Appendix A

DNA SEQUENCES AND PLASMID MAPS

TABLE OF CONTENTS

A.1 pETcyl-phes* plasmid information

Submitted by: Stacey Maskarinec, Marissa Mock **Date 4/22/04**

Strain name: **XL-1 blue (***SupE*44 *hsdR*17 *recA*1 *endA*1 *gyrA*46 *thi relA*1 *lac*$^-$ F'[*proAB*$^+$ *lacI*^{q} l*acZ*ΔM15 Tn*10*(*tet*^{r})] (From Stratagene) **) / pET28cyl-phes***

Vector (kb): **pET28cyl-phes* (7.0)**

Cloning site: See plasmid map.

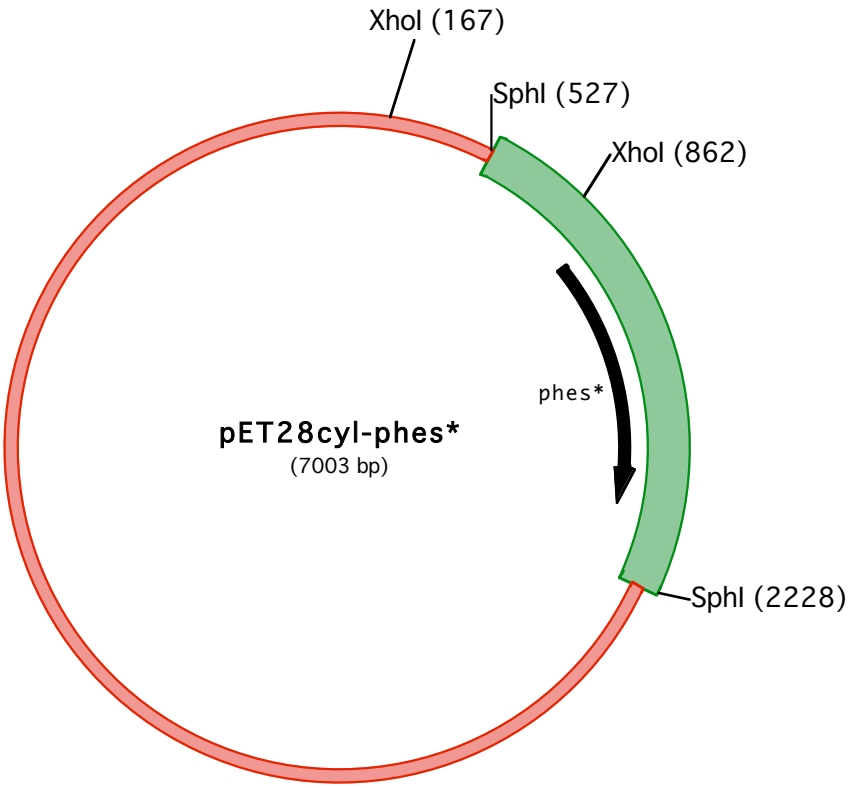
Construction of pET28cyl-phes* plasmid:

The phes***** gene was cut from pUC19-FS (Nandita Sharma) with Sph I. It was ligated into the Sph I site of pET28cyl (Charles Liu).

Source available :

