adf

Appendix A

Light-controlled *E. coli*

A.1 Plasmid Construction and Transformation

A.1.1 Plasmid Construction

Here I present the plasmid construction protocol using Gibson assembly. This method can ligate the target gene PR with the plasmid carrier pZE. We first use polymerase chain reaction (PCR) to amplify the gene to sufficient concentration for the ligation. PCR can be easily done with a standard thermo cycler, and all we need to do is to prepare a mixture according to the recipe shown in Table. A.1

The sequence of the primers can be found in Table. **??**, and are synthesized by Eurofins Scientific. After we obtain the more concentrated DNA from PCR, we use Gibson assembly to ligate the plasmid carrier pZE and the target DNA sequence PR. Gibson assembly is done in the following steps:

- prepare the Gibson assembly mix accoring to the Table. A.2.
- add the mix into a 200 μl centrifuge tube, then put the tube into a thermal cycler and run the Gibson assembly program (50-55 ℃ for 1 hour).

Ingredients	Amount (μl)
ddH_2O	37
5X HF buffer (Phusion)	10
dNTP	1
Primer-forward	0.5
Primer-reverse	0.5
Template	small amount of cell culture / 0.5 μl plasmid solution
Phusion polymerase	0.5
Sum	50

Table A.1: PCR recipe using Phusion polymerase

- thaw the chemical competent cells in ice (XL10-Gold, 3-4 min).
- mix cells and plasmid solution for 10 min in ice.
- spread the mix on a selective plate.

Ingredients	Amount (μl)
PCR product	3.3
pZE solution	1.7
2X Gibson assembly mix	5
Sum	10

Table A.2: Gibson assembly recipe.

A.1.2 Extraction and Purification

After growing the chemical competent cells into concentrated liquid cultures, we extract the plasmid from the cells and purify them. We follow the steps in the Zyppy Plasmid Miniprep Kit manual and Zymoclean Gel DNA Recovery Kit manual.

A.1.3 Transformation

Make electrocompetent cells

- Inoculate target cells (*E. coli*, 1 ml) into a test tube and incubate at 37 ℃ for 16-18 hours.
- Prepare a box of ice and put 10% glycerol on it.
- Place the cell culture on ice for 15 minutes (Note: from this step, the cell culture must be on ice all the time).
- Centrifuge at 4000 rpm for 5 min, at 4 ℃ (use the low temperature centrifuge)
- Discard the supernatant, aspirate the residual broth and add 1 ml glycerol. Re-suspend the cells by pipetting up and down.
- Repeat step 4 and 5 for 2 more times (Note: for the last time, add ~ 50 μl glycerol instead of 1 ml).

Electroporation

- Thaw the cell suspensions prepared in the previous step on ice. Place a 1.5 ml centrifuge tube and a 0.1 cm cuvette on ice.
- Add 40 μl cell suspension and 1 μl DNA solution to the 1.5 ml centrifuge tube.
- Set the electroporator mode to Ec1.
- Transfer the mixture of cells and DNA to the cuvette and tap it to make the cell suspension go to the bottom of the cuvette.
- Put the cuvette into the electroporator, and press the pulse button once.

- Remove the cuvette from the electroporator and spread the cell suspension on a selective plate.

A.2 DNA Sequences

A.2.1 Proteorhodopsin (PR)

```
   1 ATGAAGTTGT TGTTGATCTT GGGATCTGTG ATTGCACTTC CGACGTTCGC TGCCGGCGGC
  61 GGGGATTTGG ACGCATCAGA CTACACAGGG GTTTCGTTTT GGCTGGTCAC AGCCGCGCTG
 121 TTAGCGTCCA CCGTCTTTTT TTTCGTTGAA CGGGATAGAG TCTCAGCTAA GTGGAAGACA
 181 TCGTTGACCG TCTCAGGCTT GGTTACCGGC ATTGCCTTCT GGCATTATAT GTACATGCGT
 241 GGTGTCTGGA TTGAAACGGG TGACAGCCCG ACGGTGTTCC GTTATATCGA CTGGCTTTTA
 301 ACCGTTCCCC TTCTGATTTG TGAGTTTTAT TTAATATTGG CGGCAGCAAC GAATGTGGCC
 361 GGTTCACTGT TCAAGAAGCT TCTTGTAGGA AGTTTAGTTA TGTTGGTTTT CGGCTACATG
 421 GGAGAGGCAG GGATAATGGC GGCCTGGCCG GCGTTCATAA TTGGTTGCTT GGCTTGGGTG
 481 TACATGATCT ACGAGCTGTG GGCAGGAGAA GGCAAGTCTG CGTGCAACAC AGCATCGCCA
 541 GCAGTTCAAT CCGCATATAA TACGATGATG TATATAATTA TCTTTGGTTG GGCAATTTAC
 601 CCGGTCGGAT ACTTCACCGG CTATCTTATG GGCGACGGGG GCTCTGCCTT GAACTTGAAT
 661 CTTATATATA ACCTGGCCGA TTTCGTGAAC AAGATTTTGT TTGGACTTAT AATATGGAAC
 721 GTAGCCGTGA AAGAGTCATC GAACGCATAA
```

