Interchim Innovations

Interbiotech - BioScience Innovations

M06

Green Fluorescent Proteins (GFP)

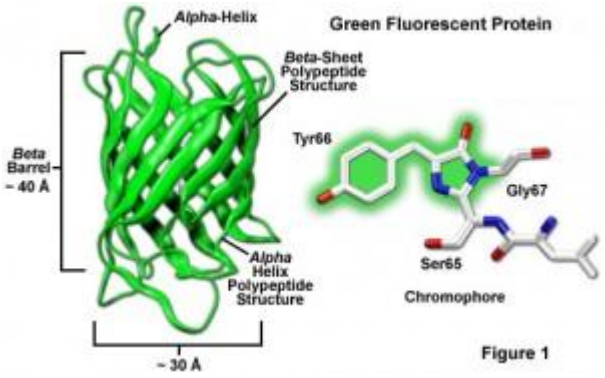


Figure 1

Split GFP and **SuperFolder GFP,** an innovative green fluorescent protein:

Green Fluorescent Protein (GFP) has been around for many years and has been used in a lot of creative ways. ***Split GFP and SuperFolder GFP*** allows for improved standard applications and innovative ones:

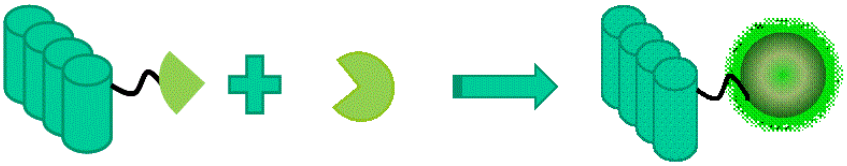
Quantify the expression level of a target protein, more accurately

Localize the expression of target protein in cells,with more natural distribution

Determine a target protein's solubility, **Discover** which domains of a protein are soluble

Evaluate how a protein interacts with other proteins (protein-protein interaction)

Reveal the effect of a small molecule on the protein's folding



A properly folded protein with exposed GFP S11 fragment

Self assembles to a GFP S1-10 detection fragment

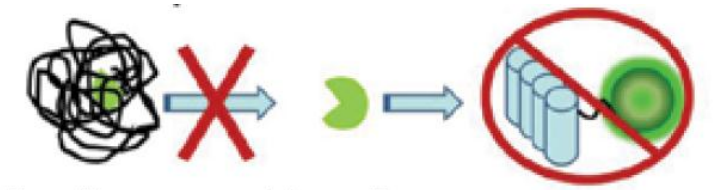
Resulting in a fluorescent GFP fusion protein

●Ordering information:

In Vitro Mammalian Optimized Split GFP Fold-N-Glow Solubility Assay - Mammalian S11 Single Plasmid	ref.22004003
In Vitro Mammalian Optimized Split GFP Fold-N-Glow Solubility Assay Kit	ref.25004001, 84 tests
In Vitro Bacterial Split GFP Fold-N-Glow Solubility Assay - Bacterial S11	ref.21004003
In Vitro Bacterial Split GFP Fold-N-Glow Solubility Assay Kit	ref.20004001, 84 tests
In Vitro Split GFP Fold-N-Glow Solubility Assay - Positive Control	ref.21004002
In Vitro Split GFP Fold-N-Glow Solubility Assay - Universal Detection Reagent	ref.21004001
Superfolder GFP Expression Plasmid	ref.23004006

●More information:

| Monitor Protein Expression | Determine Solubility | Find Soluble Domains
| Protein-Protein Interactions | Drug Discovery | References
| Comparison with other tagging methods | History



Une protéine anormalement repliée avec un fragment GFP S11 masqué

Auto-assemblage à un fragment de détection GFP S1-10 est empêché

Obtention d'une protéine non fluorescente

Comparison with other tagging methods

Technology	No false positives	Time <3hr	1-2 steps	Multiple colors	Applicable to Peptides	Eliminates Antibodies	Commercial
Split & SuperFolder GFP	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fluorescent Tags	No	n/a	No	Yes	No	No	Yes
Whole GFP	No	No	No	Yes	No	No	Yes
ELISA	No	NO	No	No	Yes	No	Yes

Detailed applications

Using GFP to Monitor Protein Expression (SuperFolder GFP, Split GFP)

Overview

A highly engineered version of Green Fluorescent Protein (GFP), can monitor the expression level of a target protein. Named "SuperFolder," this engineered GFP quantifies the expression level of a target protein because the amount of fluorescence is directly correlated with the amount of expressed protein. SuperFolder makes this possible because it folds and fluoresces no matter how poorly the target protein is expressed or how insoluble it is. If there is no fluorescence using SuperFolder, the researcher can **safely conclude that the protein is not being expressed**. Therefore, SuperFolder is a powerful tool to quickly and easily measure protein expression with a high level of confidence.

Split GFP can also be used to monitor protein expression that are properly expressed and conformed. Compared with full-length GFP, the smaller GFP11 tag **prevents non-natural distribution of proteins in cells** (see H-Ras in figure A)

Key Characteristics

- Robust
- Reliable
- Quantifiable
- Fast and easy
- Minimize perturbations to protein folding and distribution in cells

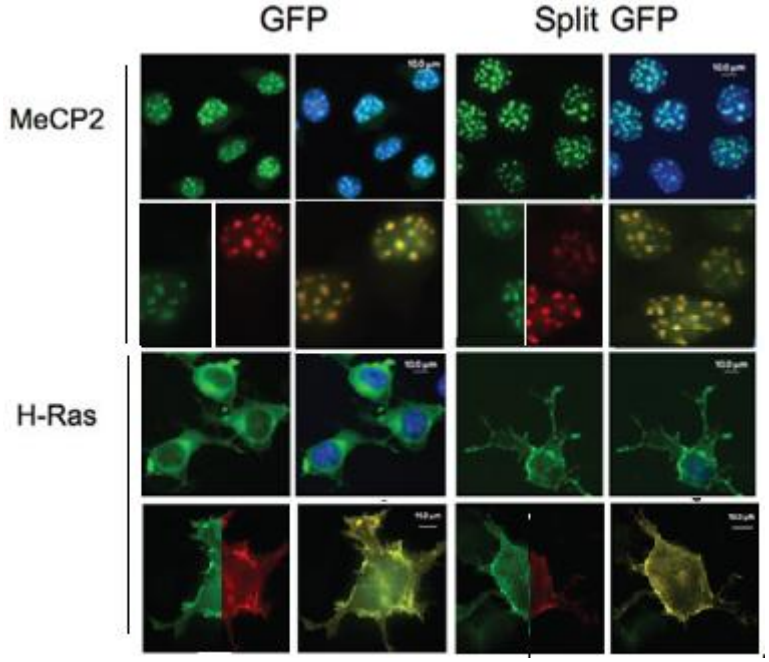


Figure A [1]: GFP(1-10) staining for studying protein localization by fluorescence microscopy.
Upper pannels: Two proteins, MeCP2 and H-Ras, were expressed in N2A cells, as GFP fusions (left panels) or GFP 11 fusions (right panels). MeCP2–GFP 11and GFP 11–H-Ras were stained with recombinant GFP 1-10 reagent before mounting on microscope slides.
Left image, green fluorescence at 488-nm excitation. Upper right image, overlay of green fluorescence and DAPI nuclear staining (blue). Scale bars, 10 µm.

Lower panels: Double staining experiments
Left: Anti-MeCP2 sera (half right field, red) and GFP 1-10 staining (half left field, green) were performed on N2A-MeCP2-GFP 11 cells and compared with N2A cells expressing MeCP2-GFP fusions.
Right: similar staining with anti-H-Ras on HEK–GFP 11–H-Ras and GFP–H-Ras HEK cells.
FITC emission channel at 530 nm with excitation at 488 nm (green), rhodamine emission channel detected at 590 nm with excitation at 545 nm (red), superimposition of both images (right/yellow).