12 % Glycerol culture in Marissa-3 freezer box
Culture conditions: 2xYT, 37°C

A.1.1 Plasmid map



A.1.2 Position of elements

	bp
Vector size	7003
Full insert region	532-2233
PheRS* gene	992-1073

A.1.3 Full sequence (pET28cyl-phes)*

```
  1 ATCCGGATATAGTTCCTCCTTTCAGCAAAAAACCCCTCAAGACCCGTTTAGAGGCCCCAAGGGGTTATGCTAGTTATT
    TAGGCCTATATCAAGGAGGAAAGTCGTTTTTTGGGGAGTTCTGGGCAAATCTCCGGGGTTCCCCAATACGATCAATAA

 79 GCTCAGCGGTGGCAGCAGCCAACTCAGCTTCCTTTCGGGCTTTGTTAGCAGCCGGATCTCAGTGGTGGTGGTGGTGGT
    CGAGTCGCCACCGTCGTCGGTTGAGTCGAAGGAAAGCCCGAAACAATCGTCGGCCTAGAGTCACCACCACCACCACCA

                Xhol (167)
157 GCTCGA ctt act cga gaa gct t CAT GTG GTG GTG GTG GTG GTG ACC CAT TTG CTG TCC
    CGAGCT gaa tga gct ctt cga a GTA CAC CAC CAC CAC CAC CAC TGG GTA AAC GAC AGG

215 ACC AGT CAT GCT CGC CAT gGTATATCTCCTTCTTAAAGTTAAACAAAATTATTTCTAGAGGGGAATTGTTAT
    TGG TCA GTA CGA GCG GTA cCATATAGAGGAAGAATTTCAATTTGTTTTAATAAAGATCTCCCCTTAACAATA

287 CCGCTCACAATTCCCCTATAGTGAGTCGTATTAATTTCGCGGGATCGAGATCTCGATCCTCTACGCCGGACGCATCGT
    GGCGAGTGTTAAGGGGATATCACTCAGCATAATTAAAGCGCCCTAGCTCTAGAGCTAGGAGATGCGGCCTGCGTAGCA

365 GGCCGGCATCACCGGCGCCACAGGTGCGGTTGCTGGCGCCTATATCGCCGACATCACCGATGGGGAAGATCGGGCTCG
    CCGGCCGTAGTGGCCGCGGTGTCCACGCCAACGACCGCGGATATAGCGGCTGTAGTGGCTACCCCTTCTAGCCCGAGC

443 CCACTTCGGGCTCATGAGCGCTTGTTTCGGCGTGGGTATGGTGGCAGGCCCCGTGGCCGGGGGACTGTTGGGCGCCAT
    GGTGAAGCCCGAGTACTCGCGAACAAAGCCGCACCCATACCACCGTCCGGGGCACCGGCCCCCTGACAACCCGCGGTA

          SphI (527)
521 CTCCTTGCATGccgaacgaccgagcgcagcgagtcagtgagcgaggaagcggaagagcgcccaatacgcaaaccgcct
    GAGGAACGTACggcttgctggctcgcgtcgctcagtcactcgctccttcgccttctcgcgggttatgcgtttggcgga

599 ctccccgcgcgttggccgattcattaatgcagctggcacgacaggtttcccgactggaaagcgggcagtgagcgcaac
    gaggggcgcgcaaccggctaagtaattacgtcgaccgtgctgtccaaagggctgacctttcgcccgtcactcgcgttg

677 gcaattaatgtgagttagctcactcattaggcaccccaggctttacactttatgcttccggctcgtatgttgtgtgga
    cgttaattacactcaatcgagtgagtaatccgtggggtccgaaatgtgaaatacgaaggccgagcatacaacacacct

755 attgtgagcggataacaatttcacacaggaaacagctatgaccatgattacgccaagctcgaaattaaccctcactaa
    taacactcgcctattgttaaagtgtgtcctttgtcgatactggtactaatgcggttcgagctttaattgggagtgatt

                                Xhol (862)
833 agggaacaaaagctggtaccgggccccccctcgaggtcgacggtatcgataagcttgatatcgaattcccccgggacc
    tcccttgttttcgaccatggcccgggggggagctccagctgccatagctattcgaactatagcttaaggggccctgg

911 aaaatggcaagtaaaatagcctgatgggataggctctaagtccaacgaaccagtgtcaccactg
    ttttaccgttcattttatcggactaccctatccgagattcaggttgcttggtcacagtggtgac
```

```
 975 acacaatgaggaaaaccatgtcacatctcgcagaactggttgccagtgcgaaggcggccattagccaggcgtcagat
     tgtgttactccttttggtacagtgtagagcgtcttgaccaacggtcacgcttccgccggtaatcggtccgcagtcta
                   1▶ M  S  H  L  A  E  L  V  A  S  A  K  A  A  I  S  Q  A  S  D

1052 gttgccgcgttagataatgtgcgcgtcgaatatttgggtaaaaaagggcacttaacccttcagatgacgaccctgcg
     caacggcgcaatctattacacgcgcagcttataaacccatttttcccgtgaattgggaagtctactgctgggacgc
  21▶ V  A  A  L  D  N  V  R  V  E  Y  L  G  K  K  G  H  L  T  L  Q  M  T  T  L  R

1129 tgagctgccgccagaagagcgtccggcagctggtgcggttatcaacgaagcgaaagagcaggttcagcaggcgctga
     actcgacggcggtcttctcgcaggccgtcgaccacgccaatagttgcttcgctttctcgtccaagtcgtccgcgact
  46▶  E  L  P  P  E  E  R  P  A  A  G  A  V  I  N  E  A  K  E  Q  V  Q  Q  A  L

1206 atgcgcgtaaagcggaactggaaagcgctgcactgaatgcgcgtctggcggcggaaacgattgatgtctctctgcca
     tacgcgcatttcgccttgacctttcgcgacgtgacttacgcgcagaccgccgcctttgctaactacagagagacggt
  72▶N  A  R  K  A  E  L  E  S  A  A  L  N  A  R  L  A  A  E  T  I  D  V  S  L  P

1283 ggtcgtcgcattgaaaacggcggtctgcatccggttacccgtaccatcgaccgtatcgaaagtttcttcggtgagct
     ccagcagcgtaacttttgccgccagacgtaggccaatgggcatggtagctggcatagctttcaaagaagccactcga
  98▶ G  R  R  I  E  N  G  G  L  H  P  V  T  R  T  I  D  R  I  E  S  F  F  G  E  L

1360 tggctttaccgtggcaaccgggccggaaatcgaagacgattatcataacttcgatgctctgaacattcctggtcacc
     accgaaatggcaccgttggcccggcctttagcttctgctaatagtattgaagctacgagacttgtaaggaccagtgg
 123▶  G  F  T  V  A  T  G  P  E  I  E  D  D  Y  H  N  F  D  A  L  N  I  P  G  H

1437 acccggcgcgcgctgaccacgacactttctggtttgacactacccgcctgctgcgtacccagacctctggcgtacag
     tgggccgcgcgcgactggtgctgtgaaagaccaaactgtgatgggcggacgacgcatgggtctggagaccgcatgtc
 149▶H  P  A  R  A  D  H  D  T  F  W  F  D  T  T  R  L  L  R  T  Q  T  S  G  V  Q

1514 atccgcaccatgaaagcccagcagccaccgattcgtatcatcgcgcctggccgtgtttatcgtaacgactacgacca
     taggcgtggtactttcgggtcgtcggtggctaagcatagtagcgcggaccggcacaaatagcattgctgatgctggt
 175▶ I  R  T  M  K  A  Q  Q  P  P  I  R  I  I  A  P  G  R  V  Y  R  N  D  Y  D  Q

1591 gactcacacgccgatgttccatcagatggaaggtctgattgttgataccaacatcagctttaccaacctgaaaggca
     ctgagtgtgcggctacaaggtagtctaccttccagactaacaactatggttgtagtcgaaatggttggactttccgt
 200▶  T  H  T  P  M  F  H  Q  M  E  G  L  I  V  D  T  N  I  S  F  T  N  L  K  G

1668 cgctgcacgacttcctgcgtaacttctttgaggaagatttgcagattcgcttccgtccttcctacttcccgtttacc
     gcgacgtgctgaaggacgcattgaagaaactccttctaaacgtctaagcgaaggcaggaaggatgaagggcaaatgg
 226▶T  L  H  D  F  L  R  N  F  F  E  E  D  L  Q  I  R  F  R  P  S  Y  F  P  F  T

1745 gaaccttctgcagaagtggacgtcatgggtaaaaacggtaaatggctggaagtgctgggctgcgggatggtgcatcc
     cttggaagacgtcttcacctgcagtacccatttttgccatttaccgaccttcacgacccgacgccctaccacgtagg
 252▶ E  P  S  A  E  V  D  V  M  G  K  N  G  K  W  L  E  V  L  G  C  G  M  V  H  P

1822 gaacgtgttgcgtaacgttggcatcgacccggaagtttactctggtttcggcttcgggatggggatggagcgtctga
     cttgcacaacgcattgcaaccgtagctgggccttcaaatgagaccaaagccgaagccctacccctacctcgcagact
 277▶  N  V  L  R  N  V  G  I  D  P  E  V  Y  S  G  F  G  F  G  M  G  M  E  R  L

1899 ctatgttgcgttacggcgtcaccgacctgcgttcattcttcgaaaacgatctgcgtttcctcaaacagtttaaataa
     gatacaacgcaatgccgcagtggctggacgcaagtaagaagcttttgctagacgcaaaggagtttgtcaaatttatt
 303▶T  M  L  R  Y  G  V  T  D  L  R  S  F  F  E  N  D  L  R  F  L  K  Q  F  K

1976 ggcaggaatagattatgaaattcagtgaactgtggttacgcgaatgggtgaacccggcgattgatagcgatgcgctg
     ccgtccttatctaatactttaagtcacttgacaccaatgcgcttacccacttgggccgctaactatcgctacgcgac
```

```
1976 ggcaggaatagattatgaaattcagtgaactgtggttacgcgaatgggtgaacccggcgattgatagcgatgcgctg
     ccgtccttatctaatactttaagtcacttgacaccaatgcgcttacccacttgggccgctaactatcgctacgcgac

2053 gcaaatcaaatcactatggcgggcctggaagttgggggatccactagttctagagcggccgccaccgcggtggagct
     cgtttagtttagtgataccgcccggaccttcaacccccctaggtgatcaagatctcgccggcggtggcgccacctcga

2130 ccaattcgccctatagtgagtcgtattacaattcactggccgtcgttttacaacgtcgtgactgggaaaaccctggc
     ggttaagcgggatatcactcagcataatgttaagtgaccggcagcaaaatgttgcagcactgacccttttgggaccg

                          SphI (2228)
2207 gttacccaacttaatcgccttgcatgCACCATTCCTTGCGGCGGCGGTGCTCAACGGCCTCAACCTACTACTGGGCT
     caatgggttgaattagcggaacgtacGTGGTAAGGAACGCCGCCGCCACGAGTTGCCGGAGTTGGATGATGACCCGA

2284 GCTTCCTAATGCAGGAGTCGCATAAGGGAGAGCGTCGAGATCCCGGACACCATCGAATGGCGCAAAACCTTTCGCGG
     CGAAGGATTACGTCCTCAGCGTATTCCCTCTCGCAGCTCTAGGGCCTGTGGTAGCTTACCGCGTTTTGGAAAGCGCC

2361 TATGGCATGATAGCGCCCGGAAGAGAGTCAATTCAGGGTGGTGAATGTGAAACCAGTAACGTTATACGATGTCGCAG
     ATACCGTACTATCGCGGGCCTTCTCTCAGTTAAGTCCCACCACTTACACTTTGGTCATTGCAATATGCTACAGCGTC

2438 AGTATGCCGGTGTCTCTTATCAGACCGTTTCCCGCGTGGTGAACCAGGCCAGCCACGTTTCTGCGAAAACGCGGGAA
     TCATACGGCCACAGAGAATAGTCTGGCAAAGGGCGCACCACTTGGTCCGGTCGGTGCAAAGACGCTTTTGCGCCCTT

2515 AAAGTGGAAGCGGCGATGGCGGAGCTGAATTACATTCCCAACCGCGTGGCACAACAACTGGCGGGCAAACAGTCGTT
     TTTCACCTTCGCCGCTACCGCCTCGACTTAATGTAAGGGTTGGCGCACCGTGTTGTTGACCGCCCGTTTGTCAGCAA

2592 GCTGATTGGCGTTGCCACCTCCAGTCTGGCCCTGCACGCGCCGTCGCAAATTGTCGCGGCGATTAAATCTCGCGCCG
     CGACTAACCGCAACGGTGGAGGTCAGACCGGGACGTGCGCGGCAGCGTTTAACAGCGCCGCTAATTTAGAGCGCGGC

2669 ATCAACTGGGTGCCAGCGTGGTGGTGTCGATGGTAGAACGAAGCGGCGTCGAAGCCTGTAAAGCGGCGGTGCACAAT
     TAGTTGACCCACGGTCGCACCACCACAGCTACCATCTTGCTTCGCCGCAGCTTCGGACATTTCGCCGCCACGTGTTA

2746 CTTCTCGCGCAACGCGTCAGTGGGCTGATCATTAACTATCCGCTGGATGACCAGGATGCCATTGCTGTGGAAGCTGC
     GAAGAGCGCGTTGCGCAGTCACCCGACTAGTAATTGATAGGCGACCTACTGGTCCTACGGTAACGACACCTTCGACG

2823 CTGCACTAATGTTCCGGCGTTATTTCTTGATGTCTCTGACCAGACACCCATCAACAGTATTATTTTCTCCCATGAAG
     GACGTGATTACAAGGCCGCAATAAAGAACTACAGAGACTGGTCTGTGGGTAGTTGTCATAATAAAAGAGGGTACTTC

2900 ACGGTACGCGACTGGGCGTGGAGCATCTGGTCGCATTGGGTCACCAGCAAATCGCGCTGTTAGCGGGCCCATTAAGT
     TGCCATGCGCTGACCCGCACCTCGTAGACCAGCGTAACCCAGTGGTCGTTTAGCGCGACAATCGCCCGGGTAATTCA

2977 TCTGTCTCGGCGCGTCTGCGTCTGGCTGGCTGGCATAAATATCTCACTCGCAATCAAATTCAGCCGATAGCGGAACG
     AGACAGAGCCGCGCAGACGCAGACCGACCGACCGTATTTATAGAGTGAGCGTTAGTTTAAGTCGGCTATCGCCTTGC

3054 GGAAGGCGACTGGAGTGCCATGTCCGGTTTTCAACAAACCATGCAAATGCTGAATGAGGGCATCGTTCCCACTGCGA
