Supporting Information

High-throughput instant quantification of protein expression and purity based on photoactive yellow protein turn off/on label (POOL)

Youngmin Kim[1,2], Prabhakar Ganesan[1,2] and Hyotcherl Ihee[1,2*]

[1]Center for Nanomaterials and Chemical Reactions, Institute for Basic Science, Daejeon, 305-701, Republic of Korea

[2]Department of Chemistry, KAIST, Daejeon, 305-701, Republic of Korea

Materials and Methods

POOL-calmodulin cloning

The PYP gene was subcloned in the pQE80L vector using EZ-Cloning (enzynomics™). Initially the PYP gene with BamHI and SalI restriction enzyme sites was amplified using PCR. The primers contained the BamHI and SalI restriction enzyme sites are as follows.

5’-CACCATCACGGATCCATGGAACACGTAGCCTTCGG -3’

5’-TGGCTGCAGGTCGACGACGCGCTTGACGAAGACCC-3’

In this PCR step, the stop codon of the end of the PYP gene was removed. The amplified PCR product with BamHI and SalI restriction enzyme sites was inserted into the pQE80L vector. Then, the calmodulin sequence from human cDNA with a thrombin protease recognition site was amplified by PCR with the following primers.

5’-AAGCGCGTCGTCGACGGCAGCCTGGTGCCGCGCGGCAGCATGGCTGATCA

GCTGACCGA-3’

5’-TGGCTGCAGGTCGACCTATTTTGCAGTCATCATCTGTACGAA -3’

The amplified calmodulin gene was inserted into the pQE80L-PYP plasmid at the SalI restriction enzyme site using EZ cloning kit. The same approach was used to make a construct with a different vector (the pQE30-PYP-calmodulin construct) except that the SalI restriction enzyme site was substituted with the BglII restriction enzyme site.

Calmodulin cloning

The PCR amplified calmodulin gene with the thrombin protease recognition site was inserted into the pQE80L vector. The primers for the calmodulin gene containing BamHI and SalI restriction enzyme sites are as follows.

5’-AAGCGCGTCGTCGACGGCAGCCTGGTGCCGCGCGGCAGCATGGCTGATCA

GCTGACCGA-3’

5’-TGGCTGCAGGTCGACCTATTTTGCAGTCATCATCTGTACGAA-3’

The amplified thrombin calmodulin gene was subcloned into the pQE80L vector using EZ-Cloning kit (Enzynomics^{TM}). The same approach was also used to make a construct with a different vector (pQE30) except that the SalI restriction enzyme site was substituted with the BglII restriction enzyme site.

POOL-horse myoglobin cloning

We performed three-fragments cloning using EZ-Cloning kit (Enzynomics^{TM}). The pQE30-PYP-calmodulin was cut by both BamHI and BglII to make a linear pQE30 vector (first fragment). The PYP gene (second fragment) was amplified without the stop codon using PCR

with the following primers containing the BamHI restriction enzyme site.

5’-CACCATCACGGATCCGATGACGATGACAAAATGGAACACGTAGCCTTCGG-3’

5’-GCTGCCGCGCGGCACCAGGACGCGCTTGACGAAGACCC-3’

The reverse primer contains the thrombin protease recognition site.

The horse Mb gene (third fragment) was amplified by PCR with the following primers containing the BglII restriction enzyme recognition site. The forward primer contains the thrombin protease recognition site at the 5’ region to overlap with the amplified PYP gene.

5’-CTGGTGCCGCGCGGCAGCGGCCTGAGCGATGGCGAATG-3’

5’-CTCTTGATCAGATCTTTAGCCCTGAAAGCCCAGTTCTTTA-3’

SUMO-POOL cloning

First, the SUMO gene was cloned in the pQE80 vector. The kpnI restriction enzyme recognition site was inserted next to the SUMO gene using site-directed mutagenesis with the following primers.

5'-GGTACCGTCGACCTGCAGCCAAGCTTAATTAG-3’

5'-CCACCAATCTGTTCTCTGTGAGCCTCAATAATATCG-3’

Then, the plasmid was treated with KpnI restriction enzyme. The PYP gene was amplified by PCR with the following primers.

