Laboratory 3:
Polymerase Chain Reaction (PCR)

1. Polymerase chain reaction

Objectives

- To amplify the Amy2 coding sequence for cloning
- To analyse PCR products by agarose gel electrophoresis

Introduction

For cloning, DNA sequences can be amplified by **PCR** from a tissue- or species-specific **cDNA pool**. Amplification primers contain restriction sites for cloning of the amplicon into a suitable **plasmid vector**. The purpose of Polymerase Chain Reaction (PCR) is to amplify DNA fragments (also called amplicons). The reaction leads to the exponential amplification of specific DNA sequences and is routinely used in laboratory research (cloning, sequencing etc) and diagnostics (COVID-19, genetic testing etc). The impact of PCR in molecular biology and thus on development society has been profound. In 1993 Kary B. Mullis received the Nobel Prize in Chemistry for the "**invention of the polymerase chain reaction (PCR) method".**

The cycling steps

There are three major steps in a PCR, which are usually repeated for 20-40 cycles, using an automated thermal cycler that heats and cools the reaction quickly.

1. **Denaturation at 94-98°C** typically 10-30 s
   During denaturation, the double stranded template melts and forms single strands. The polymerase dissociates from the template and is inactive. Most programs use an initial denaturation step preceding the cycles

2. **Annealing 50-60°C**, typically 10-30 s
   The annealing temperature depends on the sequence and length of the primers. This parameter is most frequently optimised. Start with an annealing temperature 5°C below the lowest primer's Tm. If no/weak amplification is observed lower the annealing temperature, if several unspecific bands are observed increase the annealing temperature.

3. **Extension at 72°C,** typically 0.5-3 min
   This is the ideal working temperature for the thermostable DNA polymerase and duration depends on amplicon length and extension rate (typically 1000 bp/min). The primer-template interaction allows the polymerase to bind and start to copy the template. The polymerase adds dNTPs to the 3' end of the primer and therefore has 5'→ 3' polymerase activity. Primers that are bound non-specifically to the template are dissociated at the higher temperature. Many commercial thermostable polymerases exist that differ for instance in error rate, proof-reading activity, extension rate and extension temperature.

Because both strands of the template are copied during PCR, there is an **exponential** increase of the number of copies of the cDNA template. Starting from one copy of the template, after one cycle there will be 2 copies, after two cycles there will be 4 copies, and three cycles will

result in 8 copies and so on. In a typical reaction there are many templates present at the start of the PCR that will be amplified exponentially.

Note that standard PCR is not quantitative. In order to measure the relative amount of a specific target, one needs to perform quantitative real-time PCR (qPCR), where the **amplification** product is **measured at each PCR cycle**. Two methods are commonly used:

1. Incorporation of a fluorescent non-specific DNA-binding dye (SYBR Green) into the PCR reaction.
2. Sequence-specific DNA oligonucleotides (TaqMan probes) labelled with a fluorescent reporter, allowing detection only after hybridization of the probe with its complementary DNA target.

Computer software will calculate the relative target expression, compared to a constitutively expressed gene (housekeeping gene).

1.1 PCR amplification of *Amy2* coding sequence

In this experiment you will amplify the coding region of *Amy2* from the **pancreas-specific cDNA pool**. In the next session you will clone the PCR product as a **translational fusion** into the mammalian expression vector **pcDNA6/*myc*-His A**. PCR amplification of the coding sequence will be done with a high-fidelity DNA polymerase (Phusion DNA polymerase), which has proofreading activity.

Exercises can be done in Benchling while the reaction is in the thermocycler (part 2). To retrieve the reference sequence from **NCBI** (National Center for Biotechnology Information) database:

- Select in the drop down menu the **nucleotide** database
- Enter the accession number (e.g. NCBI reference sequence for *Amy2:* NM_001042711.2)
- Find the coding sequence (CDS).

