Primer design

Where do primers come from?

- generally purchased from a company, who makes them by chemical synthesis

How do you design them?

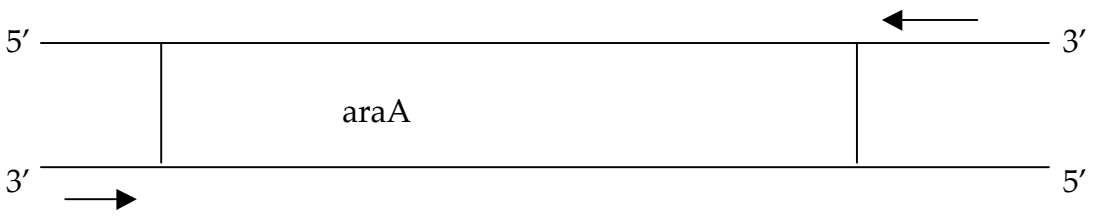
There are computer programs that will help you design primers for a given DNA sequence, but these programs help you find primers that follow the following "rules":

1. Primers should **flank** the DNA that you want to amplify (i.e. one on either side), such that the exponentially amplified product consists of the primer sequences and everything in between them.
2. Primers are generally between **18-25 basepairs** long
3. Each primer should have a Tm between **55-65˚C** and **G/C** content of 50-60%
4. Each primer should have a **3' end** that hybridizes very well to the template, but the **5' end** can be **initially** less complementary (or non complementary) to template
5. The primers should not form **"primer dimers"** or **"hairpins"**

Expand on each of these:

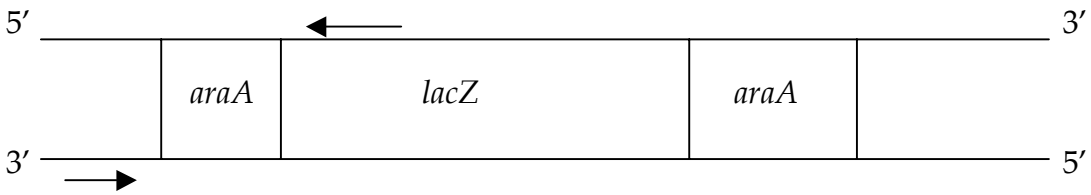
1. Location matters: Primers should flank the DNA you want to amplify

- example 1 (preparative):
  You want to amplify an entire gene for expression and protein production; thus, you'd want the PCR product to include the whole gene (from start codon to stop codon)



1. Location matters (continued)

- example 2: (analytical)
    - As in PCR, you want to determine if an insertion occurs and in what *ara* gene it occurred—getting information, but not necessarily to "do" anything with that DNA later



2. Length of primer determines <u>specificity</u> of binding to template DNA

*Probability of finding a given sequence "at random" <u>decreases</u> as the length of that sequence <u>increases</u>.

4 bases:

`5′-GATC-3′`
`3′-CTAG-5′`

Probability is $(1/4)^4 = 1/256$ (pretty common)

20 bases:

`5′-GATCCTAGATTTGATCGCGC-3′`
`3′-CTAGGATCTAAACTAGCGCG-5′`

Probability is $(1/4)^{20} = 1/1x10^{12}$ (pretty common)

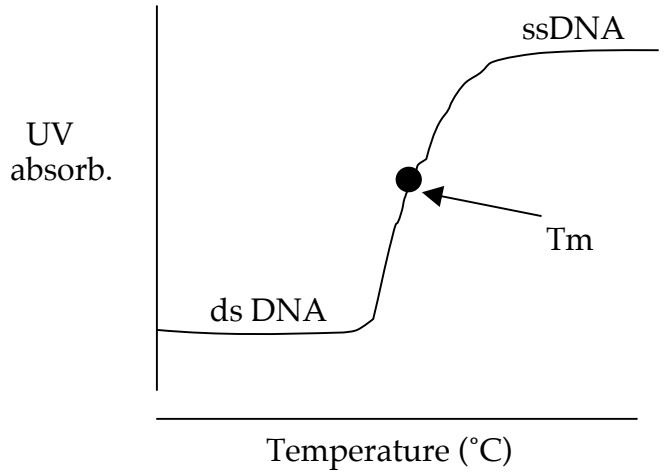
Compared to size of *E. coli* genome (4.6×10^6 bp) or human genome (3×10^9 basepairs)??

