A Thousand Proteins of Light: 15 Years of Advances in Fluorescent Proteins

George McNamara*,[1], and **Carl A. Boswell[2]**

[1]Analytical Imaging Core, Miller School of Medicine, University of Miami, 1450 NW 10th Ave (R-134), Miami, FL 33136

[2]Molecular and Cellular Biology and Arizona Cancer Center, University of Arizona,1007 E. Lowell St., Tucson, Arizona 85721

This chapter describes a database containing over 1000 wild-type and variants of fluorescent proteins derived promarily from jellyfish, corals and other cniderians. Within the database are details on fluorescent properties of the proteins, as well as specifics regarding amino acid substitutions and references for each one. Also included are data for luminescent and photoproteins.

Keywords fluorescent proteins; Aequorea; GFP; FRET;

1. Introduction

The cloning of the jellyfish *Aequorea victoria* GFP cDNAs by Prasher et al (1992) and heterologous expression by Chalfie et al (1994), has revolutionized cell biology by making it easy to make proteins, cells, tissues, and even whole organisms glow in the dark. GFP is the light emitter, following resonance energy transfer from the calcium ion sensitive photoprotein aequorin, whose cDNA was cloned by Prasher et al (1985). Around the same time, de Wet et al (1985) cloned a firefly luciferase, which itself has had a huge impact as an ATP biosensor for industrial (no ATP implies no bacterial growth), cell, and whole organism testing. Early revisions of GFP included codon optimization for use in human cells, E.coli, yeast, and multicellular organisms, improved folding mutations and blue, cyan and yellow color changes (reviewed by Tsien, 1998). Efforts to isolate red, or even orange, fluorescent protein variants of *Aequorea* GFP, or from other bioluminescent beasts that contained a luciferases or photoprotein, were unsuccessful. The red color barrier was finally broken by Matz et al (1999), who had the revolutionary idea to search for fluorescent proteins in non-bioluminescent anthozoa (sea anemones and corals) from a local salt water aquarium enthusiast. The same group later isolated chromoproteins (non-fluorescent but strongly colored proteins) from many other organisms, including crustacea (copepods) and mutagenized several to make new fluorescent proteins (Gurskaya et al 2001). The early history of bioluminescence and fluorescent proteins have been reviewed in books by Shimomura (2006), Pieribone and Gruber (2006) and Zimmer (2005).

The intent of this report is to summarize an electronic dataset of the more than one thousand fluorescent proteins and chromoproteins we have collected from the literature that has been published in the 15 years since Prasher et al (1992). In the interest of space, the >250 references are listed online. We also briefly refer to the over 200 eukaryotic luciferases (the number of *Vibrio* and other bacterial luciferases are countless), over 300 phytochromes and related molecules and one fluorescent phytofluor mutant, and over 400 phycobiliproteins (phycroerythrins, allophycocyanins). We leave for others to summarize tissue autofluorophores (NAD, flavoproteins, collagen, elastin) and plant chlorophyll and its accessory molecules[1].

*Corresponding authors: George McNamara, geomcnamara@earthlink.net, gmcnamara@med.miami.edu, 305-243-8436; Carl Boswell, cboswell@email.arizona.edu, 520-954-7053.

2. Organization and content of the database

2.1 Fluorescent Protein Data worksheet

This is the main worksheet in the dataset. It lists over 1000 fluorescent and chromoproteins, more than 100 functional fusion proteins, and over 250 references. Entries are added to the table as they get published and as we obtain full text access to older references. Hence, the reference numbers are not in chronological order. We often insert new variants next to their parental sequence, so the data is not in any particular order. The Summary worksheet (discussed in section 2.3) has been stripped of section headers to facilitate sorting. The Data worksheet includes spectral details, GeneBank references and alterations from the original sequence. Readers are encouraged to download the file and explore it.