We will use the following *Amy2* primers designed for cloning of the target sequence and subsequent protein expression in mammalian cells:

Mm-Amy2_for:	5'- CA***ggatcc***ACCATGAAGTTCGTTCTGCTGCTTTC -3'
Mm-Amy2_rev:	5'- GG***ctcgag***TTC**A**TACAATTTTGAGTCAGCATGGATTG -3'

Forward and reverse primers contain respectively *BamH*I and *Xho*I restriction sites (bold italics) at their 5' end for directional cloning into the expression plasmid (for most enzymes + 2-3 extra bases greatly increase the efficiency of digestion). The forward primer contains a Kozak sequence (underlined) for efficient translation initiation in mammalian cells, including the endogenous initiation codon ATG. The reverse primer contains a mutated stop codon (TAA>TA**T**) to allow translation of the tag.

Indicate the part of the primer sequences that is gene specific (binding the template)

Mm-Amy2_for: _____

Mm-Amy2_rev: _____

Calculate the expected size of the PCR amplicon (can be solved by hand or with Benchling)

Amy2 CDS + sequences added by the primer = _____ bp

General guidelines for designing PCR primers:

- Standard primers complementary to the target sequences are **~18 - 25 nucleotides long**.
- Composition ~**40-60% GC**.
- Design primer pair (F+R) with **similar melting point temperature (Tm) ~55-65 °C**.
- **GC Clamp:** The presence of G or C bases within the last five bases from the 3' end of primers promotes stronger bonding (G and C have stronger bonds than A and T).
- Avoid sequences that form primer-dimers or hairpins.
- **Forward primer** is **complementary** to the **non-coding strand,** The primer sequence corresponds to the coding strand sequence retrieved from the NCBI nucleotide database.
- **Reverse primer** is the **reverse complement** of the **coding strand**.
- **Primer sequences** (forward and reverse) are always written from **5' to 3'**.
- **Primers are chemically synthesized** and can be ordered from an oligonucleotide synthesis facility (for example Microsynth).

Primer guidelines for restriction enzyme cloning

1. Choose restriction enzymes that do NOT cut the sequence to be cloned (Benchling or other software).
2. Add suitable restriction sites to primers at the 5' end. Add 2-3 extra bases to the 5' end of the restriction sites for optimal cleavage.
3. For **fusion proteins** (e.g. Myc-tag) the coding sequences of insert and tag must be **in-frame**. For C-terminal fusions remove (or mutate) the stop codon of template (example *Amy2*) in the reverse primer.
4. **Add** the **Kozak sequence** (e.g. ACC**ATG**G; see http://en.wikipedia.org/wiki/Kozak_consensus_sequence) for optimal **protein expression in mammalian cells** in the 5' primer (if not present naturally or provided by the vector). Note that for bacterial expression a ribosome binding site (RBS) sequence is essential.

Materials

- cDNA pool from mouse pancreas 1:20 dilution from previous lab
  - (+) RT sample (with reverse transcriptase)
  - (−) RT control (without reverse transcriptase)
- Positive control PCR (plasmid containing *Amy2* coding sequence without myc-HIS tag)
- Forward and reverse primers (5 µM each; see sequences above)
- 10 mM dNTPs Nuclease-free water
- Phusion High-Fidelity (HF) DNA polymerase (2 U/µl)
- 5X PCR (HF) buffer
- Thermal cycler

Procedure

Each group will prepare four PCR reactions in 0.2 ml tubes. First prepare a **mastermix** in 1.5 ml tubes with ingredients common for all reactions. Then distribute the appropriate volume to the four 0.2 ml tubes, thus reducing the number of pipetting steps, errors, pipette tips and time. The experimental setup includes a positive and negative control that allows interpretation of the PCR.

1. As you are working with small volumes, quick spin all tubes before pipetting. Pay attention to the liquid in the tip during pipetting.
2. Prepare the **mastermix in a 1.5 ml tube** according to the table below (right column). To ensure that we have enough volume for four reactions we prepare slightly more (here for 5 reactions). To prevent contamination, always change tips between each pipetting step. The theoretical volumes per tube are shown in light grey.

Amounts in (µl)	**A** (+) RT sample	**B** (-) RT control	**C** positive control PCR	**D** negative control PCR	**mastermix (for 5 rx)**
H_2O	15.8	15.8	15.8	15.8	**79**
5X PCR HF buffer	5	5	5	5	**25**
10 mM dNTPs	0.5	0.5	0.5	0.5	**2.5**
Primer mix*(5µM)	2.5	2.5	2.5	2.5	**12.5**
Phusion® HF DNA polymerase	0.2	0.2	0.2	0.2	**1**

*Contains both Amy2-for and Amy2-rev primers

3. Vortex mastermix and spin down.
4. Label **four 0.2 ml tubes** with A-D (and your group number) and **distribute 24 µl** of the mastermix to each.
5. Add **1 µl** of the corresponding **template** to each tube. Vortex and spin down.

