

SynBio Project Tutorials

[Download the Full Tutorial as a PDF](#)

This site is a structured, modular guide for learning to design, execute, and manage synthetic biology experiments. It emphasizes practical skills, design from first principles, and a deep understanding of the molecular biology and chemistry behind each step.

The tutorials are organized into three parts:

- **Part A: Wetlab** — Step-by-step execution of a cloning experiment, including molecular biology concepts, lab protocols, documentation, and data interpretation.
- **Part B: Construction** — How to design cloning workflows, simulate outcomes, and represent cloning plans using standardized formats.
- **Part C: Project Planning** — Strategies for organizing experiments, selecting parts with informatics tools, and managing experimental workflows.

Each tutorial includes embedded quizzes, videos, and progress tracking. This platform supports students, researchers, and anyone seeking a rigorous, hands-on introduction to synthetic biology.

Part A: Wetlab

1. [pP6 Experiment Introduction](#)

Learn the goal of the pP6 library experiment, the concept of consensus promoters, and why randomized promoter libraries are used.

2. [Pipetting](#)

Learn how to use a pipette properly, interpret volume markings, and avoid common errors.

3. [Polymerase Chain Reaction](#)

Instructions for setting up the PCR using degenerate oligos to amplify the pP6 library.

4. [Gel Electrophoresis](#)

Run a gel to verify the PCR product, and interpret results.

5. [Zymo Cleanup](#)

Purify the PCR product to prepare it for assembly.

pP6 Intro: Finding Strong Promoters in *E. coli*

Goal of the Experiment

Your objective is to create and test a library of synthetic promoters to discover variants with high transcriptional strength. The goal is to create a family of distinct promoter parts that span a range of strengths, suitable for use in multi-gene constructs without sequence redundancy or recombination risk.

Getting Started in the Lab

To perform this experiment, you'll need access to a fully equipped molecular biology lab. This includes basic tools like a set of pipetman, a thermocycler, an incubator, a shaker, and standard molecular biology reagents. You'll also need the pJ12 plasmid template and two synthetic oligos ordered from IDT that introduce the promoter library variability.

Before going to the bench, download and fill out the pP6 LabSheet Workbook, which will guide your workflow:

[pP6 LabSheet Workbook](#)

To prepare:

1. Make a copy of the LabSheet (as a Google Sheet).
2. Enter your name and assigned ID on the ‘Entry’ worksheet.
3. Either print out the sheets from ‘PCR’ through ‘MiniSeq’ or bring them on a laptop, tablet, or phone.
4. Watch the demo video for each task before performing it in the lab.

These sheets contain the bench-level instructions you'll follow throughout the experiment.

Pipetting

Recombinant DNA technology is technically easy to do once you have mastered a few basic manual techniques. The most essential of these is the transfer of specified volumes of liquid accurately from one container to another. In molecular biology, we frequently work with microliter (μL) volumes, often using multi-step protocols where precision and care are essential.

Micropipette Basics

Micropipettes are precision tools used to measure and transfer small volumes of liquid, typically between 0.5 μL and 1000 μL . Each pipette is optimized for a specific volume range and should only be used within that range for accurate results.

Common pipette sizes:

- **P10:** 0.5–10 μL
- **P20:** 2–20 μL
- **P200:** 20–200 μL
- **P1000:** 100–1000 μL

Both the P10 and P20 can be used to deliver 0.5 μL , though accuracy at this volume depends more on careful technique than on the specific pipette. You'll need to visually confirm how far the liquid rides up the tip.

Each pipette uses tips that must be changed between uses to avoid contamination.

LTS vs. Universal Tips

In this lab, we use **LTS (LiteTouch System)** pipettes and tips. Unlike universal-fit pipettes, LTS tips require less force to attach and eject and help reduce repetitive strain. LTS pipettes have a cylindrical shaft with a defined stop, making tip seating more consistent. Do not attempt to use universal tips with LTS pipettes—they will not seal properly and may result in inaccurate volumes.

Polymerase Chain Reaction

In this step, you'll perform a PCR reaction to generate a pool of DNA molecules, each containing a different variant of a synthetic promoter.

This protocol uses **PrimeSTAR GXL DNA Polymerase** and the **PG4K55** thermal cycler program. You'll use degenerate primers that introduce many random bases, meaning each DNA product will contain a different promoter sequence that nobody has created before you.

What You're Doing

In this step, you're performing an EIPCR (Error-prone Inverse PCR) reaction using the plasmid **pJ12** as a template. The goal is to replace the weak promoter in pJ12 with a library of randomized sequences using specially designed degenerate primers.

This reaction produces a linear ~3.6 kb PCR product with terminal Bsal sites, which will later be closed into a circular plasmid by Golden Gate assembly.

Primers Used

The primers introduce Bsal sites and randomized sequence regions. N represents a position where any base could be present. These are generated by mixing the bases during oligo synthesis.

Forward Primer (P6LibF2)

Reverse Primer (P6LibR2)

Gel Electrophoresis

After PCR, it's important to confirm that the reaction worked by checking for the presence and size of the product. We do this by running a small portion of each PCR reaction on an **analytical agarose gel**.

Why We Run a Gel

PCR can fail in a variety of ways (no product, wrong size, smearing, etc.). Running a gel lets us visually check:

- Whether a product was formed
- Whether it's the correct size (~3583 bp for pP6)
- Whether there's nonspecific amplification

This is an **analytical gel**, meaning we are only checking—**not purifying**—the DNA.

What Is a Gel?

Agarose gel electrophoresis separates DNA fragments by size. DNA is negatively charged and moves toward the red (positive) electrode when voltage is applied.

- "**Run to red**": always load your DNA at the black (negative) electrode end.
- DNA travels through a gel matrix—**smaller pieces move faster**.
- A loading dye is added to weigh down the sample and track progress visually.

Gels are made from agarose, a gelling agent purified from seaweed. To prepare a gel, agarose is dissolved in 1× TAE buffer at 1% weight/volume by heating (typically in a microwave until boiling), then poured into a mold with combs to form wells. After it sets, we store the gels in bulk in a sealed container in the fridge. In lab, you'll cut a section from a pre-made gel with enough wells for your samples.