2.2 Sections within the Fluorescent Protein Data worksheet

2.2.1 Header; rows 1-2

Row 1 is the column headers for the table, while row 2 identifies the position of beta sheets within the molecule. The first 13 columns in the worksheet include the reference number, year of publication, FP name, amino acid changes and comments, absorption or excitation maximum wavelength, emission maximum wavelength, extinction coefficient, quantum yield, fluorescence lifetime (Tau1, Tau2, in nanoseconds), brightness index (extinction coefficient * quantum yield / 1000), Stokes shift (emission maximum – excitation maximum), and number of amino acid changes from the *Aequorea victoria* GFP (AvGFP) reference cDNA. Columns N-IR (239 total) identify the amino acid for each position in the protein. Columns IS-IV include the taxonomic category, subunit composition and extended comments.

2.2.2 Aequorea victoria GFP wt cDNA; rows 3-10

Several cDNAs were isolated independently by the groups of Prasher (#1, #68, #70) Inouye and Tsuji (#244), Watkins and Campbell (#245). An additional cDNA sequence was later deposited by Kono (#247). We have added a "consensus" sequence, identified in the table as AvGFP(consensus sequence). Most AvGFP mutants are based on Prasher's sequences, and commonly contain a Q80R substitution that is apparently a phenotypically silent substitution that appeared in the initial cDNA (the jellyfish presumably has Q80). It is unknown how many alleles or genetic loci the jellyfish has (the *Montastraea cavernosa* coral has a multi-gene family encoding different colors).

2.2.3 *Aequorea victoria* GFP variants; rows 13-641

With over 600 variants listed, this is the single largest section of the dataset (though the DsRed *Discosoma* is catching up). There are a few duplicate entries, with distinct amino acid sequences (such as for ECFP, where we report the "as-published" sequence from each source. Codons optimized for different species are listed subsequently. This table emphasizes the amino acid changes, not the nucleotides in the codons. Some optimizations may improve mRNA stability but decrease protein translation rate, or vice versa. For instance, the "human" codon optimization was discovered to work well in plants (i.e. *Arabidopsis*) by eliminating a cryptic splice site that was specific to plants [#19, mGFP4]. We have not tabulated destabilized forms (i.e. GFP with a PEST sequence) – several of these are commercially available, but it behooves the researcher to verify that the GFP is really degraded, since sometimes just the linker is cleaved while the released FP retains activity. The destabilized protein forms may be more useful than standard forms for kinetic assays, such as how quickly an mRNA is produced and protein translated – and in the case of GFP, matured – following transcriptional activation. mRNA destabilization is also possible (i.e. specific 3' untranslated regions). GFP, especially EGFP, has been a

favorite target for testing siRNA, though different labs and companies may have chosen different target sequences.

By far the most commonly used FP in biomedical research is EGFP. This is not to say that EGFP is the best in any quantitative sense. Like fluorescein isothiocyanate (FITC) for immunofluorescence applications, EGFP may not be ideal, but it has been good enough for most users. Shaner et al (2005) argue for Emerald as the best green FP, with mEmerald being the currently preferred monomeric Emerald. The Emerald's do have a very fast photobleaching component that will make quantitation susceptible to subtle experiment variations. For any instrument with the optimal illumination source and filters, mCitrine or mVenus should outperform EGFP in nearly all assays. mVenus is the preferred starting point for split fluorescent proteins in bimolecular complementation assays (biFC).

The ingenuity of researchers to modify GFP is almost endless. Color, improved folding, improved thermotolerance, and many more mutations have been introduced. For instance circular permutation (fusing the original N-and C-termini and making new termini elsewhere) (ref), splitting into separate polypeptides (and providing ways for the parts to find one another) (ref), removal of all cysteines) (ref),, re-addition of cysteines elsewhere to make improved redox sensors) (ref),, insertion of additional amino acids into loops) (ref), and lopping off the N- and/or C-termini) (ref), have all been published.

2.2.4 Other *Aequorea* species GFPs; rows 645-669

Along with *Aequorea victoria*, there are over 20 additional *Aequorea* species of jellyfish from which a GFP gene or cDNA has been isolated.

