATCGNA 3'
   ||  | |
3' TA--CGT 5'
```

is not considered. Scores are non-negative, and a score of 0.00 indicates that there is no reasonable local alignment between two oligos.

Max 3' Complementarity

The maximum allowable 3'-anchored global alignment score when testing a single primer for self-complementarity, and the maximum allowable 3'-anchored global alignment score when testing for complementarity between left and right primers. The 3'-anchored global alignment score is taken to predict the likelihood of PCR-priming primer-dimers, for example

```
5' ATGCCCTAGCTTCCGGATG 3'
            ||| |||||
        3' AAGTCCTACATTTAGCCTAGT 5'
```

or

```
5` AGGCTATGGGCCTCGCGA 3'
             ||||||
         3' AGCGCTCCGGGTATCGGA 5'
```

The scoring system is as for the Max Complementarity argument. In the examples above the scores are 7.00 and 6.00 respectively. Scores are non-negative, and a score of 0.00 indicates that there is no reasonable 3'-anchored global alignment between two oligos. In order to estimate 3'-anchored global alignments for candidate primers and primer pairs, Primer assumes that the sequence from which to choose primers is presented 5'->3'. It is nonsensical to provide a larger value for this parameter than for the Maximum (local) Complementarity parameter because the score of a local alignment will always be at least as great as the score of a global alignment.

Go to GeneWalker and analyze one pair of primers in detail

GeneWalker: analyze potential for primer artifacts/interactions
http://www.cybergene.se/primerdesign/genewalker/genewalker11.html
GeneWalker Manual
http://www.cybergene.se/primerdesign/help.html

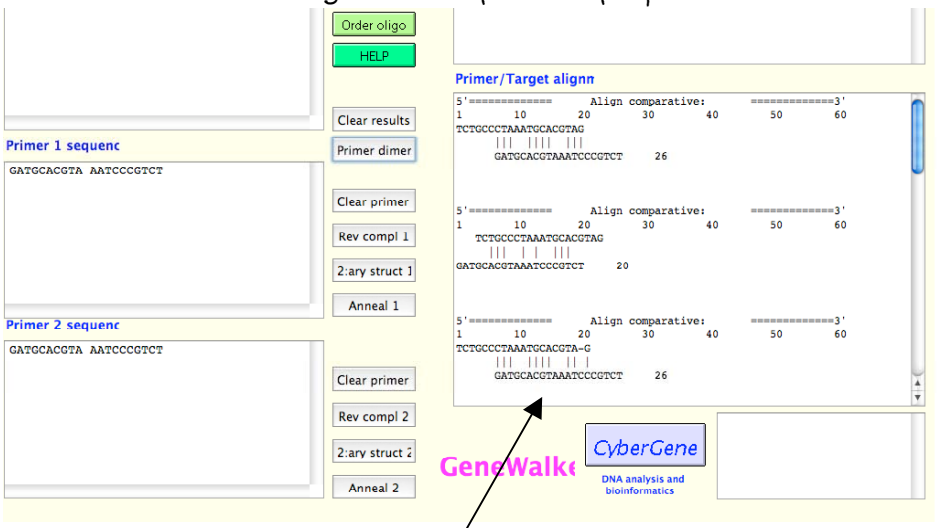
Pick a primer pair that with non-zero values here and with primers that have different any and 3' values. Have Genewalker show you primer dimer interactions.