TAE stands for Tris-Acetate-EDTA. It is a buffer that maintains pH and ionic strength during electrophoresis. Standard 1× TAE contains 40 mM Tris, 20 mM Acetate, and 1 mM EDTA at pH 8.3.

Zymo Cleanup

After confirming your PCR reaction on a gel, the next step is to purify the DNA. This is essential for most cloning workflows, especially **Golden Gate** and **restriction enzyme-based** cloning. Although it is **optional for Gibson**, cleanup helps remove components from the PCR that may interfere with downstream steps.

Why Cleanup Is Necessary

PCR reactions contain:

- Unreacted dNTPs
- Active polymerase
- Buffers and salts

These can interfere with cloning enzymes, particularly enzymes like Bsal that generate sticky ends. Polymerase can fill in sticky overhangs, preventing proper ligation. Cleanup eliminates these components.

How It Works

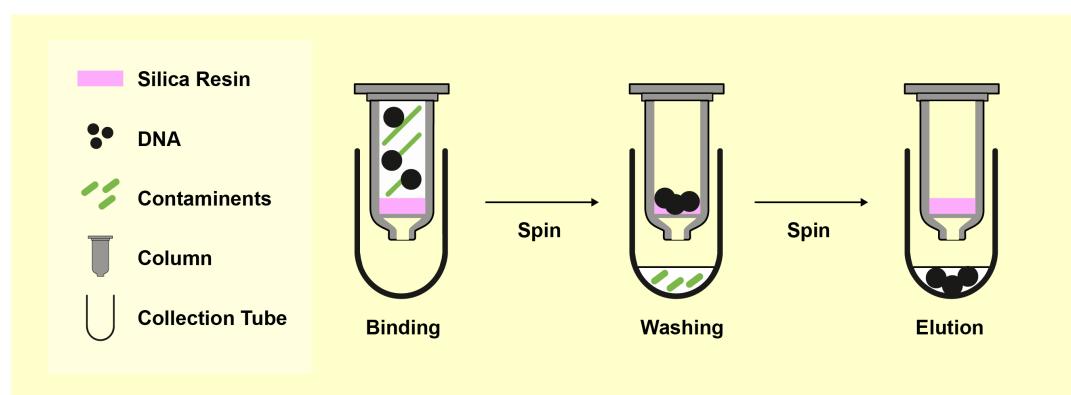


Figure: Stepwise diagram of silica column cleanup showing binding, washing, and elution. DNA binds to the pink resin in the column, contaminants are washed away, and purified DNA is eluted into a clean tube.

Assembly

In molecular cloning, “assembly” refers to joining two or more DNA molecules together into one. Two of the most common methods are **Golden Gate Assembly** and **Gibson Assembly**—together they cover nearly all modern cloning needs.

Assembly Methods Overview

Golden Gate Assembly uses a **type IIS restriction enzyme** (e.g., *BsaI*, *BsmBI*, *BbsI*) to cut DNA at defined locations and generate sticky ends. These are then joined by **T4 DNA Ligase**.

Gibson Assembly uses an exonuclease to chew back DNA ends, allowing single-stranded overlaps to anneal. A polymerase fills in gaps and a ligase seals the nicks.

Both methods:

- Are performed in a single tube (one-pot)
 - Run in a thermocycler
 - Require specific enzyme buffers and careful setup
-

What We're Doing Here

In the **pP6 experiment**, we only have one DNA fragment: the linear PCR product from earlier. Our goal is to **re-circularize it** by joining its ends together using **Golden Gate**.

This works because the PCR primers added ***BsaI* recognition sites**, which generate **complementary sticky ends** at each end of the linear product.

Transformation

After assembly, the circular plasmid has been generated, but it is mixed with other DNAs—some incomplete or mutant. By passing the material through cells, we isolate a single one of these sequences and amplify it a billion-fold.

Transformation is the process in which a bacterium takes up DNA from the environment. While some bacteria (like *B. subtilis*) do this naturally, *E. coli* requires preparation to become “competent.”

There are two common methods:

Electroporation

Electroporation involves preparing cells in salt-free, ice-cold water or 10% glycerol and shocking them with a brief electric pulse in a special cuvette. This creates pores in the membrane, allowing DNA to enter. It is:

- ~100× more efficient than heat shock
 - More expensive (cuvettes are single-use)
 - Sensitive to salt contamination
 - Ideal for library transformations needing high colony counts
-

Heat Shock (used in pP6)

We use the **TSS method** for heat shock, where cells are suspended in a PEG-salt mix. You add DNA and a small amount of KCM buffer, chill, then heat shock at 42°C for 90 seconds.

Full TSS protocol

TSS cells can be frozen in aliquots and remain competent for years. TSS cells are easier to work with during transformation and are well-suited for routine cloning workflows.

In dilute conditions, as in pP6, most cells take up just **one plasmid**, resulting in unique colonies.

Colony Picking

After plating your transformation and incubating overnight, if everything goes well you will see colonies that look something like this:

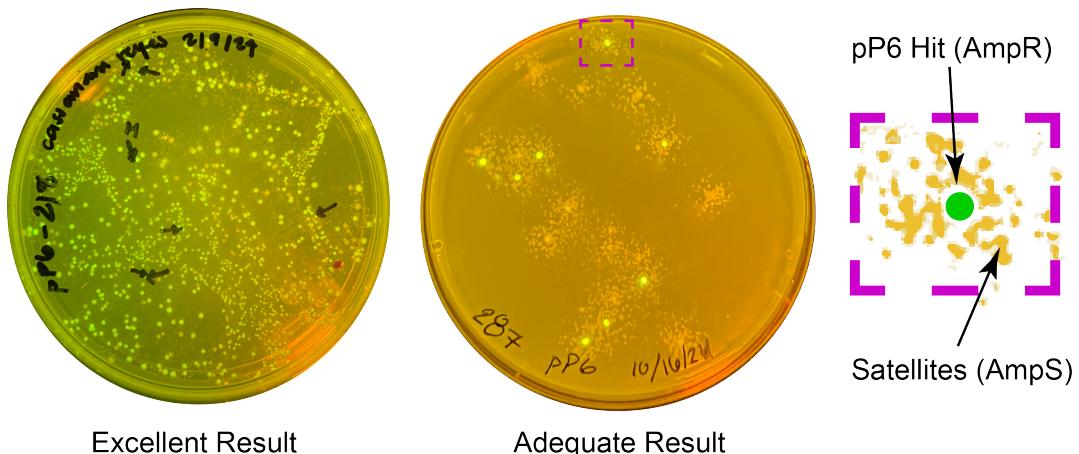


Figure: Results from two students' pP6 transformations under blue light illumination to excite amilGFP. Left: an ideal plate with many green colonies. Right: a minimally usable plate. The zoomed region shows a green colony (a true pP6 transformant) surrounded by smaller white "satellite colonies"—non-transformed cells growing in an antibiotic-depleted zone. Only the green colony contains the plasmid and is ampicillin-resistant.