- If probability of finding a sequence is only 1/256, that sequence will be found in many places in the genome—primer is not specific for your "gene of interest"

- If probability of finding a sequence is 1/1012, you will only find that sequence near your "gene of interest" (and not just "at random")

3. Annealing temperature is dependent on the Tm of the primers

Melting curve of DNA (Prof. Rajbhandary lecture)



@ Tm = 1/2 of DNA is single stranded
1/2 of DNA is double stranded

In PCR:

- "Double stranded DNA" is primer bound to template
- "Single stranded DNA" is primer NOT bound to template

Above the Tm-->most DNA is single stranded
Below the Tm-->most DNA is double stranded

So...choose a primer **annealing temperature** about 5˚C **below** the Tm of your primers

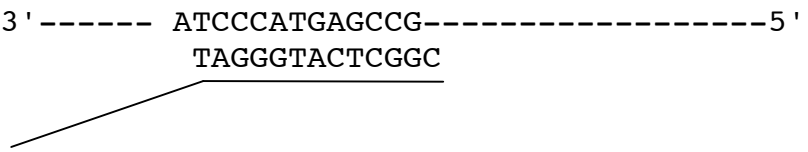
- choosing a much lower temperature for annealing leads to "mispriming" (binding of primers to sequences that are not 100% complementary to the primer)-->this reduces specificity

Calculating Tm:

For DNA between 18-25 basepairs:

Tm = 4˚C(# of GC basepairs) + 2˚C(# of AT basepairs)

4. Each primer should have a 3' end that hybridizes very well to the template, but the 5' end can be <u>initially</u> less complementary (or non complementary)



Why does this hybridize "well" to the template?

- The 3' end of the primer has a few G or C nucleotides. Since G/C form 3 hydrogen bonds with the template, it makes the primer/template complex stable. This is important for DNA polymerase to efficiently add nucleotides to the 3' OH of the primer

5. Do not form "primer dimers" or "hairpin" structures

<u>Primer dimer:</u>

- when the 3' ends of two primers can basepairs with each other as well as with the template DNA

Template DNA:

```
5'----------------------------------GTACTCGGC----3'
3'----CATGAGCCG-----------------------------------5'
```

```
Forward primer:                      Reverse primer:

5'---------GTACTCGGC-3'       5'----GCCGAGTAC-3'
```

BUT....primers can also basepair with each other:

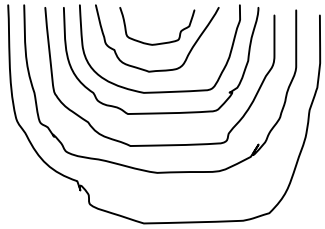
```
      5'---------GTACTCGGC-3'
               3'-CATGAGCCG----------5'
```

- Since in a PCR reaction, primers are in vast molar excess, you end up "losing" primers to this secondary reaction.
- If PCR primers form dimers with each other, they can't bind to template-->reducing the ability to copy the template and exponentially amplify the DNA-->lower yield of desired product

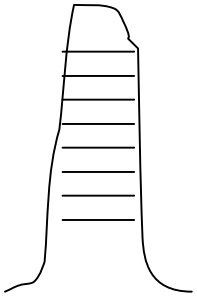
Hairpin:

- when the primer can form basepairs with itself (within its own sequence)

5'—GAGCCGTATGGGATACGGCAC—3'



Get a structure that looks like a hairpin:



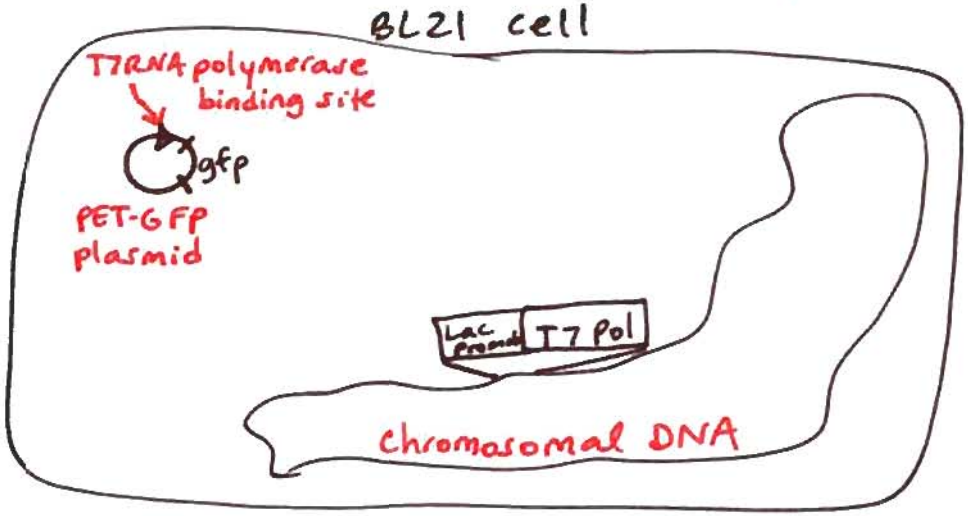
- If PCR primer forms a hairpin, it won't bind to template-->reducing the ability to copy the template and exponentially amplify the DNA-->lower yield of desired product

RDM Day 6

- Transformation again! BL21 cells this time

How do BL21 differ from AG1111?

1) BL21 has lower transformation efficiency
2) BL21 contains gene for T7 RNA polymerase (under control of lac promoter) so can express T7 RNA pol
3) BL21 deficient in (protease) - greater foreign protein stability

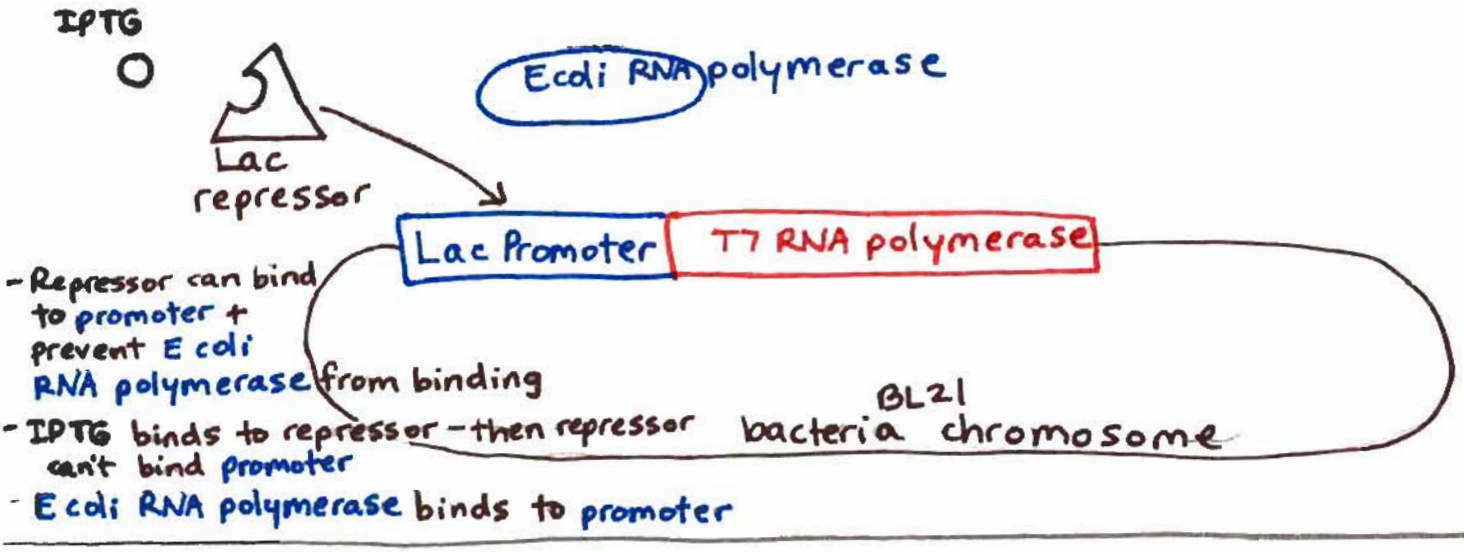


Lac Promoter "leaky" - allows low expression of T7 RNA polymerase without adding IPTG.