     CCTTCCGCTGACCTCACGGTACAGGCCAAAAGTTGTTTGGTACGTTTACGACTTACTCCCGTAGCAAGGGTGACGCT

3131 TGCTGGTTGCCAACGATCAGATGGCGCTGGGCGCAATGCGCGCCATTACCGAGTCCGGGCTGCGCGTTGGTGCGGAT
     ACGACCAACGGTTGCTAGTCTACCGCGACCCGCGTTACGCGCGGTAATGGCTCAGGCCCGACGCGCAACCACGCCTA

3208 ATCTCGGTAGTGGGATACGACGATACCGAAGACAGCTCATGTTATATCCCGCCGTTAACCACCATCAAACAGGATTT
     TAGAGCCATCACCCTATGCTGCTATGGCTTCTGTCGAGTACAATATAGGGCGGCAATTGGTGGTAGTTTGTCCTAAA

3285 TCGCCTGCTGGGGCAAACCAGCGTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGGCGGTGAAGGGCAATCAGCTGT
     AGCGGACGACCCCGTTTGGTCGCACCTGGCGAACGACGTTGAGAGAGTCCCGGTCCGCCACTTCCCGTTAGTCGACA

3362 TGCCCGTCTCACTGGTGAAAAGAAAAACCACCCTGGCGCCCAATACGCAAACCGCCTCTCCCCGCGCGTTGGCCGAT
     ACGGGCAGAGTGACCACTTTTCTTTTTGGTGGGACCGCGGGTTATGCGTTTGGCGGAGAGGGGCGCGCAACCGGCTA

3439 TCATTAATGCAGCTGGCACGACAGGTTTCCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTAAGTTAGC
     AGTAATTACGTCGACCGTGCTGTCCAAAGGGCTGACCTTTCGCCCGTCACTCGCGTTGCGTTAATTACATTCAATCG

3516 TCACTCATTAGGCACCGGGATCTCGACCGATGCCCTTGAGAGCCTTCAACCCAGTCAGCTCCTTCCGGTGGGCGCGG
     AGTGAGTAATCCGTGGCCCTAGAGCTGGCTACGGGAACTCTCGGAAGTTGGGTCAGTCGAGGAAGGCCACCCGCGCC

3593 GGCATGACTATCGTCGCCGCACTTATGACTGTCTTCTTTATCATGCAACTCGTAGGACAGGTGCCGGCAGCGCTCTG
     CCGTACTGATAGCAGCGGCGTGAATACTGACAGAAGAAATAGTACGTTGAGCATCCTGTCCACGGCCGTCGCGAGAC

3670 GGTCATTTTCGGCGAGGACCGCTTTCGCTGGAGCGCGACGATGATCGGCCTGTCGCTTGCGGTATTCGGAATCTTGC
     CCAGTAAAAGCCGCTCCTGGCGAAAGCGACCTCGCGCTGCTACTAGCCGGACAGCGAACGCCATAAGCCTTAGAACG

3747 ACGCCCTCGCTCAAGCCTTCGTCACTGGTCCCGCCACCAAACGTTTCGGCGAGAAGCAGGCCATTATCGCCGGCATG
     TGCGGGAGCGAGTTCGGAAGCAGTGACCAGGGCGGTGGTTTGCAAAGCCGCTCTTCGTCCGGTAATAGCGGCCGTAC

3824 GCGGCCCCACGGGTGCGCATGATCGTGCTCCTGTCGTTGAGGACCCGGCTAGGCTGGCGGGGTTGCCTTACTGGTTA
     CGCCGGGGTGCCCACGCGTACTAGCACGAGGACAGCAACTCCTGGGCCGATCCGACCGCCCCAACGGAATGACCAAT

3901 GCAGAATGAATCACCGATACGCGAGCGAACGTGAAGCGACTGCTGCTGCAAAACGTCTGCGACCTGAGCAACAACAT
     CGTCTTACTTAGTGGCTATGCGCTCGCTTGCACTTCGCTGACGACGACGTTTTGCAGACGCTGGACTCGTTGTTGTA

3978 GAATGGTCTTCGGTTTCCGTGTTTCGTAAAGTCTGGAAACGCGGAAGTCAGCGCCCTGCACCATTATGTTCCGGATC
     CTTACCAGAAGCCAAAGGCACAAAGCATTTCAGACCTTTGCGCCTTCAGTCGCGGGACGTGGTAATACAAGGCCTAG
```

```
4055 TGCATCGCAGGATGCTGCTGGCTACCCTGTGGAACACCTACATCTGTATTAACGAAGCGCTGGCATTGACCCTGAGT
     ACGTAGCGTCCTACGACGACCGATGGGACACCTTGTGGATGTAGACATAATTGCTTCGCGACCGTAACTGGGACTCA

4132 GATTTTTCTCTGGTCCCGCCGCATCCATACCGCCAGTTGTTTACCCTCACAACGTTCCAGTAACCGGGCATGTTCAT
     CTAAAAAGAGACCAGGGCGGCGTAGGTATGGCGGTCAACAAATGGGAGTGTTGCAAGGTCATTGGCCCGTACAAGTA

4209 CATCAGTAACCCGTATCGTGAGCATCCTCTCTCGTTTCATCGGTATCATTACCCCCATGAACAGAAATCCCCCTTAC
     GTAGTCATTGGGCATAGCACTCGTAGGAGAGAGCAAAGTAGCCATAGTAATGGGGGTACTTGTCTTTAGGGGGAATG

4286 ACGGAGGCATCAGTGACCAAACAGGAAAAAACCGCCCTTAACATGGCCCGCTTTATCAGAAGCCAGACATTAACGCT
     TGCCTCCGTAGTCACTGGTTTGTCCTTTTTTGGCGGGAATTGTACCGGGCGAAATAGTCTTCGGTCTGTAATTGCGA

4363 TCTGGAGAAACTCAACGAGCTGGACGCGGATGAACAGGCAGACATCTGTGAATCGCTTCACGACCACGCTGATGAGC
     AGACCTCTTTGAGTTGCTCGACCTGCGCCTACTTGTCCGTCTGTAGACACTTAGCGAAGTGCTGGTGCGACTACTCG

4440 TTTACCGCAGCTGCCTCGCGCGTTTCGGTGATGACGGTGAAAACCTCTGACACATGCAGCTCCCGGAGACGGTCACA
     AAATGGCGTCGACGGAGCGCGCAAAGCCACTACTGCCACTTTTGGAGACTGTGTACGTCGAGGGCCTCTGCCAGTGT

4517 GCTTGTCTGTAAGCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTTGGCGGGTGTCGGGGCGC
     CGAACAGACATTCGCCTACGGCCCTCGTCTGTTCGGGCAGTCCCGCGCAGTCGCCCACAACCGCCCACAGCCCCGCG

4594 AGCCATGACCCAGTCACGTAGCGATAGCGGAGTGTATACTGGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAG
     TCGGTACTGGGTCAGTGCATCGCTATCGCCTCACATATGACCGAATTGATACGCCGTAGTCTCGTCTAACATGACTC

4671 AGTGCACCATATATGCGGTGTGAAATACCGCACAGATGCGTAAGGAGAAAATACCGCATCAGGCGCTCTTCCGCTTC
     TCACGTGGTATATACGCCACACTTTATGGCGTGTCTACGCATTCCTCTTTTATGGCGTAGTCCGCGAGAAGGCGAAG

4748 CTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGT
     GAGCGAGTGACTGAGCGACGCGAGCCAGCAAGCCGACGCCGCTCGCCATAGTCGAGTGAGTTTCCGCCATTATGCCA

4825 TATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAG
     ATAGGTGTCTTAGTCCCCTATTGCGTCCTTTCTTGTACACTCGTTTTCCGGTCGTTTTCCGGTCCTTGGCATTTTTC

4902 GCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTG
     CGGCGCAACGACCGCAAAAAGGTATCCGAGGCGGGGGGACTGCTCGTAGTGTTTTTAGCTGCGAGTTCAGTCTCCAC

4979 GCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCC
     CGCTTTGGGCTGTCCTGATATTTCTATGGTCCGCAAAGGGGGACCTTCGAGGGAGCACGCGAGAGGACAAGGCTGGG

5056 TGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTAT
     ACGGCGAATGGCCTATGGACAGGCGGAAAGAGGGAAGCCCTTCGCACCGCGAAAGAGTATCGAGTGCGACATCCATA

5133 CTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTT
     GAGTCAAGCCACATCCAGCAAGCGAGGTTCGACCCGACACACGTGCTTGGGGGGCAAGTCGGGCTGGCGACGCGGAA

5210 ATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGA
     TAGGCCATTGATAGCAGAACTCAGGTTGGGCCATTCTGTGCTGAATAGCGGTGACCGTCGTCGGTGACCATTGTCCT

5287 TTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACA
     AATCGTCTCGCTCCATACATCCGCCACGATGTCTCAAGAACTTCACCACCGGATTGATGCCGATGTGATCTTCCTGT

5364 GTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAAC
     CATAAACCATAGACGCGAGACGACTTCGGTCAATGGAAGCCTTTTTCTCAACCATCGAGAACTAGGCCGTTTGTTTG

5441 CACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTT
     GTGGCGACCATCGCCACCAAAAAAACAAACGTTCGTCGTCTAATGCGCGTCTTTTTTTCCTAGAGTTCTTCTAGGAA

5518 TGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAACAATAAAACT
     ACTAGAAAAGATGCCCCAGACTGCGAGTCACCTTGCTTTTGAGTGCAATTCCCTAAAACCAGTACTTGTTATTTTGA

5595 GTCTGCTTACATAAACAGTAATACAAGGGGTGTTATGAGCCATATTCAACGGGAAACGTCTTGCTCTAGGCCGCGAT
     CAGACGAATGTATTTGTCATTATGTTCCCCACAATACTCGGTATAAGTTGCCCTTTGCAGAACGAGATCCGGCGCTA

5672 TAAATTCCAACATGGATGCTGATTTATATGGGTATAAATGGGCTCGCGATAATGTCGGGCAATCAGGTGCGACAATC
     ATTTAAGGTTGTACCTACGACTAAATATACCCATATTTACCCGAGCGCTATTACAGCCCGTTAGTCCACGCTGTTAG

5749 TATCGATTGTATGGGAAGCCCGATGCGCCAGAGTTGTTTCTGAAACATGGCAAAGGTAGCGTTGCCAATGATGTTAC
     ATAGCTAACATACCCTTCGGGCTACGCGGTCTCAACAAAGACTTTGTACCGTTTCCATCGCAACGGTTACTACAATG

5826 AGATGAGATGGTCAGACTAAACTGGCTGACGGAATTTATGCCTCTTCCGACCATCAAGCATTTTATCCGTACTCCTG
     TCTACTCTACCAGTCTGATTTGACCGACTGCCTTAAATACGGAGAAGGCTGGTAGTTCGTAAAATAGGCATGAGGAC

5903 ATGATGCATGGTTACTCACCACTGCGATCCCCGGGAAAACAGCATTCCAGGTATTAGAAGAATATCCTGATTCAGGT
     TACTACGTACCAATGAGTGGTGACGCTAGGGGCCCTTTTGTCGTAAGGTCCATAATCTTCTTATAGGACTAAGTCCA

5980 GAAAATATTGTTGATGCGCTGGCAGTGTTCCTGCGCCGGTTGCATTCGATTCCTGTTTGTAATTGTCCTTTTAACAG
     CTTTTATAACAACTACGCGACCGTCACAAGGACGCGGCCAACGTAAGCTAAGGACAAACATTAACAGGAAAATTGTC

6057 CGATCGCGTATTTCGTCTCGCTCAGGCGCAATCACGAATGAATAACGGTTTGGTTGATGCGAGTGATTTTGATGACG
     GCTAGCGCATAAAGCAGAGCGAGTCCGCGTTAGTGCTTACTTATTGCCAAACCAACTACGCTCACTAAAACTACTGC
```

```
6134 AGCGTAATGGCTGGCCTGTTGAACAAGTCTGGAAAGAAATGCATAAACTTTTGCCATTCTCACCGGATTCAGTCGTC
     TCGCATTACCGACCGGACAACTTGTTCAGACCTTTCTTTACGTATTTGAAAACGGTAAGAGTGGCCTAAGTCAGCAG

6211 ACTCATGGTGATTTCTCACTTGATAACCTTATTTTTGACGAGGGGAAATTAATAGGTTGTATTGATGTTGGACGAGT
     TGAGTACCACTAAAGAGTGAACTATTGGAATAAAAACTGCTCCCCTTTAATTATCCAACATAACTACAACCTGCTCA

6288 CGGAATCGCAGACCGATACCAGGATCTTGCCATCCTATGGAACTGCCTCGGTGAGTTTTCTCCTTCATTACAGAAAC
     GCCTTAGCGTCTGGCTATGGTCCTAGAACGGTAGGATACCTTGACGGAGCCACTCAAAAGAGGAAGTAATGTCTTTG

6365 GGCTTTTTCAAAAATATGGTATTGATAATCCTGATATGAATAAATTGCAGTTTCATTTGATGCTCGATGAGTTTTTC
     CCGAAAAAGTTTTTATACCATAACTATTAGGACTATACTTATTTAACGTCAAAGTAAACTACGAGCTACTCAAAAAG

6442 TAAGAATTAATTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTC
     ATTCTTAATTAAGTACTCGCCTATGTATAAACTTACATAAATCTTTTTATTTGTTTATCCCCAAGGCGCGTGTAAAG

6519 CCCGAAAAGTGCCACCTGAAATTGTAAACGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTCA
     GGGCTTTTCACGGTGGACTTTAACATTTGCAATTATAAAACAATTTTAAGCGCAATTTAAAAACAATTTAGTCGAGT

6596 TTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGAGATAGGGTTGAGTGTTGT
     AAAAAATTGGTTATCCGGCTTTAGCCGTTTTAGGGAATATTTAGTTTTCTTATCTGGCTCTATCCCAACTCACAACA

6673 TCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTCCAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCG
     AGGTCAAACCTTGTTCTCAGGTGATAATTTCTTGCACCTGAGGTTGCAGTTTCCCGCTTTTTGGCAGATAGTCCCGC

6750 ATGGCCCACTACGTGAACCATCACCCTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCT
     TACCGGGTGATGCACTTGGTAGTGGGATTAGTTCAAAAAACCCCAGCTCCACGGCATTTCGTGATTTAGCCTTGGGA

6827 AAAGGGAGCCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAAAGCGAAAGG
     TTTCCCTCGGGGGCTAAATCTCGAACTGCCCCTTTCGGCCGCTTGCACCGCTCTTTCCTTCCCTTCTTTCGCTTTCC

6904 AGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCACCACACCCGCCGCGCTTAATGCGCCGC
     TCGCCCGCGATCCCGCGACCGTTCACATCGCCAGTGCGACGCGCATTGGTGGTGTGGGCGGCGCGAATTACGCGGCG

6981 TACAGGGCGCGTCCCATTCGCCA
     ATGTCCCGCGCAGGGTAAGCGGT
```