"The split GFP detection system is particularly well-suited for protein tagging and detection in eukaryotic cells using multiple formats. Localization and quantitative expression can be simultaneously performed either in fixed models or in living cells with transient or stable expression of GFP 1-10.
The main advantages over existing epitope tags are the **high specificity** and **quantitative recognition** between GFP 11 and GFP 1-10 fragments and the **absence of fluorescence of the GFP 1-10 protein**. This confers **very low background** signals and facilitates staining procedures, as it **does not require extensive washing steps** compared with classical immunostaining methods. The small size of the GFP 11 fragment (15 amino acids) should be **less perturbing than the bulky GFP**, and GFP 1-10 staining can be performed in combination with other immunostaining procedures as for GFP. The versatility of the system will be further enhanced by the possibilities of combining the split GFP method with chromatic variants of GFP, such as cyan and yellow"

"For GFP–H-Ras, we noticed the presence of additional intracellular vesicular structures in both cell types, as previously reported showing that fusion of H-Ras to the whole GFP molecule resulted in its redistribution toward the ER and Golgi membrane, whereas unmodified H-Ras was found mostly at the plasma membrane (17). … the smaller size of the GFP 11 tag may alter less than full-length GFP, the natural distribution of the tagged protein,… the staining with split GFP is nonperturbing and correlates accurately with true protein localization, as evidenced by double staining experiments"

References● Protein Expression Level [Top]

Cava et al. "Expression and use of superfolder green flourescent protein at high temperatures in vivo; a tool to study extreme thermophile biology," *Environ Microbiol.* **10**(3), 605-13, March 2008.
Pédelacq et al. "Engineering and characterization of a superfolder green fluorescent protein,"*Nature Biotechnology* **24**(1), 79–88, January 2006.

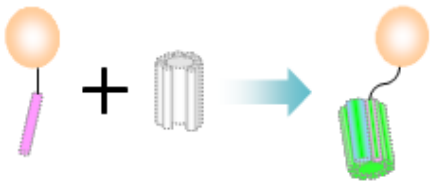
Using LANL GFP to Determine Solubility

Overview

LANL's GFP can be used to quickly determine whether an expressed protein is soluble. It does this very simply and elegantly. First, a short section of GFP called a "tether" or "tag" is appended onto the nucleotide sequence of the protein of interest. The tether is about 14 amino acids (a.k.a. S11, strand 11 of GFP) which is too short to effect the dynamics of your expressed protein.

Using split for solubility studies

Figure 1. Split GFP: how a soluble expressed protein will behave in the system.



A vector containing the remaining portion of GFP (a.k.a S1-10, strand 1-10 of GFP) called the "detector" is also inserted into the host cell. Once the target protein and S11, and the remaining portion of GFP, S1-10 are inserted into the host, determining the solubility of the target protein is easy: express your protein of interest (which has the S11 tag) (Fig. 1, #1), followed by inducing the expression of the S1-10 "detector" (Fig. 1, #2).

Split non-soluble

Figure 2. Split GFP: how an insoluble expressed protein will behave in the system.

If your protein is soluble, aggregation will not occur (Fig. 2, #1) and the S1-10 detector will be able to bind to the S11 tag to create a fully functioning, fluorescing GFP (Fig. 1, #3). However, if the expressed protein is not soluble, it will aggregate (Fig. 2, #1), not allowing the S11 tag to interact with the S1-10 detector and no fluorescense will occur (Fig. 2, #3).

Significant efforts have been directed into an automated, robotic-based system to process hundreds of protein samples in a parallel. Please inquire.

Key Characteristics

- Fast
- Cheap
- Reliable
- Quantifiable
- Automation and high-throughput sample processing

References* ● *Protein Solubility ***[Top]***

Cabantous et al. "New Molecular Reporters for Rapid Protein Folding Assays," PLOSone. June 11, 2008.
Chun et al. "Split GFP complementation assay: a novel approach to quantitatively measure aggregation of tau in situ: effects of GSK3beta activation and caspase 3 cleavage"J Neurochem. **103**(6):2529-39, December 2007.
Andrews BT, Schoenfish AR, Waldo GS, Jennings PA "The rough energy landscape of sfGFP is linked to the chromophore" (2007) J Mol Biol. **2**:476-90, October 2007
Cabantous and Waldo. "In vivo and in vitro protein solubility assays using split GFP," Nature Methods **3**(10), 845–854, October 2006.
Cabantous et al. "Recent advances in GFP folding reporter and split-GFP solubility reporter technologies. Application to improving the folding and solubility of recalcitrant proteins from Mycobacterium tuberculosis," Journal of Structural and Functional Genomics **6:**113–119, 2005.
Cabantous et al. "Protein tagging and detection with engineered self-assembling fragrments of green fluorescent protein," Nature Biotechnology **23,** 102-107, December 2004.
Yang et al. "Directed evolution approach to a structural genomics project: Rv2002 from Mycobacterium tuberculosis," PNAS **100**: 455-460, 2003
Waldo GS. "Genetic screens and directed evolution for protein solubility" Curr Opin Chem Biol 7(1), 33-8. 2003.
Pédelacq et al. "Engineering soluble proteins for structural genomics," Nature Biotechnology **20,** 927-932, September 2002.
Waldo et al. "Rapid protein-folding assay using green flourescent protein," Nature Biotechnology **17**, 691 - 695, July 1999.

Using LANL GFP to find Soluble Domains in Proteins

Overview

The motivation behind inventing a fast reliable system for identifying the individual soluble domains in proteins was guided by the following facts:

° Many proteins have multiple domains within them;
° Individual domains often fold independently of the other domains within the same protein;
° Individual domains, rather than entire proteins, are easier investigate;
° While bioinformatics provides powerful tools to assist with identifying domains, it cannot reliably predict which domains will be optimally expressed.

LANL's solution is to use our highly evolved version of GFP to approach this problem differently. Waldo has designed a system which is based on a library of fragmented genes that are then expressed (Fig. 1, # 1 and #2). Using a version of our GFP to determine solubility, the resulting proteins from the library can then be rapidly screened to find those that are soluble (Fig. 1, #3).

Figure 1: Making protein analysis easier

Figure 1. Proteins are easier to investigate if they are divided into individual domains.

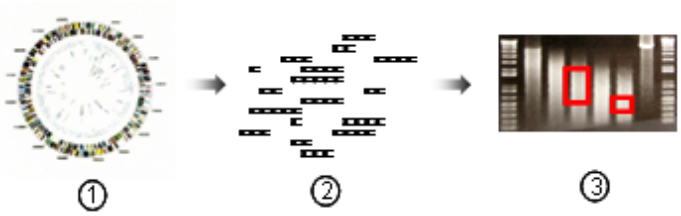
Our process involved four different steps that takes approximately seven business days to go from gene sequence(s) to identified soluble domains. This turnaround time is unbelievable considering the amount of time and effort other processes require to discover the same information. The four individual steps are described in great detail below.

Step 1 (Library size ≥107)

Figure 2: Step 1 of the LANL process

The first step of the LANL process is the fragmentation of a gene or genes of interest. Keep in mind, with the assistance of robotics, this process could easily be scaled to full genome analysis.

Figure 2. Step 1 of 4 of LANL's process for easily finding soluble domains. This step involves fragmenting gene sequences:

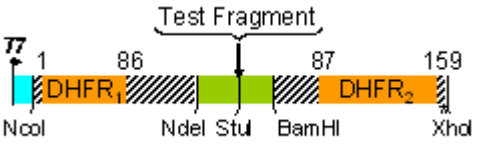


The first step of our process is to take a set of genes or ORFs and fragment them using any number of methods such as restriction enzyme treatment (e.g. DNAse) or mechanically disrupting them with shearing forces (Fig 2, #2). If the size is known of the gene or set of genes, the gene fragments are run on an agarose gel and the fragments of the correct size can be cut out and taken into the next step of the process for focused study.

●Step 2

Fragments are then blunt cloned and screened. However, since >95% of expressed proteins are not in the native frame and likely have a stop codon (Fig. 3, #1), the DNA fragments are cloned into the LANL ORF selector before spending valuable time and effort on the GFP screen. Using the LANL ORF selector, the test fragment is inserted between two halves of dihydrofolate reductase enzyme (DHFR) to determine which inserts are out-of-frame (Fig. 3).