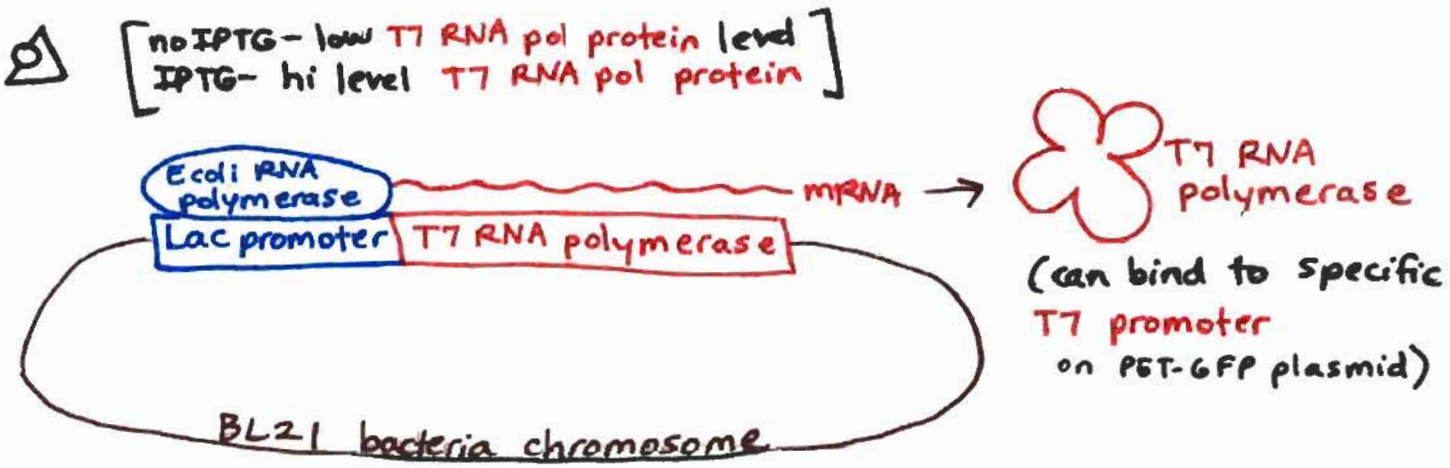
This little bit of T7 RNA pol is enough to cause transcription of gfp mRNA → makes protein → glows green!

Note- if you add IPTG, so much T7 Pol would be made that a HUGE amount of GFP protein would be produced- cells would get stressed and put the GFP into inclusion bodies (unfolded so don't glow)

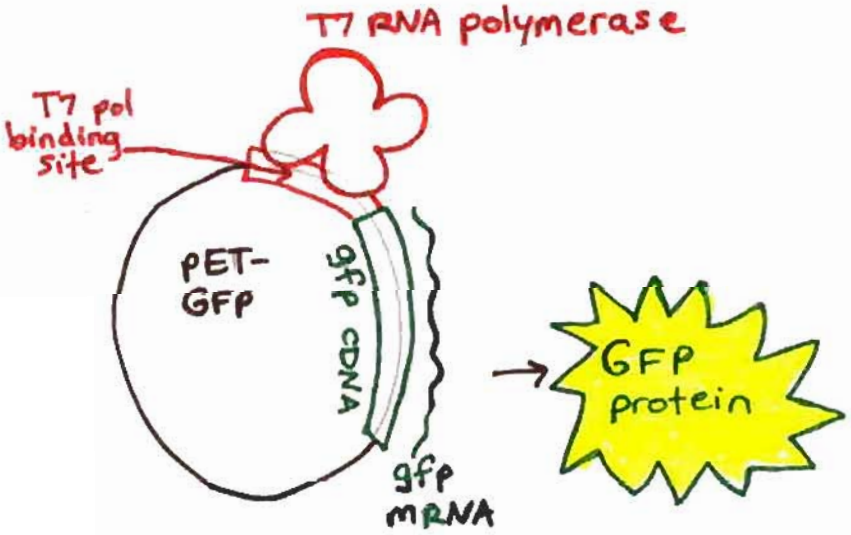
/
sequesters proteins



- Repressor can bind to promoter + prevent E coli RNA polymerase from binding
- IPTG binds to repressor - then repressor can't bind promoter
- Ecoli RNA polymerase binds to promoter



(can bind to specific T7 promoter on PET-GFP plasmid)



low level GFP - bacteria glows

too much GFP - GFP sequestered in inclusion bodies, unfolded - don't glow