Fluorescent Reporter Proteins

By David Form,
Nashoba Regional High School,
Bolton, MA

Introduction:

Green fluorescent protein (GFP) is naturally found in Aquoria victoria, a species of jellyfish found in the waters of the northern Pacific Ocean. It has the property of absorbing blue light and then emitting green light, making it fluorescent. By fusing the GFP gene with another gene, scientists can use GFP as reporter gene in biological studies, signaling gene expression or the location of a protein in a cell. We have used GFP to signal the expression of particular genes during microbial biofilm development.

In green fluorescent protein, the amino acid chain folds into a barrel-like structure with an amino acid coil extending into the barrel. This coil contains three amino acids, serine, tyrosine and glycine (SYG). The serine and glycine interact with the tyrosine to form the chromophore, or light absorbing and emitting structure of the protein. The light emitting properties of GFP can be changed by altering the structure or stability of the chromophore.

Scientists have developed mutant forms of GFP which emit light of different wavelengths, such as blue, yellow, cyan and cerulean, allowing them to visualize the expression of more than one gene or the interaction of several proteins during biological processes. Other small fluorescent proteins, such as red fluorescent protein have been isolated from other species of coelenterates. In this activity, you will explore structure/function relationships in these fluorescent proteins by comparing their amino acid sequences and their three dimensional structures.

Procedure:

1. Using Biology Workbench to construct and analyze the primary amino acid sequences of GFP, its color variants and other fluorescent proteins isolated from coelenterates.

Our multiple alignment will be prepared and analyzed using a collection of bioinformatics tools available through the ***Biology Workbench*** @workbench.sdsc.edu/. Your first step is to set up an

account, by clicking on ***set up an account*** and following the directions. The amino acid sequences for GFP, its variants and other fluorescent reporter proteins are in a text file, green fluorescent protein.txt.

Step 1: Start a session

1. Scroll down until you see a series of boxes.
2. Click on the box labeled ***Session Tools***. Now look through the menu of the long rectangular box and select ***Start New Session***. Name the session.
3. Select the box labeled ***Protein Tools***.

Step 2: You are now ready to upload the fluorescent reporter sequences.

4. From the menu, select ***Add a New Sequence.*** A new box will appear on your screen.
5. Select ***Browse***, then find the file named green fluorescent protein.txt and open it.
6. Select the **Upload** button. The sequences will be uploaded into your session window.
7. Go to the top or bottom of the screen and click on the ***Save*** button. You should now see a list of your sequences.
8. Click to the left of the sequence name to select it. Select those sequences `which you will be comparing. For your first alignment, select all of the sequences other than the two rfp sequences.
9. From the menu box, select **ClustalW**. Select ***submit*** from the box on the screen. When the program has finished running, scroll down to view your alignment.
10. Find the window labeled ***Import Alignments*** and click on it.
11. Under the rectangle labeled Alignment Tools, select ***TextShade.***
12. Click on the small box next to the list of aligned sequences and select **submit**.
13. Scroll down to see your multiple alignment. Look for the three amino acids, SYG (serine, tyrosine and glycine), characteristic of the chromophore of GFP. The tyrosine should be at position 66. Positions at which there are no amino acid differences are shaded in blue. Positions at which there are different amino acids are unshaded.
14. Using FirstGlance in JMOL to examine the 3-D structure of GFP.

a. Go FirstGlance in JMOL @ http://molvis.sdsc.edu/fgij/

b. Enter the PDB id number (in this case, gfl) and click on **Submit**.

c. When the model loads, click on background and on spin. You can now use your mouse/cursor to rotate the protein model. You should see two molecules of GFP. Each molecule is barrel shaped with a ribbon entering the top of the barrel. The end of the ribbon contains the three amino acids that comprise the chromophore of GFP.

Critical Thinking Exercises:

1. Construct a multiple alignment of the amino acid sequences of the original GFP and its fluorescent color variants, blue fluorescent protein (BFP), yellow fluorescent protein (YFP) , cyan fluorescent protein and cerulean fluorescent protein. Using Textshade to examine your multiple alignment:

   a. Do you see any evidence that altering the three amino acids that make up the chromophore can change the fluorescence color of the protein?

   b. Do you see any evidence that altering amino acids outside of the chromophore can alter the stability of the chromophore in such a way as to change the color of the fluorescence.

2. Use Star Biochem to compare the 3-D structures of the original GFP with one of its fluorescent color variants. Do they share a common structure? Locate the positions of the amino acid changes. Where are these located on the molecule.

3. Construct a multiple alignment of the amino acid sequences of the original GFP and the two red fluorescent proteins (RFP), isolated from the coelenterates, Zooanthus sp. and Entremacaea quadricolor.

   a. Do these fluorescent proteins appear to be homologous, or share a common ancestor?

   b. Compare the 3-D structures of these three fluorescent proteins. Do they share a common structure?

Protein Data Bank (PDB) id numbers for 3D structures of Fluorescent Reporter Proteins

Protein	**PDB id**
GFP (green fluorescent protein)	1GFL
TFP (yellow fluorescent protein)	3DPW
BFP (blue fluorescent protein)	1BFP
Cerulean fluorescent protein	2WSO
Cyan fluorescent protein	2WSN
RFP (red fluorescent protein from Zooanthus sp.	2ICR
RFP (red fluorescent protein from Entremacaea quadricolor)	2PJB

Primary Amino Acid Sequences of Fluorescent Reporter Proteins

> GFP original
MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTLVTTFSYGVQCFSRYPDHMKQ
HDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNG
IKVNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK

> YFP
KGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTLVTTFXLQCFARYPDHMKRHDFF
KSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVN
FKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSYQSALSKDPNEKRDHMVLLEFVTAAGI

> BFP
MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTLVTTFXVQCFSRYPDHMKRHD
FFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNFNSHNVYIMADKQKNGIK
VNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK

> CFP (cerulean)
MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTLVTTLXVQCFARYPDHMKQ
HDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNAISDNVYITADKQKNGI
KANFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITLGMDELYK

> CFP (cyan)
MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTLVTTLXVQCFSRYPDHMKQ
HDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYISHNVYITADKQKNGIK
ANFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITLGMDELYK

> RFP (Zooanthus)
MRGSHHHHHHGSAHGLTDDMTMHFRMEGCVDGHKFVIEGNGNGNPFKGKQFINLCVIEGGPLPFSEDILSAAFXNRL
FTEYPEGIVDYFKNSCPAGYTWHRSFRFEDGAVCICSADITVNVRENCIYHESTFYGVNFPADGPVMKKMTTNWEPSCE
KIIPINSQKILKGDVSMYLLLKDGGRYRCQFDTIYKAKTEPKEMPDWHFIQHKLNREDRSDAKNQKWQLIEHAIASRSALP

>RFP (Entremacaea)
MSELIKENMHMKLYMEGTVNNHHFKCTSEGEGKPYEGTQTMKIKVVEGGPLPFAFDILATSFXSKTFINHTQGIPDFFKQ
SFPEGFTWERITTYEDGGVLTATQDTSLQNGCIIYNVKINGVNFPSNGSVMQKKTLGWEANTEMLYPADGGLRGHSQ
MALKLVGGGYLHCSFKTTYRSKKPAKNLKMPGFHFVDHRLERIKEADKETYVEQHEMAVAKYCDLPSKLGHR