Supporting Information

Dual display of proteins on yeast cell surface simplifies quantification of binding interactions and enzymatic bioconjugation reactions

Sungwon Lim[1], Jeff E. Glasgow[1,2,3], Maria Filsinger Interrante[1], Erica M. Storm[4], Jennifer R. Cochran[1,5]

[1]Dept. of Bioengineering, Schools of Engineering and Medicine, Stanford University, Stanford, CA, USA

[2]Joint Initiative for Metrology in Biology, Stanford, CA, USA

[3]Genome-scale Measurements Group, National Institute of Standards and Technology, Stanford, CA, USA

[4]School of Medicine, Stanford University, Stanford, CA, USA

[5]Dept. of Chemical Engineering, School of Engineering, Stanford University, Stanford, CA, USA

Corresponding author:

*Email: jennifer.cochran@stanford.edu

SI Materials and Methods

1. Plasmids construction

The parent pTMY vector [1] was first modified to express a new $(Gly_4Ser)_3$ linker at the C-terminus of Aga2p as described in the manuscript. The fusion construct expression cassettes were then engineered as follows to streamline binding assays and enzymatic bioconjugation studies, and also to improve usability in future applications.

1.1 pCL vectors for binding assays

For binding assays using pCL plasmids, yEGFP was expressed as a fusion to the N- or C-terminus of Aga2p, and used as an indicator for expression levels of the fusion construct on yeast surface. Cloning was used to introduce yEGFP at the N-terminus of Aga2p through *Nhe*I and *Mlu*I restriction sites (pCL-nGFP; Supplementary Fig. 2A), leaving the C-terminus of Aga2p available to display a protein of interest. When designing pCL cloning primers containing a *Mlu*I site, it is important to include an additional base pair (guanine was used in this study) immediately preceding the *Mlu*I site to prevent a reading frame shift. The C-terminal portion of Aga2p was modified to enable display of wild-type Axl Ig1, an engineered Axl variant MYD1 [2] or scFv D1.3 [3, 4]. Open reading frames encoding for these proteins were cloned downstream of a second $(Gly_4Ser)_3$ linker, between *Avr*II and *Spe*I restriction sites. A c-Myc tag was included at the C-terminus of the protein of interest to generate pCL-nGFP-Aga2p-Axl (abbreviated pCL-Axl) or pCL-nGFP-Aga2p-D1.3. (abbreviated pCL-D1.3). In addition, a Factor Xa cleavage site and a HA tag were included as handles for protein characterization, if desired. For general use, yEGFP can be replaced by another yeast-optimized fluorescent protein [5] using the *Nhe*I and *Mlu*I restriction sites, and an alternative protein of interest can be cloned into the vector through the *Avr*II and *Spe*I restriction sites.

To display a protein of interest at the N-terminus of Aga2p, yEGFP was first cloned at the C-terminus of Aga2p, after the $(Gly_4Ser)_3$ linker, using *Xma*I and *Spe*I restriction sites (pCL-cGFP; Supplementary Fig. 2B). NK1 was cloned upstream of the first $(Gly_4Ser)_3$ linker located N-terminal to Aga2p, using *Nhe*I and *Mlu*I restriction sites to generate pCL-NK1-Aga2p-cGFP (abbreviated pCL-NK1). The vector included a HA tag located upstream and a c-Myc tag located downstream of NK1 to compare protein expression levels measured from epitope-binding antibodies or yEGFP. For general use, yEGFP can be replaced by another yeast-optimized fluorescent protein [5] using the *Xma*I and *Spe*I restriction sites, and an alternative protein of interest can be cloned into the vector through the *Nhe*I and the *Mlu*I restriction sites.

1.2 pCL vectors for enzymatic assays

The engineered sortase variant 7M [6] and its substrate sequence LPETGG were cloned into the pCL vector along with different linker strategies: a short linker (15 aa; $(Gly_4Ser)_3$), a long linker (42 aa; $(Gly_4Ser)_3$–$(Gly_2Ser)_9$), or yEGFP plus a long linker (($(Gly_4Ser)_3$–yEGFP–$(Gly_2Ser)_9$) (Figure 3B). The open reading frame encoding for sortase 7M was cloned into pCL-cGFP (described above), upstream of Aga2p using *Nhe*I and *Mlu*I restriction sites. Then, each linker and the LPETGG sequence were cloned downstream of Aga2p between *Xma*I and *Spe*I restriction sites. For plasmids with a short or long linker, a FLAG epitope tag was included to validate expression of the linker and LPETGG fusion. Here, nucleotides encoding for the LPETGG and the FLAG tag sequences were codon-optimized for *S. cerevisiae* as TTGCCAGAAACTGGTGGT and GACTACAAAGACGATGATGACAAG, respectively. The resulting recombinant plasmids are termed pCL-Srt-SS (Short linker plus Substrate sequence), pCL-Srt-LS (Long linker plus Substrate sequence), and pCL-Srt-cGFP-LS (C-terminal GFP plus Long linker and Substrate sequence) (Fig. 3B).

1.3 pCL2: An optimized pCL vector for facile and modular cloning and library construction

The pCL vector was further optimized to add versatility and enhance compatibility with current yeast display platforms. This upgraded vector is called pCL2. As an example, the vector pCL-Srt-cGFP-LS was rebuilt to generate the pCL2-Srt-cGFP-LS vector (Supplementary Fig. 8). In the pCL2 construct, the *Mlu*I restriction site originally located downstream of the sortase sequence was exchanged with *Bam*HI, a restriction site that is compatible with the widely used yeast display vector pCTCON2 [7]. This exchange also eliminates the need for an extra base pair preceding the former *Mlu*I site, which was required to keep the translated gene in frame. In addition, two *Avr*II restriction sites were inserted flanking the c-Myc tag, which allows optional removal of the tag using a single restriction enzyme digestion and re-ligation. Insertion of a *Spe*I restriction site downstream of the yEGFP and insertion of an *Eco*RI site upstream of the sortase substrate sequence simplifies exchange of fluorescent proteins or C-terminal elements, respectively. The HA tag was removed from the N-terminus of the entire expression cassette to streamline the construct and to facilitate engineering of the N-terminal domain of a protein-of-interest. Finally, to minimize potential homologous recombination between the $(Gly_4Ser)_3$ linker regions, these codon sequences were randomized to reduce nucleotide sequence similarity.

2. Binding assays

2.1 scFv D1.3 binding to lysozyme. To evaluate the binding affinity of lysozyme to scFv D1.3 expressed using the pCL or pCT yeast display systems, 1×10^5 induced yeast cells were incubated with varying concentrations of biotinylated lysozyme (Sigma, L0289) in PBSA for 12 h at room temperature. Yeast expressing scFv D1.3 using the pCT-D1.3 vector were incubated with a 1:250 dilution of chicken anti-c-Myc antibody (Thermo Fisher Scientific, A21281) in PBSA for 30 min at 4 °C. Cells were washed with PBSA and incubated with a 1:100 dilution of AlexaFluor 488-labeled goat anti-chicken IgY (Thermo Fisher Scientific, A11039) and a 1:50 dilution of streptavidin-PE (BioLegend, 405204) for 15 min at 4 °C. Cells were then washed with PBSA prior to analysis by flow cytometry. Yeast expressing the pCL-nGFP-Aga2p-D1.3 (pCL-D1.3) plasmid were stained with a 1:50 dilution of streptavidin-PE for 15 min at 4 °C and washed with PBSA prior to flow cytometry analysis.

2.2 Axl Ig1 binding to Gas6. Similar to the scFv D1.3–lysozyme binding assay, the Gas6 binding affinity of Axl Ig1 expressed using the pCL or pCT vectors were measured by incubating 1×10^5 induced yeast cells with varying concentrations of His_6-tagged Gas6 [8] in PBSA for 17 h at room temperature. For yeast expressing Axl Ig1 using the pCL vector, cells were incubated with a 1:100 dilution of anti-His Tag IgG Hilyte Fluor 555 (Anaspec, AS-61250-H555) for 20 min at 4 °C, washed in PBSA, and analyzed by flow cytometry. For yeast transformed with pCT-nGFP-Aga2p-Axl (pCL-Axl), induced cells were incubated with a 1:500 dilution of chicken anti-c-Myc antibody (Thermo Fisher Scientific, A21281) for 30 min at 4 °C. Cells were washed with PBSA and secondary antibody labeling was carried out by incubating with a 1:100 dilution of mouse anti-His Tag IgG Hilyte Fluor 555 and a 1:100 dilution of goat anti-chicken IgY AlexaFluor 555 (Thermo Fisher Scientific, A21437) for 20 min at 4 °C. Cells were then washed with PBSA and analyzed by flow cytometry.

