MOLECULAR CLONING

The technique described in this protocol is used to perform directional cloning of a gene of interest (GOI) into an expression vector. It utilizes conventional cloning techniques, requiring type II restriction enzymes for generating sticky ends, ligation of sticky ends into the desired expression vector and transformation for cloning. This technique can either be performed on a synthetic gene block or on a PCR amplicon originating from the organism's genomic DNA (gDNA). This protocol is divided into 3 sections, where two of these (A1 and A2) follow similar steps but differ based on the origin of the GOI (i.e. synthetic gene block vs PCR amplicon).

A1: designing a synthetic gene block of your GOI

As an example, the coding sequence for the rhizomide (Rzm) thioesterase peptide sequence was used and shown in Courrier New font below each relevant step. It was cloned into a pET-28a(+) expression vector for an N-terminal 6xHis tag.

1. Copy the desired DNA coding sequence from the NCBI database (or other suitable database) and paste into a word document. If the peptide sequence is taken, reverse translate the sequence into its corresponding coding DNA sequence using EMBOSS backtranseq tool. This will generate a codon optimized sequence for *E. coli*.

```
>WP_013428324.1:7387-end non-ribosomal peptide synthetase [Mycetohabitans
rhizoxinica]
LARHLVKRDGIQENSFSVLLPLKPTGSRPALFCIHPGFGLSWSYMGLSSYLHPDQPLYGLQARGLDGTSP
LASTLDEMVSDYLEHMRRVQPKGPYCLLGWSFGGNVAHSMAVRLEHQGEKVALLALLDSTPLVGSNAFAE
KENQLDQAMVRDVFAYHYSNELVSAMDEQLLENTEEIAKNNLCIVRDYSPSQYGGNALLFRATVAEAGCE
TLVSPDAWNPYVFGGIEVHDVHCSHGEMLKPEPTAIIGSTLACKLDKWESQQAQKVNTEELK-

>Reverse translated using EMBOSS backtranseq tool (codon optimized)
CTGGCGCGCCATCTGGTGAAACGCGATGGCATTCAGGAAAACAGCTTTAGCGTGCTGCTG
CCGCTGAAACCGACCGGCAGCCGCCCGGCGCTGTTTTGCATTCATCCGGGCTTTGGCCTG
AGCTGGAGCTATATGGGCCTGAGCAGCTATCTGCATCCGGATCAGCCGCTGTATGGCCTG
CAGGCGCGCGGCCTGGATGGCACCAGCCCGCTGGCGAGCACCCTGGATGAAATGGTGAGC
GATTATCTGGAACACATGCGCCGCGTGCAGCCGAAAGGCCCGTATTGCCTGCTGGGCTGG
AGCTTTGGCGGCAACGTGGCGCATAGCATGGCGGTGCGCCTGGAACATCAGGGCGAAAAA
GTGGCGCTGCTGGCGCTGCTGGATAGCACCCCGCTGGTGGGCAGCAACGCGTTTGCGGAA
AAAGAAAACCAGCTGGATCAGGCGATGGTGCGCGATGTGTTTGCGTATCATTATAGCAAC
GAACTGGTGAGCGCGATGGATGAACAGCTGCTGGAAAACACCGAAGAAATTGCGAAAAAC
AACCTGTGCATTGTGCGCGATTATAGCCCGAGCCAGTATGGCGGCAACGCGCTGCTGTTT
CGCGCGACCGTGGCGGAAGCGGGCTGCGAAACCCTGGTGAGCCCGGATGCGTGGAACCCG
TATGTGTTTGGCGGCATTGAAGTGCATGATGTGCATTGCAGCCATGGCGAAATGCTGAAA
CCGGAACCGACCGCGATTATTGGCAGCACCCTGGCGTGCAAACTGGATAAATGGGAAAGC
CAGCAGGCGCAGAAAGTGAACACCGAAGAACTGAAATAA
```

2. Add an NdeI site at the 5' end and an EcoRI site at the 3' end of the GOI. This will allow for cloning into pET-28a(+) for an N-terminal 6xHis tag for protein purification. Add a STOP codon (TAA) before the EcoRI site if there is not one already in the GOI. If a C-terminal 6xHis tag is desired, clone into

pET-21c(+) using the same restriction sites, except remove the STOP codon at the 3' end of your GOI.

NOTE 1: Given that restriction enzymes might need to cut linear DNA during step **B**, a few DNA bases must be added at the ends of the cut sites to ensure optimal activity (shown in bold). Look at NEB cutting close to DNA ends for how many bases are needed for each enzyme.

NOTE 2: Other restriction sites can be used for cloning into the pET vector system, so long as they are (1) present in the multiple cloning site (MCS) and (2) only cut once in the vector. NdeI and EcoRI are routinely used for directional cloning while maintaining the desired reading frame. If other restriction sites are used, double check the sequence is translated properly using Expasy translate tool. Otherwise, adjust the reading frame by adding bases or switching the restriction enzymes used.