2.2.5 Other jellyfish FPs; rows 673-682

Six fluorescent proteins have been isolated from several other genera of jellyfish, mostly from *Phialidium* species, not all of which are green. PhiYFFP (525ex/537em) is yellow (or at least yellow-ish – color appearance to the human eye of all green vs yellow FPs depends on what filter set is used) and J-red (584ex/610em, a.k.a. MaxGFP-Red) is red.

2.2.6 Other Cnidarian FPs; rows 686-733

This section includes *Renilla reniformis* and *Renilla mulleri*, which have long histories in bioluminescence research because both species of sea pansy have both a luciferase (RLuc) and a GFP (RrGFP, RmGFP). The *Ptilosarcus gurneyi* sea pen GFP is also listed in this section.

2.2.7 *Discosoma* coral FPs; rows 738-910

The isolation of DsRed and other fluorescent proteins by Matz et al (ref #), from non-bioluminescent corals has triggered a second revolution, focused mostly in the red. The progenitor of this series, DsRed, was isolated from *Discosoma striata***.** With over 160 variants, including the popular fruity FP series, *Discosoma* certainly rates it own section.

2.2.8 Other Coral FPs (non-*Discosoma*); rows 914-1057

Matz et al (ref #) and many other labs have isolated over 180 fluorescent proteins from corals other than *Discosoma*. This includes several polymorphisms and mutational variants of FPs from *Montastraea cavernosa* (McFPs), the great star coral. Additional McFPs appear below in the EvoFP section.

2.2.9 Sea anemone FPs; rows 1061-1140

Sea anemones are essentially soft corals without the calcium carbonate. Over 50 fluorescent protein variants are listed. The discovery by Matz et al (1999, #115) of FPs in the sea anemone *Anemonia*

majano (as well as DsRed from *Discosoma striata* coral) resulted from their hypothesis that FPs did not require luciferase proteins – that is, FPs could exist in non-bioluminescent organisms. It is noteworthy that Matz found these Indo-pacific organisms bioprospecting in a local (Moscow), privately maintained, salt water aquaria.

2.2.10 EvoFPs rows; 1158-1203

The EvoFP series (#240, #241) are from Matz et al's synthesis of possible ancestral forms of *Montastraea cavernosa* great star coral FPs (the non-EvoFP McFPs are in the coral FPs section).

2.2.11 Copepod FPs; 1144-1154

Copepods, are planktonic and benthic arthropods, from sea and fresh water habitats. The 11 variants listed are salt to brackish water plankton. The discovery of FPs in arthropods suggests that nearly any animal could have an FP gene.

2.2.12 Unknown phylogeny FPs; rows 1207-1208

Diversa Corp. sampled an ocean environment and deliberately did not attempt to identify source organism (one way to avoid patents?). The DiscoveryPoint cyan and green FPs are the only fruit of this effort. We encourage bio-prospectors to bring someone along on expeditions to isolate, identify and extract organisms that fluoresce. This could be a fruitful resource for novel fluorescent molecules

2.2.13 Fusion proteins; rows 1212-1346

Over 140 fusion proteins are listed. These are mostly FP-X-FP fusions, where X can be practically anything! Or nothing, as in CY11.5, C5Y and C5A (discussed below with respect to FRET). If two FPs are good, more might be better, so we mention 3xGFP, 4xGFP and 5xGFP (rows 1290-1299) as examples of putting more than two FPs together. Part of the value of three or more is to increase the size of the polypeptide to make it bigger than the ~60 kDa limit for free diffusion into the nucleus (of eukaryotyic cells). Of course if you want to target a 4xGFP to the nucleus, a nuclear import sequence (or more than one) can be added to the construct. We included the yeast Cdc3-GFP series (septin-GFP, #212) because some forms, with a rigid rod, proved to be nice fluorescence anisotropy probes during cell division.