What Are You Looking For?

In the pP6 experiment, you're trying to identify the **strongest promoter variants**. These drive expression of amilGFP, which fluoresces green.

- **Bright green colonies** indicate successful assembly of a strong promoter driving GFP.
- Colonies with no fluorescence are likely the template plasmid (pJ12) or inactive variants.

Some bright colonies are so strong that you can see yellowish color even without blue light. These are your best candidates.

Miniprep: Plasmid DNA Purification

Once you've picked and grown up an isolated colony, you now have a saturated culture containing billions of identical cells. Each cell carries multiple copies of the plasmid, and a miniprep allows you to extract and purify that plasmid DNA.

Why Miniprep?

1. **Storage** — Purified plasmid can be stored at -20°C indefinitely.
 2. **DNA as a Building Block** — You often need clean DNA to build or clone further constructs.
 3. **Sequencing** — Verifying the plasmid sequence requires pure template.
 4. **Transformation into Other Cells** — Requires isolated DNA.
 5. **General Use** — Quantification, restriction mapping, and other analytical methods.
-

Cycle Sequencing

After you've picked colonies and completed your minipreps, you've finished the fabrication phase. The next question: **what did you actually make?** This tutorial will guide you through how to determine if your plasmid contains the correct sequence.

What Is Cycle Sequencing?

Cycle sequencing, also often called Sanger sequencing, uses DNA polymerase and chain-terminating nucleotides (ddNTPs) to create truncated DNA fragments, each ending in a labeled base. These fragments are then separated by capillary electrophoresis to reveal the DNA sequence.

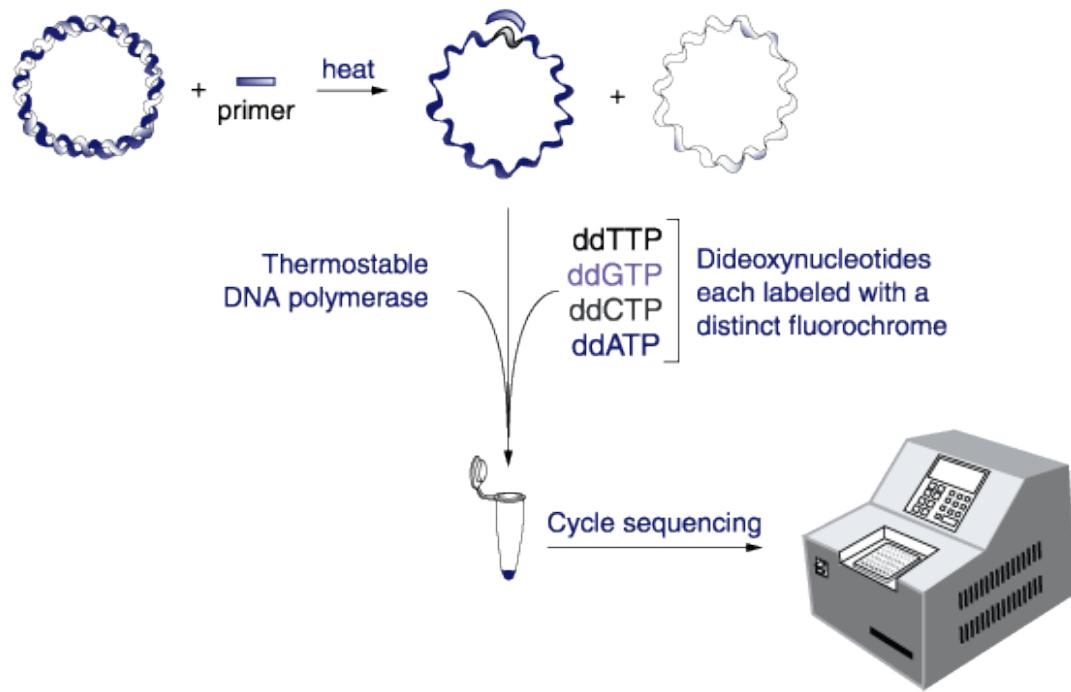


Figure: Cycle sequencing overview. Fluorescently-labeled ddNTPs stop extension at every base type. The resulting DNA fragments are separated and analyzed.

Why Not Just Sequence the Whole Plasmid?

Full-plasmid sequencing services (e.g., Plasmidsaurus) cost ~\$15/sample and return the entire sequence. This is useful in some cases, but often overkill. In the pP6

BestP: Measuring Fluorescence

Overview

The **BestP** experiment marks a shift from building DNA to **measuring its activity**. The goal is to assign a quantitative value—**Relative Promoter Units (RPU)**—to each promoter you identified in pP6. This allows us to compare their strengths under consistent conditions.

Why RPU?

Different experiments can yield different fluorescence values due to instrument settings, media conditions, or cell growth. To standardize promoter activity, we compare each sample's fluorescence to a **reference promoter: J23101**, a commonly used medium-strength promoter from the Anderson library. Its activity is defined as **1 RPU**.

Reference Plasmids

We use three plasmids from the Anderson promoter library:

- **pJ12** – Contains **J23112**, a very weak promoter
- **pJ01** – Contains **J23101**, the **standard reference**
- **pJ19** – Contains **J23119**, a very strong promoter

Each has the same vector backbone and reporter gene (amilGFP) as your pP6 clones, ensuring fair comparison.