Figure 3. DHFR insertion site for fragment screening:



If the inserted fragment is out of frame, it is likely to have one or more stop codons, so the second half of DHFR selection gene will not be expressed. The result is that when the host E. coli cells are plated on a low concentration of trimethoprim, the clones with out-of-frame constructs cannot survive. Clones without stop codons (the fragment is in-frame) can produce DHFR to metabolize trimethoprim and survive (Fig. 4, image 2). These plasmids are then recovered and investigated further (Fig. 4, image 3, the plasmids with green inserts).

Step 3

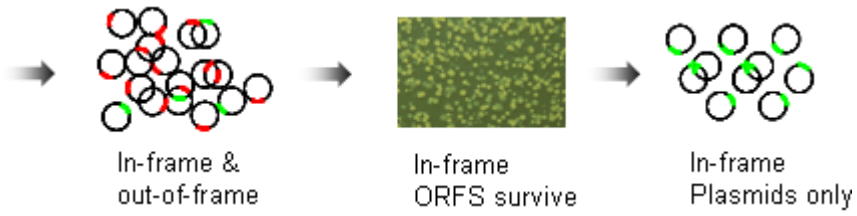


Figure 4. Step 2 of 4 of LANL's process for easily finding soluble domains.

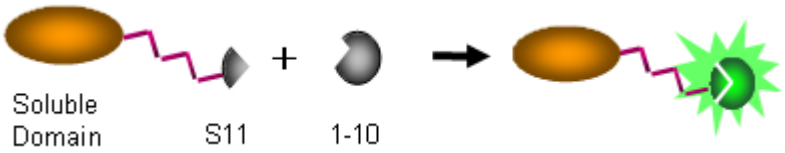
Keep in mind, in this step solubility has not been determined yet. This step simply removes those DNA fragments which do not encode authentic protein domains.

●Step 3 (Library size approx. 10^5)

Once the collection of fragments that are known to express viable protein domains is identified, the in-frame inserts can then be cloned into LANL's in vivo Split GFP system to determine which in-frame, expressed domains are soluble (Fig. 4, #1). The Split GFP constructs are transformed and grown in E. coli and single colonies are screened by plating on agar plates. First, expression of the S11-tagged protein is induced, followed by induction of the GFP detector fragment (S1-10)(Fig. 5, left).

Figure 5. Step 3 of 4 of LANL's process for easily identifying soluble domains:

Under these conditions, brighter the clones indicate that the S11-tagged, soluble protein is interacting with the S1-10 detector fragment and fluorescence is being produced. The brightest clones are then picked from the agar plates and grown in 96 well liquid cultures. By inducing only the AnTET promoter in the 96-well plate cultures, only the S11-tagged protein is expressed. The cells are lysed, and the soluble and insoluble protein products quantified by adding (not inducing the expression of) the in vitro S1-10 detector fragment.



Briefly, the Split GFP system is based on fragmenting a highly modified GFP into two separate, soluble pieces: GFP strand 11 (S11), also known as the "tether" of "tag" and GFP strands 1-10 (S1-10), known as the "detector fragment" (Fig. 6).

Figure 6. The LANL Split GFP system.

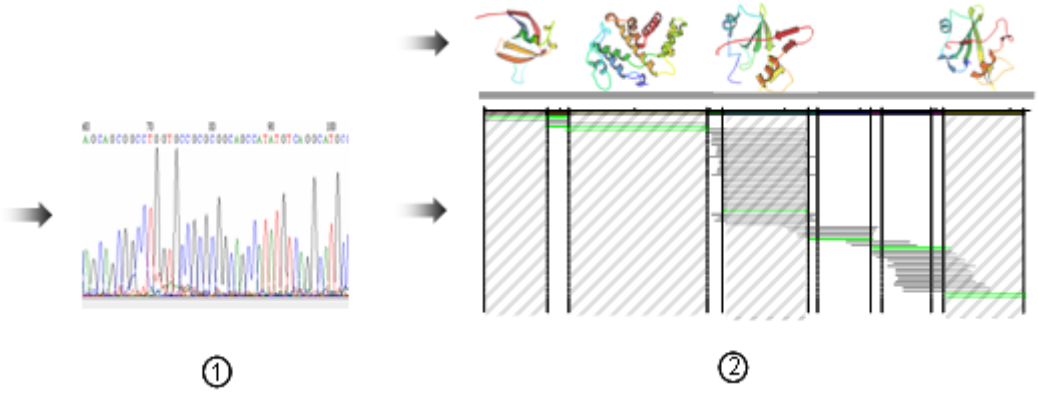
First, the S11 tagged protein domain is expressed using AnTET induction. Then the S1-10 detector fragment is expressed using IPTG induction. If the expressed domain is soluble, it will allow the S11 "tether" to interact with the larger GFP fragment, S1-10. The interaction of S11 and S1-10 allows GFP to fluoresce (Fig. 6).

The beauty of the LANL GFP system is that there is no need to reclone the construct to quantify expression. To express the S11 tag alone, one simply adds only AnTET which induces expression of the S-11 tagged protein and the S1-10 can be added as an in vitro reagent!

●Step 4 (Library size 10^2)

In the final step of LANL's process, clones of interest are selected from the 96-well plate used in Step 3 and are sequenced (Fig. 7, # 1). At LANL, typically the entire 96-well plate is sequenced so that the position of all the fragment clones can be mapped to the parent gene.

Figure 7. Step 4 of 4 of LANL's process for easily identifying soluble domains:



In silico, the sequence of the gene fragments are aligned onto the full parent gene and color coded by solubility. At LANL, we use the color scheme, red-orange-yellow-green-blue-bluish white, where the red side of the spectrum identifies the least soluble domains.). This makes finding the experimental domain boundaries very clear and makes it easy to identify compact members of the groups (a.k.a the minimal tiling path) which are subcloned for scale up (Fig. 7, #2). For crystallographic applications, the clone is subcloned without the S11 tag, or the tag can be included as a way to track purification, and then cleaved off.

●Crystallizing a 2200 aa protein from Mycobacterium tuberculosis

If the intent is to crystallize the identified soluble domains as we did with a 2200 aa protein called ssPC, LANL recommends selecting multiple compact versions of each soluble domain to increase the probability of finding crystallizable, diffracting constructs (Typical yields 15-30 mg/l, concentrate to >40 mg/ml). Figure 8 shows an SDS gel of some of the most compact (smallest in set) soluble fragments or soluble domains representing the 6 predicted domains of ppSC, subcloned into a pET N-terminal 6HIS vector. Soluble (S), insoluble pellet fraction (P) of E. coli lysates.
LANL's Split GFP