Amounts in (µl)	**A**	**B**	**C**	**D**
Mastermix	24	24	24	24
Template	(+) RT sample 1	(−) RT control 1	Plasmid DNA 1	H_2O 1
Total volume	25	25	25	25

6. Place your tubes in the PCR cycler and start the program (several groups share a block).
7. Thermal cycler program (takes about 1h 10 min)

Initial denaturation	94°C	2 min	1x
denaturation	94°C	15 sec	22 x
primer annealing	62°C	15 sec	
extension	72°C	60 sec	
final extension	72°C	5 min	1 x
end of reaction	10°C	∞ hold	1 x

It is good practice to anticipate the expected result of the agarose gel electrophoresis before receiving the data, and in addition it helps you later to interpret the results. The size marker, also called DNA ladder is loaded on the same gel as the samples. It is a solution of DNA molecules with known size that allows to estimate the size of unknown DNA molecules that were separated based on their mobility during gel electrophoresis.

While the PCR is running use the box below to draw the expected outcome for samples A-D, M=marker; then discuss with your group members and assistant. Which samples show an amplification product? Of what size? Which samples should not show an amplification product?

draw expected gel image
M A B C D

1.2 Analysis of PCR products by agarose gel electrophoresis

Agarose gel electrophoresis is used to separate DNA fragments by size. Similar to RNA, the negatively charged DNA migrates towards the positive pole within an electric field, with smaller molecules migrating faster than longer ones. By running a size marker (called ladder consisting of several DNA fragments with known length) on the same gel we can estimate the size of DNA fragments (for example PCR product or restriction fragment).

The DNA is visualized by a specific dye for nucleic acids (e.g. GelRed or Ethidium bromide) in the agarose gel. Because of its chemical structure, the dye intercalates between DNA base pairs that fluoresce in UV light. To follow the migration during the run by eye the loading dye contains several visible dyes. For instance, bromophenol blue comigrates with 300 bp fragments and xylene cyanol FF comigrates with 4000 bp in 1% gel.

Task

We will run PCR products on an agarose gel to check amplification of *Amy2* coding sequence.

Material

- PCR reactions tubes A-D- the green PCR buffer contains already loading dye
- 1 % agarose gel in Tris-acetate-EDTA (TAE) buffer (one gel for two groups)
- 1X TAE buffer
- GelRed
- 1 kb ladder (Generuler, Fermentas)

Preparation of agarose gel

1. Prepare 50 ml of 1% agarose in 1X TAE buffer (Tris-acetate-EDTA) buffer in a 200 ml Erlenmeyer. Melt the agarose (0.5 g /50 ml) in a microwave oven (no aluminium foil or Parafilm!). Mix until completely dissolved.
2. Add 1:10’000 parts GelRed dye (5 µl/ 50 ml). Mix well. Let cool down for a few minutes.
3. Set up the gel-casting chamber into the frame. Place **two combs** onto the frame at the top and the middle of the chamber.
4. Pour the melted agarose into the gel-cast (don’t overfill) and wait until the gel solidifies before loading your samples (1 hour).
5. Cover the gel with 1X TAE buffer and CAREFULLY remove the combs

Loading of agarose gel

1. Load your samples as indicated below, note the different volumes. Since the green PCR buffer contains already a loading dye, the samples can be directly loaded. We suggest you keep an empty well between samples to avoid carryover. In case you load differently, make sure to write it down in your lab notebook.

Upper wells (group _____)

Lane	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Sample	M*	_	A	_	B	_	C	_	D	_
µl	**3.5**	_	10	_	10	_	10	_	10	_

Middle wells (group _____)

Lane	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Sample	M*	_	A	_	B	_	C	_	D	_
µl	**3.5**	_	10	_	10	_	10	_	10	_

2. Run the gel at 120 Volts for 20 minutes. Careful: the dye front of the upper wells must not enter the middle wells!
3. Remove the gel from the chamber and take a photograph at the UV transilluminator. Save image with group number; results will be put onto Moodle.