OLIGO	start	len	tm	gc%	any	3'	seq
LEFT PRIMER	5309	20	59.96	50.00	6.00	0.00	GATGCACGTAAATCCCGTCT
RIGHT PRIMER	5808	20	60.00	55.00	8.00	1.00	GTATGAGCCGGGTCACTGTT

SEQUENCE SIZE: 48502
INCLUDED REGION SIZE: 8000

PRODUCT SIZE: 500, PAIR ANY COMPL: 5.00, PAIR 3' COMPL: 2.00

LEFT Primer annealing with itself -- ask for primer dimer

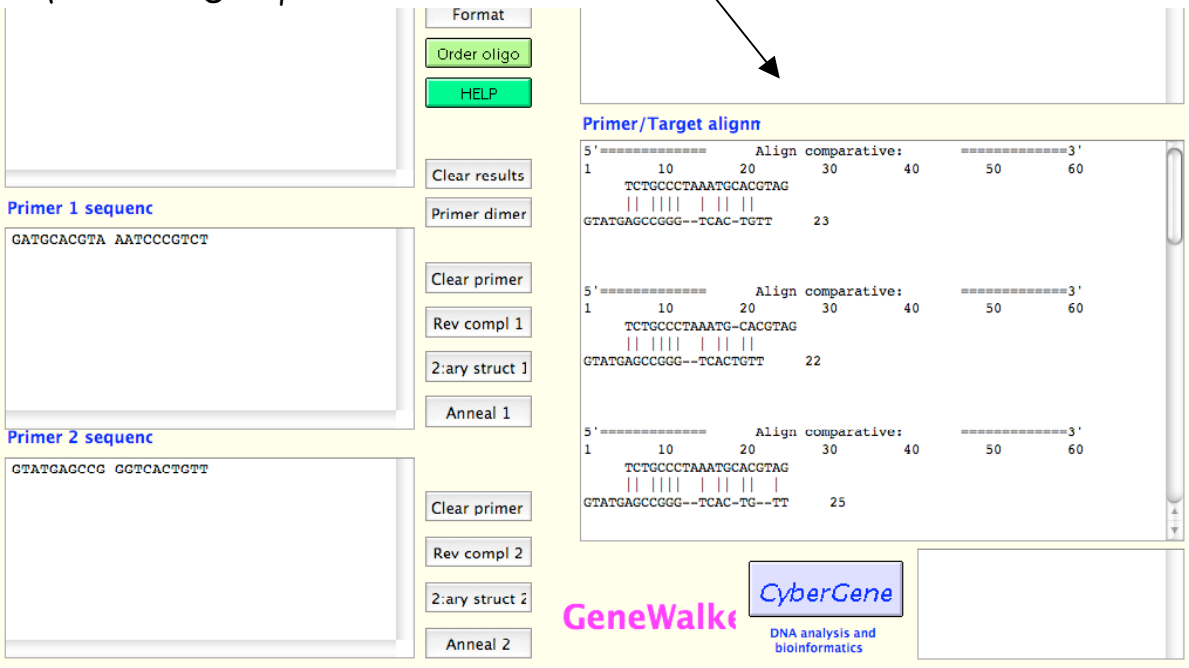


For each indicate the 5' and 3' ends

Which interaction would amplify in a PCR reaction?

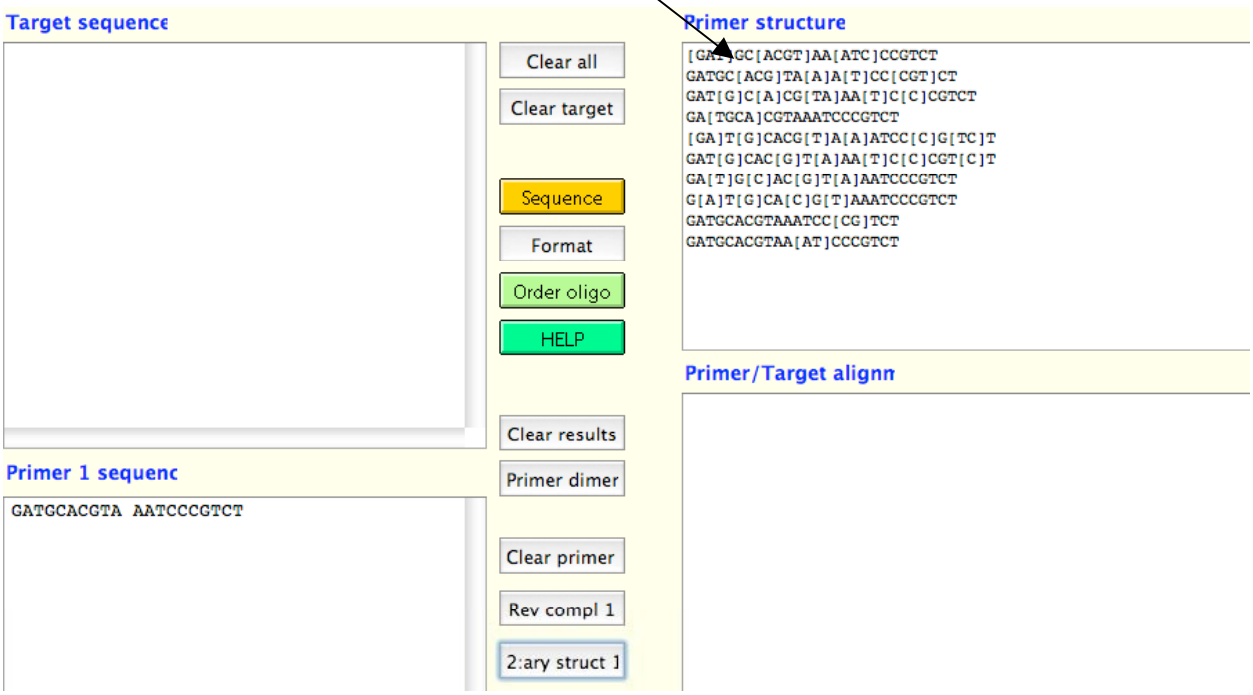
Repeat with RIGHT Primer annealing with itself

Left and right primers with each other



PRIMER intrastrand STRUCTURE

Paste in your primer sequence and ask for primer structure
The structure will be coded as shown
Go to genewalker help and draw out the structural meannig of this code –
first two lines



OPTIONAL for those who want to do more

Go to IDT-DNA Another site to analyze primer-primer interactions
