SUPPLEMENTAL INFORMATION

Combining random gene fission and rational gene fusion to discover near-infrared fluorescence protein fragments that report on protein-protein interactions

Naresh Pandey[1], Christopher L. Nobles[2], Lynn Zechiedrich[3,4], Anthony W. Maresso[2], and Jonathan J. Silberg[1,*]

Author affiliations:

1. Department of Biosciences, Rice University, Houston, TX 77005, USA
2. Department of Molecular Virology and Microbiology, Baylor College of Medicine, Houston, TX 77030, USA
3. Department of Pharmacology, Baylor College of Medicine, Houston, TX 77030, USA
4. Verna and Marrs McLean Department of Biochemistry and Molecular Biology, Baylor College of Medicine, Houston, TX 77030, USA

*To whom correspondence should be addressed: Tel: 713-348-3849; Fax: 713-348-5154; Email: joff@rice.edu

		+EK		–fusion		+K		+AY	
ORF1 (residues)	**ORF2** (residues)	λ_{ex} (nm)	λ_{em} (nm)	λ_{ex} (nm)	λ_{em} (nm)	λ_{ex} (nm)	λ_{em} (nm)	λ_{ex} (nm)	λ_{em} (nm)
1-108	108-321	683	709	682	709	---	---	682	712
1-111	111-321	684	712	682	712	---	---	683	712
1-117	117-321	682	711	nd	708	nd	710	683	711
1-118	118-321	683	713	nd	711	nd	715	683	711
1-129	129-321	682	712	682	711	683	709	683	711
1-131	131-321	682	711	682	709	681	712	683	712
1-135	135-321	683	710	682	711	---	---	683	709
1-140	140-321	682	711	683	710	684	709	682	712
1-141	141-321	681	710	684	711	---	---	683	711
1-142	142-321	683	710	683	715	681	709	684	710
1-143	143-321	683	711	683	709	---	---	682	711
1-144	144-321	683	711	682	713	682	714	683	712
1-147	147-321	682	711	683	712	---	---	683	711

Table S1. Emission and excitation properties of fragmented IFP. Emission and excitation spectra were measured using *E. coli* expressing each fragmented IFP to determine the wavelengths that yield maximum emission (λ_{em}) and excitation (λ_{ex}) at 37°C. For each variant discovered in the EK and AY libraries, spectral analysis was performed with fragmented IFP fused to IAAL-E3 and IAAL-K3 (+EK), CheA and CheY (+AY), or in the absence of any fusions (-fusion). A subset of the fragmented IFPs were also analyzed with only one fragment fused to the IAAL-K3 peptide (+K). ORF1 represents the IFP residues encoded by the open reading frame that precedes the peptide backbone cleavage site, whereas ORF2 represents the IFP residues encoded by the open reading frame that follows the cleavage site. A sharp excitation spectrum was not detected (nd) with several of the fragmented IFP upon removal of IAAL-E3 or IAAL-E3 and IAAL-K3 due to the weak signal. Cells expressing full-length IFP yielded maximal emission upon excitation at 683 nm and maximal excitation when emission was monitored at 711 nm.

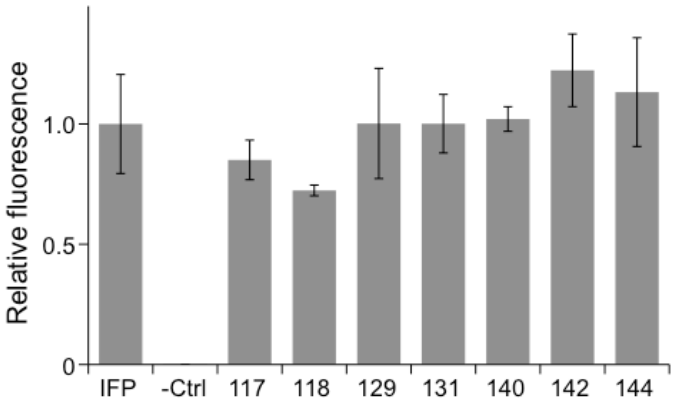

Figure S1. Fluorescence of fragmented IFP at 23°C. The fluorescence (λ_{ex} = 684 nm; λ_{em} = 710 nm) of *E. coli* expressing fragmented IFP identified in the EK library was measured at 23°C, normalized to cell density, and reported relative to the fluorescence of cells expressing full-length IFP and lacking a fluorescent protein (-Ctrl). Error bars represent $\pm 1\sigma$. The fluorescence intensity obtained with cells expressing each split variant was significantly different from cells lacking IFP (two tailed t test; $p<0.005$).