A.2.2 pZE-PR

```
   1 AGGCGTATCA CGAGGCCCTT TCGTCTTCAC CTCGAGAATT GTGAGCGGAT AACAATTGAC
  61 ATTGTGAGCG GATAACAAGA TACTGAGCAC ATCAGCAGGA CGCACTGACC GAATTCATTA
 121 AAGAGGAGAA AGGTACCATG AAGTTGTTGT TGATCTTGGG ATCTGTGATT GCACTTCCGA
 181 CGTTCGCTGC CGGCGGCGGG GATTTGGACG CATCAGACTA CACAGGGGTT TCGTTTTGGC
 241 TGGTCACAGC CGCGCTGTTA GCGTCCACCG TCTTTTTTTT CGTTGAACGG GATAGAGTCT
```

```
 301 CAGCTAAGTG GAAGACATCG TTGACCGTCT CAGGCTTGGT TACCGGCATT GCCTTCTGGC
 361 ATTATATGTA CATGCGTGGT GTCTGGATTG AAACGGGTGA CAGCCCGACG GTGTTCCGTT
 421 ATATCGACTG GCTTTTAACC GTTCCCCTTC TGATTTGTGA GTTTTATTTA ATATTGGCGG
 481 CAGCAACGAA TGTGGCCGGT TCACTGTTCA AGAAGCTTCT TGTAGGAAGT TTAGTTATGT
 541 TGGTTTTCGG CTACATGGGA GAGGCAGGGA TAATGGCGGC CTGGCCGGCG TTCATAATTG
 601 GTTGCTTGGC TTGGGTGTAC ATGATCTACG AGCTGTGGGC AGGAGAAGGC AAGTCTGCGT
 661 GCAACACAGC ATCGCCAGCA GTTCAATCCG CATATAATAC GATGATGTAT ATAATTATCT
 721 TTGGTTGGGC AATTTACCCG GTCGGATACT TCACCGGCTA TCTTATGGGC GACGGGGGCT
 781 CTGCCTTGAA CTTGAATCTT ATATATAACC TGGCCGATTT CGTGAACAAG ATTTTGTTTG
 841 GACTTATAAT ATGGAACGTA GCCGTGAAAG AGTCATCGAA CGCATAATCT AGAGGCATCA
 901 AATAAAACGA AAGGCTCAGT CGAAAGACTG GGCCTTTCGT TTTATCTGTT GTTTGTCGGT
 961 GAACGCTCTC CTGAGTAGGA CAAATCCGCC GCCCTAGACC TAGGCGTTCG GCTGCGGCGA
1021 GCGGTATCAG CTCACTCAAA GGCGGTAATA CGGTTATCCA CAGAATCAGG GGATAACGCA
1081 GGAAAGAACA TGTGAGCAAA AGGCCAGCAA AAGGCCAGGA ACCGTAAAAA GGCCGCGTTG
1141 CTGGCGTTTT TCCATAGGCT CCGCCCCCCT GACGAGCATC ACAAAAATCG ACGCTCAAGT
1201 CAGAGGTGGC GAAACCCGAC AGGACTATAA AGATACCAGG CGTTTCCCCC TGGAAGCTCC
1261 CTCGTGCGCT CTCCTGTTCC GACCCTGCCG CTTACCGGAT ACCTGTCCGC CTTTCTCCCT
1321 TCGGGAAGCG TGGCGCTTTC TCAATGCTCA CGCTGTAGGT ATCTCAGTTC GGTGTAGGTC
1381 GTTCGCTCCA AGCTGGGCTG TGTGCACGAA CCCCCCGTTC AGCCCGACCG CTGCGCCTTA
1441 TCCGGTAACT ATCGTCTTGA GTCCAACCCG GTAAGACACG ACTTATCGCC ACTGGCAGCA
1501 GCCACTGGTA ACAGGATTAG CAGAGCGAGG TATGTAGGCG GTGCTACAGA GTTCTTGAAG
1561 TGGTGGCCTA ACTACGGCTA CACTAGAAGG ACAGTATTTG GTATCTGCGC TCTGCTGAAG
1621 CCAGTTACCT TCGGAAAAAG AGTTGGTAGC TCTTGATCCG GCAAACAAAC CACCGCTGGT
1681 AGCGGTGGTT TTTTTGTTTG CAAGCAGCAG ATTACGCGCA GAAAAAAAGG ATCTCAAGAA
1741 GATCCTTTGA TCTTTTCTAC GGGGTCTGAC GCTCAGTGGA ACGAAAACTC ACGTTAAGGG
1801 ATTTTGGTCA TGACTAGTGC TTGGATTCTC ACCAATAAAA AACGCCCGGC GGCAACCGAG
```