5'-GAACAGATTGGTGGGATGGAACACGTAGCCTTCGG-3’

5'-GCAGGTCGACGGTACCTAGACGCGCTTGACGAAGAC-3’

Finally, by using EZ-cloning kit, the amplified PYP gene was subcloned to the vector treated with kpnI restriction enzyme.

MerP-POOL cloning

The MerP was inserted in front of the PYP gene by using EZchange™ Site-directed Mutagenesis kit (Enzynomics^{TM}) with the following primers.

5’-ACCCTGGCGGTGCCGGGCATGACCTGCGCGGCGTGCCCGATTACCGTGAAAA AAGCGGCGATGGAACACGTAGCCTTCGGTAGC-3’

5’-TTTGTCATCGTCATCGGATCCGTG-3’

Protein expression, purification and instant quantification of the target protein

The gene constructs were transformed into the E. Coli BL21 (DE3) cell line. The bacterial cultures were grown at 37 ℃ with ampicillin and vigorous shaking until a mid-exponential phase (OD 0.6 at 600 nm) was reached. Then in some cases temperature was lowered to 18 ℃. The expression was induced with 1 mM or 0.2 mM IPTG. Cultures were grown for an additional 18 h and were harvested by centrifugation at 6000 g for 10 minutes. The pellet was sonicated in lysis buffer (50 mM NaPO4 buffer pH 7.4 with 0.1 M NaCl). The characteristic yellow color of PYP is absent in this stage, but can be turned on by binding its chromophore, p-coumaric acid. The expression of the POOL labeled target protein was checked by adding the chromophore precursor to the cell lysate and inspecting the color change. The concentration of the expressed target protein can be quantified by comparing its yellow color with those of the standard reference solutions of known concentrations. The cell lysate was centrifuged at 15000 g for 1 hour and the supernatant was applied to a nickel affinity column that was previously equilibrated with the lysis buffer. The protein bound to the nickel affinity column was additionally washed with a lysis buffer containing 40 mM imidazole and was subsequently eluted using a lysis buffer containing 200 mM imidazole. After dialyzing the

eluted protein with 20 mM NaPO4 buffer pH 7.4, the dialyzed protein was applied to ion exchange column (HiTrap, GE Healthcare Life Sciences). Then the protein was eluted in a gradient method by slowly increasing the concentration of NaCl. The amount of the target protein was checked in all the eluted fractions by comparing the color with the standard reference solutions of known concentrations.

Purity test of the target protein for fraction selection after purification work

The yellow colored fractions eluted from the ion-exchange column were further used to measure the concentration and purity. By using a UV-VIS spectrometer or a microplate absorption reader after transferring the fractions to a new 24 well plates, the UV-VIS spectra of the selected fractions were measured to assess the exact amount of the concentration and purity which can be determined by using Equations (1) and (2), respectively.

Engineering mini-photoactive yellow protein

We used the pQE80L-SUMO-POOL plasmid for engineering the mini-photoactive yellow protein. The deletion mutagenesis was performed by EZchange™ Site-directed Mutagenesis kit (Enzynomics^{TM}). We followed the procedures according to the EZchange™ Site-directed Mutagenesis kit manual. We made a total of 12 deletion mutants; eight types of N-terminal deletion mutants (deletion of residues 1~5 (N5), deletion of residues 1~10 (N10), deletion of residues 1~18 (N18), deletion of residues 1~27 (N27), deletion of residues 1~31 (N31), deletion of residues 1~40 (N40), deletion of residues 1~42 (N42), and deletion of residues 1~46 (N46)) and three types of C-terminal deletion mutants (deletion of residues 73~125 (C53), deletion of residues 107~125 (C18), and deletion of residues 117~125 (C8)). In

addition, a mutant containing both N-terminal deletion (deletion of residues 1~27) and C-terminal deletion (deletion of residues 107~125) was made

The primers used for making the deletion mutants are as follows.

