Bioproduction

This week, I created a custom plasmid for bioproduction. I then ordered the relevant gene from Twist. A strong contender for my final project is to use E coli to produce EGFL7, epidermal growth factor-like protein 7 in humans. To do this, I first optimized the protein codons for human:

Codon Optimization Tool

RESTRICTION SITE INFORMATION ⓘ

Sequence Type: ○ DNA Bases ⦿ Amino Acids

Product Type: gBlocks® Gene Fragments

Organism: Escherichia coli K12

Single Entry   **Bulk Entry**

MRGSQEVLLMWLLVLAVGGTEHAYRPGRRVCAVRAHGDPVSESFVQRVYQPFLTTCDGHRACSTYRTIYRTAYRRSPGLAPARPRYACCPGWKRTSGLPGACGAAICQPPCRNGGSCVQPGRC
RCPAGWRGDTCQSDVDECSARRGGCPQRCVNTAGSYWCQCWEGHSLSADGTLCVPKGGPPRVAPNPTGVDSAMKEEVQRLQSRVDLLEEKLQLVLAPLHSLASQALEHGLPDPGSLLVHSFQQ
LGRIDSLSEQISFLEEQLGSCSCKKDS

# of bases: 273

MAP CODONS ONLY   OPTIMIZE

Results   ORDER 1 ITEM

MANUAL OPTIMIZATION

ATG CGT GGG TCC CAG GAG GTC TTA TTG ATG TGG CTG CTT GTC TTG GCG GTT GGG GGT ACC GAA CAT GCC TAC CGT CCG GGA CGC
CGC GTT TGT GCC GTG CGT GCT CAC GGT GAT CCT GTG TCC GAA AGT TTC GTC CAG CGT GTC TAC CAG CCT TTT TTA ACG ACG TGC
GAC GGG CAT CGT GCC TGC AGC ACG TAT CGC ACC ATT TAC CGT ACT GCG TAC CGC CGC TCC CCT GGT TTA GCA CCG GCT CGT CCG
CGC TAT GCG TGT TGT CCC GGC TGG AAA CGC ACT TCT GGC TTA CCT GGT GCT TGT GGC GCT GCA ATT TGT CAA CCG CCT TGT CGC
AAT GGT GGG TCT TGC GTA CAG CCG GGT CGC TGC CGT TGT CCT GCT GGT TGG CGC GGA GAT ACA TGT CAA TCG GAC GTG GAC GAG
TGT AGC GCC CGC CGC GGA GGG TGT CCA CAA CGT TGC GTG AAC ACC GCA GGA TCG TAT TGG TGC CAG TGC TGG GAG GGT CAC TCT
CTT AGT GCT GAT GGC ACT TTG TGC GTG CCG AAA GGG GGA CCT CCC CGT GTC GCC CCG AAT CCC ACC GGA GTT GAC TCA GCA ATG
AAA GAG GAA GTA CAG CGT CTT CAA TCC CGT GTG GAT TTG CTT GAG GAA AAG CTG CAA CTT GTT CTG GCG CCT TTA CAC TCC TTG
GCT TCA CAG GCT CTG GAA CAT GGT CTT CCA GAC CCA GGT TCC CTG CTG GTG CAC TCC TTC CAA CAG TTG GGA CGT ATC GAT AGC
TTA AGC GAG CAG ATT TCG TTT TTG GAA GAA CAG TTA GGG AGC TGT TCG TGC AAA AAA GAT AGC

Accepted - Low Complexity (Scores less than 7)

Some complexities exist but we do not anticipate problems with this sequence.

Total Complexity Score: 0.8

Complexity Description	Score
This sequence contains a window of 100 bases starting at base 340 with a GC content of 66%. Solution: Redesign this region to have a GC content less than 65%.	0.8

Possible restriction sites are:

The following restriction enzyme sites have been found in the selected reading frame:
AclI (AACGTT)
AflII (CTTAAG)
ApaLI (GTGCAC)
ClaI (ATCGAT)
KpnI (GGTACC)
NarI (GGCGCC)
PciI (ACATGT)
PstI (CTGCAG)
SacII (CCGCGG)

Restriction Site Information

I then placed a RBS 8bp before the ATG start codon.

In order to more easily monitor protein production, I used a simple, flexible linker found on the registry of standard biological parts for iGEM:

Registry of Standard Biological Parts

iGEM   tools   catalog   repository   assembly   protocols   help   search   [BBa_____]

main page   design   experience   information   part tools   edit

Part:BBa_K648007
Designed by: Jim Rose Group: iGEM11_Penn_State (2011-07-03)

Protein_Domain

| Not Released |
| Sample Not in stock |
| Experience: None |
| 4 Uses |
| Get This Part |

Medium 6AA Fusion Protein Linker: GGSGGS with Standard 25 Prefix/Suffix

This is the mid-sized linker of the three fusion protein linkers created by the Penn State 2011 iGEM team. It encodes for the six amino acids GGSGGS and has a prefix and suffix compatible with assembly standard 25. It is used in one of the variants of our Fast-Fusion protein reporter system.

Sequence and Features

Subparts | **Ruler** | SS | DS   Scars: Show | **Hide**   Vertical: Show | Hide   **Length: 18 bp**   View plasmid ◯ Get part sequence

1

Assembly Compatibility: 10 12 21 23 25 1000

[edit]

Parameters	Categories
None	//proteindomain/linker

And linked this to GFP, found through the addgene website.