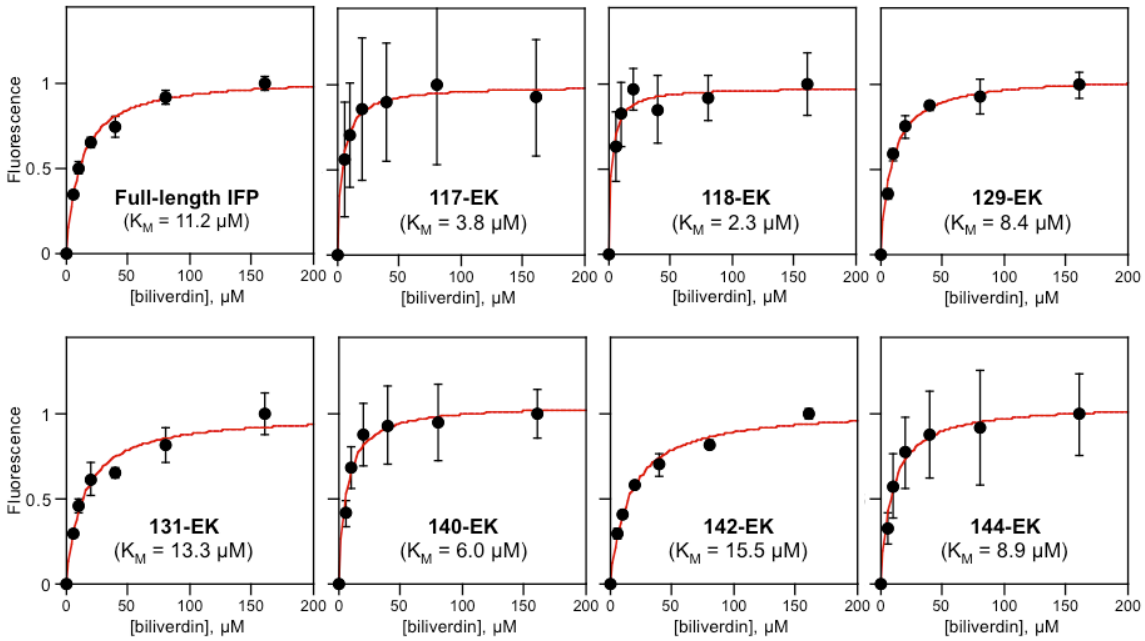

Figure S2. Effect of BV concentration on the fluorescence of fragmented IFP fused to IAAL-E3 and IAAL-K3 peptides. The fluorescence (λ_{ex} = 684 nm; λ_{em} = 710 nm) of *E. coli* expressing different fragmented IFP discovered in the EK library was measured in cells grown in medium supplemented with varying BV amounts (0, 5, 10, 20, 40, 80, and 160 μM). Fluorescence normalized to the maximum value obtained for each variant was fit to a hyperbolic saturation function to obtain the concentration required for half maximal signal (K_M) within whole cells. Error bars represent ±1σ.