For the purpose of this table, functional fusions are defined as being either between two fluorescent proteins, with or without bionsensor sequences, or fusions that result in novel fluorescence characteristics. Many of these constructs rely on changes in fluorescence resonance energy transfer (FRET) efficiency. For example, a CFP-protease cleavage site-YFP, can be expected to have a moderate to high FRET efficiency when intact, but very low FRET once cleavage has occurred and the subunits diffuse apart. For maximum FRET, a small distance (short linker) and optimum orientation of the donor and acceptor dipole moments are essential. Examples of high FRET CFP-YFP constructs include CyPet-YPet, CY11.5, C5Y and C5A. C5A (mCerulean-5aa linker-Amber) is notable for using Amber, a non-fluorescent derivative of YFP, as a "dark quencher". A key to success in FP-FP FRET is both the short linker and the orientation. One can even imagine FP-FP designs where dipole moment orientations are deliberately designed to make the FRET orientation factor ($Kappa^2$) very small, resulting in deliberately very low FRET efficiency (FRET efficiency is high with parallel or linear dipole moments and decreases as the dipole moment approaches perpendicular).

Many FP-protease cleavage site-FP constructs populate the fusion table. This reflects both the ease of testing for loss of FRET in vitro (just add trypsin or whatever purified protease is desired) and the publishability of caspase/apoptosis research in cells, upon which many of these references are based.

A large family of CaMeleon variants also populate the fusion table. These are named for the prototypical ECFP-Cameleon-M13-YFP, where binding of calcium to a calmodulin domain results in

binding to the adjacent M13 peptide, resulting in an increase in FRET. Variations have included changing the FRET donor from ECFP to Cerulean (ref#) , changing EYFP to circularly permuted Venus (cpVenus) (ref#), mutagenizing the calmodulin domain to adjust calcium affinity (over several log Kd) (ref#), replacement of M13 with other binding partners(ref#), and replacement of calmodulin with other calcium binding proteins (i.e. troponin) (ref#).

Clomeleon is an example of a self-contained chloride and other halide ion biosensor, that takes advantage of CFP being halide insensitive, and YFP (the Topaz variant in this case) being halide sensitive. A nice touch is the inclusion of a TEV protease site, to enable cleavage to verify loss of FRET. The downside of this linker is that the FRET efficiency is only 35% (due to the C24T substitution (row 1212, 1326). Clomeleon is referenced twice in the table, once as itself (#34) and once as C24T (#219), because it is being used for different purposes (chloride sensor and FRET reporter prototype, respectively). If someone wanted to verify that a Clomeleon response was really due to halide response, a halide insensitive version would be a CFP-mVenus or CFP-mCitrine fusion.

2.2.14 Single FP Sensors; rows 1349-1357

A number of simple FP sensors have been made, including FlaSh and SPARC, made with a fluorescent protein fused to the Shaker potassium channel or a voltage gated sodium channel, respectively, several septin-GFP fusions as fluorescence anisotropy reporters (and controls!), and a heme binding domain fused to EGFP to report on changes in NO (nitric oxide).

2.2.15 Summary count table rows 1361-1379

The summary count lists the current totals for each section (over 1170 FPs and over 140 fusions, for over 1220 total as of March 2007). This section illustrates uses of the =ROW and =COUNTA commands of Microsoft Excel (column B; see Excel Tips worksheet).

2.2.16 Amino acid count; rows 1384-1459

We tabulate amino acid usage, on a per position basis, for *Aequorea* cDNAs, *Aequorea* variants, and *Discosoma* FPs (interested readers can easily duplicate and edit the formulas to obtain totals for other sections where amino acid composition is available).

2.2.17 References; rows 1464-1739

The References section contains all 276 references that have been used to construct the FPaa data table.

2.3 Fluorescent Protein summary table (sortable) worksheet

The purpose of this worksheet is to enable sorting by whatever criteria the user chooses. Sorting by amino acid substitutions is easily done by making a new worksheet (Insert>worksheet) and copying the contents of the Fluorescent Protein data worksheet into it.

The summary worksheet has been stripped down by removing most columns, including the fluorescent lifetimes and amino acid substitutions . The spacer rows, references and notes have also been removed. A column, has been added (replacing the right-most“Comments“ column) to make it easy to sort the rows back to the same order as in the FPaa Data worksheet. We used Copy -> Paste Special->Values to convert all formulas into values. This avoids the risk of sorting recalucalating the data.