Before you insert the gel picture into your notebook, crop parts that contain no information, label the image outside of the image (do not cover data, no handwriting) by indicating lanes, size marker, amplicon size. You may use any suitable program (Powerpoint or Inkscape). Look at the data sheet of the DNA molecular weight standard (1 kb ladder; available in SLIMS) to determine the approximate length of your PCR amplicons. Remember to add a figure legend below the image. Describe the pattern of the bands that you observe as well as the sizes. Do you see the expected fragments? What do you see in the controls? Discuss the results in your lab notebook.

2. Recombinant DNA Technology: Plasmid Design

Objectives

- To design primers suitable for cloning a target cDNA into an expression plasmid
- To use Benchling software for plasmid design

Introduction

Plasmids are small circular extrachromosomal DNA elements found in a wide variety of bacterial species. These naturally occurring DNA elements have been engineered to optimize their use as vectors in DNA cloning. They replicate and are inherited independently of the bacterial chromosome.

For cloning, the PCR product (insert) as well as the target plasmid (vector) are both digested with the corresponding **restriction enzymes** to provide compatible ends for subsequent **ligation**. The resulting **recombinant plasmid** will be transformed into *E. coli* and can then be used for expression of a protein of interest in bacteria or eukaryotic cells (example: insulin, antibody).

Plasmid Key Features

Antibiotic resistance gene: used for selection of positive cells (bacteria or eukaryotic cells); for example ampicillin for selection in bacteria.

Multiple cloning site (MCS): contains several unique restriction sites to facilitate cloning.

Origin of replication (ori): allows replication of the plasmid within the host cell. The number of plasmids per cell is variable and depends on the type of ori (high vs low copy numbers). Mammalian expression plasmids contain an additional eukaryotic ori such as SV40.

Expression plasmids (also called expression vector) are designed for **protein expression** in prokaryotic or eukaryotic cells. They are engineered to contain regulatory sequences (see below) for efficient transcription and translation of the cloned DNA sequence in the host cell.

Plasmid Features for Expression (Transcription & Translation)

1. **Transcription**

Promoter: different promoters are used for transcription in bacteria and eukaryotic cells (they can be on the same plasmid). Cytomegalovirus (CMV) and SV40 promoters are commonly used in mammalian expression vectors. The T7 promoter drives transcription in bacteria by T7 RNA polymerase and is often used for sequencing the insert with standard primers (T7 forward primer).

Polyadenylation signal: Polyadenylation occurs on a consensus sequence (AAUAAA) on the 3' end of the mRNA. Here, the polyA polymerase adds a string of adenine residues, typically 50-250 in higher eukaryotes. Frequently used polyadenylation signals on commercial plasmids are for example Bovine growth hormone (BGH) and SV40. The BGH sequence may be used for sequencing the insert with standard primers (BGH reverse primer).

2. **Translation**

Translation initiation consensus sequence for efficient translation:

- Bacterial expression: **ribosome binding site (RBS)** (also called Shine-Dalgarno sequence)
- Mammalian expression: **Kozak** sequence

Start codon (ATG): required for protein translation, present in the plasmid or the insert sequence.

Stop codon (TAA, TAG, TGA): required for translation termination, present in the plasmid or the insert sequence.

Protein Tags

Protein tags are sequences frequently contained in the expression plasmid that allow purification (example His-tag, GST) or detection (Myc-tag, GFP) of the fusion protein. The coding sequence of your gene of interest is inserted in-frame relative to the tag (at the N-terminus or C-terminus) and upon expression a fusion protein containing the tag is synthesized. This can be very useful to observe the tagged protein in cells, if there are no specific antibodies against the protein of interest or to distinguish the recombinant protein from the endogenous one. When designing C-terminal fusion proteins the endogenous Stop codon must be removed in order to translate the tag.
Some commonly used tags:

- **His-tag** (polyhistidine, 1 kDa): consists of 6 histidine residues that bind with high affinity to metal ions such as cobalt or nickel.
- **Myc-tag** (1.2 kDa): for immunodetection and protein purification.
- **GFP** (green fluorescent protein, 26.9 kDa): for life imaging (fluorescent microscope) and immunodetection.
- **GST** (glutathione-S-transferase, 26 kDa): has a high affinity for glutathione and GST-tagged proteins can be purified from crude cell lysate on glutathione-coated agarose beads.