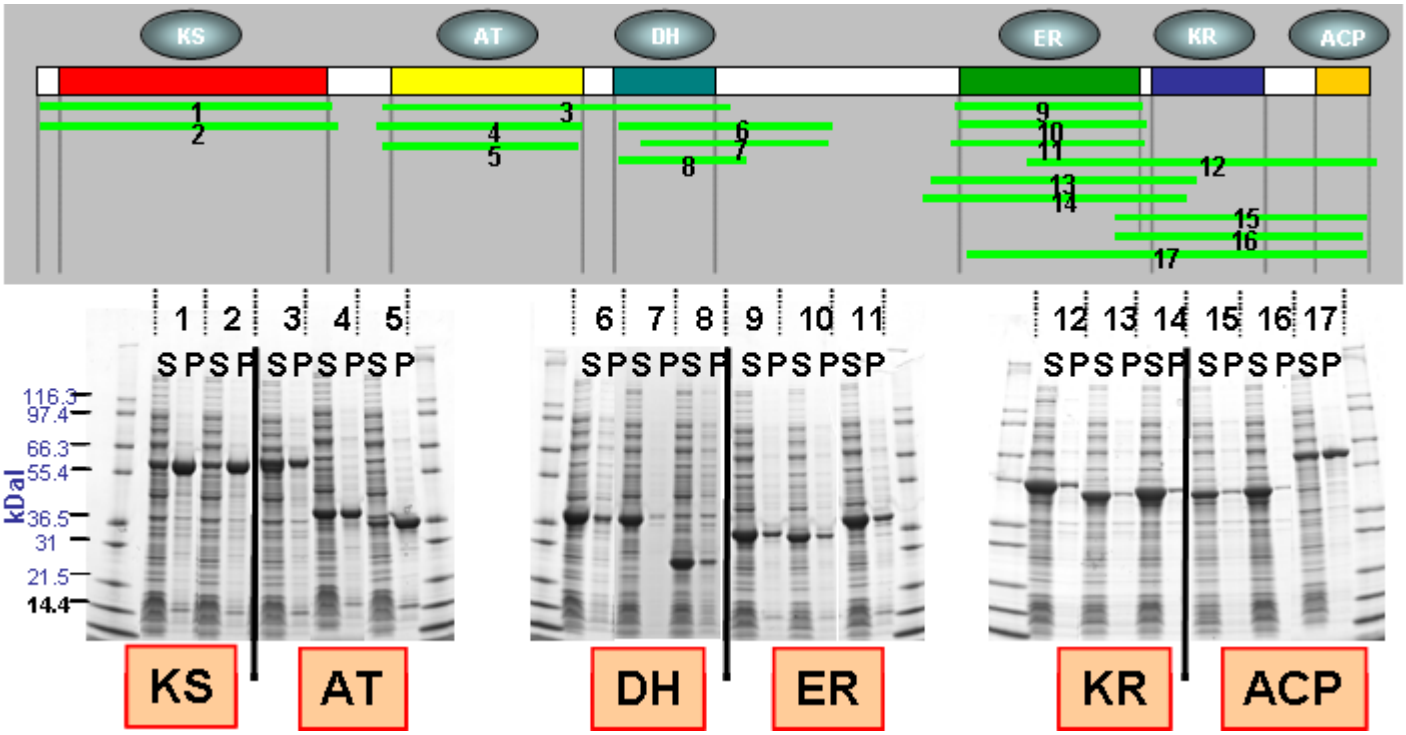


Figure 8. For crystallization, select multiple compact versions of each domain.

One of the advantages of having a dense sampling of fragment position and solubility is that it is easy to identify the boundaries of a domain. As fragments get progressively shorter, they become soluble near the boundary domains, then suddenly become less soluble as the fragments are further truncated. This makes it easier to select a small subset of compact clones for detailed study. Shown in Fig. 9 are some compact 'double domain' constructs containing the KR+ACP domains of the large polyketide synthetase from Mycobacterium tuberculosis (M. tb). The protein is nicely monodisperse by gel filtration chromatography and runs as a monomer without significant aggregation at >10 mg/ml. LANL's Split GFP

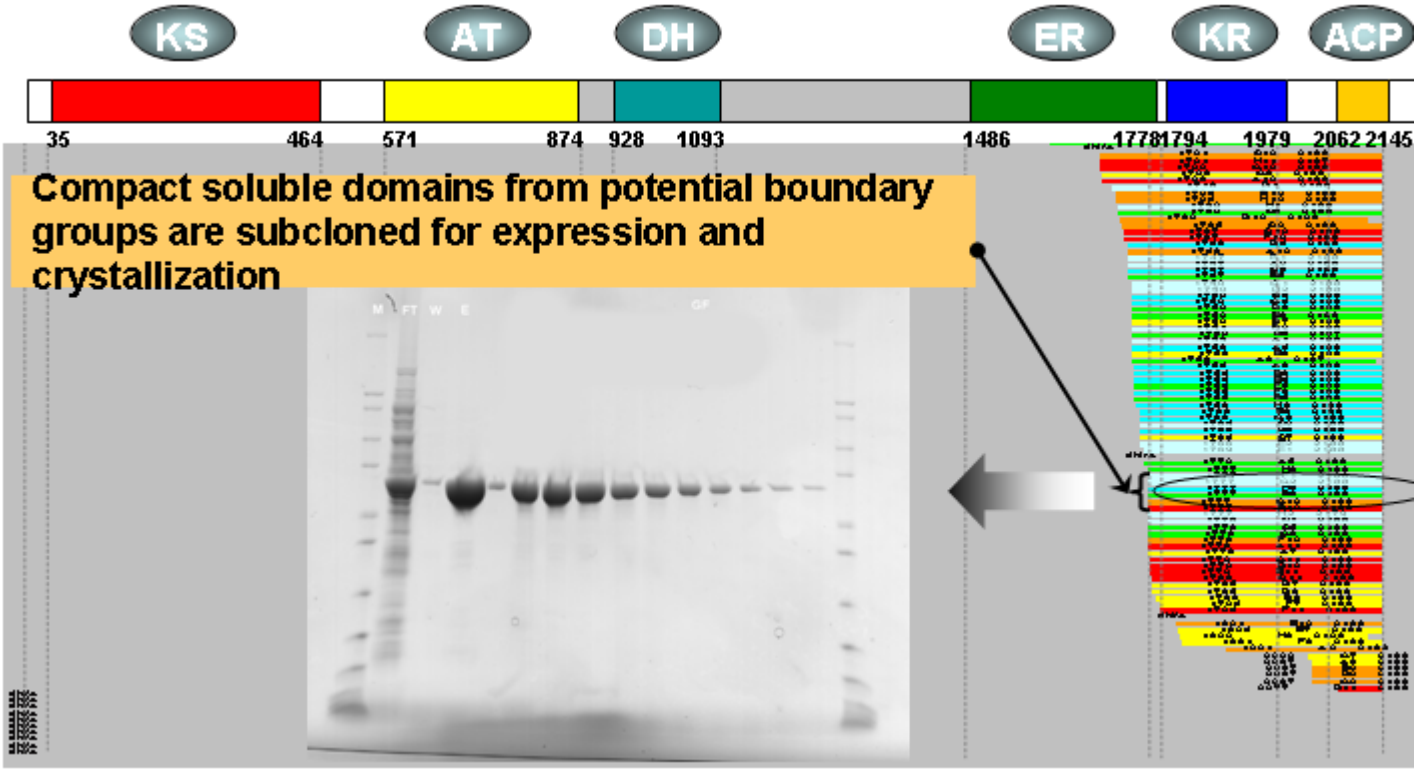


Figure 9. For crystallization, select multiple compact versions of each domain.

Compact clones are less likely to contain disordered ends, and having several choices near a given boundary size increases the chance that at least one will crystallize and give diffraction-quality crystals. The split GFP domain trapping protocol readily identifies two sets of fragments, one focusing on a larger version of the ER-containing domain from ca. 1480-1755, and another more compact version focusing down to amino acids 1558-1750. This fragment crystallized and is very similar to the previously published construct Shapiro et. al used to solve the structure (1PQW). (Fig. 10)

