Bioinformatics 2

Authors:
Mark Sibbald
Jeanine de Keyzer
Jurre Hageman

Version:
1-12-2022

Lesson:
Primer design and PCR with Benchling

Learning outcomes

After this lesson you can:

- Select restriction sites that are suitable for cloning a PCR product into a plasmid
- Design primers that introduce these restriction sites at the 5'and 3' end of the PCR product
- Perform an *in silico* PCR
- Use restriction enzymes to cut DNA sequences
- Ligate compatible ends of DNA sequences
- Perform restriction analysis on a construct
- Select suitable sequencing primers from a primer list

Introduction

F_1F_O-ATPase is an ATP synthase that is found in the cytoplasmic membrane of bacteria. It uses the proton motive force to drive ATP synthesis. ATP is formed using the energy that is released when protons move across the membrane down their electrochemical gradient.

F_1F_O-ATPase consists of a peripheral F_1 domain and a transmembrane F_O domain. Both domains are composed of multiple subunits (Figure 1). F_Oc is one of the subunits of the transmembrane domain. In *Escherichia coli*, F_Oc is encoded by the *atpE* gene.

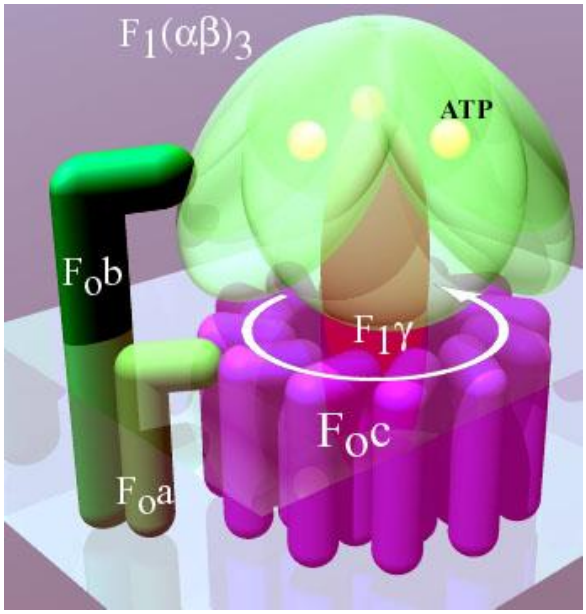


Figure 1. *The F_1F_O ATPase*

In this lesson, you will use Benchling to design a strategy to clone the *Escherichia coli atpE* gene in plasmid pBluescript II KS.

In the previous lesson, you simulated the cloning of the CFTR gene into a plasmid by manually adding restriction sites to the 5' and 3' ends of the CFTR open reading frame (ORF) and using these sites to insert the ORF into a plasmid.
Of course, in the laboratory you cannot simply add nucleotides to a piece of DNA. Instead, you need to introduce the restriction sites using a PCR with primers that are complementary to the insert and contain the desired restriction sites. When the PCR is performed with these primers the restriction sites are introduced at the ends of the insert (Figure 2).

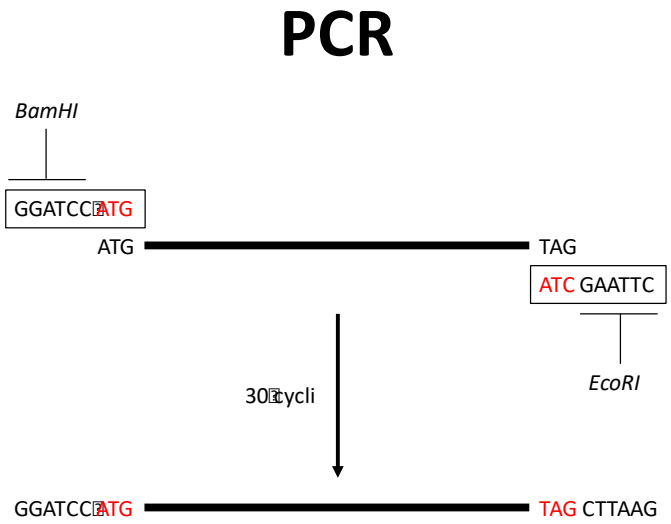


Figure 2. *Introduction of restriction sites in DNA-sequence using PCR. Own work CC-BY-SA.*

In this lesson, you will use Benchling to design PCR primes, test them in an *in silico* PCR and insert the PCR product in a plasmid. The approach consists of the following steps:

1) Selection of a set of suitable restriction sites
2) Primer design
3) *In silico* PCR
4) *In silico* cloning of the PCR-product into a plasmid
5) Control restrictions

Step 1: restriction site selection

On Blackboard you will find a FASTA file of the *Escherichia coli atpE* gene.