2.3 NK1 binding to Met. To measure the binding affinity of Met receptor to NK1 expressed using the pCL or pTMY vectors, 5×10^4 induced yeast cells were incubated with various concentrations of recombinant human Met-Fc (R&D Systems, 358-MT-100) for 6 h at 4 °C in PBSA100 (PBSA supplemented with additional 100 mM NaCl). After incubation, yeast cells displaying NK1 using the pCL-NK1-Aga2p-cGFP (pCL-NK1) plasmid were incubated with PBSA100 containing a 1:50 dilution of AlexaFluor 647 labeled goat anti-human IgG (Thermo Fisher Scientific, A21445) for 20 min at 4 °C, washed and analyzed by flow cytometry. For yeast transformed with pTMY-NK1, after

incubation with Met-Fc, cells were incubated with PBSA100 containing a 1:20 dilution of mouse anti-HA antibody (Cell Signaling Technology, 2367) for 1 h at 4 °C. Cells were washed and then incubated in PBSA100 containing a 1:20 dilution of goat anti-mouse IgG PE (Sigma, P9287) for 20 min at 4 °C, followed by incubation with PBSA100 containing a 1:50 dilution of AlexaFluor 647 goat anti-human IgG for 20 min at 4 °C. Cells were washed and analyzed by flow cytometry.

3. Antibody and target protein dissociation with washing steps

To compare changes in expression signals due to multiple washing steps performed during binding assays, experiments were designed to mimic various antibody staining strategies: the original condition with a minimal number of staining/washing steps; 1 additional wash for a two-step secondary antibodies staining; and 2 additional washes for a tertiary antibody staining strategy. Here, pCT-Axl and pCL-Axl vectors were used for demonstration.

Yeast cells transformed with pCT-Axl were induced for expression and incubated with Gas6 for 17 h at room temperature. After incubation, cells were washed one time with 1 ml of PBSA and then incubated with 50 µl of PBSA containing a 1:500 dilution of chicken anti-c-Myc antibody (Thermo Fisher Scientific, A21281) for 30 min at 4 °C. Following another wash with 1 ml of PBSA, secondary antibody labeling was carried out in 50 µl PBSA containing a 1:100 dilution of mouse anti-His Tag IgG Hilyte Fluor 555 (Anaspec, AS-61250-H555) and AlexaFluor 555 goat anti-chicken IgY (Thermo Fisher Scientific, A21437) for 20 min at 4 °C. Cells were then analyzed by flow cytometry (Accuri C6, BD Biosciences) after the final wash (three washes in total). For the additional washing conditions, one or two incubation/wash steps were added (incubated in 50 µl PBSA for 20–30 min at 4 °C and washed with 1 ml PBSA) between the anti-c-Myc primary antibody staining step and the secondary antibodies staining step.

Due to constitutive yEGFP expression, antibody staining conditions are simplified using the pCL vector. Yeast cells containing pCL-Axl were induced for expression, incubated with Gas6 for 17 h at room temperature, washed one time, and then labeled with secondary antibody in 50 µl PBSA containing a 1:100 dilution of mouse anti-His Tag IgG Hilyte Fluor 555. Only two wash steps are necessary for binding assays performed on proteins displayed using the pCL system, including the final wash before flow cytometry analysis. For the additional washing conditions, one or two incubation/wash steps were added between the target-binding step and the secondary antibody staining step.

To compare the expression signal changes in the pCT and pCL groups, the mean signal values were normalized to the original condition and shown as bar graphs (Supplementary Fig. 6B). For statistical analysis, p-values were calculated by Student's t-test (one-tailed, paired).

4. Stability of yEGFP expression levels

To measure the stability of the yEGFP signal over time, the assay was performed with 5×10^4 induced yeast cells transformed with pCL-nGFP-Aga2p-Axl. The cells were incubated in 50 μl of PBSA at room temperature and green fluorescence from yEGFP was measured using flow cytometry (Guava, Millipore) at 0, 12, 24, 36, 48, and 72 h in triplicate.

SI References

[1] Jones, D.S., Tsai, P.-C., Cochran, J.R., Engineering hepatocyte growth factor fragments with high stability and activity as Met receptor agonists and antagonists. *Proc. Natl. Acad. Sci. U. S. A.* 2011, *108*, 13035–13040.

[2] Kariolis, M.S., Miao, Y.R., Jones, D.S., Kapur, S., et al., An engineered Axl "decoy receptor" effectively silences the Gas6-Axl signaling axis. *Nat. Chem. Biol.* 2014, *10*, 977–983.

[3] VanAntwerp, J.J., Wittrup, K.D., Fine Affinity Discrimination by Yeast Surface Display and Flow Cytometry. *Biotechnol. Prog.* 2000, *16*, 31–37.

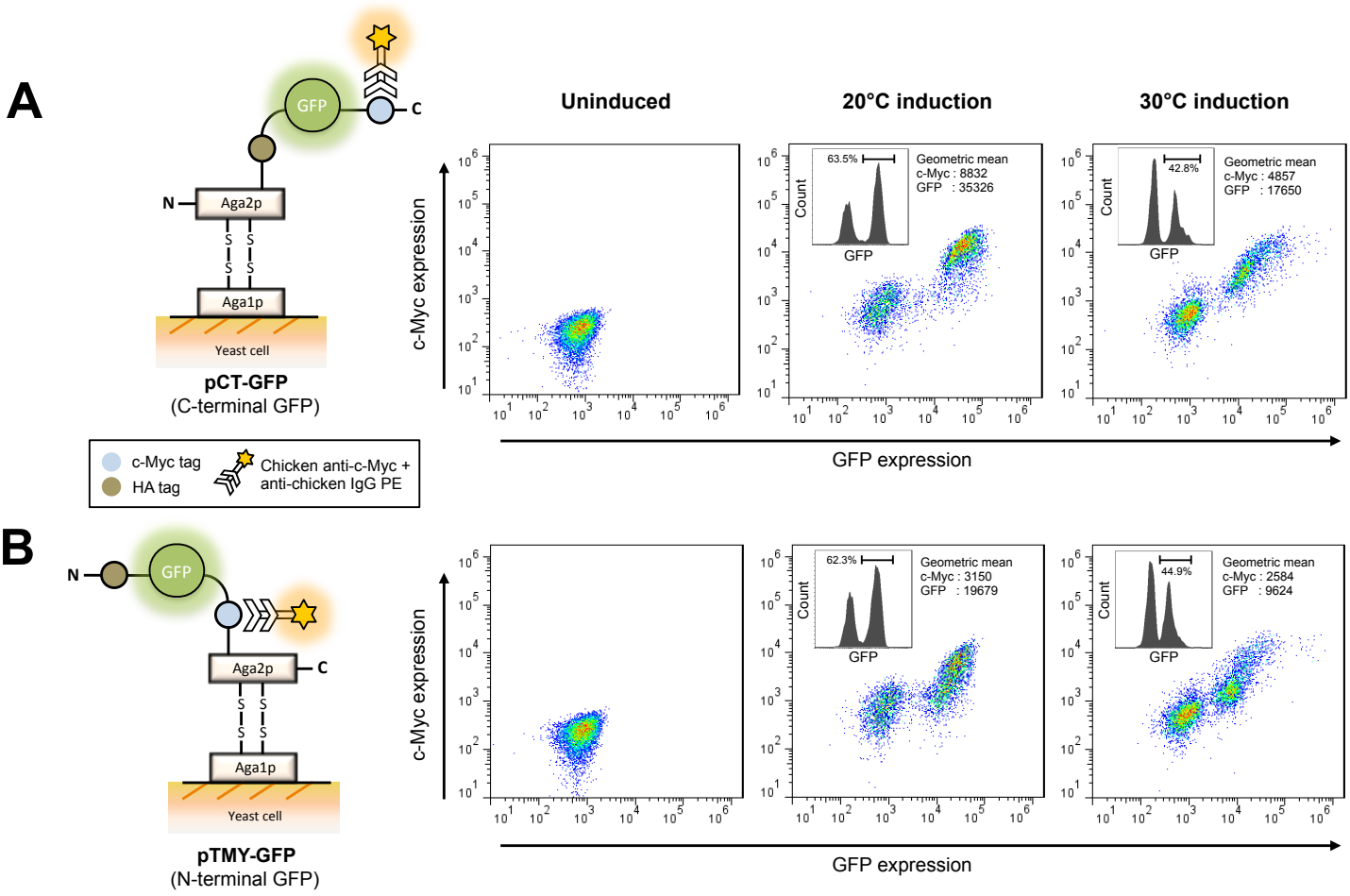
[4] Hawkins, R.E., Russell, S.J., Baier, M., Winter, G., The Contribution of Contact and Non-contact Residues of Antibody in the Affinity of Binding to Antigen: The Interaction of Mutant D1.3 Antibodies with Lysozyme. *J. Mol. Biol.* 1993, *234*, 958–964.

