**\*\* ONLY TEXT AND IMAGES APPEARING INSIDE THE RED BOXES WILL BE GRADED\*\***

**Introduction**

1. In this assignment you will design and clone a DNA gene that encodes your protein. This would allow you to overexpress, purify, and analyze the protein..
2. Here you will:
   1. design a completely synthetic DNA gene sequence that encodes your protein, but has codons optimized for expression in *E. coli*,
   2. purchase (not really) the synthetic DNA gene sequence from a company ([Eurofins](https://eurofinsgenomics.com/en/products/gene-synthesis/ordering-guide/) or [IDT](https://www.idtdna.com/pages/products/genes-and-gene-fragments/custom-gene-synthesis))
   3. clone the purchased gene to a cloning vector designed for tagged overexpression in *E. coli*,
   4. overexpress and purify the protein, and
   5. analyze it on an SDS-PAGE gel.

**Synthetic DNA sequence design and cloning**

1. Give your 4 character pdb Entry Code, protein name, and organism.

|  |
| --- |
| Insert here |

1. Go to Go the NCBI web page (<https://www.ncbi.nlm.nih.gov/>), pull down the protein tab and get the amino acid sequence of your protein in FASTA format (it does not have to be the same organism as your PDB entry). You already did this in a previous assignment. Paste your amino acid sequence in the box below.

|  |
| --- |
| INSERT FASTA SEQUENCE HERE |

1. Create you’re the sequence of your synthetic gene with the JCat Codon Adaptation Tool ([www.jcat.de](http://www.jcat.de)). Enter the protein sequence from your FASTA file into JCat as shown below.



<< Avoid NdeI and XhoI sites!

<< Choose E. coli strain K-12

<< Choose ‘Protein sequence’

<< Enter your protein’s amino acid sequence

1. You will see something like this



1. Copy and paste your DNA sequence here set the font to Courier.

|  |
| --- |
| Insert DNA sequence here |

Add the *NdeI* recognition sequence to the 5’ end of your gene and the *XhoI* recognition sequence to the 3’ end. Obtain those sequences from the [NEB Enzyme Finder](https://enzymefinder.neb.com/" \l "!). Color the *Nde* I sequence green and the *Xho* I sequence blue.

An example using different restriction sites is shown below.

For example, for *Bam* HI (green) and *Sac* I (blue) recognition sequences:

5’-GGATCCGTCGGTGGCTCTGAGCTC-3’

Change the font to Courier and paste the resulting DNA sequence in the box below.

This is the sequence that you will order from the company (IDT) (not really).

|  |
| --- |
| Insert DNA sequence including restriction sites here |

If IDT charges $0.39 per base pair, how much will your gene cost?

|  |
| --- |
|  |

1. Look the plasmid (cloning vector map) (on the next page). It is the widely used cloning vector pET-28a+. We will use this for cloning and expression of your protein. The vector has an origin of replication, an multiple cloning site, and many other engineered features.

The dark black arrow on the circular representation, pointing around the plasmid to the left, corresponds to the sequence in the box at the bottom the pET-28a-c\_map. The tail of the arrow is the start of the sequence in the box; while the head of the arrow is the end of the sequence. Notice that only the sense (or coding strand) sequence of DNA is shown in the box. Under the DNA sequence are the three letter codes for the amino acids encoded by each codon.

1. pET-28 is a family of **expression vectors**, which means that they are cloning vectors with the transcription and translation signals needed for protein expression (protein production) from a cloned gene. Look carefully at the pET-28a+ sequence.

What is the translation start codon?

|  |
| --- |
|  |

What amino acid does it encode?

|  |
| --- |
|  |

What is the translation stop codon?