3. Using the database

An example of what can be done with the data is shown in Figure 1, in which the relationship of absorption or excitation maximum vs the emission maximum is shown. Any single fluorophore or

fluorescent protein maxima are useful to appreciate what light source, filters and detector to use. The laws that govern single photon fluorescence dictate that the lower right corner must be empty because excitation maximum wavelength will always be less than the emission maximum. Two photon and three photon excitation is possible, but that discussion is beyond the scope of this work.

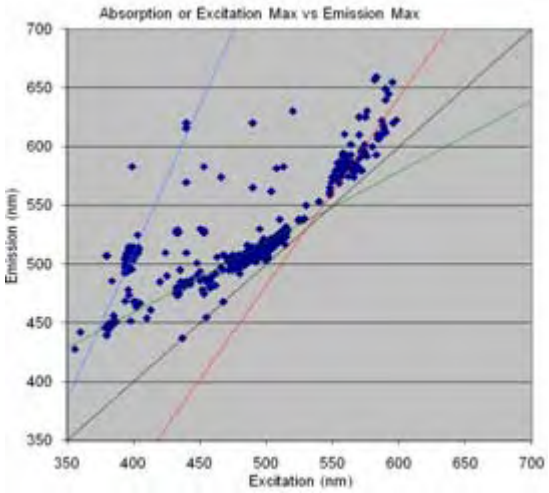

Fig. 1 Absorption/Excitation Maximum vs Emission Maximum. Cluster analysis indicates groups of data clustered in the blue, green and red ranges of emission.

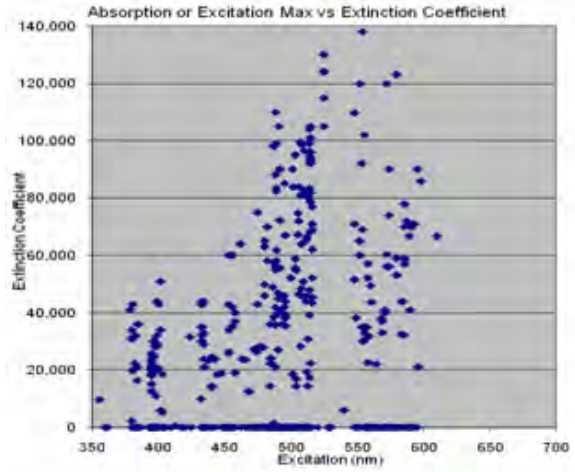

Fig. 2. Absorption/Excitation vs Extinction Cooefficient. Fluorescent proteins exhibit a wide range of extinction coefficients. All the fluorescent proteins above (and some below) Y=100,000 $M^{-1}cm^{-1}$ are dimers or tetramers. Only those fluorescent proteins with published extinction coefficients have been plotted.

Figure 2 shows the comparison between absorption/excitation vs extinction coefficient. All the *Aequorea* fluorescent proteins are to the left of this gap; the majority of the *Discosoma* and other coral and anemone fluorescent proteins are to the right. The gap that appears between approximately 530-550nm suggests that it may be biologically or enzymatically unfavorable to synthesize fluorescent proteins that have peak absorption (excitation) at these energy levels.

A similar relationship is suggested by the plot of emisison maximum vs. quantum yield (Figure 3.)

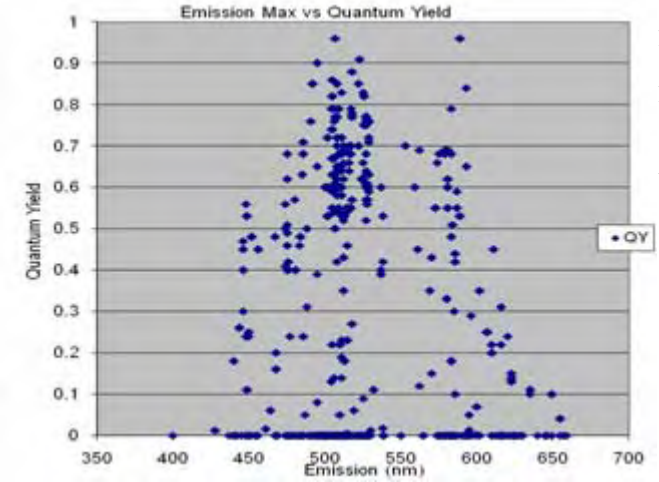

Fig. 3. Emission Max vs Quantum Yield. Only those fluorescent proteins with published quantum yield have been plotted. A yellow-orange gap around 530-570 nm suggests that it is difficult to obtain fluorescent proteins that have peak emission at these wavelengths.