[5] Lee, S., Lim, W.A., Thorn, K.S., Improved Blue, Green, and Red Fluorescent Protein Tagging Vectors for S. cerevisiae. *PLoS One* 2013, *8*, e67902.

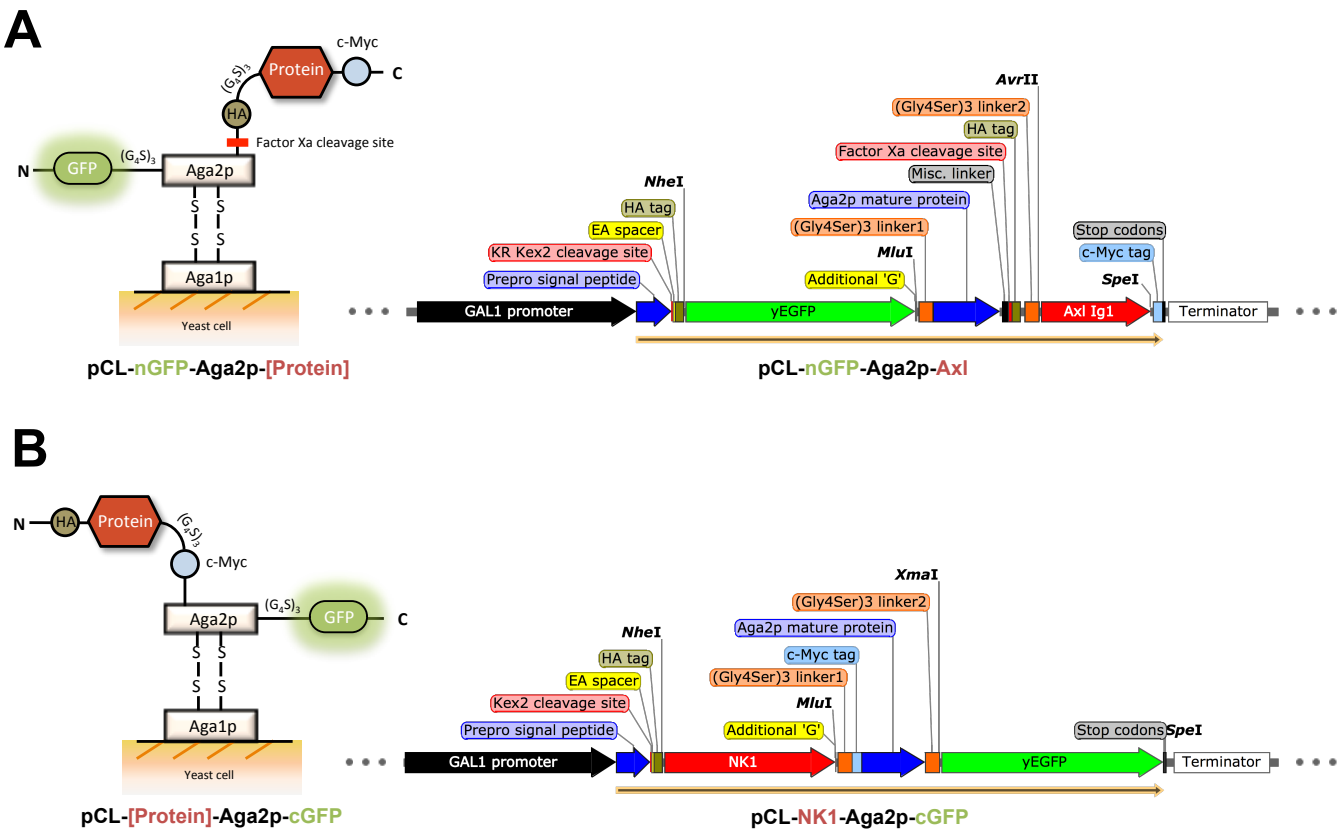
[6] Chen, I., Dorr, B.M., Liu, D.R., A general strategy for the evolution of bond-forming enzymes using yeast display. *Proc. Natl. Acad. Sci. U. S. A.* 2011, *108*, 11399–11404.

[7] Chao, G., Lau, W.L., Hackel, B.J., Sazinsky, S.L., et al., Isolating and engineering human antibodies using yeast surface display. *Nat. Protoc.* 2006, *1*, 755–768.

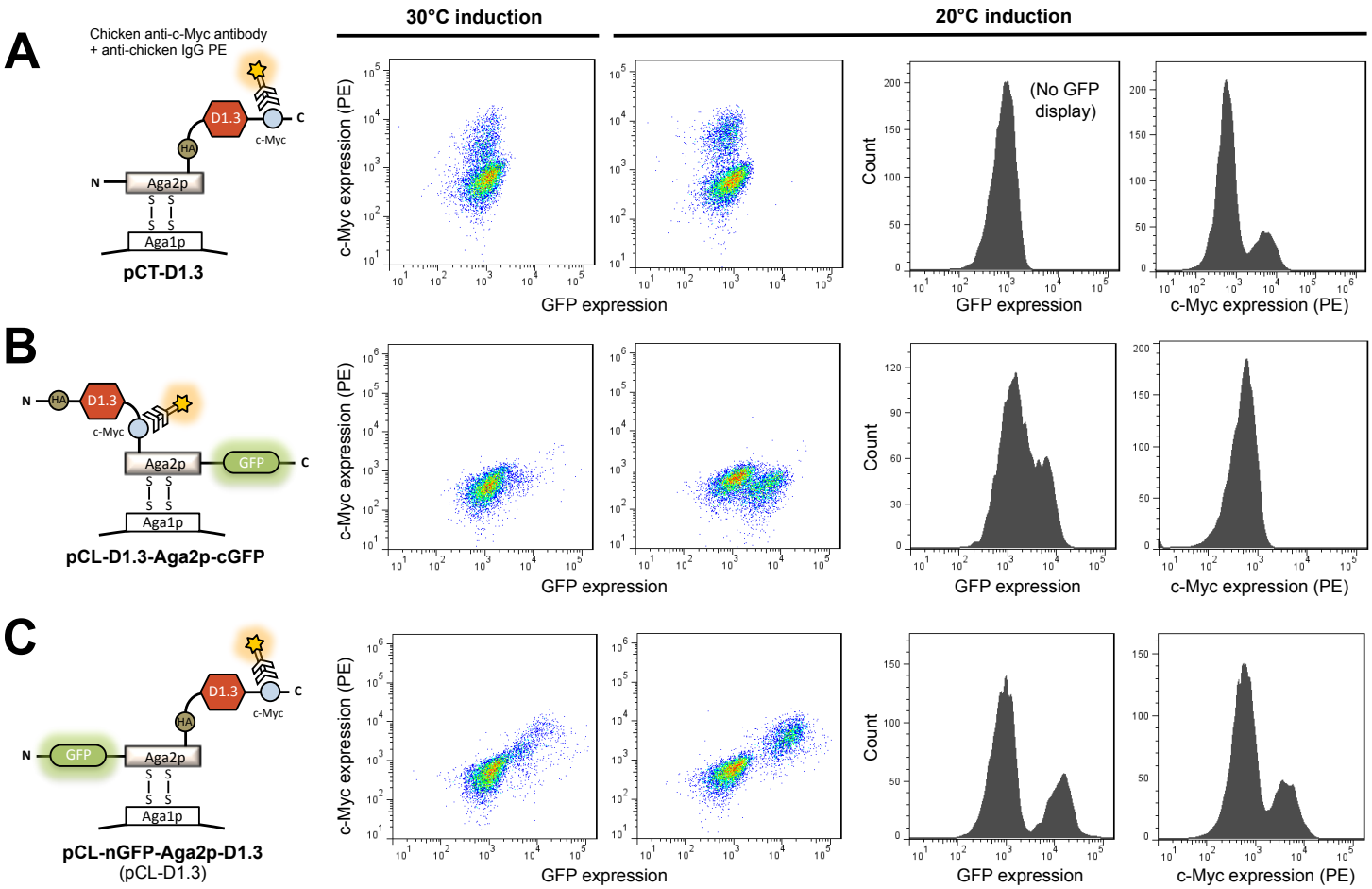
[8] Chen, B., Lim, S., Kannan, A., Alford, S.C., et al., High-throughput analysis and protein engineering using microcapillary arrays. *Nat. Chem. Biol.* 2015, *12*, 76–81.