1861 CGTTCTGAAC AAATCCAGAT GGAGTTCTGA GGTCATTACT GGATCTATCA ACAGGAGTCC

1921 AAGCGAGCTC TCACTGCCCG CTTTCCAGTC GGGAAACCTG TCGTGCCAGC TGCATTAATG

1981 AATCGGCCAA CGCGCGGGGA GAGGCGGTTT GCGTATTGGG CGCCAGGGTG GTTTTTCTTT

2041 TCACCAGTGA GACGGGCAAC AGCTGATTGC CCTTCACCGC CTGGCCCTGA GAGAGTTGCA

2101 GCAAGCGGTC CACGCTGGTT TGCCCCAGCA GGCGAAAATC CTGTTTGATG GTGGTTAACG

2161 GCGGGATATA ACATGAGCTG TCTTCGGTAT CGTCGTATCC CACTACCGAG ATATCCGCAC

2221 CAACGCGCAG CCCGGACTCG GTAATGGCGC GCATTGCGCC CAGCGCCATC TGATCGTTGG

2281 CAACCAGCAT CGCAGTGGGA ACGATGCCCT CATTCAGCAT TTGCATGGTT TGTTGAAAAC

2341 CGGACATGGC ACTCCAGTCG CCTTCCCGTT CCGCTATCGG CTGAATTTGA TTGCGAGTGA

2401 GATATTTATG CCAGCCAGCC AGACGCAGAC GCGCCGAGAC AGAACTTAAT GGGCCCGCTA

2461 ACAGCGCGAT TTGCTGGTGA CCCAATGCGA CCAGATGCTC CACGCCCAGT CGCGTACCGT

2521 CTTCATGGGA GAAAATAATA CTGTTGATGG GTGTCTGGTC AGAGACATCA AGAAATAACG

2581 CCGGAACATT AGTGCAGGCA GCTTCCACAG CAATGGCATC CTGGTCATCC AGCGGATAGT

2641 TAATGATCAG CCCACTGACG CGTTGCGCGA GAAGATTGTG CACCGCCGCT TTACAGGCTT

2701 CGACGCCGCT TCGTTCTACC ATCGACACCA CCACGCTGGC ACCCAGTTGA TCGGCGCGAG

2761 ATTTAATCGC CGCGACAATT TGCGACGGCG CGTGCAGGGC CAGACTGGAG GTGGCAACGC

2821 CAATCAGCAA CGACTGTTTG CCCGCCAGTT GTTGTGCCAC GCGGTTGGGA ATGTAATTCA

2881 GCTCCGCCAT CGCCGCTTCC ACTTTTTCCC GCGTTTTCGC AGAAACGTGG CTGGCCTGGT

2941 TCACCACGCG GGAAACGGTC TGATAAGAGA CACCGGCATA CTCTGCGACA TCGTATAACG

3001 TTACTGGTTT CATGGTATAT CTCCTTCGAG CTCGTAAACT TGGTCTGACA GTTACCAATG