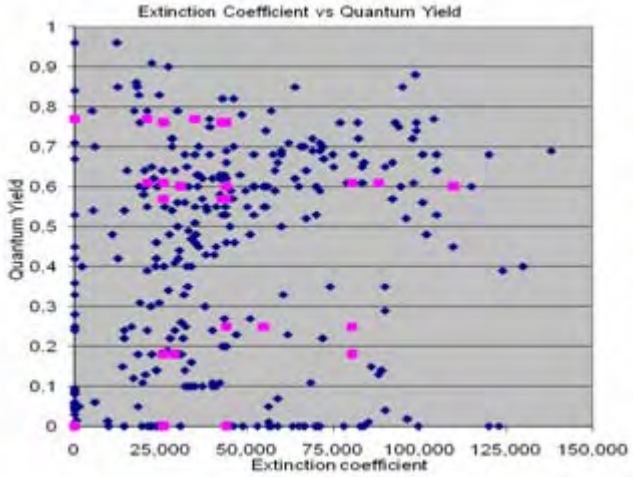

Fig. 4. Extinction Coefficient vs Quantum Yield.

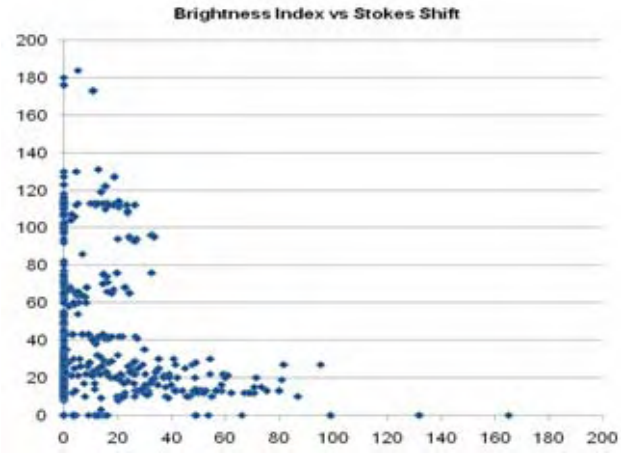

Fig. 5. Brightness Index vs Stokes Shift.

Figure 4 shows a wide scatter in extinction coefficient vs quantum yield. Chomoproteins, by definition, have low fluorescence quantum yield, but absorb light well enough to get noticed and be published. The FPs at zero extinction coefficient are an artifact of insufficisnt data: a zero is entered into the E.c. table entry if the data is not available. The Brightness Index is extinction coefficient * Quantum yield / 1000. This provides a simple statistic that can be used to plot relative brightness (ignoring photobleaching rate, halide quenching, etc) vs other parameters. Figure 5, reveals that a bright FP (large values on X-axis), or a large Stokes shift (large values on Y-axis), are mutually exclusive. The brightness index (Extinction coefficient * quantum yield / 1000) can be thought of as the maximum performance of a fluorescent protein or dye. FITC and Alexa Fluor 488 (a pH independent replacement for FITC for most applications) have a brightness index on the order of 90. Some FRET constructs, such as CY11.5 and C5Y, are one way to increase Stokes shift. Perhaps researchers will create an optimized GFPuv-mOrange fusion to generate a large Stokes shift fusion protein that is bright. The brightness index is really a best case scenario, since it tells us only about fluorescence performance, not the propensity to be quenched by molecular oxygen, halides or other molecules, to photobleach or undergo other photochemical reactions or time spent in blink-off or other dark states.

The original graphs, with data columns can be found in the Fluorescent Proteins Graph worksheet.