A.2 pEC2-SC5, -RGD, -RDG plasmid information

Submitted by Marissa Mock **Date 4/22/04**

Strain name/ **XL-1 blue (***SupE*44 *hsdR*17 *recA*1 *endA*1 *gyrA*46 *thi relA*1 *lac*$^-$ F'[*proAB*$^+$ *lacI*^{q} l*acZ*ΔM15 Tn*10*(*tet*^{r})] (From Stratagene) **)/ pEC2-SC5**
pEC2-RGD
pEC2-RDG

Vector (kb) : **pEC2-SC5**
pEC2-RGD
pEC2-RDG

Cloning site: See plasmid map.

Construction of pEC2-CBD plasmids

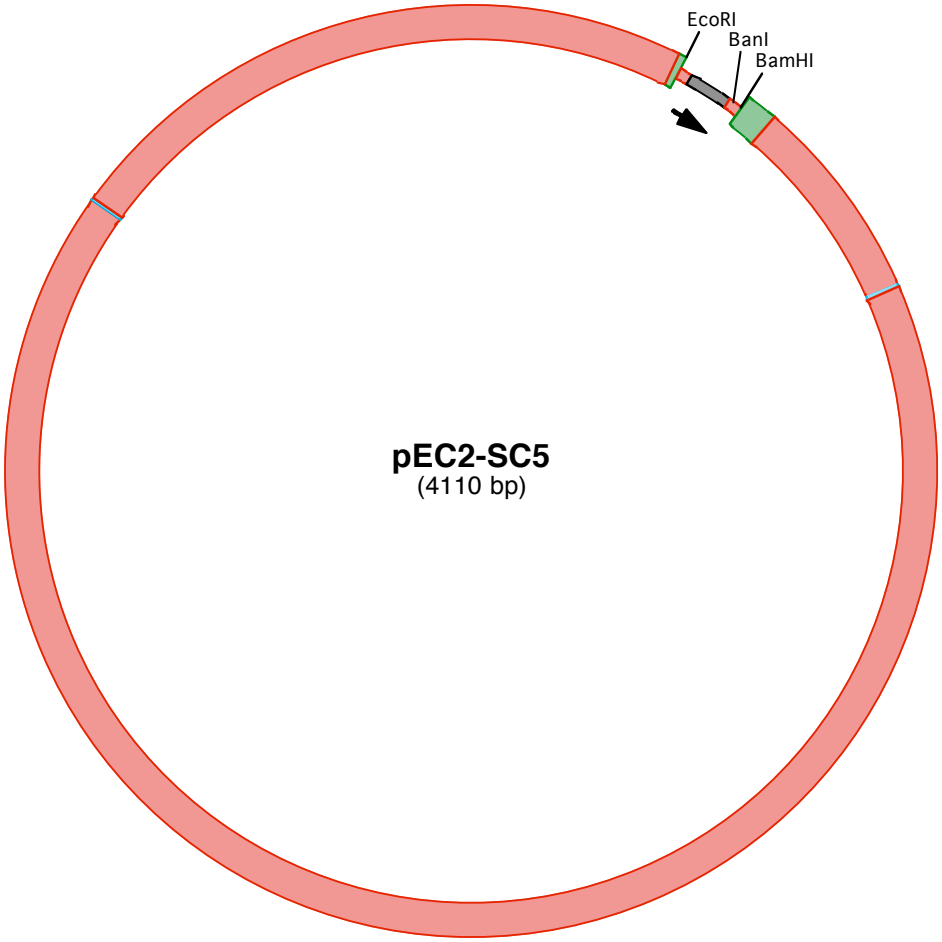
The cell binding domain region was cut out of the appropriate pUC19-SC5, pUC19-RGD, or pUC19-RDG vector with EcoR I/ BamH I digestion. The pEC2 (Eric Cantor) vector was cut with EcoR I/BamH I and ligated with the inserts to produce pEC2-SC5, pEC2-RGD, and pEC2-RDG.

Plasmid map and full sequence are shown only for pEC2-SC5; others are similar.

Source available :