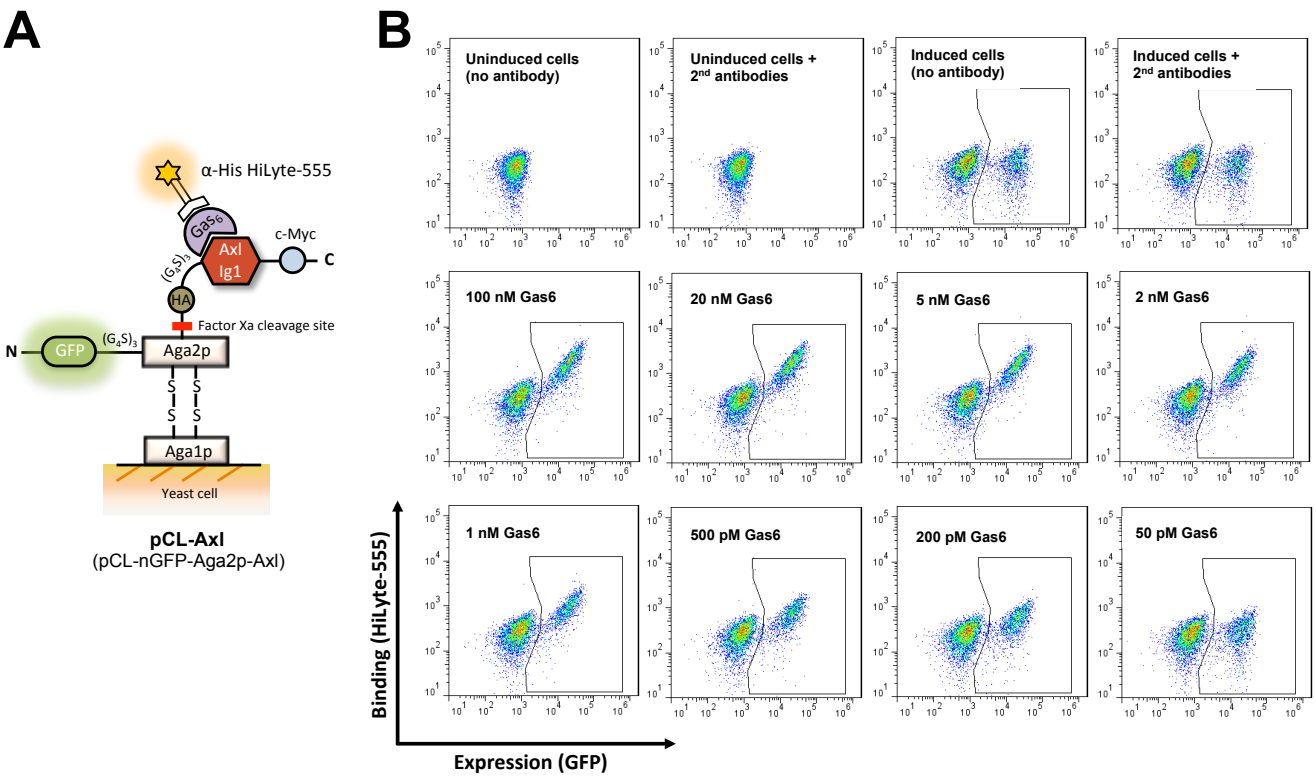
Supplementary Figure 1. Yeast-codon-optimized enhanced GFP (yEGFP) expression when fused at the C-terminus (A) or N-terminus (B) of Aga2p under different induction temperatures (20 °C or 30 °C). Yeast cells were stained with anti-c-Myc primary antibody (Thermo Fisher Scientific, A21281) followed by PE-labeled secondary antibody (Santa Cruz, sc-3730) to measure c-Myc expression levels. Insets show histogram representations of data, including quantification of geometric mean for c-Myc and GFP fluorescence.



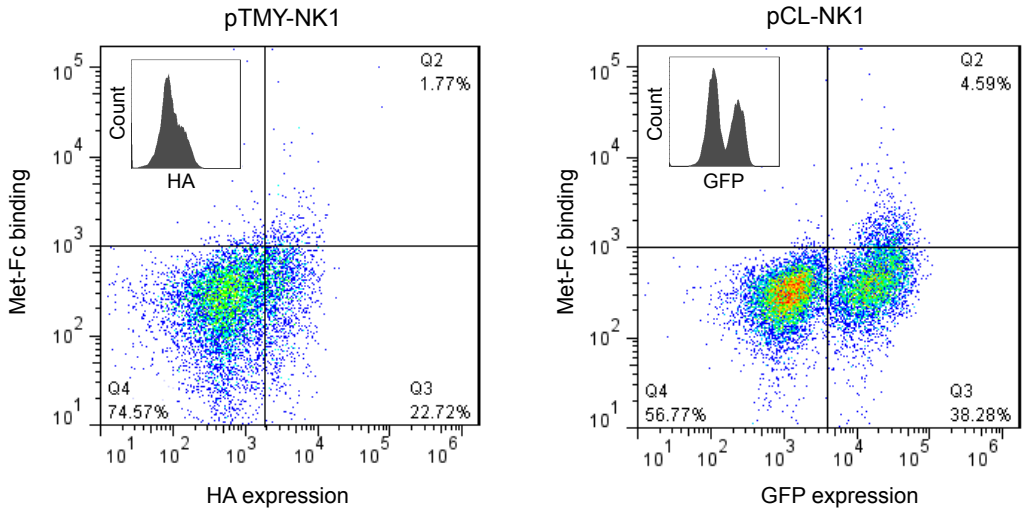
Supplementary Figure 2. Schematic diagrams and plasmid maps of pCL-nGFP (A) and pCL-cGFP (B). The Axl Ig1 domain (Axl) was cloned into pCL-nGFP and the NK1 domain of human HGF was cloned into pCL-cGFP as model proteins for binding assays.



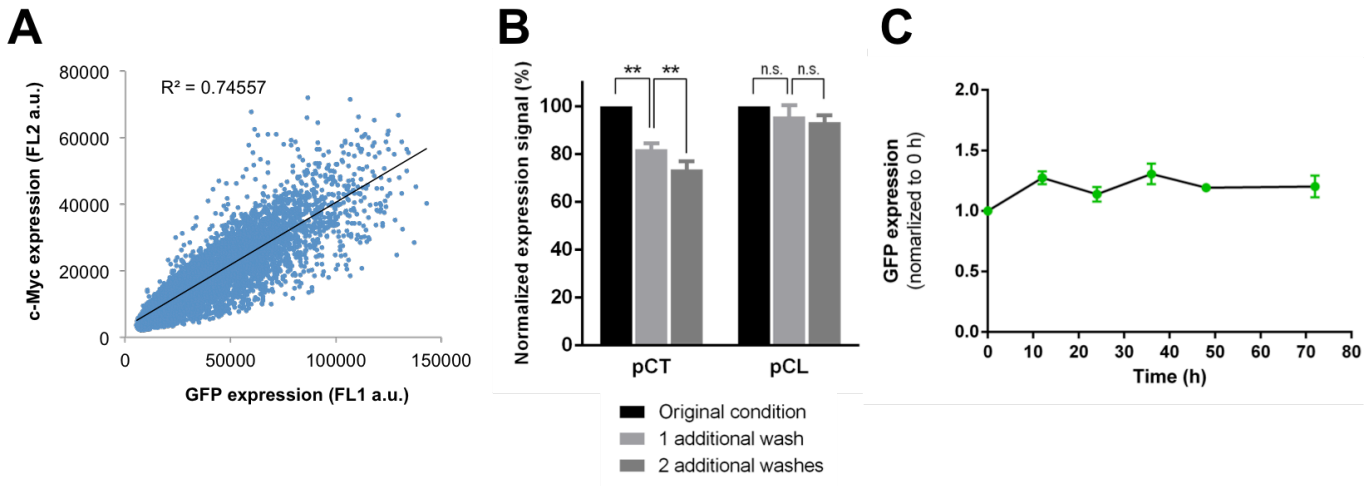
Supplementary Figure 3. Comparison of scFv D1.3 expression using pCT (A), pCL-cGFP (B), or pCL-nGFP (C) at different induction temperatures. C-terminal expression of the scFv D1.3 protein under induction at 20 °C was deemed most optimal.