Experimental Workflow

You'll choose **4 pP6 clones** to characterize—either your own from sequencing or others from the TPcon6B box. You'll also measure the reference plasmids above. Here's the full procedure:

DNA Basics and Molecular Biology Refresher

DNA Structure and the Reverse Complement Function

DNA is a double-stranded molecule arranged in an **antiparallel** fashion, meaning one strand runs in the 5' to 3' direction while its complementary strand runs in the 3' to 5' direction. The two strands are held together by **base pairing**: Adenine (A) pairs with Thymine (T), and Cytosine (C) pairs with Guanine (G). These fundamental properties allow for key sequence operations like reversing and complementing strands. To understand this structure visually, watch the video below:



Reverse

To reverse a DNA sequence, simply write it backward. For example:

- **Original:** 5' - ATGCAG -3'
- **Reversed:** 5' - GACGTA -3'

Complement

To complement a DNA sequence, replace each base with its pairing base:

- **A (Adenine)** ↔ **T (Thymine)**
- **C (Cytosine)** ↔ **G (Guanine)**

Example:

- **Original:** 5' - ATGCAG -3'
- **Complement:** 5' - TACGTC -3'

Reverse Complement

The **reverse complement** is simply applying the complement rule **after reversing** the sequence:

1. Reverse the sequence.
2. Replace each base with its complementary base.

Sequence Tools

Introduction

Sequence editing tools are essential for molecular biology and bioinformatics, allowing researchers to manipulate and analyze DNA sequences effectively. These tools range from graphical interfaces designed for ease of use to powerful command-line utilities and programming libraries.

Tools Overview

Tool Name	Pros	Cons
ApE (A Plasmid Editor)	<ul style="list-style-type: none">- Course demonstrations will be done in ApE.- Simple interface- Efficient auto-annotation- Supports multiple file formats	<ul style="list-style-type: none">- Lacks cloud collaboration features- Limited automation capabilities.
Benchling	<ul style="list-style-type: none">- Real-time collaboration- Extensive documentation- Integrates well with other laboratory tools.	<ul style="list-style-type: none">- Requires an internet connection- Subscription required for advanced features.
SnapGene	<ul style="list-style-type: none">- User-friendly- Excellent visualization tools- Supports automatic primer design and cloning workflows.	<ul style="list-style-type: none">- Paid software with limited free version capabilities.
UGENE	<ul style="list-style-type: none">- Free- Supports a wide range of sequence analysis tools- Allows scripting.	<ul style="list-style-type: none">- Interface can be complex for beginners.
Biopython	<ul style="list-style-type: none">- Highly customizable- Integrates well with other computational tools.	<ul style="list-style-type: none">- Requires programming knowledge.

Manual Product Prediction

In this module, you'll learn how to predict the results of molecular biology operations by reasoning through the mechanisms. This manual prediction process is foundational for verifying whether a cloning strategy will work as intended. While there are many tools and algorithms to help you design a cloning plan, here we are learning to do it based on first principles. It's a skill you'll often use in the lab when results are unclear and you need to troubleshoot. Going through this process also deepens your understanding of how molecular biology works and how different biochemical elements interact.

Before You Try This...

Want a refresher on how PCR works at a mechanistic level? Watch this short, animated video:

 [PCR Animation – The Polymerase Chain Reaction](#)

Guided Walkthrough: Predicting PCR Products

We'll begin with a simple scenario using a structured table format called `cf_shorthand`. You'll encounter this notation in later tutorials, but here we'll treat it as a straightforward summary of oligos and templates.

PCR Example

operation	primer1	primer2	template	product
PCR	exFor	exRev	pTemp1	pcrpdt
oligo	exFor	CAGCGGATCGGATCGCGAC		
oligo	exRev	CGGTTGTGGGGCGGAAC		
plasmid	pTemp1			
	CTGGTGACCCAGCGGATCGGATCGCGACCCAAAGCGCCTGGTTCCGCCGCACAACCGCGA			

We have two primers (`exFor` and `exRev`) and a circular template plasmid (`pTemp1`). The task is to predict the PCR product.

1. Find `exFor` exactly in the template.
2. Find the reverse complement of `exRev` and match that in the template.

Simulation Tools

Several tools can automate molecular cloning workflows. ApE and Benchling offer graphical interfaces for simulating steps like PCR, Golden Gate, Gibson, and digestion reactions. Alternatively, C6-Tools provides a scripting-based approach using Construction File (CF) shorthand, and is available both as a web tool and as a Google Sheets plugin:

 [Open C6-Tools in Google Sheets](#)

The Google Sheets version includes a library of DNA design functions that you can access directly from spreadsheet cells. These include methods for `PCR`, `Digest`, `Ligate`, `GoldenGate`, and `Gibson`, as well as functions for parsing and simulating CF scripts. Visit the linked page and follow the instructions to get started.

Understanding CF Syntax

The Construction File (CF) format is a streamlined way to describe molecular cloning procedures using a structured table-like syntax. Each line specifies either:

- An **operation** to perform (e.g., `PCR`)
- A **sequence element** involved in the operation (e.g., an `oligo` or `plasmid`)

Each line begins with a **keyword** (like `PCR`, `oligo`, or `plasmid`) followed by a set of fields separated by either tabs or multiple spaces. These fields define the names and sequences needed for simulation.

PCR Line Breakdown

For example, the line:

PCR	exFor2	exRev2	pTemp1	pcrpdt2
-----	--------	--------	--------	---------

...defines a **PCR** reaction using:

- a forward primer called `exFor2`
- a reverse primer called `exRev2`
- a circular DNA template called `pTemp1`

Basic Cloning: Overexpression of Human Insulin

Overview

In this tutorial, you'll learn to design oligonucleotides and plan a traditional restriction/ligation cloning experiment. We'll walk through cloning the human insulin cDNA into a pET expression vector for high-level expression in *E. coli*. This process mimics a landmark achievement in biotechnology: the recombinant production of human insulin, which today powers a multi-billion-dollar global industry.

Insulin was the first therapeutic protein produced using recombinant DNA technology, replacing animal-derived sources and transforming diabetes care. In this example, you'll see how tools like PCR, restriction enzymes, and plasmid vectors can be used to enable scalable, microbial production of life-saving medicines.

