ACS SyntheticBiology

pubs.acs.org/synthbio

Research Article

YESS 2.0, a Tunable Platform for Enzyme Evolution, Yields Highly Active TEV Protease Variants

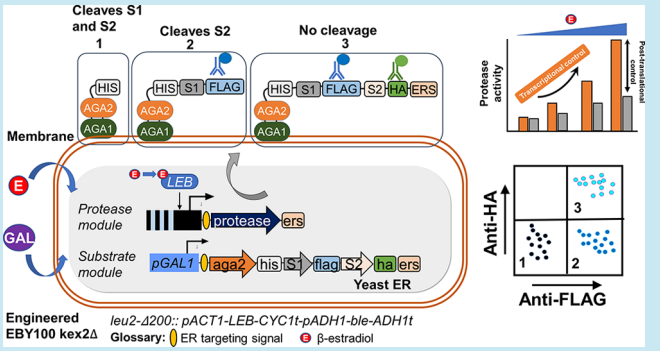
Carl A. Denard,* Chelsea Paresi, Rasha Yaghi, Natalie McGinnis, Zachary Bennett, Li Yi, George Georgiou, and Brent L. Iverson*

Cite This: https://dx.doi.org/10.1021/acssynbio.0c00452

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Here we describe YESS 2.0, a highly versatile version of the yeast endoplasmic sequestration screening (YESS) system suitable for engineering and characterizing protein/peptide modifying enzymes such as proteases with desired new activities. By incorporating features that modulate gene transcription as well as substrate and enzyme spatial sequestration, YESS 2.0 achieves a significantly higher operational and dynamic range compared with the original YESS. To showcase the new advantages of YESS 2.0, we improved an already efficient TEV protease variant (TEV-EAV) to obtain a variant (eTEV) with a 2.25-fold higher catalytic efficiency, derived almost entirely from an increase in turnover rate (k_{cat}). In our analysis, eTEV specifically digests a fusion protein in 2 h at a low 1:200 enzyme to substrate ratio. Structural modeling indicates that the increase in catalytic efficiency of eTEV is likely due to an enhanced interaction between the catalytic Cys151 with the P1 substrate residue (Gln). Furthermore, the modeling showed that the ENLYFQS peptide substrate is buried to a larger extent in the active site of eTEV compared with WT TEV. The new eTEV variant is functionally the fastest TEV variant reported to date and could potentially improve efficiency in any TEV application.



Proteases are ubiquitous in biology, playing major roles in initiating, regulating, and terminating cellular processes. The diversity and breadth of protease substrate preferences have been harnessed for a variety of applications, including protease therapeutics,[1] protein purification,[2] and mass spectrometry-based proteomics.[3] Proteases are also being utilized for interrogating protein−protein interactions,[4,5] imaging newly synthesized proteins,[6] and designing synthetic genetic[7,8] and protein circuits.[9−11]

To successfully repurpose proteases for these applications, one often needs to engineer protease catalytic properties using Directed Evolution. Protease variants with the desired phenotypes are isolated from a large pool, so a high-throughput protease screening system should ideally exhibit a high operational range (sensitivity over a large variation in input), and a high dynamic range (signal-to-noise ratio). Also, it should avoid laborious design−build−test cycles that slow progress.[7,12−19]

Previously, our lab developed the **Y**east **E**ndoplasmic **R**eticulum (ER) **S**equestration **S**creening (YESS) system,[20−24] which is a robust yeast surface display (YSD) approach that combines for the first time a protease engineering and a comprehensive substrate specificity profiling platform (Figure 1A). In YESS, an aga2-substrate cassette fusion designed for surface display contains both desired substrates and counter-selection substrates flanked by strategically placed epitope tags. A protease is coexpressed with the substrate cassette, both of which are targeted to the ER of *S. cerevisiae*. Upon coexpression in the ER, the protease interacts with the transiting substrate fusion, resulting in four possible outcomes: no interaction, cleavage of the selection substrate (SS), cleavage of the counter-selection substrate (CS), or cleavage of both substrates (Figure 1B). These phenotypes can be directly quantified by staining the cells with fluorophore-labeled antibodies, allowing visualization and selection by fluorescence-activated cell sorting (FACS). In contrast to modified YSD methods for enzyme engineering,[25−30] the YESS system[21,22] marks a significant step in converting YSD, conventionally used to engineer protein binding affinity, thermostability, and solubility into a platform that quantifies catalytic turnovers.