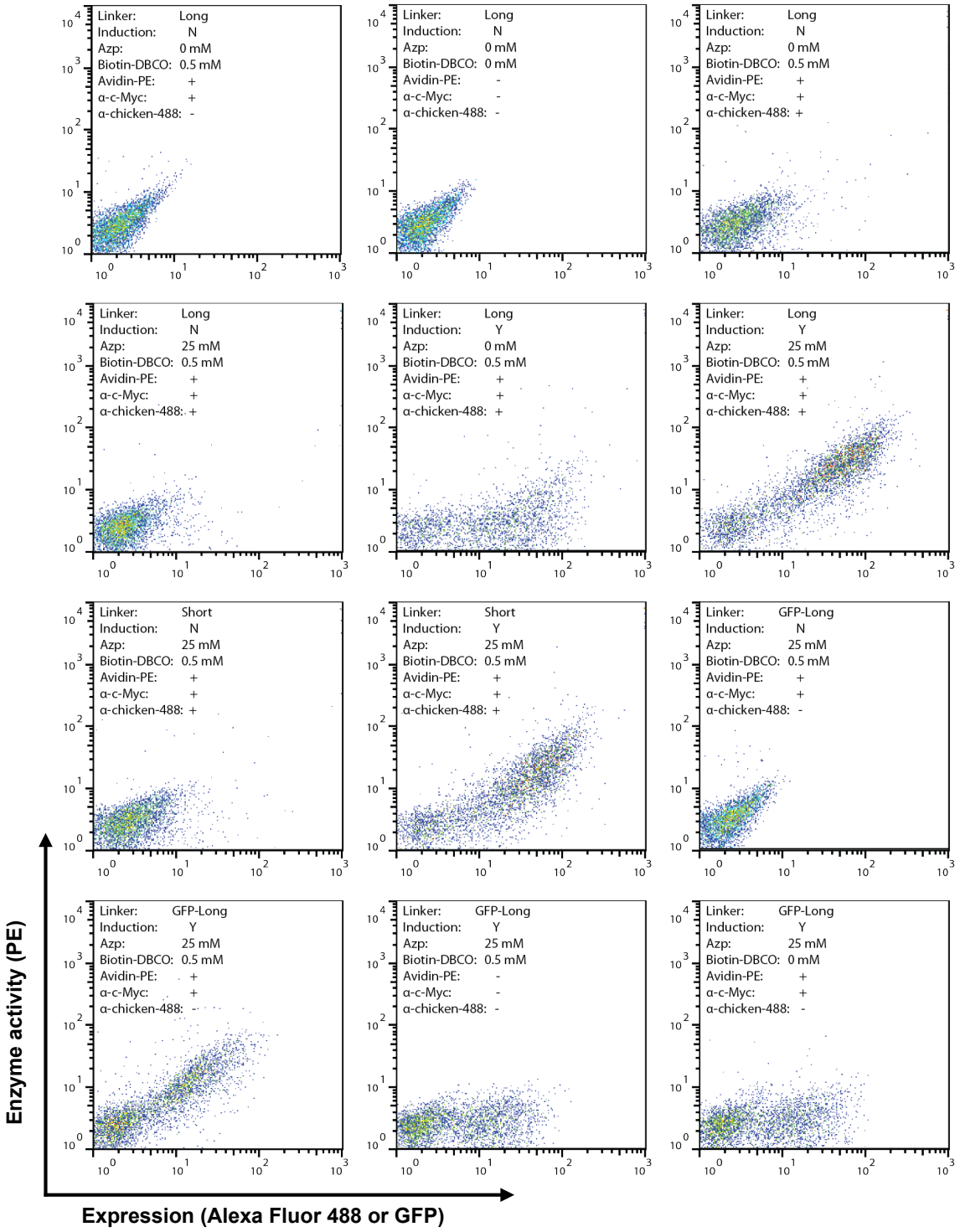
Supplementary Figure 4. (A) Antibody staining strategy for proteins expressed using pCL-nGFP-Aga2p-Axl (pCL-Axl). (B) Flow cytometry scatter plots of negative controls and cells incubated with different Gas6 concentrations ranging from 50 pM to 100 nM. In each scatter plot, the GFP-positive population is gated to calculate a geometric mean of binding signal at the designated Gas6 concentration.



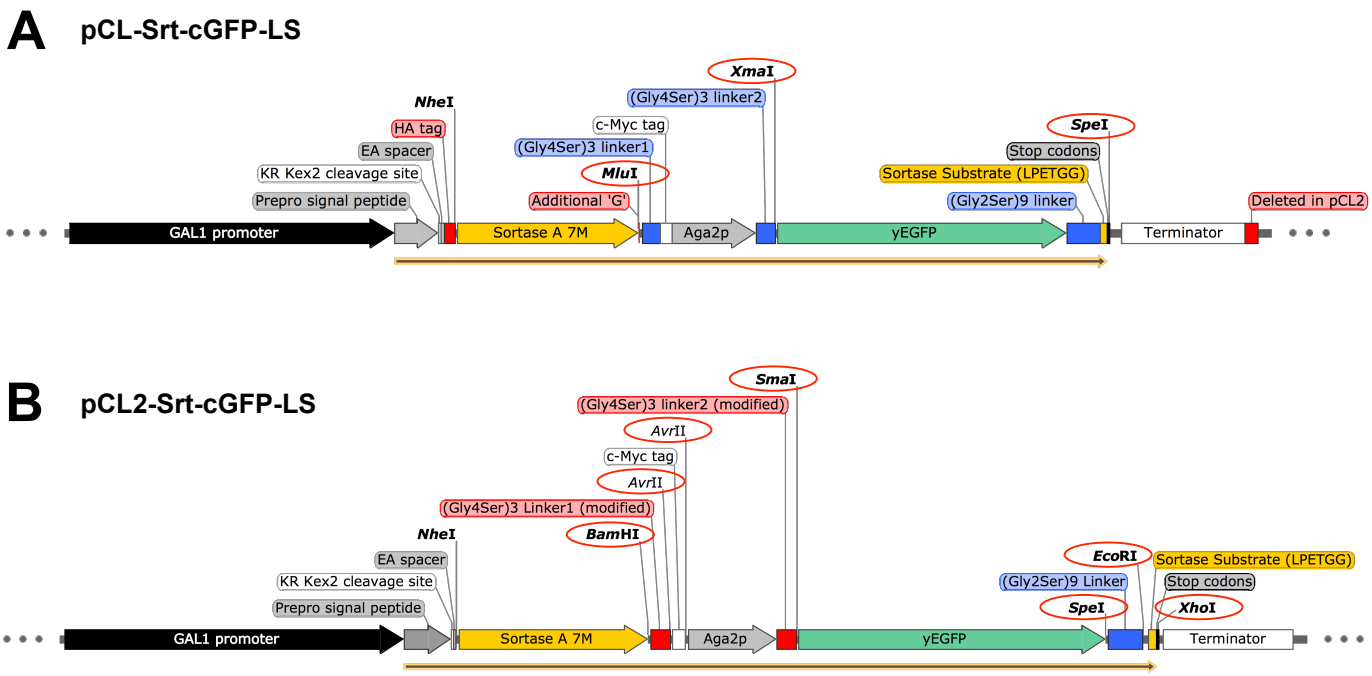
Supplementary Figure 5. A distinct yeast-displayed NK1 population is observed through detection of GFP-expression of the pCL vector compared to low NK1 expression observed with the pTMY vector, measured through antibody staining of the HA epitope tag. In both cases, expression signals were measured after incubating the yeast cells with 20 nM Met-Fc. With pCL-NK1, Met-Fc binding levels are also higher, suggesting that this construct enables improved expression of functional NK1 protein.



Supplementary Figure 6. Yeast-displayed yEGFP expression correlates with epitope tag expression on the same yeast cell and its signal intensity remains constant over time. (A) yEGFP fluorescence measured in yeast cells transformed with pCL-nGFP-Aga2p-Axl (pCL-Axl), correlated with fluorescence of C-terminal c-Myc expression measured by antibody staining. Cells were incubated with PBSA containing chicken anti-c-Myc antibody for 30 min at 4 °C, washed with 1 ml of PBSA, and then stained with PBSA containing AlexaFluor 555 goat anti-chicken IgY for 20 min at 4 °C. R^2 was calculated from simple linear regression of all the clones in an expression-positive population. (B) Comparison of expression signals between Axl Ig1 displayed using pCT or pCL vectors upon increasing the number of antibody-staining/washing steps. In each group, signals are normalized to the fluorescence intensity of the original condition. Error bars correspond to the standard deviation of six independent measurements (n = 6 for each of pCT and pCL groups). One-tailed, paired Student's t-tests were performed to analyze significance of the expressing signal change. **$p \leq 0.01$; n.s. for $p > 0.05$. (C) yEGFP yeast surface expression levels measured over 70 h incubation of yeast at room temperature. In these studies, the Axl Ig1 protein is expressed using the pCT or the pCL-nGFP-Aga2p-Axl vector. Error bars correspond to the standard deviation of triplicate samples.