→ Open the file in Benchling and annotate the ORF (see previous lessons)

Question:

1) What is the length of the ORF (in bp)?

On Blackboard you will find a genbank file of the plasmid pBluescript II.

→ Open the file in Benchling

Initially, pBluescript II KS will open as a linear sequence.

→ Change the sequence into a circular sequence using the *information* button on the right side panel. (hint: change the topology from linear to circular and press *update information*)

As you can see in the map of pBluescript II, this plasmid contains an extensive multiple cloning site (MCS). Recall from the previous lesson that a strategy to ensure the orientation of the insert in the plasmid, is the use of two different restriction enzymes on each side of the insert. In this way, the insert will fit only in one orientation in the plasmid. Each restriction enzymes should not cut in the insert (why?) and must cut only once in the plasmid at the multiple cloning site (MCS).

Questions:

2) Which restriction enzymes are the first and last of the MCS?

3) Imagine that the following restriction enzymes are available in the lab:
   DrdI, HincII, NotI, PciI, TatI, XhoI.
   Use your knowledge from the previous lesson to choose a pair of enzymes that is suitable for the introduction of the *atpE* gene into pBluescript II (hint: notice the transcription direction form the lac promoter). Which enzymes do you choose?

4) What are the sequences of the restriction sites of these enzymes? Example notation: GAATTC;AAGCTT

5) In the plasmid map, the Lac-promoter is indicated. If you would like the *atpE* ORF to end up in pBluescript II in an orientation that is similar to the orientation

of the *lac*-promoter, which restriction site should be introduced at the 5’ end of the *atpE* ORF?

6) Which restriction site should be introduced at the 3’ end of the atpE ORF?

Step 2: primer design

Next you will design a primer set to introduce these restriction sites to the PCR product.

→ Go back to the annotated *atpE* sequence.

→ In the *primers* menu (red arrow in Figure 3) click *create primers* → *manual*

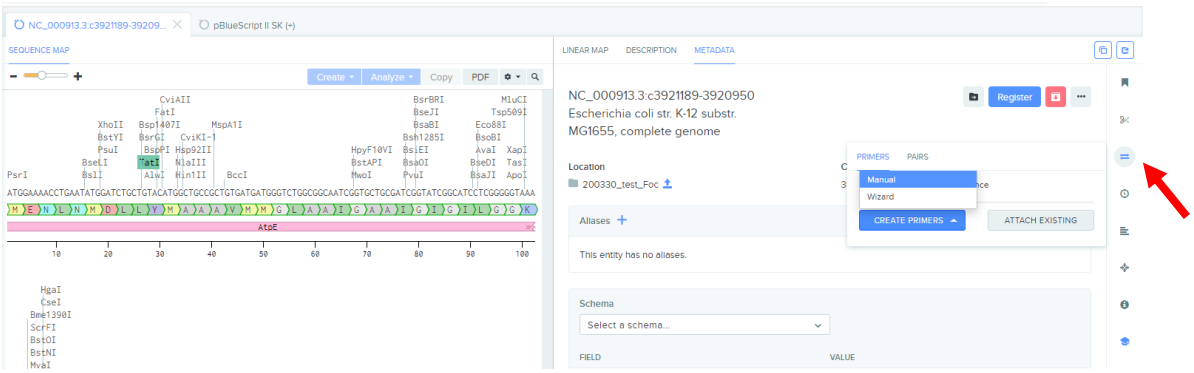


Figure 3. *Primers (oligos) menu in Benchling.*

First design the forward primer:

→ In the *sequence map*, select the first 20 nucleotides of the ORF (start at the start codon).