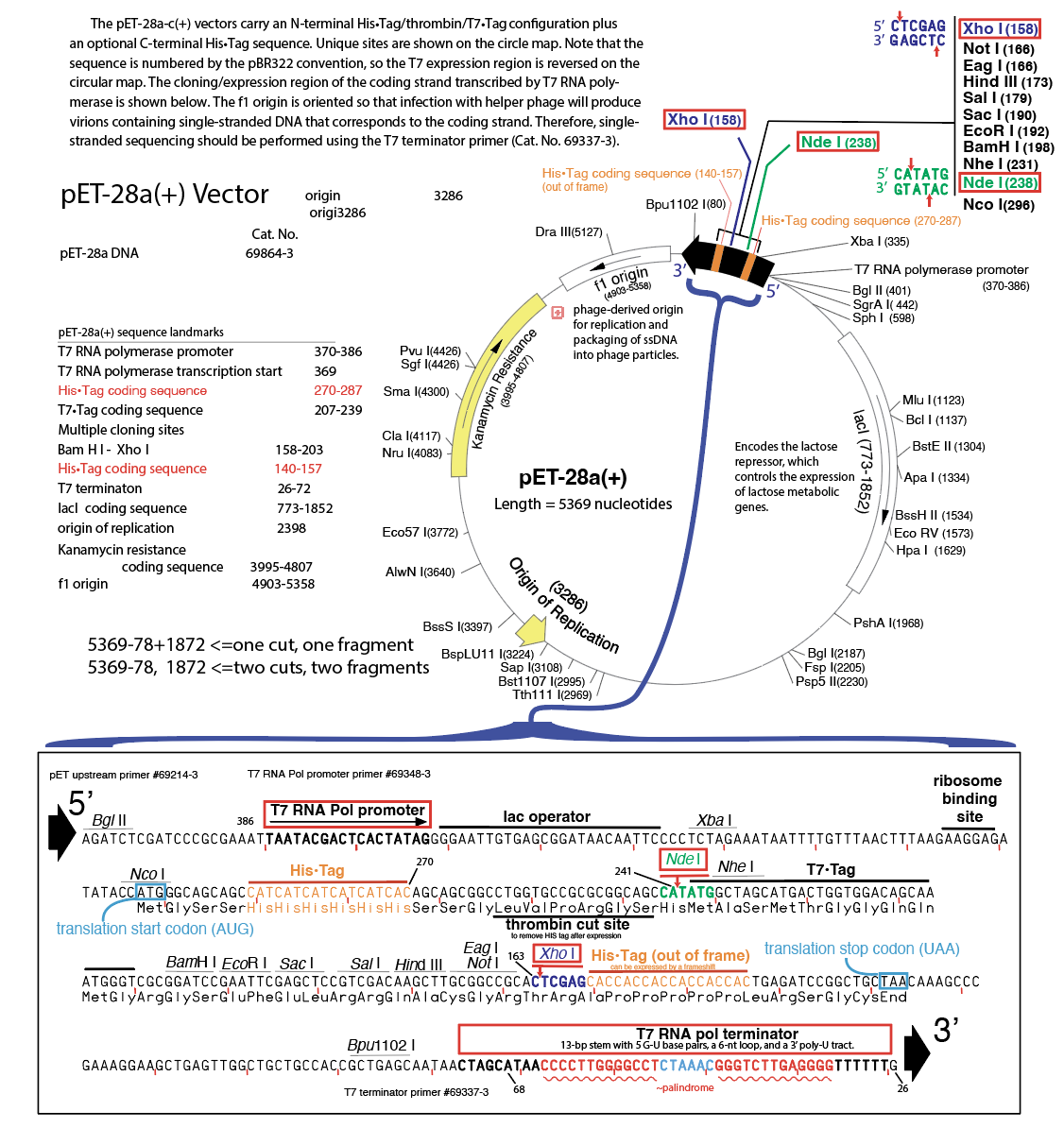
|  |
| --- |
|  |

What amino acid does it encode?

|  |
| --- |
|  |

Name two other transcription or translation signals in this sequence.

|  |
| --- |
|  |



1. Clone your gene to the pET28 expression vector so you can express the protein. Use the *NdeI* and *XhoI* restriction sites that are highlighted in the map.

First, digest your gene with *Nde*I and *Xho*I. This will leave distinct sticky ends on gene.

In Courier font, write both the sense and antisense strands of the *Nde* I recognition sequence, followed by a string of 12 N’s (pretend this is your gene), then the *Xho* I recognition sequence. Color the *Nde* I site green and the *Xho* I site blue. Indicate the 5’ and 3’ ends of each strand.

The cut the DNA with NdeI and XhoI. You can see the sequences of the sticky ends (after cutting) by looking them up at [NEB Enzyme Finder](https://enzymefinder.neb.com/#!). Indicate the exact sense and antisense sequences that will remain with your gene as ‘sticky ends’ after digestion with *Nde* I and *Xho* I by highlighting them yellow. An example with differerent restriction sites (*Bam*HI and *Sac*I) is shown below.

For example, DNA with *Bam*HI and *Sac*I recognition sites before restriction (cutting):

*BamHI* *SacI*

5’-**GGATCC…NNNNNNNNNNNN…GAGCTC**-3’

3’-**CCTAGG…NNNNNNNNNNNN…CTCGAG**-5’

The DNA after restriction (cutting) with *BamHI* and SacI:

5’- **GATCC…NNNNNNNNNNNN…GAGCT**-3’

3’- **G…NNNNNNNNNNNN…C** -5’

*BamHI* leaves a 5′ GATC sticky end

*SacI* leaves a 5′ AGCT sticky end

Insert the gene sequence before and after cutting with NdeI and XhoI in the box below. Use 12 N’s to indicate the plasmid sequence. Use Courier font.

|  |
| --- |
| Before cutting  5’- CATATG**…NNNNNNNNNNNN…**CTCGAG-3’  3’- GTATAC**…NNNNNNNNNNNN…**GAGCTC-5’  After cutting  5’- TATG**…NNNNNNNNNNNN…**C -3’  3’- AC**…NNNNNNNNNNNN…**GAGCT -5’ |

Second, digest the pET28 cloning vector, also with *Nde* I and *Xho* I. What is the length of the DNA fragment that gets excised (cut out) of the plasmid by this double digestion (look at the map)

Digestion of the cloning vector with NdeI and XhoI will create sticky ends complementary to the gene’s sticky ends.

Third, mix the digested pET28 cloning vector and the digested gene with ligase enzyme in buffer to covalently connect the gene and vector at the sticky ends. In the box below, delete the sequence between the *Nde* I and *Xho* I sites and replace it with your gene sequence. You have now ligated your gene into the protein expression vector pET28a+.