12 % Glycerol culture in Marissa freezer box
Culture conditions: 2xYT, 37°C

A.2.1 Plasmid map



A.2.2 Position of elements

	bp
Vector size	4110
SC5 insert	314-420
RGD insert	314-414
RDG insert	314-414

A.2.3 Full sequence (pEC2-SC5)

```
   1 GTTGACGCCGGGCAAGAGCAACTCGGTCGCCGCATACACTATTCTCAGAATGACTTGGTTGAGTACTCACCAGTCACAGAAA
     CAACTGCGGCCCGTTCTCGTTGAGCCAGCGGCGTATGTGATAAGAGTCTTACTGAACCAACTCATGAGTGGTCAGTGTCTTT

  83 AGCATCTTACGGATGGCATGACAGTAAGAGAATTATGCAGTGCTGCCATAACCATGAGTGATAACACTGCGGCCAACTTACT
     TCGTAGAATGCCTACCGTACTGTCATTCTCTTAATACGTCACGACGGTATTGGTACTCACTATTGTGACGCCGGTTGAATGA

 165 TCTGACAACGATCGGAGGACCGAAGGAGCTAACCGCTTTTTTGCACAACATGGGGGATCATGTAACTCGCCTTGATCGTTGG
     AGACTGTTGCTAGCCTCCTGGCTTCCTCGATTGGCGAAAAAACGTGTTGTACCCCCTAGTACATTGAGCGGAACTAGCAACC

                                                                 PstI            EcoRI
 247 GAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTGACACCACGATGCCTGCAgatctgatcag aat tcg tcg
     CTTGGCCTCGACTTACTTCGGTATGGTTTGCTGCTCGCACTGTGGTGCTACGGACGTctagactagtc tta agc agc

       NheI
 324 acg cta gct tcc tcg atg gtg aag aga tcc aga tcg gcc aca tcc cgc gtg aag ttg atg
     tgc gat cga agg agc tac cac ttc tct agg tct agc cgg tgt agg gcg cac ttc aac tac
               1▶  L   D   G   E   E   I   Q   I   G   H   I   P   R   E   V   D

                         NheI         BanI       XhoI      BamHI
 384 att acc acc tgt acg cta gcg cgg tgc cgc tcg agg gatccatctagagtcgacgtcggccgttaaccta
     taa tgg tgg aca tgc gat cgc gcc acg gcg agc tcc ctaggtagatctcagctgcagccggcaattggat
  17▶D   Y   H   L   Y

           PstI
 454 ggagatctgcaGCAATGGCAACAACGTTGCGCAAACTATTAACTGGCGAACTACTTACTCTAGCTTCCCGGCAACAATTAAT
     cctctagacgtCGTTACCGTTGTTGCAACGCGTTTGATAATTGACCGCTTGATGAATGAGATCGAAGGGCCGTTGTTAATTA

 536 AGACTGGATGGAGGCGGATAAAGTTGCAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAATCT
     TCTGACCTACCTCCGCCTATTTCAACGTCCTGGTGAAGACGCGAGCCGGGAAGGCCGACCGACCAAATAACGACTATTTAGA

 618 GGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACA
     CCTCGGCCACTCGCACCCAGAGCGCCATAGTAACGTCGTGACCCCGGTCTACCATTCGGGAGGGCATAGCATCAATAGATGT

 700 CGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCgtgcCTCACTGATTAAGCATTGGT
     GCTGCCCCTCAGTCCGTTGATACCTACTTGCTTTATCTGTCTAGCGACTCTATCCACGcacgGAGTGACTAATTCGTAACCA

 782 AACTGTCAGACCAAGTTTACTCATATATACTTTAGATTGATTTAAAACTTCATTTTTAATTTAAAAGGATCTAGGTGAAGAT
     TTGACAGTCTGGTTCAAATGAGTATATATGAAATCTAACTAAATTTTGAAGTAAAAATTAAATTTTCCTAGATCCACTTCTA

 864 CCTTTTTGATAATCTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCTTAATAAGATGATCT
     GGAAAAACTATTAGAGTACTGGTTTTAGGGAATTGCACTCAAAAGCAAGGTGACTCGCAGTCTGGGGAATTATTCTACTAGA

 946 TCTTGAGATCGTTTTGGTCTGCGCGTAATCTCTTGCTCTGAAAACGAAAAAACCGCCTTGCAGGGCGGTTTTTCGAAGGTTC
     AGAACTCTAGCAAAACCAGACGCGCATTAGAGAACGAGACTTTTGCTTTTTTGGCGGAACGTCCCGCCAAAAAGCTTCCAAG

1028 TCTGAGCTACCAACTCTTTGAACCGAGGTAACTGGCTTGGAGGAGCGCAGTCACCAAAACTTGTCCTTTCAGTTTAGCCTTA
     AGACTCGATGGTTGAGAAACTTGGCTCCATTGACCGAACCTCCTCGCGTCAGTGGTTTTGAACAGGAAAGTCAAATCGGAAT

1110 ACCGGCGCATGACTTCAAGACTAACTCCTCTAAATCAATTACCAGTGGCTGCTGCCAGTGGTGCTTTTGCATGTCTTTCCGG
     TGGCCGCGTACTGAAGTTCTGATTGAGGAGATTTAGTTAATGGTCACCGACGACGGTCACCACGAAAACGTACAGAAAGGCC

1192 GTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTCGGACTGAACGGGGGGTTCGTGCATACAGTCCAGCTTGGAG
     CAACCTGAGTTCTGCTATCAATGGCCTATTCCGCGTCGCCAGCCTGACTTGCCCCCAAGCACGTATGTCAGGTCGAACCTC

1274 CGAACTGCCTACCCGGAACTGAGTGTCAGGCGTGGAATGAGACAAACGCGGCCATAACAGCGGAATGACACCGGTAAACCGA
     GCTTGACGGATGGGCCTTGACTCACAGTCCGCACCTTACTCTGTTTGCGCCGGTATTGTCGCCTTACTGTGGCCATTTGGCT
```

```
1356 AAGGCAGGAACAGGAGAGCGCACGAGGGAGCCGCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTCGGGTTTCGCCACC
     TTCCGTCCTTGTCCTCTCGCGTGCTCCCTCGGCGGTCCCCCTTTGCGGACCATAGAAATATCAGGACAGCCCAAAGCGGTGG

1438 ACTGATTTGAGCGTCAGATTTCGTGATGCTTGTCAGGGGGGCGGAGCCTATGGAAAAACGGCTTTGCCGCGGCCCTCTCACT
     TGACTAAACTCGCAGTCTAAAGCACTACGAACAGTCCCCCCGCCTCGGATACCTTTTTGCCGAAACGGCGCCGGGAGAGTGA

1520 TCCCTGTTAAGTATCTTCCTGGCATCTTCCAGGAAATCTCCGCCCCGTTCGTAAGCCATTTCCGCTCGCCGCAGTCGAACGA
     AGGGACAATTCATAGAAGGACCGTAGAAGGTCCTTTAGAGGCGGGGCAAGCATTCGGTAAAGGCGAGCGGCGTCAGCTTGCT

1602 CCGAGCGTAGCGAGTCAGTGAGCGAGGAAGCGGAATATATCCTGTATCACATATTCTGCTGACGCACCGGTGCAGCCTTTTT
     GGCTCGCATCGCTCAGTCACTCGCTCCTTCGCCTTATATAGGACATAGTGTATAAGACGACTGCGTGGCCACGTCGGAAAAA

1684 TCTCCTGCCACATGAAGCACTTCACTGACACCCTCATCAGTGCCAACATAGTAAGCCAGTATACACTCCGCTAGCGCTGAGG
     AGAGGACGGTGTACTTCGTGAAGTGACTGTGGGAGTAGTCACGGTTGTATCATTCGGTCATATGTGAGGCGATCGCGACTCC

1766 TCTGCCTCGTGAAGAAGGTGTTGCTGACTCATACCAGGCCTGAATCGCCCCATCATCCAGCCAGAAAGTGAGGGAGCCACGG
     AGACGGAGCACTTCTTCCACAACGACTGAGTATGGTCCGGACTTAGCGGGGTAGTAGGTCGGTCTTTCACTCCCTCGGTGCC

1848 TTGATGAGAGCTTTGTTGTAGGTGGACCAGTTGGTGATTTTGAACTTTTGCTTTGCCACGGAACGGTCTGCGTTGTCGGGAA
     AACTACTCTCGAAACAACATCCACCTGGTCAACCACTAAAACTTGAAAACGAAACGGTGCCTTGCCAGACGCAACAGCCCTT

1930 GATGCGTGATCTGATCCTTCAACTCAGCAAAAGTTCGATTTATTCAACAAAGCCACGTTGTGTCTCAAAATCTCTGATGTTA
     CTACGCACTAGACTAGGAAGTTGAGTCGTTTTCAAGCTAAATAAGTTGTTTCGGTGCAACACAGAGTTTTAGAGACTACAAT

2012 CATTGCACAAGATAAAAATATATCATCATGAACAATAAAACTGTCTGCTTACATAAACAGTAATACAAGGGGTGTTATGAGC
     GTAACGTGTTCTATTTTTATATAGTAGTACTTGTTATTTTGACAGACGAATGTATTTGTCATTATGTTCCCCACAATACTCG

2094 CATATTCAACGGGAAACGTCTTGCTCGAGGCCGCGATTAAATTCCAACATGGATGCTGATTTATATGGGTATAAATGGGCTC
     GTATAAGTTGCCCTTTGCAGAACGAGCTCCGGCGCTAATTTAAGGTTGTACCTACGACTAAATATACCCATATTTACCCGAG

2176 GCGATAATGTCGGGCAATCAGGTGCGACAATCTATCGATTGTATGGGAAGCCCGATGCGCCAGAGTTGTTTCTGAAACATGG
     CGCTATTACAGCCCGTTAGTCCACGCTGTTAGATAGCTAACATACCCTTCGGGCTACGCGGTCTCAACAAAGACTTTGTACC

2258 CAAAGGTAGCGTTGCCAATGATGTTACAGATGAGATGGTCAGACTAAACTGGCTGACGGAATTTATGCCTCTTCCGACCATC
     GTTTCCATCGCAACGGTTACTACAATGTCTACTCTACCAGTCTGATTTGACCGACTGCCTTAAATACGGAGAAGGCTGGTAG

2340 AAGCATTTTATCCGTACTCCTGATGATGCATGGTTACTCACCACTGCGATCCCCGGGAAAACAGCATTCCAGGTATTAGAAG
     TTCGTAAAATAGGCATGAGGACTACTACGTACCAATGAGTGGTGACGCTAGGGGCCCTTTTGTCGTAAGGTCCATAATCTTC

2422 AATATCCTGATTCAGGTGAAAATATTGTTGATGCGCTGGCAGTGTTCCTGCGCCGGTTGCATTCGATTCCTGTTTGTAATTG
     TTATAGGACTAAGTCCACTTTTATAACAACTACGCGACCGTCACAAGGACGCGGCCAACGTAAGCTAAGGACAAACATTAAC

2504 TCCTTTTAACAGCGATCGCGTATTTCGTCTCGCTCAGGCGCAATCACGAATGAATAACGGTTTGGTTGATGCGAGTGATTTT
     AGGAAAATTGTCGCTAGCGCATAAAGCAGAGCGAGTCCGCGTTAGTGCTTACTTATTGCCAAACCAACTACGCTCACTAAAA

2586 GATGACGAGCGTAATGGCTGGCCTGTTGAACAAGTCTGGAAAGAAATGCATAAGCTTTTGCCATTCTCACCGGATTCAGTCG
     CTACTGCTCGCATTACCGACCGGACAACTTGTTCAGACCTTTCTTTACGTATTCGAAAACGGTAAGAGTGGCCTAAGTCAGC

2668 TCACTCATGGTGATTTCTCACTTGATAACCTTATTTTTGACGAGGGGAAATTAATAGGTTGTATTGATGTTGGACGAGTCGG
     AGTGAGTACCACTAAAGAGTGAACTATTGGAATAAAAACTGCTCCCCTTTAATTATCCAACATAACTACAACCTGCTCAGCC

2750 AATCGCAGACCGATACCAGGATCTTGCCATCCTATGGAACTGCCTCGGTGAGTTTTCTCCTTCATTACAGAAACGGCTTTTT
     TTAGCGTCTGGCTATGGTCCTAGAACGGTAGGATACCTTGACGGAGCCACTCAAAAGAGGAAGTAATGTCTTTGCCGAAAAA

2832 CAAAAATATGGTATTGATAATCCTGATATGAATAAATTGCAGTTTCATTTGATGCTCGATGAGTTTTTCTAATCAGAATTGG
     GTTTTTATACCATAACTATTAGGACTATACTTATTTAACGTCAAAGTAAACTACGAGCTACTCAAAAAGATTAGTCTTAACC
```

```
2914 TTAATTGGTTGTAACACTGGCAGAGCATTACGCTGACTTGACGGGACGGCGGCTTTGTTGAATAAATCGAACTTTTGCTGAG
     AATTAACCAACATTGTGACCGTCTCGTAATGCGACTGAACTGCCCTGCCGCCGAAACAACTTATTTAGCTTGAAAACGACTC

2996 TTGAAGGATCAGATCACGCATCTTCCCGACAACGCAGACCGTTCCGTGGCAAAGCAAAAGTTCAAAATCACCAACTGGTCCA
     AACTTCCTAGTCTAGTGCGTAGAAGGGCTGTTGCGTCTGGCAAGGCACCGTTTCGTTTTCAAGTTTTAGTGGTTGACCAGGT

3078 CCTACAACAAAGCTCTCATCAACCGTGGCTCCCTCACTTTCTGGCTGGATGATGGGGCGATTCAGGCCTGGTATGAGTCAGC
     GGATGTTGTTTCGAGAGTAGTTGGCACCGAGGGAGTGAAAGACCGACCTACTACCCCGCTAAGTCCGGACCATACTCAGTCG

3160 AACACCTTCTTCACGAGGCAGACCTCAGCGCTCAAAGATGCAGGGGTAAAAGCTAACCGCATCTTTACCGACAAGGCATCCG
     TTGTGGAAGAAGTGCTCCGTCTGGAGTCGCGAGTTTCTACGTCCCCATTTTCGATTGGCGTAGAAATGGCTGTTCCGTAGGC

3242 GCAGTTCAACAGATCGGGAAGGGCTGGATTTGCTGAGGATGAAGGTGGAGGAAGGTGATGTCATTCTGGTGAAGAAGCTCGA
     CGTCAAGTTGTCTAGCCCTTCCCGACCTAAACGACTCCTACTTCCACCTCCTTCCACTACAGTAAGACCACTTCTTCGAGCT

3324 CCGTCTTGGCCGCGACACCGCCGACATGATCCAACTGATAAAAGAGTTTGATGCTCAGGGTGTAGCGGTTCGGTTTATTGAC
     GGCAGAACCGGCGCTGTGGCGGCTGTACTAGGTTGACTATTTTCTCAAACTACGAGTCCCACATCGCCAAGCCAAATAACTG

3406 GACGGGATCAGTACCGACGGTGATATGGGGCAAATGGTGGTCACCATCCTGTCGGCTGTGGCACAGGCTGAACGCCGGAGGA
     CTGCCCTAGTCATGGCTGCCACTATACCCCGTTTACCACCAGTGGTAGGACAGCCGACACCGTGTCCGACTTGCGGCCTCCT

3488 TCgatcCTAGAGCGCACGAATGAGGGCCGACAGGAAGCAAAGCTGAAAGGAATCAAATTTGGCCGCAGGCGTACCGTGGACA
     AGctagGATCTCGCGTGCTTACTCCCGGCTGTCCTTCGTTTCGACTTTCCTTAGTTTAAACCGGCGTCCGCATGGCACCTGT

3570 GGAACGTCGTGCTGACGCTTCATCAGAAGGGCACTGGTGCAACGGAAATTGCTCATCAGCTCAGTATTGCCCGCTCCACGGT
     CCTTGCAGCACGACTGCGAAGTAGTCTTCCCGTGACCACGTTGCCTTTAACGAGTAGTCGAGTCATAACGGGCGAGGTGCCA

3652 TTATAAAATTCTTGAAGACGAAAGGGCCTCGTGATACGCCTATTTTTATAGGTTAATGTCATGATAATAATGGTTTCTTAGA
     AATATTTTAAGAACTTCTGCTTTCCCGGAGCACTATGCGGATAAAAATATCCAATTACAGTACTATTATTACCAAAGAATCT

3734 CGTCAGGTGGCACTTTTCGGGGAAATGTGCGCGGAACCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCT
     GCAGTCCACCGTGAAAAGCCCCTTTACACGCGCCTTGGGGATAAACAAATAAAAAGATTTATGTAAGTTTATACATAGGCGA

3816 CATGAGACAATAACCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGAGTATTCAACATTTCCGTGTCGCCCTT
     GTACTCTGTTATTGGGACTATTTACGAAGTTATTATAACTTTTTCCTTCTCATACTCATAAGTTGTAAAGGCACAGCGGGAA

3898 ATTCCCTTTTTTGCGGCATTTTGCCTTCCTGTTTTTGCTCACCCAGAAACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGT
     TAAGGGAAAAAACGCCGTAAAACGGAAGGACAAAAACGAGTGGGTCTTTGCGACCACTTTCATTTTCTACGACTTCTAGTCA

3980 TGGGTGCACGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTCC
     ACCCACGTGCTCACCCAATGTAGCTTGACCTAGAGTTGTCGCCATTCTAGGAACTCTCAAAAGCGGGGCTTCTTGCAAAAGG

4062 AATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGT
     TTACTACTCGTGAAAATTTCAAGACGATACACCGCGCCATAATAGGGCA
```