START codon

STOP codon

GTTACATATGCTGGCGCGCCATCTGGTGAAACGCGATGGCATTCAGGAAAACAGCTTTAG
CGTGCTGCTGCCGCTGAAACCGACCGGCAGCCGCCCGGCGCTGTTTTGCATTCATCCGGG
CTTTGGCCTGAGCTGGAGCTATATGGGCCTGAGCAGCTATCTGCATCCGGATCAGCCGCT
GTATGGCCTGCAGGCGCGCGGCCTGGATGGCACCAGCCCGCTGGCGAGCACCCTGGATGA
AATGGTGAGCGATTATCTGGAACACATGCGCCGCGTGCAGCCGAAAGGCCCGTATTGCCT
GCTGGGCTGGAGCTTTGGCGGCAACGTGGCGCATAGCATGGCGGTGCGCCTGGAACATCA
GGGCGAAAAAGTGGCGCTGCTGGCGCTGCTGGATAGCACCCCGCTGGTGGGCAGCAACGC
GTTTGCGGAAAAAGAAAACCAGCTGGATCAGGCGATGGTGCGCGATGTGTTTGCGTATCA
TTATAGCAACGAACTGGTGAGCGCGATGGATGAACAGCTGCTGGAAAACACCGAAGAAAT
TGCGAAAAACAACCTGTGCATTGTGCGCGATTATAGCCCGAGCCAGTATGGCGGCAACGC
GCTGCTGTTTCGCGCGACCGTGGCGGAAGCGGGCTGCGAAACCCTGGTGAGCCCGGATGC
GTGGAACCCGTATGTGTTTGGCGGCATTGAAGTGCATGATGTGCATTGCAGCCATGGCGA
AATGCTGAAACCGGAACCGACCGCGATTATTGGCAGCACCCTGGCGTGCAAACTGGATAA
ATGGGAAAGCCAGCAGGCGCAGAAAGTGAACACCGAAGAACTGAAA<u>TAA</u>GAATTC**ATCT**

3. Verify the translated protein sequence produces the desired protein by using the Expasy translate tool.
4. Verify the chosen restriction enzymes (NdeI and EcoRI by default) do not cut in the coding DNA sequence (i.e. between the yellow and green sites) by using the NEBcutter v2.0 tool. If the restriction enzyme cuts in the coding DNA sequence, use different enzymes for cloning or generate a silent mutation that abolishes the cut site while maintaining the desired translated peptide sequence using this Silent mutator tool.
5. Order the generated synthetic gene block through your favourite manufacturer. Typically, we order our gene blocks through ThermoFischer's GeneArt services (connect via WorkDay punchout).
6. Proceed to section **B** once the gene block arrives.

NOTE 1: When ordering your gene block, the manufacturer offers codon optimization and silent mutation services to optimize the coding sequence of your GOI prior to ordering. It is recommended to utilize these services to ensure the correct sequence is ordered. Do not forget to "lock" your restriction sites to ensure they are not codon optimized, as this would otherwise destroy these.

NOTE 2: It is recommended to order the GOI pre-cloned into a shuttle vector. This removes the need for a PCR step to stock up on the gene block. Instead, the pre-cloned GOI can be transformed into chemically competent cells to generate stocks for subsequent cloning into your desired expression vector.

A2: amplifying the GOI from gDNA

1. Using the NCBI database (or other suitable database), find the GOI within the organism's sequenced genome.
2. Once the GOI is found, design a forward primer to be positioned at the 5' and a reverse primer to be positioned at the 3' of the GOI.

NOTE 1: The notation forward and reverse is dependent on the direction of the reading frame. Forward means the primer is positioned at the beginning of the reading frame while the reverse is positioned at the end (i.e. after the STOP codon or C-terminus of the protein target).

NOTE 2: Ensure that the Tm for both primers are within 5 °C from each other with 50% GC content for easier amplification. Primers should be designed to contain a minimum of 20 bp that anneal to the template. Use NEB's Tm calculator to estimate the primers' Tm using Q5 as the polymerase choice.

3. Order the primers from your favourite manufacturer. Typically, idtDNA is used to order our oligos.
4. Isolate gDNA from the organism using an appropriate DNA isolation kit. NEB's Monarch genomic DNA extraction kit (#T3010) can isolate gDNA from yeast, bacteria, blood and organ tissue samples. Follow the manufacturer's recommended instructions for optimal extraction.

NOTE: Sometimes, the kit may not be able to isolate enough gDNA for PCR. If this is the case, an alternative isolation method would be needed. Alternatively, colony PCR can be used to amplify GOI from bacterial, yeast or fungal colonies.