Context: Recombinant Insulin Production

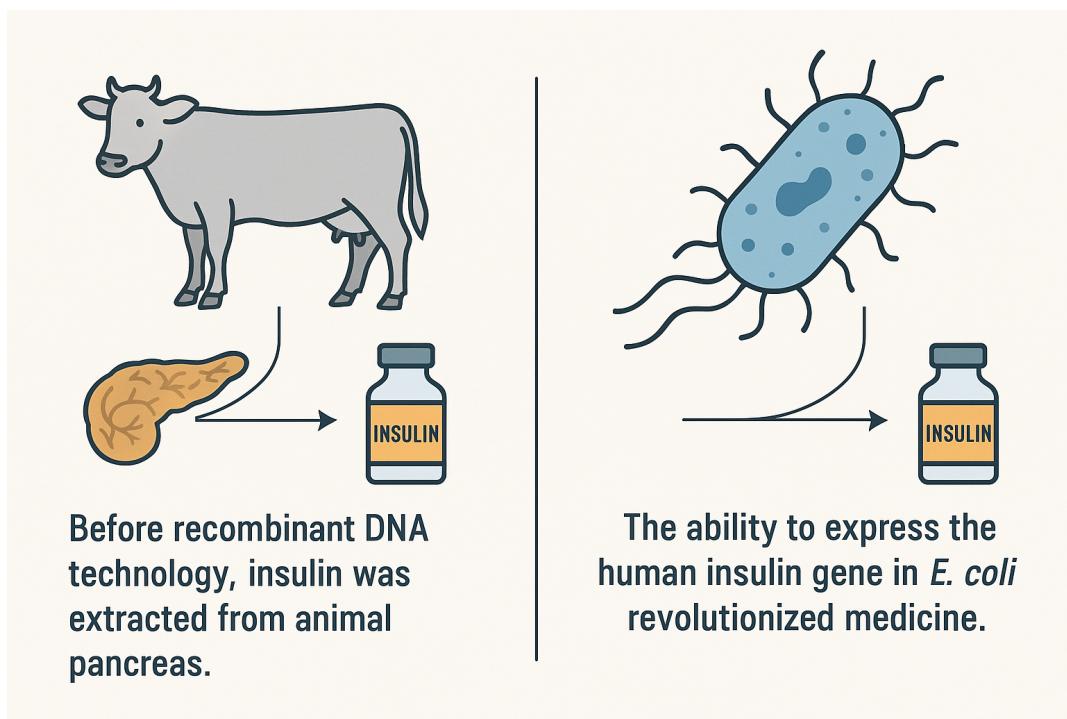


Figure: Before 1982, insulin was extracted from thousands of pig or cow pancreases, yielding limited supply and often triggering immune reactions. Recombinant DNA

Gibson Assembly

Why Gibson Assembly?

In the previous tutorial, you used restriction enzymes to insert the INS gene into a pET vector. This is a reliable method, but it has a key limitation: it depends on restriction sites being in the right places.

In synthetic biology, you often need to assemble multiple parts with precise control over sequence. This is where **Gibson Assembly** shines. It allows you to join DNA fragments without restriction sites—using only sequence overlaps.

🔬 How Gibson Assembly Works

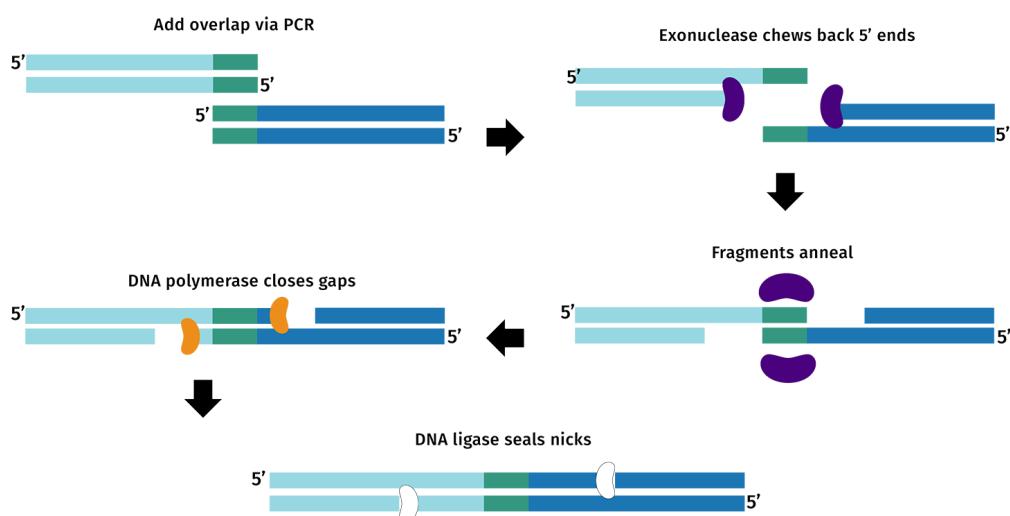


Figure: Gibson Assembly joins DNA fragments with overlapping ends in a seamless, scarless way. An exonuclease chews back 5' ends, exposing complementary overhangs. These anneal, and a polymerase fills in gaps. A ligase seals the nicks, yielding a continuous double-stranded product.