Hey, this site will gives the ΔG's of the interactions!
http://www.idtdna.com/analyzer/Applications/OligoAnalyzer/Default.aspx

Appendix A What is Tm and why is it important?

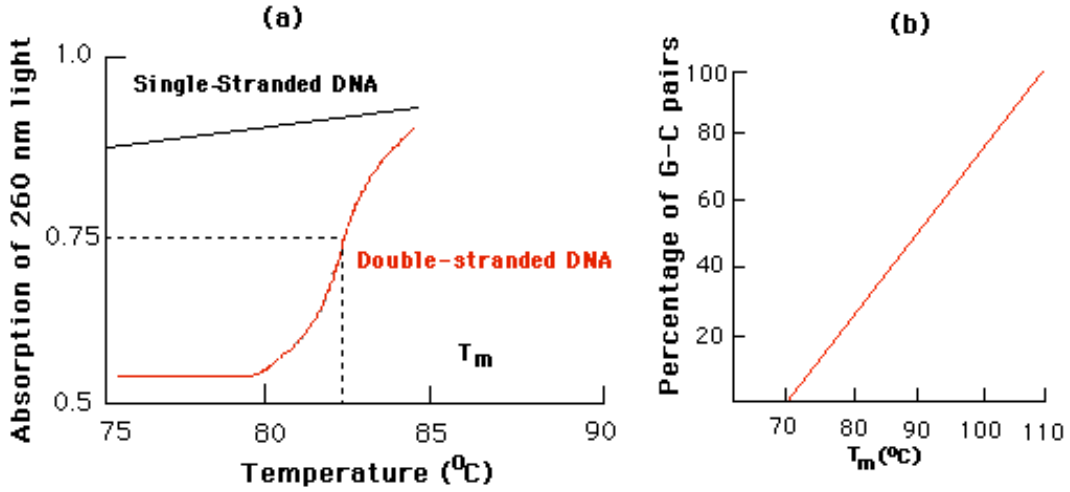


Fig. (a) The absorption of ultraviolet light of 260-nm wavelength by solutions of single-stranded and double stranded DNA. As regions of ds-DNA unpair, the absorption of light by those regions increases almost twofold. The temperature at which half the bases in a ds-DNA sample have denatured is denoted T_m (for temperature melting). Light absorption by single-stranded DNA changes much less as the temperature is increased. (b) The T_m is a function of the G + C content of the DNA; the higher the G + C percentage, the greater the T_m.

For primers shorter than 20 bases, T_m can be calculated as
$T_m = 4(G + C) + 2(A + T)$.

- Single base mismatches can significantly lower the Tm of the primer-template.