5. Once primers and gDNA has been acquired, prepare a 50 μL PCR reaction and a (-) control as follows (can be scaled down to 25 μL):

Reagent	Sample (μL)	(-) control (μL)
H_2O	32.5	33
Q5 reaction buffer [5X]	10	10
dNTPs mix [10 mM/ea.]	1	1
gDNA template [1 ng/μL]*	1	1
Forward primer [10 μM]	2.5	2.5
Reverse primer [10 μM]	2.5	2.5
Q5 High-Fidelity DNA polymerase	0.5	-

* Template DNA can be replaced with a single colony picked from a petri dish. Cells from colony will release gDNA during denaturation (colony PCR).

NOTE: PCR is done using NEB's Q5 High-Fidelity DNA polymerase. Any high-fidelity DNA polymerase will work. If using a different polymerase, please look at the manufacturer's recommended protocol for PCR amplification, primer Tm calculation and reaction mixture preparation.

6. Design an appropriate PCR protocol using the below protocol as a template.

Initial Denaturation	98 °C	30 seconds
Denaturation	98 °C	10 seconds
Annealing	50-74 °C [a]	30 seconds
Extension	72 °C	30 seconds/kb [b]
Final Extension	72 °C	2 minutes
Hold	12 °C	∞

29-35 x (applies to Denaturation, Annealing, Extension)

[a] Recommend to use NEB's Tm Calculator to calculate theoretical annealing temperature and use the gradient heatblock feature to probe optimal annealing temperature for PCR.

[b] Ensure the extension time allows for amplification of the entire GOI.

7. Mix the entire PCR reaction with 10 µL of Green GoTaq [5X] reaction buffer, run the entire mixture on an agarose gel.
8. Once you have confirmed amplification worked, perform a gel extraction of the positive sample. Proceed to section **B**.

B: digestion, ligation and transformation

1. Prepare the following 50 µL digestion reactions:

Reagent	Amplicon (µL)	Vector (µL)	NdeI control (µL)	EcoRI-HF control (µL)	No digest control (µL)
H_2O	3	38	39	39	40
GOI (pre-cloned or PCR)[a]	30	0	0	0	0
Vector[b]	0	5	5	5	5
CutSmart buffer [10X]	5	5	5	5	5
NdeI	1	1	1	0	0
EcoRI-HF	1	1	0	1	0

[a] If the GOI is pre-cloned, 5 µL of it will suffice. Adjust the volume of water for a 50 µL final reaction volume.

[b] Vector used for cloning is typically either pET-28a(+) or pET-21c(+). Any other expression vector can be used, however, the restriction enzymes needed for cloning may differ.

2. Incubate reactions for a minimum of 1 hour at 37 °C.
3. Mix 5 µL of each reaction with 1 µL of Green GoTaq [5X] reaction buffer, run all 6 µL on an agarose gel.
4. Once you have confirmed digestion was successful, perform a PCR clean-up of the double digested samples.

NOTE: To further minimize vector recircularization, you can dephosphorylate the 5' ends of the doubly digested vector with Shrimp Alkaline Phosphatase (NEB #M0371). Follow the manufacturer's recommended instructions, then purify the DNA using a PCR clean-up kit.

5. Measure the concentration of both digested vector and amplicon samples either via gel electrophoresis (estimation) or via Nanodrop (accurate).
6. Prepare a 20 µL ligation reaction (T4 DNA ligase) using a 3:1 molar ratio of digested vector to amplicon with 40 ng of vector. Use NEB's ligation calculator to calculate the amount of amplicon

(in ng) required. Prepare two negative controls, one containing digested vector and ligase, and another containing only the digested vector.

NOTE: If no concentration could be acquired from the samples, prepare the 20 µL ligation reactions as shown in the table below:

Reagent	Sample (µL)	(-1) control (µL)	(-2) control (µL)
H_2O	9.5	11.5	12
Digested amplicon	2	0	0
Digested vector	6	6	6
T4 DNA ligase buffer [10 X]	2	2	2
T4 DNA ligase	0.5	0.5	0

7. Incubate ligation reactions for a minimum of 10 minutes on the benchtop.

NOTE: Ligation reactions can also be incubated overnight at 16 °C or in the fridge (4 °C). However, ligation works best on the benchtop.

8. After the ligation incubation is complete, transform 5-10 µL into the appropriate *E. coli* chemically competent cells (XL1-Blue, DH5α or Top10). Follow the transformation protocol found on the Boddy lab website. Plate the whole tube after transformation is complete.
9. Count colonies on sample and negative control plates, a ratio of > 1.0 (sample/negative control) indicates cloning is successful. If the ratio is < 1.0, repeat entire protocol.
10. Pick 5-10 clones from the sample plate, then perform miniprep to isolate plasmid DNA and confirm successful cloning via restriction digest.
11. Send positive clones identified in step 10 for sequencing.