🔬 Diagram credit: [SnapGene Gibson Guide](#)

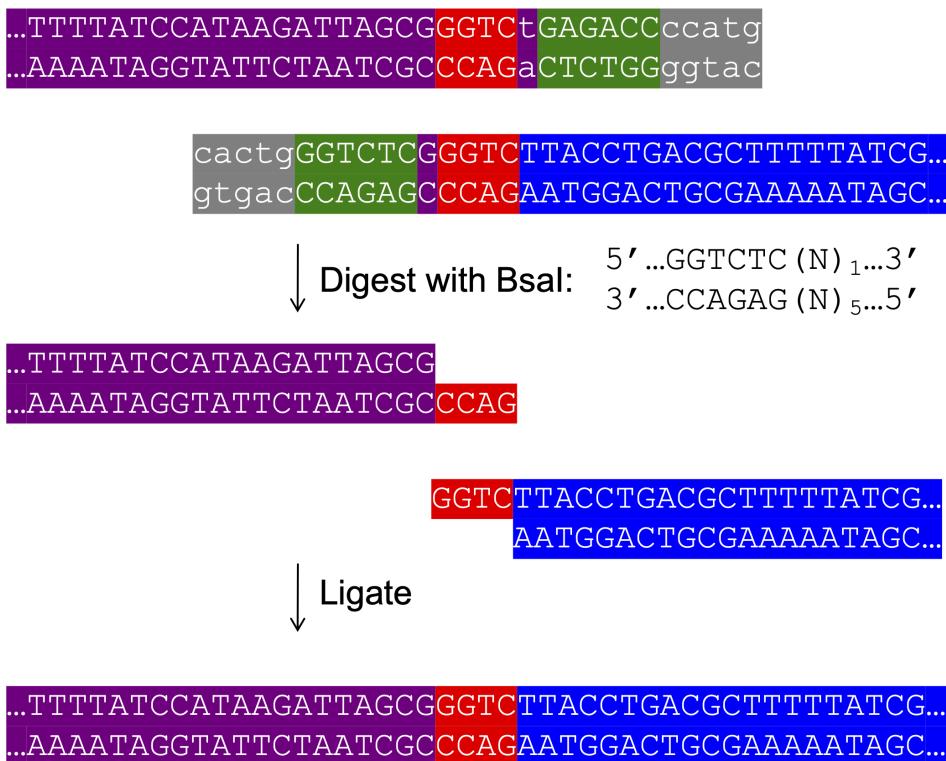
The enzyme mix includes three key components:

- 1. 5' Exonuclease:** Creates single-stranded 3' overhangs by chewing back the 5' ends.
- 2. DNA Polymerase:** Fills in gaps after annealing.

Golden Gate Assembly

In previous tutorials, we built the **pET-INS** plasmid using both traditional restriction enzyme cloning and Gibson Assembly. In this tutorial, you'll use **Golden Gate Assembly** to build the same construct—while learning how this method enables precise, scar-controlled, multi-part DNA construction.

What is Golden Gate Assembly?



Golden Gate Assembly is a method for joining DNA fragments using **Type II** restriction enzymes like **Bsal**, **BsmBI**, **BbsI**, and **SapI**. These enzymes cut a fixed number of bases away from their recognition sites, which allows the creation of custom 4 bp overhangs that control exactly where and how parts join together.

Golden Gate is powerful because it allows:

- Custom-designed sticky ends for precise, seamless (or intentionally scarred) ligation
- Single-pot digestion and ligation, increasing efficiency

Site-Directed Mutagenesis

Site-directed mutagenesis refers to any cloning technique where you introduce a specific, localized change into a DNA sequence. These changes may include:

- Substituting a base or codon
- Inserting or deleting a small sequence (typically <30 bp)
- Replacing a defined region
- Creating a variant library at a defined site

In all cases, you're targeting a specific location within a plasmid and rewriting a short region while leaving the rest unchanged.

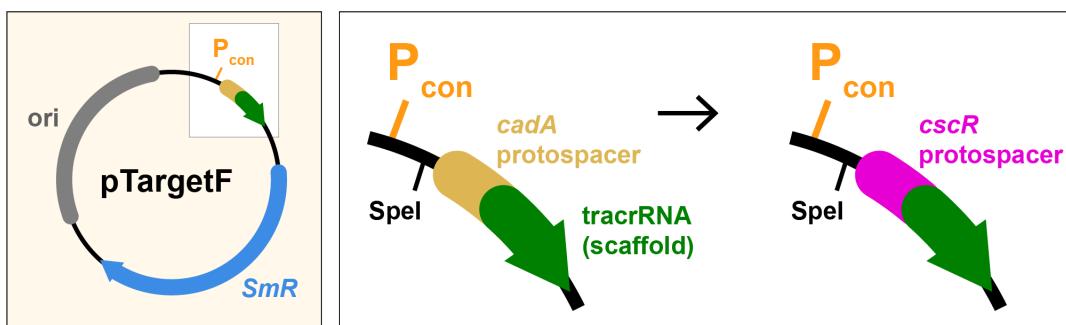


Figure: Site-directed mutagenesis of the pTargetF plasmid to retarget the gRNA. The original cadA-targeting protospacer (orange) is replaced with a new cscR-targeting protospacer (pink), using a Spel-flanked region upstream of the gRNA scaffold (green tracrRNA).

Overview of Mutagenesis Strategies

Most mutagenesis methods resemble the PCR-based cloning techniques you've already seen. The general pattern is:

1. PCR the entire plasmid using primers that encode your intended mutations.
2. Re-close the product using ligation or Gibson-like mechanisms.
3. Transform and screen for correct clones.

Conceptually, it's like the limiting case of assembly: just one fragment, but edited.

Sequencing Confirmation

After you've completed your cloning, your next step is to verify the integrity of your plasmid. This tutorial walks through different sequencing strategies and how to choose and apply the right one for your purpose.

Choosing the Right Strategy

There are three main sequencing strategies available, each suited to different goals and budgets:

1. **Cycle sequencing** (like you used in the pP6 experiment) gives a ~1 kb window of high-quality sequence starting about 50 bp downstream of a primer. It's ideal for checking a specific region of a plasmid—like an insert or promoter.
2. **Full-plasmid sequencing** from providers like Plasmidsaurus uses long-read nanopore technology to return the complete sequence of your plasmid. This is more expensive (~\$15/sample), but you get the entire plasmid and can catch unexpected rearrangements or background DNA.
3. **NGS** (Next-Generation Sequencing) is a high-throughput approach designed for analyzing complex populations. You PCR-amplify your targets with adapters and submit the pool for deep sequencing. This is powerful, but expensive, and more suitable for library screens than for single plasmid validation.