→ In *design primer*, click *set from selection* → *forward* (red arrow in Figure 4)

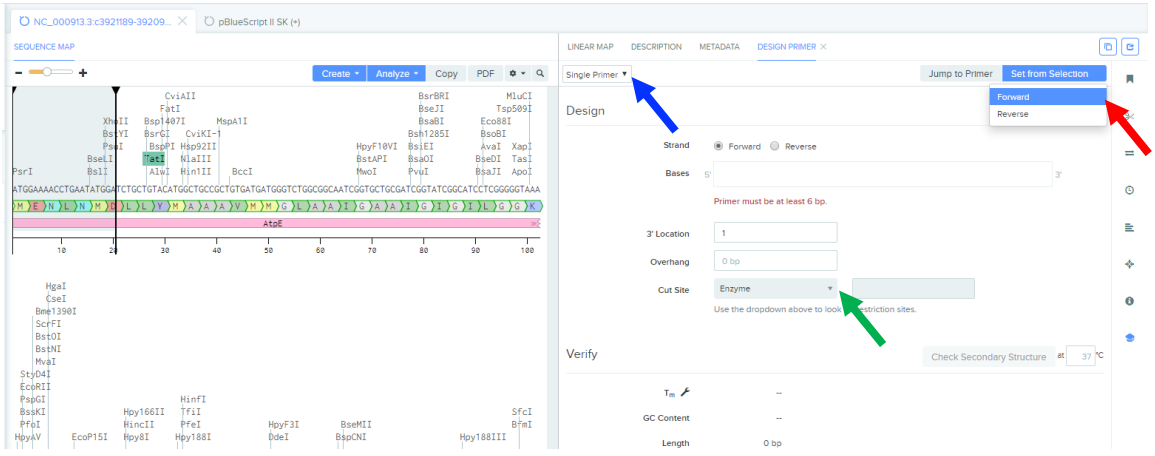


Figure 4. *Design primer from selected nucleotides in the sequence.*

→ At *cut site* (green arrow in Figure 4), choose the desired restriction site (see question 4)

→ Copy the sequence of the restriction site and paste it before the 5' end of the primer sequence

You have now created an **overhang**, a short sequence at the 5' end of the primer that does not anneal to the template.

→ Indicate the length of the overhang in *design primer*.

Now design the reverse primer:

→ In *design primer* switch from *single primer* to *primer pair* (blue arrow in Figure 4)

→ Select the last 20 nucleotides of the ORF, including the stop codon.

→ In *design primer*, click *set from selection* → *reverse*.
(Note that Benchling automatically converts the selected sequence into the reverse complement)

→ Choose the desired restriction site (see question 4) and paste it before the **5' end** of the reverse primer sequence.

→ Indicate the length of the overhang.

Questions

7) What is the melting temperature of the forward primer (in degrees Celsius)?
8) What is the GC content of the reverse primer (in %)?
9) What is the length of the forward primer (in nucleotides)?
10) What is the length of the reverse primer (in nucleotides)?

Save the primers in your project map. Tips:

– Use unique, recognizable names (NOT general names such as *forward* or *primer1*)
– Make sure that the primers are saved in the correct project map.

The primers will be linked to the template molecule automatically.

Step 3: in silico PCR

Now that you have designed the primers, you can check them using an *in silico* PCR.

→ Go to the *primers* menu

→ In the *pairs* tab (red arrow in Figure 5) click *create PCR product*, followed by *copy.* (Make sure you select the correct folder).

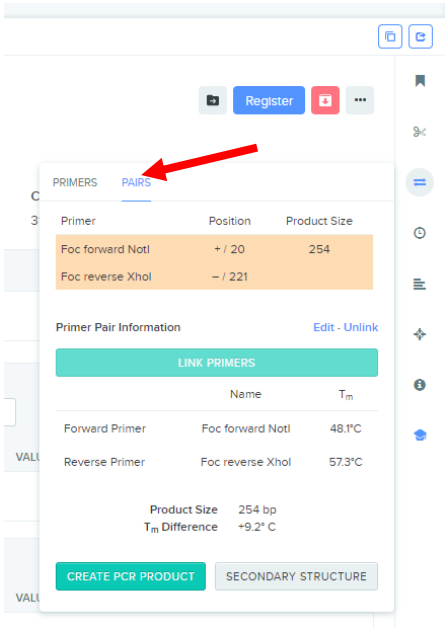


Figure 5. *Find primers to create PCR product in the tab 'Pairs'.*

Question:

11)What is the size of the PCR-product (in bp)?

Step 4: ligation of the plasmid and the insert

Use your knowledge from the previous lesson to ligate the PCR fragment in pBluescript II.

Tip: if Benchling gives the error warning *'The left sticky end does not match. The right sticky end does not match'* click *reverse orientation*. If you designed your primers well, the error message should disappear.

Question:

12) What is the size of the product (plasmid+insert) (in bp)?