Here we report YESS 2.0, a markedly improved version of YESS that incorporates features to independently modulate protease and substrate transcription, substrate and protease

Received: August 27, 2020

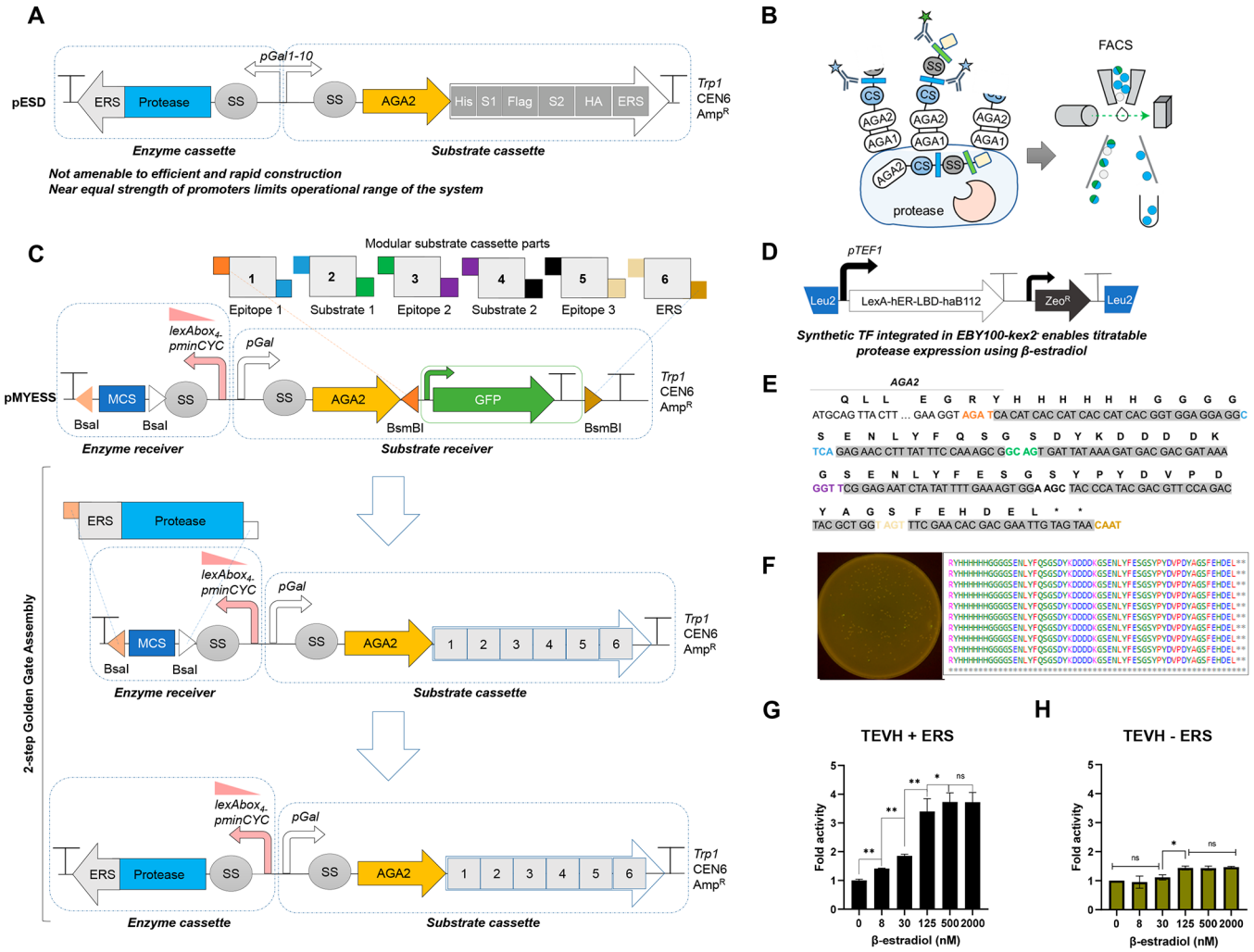


Figure 1. YESS 2.0: an improved YESS system for the engineering and profiling protease specificity. (A) Schematic of the YESS plasmid and process for protease engineering. (B) Phenotypes of surface-displayed substrate cassettes after interaction with protease in yeast. (C) Rapid assembly of substrate and protease parts in YESS 2.0. In the pYESS2 plasmid, the enzyme is under the control of the β-estradiol titratable promoter $(\text{LexA-box})_4$-pminCYC). On the substrate side, a GFP expression cassette expressed constitutively in *E. coli* and flanked by BsmBI restriction sites is placed directly downstream of aga2. A BsmBI-mediated golden gate reaction enables the seamless and rapid assembly of a polypeptide or protein substrate display cassette. The enzyme and substrate cassettes can be added independently or sequentially in the pYESS2 scaffold using BsaI or BsmBI assemblies, respectively. Arrows indicate the directionality of Type IIs restriction sites. (D) Construction of the EBY100-Tune strain. (E) Breakdown of aga2 substrate cassette to enable assembly of substrate parts using Golden Gate. The overhangs are made unique (two hamming distances between any two overhangs) by taking advantage of the redundancy in DNA sequences that translate into the GS linker between parts, effectively rendering parts only position-dependent. (F) Assembly of a six-part substrate cassette into pYESS2. The order of parts was chosen to be HIS-ENLYFQS-FLAG-ENLYFES-HA-ERS. However, parts can be assembled in any order as long as the correct overhangs are chosen. The plate shown highlights the efficiency of this reaction as more than 90% of colonies show GFP drop out. Moreover, all 10 white colonies that were selected were found to have the correct sequence. (G) β-Estradiol dose-dependent induction of TEV-H activity in the presence of an ERS in the enzyme cassette. To calculate fold activity, the ratio of FLAG/HA signal from each β-E induction concentration is normalized to the FLAG/HA ratio of the control sample (no β-E induction). Reactions were run in duplicates. Statistical two-tailed *t* test was performed ($p > 0.05$: ns; $p \leq 0.001$: ***; $p \leq 0.01$: **; $p \leq 0.05$: *). (H) β-Estradiol dose-dependent induction of TEV-H activity in the absence of an ERS in the enzyme cassette. Data was analyzed as in 1G).