A.3 pEC2-SC5-ELF$_5$ and -ELF$_6$ information

Submitted by Marissa Mock **Date 4/22/04**

Strain name: **JM109** ***(****e14–(McrA–) recA1 endA1 gyrA96 thi-1 hsdR17 (rK– mK+) supE44 relA1 £G(lac-proAB) [F′ traD36 proAB lacIqZ£GM15)*. (From Zymo Research) **)/ pEC2-SC5-ELF$_5$**
or / pEC2-SC5-ELF$_6$

Vector (kb): **pEC2-SC5-ELF$_5$ (X)**
pEC2-SC5-ELF$_6$ (X)

Construction of pEC2-CBD plasmids

ELF monomer (see sequence below) was obtained through Ban I digestion of pUC19-ELF (Nandita Sharma).

pEC2-SC5-ELF$_5$:
The 75 bp fragment was multimerized through ligation with T4 DNA ligase for 3 minutes at 0°C. The ligation mixture was run on a 2% agarose gel, and the band corresponding to pentamer (375 bp) was cut out and extracted from the gel. The pentamer DNA was ligated with pEC2-SC5 vector that had been digested with Ban I and dephosphorylated with CIP, to yield pEC2-SC5-ELF$_5$.

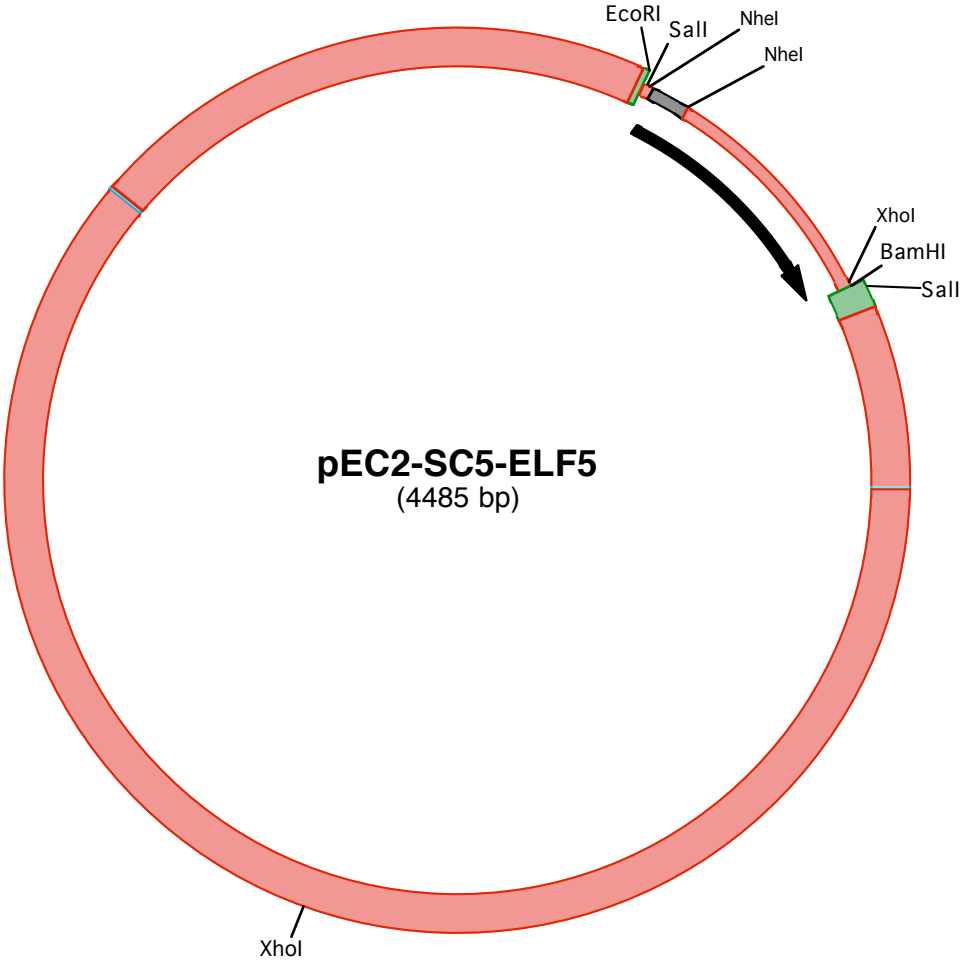
pEC2-SC5-ELF$_6$:
The 75 bp ELF fragment was mixed with T4 ligase at 0°C for 30 s and added directly to pEC2-SC5 vector that had been digested with Ban I and dephosphorylated with CIP. The entire ligation mixture was transformed in JM109 competent cells; screened colonies revealed a strain containing pEC2-SC5-ELF$_6$.

ELF monomer sequence:
ggtgccgggtgtgggcgttccgggcgtgggtgtaccgggcttcggtgtcccgggcgtaggtgttccgggtgtcggggtgcc

Plasmid map and full sequence shown only for pEC2-SC5-ELF$_5$; other is similar.