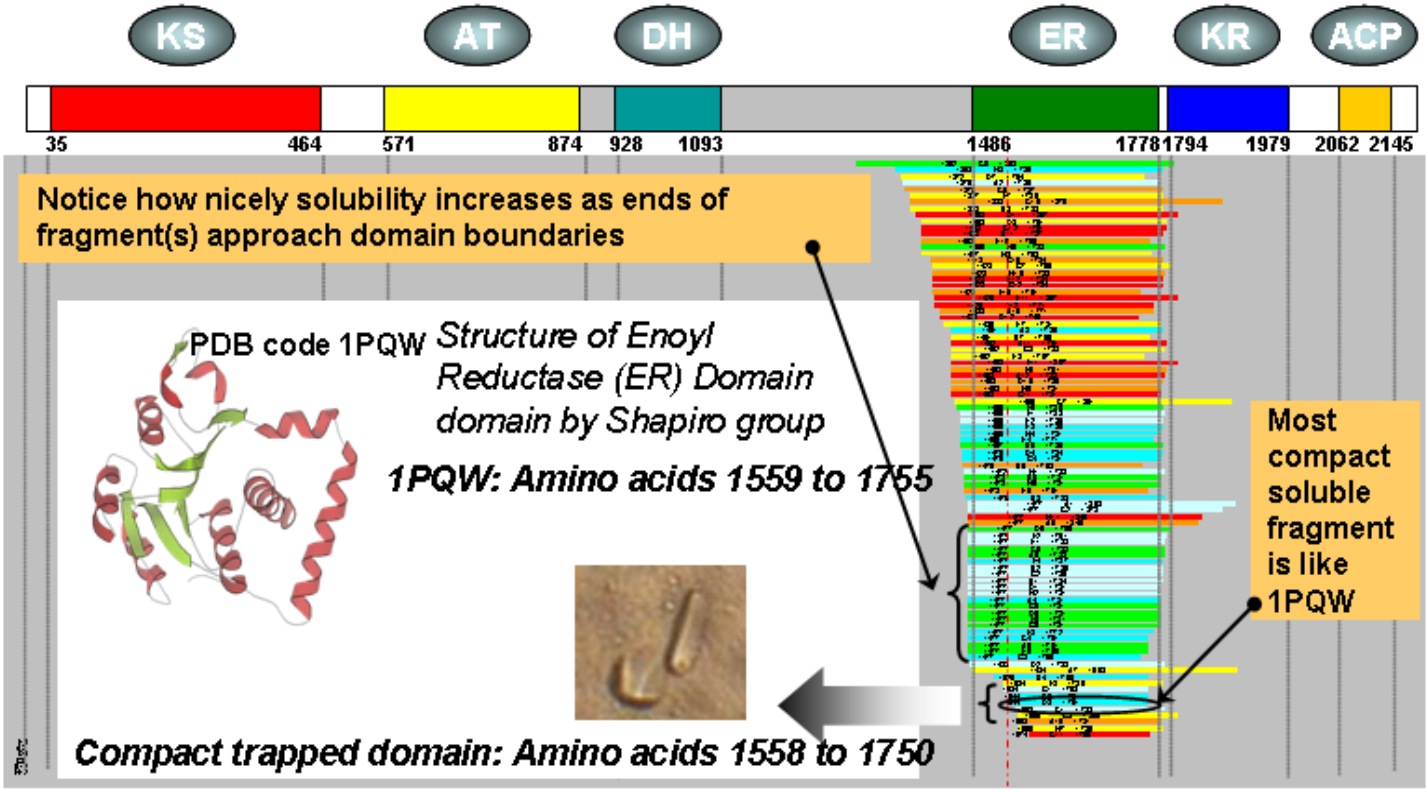


Figure 10. For crystallization, select multiple compact versions of each domain.

Using LANL's GFP to determine protein-protein interactions

Overview

Typical protein complementation assays (PCA) work by spliting a reporter molecule such as a fluorescent protein or enzyme into two pieces (see Michnick's work). Alone, the pieces cannot fold and are inactive. However, if the pieces are attached to interacting proteins, their interaction forces the two pieces of the reporter molecule together and its activity is restored. This sounds appealing in principle but is difficult in practice. Key to understanding the limitations is understanding that typical reporter fragments generated by splitting a reporter molecule, such as GFP, are relatively large, have poor solubility, and can perturb the normal behavior of the attached test proteins (described by Lynne Regan et al.). In the typical case for GFP, the protein is split into two large pieces that cannot spontaneously fold and aggregate on their own (see Fig. 1, #1 and #2).

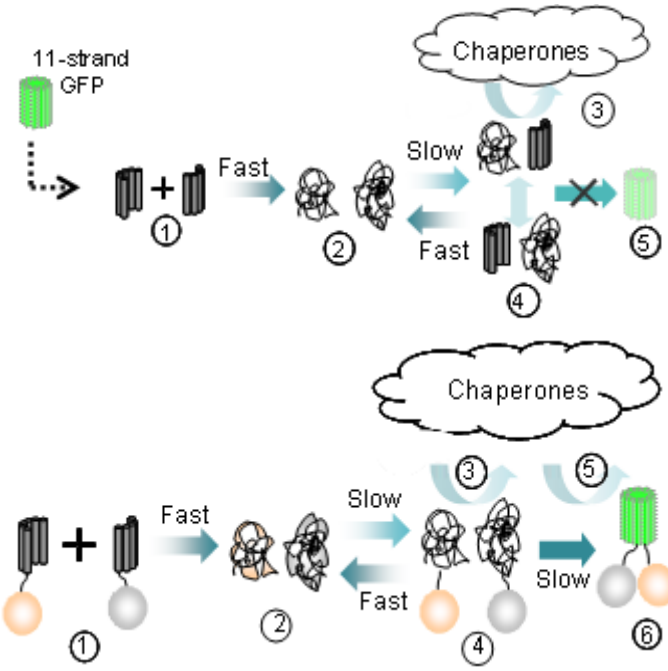


Figure 1. GFP that is randomly split does not spontaneously recombine.

Chaperones try to refold them, but each piece unfolds and aggregates before it "finds" its compliment (Fig. 1, #3), and the GFP fragments are not usually close enough together to recombine (Figure 1, #4). As a result, very little of the active, fluorescent GFP forms (Figure 1, #5). Many researchers hope that by attaching the GFP fragments to potentially interacting "passenger" proteins (Fig. 2, #1), they will overcome the solubility issues with the GFP fragments. Unfortunately, this is not the case. With interacting passenger proteins attached to the cumbersome GFP fragments, folding interference leads even more strongly to misfolding, not only of the GFP pieces but also the fused passengers. (Fig. 2, #2).

Figure 2. GFP that is randomly split does not spontaneously recombine even when attached to interacting "passenger" proteins.

In a small number of cases, chaperones will succeed in refolding the passenger domain(s), but the process is very inefficient and depends entirely on the host cell's folding machinery (Fig. 2, #3). As a result, a small percentage of the correctly folded proteins can interact, which brings the two fragments of GFP close enough to recombine (Fig. 2, #5). Since folded GFP is stable, it remains folded and eventually (slowly) the GFP builds up in a way consistent with Le Chatelier's Principle and can be detected (Fig. 2, #6).

LANL's protein interaction detector is different than any others currently available. Instead of large, poorly soluble fragments of GFP, we have engineered small "tethers" or "tags" (fragments of GFP that are 14 amino acids long) to tag the interacting test proteins. Detection of interacting proteins is then accomplished with a third fragment of GFP, the "detector".

The general idea of the LANL approach is presented in Figure 3. Initial work was done to express GFP as two domains (Fig. 3, #1). The first domain, a hairpin structure, interacts instantly with the second expressed domain, the GFP strand 1-9 "detector" (Fig. 3, #2) to form a fully functional GFP (Fig. 3, #3). Taking the next step, expressing the first domain as two separate strands (Fig. 4, #5) does not create a funtional GFP even with the expression of the "detector" (Fig. 3, #5).

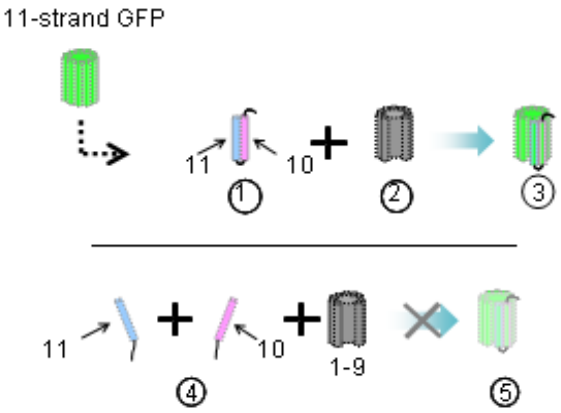
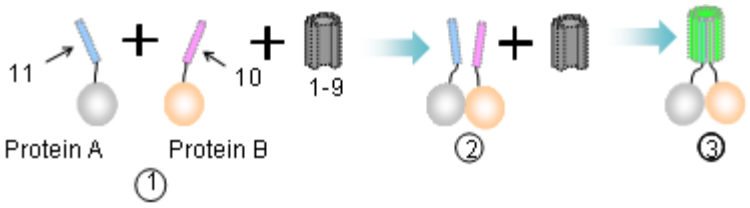


Figure 3. LANL's novel approach with GFP:

The LANL approach is applied to protein-protein interactions by doing the following. Strand 10 of GFP(S10) is attached to the first test protein, say protein A. Strand 11 of GFP (S11) is attached to the second test protein, call it protein B in this example (Fig. 4, #1). Interaction of protein A with protein B is detected using GFP strands 1-9, the "GFP 1-9 detector." This system works by reducing entropy. Normally, S10 and S11 don't interact by themselves and the so-called 'three-body interaction' between S10, S11, and 1-9 is not energetically favorable and cannot happen.