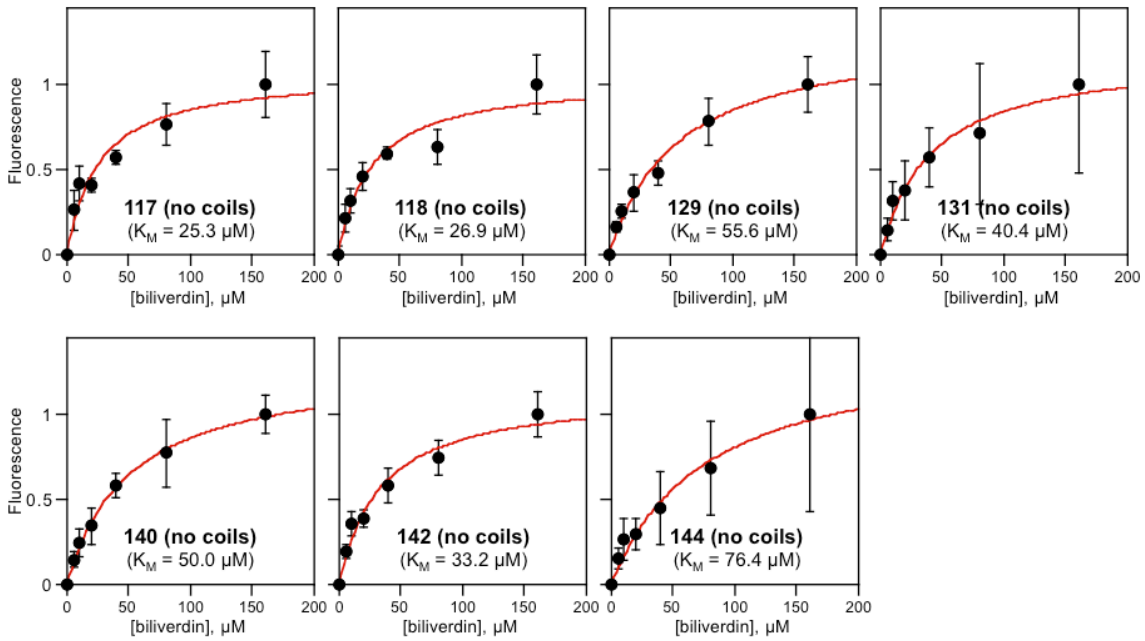

Figure S3. Effect of biliverdin concentration on the fluorescence of fragmented IFP lacking fusion to any peptides or proteins. The fluorescence (λ_{ex} = 684 nm; λ_{em} = 710 nm) of *E. coli* expressing fragmented IFP that lack fusion to IAAL-E3 and IAAL-K3 peptides was analyzed in medium supplemented with 0, 5, 10, 20, 40, 80, and 160 μM BV. Fluorescence normalized to the maximum value obtained for each variant was fit to a hyperbolic saturation function to obtain the concentration required for half maximal signal (K_M) within whole cells. Error bars represent ±1σ.