Cloning using Restriction Enzymes

Restriction enzymes are important tools for the cloning of DNA fragments into a plasmid vector. The restriction enzymes from bacteria recognize a specific sequence within DNA and cleaves it. This primitive “defense mechanism” allows bacteria to destroy foreign DNA (for example from phages). The **recognition sequence** is often four or six base pairs (bp) in length and palindromic that can produce sticky ends (overhang) or blunt ends (no overhang). For example, the recognition site for *Eco*RI is GAATTC as shown below. After binding to the recognition site, the enzyme cleaves double stranded DNA (Figure 3). Hundreds of restriction enzymes are commercially available and optimized for cloning (have now been identified (e.g. have a look at a restriction enzyme database; example *Bam*HI).

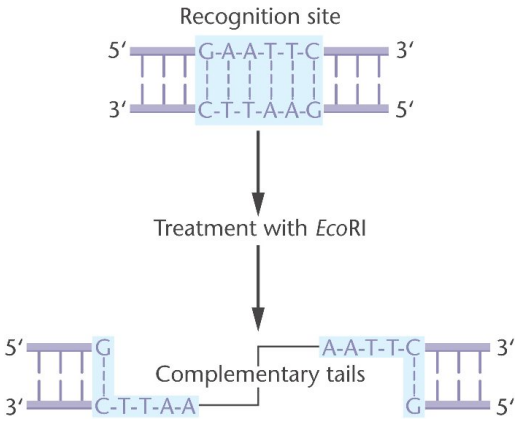


Figure 1. A schematic representation of the *Eco*RI restriction site is shown. The enzyme recognizes the sequence GAATTC and cleaves each strand of DNA between G and A thereby producing overhanging / sticky ends.

When a single restriction enzyme is used for cloning, the DNA fragment will be inserted in two directions in a plasmid. To clone inserts with a defined direction (**directional cloning**), we use two restriction enzymes that create different 3' or 5' overhangs that are not compatible with each other (example *Eco*RI and *Bam*HI). The choice of restriction enzymes depends on their presence in the MCS of the plasmid which generally offers multiple strategies. Ligation of fragments with blunt ends (example *Eco*RV) is a comparatively inefficient reaction and will generate plasmids containing inserts in both directions.

During ligation the annealing of complementary cohesive (sticky) ends allows DNA ligase to catalyze the formation of phosphodiester bonds between 5'-phosphate and the 3'-hydroxyl residues on vector and target DNA.

Other Cloning Strategies

There are many different cloning methods, and often several approaches may be suitable for a given project. You may choose based on cost, speed, availability of starting material or prior experience. Cloning strategies that do not employ restriction enzymes are very useful, since they offer more flexibility:

PCR cloning (T/A cloning): PCR products may contain a single nucleotide (Adenine) extension at the 3' end that is added by some polymerases (Taq) in a template-independent mechanism. Such PCR products with A-overhangs can be directly ligated into commercial plasmids with a complementary T-overhang (provided in the open form). The resulting plasmids contain inserts in both directions.

Seamless cloning (Gibson Assembly): To assemble the first synthetic genome Daniel Gibson at the Craig Ventor Institute developed a method that allows assembly of multiple DNA fragments regardless of their size and end compatibility (Gibson et al. 2009, Gibson et al. 2010). Since this technique can assemble any sequence (PCR or synthetic) and is quick and easy (single tube reaction, 1h 50°C), it was adopted for routine cloning. Linear DNA fragments (such as insert and plasmid) containing overlapping sequences at their end (30-60 bp) can be joined using a mix of three different enzymes. Each enzyme has a specific function: first a 5' exonuclease creates long overhangs. After annealing of two fragments with complementary sequences a polymerase fills in the single stranded gaps, and finally a ligase joins the fragments.

Benchling Exercises

You will design cloning strategies for sequences of interest using Benchling software. See Moodle for details.

Exercise 1 Cloning with primer sequences provided

The virtual step-by-step assembly follows the same procedure you do during the practical. In practice the virtual cloning is always completed BEFORE starting the wet lab cloning and is an important quality control to ensure that everything is correct in the final assembly.