Source available :
12 % Glycerol culture in Marissa freezer box
Culture conditions: 2xYT, 37°C

A.3.1 Plasmid map



A.3.2 Position of elements

	bp
Vector size	4485
SC5 cell binding domain	314-406
ELF pentamer	407-782

A.3.3 Full sequence (pEC2-SC5-ELF$_5$)

```
   1 GTTGACGCCGGGCAAGAGCAACTCGGTCGCCGCATACACTATTCTCAGAATGACTTGGTTGAGTACTCACCAGTCACA
     CAACTGCGGCCCGTTCTCGTTGAGCCAGCGGCGTATGTGATAAGAGTCTTACTGAACCAACTCATGAGTGGTCAGTGT

  79 GAAAAGCATCTTACGGATGGCATGACAGTAAGAGAATTATGCAGTGCTGCCATAACCATGAGTGATAACACTGCGGCC
     CTTTTCGTAGAATGCCTACCGTACTGTCATTCTCTTAATACGTCACGACGGTATTGGTACTCACTATTGTGACGCCGG

 157 AACTTACTTCTGACAACGATCGGAGGACCGAAGGAGCTAACCGCTTTTTTGCACAACATGGGGGATCATGTAACTCGC
     TTGAATGAAGACTGTTGCTAGCCTCCTGGCTTCCTCGATTGGCGAAAAAACGTGTTGTACCCCCTAGTACATTGAGCG

 235 CTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTGACACCACGATGCCTGCAgatctgatc
     GAACTAGCAACCCTTGGCCTCGACTTACTTCGGTATGGTTTGCTGCTCGCACTGTGGTGCTACGGACGTctagactag

          EcoRI            NheI
 313 ag aat tcg tcg acg cta gct tcc tcg atg gtg aag aga tcc aga tcg gcc aca tcc cgc
     tc tta agc agc tgc gat cga agg agc tac cac ttc tct agg tct agc cgg tgt agg gcg
                              1▶ L  D  G  E  E  I  Q  I  G  H  I  P

                                        NheI
 372 gtg aag ttg atg att acc acc tgt acg cta gcg cg gtgccgggtgtgggcgttccgggcgtgggtg
     cac ttc aac tac taa tgg tgg aca tgc gat cgc gc cacggcccacacccgcaaggcccgcacccac
  13▶R  E  V  D  D  Y  H  L  Y  A  S  A  V  P  G  V  G  V  P  G  V  G

 438 taccgggcttcggtgtcccgggcgtaggtgttccgggtgtcggggtgccgggtgtgggcgttccgggcgtgggtgtac
     atggcccgaagccacagggcccgcatccacaaggcccacagccccacggcccacacccgcaaggcccgcacccacatg
  35▶V  P  G  F  G  V  P  G  V  G  V  P  G  V  G  V  P  G  V  G  V  P  G  V  G  V

 516 cgggcttcggtgtcccgggcgtaggtgttccgggtgtcggggtgccgggtgtgggcgttccgggcgtgggtgtaccgg
     gcccgaagccacagggcccgcatccacaaggcccacagccccacggcccacacccgcaaggcccgcacccacatggcc
  61▶P  G  F  G  V  P  G  V  G  V  P  G  V  G  V  P  G  V  G  V  P  G  V  G  V  P

 594 gcttcggtgtcccgggcgtaggtgttccgggtgtcggggtgccgggtgtgggcgttccgggcgtgggtgtaccgggct
     cgaagccacagggcccgcatccacaaggcccacagccccacggcccacacccgcaaggcccgcacccacatggcccga
  87▶G  F  G  V  P  G  V  G  V  P  G  V  G  V  P  G  V  G  V  P  G  V  G  V  P  G

 672 tcggtgtcccgggcgtaggtgttccgggtgtcggggtgccgggtgtgggcgttccgggcgtgggtgtaccgggcttcg
     agccacagggcccgcatccacaaggcccacagccccacggcccacacccgcaaggcccgcacccacatggcccgaagc
 113▶F  G  V  P  G  V  G  V  P  G  V  G  V  P  G  V  G  V  P  G  V  G  V  P  G  F

                                                  XhoI      BamHI
 750 gtgtcccgggcgtaggtgttccgggtgtcggg gtg ccg ctc gag g gatccatctagagtcgacgtcggccgt
     cacagggcccgcatccacaaggcccacagccc cac ggc gag ctc c ctaggtagatctcagctgcagccggca
 139▶G  V  P  G  V  G  V  P  G  V  G

 822 taacctaggagatctgcaGCAATGGCAACAACGTTGCGCAAACTATTAACTGGCGAACTACTTACTCTAGCTTCCCGG
     attggatcctctagacgtCGTTACCGTTGTTGCAACGCGTTTGATAATTGACCGCTTGATGAATGAGATCGAAGGGCC

 900 CAACAATTAATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTT
     GTTGTTAATTATCTGACCTACCTCCGCCTATTTCAACGTCCTGGTGAAGACGCGAGCCGGGAAGGCCGACCGACCAAA

 978 ATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCC
     TAACGACTATTTAGACCTCGGCCACTCGCACCCAGAGCGCCATAGTAACGTCGTGACCCCGGTCTACCATTCGGGAGG

1056 CGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCg
     GCATAGCATCAATAGATGTGCTGCCCCTCAGTCCGTTGATACCTACTTGCTTTATCTGTCTAGCGACTCTATCCACGc
```

```
1134 tgcCTCACTGATTAAGCATTGGTAACTGTCAGACCAAGTTTACTCATATATACTTTAGATTGATTTAAAACTTCATTT
     acgGAGTGACTAATTCGTAACCATTGACAGTCTGGTTCAAATGAGTATATATGAAATCTAACTAAATTTTGAAGTAAA

1212 TTAATTTAAAAGGATCTAGGTGAAGATCCTTTTTGATAATCTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCA
     AATTAAATTTTCCTAGATCCACTTCTAGGAAAAACTATTAGAGTACTGGTTTTAGGGAATTGCACTCAAAAGCAAGGT

1290 CTGAGCGTCAGACCCCTTAATAAGATGATCTTCTTGAGATCGTTTTGGTCTGCGCGTAATCTCTTGCTCTGAAAACGA
     GACTCGCAGTCTGGGGAATTATTCTACTAGAAGAACTCTAGCAAAACCAGACGCGCATTAGAGAACGAGACTTTTGCT

1368 AAAAACCGCCTTGCAGGGCGGTTTTTCGAAGGTTCTCTGAGCTACCAACTCTTTGAACCGAGGTAACTGGCTTGGAGG
     TTTTTGGCGGAACGTCCCGCCAAAAAGCTTCCAAGAGACTCGATGGTTGAGAAACTTGGCTCCATTGACCGAACCTCC

1446 AGCGCAGTCACCAAAACTTGTCCTTTCAGTTTAGCCTTAACCGGCGCATGACTTCAAGACTAACTCCTCTAAATCAAT
     TCGCGTCAGTGGTTTTGAACAGGAAAGTCAAATCGGAATTGGCCGCGTACTGAAGTTCTGATTGAGGAGATTTAGTTA

1524 TACCAGTGGCTGCTGCCAGTGGTGCTTTTGCATGTCTTTCCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGC
     ATGGTCACCGACGACGGTCACCACGAAAACGTACAGAAAGGCCCAACCTGAGTTCTGCTATCAATGGCCTATTCCGCG

1602 AGCGGTCGGACTGAACGGGGGGTTCGTGCATACAGTCCAGCTTGGAGCGAACTGCCTACCCGGAACTGAGTGTCAGGC
     TCGCCAGCCTGACTTGCCCCCCAAGCACGTATGTCAGGTCGAACCTCGCTTGACGGATGGGCCTTGACTCACAGTCCG

1680 GTGGAATGAGACAAACGCGGCCATAACAGCGGAATGACACCGGTAAACCGAAAGGCAGGAACAGGAGAGCGCACGAGG
     CACCTTACTCTGTTTGCGCCGGTATTGTCGCCTTACTGTGGCCATTTGGCTTTCCGTCCTTGTCCTCTCGCGTGCTCC

1758 GAGCCGCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTCGGGTTTCGCCACCACTGATTTGAGCGTCAGATTTCG
     CTCGGCGGTCCCCCTTTGCGGACCATAGAAATATCAGGACAGCCCAAAGCGGTGGTGACTAAACTCGCAGTCTAAAGC

1836 TGATGCTTGTCAGGGGGGCGGAGCCTATGGAAAAACGGCTTTGCCGCGGCCCTCTCACTTCCCTGTTAAGTATCTTCC
     ACTACGAACAGTCCCCCCGCCTCGGATACCTTTTTGCCGAAACGGCGCCGGGAGAGTGAAGGGACAATTCATAGAAGG

1914 TGGCATCTTCCAGGAAATCTCCGCCCCGTTCGTAAGCCATTTCCGCTCGCCGCAGTCGAACGACCGAGCGTAGCGAGT
     ACCGTAGAAGGTCCTTTAGAGGCGGGGCAAGCATTCGGTAAAGGCGAGCGGCGTCAGCTTGCTGGCTCGCATCGCTCA

1992 CAGTGAGCGAGGAAGCGGAATATATCCTGTATCACATATTCTGCTGACGCACCGGTGCAGCCTTTTTTCTCCTGCCAC