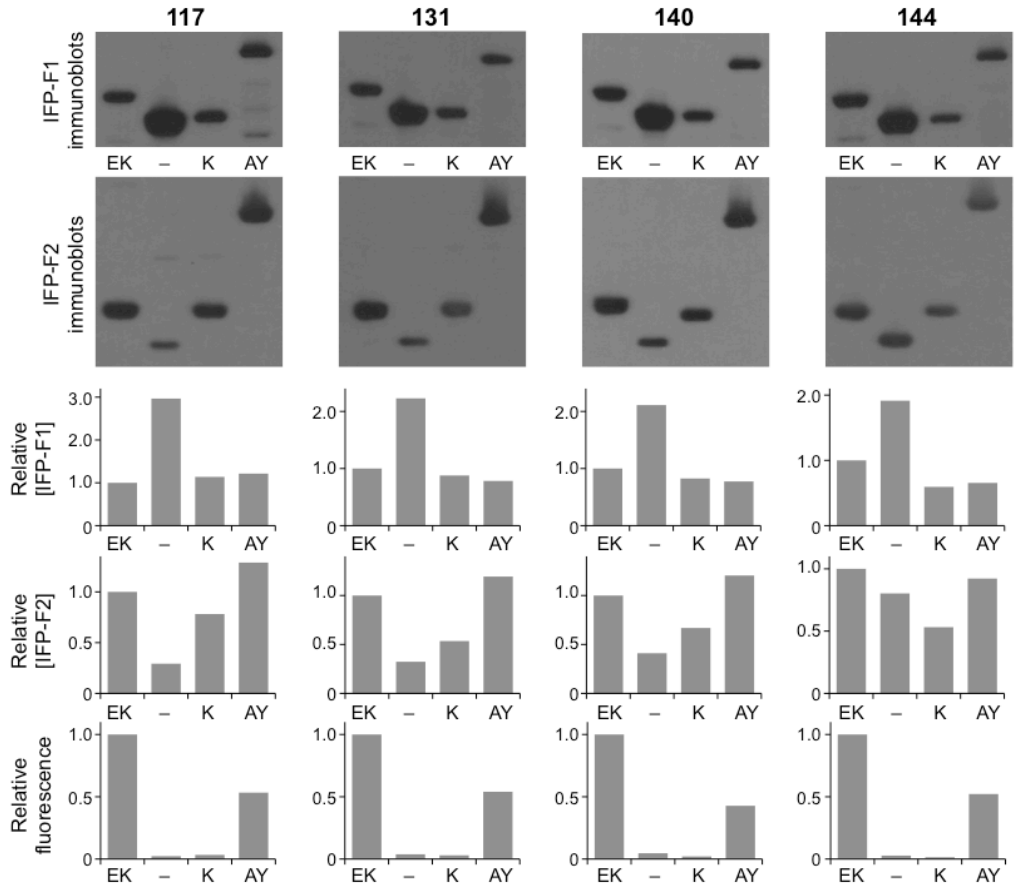

Figure S4. Effect of changing protein fusion on the levels of IFP fragments within cells. Western blot analysis was used to analyze how changing the peptides and proteins fused to different fragmented IFP affects the steady-state levels of the fragments under conditions where fluorescence measurements were performed. For four of the fragmented IFP discovered in the EK library, we compared the relative expression of the N- and C-terminal fragments (IFP-F1 and IFP-F2, respectively) of each variant with IAAL-E3 and IAAL-K3 (EK) fused to the different IFP fragments, CheA and CheY (AY) fused to the different IFP fragments, IAAL-K3 (K) fused to only one of the fragments, and no peptides or proteins (–) fused to either fragment. ImageJ was used to quantify the levels of the N- and C-terminal fragments in each blot, which are reported relative to the levels of the two-piece IFP fused to IAAL-E3 and IAAL-K3. For comparison, the bottom panels illustrates the effects of peptide and protein fusion on fluorescence (λ_{ex} = 684 nm; λ_{em} = 710 nm).

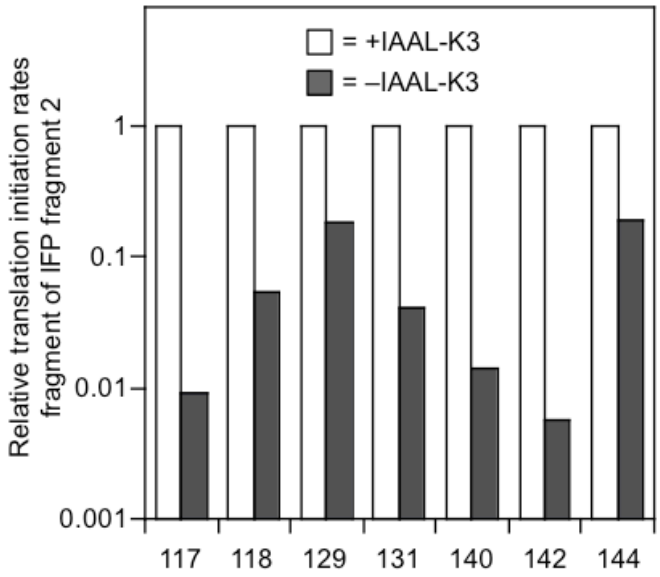

Figure S5. Calculated translation initiation rates for C-terminal fragments containing or lacking fusion to IAAL-K3. A thermodynamic model for translation initiation was used to calculate the relative translation initiation rates of the RBS that controlled the expression of IFP fragments containing (open bars) or lacking (closed bars) IAAL-K3 peptides. Translation initiation rates were normalized to the maximum value obtained with IAAL-K3 peptides. All calculations were performed using the DNA sequence that flanked the start codon of the ORFs encoding the IFP fragments, including 50 bases before and after the predicted start codons.