4. Additional worksheets

4.1 Aequorea FPaa changes effects worksheet

This worksheet has three parts. Part 1 (row 6-248) tabulates in column order the *Aequorea victoria* GFP amino acid changes (2006 data – see FPaa Data, part 3, for current summaries of *Aequorea* and *Discosoma* totals) Part 2 (rows 253-480) summarizes information by position, on the effects of amino acid substitutions.

4.2 Luciferase list worksheet

This worksheet summarizes over 200 eukaryotic luciferases (part 1) and 13 luciferases-FP and other fusion proteins (many more fusions exist). Many bacterial luciferases exist from *Vibrio* and other genera, which we have not tried to enumerate. We look forward to full details on destabilized firefly luciferases and circularly permuted firefly luciferases (Promega, 2006 presentation at the www.BLCL14.org meeting). Many more details on the variety of eukaryotic luciferases can be found in Shimomura (2006).

4.3 Luciferase Photoprotein summary worksheet

Summarizes information on the various protein families that absorb and or emit light. Much more information on the luciferases is found in Shimomura (2006) and in the luciferase lists worksheet. We note here a recent report by Inouye and Sasaki (2007) that the "complex of Ca(2+)-bound apoaequorin and coelenteramide" (the photoproduct) is blue fluorescent ("BFP-aq"), that removal of the calcium ion converts this to a green fluorescent, and – most interestingly– that BFP-aq is itself a luciferase that catalyzes the oxidation of coelenterazine by molecular oxygen, with the products being a second coelenteramide and a photon. If verified, this makes one wonder why Aequorea bothered with GFP, other than to give Doug Prasher something to do.

4.4 Phycobiliproteins worksheet

Phycobiliproteins include allophycocyanins (APC), phycocyanins (PC), B-, C- and R-phycoerthyrins (B-PE, C-PE, R-PE), and are found in many marine algae (including some fluorescent protein coral's zoanthellate symbionts). These are multi-subunit proteins that act as scaffolds for small molecule fluorophores. The result is a large molecule with a large extinction coefficient ("absorbability"). The

extinction coefficient can reach 1,000,000 $M^{-1}cm^{-1}$, vs ~90,000 for FITC, Alexa Fluor 488 or EGFP. APC and R-PE are popular flow cytometry reagents and can be modified in a test tube into 'tandem dyes' containing far red (ex. Texas Red, Cy5) and near infrared (ex. Cy5.5, Cy7) dyes, that use fluorescence resonance energy transfer (FRET) to shift the emission wavelength to enable multiplexing experiments. Unfortunately, APC and R-PE photobleach very quickly, to the point of being very difficult to use in fluorescence microscopy.

4.5 Phytochromes and Phytofluors worksheet

This worksheet contains two parts. Part 1 is a data table from Rockwell et al (2006), that lists 317 phytochromes: 147 wild-type (from nature) phytochromes, 50 Phr family members (phytochromes related GAF proteins) and 120 plant phytochromes mutants. The bottom of part 1 includes all the table's references. Phytochromes contain a (linear) tetrapyrrole prosthetic group.

Part 2 lists papers on the first Phytofluor – a mutant phytochrome that fluoresces bright red. Unfortunately, the first phytochromes structure was shown to have a trefoil knot (Wagner et al 2005), which may make development of fusion proteins difficult.

4.6 Excel tips worksheet

This worksheet summarizes some of the commands and tricks we use to be productive with Microsoft Excel. Many of these were used in the FPaa data worksheet. We include this worksheet to facilitate learning Excel.

4.7 aa 1 and 3 letter code table worksheet

Lists the one and three letter amino acid assignments, since we have not memorized them.

4.8 FP Taxonomy

We extracted taxonomic information from the NCBI-Entrez taxonomy database (http://www.ncbi.nlm.nih.gov/Taxonomy/Browser/wwwtax.cgi) for many of the FP expressing organisms.