Supplementary Figure 7. Flow cytometry scatter plots under various conditions for sortase-LPETGG bioconjugation using pCL-Srt-LS (substrate tethered using a long linker), pCL-Srt-SS (substrate tethered using a short linker), and pCL-Srt-cGFP-LS (substrate tethered through C-terminal GFP plus a long linker).



Supplementary Figure 8. Comparison between the original pCL vector (SEQ.1) and pCL2 (SEQ.2), which has been optimized for facile and modular cloning and library construction. Modified features are highlighted in red blocks (removal of N-terminal HA epitope tag and the need for an extra base pair, and further codon-optimization of the linkers) and red ovals (restriction site additions and modifications).

SEQ.1: pCL-nSrt-cGFP-LS

gtaataacctatttcttagcatttttgacgaaatttgctattttgttagagtcttttacaccatttgtctccacacctccgcttacatcaacaccaataacgc
catttaatctaagcgcatcaccaacattttctggcgtcagtccaccagctaacataaaatgtaagctttcggggctctcttgccttccaacccagtca
gaaatcgagttccaatccaaaagttcacctgtcccacctgcttctgaatcaaacaagggaataaacgaatgaggtttctgtgaagctgcactgagt
agtatgttgcagtcttttggaaatacgagtcttttaataactggcaaaccgaggaactcttggtattcttgccacgactcatctccatgcagttggacg
atatcaatgccgtaatcattgaccagagccaaaacatcctccttaggttgattacgaaacacgccaaccaagtatttcggagtgcctgaactatttttt
atatgcttttacaagacttgaaattttccttgcaataaccgggtcaattgttctctttctattgggcacacatataatacccagcaagtcagcatcggaa
tctagagcacattctgcggcctctgtgctctgcaagccgcaaactttcaccaatggaccagaactacctgtgaaattaataacagacatactccaa
gctgcctttgtgtgcttaatcacgtatactcacgtgctcaatagtcaccaatgccctccctcttggccctctccttttctttttttcgaccgaattaattctta
atcggcaaaaaaagaaaagctccggatcaagattgtacgtaaggtgacaagctatttttcaataaagaatatcttccactactgccatctggcgtca
taactgcaaagtacacatatattacgatgctgttctattaaatgcttcctatattatatatatagtaatgtcgtgatctatggtgcactctcagtacaatct
gctctgatgccgcatagttaagccagccccgacacccgccaacacccgctgacgcgccctgacgggcttgtctgctcccggcatccgcttaca
gacaagctgtgaccgtctccgggagctgcatgtgtcagaggttttcaccgtcatcaccgaaacgcgcgagacgaaagggcctcgtgatacgcc
tatttttataggttaatgtcatgataataatggtttcttagacggatcgcttgcctgtaacttacacgcgcctcgtatcttttaatgatggaataatttggg
aatttactctgtgtttatttatttttatgttttgtatttggattttagaaagtaaataaagaaggtagaagagttacggaatgaagaaaaaaaaataaacaa
aggtttaaaaaatttcaacaaaaagcgtactttacatatatatttattagacaagaaaagcagattaaatagatatacattcgattaacgataagtaaa
atgtaaaatcacaggattttcgtgtgtggtcttctacacagacaagatgaaacaattcggcattaatacctgagagcaggaagagcaagataaaa
ggtagtatttgttggcgatccccctagagtcttttacatcttcggaaaacaaaaactatttttttctttaatttctttttttactttctatttttaatttatatatttat
attaaaaaatttaaattataattatttttatagcacgtgatgaaaaggacccaggtggcacttttcggggaaatgtgcgcggaacccctatttgtttattt
ttctaaatacattcaaatatgtatccgctcatgagacaataaccctgataaatgcttcaataatattgaaaaaggaagagtatgagtattcaacatttcc
gtgtcgcccttattcccttttttgcggcattttgccttcctgtttttgctcacccagaaacgctggtgaaagtaaaagatgctgaagatcagttgggtgc
acgagtgggttacatcgaactggatctcaacagcggtaagatccttgagagttttcgccccgaagaacgttttccaatgatgagcacttttaaagtt
ctgctatgtggcgcggtattatcccgtattgacgccgggcaagagcaactcggtcgccgcatacactattctcagaatgacttggttgagtactca
ccagtcacagaaaagcatcttacggatggcatgacagtaagagaattatgcagtgctgccataaccatgagtgataacactgcggccaacttact
tctgacaacgatcggaggaccgaaggagctaaccgcttttttgcacaacatgggggatcatgtaactcgccttgatcgttgggaaccggagctg
aatgaagccataccaaacgacgagcgtgacaccacgatgcctgtagcaatggcaacaacgttgcgcaaactattaactggcgaactacttactc
tagcttcccggcaacaattaatagactggatggaggcggataaagttgcaggaccacttctgcgctcggcccttccggctggctggtttattgctg
ataaatctggagccggtgagcgtgggtctcgcggtatcattgcagcactggggccagatggtaagccctcccgtatcgtagttatctacacgacg
gggagtcaggcaactatggatgaacgaaatagacagatcgctgagataggtgcctcactgattaagcattggtaactgtcagaccaagtttactc
atatatactttagattgatttaaaacttcatttttaatttaaaaggatctaggtgaagatccttttgataatctcatgaccaaaatcccttaacgtgagtttt
cgttccactgagcgtcagaccccgtagaaaagatcaaaggatcttcttgagatcctttttttctgcgcgtaatctgctgcttgcaaacaaaaaaacc
accgctaccagcggtggtttgtttgccggatcaagagctaccaactctttttccgaaggtaactggcttcagcagagcgcagataccaaatactgt
tcttctagtgtagccgtagttaggccaccacttcaagaactctgtagcaccgcctacatacctcgctctgctaatcctgttaccagtggctgctgcca
gtggcgataagtcgtgtcttaccgggttggactcaagacgatagttaccggataaggcgcagcggtcgggctgaacgggggttcgtgcacac
agcccagcttggagcgaacgacctacaccgaactgagatacctacagcgtgagctatgagaaagcgccacgcttcccgaagggagaaaggc
ggacaggtatccggtaagcggcagggtcggaacaggagagcgcacgagggagcttccagggggaaacgcctggtatctttatagtcctgtcg
ggtttcgccacctctgacttgagcgtcgatttttgtgatgctcgtcaggggggcggagcctatggaaaaacgccagcaacgcggccttttacggt
tcctggccttttgctggccttttgctcacatgttctttcctgcgttatcccctgattctgtggataaccgtattaccgcctttgagtgagctgataccgctc
gccgcagccgaacgaccgagcgcagcgagtcagtgagcgaggaagcggaagagcgcccaatacgcaaaccgcctctccccgcgcgttgg
ccgattcattaatgcagctggcacgacaggtttcccgactggaaagcgggcagtgagcgcaacgcaattaatgtgagttagctcactcattaggc
accccaggctttacactttatgcttccggctcgtatgttgtgtggaattgtgagcggataacaatttcacacaggaaacagctatgaccatgattacg
ccaagctcgaaattaaccctcactaaagggaacaaaagctggtacccctcgatcgaggtcgacggtatcgatgagcttgatatcgaattccttga
attttcaaaaattcttacttttttttggatggacgcaaagaagtttaataatcatattacatggcattaccaccatatacatatccatatacatatccatatc
taatcttacttatatgttgtggaaatgtaaagagccccattatcttagcctaaaaaaaccttctctttggaactttcagtaatacgcttaactgctcattgc
tatattgaagtacggattagaagccgccgagcgggtgacagccctccgaaggaagactctcctccgtgcgtcctcgtcttcaccggtcgcgttcc
tgaaacgcagatgtgcctcgcgccgcactgctccgaacaataaagattctacaatactagcttttatggttatgaagaggaaaaattggcagtaac
ctggccccacaaaccttcaaatgaacgaatcaaattaacaaccataggatgataatgcgattagttttttagccttatttctggggtaattaatcagcg
aagcgatgatttttgatctattaacagatatataaatgcaaaaactgcataaccactttaactaatactttcaacattttcggtttgtattacttcttattcaa
atgtaataaaagtatcaacaaaaaattgttaatatacctctatactttaacgtcaaggagaaaaaaccccggatcatgaaggttttgattgtcttgttg