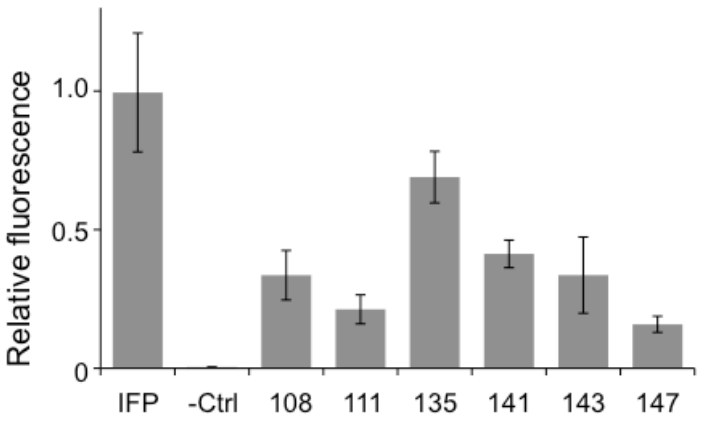

Figure S6. Fragmented IFP discovered in the AY library. The fluorescence (λ_{ex} = 684 nm; λ_{em} = 710 nm) of each fragmented IFP upon expression in *E. coli* at 37°C is shown relative to the fluorescence of cells expressing full-length IFP and lacking IFP (-Ctrl). Fluorescence is normalized to cell density. Error bars represent ±1σ. The fluorescence intensity obtained with cells expressing each split variant was significantly different from cells lacking IFP (two tailed t test; $p<0.005$ for all variants except 143 which displayed a $p<0.02$).