Primer name	Sequence
Reverse primer for all N-terminal deletion mutant	5’-CCCACCAATCTGTTCTCTGTGAGCCTCAATAATATC
Forward primer for deletion of residues 1~5	5’-TTCGGTAGCGAGGACATCGAGAACACCC
Forward primer for deletion of residues 1~10	5’-ATCGAGAACACCCTCGCCAAGATGGAC
Forward primer for deletion of residues 1~18	5’-GACGACGGCCAGCTCGACGGC
Forward primer for deletin of residues 1~27	5’-TCCGGCGCCATCCAGCTCGACG
Forward primer for deletion of residues 1~31	5’-CAGCTCGACGGCGACGGCAACATC
Forward primer for deletion of residues 1~40	5’-CAGTACAACGCCGCGCAGGGCG
Forward primer for deletion of residues 1~42	5’-AACGCCGCGCAGGGCGACAT
Forward primer for deletion of residues 1~46	5’-GGCGACATCACCGGCCGCGA
Forward primer for deletion of residues 73~125	5’-GTCTAGGTACCGTCGACCTGCAGCCAAG
Reverse primer for deletion of residues 73~125	5’-GCTGTCAGTGCACGGGGCCACG
Forward primer for deletion of residues 107~125	5’-GTCTAGGTACCGTCGACCTGCAGCCAAG
Reverse primer for deletion of residues 107~125	5’-CTTCACCTTCGTGGGCGTCATTTGGTAAT
Forward primer for deletion of residues 118~125	5’-TAGGTACCGTCGACCTGCAGCCAAGC
Reverse primer for deletion of residues 118~125	5’-GCTGTCGCCGGAGAGGGCCTTC

Cloning for application of POOL to eukaryotic (mammalian and insect) cells

All cells used in this study were obtained from the American Type Culture Collection (ATCC, Manassas, VA, USA). The signal sequence and extracellular domain of human CD40 receptor (1M ~ 277Q) was cloned together with POOL into the pcDNA3.1 vector. The CD40-negative HEK 293 cells were transfected with the plasmid using PEI reagent with 1:2.5 ratio between PEI and DNA. The HEK 293 cells for coexpressing CD40 receptor (1M ~ 277Q) were cultured in DMEM containing 5% FBS, L-glutamine, penicillin, and streptomycin (Wisent). After transfection, we incubated the cells for 5 days. Then the CD40 receptor was expressed. The cells were subsequently harvested and sonicated.

The truncated TLR3 (toll-like receptor 3 without transmembrane domain) with POOL were cloned into pIEX3 (EMD Biosciences, San Diego) for transient expression in insect cells. The Spodoptera frugiperda (Sf21) insect cells were transfected with the plasmid. The protein was produced by incubating Sf21 cells for 1 hour in modified Excel 401 medium (JRH Biosciences, Lenexa, KS) supplemented with 10% dialyzed, heat-inactivated fetal bovine serum (Life Technologies, Grand Island, NY) and 50 μg/ml gentamicin. The cells were subsequently harvested and sonicated.

Recipe for making a precursor solution of chromophore (p-coumaric anhydride)

The precursor, p-coumaric anhydride, of the chromophore can be synthesized based on the methodology developed by Imamoto et al. [1]. First, 1 mM DCC (N, N'-Dicyclohexylcarbodiimide) solution in DMF (Dimethylformamide) (2.7 g in 50 ml) and 0.77 mM p-coumaric acid in DMF (2.55 g in 50 ml) are prepared by dissolving each into DMF and stirring the resulting solution at 4 ℃. After complete dissolution of each substance in

DMF, two solutions are mixed and stirred at 4 ℃ overnight. The solution can be scaled up if necessary, but the molar ratio of DCC versus p-coumaric acid has to be maintained at 1.3:1. At the end of this reaction, some white precipitation in light yellow solution appears. The white precipitation is removed by centrifugation at 12,000 rpm, 4 ℃. The supernatant is stored at -70 ℃. Typically for efficient binding of the chromophore to all the PYP molecules, an excess amount of the chromophore precursor is used. We usually use 10 ml choromophore solution for 2 liter PYP culture.