     GTCACTCGCTCCTTCGCCTTATATAGGACATAGTGTATAAGACGACTGCGTGGCCACGTCGGAAAAAAGAGGACGGTG

2070 ATGAAGCACTTCACTGACACCCTCATCAGTGCCAACATAGTAAGCCAGTATACACTCCGCTAGCGCTGAGGTCTGCCT
     TACTTCGTGAAGTGACTGTGGGAGTAGTCACGGTTGTATCATTCGGTCATATGTGAGGCGATCGCGACTCCAGACGGA

2148 CGTGAAGAAGGTGTTGCTGACTCATACCAGGCCTGAATCGCCCCATCATCCAGCCAGAAAGTGAGGGAGCCACGGTTG
     GCACTTCTTCCACAACGACTGAGTATGGTCCGGACTTAGCGGGGTAGTAGGTCGGTCTTTCACTCCCTCGGTGCCAAC

2226 ATGAGAGCTTTGTTGTAGGTGGACCAGTTGGTGATTTTGAACTTTTGCTTTGCCACGGAACGGTCTGCGTTGTCGGGA
     TACTCTCGAAACAACATCCACCTGGTCAACCACTAAAACTTGAAAACGAAACGGTGCCTTGCCAGACGCAACAGCCCT

2304 AGATGCGTGATCTGATCCTTCAACTCAGCAAAAGTTCGATTTATTCAACAAAGCCACGTTGTGTCTCAAAATCTCTGA
     TCTACGCACTAGACTAGGAAGTTGAGTCGTTTTCAAGCTAAATAAGTTGTTTCGGTGCAACACAGAGTTTTAGAGACT

2382 TGTTACATTGCACAAGATAAAAATATATCATCATGAACAATAAAACTGTCTGCTTACATAAACAGTAATACAAGGGGT
     ACAATGTAACGTGTTCTATTTTTATATAGTAGTACTTGTTATTTTGACAGACGAATGTATTTGTCATTATGTTCCCCA

2460 GTTATGAGCCATATTCAACGGGAAACGTCTTGCTCGAGGCCGCGATTAAATTCCAACATGGATGCTGATTTATATGGG
     CAATACTCGGTATAAGTTGCCCTTTGCAGAACGAGCTCCGGCGCTAATTTAAGGTTGTACCTACGACTAAATATACCC

2538 TATAAATGGGCTCGCGATAATGTCGGGCAATCAGGTGCGACAATCTATCGATTGTATGGGAAGCCCGATGCGCCAGAG
     ATATTTACCCGAGCGCTATTACAGCCCGTTAGTCCACGCTGTTAGATAGCTAACATACCCTTCGGGCTACGCGGTCTC

2616 TTGTTTCTGAAACATGGCAAAGGTAGCGTTGCCAATGATGTTACAGATGAGATGGTCAGACTAAACTGGCTGACGGAA
     AACAAAGACTTTGTACCGTTTCCATCGCAACGGTTACTACAATGTCTACTCTACCAGTCTGATTTGACCGACTGCCTT
```

```
2694 TTTATGCCTCTTCCGACCATCAAGCATTTTATCCGTACTCCTGATGATGCATGGTTACTCACCACTGCGATCCCCGGG
     AAATACGGAGAAGGCTGGTAGTTCGTAAAATAGGCATGAGGACTACTACGTACCAATGAGTGGTGACGCTAGGGGCCC

2772 AAAACAGCATTCCAGGTATTAGAAGAATATCCTGATTCAGGTGAAAATATTGTTGATGCGCTGGCAGTGTTCCTGCGC
     TTTTGTCGTAAGGTCCATAATCTTCTTATAGGACTAAGTCCACTTTTATAACAACTACGCGACCGTCACAAGGACGCG

2850 CGGTTGCATTCGATTCCTGTTTGTAATTGTCCTTTTAACAGCGATCGCGTATTTCGTCTCGCTCAGGCGCAATCACGA
     GCCAACGTAAGCTAAGGACAAACATTAACAGGAAAATTGTCGCTAGCGCATAAAGCAGAGCGAGTCCGCGTTAGTGCT

2928 ATGAATAACGGTTTGGTTGATGCGAGTGATTTTGATGACGAGCGTAATGGCTGGCCTGTTGAACAAGTCTGGAAAGAA
     TACTTATTGCCAAACCAACTACGCTCACTAAAACTACTGCTCGCATTACCGACCGGACAACTTGTTCAGACCTTTCTT

3006 ATGCATAAGCTTTTGCCATTCTCACCGGATTCAGTCGTCACTCATGGTGATTTCTCACTTGATAACCTTATTTTTGAC
     TACGTATTCGAAAACGGTAAGAGTGGCCTAAGTCAGCAGTGAGTACCACTAAAGAGTGAACTATTGGAATAAAAACTG

3084 GAGGGGAAATTAATAGGTTGTATTGATGTTGGACGAGTCGGAATCGCAGACCGATACCAGGATCTTGCCATCCTATGG
     CTCCCCTTTAATTATCCAACATAACTACAACCTGCTCAGCCTTAGCGTCTGGCTATGGTCCTAGAACGGTAGGATACC

3162 AACTGCCTCGGTGAGTTTTCTCCTTCATTACAGAAACGGCTTTTTCAAAAATATGGTATTGATAATCCTGATATGAAT
     TTGACGGAGCCACTCAAAAGAGGAAGTAATGTCTTTGCCGAAAAAGTTTTTATACCATAACTATTAGGACTATACTTA

3240 AAATTGCAGTTTCATTTGATGCTCGATGAGTTTTTCTAATCAGAATTGGTTAATTGGTTGTAACACTGGCAGAGCATT
     TTTAACGTCAAAGTAAACTACGAGCTACTCAAAAAGATTAGTCTTAACCAATTAACCAACATTGTGACCGTCTCGTAA

3318 ACGCTGACTTGACGGGACGGCGGCTTTGTTGAATAAATCGAACTTTTGCTGAGTTGAAGGATCAGATCACGCATCTTC
     TGCGACTGAACTGCCCTGCCGCCGAAACAACTTATTTAGCTTGAAAACGACTCAACTTCCTAGTCTAGTGCGTAGAAG

3396 CCGACAACGCAGACCGTTCCGTGGCAAAGCAAAAGTTCAAAATCACCAACTGGTCCACCTACAACAAAGCTCTCATCA
     GGCTGTTGCGTCTGGCAAGGCACCGTTTCGTTTTCAAGTTTTAGTGGTTGACCAGGTGGATGTTGTTTCGAGAGTAGT

3474 ACCGTGGCTCCCTCACTTTCTGGCTGGATGATGGGGCGATTCAGGCCTGGTATGAGTCAGCAACACCTTCTTCACGAG
     TGGCACCGAGGGAGTGAAAGACCGACCTACTACCCCGCTAAGTCCGGACCATACTCAGTCGTTGTGGAAGAAGTGCTC

3552 GCAGACCTCAGCGCTCAAAGATGCAGGGGTAAAAGCTAACCGCATCTTTACCGACAAGGCATCCGGCAGTTCAACAGA
     CGTCTGGAGTCGCGAGTTTCTACGTCCCCATTTTCGATTGGCGTAGAAATGGCTGTTCCGTAGGCCGTCAAGTTGTCT

3630 TCGGGAAGGGCTGGATTTGCTGAGGATGAAGGTGGAGGAAGGTGATGTCATTCTGGTGAAGAAGCTCGACCGTCTTGG
     AGCCCTTCCCGACCTAAACGACTCCTACTTCCACCTCCTTCCACTACAGTAAGACCACTTCTTCGAGCTGGCAGAACC

3708 CCGCGACACCGCCGACATGATCCAACTGATAAAAGAGTTTGATGCTCAGGGTGTAGCGGTTCGGTTTATTGACGACGG
     GGCGCTGTGGCGGCTGTACTAGGTTGACTATTTTCTCAAACTACGAGTCCCACATCGCCAAGCCAAATAACTGCTGCC

3786 GATCAGTACCGACGGTGATATGGGGCAAATGGTGGTCACCATCCTGTCGGCTGTGGCACAGGCTGAACGCCGGAGGAT
     CTAGTCATGGCTGCCACTATACCCCGTTTACCACCAGTGGTAGGACAGCCGACACCGTGTCCGACTTGCGGCCTCCTA

3864 CgatcCTAGAGCGCACGAATGAGGGCCGACAGGAAGCAAAGCTGAAAGGAATCAAATTTGGCCGCAGGCGTACCGTGG
     GctagGATCTCGCGTGCTTACTCCCGGCTGTCCTTCGTTTCGACTTTCCTTAGTTTAAACCGGCGTCCGCATGGCACC

3942 ACAGGAACGTCGTGCTGACGCTTCATCAGAAGGGCACTGGTGCAACGGAAATTGCTCATCAGCTCAGTATTGCCCGCT
     TGTCCTTGCAGCACGACTGCGAAGTAGTCTTCCCGTGACCACGTTGCCTTTAACGAGTAGTCGAGTCATAACGGGCGA

4020 CCACGGTTTATAAAATTCTTGAAGACGAAAGGGCCTCGTGATACGCCTATTTTTATAGGTTAATGTCATGATAATAAT
     GGTGCCAAATATTTTAAGAACTTCTGCTTTCCCGGAGCACTATGCGGATAAAAATATCCAATTACAGTACTATTATTA
```

```
4098 GGTTTCTTAGACGTCAGGTGGCACTTTTCGGGGAAATGTGCGCGGAACCCCTATTTGTTTATTTTTCTAAATACATTC
     CCAAAGAATCTGCAGTCCACCGTGAAAAGCCCCTTTACACGCGCCTTGGGGATAAACAAATAAAAAGATTTATGTAAG

4176 AAATATGTATCCGCTCATGAGACAATAACCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGAGTATTCA
     TTTATACATAGGCGAGTACTCTGTTATTGGGACTATTTACGAAGTTATTATAACTTTTTCCTTCTCATACTCATAAGT

4254 ACATTTCCGTGTCGCCCTTATTCCCTTTTTTGCGGCATTTTGCCTTCCTGTTTTTGCTCACCCAGAAACGCTGGTGAA
     TGTAAAGGCACAGCGGGAATAAGGGAAAAAACGCCGTAAAACGGAAGGACAAAAACGAGTGGGTCTTTGCGACCACTT

4332 AGTAAAAGATGCTGAAGATCAGTTGGGTGCACGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGA
     TCATTTTCTACGACTTCTAGTCAACCCACGTGCTCACCCAATGTAGCTTGACCTAGAGTTGTCGCCATTCTAGGAACT

4410 GAGTTTTCGCCCCGAAGAACGTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGT
     CTCAAAAGCGGGGCTTCTTGCAAAAGGTTACTACTCGTGAAAATTTCAAGACGATACACCGCGCCATAATAGGGCA
```