- Mismatch tolerance between template and primer -- is the 5' or 3' end the most critical?

Read about annealing temp and primer design at this site
http://www.mcb.uct.ac.za/pcroptim.htm

Appendix B *Here is a slightly different take on designing prime primers from Jeff Young:*

Primers

♦ Length**: Long (29 bp)**

1. G's or C's allowed at positions number 28 and 29 (3' end of primer):

Zero or One

nnnnnnnnnnnnnnnnnnnnnnnnnnn**GC**	**BAD**
nnnnnnnnnnnnnnnnnnnnnnnnnnn**GG**	**BAD**
nnnnnnnnnnnnnnnnnnnnnnnnnnn**TG**	**GOOD**
nnnnnnnnnnnnnnnnnnnnnnnnnnn**AT**	**GOOD**

1. GC content of primer: **34% to 50%**

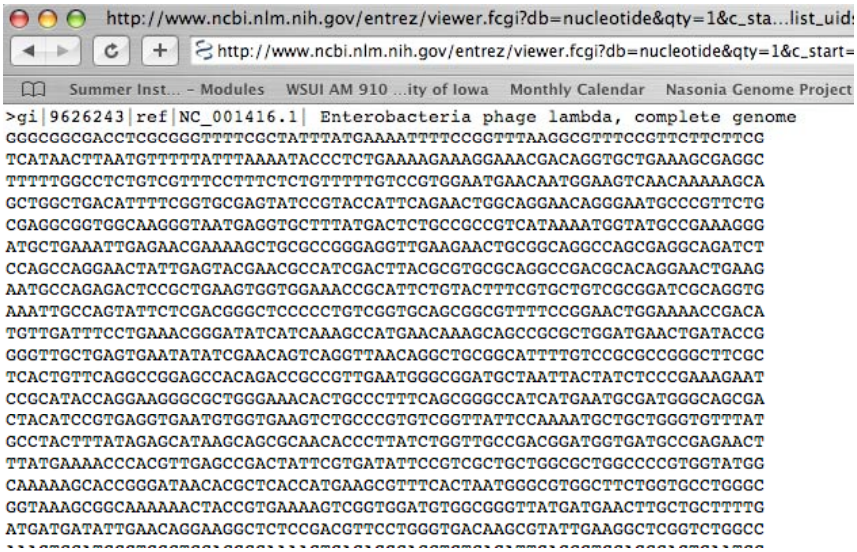
♦ TM: Don't worry about TM. If the primer is designed as above, TM will be in operable range.

2. Use 65º annealing temperatures for these primers.

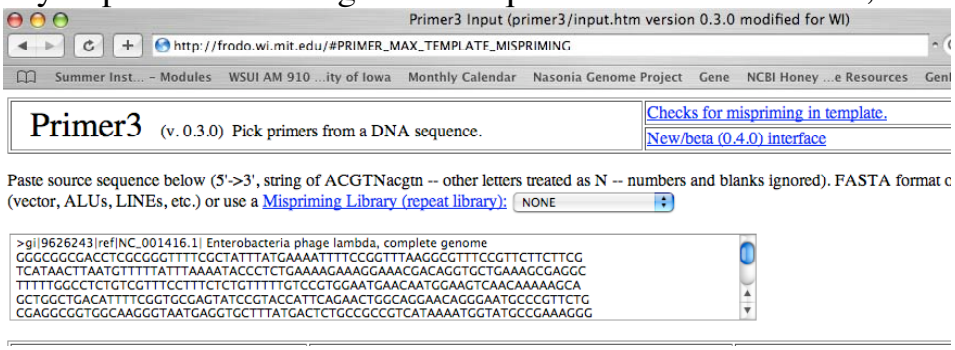
Avoid **High G+C content at the 3' end of the primer**
- avoid G+C > 50% in the 3' end (10 bp),
- avoid clusters of nucleotides, especially Gs and Cs.

Appendix C

Downloading sequence from NCBI
OR select FASTA format and send to TEXT



If you paste the entire genome sequence in IN PRIMER 3,



then you will need to limit the primer sequence to your assigned sequence block as follows:

Scroll down to other sequence input