Step 5: ligation controls

Each cloning experiment needs controls to assure that the ligation was successful and that the plasmid contains the correct insert.

Control 1: In silico analysis

Before you rush of to the lab, you should test whether the primers were designed correctly. In the previous step, you have already established that the PCR product contains the correct restriction sites and that it can be ligated in pBluescript II. Next you should check whether the insert contains the full length *atpE* gene.

As you can see the ORF is already indicated in the plasmid map. Benchling automatically copies the features that are annotated in the template into the PCR product. This is convenient, but also treacherous because Benchling also copies the annotation when only part of the ORF ends up in the PCR product.
It is therefore advisable to check whether the ORF is correct (in other words, always re-annotate the ORF).

→ Use your knowledge from the previous lesson to perform an ORF search on the new plasmid.

Question:

13) Does the length of the *atpE* ORF in the plasmid correspond to the length of the *atpE* ORF from question 1?

Control 2: Restriction analysis

After performing a ligation in the laboratory, it is necessary to check whether the plasmid contains the insert and that the ligation was successful.

→ Use your knowledge from the previous lesson to create a representable figure that only shows the restriction enzymes that cut only once in the construct.

Question:

14) Which two enzymes would you use in the laboratory to test whether the plasmid contains the insert? Example answer: KpnI, EcoRI
15) Which two enzymes would you use in the laboratory to test whether the ligation was successful? Example answer: KpnI, EcoRI

→ Cut the constructs with your choice of restriction enzyme and analyse the result check the '*Virtual Digest*' tab.

Question:

16) Which fragment sizes do you obtain with a control digestion for the presence of the insert (in bp)? Example answer: 2300, 345
17) Which fragment sizes do you obtain with a control digestion for a correct ligation (in bp)? Example answer: 2300, 345
18) Create a '*Virtual digest*' which shows the pBluescript II (negative control) and pBluescript with insert, cut with the selected enzyme pairs. Create a screenshot (check how this works for your OS). Save the picture as png file and upload in the webform.

Control 3: Sequence analysis

During a PCR errors can occur. Since most errors cannot be detected by restriction analysis, the insert is usually sequenced.

For sequencing you need so-called *sequencing primers*. Many commercial plasmids contain binding sites for standard sequencing primers, eliminating the necessity to design new primers for each sequencing project.

Imagine your lab has the following standard sequencing primers in stock:

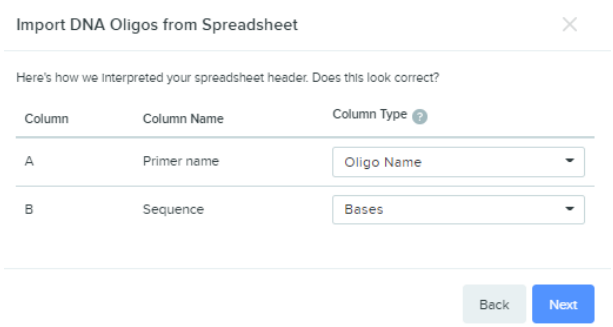
Primer 1 TGTAAAACGACGGCCAGT
Primer 2 GTAGCTCTTGATCCTTG
Primer 3 GCAGGAGAACATGTAAC
Primer 4 CAGGAAACAGCTATGAC
Primer 5 TACGCAGCAGAGGAATC

How do you select the sequencing primers that are suitable for your construct?

In Benchling, you can import a list of primers and check if and where these primers anneal to your construct.

→ Click the + *sign* at the left side of the page → *Oligo* → *Import DNA oligo from spreadsheet*

→ Select the Excel file. Check if the column names match the correct column type.



→ Click *next*

Select the pBluescript vector with the cloned insert (AtpE)

→ At the right side of the page open the primer menu and click *attach existing.*

→ Change the T_m selection range to 40-100 degrees Celsius (as one primer has a rather low T_m value):

→ Select the correct project folder and click *find binding sites.*

In the primer search result, you see which primers from your project map anneal to your construct. The table also gives information on the location of the binding site and the melting temperature. Since the Sanger sequencing results usually have poor quality in the first 15-40 bases of the sequence due to primer binding, sequence primer should anneal at least 50-60 bases upstream of your sequence of interest.

Question:

19) Why are some locations in the results table preceded by a '+' and other by a '-'?

20) Which of the five primers above can be used to sequence your insert?

The end...