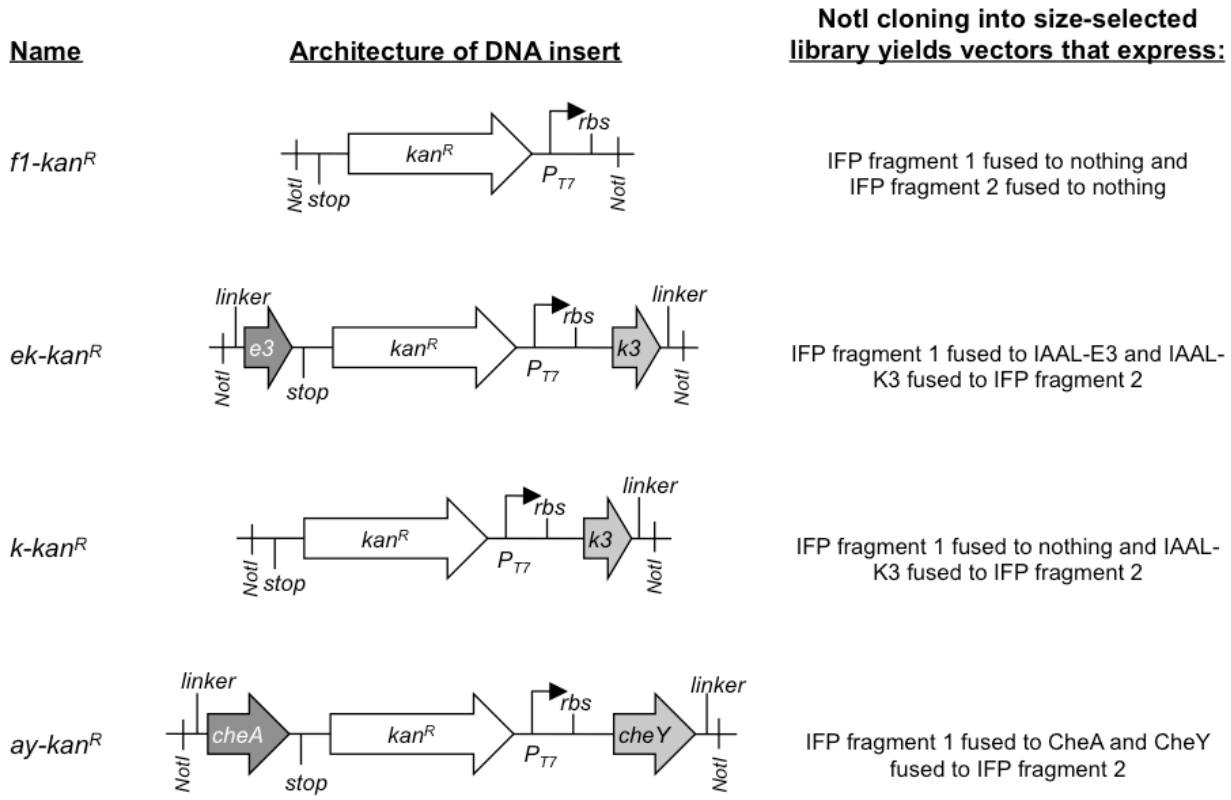

Figure S7. Synthetic DNA used for the last step of library construction. The architecture of the four DNA used to create vectors that express fragmented IFP. Upon insertion into IFP genes, all of these DNA terminate expression of the IFP fragment that precedes the original transposon insertion site and initiate expression of the IFP fragment that follows the ORF. *f1-kan*^{R} does not amend associating peptides or proteins to either IFP fragment. *ek-kan*^{R} fuses IAAL-E3 to the end of the IFP fragments that precede the fragmentation site and IAAL-K3 to the beginning of the IFP fragments that follow the fragmentation sites. *k-kan*^{R} fuses IAAL-K3 to the beginning of the IFP fragments that follow the fragmentation sites and nothing to the fragment that precedes the fragmentation site. *ay-kan*^{R} fuses CheA to the end of the IFP fragments that precede the fragmentation site and CheY to the beginning of the IFP fragments that follow the fragmentation sites.

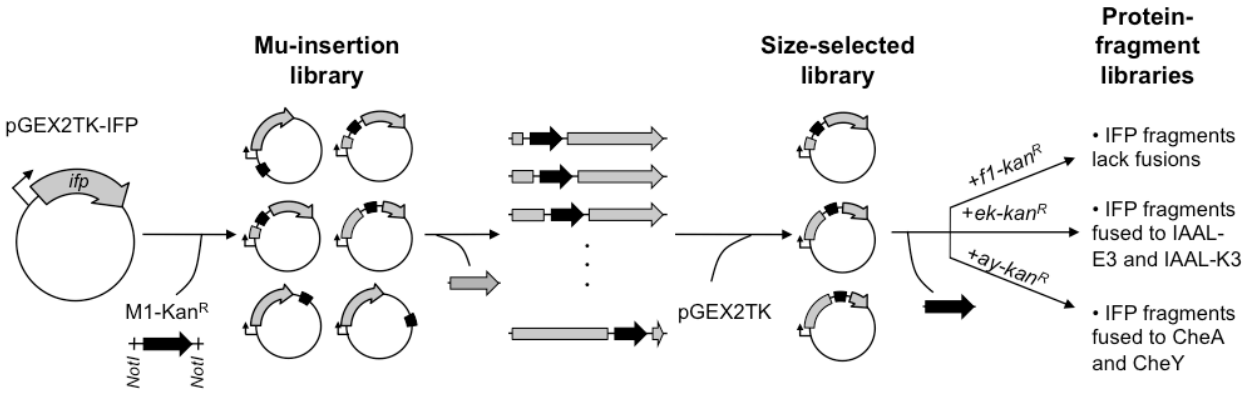

Figure S8. The method for creating libraries of fragmented proteins fused to different pairs of associating proteins. Library construction involves four steps, which included: (i) random insertion of transposons into the IFP genes using the transposase MuA, (ii) purifying IFP genes with inserted transposons away from IFP genes lacking inserted transposons, (iii) subcloning the ensemble of IFP-transposon hybrid genes into an expression vector, and (iv) subcloning synthetic DNA encoding different pairs of interacting proteins in place of the transposons. The last step can be performed using synthetic DNA that encode for interacting peptides (IAAL-E3 and IAAL-K3), proteins (CheA and CheY), or lacking additional ORFs as shown in Figure S2.