If test proteins A and B interact with each other, they bring the attached S10 and S11 close enough together to interact with the GFP 1-9 detector (Fig. 4, #2). Essentially, the interaction of protein A and protein B converts the energetically unfavorable three-body problem into an energetically favorable two-body problem (Fig. 4, #3).

Figure 4. LANL's GFP approach reduces a 3-body system to a 2-body system.

A researcher can do this by following these simple steps. First, a vector is created that includes short nucleotide section of GFP called a "microdomain" tag (strand 10 of GFP, or S10) appended onto the nucleotide sequence of the first protein of interest (protein X), to make protein X-S10. The vector also includes a different nucleotide section of GFP (strand 11 of GFP, or S11) appended onto the nucleotide sequence of another protein (protein Y). Protein Y is expressed from the same plasmid at a second ribosome binding site to give protein Y-S11. The vector is transformed into a host cell, for example E. coli, containing a second plasmid expressing the GFP 1-9 detector. Determing whether protein X is interacting with protein Y is easily done by expressing the tagged proteins at the same time as the GFP 1-9 detector. If fluorescence is observed, the two proteins interact.

This same transformation can be used to determine the solubility of the interacting proteins by staggering their expression. This is done in a similar way as LANL's solubility assay. By inducing expression of the tagged proteins first, shutting their expression off, and then inducing the expression of the GFP 1-9 detector, observing fluorescence indicates that the proteins not only interact, but they are soluble as well.

Expanding this approach, it would be easy to create a library that contains thousands of target proteins to determine if they interact with thousands of other proteins.

References ● Protein Interactions ***[Top]***

Nature Network: Neuroscience Topic "JOURNAL CLUB: Getting a GRASP on synapse location". March 2008.
Feinberg et al. "GFP reconstitution across synaptic partners (GRASP) defines cell contacts and Synapses in living nervous systems," *Neuron*. **57**(3) 353-363, Feb 2008.

Using LANL GFP for Drug Discovery

Overview

Another valuable use of LANL's GFP technology is to determine the effects of other molecules (small molecules, antibodies, aptamers, etc.) on a specific protein of interest. Determining such an effect could be done a couple of different ways depending upon which GFP technology will be employed. For example, if a researcher wants to study a protein that is naturally insoluble due to genetic mutation, that protein could be transformed into E. coli and expressed in a way similar to that described on the Solubility page of this website, which describes a split version of GFP. However, instead of immediately expressing the S1-10 tag to complement the GFP tether, one could treat the cell with a library of small molecules first, then express S1-10. If a protein flouresces after treatment of a particular small molecule, whereas it did not before the treatment, it is safe to assume that small molecule has helped to correct the protein's solubility in some way.

In a related way, our Folding Reporter could also be used to determine effects of other molecules on a protein of interest. The Folding Reporter reports on the success of a target protein's folding and solubility. When LANL's engineered version of GFP is attached to the C-Terminus of a target protein, it mimics the folding success of that target protein. Thus, when the target protein folds correctly, GFP folds correctly and fluoresces. Conversely, when the target protein folds incorrectly, GFP folds incorrectly and fails to fluoresce. The Folding Reporter mimics the folding state of the target protein and the level of fluorescence is correlated with how well the target protein is folded. At LANL, Folding Reporter's key application is evolving proteins for solubility in a high throughput format. But it could easily be used to determine the effect a particular cell treatment has on the target protein in a fast, high-throughput way. A researcher would simply need to place cells in a 96- or even 384-

well format, treat each well with a different compound and quickly determine improved flourescence. In fact, this approach has been proven successful by Michael Hecht's group at Princeton doing research on Alzheimer's (see the technical library for more information). It is also being pursued by Dr. Greg Philipps at the University of Iowa to investigate RNA-protein interactions in vivo in E. coli (read the user feedback page). Please see our technical library for a long list of citations related to our Folding Reporter.

Key Characteristics

- Robust
- Amenable to high-throughput
- Easy to determine increased solubility visually

[Top]

References● Drug Discovery

Barakat et al. "Molecular diversity in engineered protein libraries," *Chem Biol.* **11**:335–341, June 2007.
Kim et al. "A high-throughput screen for compounds that inhibit aggregation of the Alzheimer's peptide," *ACS Chem Biol*. **1**(7), 461-9, Aug. 2006
Zaragoza and Hines. "Understanding Proteins,"*Innovation*, Oct/Nov 2006.

[Top]

References ● Related Articles

Protein Complimentation Assays (PCA)

Remy et al. "Application of a protein-fragment complimentation assay in cell biology," *Biotechniques*. **42**(2):137, 139, 141 passim. February 2007.

Remy et al. "Detection of protein-protein interactions using a simple survival protein-fragment complementation assay based on the enzyme dihydrofolate reductase," *Nat Protoc*. **2**(9):2120-5, 2007.

Protein-Protein Interactions

Magliery TJ et al.. "Detecting protein-protein interactions with a green fluorescent protein fragment reassembly trap: Scope and mechanism," *J of Am Chem Soc* **127**(1):146-157, 2005.

Others

Bellinzoni, M. and G. Riccardi (2003). "Techniques and applications: The heterologous expression of Mycobacterium tuberculosis genes is an uphill road." Trends in Microbiology **11**(8): 351-358.

Wurth, C., W. Kim, et al. (2006). "Combinatorial approaches to probe the sequence determinants of protein aggregation and amyloidogenicity." Protein and Peptide Letters **13**(3): 279-286.

Delagrave, S. and D. J. Murphy (2003). "In vitro evolution of proteins for drug development." Assay and Drug Development Technologies **1**(1): 187-198.

Jayathilaka, L. P., M. Deb, et al. (2004). "Asymmetric synthesis and translational competence of L-alpha-(1-cyclobutenyl)glycine." Organic Letters **6**(21): 3659-3662.

Sorensen, H. P. and K. K. Mortensen (2005). "Soluble expression of recombinant proteins in the cytoplasm of Escherichia coli." Microbial Cell Factories **4**: 1.

Zhou, C. Z. and Y. X. Chen (2004). "Developments in structural genomics: Protein purification and function interpretation." Current Genomics **5**(1): 37-48.

Matsuura, T. and T. Yomo (2006). "In vitro evolution of proteins." Journal of Bioscience and Bioengineering **101**(6): 449-456.

Mittl, P. R. E. and M. G. Grutter (2001). "Structural genomics: Opportunities and challenges." Current Opinion in Chemical Biology **5**(4): 402-408.

Chambers, S. P. (2002). "High-throughput protein expression for the post-genomic era." Drug Discovery Today **7**(14): 759-765.

Kuhn, P., K. Wilson, et al. (2002). "The genesis of high-throughput structure-based drug discovery using protein crystallography." Current Opinion in Chemical Biology **6**(5): 704-710.

Heinemann, U., G. Illing, et al. (2001). "High-throughput three-dimensional protein structure determination." Current Opinion in Biotechnology **12**(4): 348-354.

Heinemann, U. (2001). The Berlin "Protein structure factory" initiative: A technology-oriented approach to structural genomics

Workshop on Data Mining in Structural Biology. Ernst Schering Research Foundation Workshop SO - Ernst Schering Research Foundation Workshop. Data mining in structural biology; 2001; v.34; p.101-121

Ernst Schering Research Foundation Workshop 34. I. Schlichting, U. Egner, I. Schlichting and U. Egner. BERLIN, GERMANY, 175 Fifth Avenue, New York, NY, 10010-7858, USA : Springer-Verlag New York Inc., 2001 **:** 101-121.

Sorensen, H. P. and K. K. Mortensen (2005). "Advanced genetic strategies for recombinant protein expression in Escherichia coli." Journal of Biotechnology **115**(2): 113-128.