1. Transform competent *E. coli* with the ligation mixture. What antibiotic must be used to select transformants?

|  |
| --- |
| INSERT HERE |

3521 STOP HERE

1. Transformation of *E. coli* with ligation mixture can produce many clones. Not all will contain the gene. Some will be wild-type (the pET28a+ vector without a gene inserted). Candidate clones must be screened for the presence of the gene, then verified by sequencing before proceeding to protein expression. Extra or missing bases can cause frameshift and expression of the incorrect protein. There are multiple ways to screen candidate clones for the presence of a gene. Two are digestion and colony PCR. How long (in kb) should the ligated plasmid be if ligation was successful?

|  |
| --- |
| INSERT HERE |

1. Screen for candidate clones by digesting a candidate clone and the original pET28a+ vector with *Nde* I and *Xho* I. Run the digestions side-by-side on a gel. How many fragments, of what sizes (in bp), are expected for the pET28a+ vector? How many fragments, of what lengths (in bp), are expected for a clone containing your gene? Will you be able to differentiate a successful clone from wild-type pET28a+ on a gel by digestion?

|  |
| --- |
| INSERT HERE |

1. Screen for successful clones by colony PCR of a candidate clone and the original pET28a+ vector. You are given the **T7** and **T7term** primers from [Eurofins Genomics standard primers list](https://eurofinsgenomics.com/en/products/dna-sequencing/standard-primers/). What is the expected amplification product length (in bp) for a candidate clone if it is wild-type? If it contains your gene? Will you be able to differentiate a successful clone from wild-type pET28a+ on a gel by PCR?

|  |
| --- |
| INSERT HERE |

1. Indicate the cycling conditions for the PCR performed in part 14).

|  |  |  |
| --- | --- | --- |
| **Step** | **Temperature (°C)** | **Duration (s)** |
| 1. Initial denaturation |  |  |
| 1. Denaturation |  |  |
| 1. Annealing |  |  |
| 1. Extension |  |  |
| 1. Go to 2, Cycle 30x |  |  |
| 1. Final Extension |  |  |
| 1. Hold |  |  |

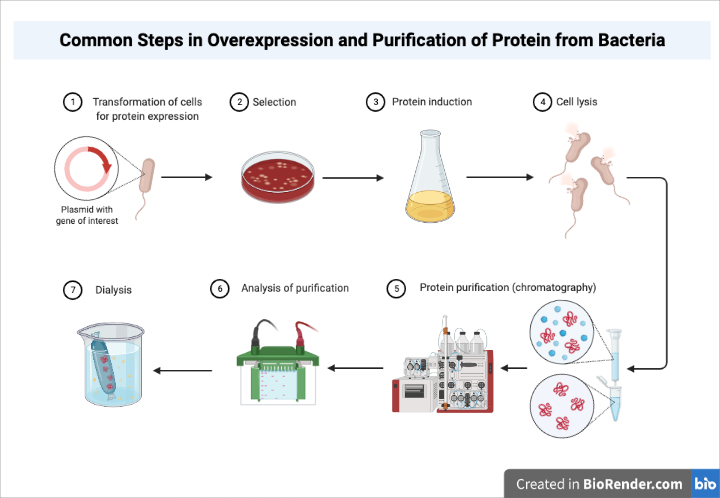
1. You send the candidate clone for sequencing. Check your sequencing data by translating the entire sequence from part 10) (b) with the translate tool <https://web.expasy.org/translate/>. In Results of Translation, the 5’3’ Frame 3 should contain the expressed protein sequence (but check that it matches). Insert a screenshot of the frame that shows the correct expression of your tagged protein from pET28a+.

|  |
| --- |
| INSERT HERE |

1. Which protein tag is in frame in the expressed protein sequence? Is it present on the N-terminal, C-terminal, or both termini of your protein? Which tag can be removed after protein expression?

|  |
| --- |
| INSERT HERE |

1. Transform *E. coli* cells engineered for protein expression (e.g., BL21) with the sequence-confirmed clone and follow the steps shown below.



1. In steps 2 and 3 (selection and protein induction), what antibiotic will you use to select for transformed *E. coli*? What molecule will you add to the culture to induce overexpression? Give the acronym and chemical name.

|  |
| --- |
| INSERT HERE |

1. In step 5 (protein purification), you will use a matrix or resin in a gravity column to bind the tag(s) covalently attached to your protein, followed by size exclusion chromatography. What does the immobilized ligand in the resin need to be to capture your tagged protein?
2. As in step 6 (analysis of purification), sketch an SDS-PAGE gel with lanes that contain i) a molecular weight marker, ii) cell lysate of *E. coli* before induction, iii) cell lysate of *E. coli* after induction, iv) resin elution fractions, and v) gel filtration / size exclusion chromatography effluent fractions. Your gel should have 10 lanes.

|  |
| --- |
| INSERT HERE |

1. (omit this question) Sketch the Circular Dichroism spectrum of your protein as discussed in class.

|  |
| --- |
| INSERT HERE |