4.9 Fluorescent protein nomenclature

There is no consensus nomenclature for naming chromoproteins or fluorescent proteins isolated from the wild or the lab. As a suggestion, if the sequence of all new FPs were deposited in Genbank (or corresponding databases), then each unique FP could at least be referred to by its accession number. Organism codon optimized forms could be referred to as, for example, Genbank # (human codon optimized M62653). EGFP is Genbank AAB02572 (pEGFP-1), and authors could cite this Genbank number plus the Genbank number for the cloning vector used (ex. AAB02575 for pEGFP-N1, AAB02576 for pEGFP-C1).

At the least, when using a new name, for example, zRFP574, the original name (zoanRFP), and Genbank accession number (if available) should by cited, along with the original publication (Labas et al 2002; #117).

4.10 Final thoughts

The Thousand Proteins of Light dataset was made to complement our PubSpectra / Fluorescence Spectra Graphing websites (McNamara et al 2006). We will add new fluorescent protein data and spectra to their respective projects as they get published or as authors send us data. We encourage readers to set up, post online, and publish, datasets of areas of interest. We look forward to the next thousand proteins of light.

References (cited)

[1] M. Chalfie, Y. Tu, G. Euskirchen, W.W. Ward, D.C. Prasher (1994) Green fluorescent protein as a marker for gene expression. *Science* 263:802-805.
[2] de Wet JR, Wood KV, Helinski DR, DeLuca M (1985) Cloning of firefly luciferase cDNA and the expression of active luciferase in Escherichia coli. *Proc Natl Acad Sci* USA 82: 7870-7873.
[3] N.G. Gurskaya, A.F. Fradkov, A. Terskikh, M.V. Matz, Y.A. Labas, V.I. Martynov, Y.G. Yanushevich, K.A. Lukyanov, S.A. Lukyanov (2001) GFP-like chromoproteins as a source of far-red fluorescent proteins. *FEBS Lett* 507: 16-20.
[4] S. Inouye, S. Sasaki (2007) Imidazole-assisted catalysis of luminescence reaction in blue fluorescent protein from the photoprotein aequorin. Biochem Biophys Res Commun 354: 650-655.
[5] M.V. Matz, A.F. Fradkov, Y.A. Labas, A.P. Savitsky, A.G. Zaraisky, M.L. Markelov, S.A. Lukyanov (1999) Fluorescent proteins from nonbioluminescent *Anthozoa* species. *Nat. Biotechnol* 17: 969-973.
[6] G. McNamara, A. Gupta, J. Reynaert, T.D. Coates, C. Boswell (2006) Spectral imaging microscopy web sites and data. *Cytometry*: 69: 863-871.
[7] V. Pieribone, D. Gruber (2006) Aglow in the dark: The revolutionary science of biofluorescence. The Belknap Press of Harvard University Press, Cambridge, MA. 288 pages.
[8] D.C. Prasher, V.K. Eckenrode, W.W. Ward, F.G. Prendergast, M.J. Cormier (1992) Primary structure of the *Aequorea victoria* green-fluorescent protein. *Gene* 111: 229-233.
[9] Prasher D, McCann RO, Cormier MJ (1985) Cloning and expression of the cDNA coding for aequorin, a bioluminescent calcium-binding protein. *Biochem Biophys Res Commun* 126: 1259-1268.
[10] N.C. Rockwell, Y.S. Su, J.C. Lagarias (2006) Phytochrome structure and signaling mechanisms. *Annu Rev Plant Biol* 57: 837-858. (Data table from Supplemental material).
[11] Shaner NC, Steinbach PA, Tsien RY. (2005) A guide to choosing fluorescent proteins. *Nat Methods* 2: 905-909.
[12] O. Shimomura (2006) Bioluminescence. Chemical principles and methods. World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 500 pages.
[13] R.Y. Tsien (1998) The green fluorescent protein. *Annu Rev Biochem* 67: 509-544.
[14] J.R. Wagner, J.S. Brunzelle, K.T. Forest, R.D. Vierstra (2005) A light-sensing knot revealed by the structure of the chromophore-binding domain of phytochrome. *Nature* 438: 325-331.
[15] M. Zimmer (2005) Glowing Genes: A Revolution In Biotechnology. Prometheus Books, Amherst, NY. 222 pages.