Magliery, T. J. and L. Regan (2004). "Combinatorial approaches to protein stability and structure." European Journal of Biochemistry **271**(9): 1595-1608.

Rubin-Pitel, S. B. and H. M. Zhao (2006). "Recent advances in biocatalysis by directed enzyme evolution." Combinatorial Chemistry & High Throughput Screening **9**(4): 247-257.

Derewenda, Z. S. (2004). "The use of recombinant methods and molecular engineering in protein crystallization." Methods (Amsterdam) **34**(3): 354-363.

Schmidt-Dannert, C. (2001). "Directed evolution of single proteins, metabolic pathways, and viruses." Biochemistry **40**(44): 13125-13136.

Terwilliger, T. C., M. S. Park, et al. (2003). "The TB structural genomics consortium: A resource for Mycobacterium tuberculosis biology." Tuberculosis (Amsterdam) **83**(4): 223-249.

Stout, T. J., P. G. Foster, et al. (2004). "High-throughput structural biology in drug discovery: Protein kinases." Current Pharmaceutical Design **10**(10): 1069-1082.

Buchanan, S. G., J. M. Sauder, et al. (2002). "The promise of structural genomics in the discovery of new antimicrobial agents." Current Pharmaceutical Design **8**(13): 1173-1188.

Bradbury, A. R. M. and J. D. Marks (2004). "Antibodies from phage antibody libraries." Journal of Immunological Methods **290**(1-2): 29-49.

Brenner, S. E. (2001). "A tour of structural genomics." Nature Reviews Genetics **2**(10): 801-809.

March, J. C., G. Rao, et al. (2003). "Biotechnological applications of green fluorescent protein." Applied Microbiology and Biotechnology **62**(4): 303-315.

Buchanan, S. G. (2002). "Structural genomics: Bridging functional genomics and structure-based drug design." Current Opinion in Drug Discovery & Development **5**(3): 367-381.

Sakanyan, V. (2005). "High-throughput and multiplexed protein array technology: Protein-DNA and protein-protein interactions." Journal of Chromatography B: Analytical Technologies in the Biomedical and Life Sciences **815**(1-2): 77-95.

Zhang, J., Z. Liu, et al. (2003). "Three-dimensional localization of divergent region 3 of the ryanodine receptor to the clamp-shaped structures adjacent to the FKBP binding sites." Journal of Biological Chemistry **278**(16): 14211-14218.

Liu, Z., J. Zhang, et al. (2002). "Three-dimensional reconstruction of the recombinant type 2 ryanodine receptor and localization of its divergent region 1." Journal of Biological Chemistry **277**(48): 46712-46719.

Hardt, M. and R. A. Laine (2004). "Mutation of active site residues in the chitin-binding domain ChBD(ChiAl) from chitinase A1 of Bacillus circulans alters substrate specificity: Use of a green fluorescent protein binding assay." Archives of Biochemistry and Biophysics **426**(2): 286-297.

Pengelley, S. C., D. C. Chapman, et al. (2006). "A suite of parallel vectors for baculovirus expression." Protein Expression and Purification **48**(2): 173-181.

Fisher, A. C., W. Kim, et al. (2006). "Genetic selection for protein solubility enabled by the folding quality control feature of the twin-arginine translocation pathway." Protein Science **15**(3): 449-458.

Liu, S., R. C. Bugos, et al. (2001). "Green fluorescent protein as a secretory reporter and a tool for process optimization in transgenic plant cell cultures." Journal of Biotechnology **87**(1): 1-16.

Benedetti, C. E., J. Kobarg, et al. (2003). "Plasmodium falciparum histidine-rich protein II binds to actin, phosphatidylinositol 4,5-bisphosphate and erythrocyte ghosts in a pH-dependent manner and undergoes coil-to-helix transitions in anionic micelles." Molecular & Biochemical Parasitology **128**(2): 157-166.

Heinemann, U., J. Frevert, et al. (2000). "An integrated approach to structural genomics." Progress in Biophysics & Molecular Biology **73**(5): 347-62.

Offredi, F., F. Dubail, et al. (2003). "De novo backbone and sequence design of an idealized alpha/beta-barrel protein: Evidence of stable tertiary structure." Journal of Molecular Biology **325**(1): 163-174.

Hirano, N., T. Sawasaki, et al. (2006). "Tolerance for random recombination of domains in prokaryotic and eukaryotic translation systems: Limited interdomain misfolding in a eukaryotic translation system." Proteins-Structure Function and Bioinformatics **64**(2): 343-354.

Pattenden, L. K., A. P. J. Middelberg, et al. (2005). "Towards the preparative and large-scale precision manufacture of virus-like particles." Trends in Biotechnology **23**(10): 523-529.

de Groot, N. S., F. X. Aviles, et al. (2006). "Mutagenesis of the central hydrophobic cluster in A beta 42 Alzheimer's pepticle - Side-chain properties correlate with aggregation propensities." FEBS Journal **273**(3): 658-668.

Muddana, S. S. and B. R. Peterson (2003). "Fluorescent cellular sensors of steroid receptor ligands." Chembiochem **4**(9): 848-855.

Kim, C. A., M. L. Phillips, et al. (2001). "Polymerization of the SAM domain of TEL in leukemogenesis and transcriptional repression." EMBO (European Molecular Biology Organization) Journal **20**(15): 4173-4182.

Dyson, M. R., S. P. Shadbolt, et al. (2004). "Production of soluble mammalian proteins in Escherichia coli: identification of protein features that correlate with successful expression." BMC Biotechnology **4**: 32.

Wigley, W. C., R. D. Stidham, et al. (2001). "Protein solubility and folding monitored in vivo by structural complementation of a genetic marker protein." Nature Biotechnology **19**(2): 131-136.

Card, P. B. and K. H. Gardner (2005). Identification and optimization of protein domains for NMR studies. Nuclear Magnetic Resonance of Biological Macromolecules, Part C SO - Methods in Enzymology; 2005; v.394; p.3-16. T. James, 525 B STREET, SUITE 1900, SAN DIEGO, CA 92101-4495 USA : ELSEVIER ACADEMIC PRESS INC, 2005 **:** 3-16.

Pritchard, J. B. and D. S. Miller (2005). "Expression systems for cloned xenobiotic transporters." Toxicology and Applied Pharmacology **204**(3): 256-262.

Blundell, T. L., H. Jhoti, et al. (2002). "High-throughput crystallography for lead discovery in drug design." Nature Reviews Drug Discovery **1**(1): 45-54.

Stevenson, J. D. and S. J. Benkovic (2002). "Combinatorial approaches to engineering hybrid enzymes." Journal of the Chemical Society-Perkin Transactions 2(9): 1483-1493.

Cotlet, M., P. M. Goodwin, et al. (2006). "A comparison of the fluorescence dynamics of single molecules of a green fluorescent protein: One- versus two-photon excitation." Chemphyschem **7**(1): 250-260.

Rupp, B. and J. W. Wang (2004). "Predictive models for protein crystallization." Methods (Amsterdam) **34**(3): 390-407.

Zagrebelny, S. N. (2005). "Controlled evolution as an approach to the design of efficient biocatalysts." Uspekhi Khimii **74**(3): 307-320.

Yao, Y. X., J. Y. Ren, et al. (2003). "Amino terminal interaction in the prion protein identified using fusion to green fluorescent protein." Journal of Neurochemistry **87**(5): 1057-1065.

Zhang, F. J. and H. N. Zhu (2006). "Intracellular conformational alterations of mutant SOD1 and the implications for fALS-associated SOD1 mutant induced motor neuron cell death." Biochimica et Biophysica Acta **1760**(3): 404-414.

Chen, M. C., C. R. Lambert, et al. (2001). "Photoisomerization of green fluorescent protein and the dimensions of the chromophore cavity." Chemical Physics **270**(1): 157-64.

Rabin, S. D. P. and A. R. Hauser (2005). "Functional regions of the Pseudomonas aeruginosa cytotoxin ExoU." Infection and Immunity **73**(1): 573-582.