Method	Cost/ sample	Output	Best For
Cycle sequencing	~\$3.50	~1 kb from a primer	Targeted region confirmation
Full plasmid (Plasmidsaurus)	~\$15	Entire plasmid sequence	Whole-plasmid verification, structural issues
NGS (deep sequencing)	\$750+	Millions of short reads	Large libraries, pooled clone analysis

Project Setup and Nomenclature

This guide outlines conventions for setting up and organizing a synthetic biology project. These conventions are intended to support clarity, reproducibility, and collaboration. The approach described here emphasizes a consistent, pre-planned structure for working with DNA and experiment metadata. While documentation systems vary across labs, any good system should have a clear scope, support structured naming, and promote project-wide consistency over time.

Defining a Project

A *project* is an independent, long-term research effort with a clear scientific or engineering goal. Each project should have a unique, single-token name (e.g., `lycopene`) that will be used consistently across documentation, file structures, plasmid names, and experiment records. Avoid informal or overlapping nicknames to ensure records remain searchable and unambiguous over time.

Projects are generally defined by scope: they may focus on a distinct application, scientific question, or system. They are often aligned with specific funding sources, collaborations, or distinct groups of people. For example:

Project: `lycopene`

Goal: Maximize lycopene production in *E. coli* by tuning the metabolic pathway.

This is distinct from other efforts that may be running in parallel, such as:

Project: `stickbug`

Goal: Build a surface-display system in *Bacillus subtilis* to test adhesion strength on plastic, glass, and food packaging materials.

Although these projects may share some techniques or tools, they target different applications and will generate separate materials and documentation.

Defining an Experiment

An *experiment* is a specific, bounded effort to test or implement one part of a larger project. Experiments are sequential and named using the project name followed by a number — for example, `lycopene33` is the 33rd experiment within the `lycopene` project.

Experimental Design Principles

Construction Files, Genbank sequences, and LabSheets are a general-purpose toolkit for documenting synthetic biology experiments. They do not tell you *what* to build—only *how to record it*. So the fundamental question at the start of every project is: **What should you build and test?**

Types of Experiments

Most synthetic biology experiments fall into two broad categories:

1. Prototyping

In the prototyping stage, you are exploring new designs that haven't been tested in your system. You might have:

- A gene from a database or literature
- A new combination of parts
- A hypothesis about how something will work in a cell

But no direct experience with the function in your hands.

Questions to Ask:

- What proteins/RNAs need to be expressed?
- What regulatory elements do I need?
- What assay tells me if it works?
- What controls are required for interpreting results?
- How many variants can I build/test given assay throughput?

2. Optimization

Once a prototype is functional, optimization improves performance: yield, growth, stability, etc.

Sequence Retrieval and Analysis

In this tutorial, you will learn how to retrieve the sequence of a gene from another organism and prepare it for cloning. Our goal is to perform an **ortholog scan** of the *dxs* gene as part of the **Lycopene33** optimization experiment. We will search for alternative versions of *dxs* from different species, extract their RBS.CDS sequences, and prepare those sequences to build RC parts that match the TP.RC part architecture.

The process includes:

- Identifying orthologs via BLAST
 - Extracting upstream and downstream flanking sequences
 - Verifying gene boundaries
 - Preparing the sequence for cloning
-

Step-by-Step Workflow

1. Identify Your Starting Sequence

Begin with a known working version of the *dxs* gene from *E. coli* MG1655. You'll need the CDS (coding DNA sequence), and ideally the 200 bp upstream (pre) and downstream (post) flanking sequences. These will help identify the ribosome binding site and define PCR primers if you're cloning from genomic DNA.

2. Translate the CDS to Protein

Using a sequence editor such as ApE, paste in the *dxs* CDS and use the "Translate" function to convert it to its corresponding amino acid sequence.

3. Run a BLASTP Search

BLASTP is a search tool that compares your protein sequence to a vast database of known protein sequences maintained by NCBI. Nearly all public genomic and protein data are included in this database, and BLASTP has indexed these sequences in a way that allows it to quickly find similar entries.

Choosing a Fabrication Strategy

Once you've finalized your RC part design (e.g., RBS.CDS), it's time to decide how you'll turn that design into actual DNA. This section walks you through the decision process for choosing a fabrication strategy and explains how to encode that strategy in a Construction File (CF).

Step 1: Organize Your Source Sequences

Before starting oligo design, ensure your GenBank files are saved and annotated. These should live in the `Maps/` folder of your GitHub repo.

For the Lycopene33 example, we'll focus on an RC part derived from the dxs ortholog of:

Trueperella pyogenes strain TP1

GenBank: CP033902.1

Coordinates: 656655–658523

[NCBI Link](#)

You've already discussed this sequence, selected a suitable RBS, and formatted it as a standardized RC part following TP.RC conventions. A visual reminder of the completed part is shown below:

TpDXS RC Part

The corresponding DNA sequence files is included in the [example Maps](#) folder for reference.

- **CP033902.seq:** The raw sequence of the region of the *Trueperella* genome containing the dxs CDS and flanking sequences
 - **gTpDXS.seq:** A GBlock encoding an RC (rbs.cds) part encoding the *Trueperella* sequence with restriction sites removed
 - **pTpDXS.seq:** A clonal gene synthesis plasmid encoding the restriction site-free part
-

Simulation Tools: CFS & C6

This page explains how to simulate DNA construction workflows using two tools: - **CFS (Construction File Simulator)** — a Java-based simulation engine. - **C6 Tools** — a spreadsheet-based interface using Google Sheets.

CFS (Construction File Simulator)

CFS is a Java tool that simulates a Construction File (CF) to verify correctness of molecular biology protocols like PCR, digestion, ligation, assembly, and transformation. It detects common design issues before you build.

Getting Started

1. **Install Java (JRE 8+).**
2. **Download the `.jar` file** from the [Releases section](#).
3. **Extract `dist.zip`**, if applicable.

Usage Options

- **Double-click** `ConstructionFileSimulator.jar` to launch the GUI.
- **Drag-and-drop** an experiment folder onto the GUI.
- **Command line (GUI launch):**

```
java -jar ConstructionFileSimulator.jar
```

- **Command line (direct simulation):**

```
java -jar ConstructionFileSimulator.jar /path/to/experiment/folder
```

File Requirements

A valid experiment folder must include:

- A text-based Construction File
- GenBank files for sequence definitions
- A TSV file with part definitions (if needed)

Inventory and LabSheets

At this stage, you've designed your oligos, written construction files, and confirmed the molecular biology is theoretically sound. With the experimental design finalized, it's time to order materials and plan lab work. This includes selecting protocols, identifying required samples, and specifying where to find them.