Comparison of POOL and GFP tag

The POOL tag has the following advantages and disadvantages compared with the GFP tag. The advantages of POOL are as follows. First, the most important aspect of the POOL tag is that the visible yellow color is initially off but can be turned on by adding the chromophore. This turn-on property is in stark contrast to the GFP tag which usually comes with its chromophore (generated by the reaction of inherent amino acid residues) already present and thus its color is already in the on-state. Therefore the color cannot be easily distinguished from other potential sources of color such as contamination and the inherent color of the target protein, which essentially makes it difficult to use the GFP tag for quantitative purpose. In contrast, in the case of POOL, the yellow color is turned on only when we add the chromophore and thus can be easily distinguished from such other potential sources of color. For example, this visual colorimetric inspection scheme works even when the target protein is myoglobin, which has an inherent strong red color. For this reason, the POOL tag can be easily used to measure the concentration and purity. Second, POOL allows the expression of the target protein to be checked even without using any equipment. GFP requires

fluorescence detection and it is well known that the GFP suffers from the bleaching problem. Finally, the POOL tag is much smaller and more soluble than the GFP tag, and thus we can expect that the former interfere with the expression of the target protein compared with the latter. One of disadvantages of the POOL tag compared with the GFP tag is that the cells need to be disrupted for clear color detection whereas the fluorescence of the GFP tag can be detected without cell disruption.

Reference

1. Imamoto, Y., et al., *Reconstitution photoactive yellow protein from apoprotein and p-coumaric acid derivatives.* FEBS Lett, 1995. **374**(2): p. 157-60.

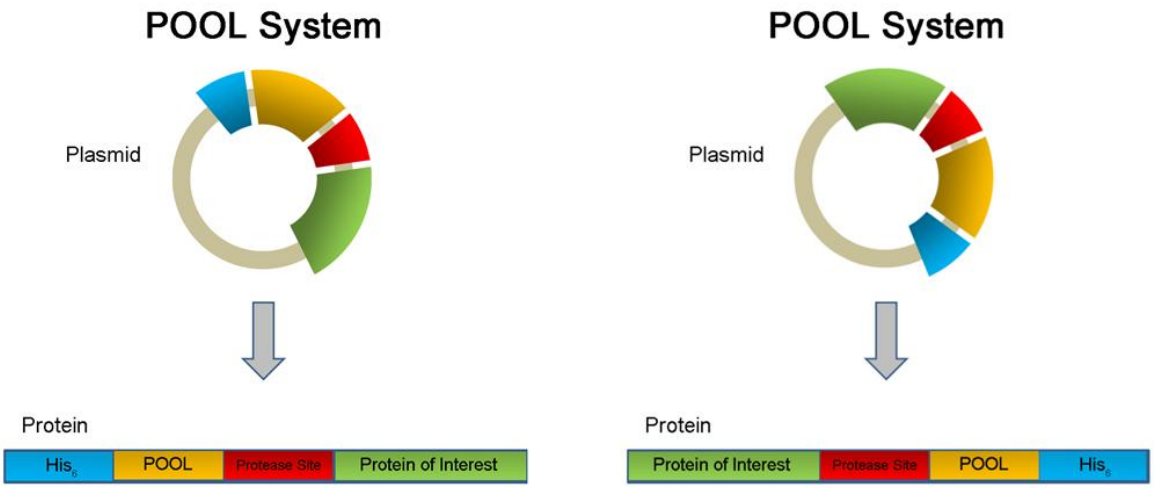

Figure S1. POOL expression vector including the PYP gene and a target protein gene. The order of genes in the POOL expression system shown in the left is as follows: poly-histidine tag, PYP, protease recognition site and target protein. Reverse order arrangement of genes in the POOL system is also feasible as shown in the right. The pQE80, pQE30 and pET15b vectors were used as the bases for designing the construct with thrombin, HRV C3 protease, TEV or enterokinase as the protease recognition site.