Schwalbach, G., A. P. Sibler, et al. (2000). "Production of fluorescent single-chain antibody fragments in Escherichia coli." Protein Expression and Purification **18**(2): 121-132.

Griffith, D. A., C. Delipala, et al. (2003). "A novel yeast expression system for the overproduction of quality-controlled membrane proteins." FEBS Letters **553**(1-2): 45-50.

Braud, S., M. Moutiez, et al. (2005). "Dual expression system suitable for high-throughput fluorescence-based screening and production of soluble proteins." Journal of Proteome Research **4**(6): 2137-2147.

Belin, P., J. Dassa, et al. (2004). "Toxicity-based selection of Escherichia coli mutants for functional recombinant protein production: application to an antibody fragment." Protein Engineering Design & Selection **17**(5): 491-500.

Zuo, X., M. R. Mattern, et al. (2005). "Expression and purification of SARS coronavirus proteins using SUMO-fusions." Protein Expression and Purification **42**(1): 100-110.

Hart, D. J. and F. Tarendeau (2006). "Combinatorial library approaches for improving soluble protein expression in Escherichia coli." Acta Crystallographica Section D Biological Crystallography **62**(Part 1): 19-26.

Cort, J. R., S. V. S. Mariappan, et al. (2001). "Solution structure of Pyrobaculum aerophilum DsrC, an archaeal homologue of the gamma subunit of dissimilatory sulfite reductase." European Journal of Biochemistry **268**(22): 5842-5850.

Pedelacq, J. D., S. Cabantous, et al. (2006). "Engineering and characterization of a superfolder green fluorescent protein." Nature Biotechnology **24**(1): 79-88.

Miyazaki, S., Y. Kuroda, et al. (2006). "Identification of putative domain linkers by a neural network - Application to a large sequence database." BMC Bioinformatics **7**: 323.

Yokoyama, S. (2003). "Protein expression systems for structural genomics and proteomics." Current Opinion in Chemical Biology **7**(1): 39-43.

Sacchetti, A., V. Cappetti, et al. (2001). "Green fluorescent protein variants fold differentially in prokaryotic and eukaryotic cells." Journal of Cellular Biochemistry: 117-128.

Kim, W. and M. H. Hecht (2005). "Sequence determinants of enhanced amyloidogenicity of Alzheimer A beta 42 peptide relative to A beta 40." Journal of Biological Chemistry **280**(41): 35069-35076.

Huang, D. and E. V. Shusta (2005). "Secretion and surface display of green fluorescent protein using the yeast Saccharomyces cerevisiae." Biotechnology Progress **21**(2): 349-357.

Japrung, D., S. Chusacultanachai, et al. (2005). "A simple dual selection for functionally active mutants of Plasmodium falciparum dihydrofolate reductase with improved solubility." Protein Engineering, Design and Selection **18**(10): 457-464.

Wall, M. A., M. Socolich, et al. (2000). "The structural basis for red fluorescence in the tetrameric GFP homolog DsRed." Nature Structural Biology **7**(12): 1133-1138.

Pedelacq, J. D., G. S. Waldo, et al. (2005). "Structural and functional features of an NDP kinase from the hyperthermophile crenarchaeon Pyrobaculum aerophilum." Protein Science **14**(10): 2562-2573.

Roodveldt, C., A. Aharoni, et al. (2005). "Directed evolution of proteins for heterologous expression and stability." Current Opinion in Structural Biology **15**(1): 50-56.

King, D. A., B. E. Hall, et al. (2006). "Domain structure and protein interactions of the silent information regulator Sir3 revealed by screening a nested deletion library of protein fragments." Journal of Biological Chemistry **281**(29): 20107-20119.

Matsuura, T., A. Ernst, et al. (2002). "Construction and characterization of protein libraries composed of secondary structure modules." Protein Science **11**(11): 2631-2643.

Wurth, C., N. K. Guimard, et al. (2002). "Mutations that reduce aggregation of the Alzheimer's A beta 42 peptide: an unbiased search for the sequence determinants of A beta amyloidogenesis." Journal of Molecular Biology **319**(5): 1279-1290.

Nordlund, H. R., O. H. Laitinen, et al. (2005). "Production of Hev b5 as a fluorescent biotin-binding tripartite fusion protein in insect cells." Biochemical and Biophysical Research Communications **336**(1): 232-238.

Nakayama, M. and O. Ohara (2003). "A system using convertible vectors for screening soluble recombinant proteins produced in Escherichia coli from randomly fragmented cDNAs." Biochemical and Biophysical Research Communications **312**(3): 825-830.

DeLucas, L. J., D. Hamrick, et al. (2005). "Protein crystallization: virtual screening and optimization." Progress in Biophysics & Molecular Biology **88**(3): 285-309.

Kaba, S. A., V. Nene, et al. (2002). "Fusion to green fluorescent protein improves expression levels of Theileria parva sporozoite surface antigen p67 in insect cells." Parasitology **125**: 497-505.

Ignatova, Z. and L. M. Gierasch (2004). "Monitoring protein stability and aggregation in vivo by real-time fluorescent labeling." Proceedings of the National Academy of Sciences of the United States of America **101**(2): 523-528.

Sandrock, T. M., B. Risley, et al. (2001). "Exogenous peptide and protein expression levels using retroviral vectors in human cells." Molecular Therapy **4**(5): 398-406.

History

●Green Fluorescent Proteins (GFP) is a new technology, nobilized in 2008 [(r)].

Osamu Shimomura first isolated GFP from the jellyfish Aequorea victoria, which drifts with the currents off the west coast of North America. He discovered that this protein glowed bright green under ultraviolet light.

Martin Chalfie demonstrated the value of GFP as a luminous genetic tag for various biological phenomena. In one of his first experiments, he coloured six individual cells in the transparent roundworm Caenorhabditis elegans with the aid of GFP.

Roger Y. Tsien contributed to our general understanding of how GFP fluoresces. He also extended the colour palette beyond green allowing researchers to give various proteins and cells different colours. This enables scientists to follow several different biological processes at the same time.

Douglas Prasher, a researcher at Woods Hole Oceanographic Institution in Massachusetts who originally isolated the gene for GFP, was not one of the Nobel recipients. Prasher freely gave the gene sequence for GFP to both Roger Tsien and Martin Chaflie. Doug Prasher's scientific colleagues at LANL would like to acknowledge Prasher for his embodiment of the concept of pure scientific collaboration.

●LANL Overview

Green Fluorescent Protein (GFP) has been around for many years and has been used in a lot of creative ways. Los Alamos National Laboratory researcher (LANL) Dr. **Geoff Waldo**, has spent the last decade improving the flexibility, usability, reliability and sensitivity of GFP by engineering it to have more desirable characteristics. His work has resulted in a GFP that fluoresces more brightly, does not perturb the protein of interest, and works reliably in a number of important scientific applications. Not only does it perform better than other tags, but it is faster and cheaper!

Related products/documents from Interchim - BioSciences Innovations:

Products HighLights Overview

Split GFP: unique GFP system to increase S/N signal, solve solubility/aggregation issues and it's study, as well as protein/protein interactions.	See PH-BB213b
LEXSY2 protein expression system – combines scalability with full-eucaryotic machinery	See PH-BB216c
EvoGlow GFP: unique GFP able to express in anaerobic cells	See PH-BB213e
Recombinant protein purification	
Desalting tools (dialysis, gelfiltration, ultrafiltration)	See B100
Protein assays, colorimetric (BC Assay, Coo Acssay), and fluorimetric (LavaPep)	See B191, BB191b
Electrophoresis analysis: biochemicals for gels, pre-cast gels, stains, MW markers	See B194-208, BB198a

For ordering information please search at http://www.interchim.eu
For more information, please ask interbiotech@interchim.com

Information inquire

Reply by Fax : +33 (0) 4 70 03 82 60 or email at interbiotech@interchim.com

☐ I wish to receive the complete documentation about: _____

_____

Name: _____ 2nd name:_____ Position:_____

Company/Institute:_____ Service, Lab:_____

Adress: _____

Zip code: _____ Town: _____

Tel_____ Fax_____ Email: _____