gctatcttcgctgctttgccattggccttagctcaaccggttatttctactaccgtcggttccgctgcagaaggctctttggacaagagagaagctta
cccatacgatgttccagattacgctgctagccaagctaaacctcaaattccgaaagataaatcaaaagtggcaggctatattgaaattccagatgct
gatattaaagaaccagtatatccaggaccagcaacacgcgaacaattaaatagaggtgtaagctttgcaaaagaaaatcaatcactagatgatca
aaatatttcaattgcaggacacactttcattgaccgtccgaactatcaatttacaaatcttaaagcagccaaaaaaggtagtatggtgtactttaaagt
tggtaatgaaacacgtaagtataaaatgacaagtataagaaacgttaagccaacagctgtagaagttctagatgaacaaaaaggtaaagataaac
aattaacattaattacttgtgatgattacaatgaagagacaggcgtttgggaaacacgtaaaatctttgtagctacagaagtcaaactcgaggacgc
gtcaggtggtggaggctctggtggcggtggatctggcggtggaggttctgaacaaaagcttatctccgaagaagacttgcaggaactgacaact
atatgcgagcaaatcccctcaccaactttagaatcgacgccgtactctttgtcaacgactactattttggccaacgggaaggcaatgcaaggagtt
tttgaatattacaaatcagtaacgtttgtcagtaattgcggttctcacccctcaacaactagcaaaggcagccccataaacacacagtatgtttttggt
ggtggtggttctggtggtggtggttctggtggtggtggttctcccgggatgtctaaaggtgaagaattattcactggtgttgtcccaattttggttgaa
ttagatggtgatgttaatggtcacaaattttctgtctccggtgaaggtgaaggtgatgctacttacggtaaattgaccttaaaatttatttgtactactgg
taaattgccagttccatggccaaccttagtcactactttcggttatggtgttcaatgttttgcgagatacccagatcatatgaaacaacatgacttttca
agtctgccatgccagaaggttatgttcaagaaagaactatttttttcaaagatgacggtaactacaagaccagagctgaagtcaagtttgaaggtg
ataccttagttaatagaatcgaattaaaaggtattgattttaaagaagatggtaacattttaggtcacaaattggaatacaactataactctcacaatgt
ttacatcatggctgacaaacaaaagaatggtatcaaagttaacttcaaaattagacacaacattgaagatggttctgttcaattagctgaccattatca
acaaaatactccaattggtgatggtccagtcttgttaccagacaaccattacttatccactcaatctaagttatccaaagatccaaacgaaaagaga
gaccacatggtcttgttagaatttgttactgctgctggtattacccatggtatggatgaattgtacaaaggtggatcaggaggttctggtggatctggt
ggttcaggtggttcaggtggttctggaggttctggaggttctggtggatctttgccagaaactggtggttaatagactagtatgtctaaaggctcga
gcgaaatctgataacaacagtgtagatgtaacaaaatcgactttgttcccactgtacttttagctcgtacaaaatacaatatacttttcatttctccgtaa
acaacatgttttcccatgtaatatccttttctatttttcgttccgttaccaactttacacatactttatatagctattcacttctatacactaaaaaactaagac
aattttaattttgctgcctgccatatttcaatttgttataaattcctataatttatcctattagtagctaaaaaaagatgaatgtgaatcgaatcctaagaga
attccggatcctctagagcggccgccaccgcggtggagctccaattcgccctatagtgagtcgtattacaattcactggccgtcgttttacaacgt
cgtgactgggaaaaccctggcgttacccaacttaatcgccttgcagcacatccccctttcgccagctggcgtaatagcgaagaggcccgcaccg
atcgcccttcccaacagttgcgcagcctgaatggcgaatggcgcgacgcgccctgtagcggcgcattaagcgcggcgggtgtggtggttacg
cgcagcgtgaccgctacacttgccagcgccctagcgcccgctcctttcgctttcttcccttcctttctcgccacgttcgccggctttccccgtcaag
ctctaaatcgggggctccctttagggttccgatttagtgctttacggcacctcgaccccaaaaaacttgattagggtgatggttcacgtagtgggcc
atcgccctgatagacggtttttcgccctttgacgttggagtccacgttctttaatagtggactcttgttccaaactggaacaacactcaaccctatctc
ggtctattcttttgatttataagggattttgccgatttcggcctattggttaaaaaatgagctgatttaacaaaaatttaacgcgaattttaacaaaatatt
aacgcttacaatttcctgatgcggtattttctccttacgcatctgtgcggtatttcacaccgcatagatcggcaagtgcacaaacaatacttaaataaa
tactactca