3061 CTTAATCAGT GAGGCACCTA TCTCAGCGAT CTGTCTATTT CGTTCATCCA TAGTTGCCTG

3121 ACTCCCCGTC GTGTAGATAA CTACGATACG GGAGGGCTTA CCATCTGGCC CCAGTGCTGC

3181 AATGATACCG CGAGACCCAC GCTCACCGGC TCCAGATTTA TCAGCAATAA ACCAGCCAGC

3241 CGGAAGGGCC GAGCGCAGAA GTGGTCCTGC AACTTTATCC GCCTCCATCC AGTCTATTAA

3301 TTGTTGCCGG GAAGCTAGAG TAAGTAGTTC GCCAGTTAAT AGTTTGCGCA ACGTTGTTGC

3361 CATTGCTACA GGCATCGTGG TGTCACGCTC GTCGTTTGGT ATGGCTTCAT TCAGCTCCGG

```
3421 TTCCCAACGA TCAAGGCGAG TTACATGATC CCCCATGTTG TGCAAAAAAG CGGTTAGCTC
3481 CTTCGGTCCT CCGATCGTTG TCAGAAGTAA GTTGGCCGCA GTGTTATCAC TCATGGTTAT
3541 GGCAGCACTG CATAATTCTC TTACTGTCAT GCCATCCGTA AGATGCTTTT CTGTGACTGG
3601 TGAGTACTCA ACCAAGTCAT TCTGAGAATA GTGTATGCGG CGACCGAGTT GCTCTTGCCC
3661 GGCGTCAATA CGGGATAATA CCGCGCCACA TAGCAGAACT TTAAAAGTGC TCATCATTGG
3721 AAAACGTTCT TCGGGGCGAA AACTCTCAAG GATCTTACCG CTGTTGAGAT CCAGTTCGAT
3781 GTAACCCACT CGTGCACCCA ACTGATCTTC AGCATCTTTT ACTTTCACCA GCGTTTCTGG
3841 GTGAGCAAAA ACAGGAAGGC AAAATGCCGC AAAAAAGGGA ATAAGGGCGA CACGGAAATG
3901 TTGAATACTC ATACTCTTCC TTTTTCAATA TTATTGAAGC ATTTATCAGG GTTATTGTCT
3961 CATGAGCGGA TACATATTTG AATGTATTTA GAAAAATAAA CAAATAGGGG TTCCGCGCAC
4021 ATTTCCCCGA AAAGTGCCAC CTGACGTCTA AGAAACCATT ATTATCATGA CATTAACCTA
4081 TAAAAAT
```

A.2.3 An Illustration of the Plasmid Structure

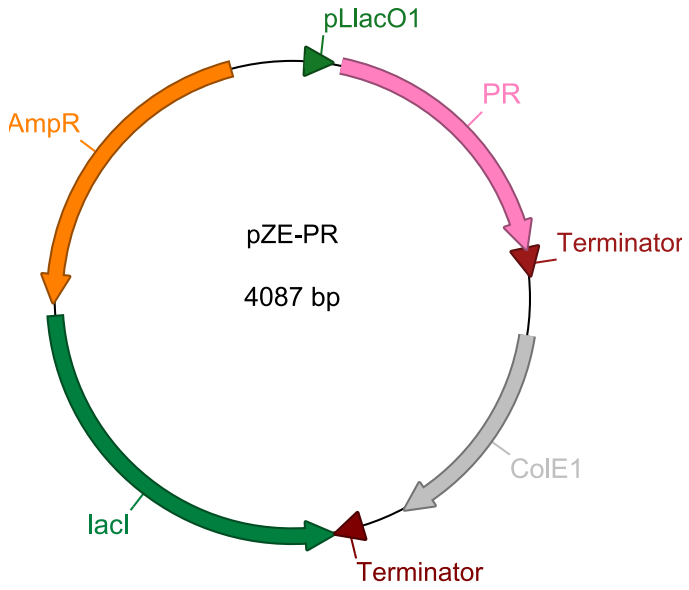

Figure A.1: An illustration of pZE-PR plasmid.