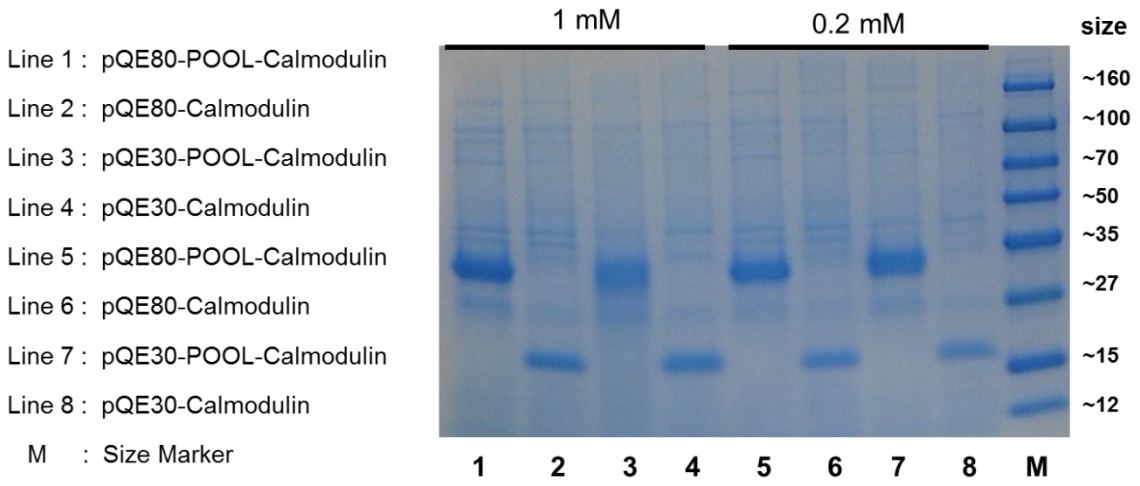

Figure S2. The observed differences in the expression rate of soluble human calmodulin with or without POOL. The samples loaded into all wells were the same amount (total 1ug). The figure shows PAGE results of the cell lysate of calmodulin expressed in18 °C. Lines 1, 3, 5 and 7 correspond to human calmodulin with POOL. Lines 2, 4, 6 and 8 correspond to human calmodulin without POOL. The band thickness can be related to the calmodulin expression rate when the size of the POOL is taken into account. The expression rate of calmodulin with POOL is either higher than or the same as that of calmodulin without POOL. These results indicate that the POOL system does not interfere with the expression of soluble target proteins and sometimes even improves the expression rate.