SEQ.2: pCL2-nSrt-cGFP-LS

gtaataacctatttcttagcatttttgacgaaatttgctattttgttagagtcttttacaccatttgtctccacacctccgcttacatcaacaccaataacgc
catttaatctaagcgcatcaccaacattttctggcgtcagtccaccagctaacataaaatgtaagctttcggggctctcttgccttccaacccagtca
gaaatcgagttccaatccaaaagttcacctgtcccacctgcttctgaatcaaacaagggaataaacgaatgaggtttctgtgaagctgcactgagt
agtatgttgcagtcttttggaaatacgagtcttttaataactggcaaaccgaggaactcttggtattcttgccacgactcatctccatgcagttggacg
atatcaatgccgtaatcattgaccagagccaaaacatcctccttaggttgattacgaaacacgccaaccaagtatttcggagtgcctgaactatttttt
atatgcttttacaagacttgaaattttccttgcaataaccgggtcaattgttctctttctattgggcacacatataatacccagcaagtcagcatcggaa
tctagagcacattctgcggcctctgtgctctgcaagccgcaaactttcaccaatggaccagaactacctgtgaaattaataacagacatactccaa
gctgcctttgtgtgcttaatcacgtatactcacgtgctcaatagtcaccaatgccctccctcttggccctctccttttctttttttcgaccgaattaattctta
atcggcaaaaaaagaaaagctccggatcaagattgtacgtaaggtgacaagctatttttcaataaagaatatcttccactactgccatctggcgtca
taactgcaaagtacacatatattacgatgctgtctattaaatgcttcctatattatatatatagtaatgtcgtttatggtgcactctcagtacaatctgctct
gatgccgcatagttaagccagccccgacacccgccaacacccgctgacgcgccctgacgggcttgtctgctcccggcatccgcttacagacaa
gctgtgaccgtctccgggagctgcatgtgtcagaggttttcaccgtcatcaccgaaacgcgcgagacgaaagggcctcgtgatacgcctatttttt
ataggttaatgtcatgataataatggtttcttaggacggatcgcttgcctgtaacttacacgcgcctcgtatcttttaatgatggaataatttgggaattt
actctgtgtttatttatttttatgttttgtatttggattttagaaagtaaataaagaaggtagaagagttacggaatgaagaaaaaaaaataaacaaaggt
ttaaaaaatttcaacaaaaagcgtactttacatatatatttattagacaagaaaagcagattaaatagatatacattcgattaacgataagtaaaatgta
aaatcacaggattttcgtgtgtggtcttctacacagacaagatgaaacaattcggcattaatacctgagagcaggaagagcaagataaaaggtag
tatttgttggcgatccccctagagtcttttacatcttcggaaaacaaaaactatttttttctttaatttctttttttactttctatttttaatttatatatttatattaaa
aaatttaaattataattatttttatagcacgtgatgaaaaggacccaggtggcacttttcggggaaatgtgcgcggaacccctatttgtttatttttctaa
atacattcaaatatgtatccgctcatgagacaataaccctgataaatgcttcaataatattgaaaaaggaagagtatgagtattcaacatttccgtgtc
gcccttattcccttttttgcggcattttgccttcctgtttttgctcacccagaaacgctggtgaaagtaaaagatgctgaagatcagttgggtgcacga
gtgggttacatcgaactggatctcaacagcggtaagatccttgagagttttcgccccgaagaacgttttccaatgatgagcacttttaaagttctgct
atgtggcgcggtattatcccgtattgacgccgggcaagagcaactcggtcgccgcatacactattctcagaatgacttggttgagtactcaccagt
cacagaaaagcatcttacggatggcatgacagtaagagaattatgcagtgctgccataaccatgagtgataacactgcggccaacttacttctga
caacgatcggaggaccgaaggagctaaccgcttttttcacaacatgggggatcatgtaactcgccttgatcgttgggaaccggagctgaatga
agccataccaaacgacgagcgtgacaccacgatgcctgtagcaatggcaacaacgttgcgcaaactattaactggcgaactacttactctagctt
cccggcaacaattaatagactggatggaggcggataaagttgcaggaccacttctgcgctcggcccttccggctggctggtttattgctgataaat
ctggagccggtgagcgtgggtctcgcggtatcattgcagcactggggccagatggtaagccctcccgtatcgtagttatctacacgacgggca
gtcaggcaactatggatgaacgaaatagacagatcgctgagataggtgcctcactgattaagcattggtaactgtcagaccaagtttactcatatat
actttagattgatttaaaacttcatttttaatttaaaaggatctaggtgaagatccttttgataatctcatgaccaaaatcccttaacgtgagttttcgttcc
actgagcgtcagaccccgtagaaaagatcaaaggatcttcttgagatcctttttttctgcgcgtaatctgctgcttgcaaacaaaaaaaccaccgct
accagcggtggtttgtttgccggatcaagagctaccaactctttttccgaaggtaactggcttcagcagagcgcagataccaaatactgtccttcta
gtgtagccgtagttaggccaccacttcaagaactctgtagcaccgcctacatacctcgctctgctaatcctgttaccagtggctgctgccagtggc
gataagtcgtgtcttaccgggttggactcaagacgatagttaccggataaggcgcagcggtcgggctgaacgggggttcgtgcacacagccc
agcttggagcgaacgacctacaccgaactgagatacctacagcgtgagcattgagaaagcgccacgcttcccgaagggagaaaggcggaca
ggtatccggtaagcggcagggtcggaacaggagagcgcacgagggagcttccagggggaacgcctggtatctttatagtcctgtcgggtttc
gccacctctgacttgagcgtcgatttttgtgatgctcgtcaggggggccgagcctatggaaaaacgccagcaacgcggccttttacggttcctg
gcctttgctggccttttgctcacatgttctttcctgcgttatcccctgattctgtggataaccgtattaccgcctttgagtgagctgataccgctcgccg
cagccgaacgaccgagcgcagcgagtcagtgagcgaggaagcggaagagcgcccaatacgcaaaccgcctctccccgcgcgttggccga
ttcattaatgcagctggcacgacaggtttcccgactggaaagcgggcagtgagcgcaacgcaattaatgtgagttacctcactcattaggcaccc
caggctttacactttatgcttccggctcctatgttgtgtggaattgtgagcggataacaatttcacacaggaaacagctatgaccatgattacgccaa
gctcggaattaaccctcactaaagggaacaaaagctgggtacccgacaggttatcagcaacaacacagtcatatccattctcaattagctctacca
cagtgtgtgaaccaatgtatccagcaccacctgtaaccaaaacaattttagaagtactttcactttgtaactgagctgtcatttatattgaattttcaaaa
attcttactttttttttggatggacgcaaagaagtttaataatcatattacatggcattaccaccatatacatatccatatctaatcttacttatatgttgtgg
aaatgtaaagagccccattatcttagcctaaaaaaaccttctctttggaactttcagtaatacgcttaactgctcattgctatattgaagtacggattag
aagccgccgagcgggtgacagccctccgaaggaagactctcctccgtgcgtcctcgtcttcaccggtcgcgttcctgaaacgcagatgtgcct
cgcgccgcactgctccgaacaataaagattctacaatactagcttttatggttatgaagaggaaaaattggcagtaacctggccccacaaaccttc
aaatgaacgaatcaaattaacaaccataggatgataatgcgattagttttttagccttatttctggggtaattaatcagcgaagcgatgatttttgatct
attaacagatatataaatgcaaaaactgcataaccactttaactaatactttcaacattttcggtttgtattacttcttattcaaatgtaataaaagtatcaa

caaaaaattgttaatatacctctatactttaacgtcaaggagaaaaaaccccggatcatgaaggttttgattgtcttgttggctatcttcgctgctttgc
cattggccttagctcaaccggttatttctactaccgtcggttccgctgcagaaggctctttggacaagagagaagctgctagccaagctaaacctc
aaattccgaaagataaatcaaaagtggcaggctatattgaaattccagatgctgatattaaagaaccagtatatccaggaccagcaacacgcgaa
caattaaatagaggtgtaagctttgcaaaagaaaatcaatcactagatgatcaaaatatttcaattgcaggacacactttcattgaccgtccgaacta
tcaatttacaaatcttaaagcagccaaaaaaggtagtatggtgtactttaaagttggtaatgaaacacgtaagtataaaatgacaagtataagaaac
gttaagccaacagctgtagaagttctagatgaacaaaaaggtaaagataaacaattaacattaattacttgtgatgattacaatgaagagacaggc
gtttgggaaacacgtaaaatctttgtagctacagaagtcaaaggatccggcggcggaggctccggtggaggcggaagcggcggtggaggta
gtcctagggaacaaaagcttatctccgaagaagacttgcctaggcaggaactgacaactatatgcgagcaaatcccctcaccaactttagaatcg
acgccgtactctttgtcaacgactactattttggccaacgggaaggcaatgcaaggagtttttgaatattacaaatcagtaacgtttgtcagtaattgc
ggttctcacccctcaacaactagcaaaggcagccccataaacacacagtatgtttttggtggtggtggttctggtggtggtggttctggtggtggtg
gttctcccgggatgtctaaaggtgaagaattattcactggtgttgtcccaattttggttgaattagatggtgatgttaatggtcacaaattttctgtctcc
ggtgaaggtgaaggtgatgctacttacggtaaattgaccttaaaatttatttgtactactggtaaattgccagttccatggccaaccttagtcactactt
tcggttatggtgttcaatgttttgcgagatacccagatcatatgaaacaacatgacttttcaagtctgccatgccagaaggttatgttcaagaaagaa
ctatttttttcaaagatgacggtaactacaagaccagagctgaagtcaagtttgaaggtgataccttagttaatagaatcgaattaaaaggtattgatt
ttaaagaagatggtaacattttaggtcacaaattggaatacaactataactctcacaatgtttacatcatggctgacaaacaaaagaatggtatcaaa
gttaacttcaaaattagacacaacattgaagatggttctgttcaattagctgaccattatcaacaaaatactccaattggtgatggtccagtcttgttac
cagacaaccattacttatccactcaatctaagttatccaaagatccaaacgaaaagagagaccacatggtcttgttagaatttgttactgctgctggt
attacccatggtatggatgaattgtacaaaactagtggtgggtccggtggtagcggagggtcaggcggaagcggaggaagcggcgggtctgg
aggaagcggaggtagtggaggatctgaattcggcggcttgccagaaactggcggctaatagctcgagatctgataacaacagtgtagatgtaa
caaaatcgactttgttcccactgtacttttagctcgtacaaaatacaatatacttttcatttctccgtaaacaacatgttttcccatgtaatatccttttctatt
tttcgttccgttaccaactttacacatactttatatagctattcacttctatacactaaaaaactaagacaatttttaattttgctgcctgccatatttcaatttg
ttataaattcctataatttatcctattagtagctaaaaaaagatgaatgtgaatcgaatcctaagagaattgagctccaattcgccctatagtgagtcgt
attacaattcactggccgtcgttttacaacgtcgtgactgggaaaaccctggcgttacccaacttaatcgccttgcagcacatccccccttcgccag
ctggcgtaatagcgaagaggcccgcaccgatcgcccttcccaacagttgcgcagcctgaatggcgaatggcgcgacgcgccctgtagcggc
gcattaagcgcggcgggtgtggtggttacgcgcagcgtgaccgctacacttgccagcgccctagcgcccgctcctttcgctttcttcccttcctttt
ctcgccacgttcgccggctttccccgtcaagctctaaatcgggggctccctttagggttccgatttagtgctttacggcacctcgaccccaaaaaa
cttgattagggtgatggttcacgtagtgggccatcgccctgatagacggtttttcgccctttgacgttggagtccacgttctttaatagtggactcttgt
tccaaactggaacaacactcaaccctatctcggtctattcttttgatttataagggattttgccgatttcggcctattggttaaaaaatgagctgatttaa
caaaaatttaacgcgaattttaacaaaatattaacgtttacaatttcctgatgcggtattttctccttacgcatctgtgcggtatttcacaccgcaggca
agtgcacaaacaatacttaaataaatactactca