Appendix B

Image Analysis implementations

B.1 Code Vectorization

B.1.1 Vectorized Code for Spatial Correlation Function

```
def corrS(X, Y, U, V):
    row, col = X.shape
    vsqrt = (U ** 2 + V ** 2) ** 0.5
    U = U - U.mean()
    V = V - V.mean()
    Ax = U / vsqrt
    Ay = V / vsqrt
    CA = np.ones(X.shape)
    CV = np.ones(X.shape)
    for xin in range(0, col):
        for yin in range(0, row):
            if xin != 0 or yin != 0:
                CA[yin, xin] = (Ax[0:row-yin, 0:col-xin] * Ax[yin:row, xin:col] + Ay[0:r
                CV[yin, xin] = (U[0:row-yin, 0:col-xin] * U[yin:row, xin:col] + V[0:row-
    return CA, CV
```

B.1.2 Non-vectorized Code for Spatial Correlation Function

```
def corrS(X, Y, U, V):
    row, col = X.shape
    vsq = 0
    CA = np.zeros((row, col))
    CV = np.zeros((row, col))
    for i in range(0, row):
        for j in  range(0, col):
            vsq += U[i, j]**2 + V[i, j]**2
    for xin in range(0, col):
        for yin in range(0, row):
            count = 0
            CAt = 0
            CVt = 0
            for i in range(0, col-xin):
                for j in range(0, row-yin):
                    ua = U[j, i]
                    va = V[j, i]
                    ub = U[j+yin, i+xin]
                    vb = V[j+yin, i+xin]
                    CAt += (ua*ub+va*vb)/((ua**2+va**2)*(ub**2+vb**2))**.5
                    CVt += ua*ub + va*vb
                    count += 1
            CA[yin, xin] = CAt / count
            CV[yin, xin] = CVt / vsq
    return CA, CV
```

B.1.3 Performance Comparison

We notice that the vectorized code has two less nested `for` loops compared to the non-vectorized code. As a result, the vectorized one runs much faster for the same task. To quantify this performance difference, we perform the spatial correlation function calculation using both code on the same velocity field, shown in Fig. B.1a. The times

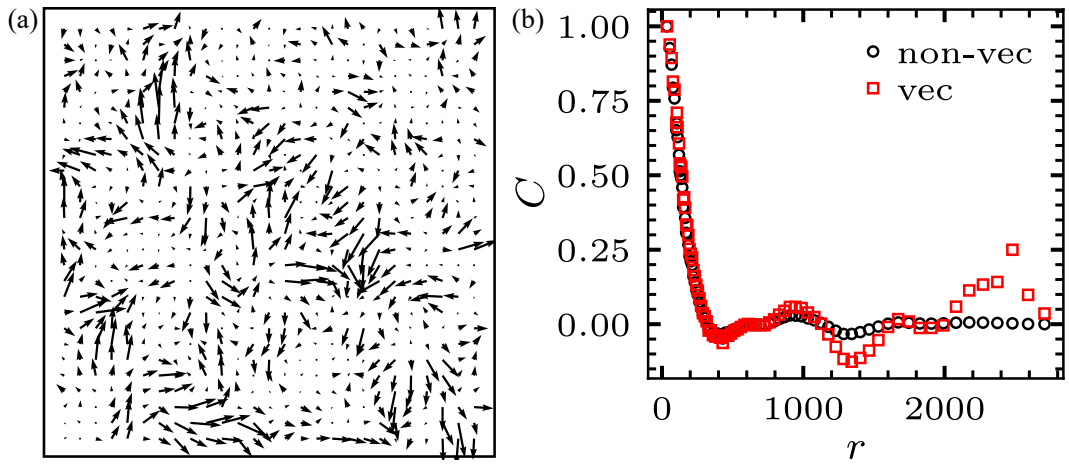

Figure B.1: **Compare the performance of vectorized and non-vectorized code.** (a) Sample velocity field. (b) Velocity correlation functions obtained from the vectorized and non-vectorized code.

taken for the two functions are:

- `Vectorized code:  0.84 s`
- `Vectorized code:  52.06 s`

The result is shown in Fig. B.1b. Although in the large r regime, two methods show descrepancies, in the meaningful small r regime, two methods give exactly the same results.

B.2 Energy Spectrum Calculation

B.3 Cross-correlation Tracking Method

B.4 Fourier Transform Based Orientation Analysis

B.5 Density Fluctuation Calculation