- Create **primers** (sequences provided) attached to a DNA template
- Create a **primer pair**
- Create a new DNA entry with the **PCR product** obtained in laboratory 2
- Create the **recombinant plasmid** containing the Amy2 PCR product in pcDNA™6/myc-His A using BamHI and XhoI sites.
- Translate and **annotate** the recombinant fusion protein.

Provide the image of the recombinant plasmid map with relevant annotations (coding sequences and restriction sites, primers) in your lab notebook with a short figure legend. The recombinant plasmid will be used in laboratory 6 for analysis of the sequencing results.

Exercise 2

You want to clone Amy2 as a fusion protein into pcDNA6™/*myc*-His **version B** (note the vectors are supplied in three reading frames; see plasmid map below). The coding sequence should be **in frame** with the **C-terminal** peptide containing a polyhistidine metal-binding (His) tag and the *myc* epitope. Primer sequences contain respectively *Hind*III and *Not*I and will be provided (on Benchling/ Moodle). Would you be able to generate a **translational fusion**?

a) Check whether the amino acid sequence between the C-terminus of your protein of interest and the myc-tag contains a STOP codon or a frameshift (the amino acid sequence of the tag is changed). Justify your answer. NCBI reference sequence for *Amy2*: NM_001042711.2
b) Can you modify the primers to obtain the correct fusion protein?
c) What happens if you clone into version A?

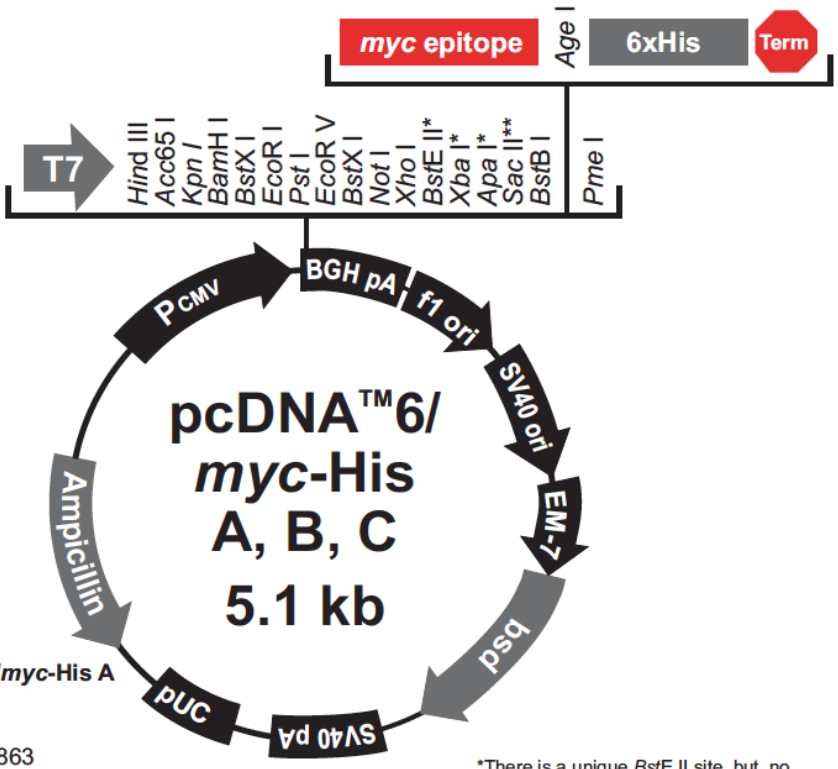
References

Gibson DG, Young L, Chuang RY, Venter JC, Hutchison CA, Smith HO. Enzymatic assembly of DNA molecules up to several hundred kilobases. Nat Methods. 2009, 6, 343-5.

Gibson DG, Glass JI, Lartigue C, Noskov VN, Chuang RY, Algire MA, Benders GA, Montague MG, Ma L, Moodie MM, *et al*. Creation of a bacterial cell controlled by a chemically synthesized genome. Science. 2010, 329, 52-6.

Plasmid maps

The figure below summarises the features of the pcDNA6/myc-His A, B, C plasmids. Details of the multiple cloning sites are shown below. Note that these plasmids are supplied in three different reading frames to allow in-frame cloning of any insert. You can download the sequences from Thermofisher.