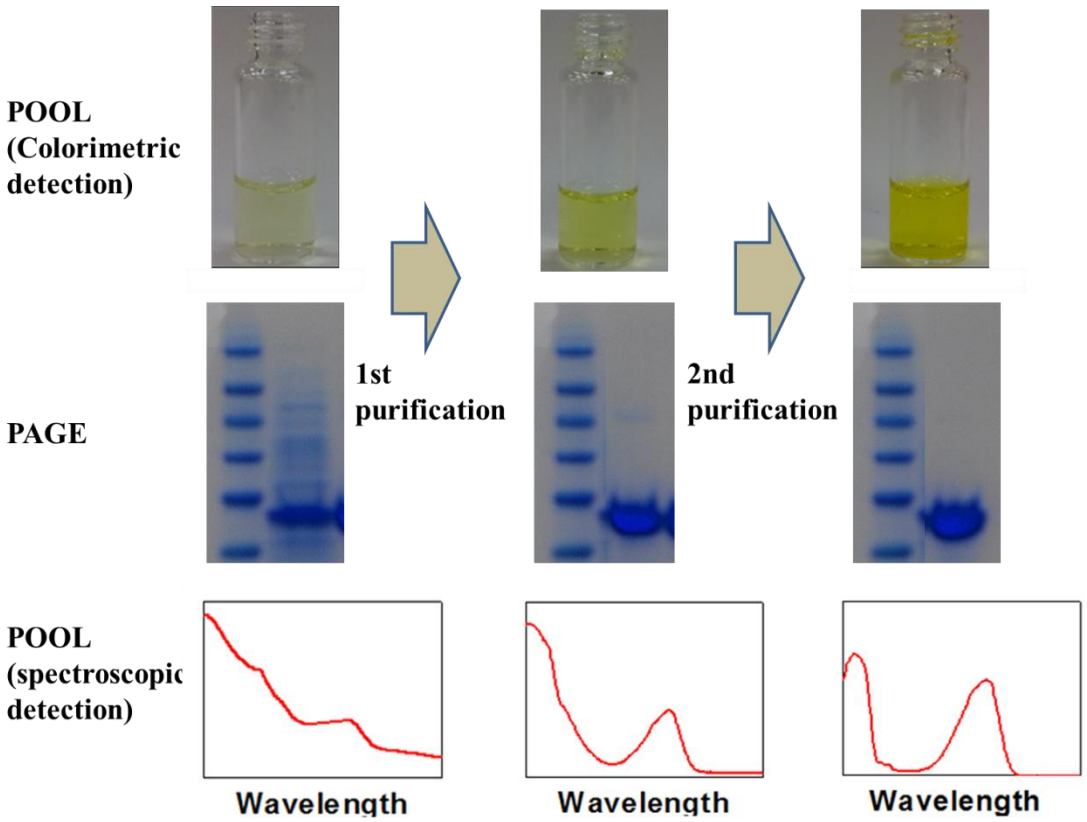

Figure S3. For eluted fractions during the purification steps, the visual colorimetric detection, PAGE gel and UV-VIS spectra are compared. One can see that the purity of the target protein (in this case, SUMO protein) increases with each purification step. The brighter yellow color of the eluted fraction indicates higher purity of the target protein, which is reflected in the PAGE gel as well as the magnitude of absorption peak at 460 nm in the UV-VIS spectra. The results indicate that the POOL system can measure the purity and concentration of the target protein in an easier and simpler manner than by PAGE.

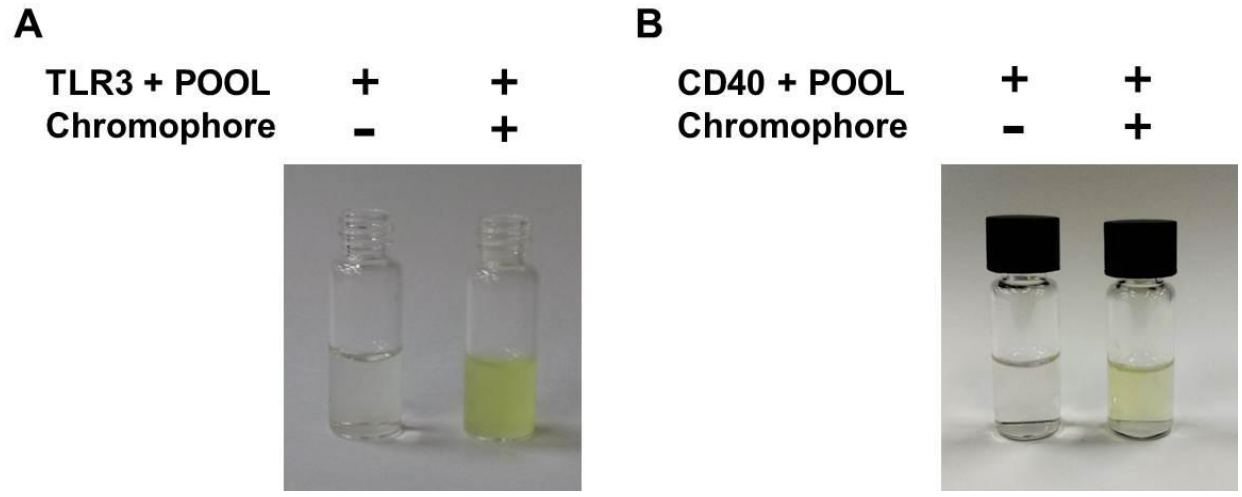

Figure S4. Applying POOL to the insect and mammalian cells. The expression of TLR3 and CD40 fused with POOL can be easily checked by adding a precursor of chromophore in the insect cell (A) and mammalian 293 cell (B), respectively. The left and right bottles correspond to before and after adding a precursor of chromophore.