You will generate two key documents:

- **Inventory:** Tracks all DNA samples and reagents used or created in the project.
- **LabSheet:** A lab-ready checklist detailing the experimental workflow.

Sample Identity and Granularity

While construct names like `pLYC2` refer to a designed DNA sequence, the inventory tracks physical samples—things in tubes. If you pick multiple clones, perform retransformation, or prepare fresh minipreps, it's essential to use names that distinguish each version.

A few naming examples: - `pLYC2` — refers to the theoretical DNA design. - `pLYC2 clone A` — a specific isolate from a transformation plate. - `pLYC2 clone A 2°` — a separate (secondary) prep or colony from the same clone lineage.

In simple cases (e.g., only one clone, one miniprep), just using the construct name may be fine. But as soon as there are multiple samples that could plausibly differ, your naming must be specific enough to distinguish them. Inventory records must include all details needed to unambiguously identify each tube.

Similarly, when documenting results (e.g., sequencing, assay data), always indicate which specific sample was used—not just the construct name. The more complexity in your project, the more critical this becomes.

Examples of the completed ones for lycopene33 are available:

[Inventory LabSheets](#)

To illustrate, consider the construction plans for the lycopene33 project:

Ortholog Scan Project — *ispA*

This capstone exercise guides you through designing and documenting a complete cloning project centered on *ispA*, a key enzyme in the isoprenoid biosynthesis pathway. You'll build an ortholog expression panel in the context of plasmid pLYC72. This gene presents more complexity than previous examples—biochemical differences among orthologs can influence the outcome, making your selection and rationale especially meaningful.

Project Overview

Your goal is to follow the full cloning pipeline we've developed throughout the tutorial series. You'll prepare a complete project folder for an experiment called `lycopene34`, which includes design files, construction plans, lab documents, and a short writeup.

Although this exercise is built around *ispA*, the tutorial doubles as a model for how to present any cloning-based project. You're encouraged to customize or extend it depending on your interests, as long as the documentation standards are met.

Key Elements

Your `lycopene34` project folder should include:

- Construction File(s) covering all planned builds
- Inventory, Oligos, and LabSheets
- Annotated GenBank files
- A `Docs/` folder with a written summary (as a `README.md`, Word doc, or PDF)

The summary should explain:

- What the project is about
- The biology and context of the *ispA* gene
- Why you chose your orthologs (biochemical rationale, functional hypotheses, etc.)
- Any literature you consulted

Use this document to guide the reader through the logic and goals of your experiment.

GenBank/APE Feature Extractor

This tool allows you to drag and drop `.seq`, `.gb`, `.gcc`, `.str`, `.genbank`, or `.ape` files and extract features (name, color, sequence) into the ApE-compatible TSV format.

Feature Extractor

Upload a File

Drop your file here or click to upload

Add a Feature Manually

Name:

Sequence:

Type:

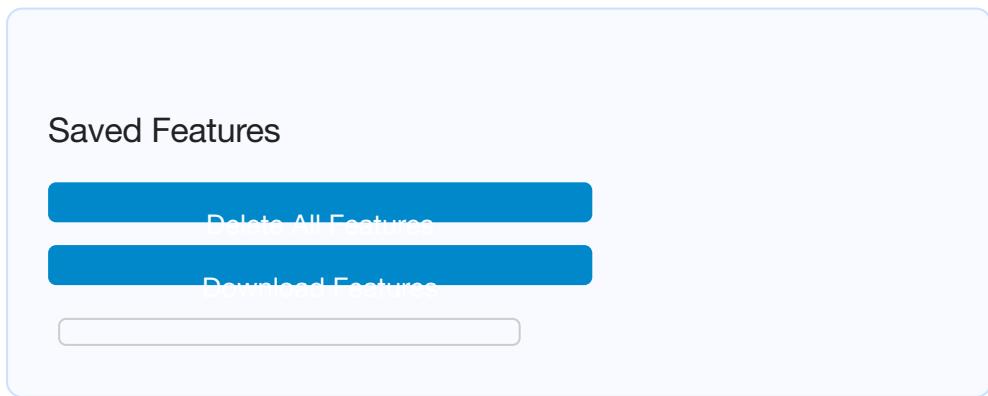
 CDS

Color:

 blue

Feature Manager

This tool manages your locally stored features extracted from GenBank/APE files or added manually. All features are stored in your browser's localStorage.



\

SuperSimulator

This tool lets you simulate a molecular biology construction file (CF), view the result, and auto-annotate the final product.

Step 1: Paste Your Construction File

[Simulate CF](#)

Step 2: Auto-Annotate Output

[Annotate and Simulate](#)

About the SynBio Project Tutorials

This site provides a modernized training resource for synthetic biology and DNA cloning workflows. It replaces and extends material previously hosted at [OpenWetWare](#), which remains available for reference.

These tutorials teach the practical and conceptual foundations of synthetic biology, with a focus on modular DNA assembly and project-based learning. The material is modular but intended to be followed sequentially.

Who This Is For

This tutorial is for anyone who wants to learn how to design and carry out DNA cloning experiments — the technical core of synthetic biology. It focuses on the foundational concepts and practical skills needed to plan and execute cloning workflows, without covering genome editing methods such as CRISPR.

The content is:

- Aimed at students, including those in BioE 140L at UC Berkeley and participants in iGEM@Berkeley and DeCal programs
- Helpful for post-baccalaureate researchers, educators, and independent learners looking for a practical, modular guide
- Designed for those with prior coursework in organic chemistry and biochemistry, but no hands-on cloning experience

The site is freely accessible, fully open-source, and suitable for both classroom use and self-paced learning.

Structure and Use

The content is organized into three parts:

- **Part A: Wetlab** — hands-on PCR, gels, cleanup, assembly, and transformation using the pP6 promoter library
- **Part B: Cloning Techniques** — design and simulation of primers and cloning strategies using tools like Benchling and ApE