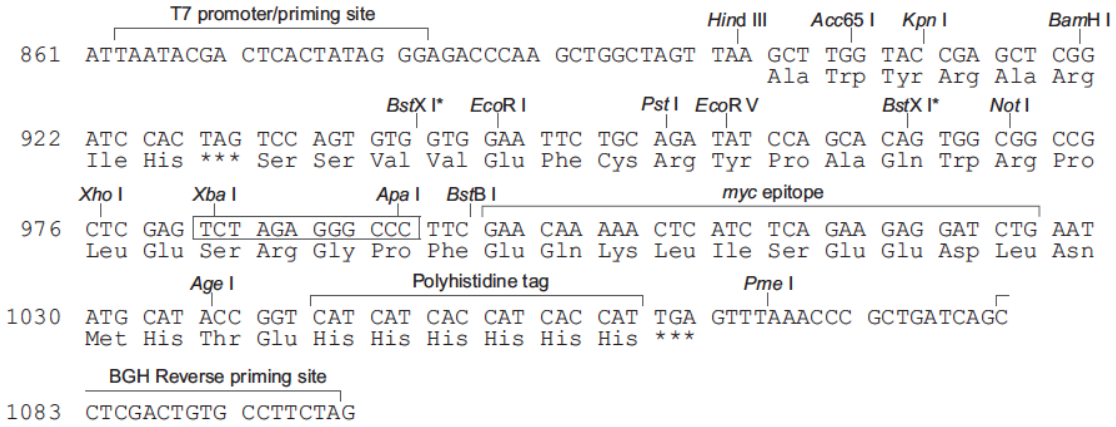
Comments for pcDNA™6/*myc*-His A
5126 nucleotides

CMV promoter: bases 209-863
T7 promoter/priming site: bases 863-882
Multiple cloning site: bases 902-999
myc epitope: bases 997-1026
Polyhistidine tag: bases 1042-1059
BGH reverse priming site: bases 1082-1099
BGH polyadenylation signal: bases 1081-1295
f1 origin: bases 1358-1771
SV40 promoter and origin: bases 1813-2121
EM-7 promoter: bases 2169-2224
Blasticidin resistance gene: bases 2249-2641
SV40 polyadenylation signal: bases 2799-2929
pUC origin: bases 3312-3985
Ampicillin resistance gene: bases 4130-4991

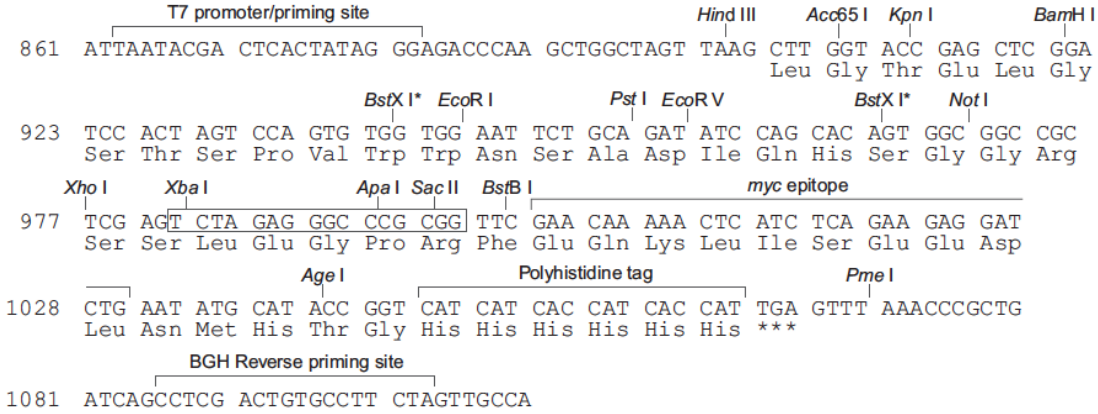
*There is a unique *Bst*E II site, but no *Xba* I or *Apa* I sites in version C.

There is a unique *Sac* II site between the *Apa* I site and the *Bst*B I site in **version B only.

Multiple cloning site pcDNA™6/myc-His Version A



Multiple cloning site pcDNA™6/myc-His Version B



Multiple cloning site pcDNA™6/myc-His Version C