spatial sequestration within a versatile, seamless, and rapid assembly method (Figure 1C). Importantly, protease activity can be tuned with significantly higher sensitivity and dynamic range in YESS 2.0 compared with the original YESS platform. Using YESS 2.0, we sought to further improve an already efficient Tomato etch virus (TEV) protease variant (TEV-EAV) evolved using YESS 1.0. By modulating both the ratio of TEV to substrate and the contact time between TEV and its substrate in the ER, several TEV variants with higher catalytic efficiency were obtained. Among them, the variant eTEV displayed a 2.25 fold higher catalytic efficiency compared to TEV-EAV, solely attributable to an increase in turnover rate (k_{cat}). eTEV also shows a noticeable improvement in efficiency under real-use conditions, making it a potentially valuable new tool for protein purification and many other applications.

MATERIALS AND METHODS

Construction of pYESS2. The plasmid pYESS2 was based on the pESD backbone previously described[21] with the following modifications. First, the BsaI and BsmBI sites found on the pESD backbone were removed by Gibson Isothermal assembly. Second, an intermediate entry vector was

https://dx.doi.org/10.1021/acssynbio.0c00452
ACS Synth. Biol. XXXX, XXX, XXX−XXX

constructed by replacing the substrate and enzyme cassettes of pESD with a gene block containing multiple type II restriction sites. These restriction sites were used to add sequentially a Gal1 promoter, a β-estradiol-inducible promoter (lexA(box)$_4$-pCYC1), *aga2*, a constitutively expressed GFP cassette flanked by BsmBI sites on the substrate side, and a multiple cloning site flanked by BsaI sites on the enzyme side. The plasmid FRP793_insul-(lexA-box)$_4$-PminCYC1-Citrine-TCYC1 used to amplify the β-estradiol-inducible promoter, was a gift from Joerg Stelling (Addgene plasmid no. 58434; http://n2t.net/addgene:58434; RRID:Addgene_58434).[31]

EBY-Tune Strain Construction. The MoClo-YTK plasmid kit was a gift from John Dueber (Addgene kit no. 1000000061). A receiver plasmid was constructed using MoClo-YTK to integrate at the Leu2 marker of EBY100 under zeocin selection (LeuIV-ZeoR). The lexA-ER-haB112 transcription factor cassette was amplified from the plasmid FRP880_PACT1(−1−520)-LexA-ER-haB112-TCYC1, a gift from Joerg Stelling (Addgene plasmid no. 58437; http://n2t.net/addgene:58437; RRID: Addgene_58437). The transcription factor was cloned in LeuIV-ZeoR via BsaI assembly. The obtained plasmid LeuIV-ZeoR-LexA_TF was linearized with NotI and transformed into EBY100-kex2- by electroporation as described elsewhere.[32]

Cassette Assemblies. For module construction, 150 pmol of each the forward and reverse primers were mixed with 1 μL of T4 ligase buffer and 0.2 μL of T4 polynucleotide kinase (T4 PNK) in a 10 μL reaction. The oligos were phosphorylated and annealed using the following PCR program: 37 °C for 30 min, 98 °C for 5 min, 0.1 °C/s ramp down to 4 °C. Annealed modules were diluted 12.5-fold in water, kept on ice to be used immediately, or stored at −20 °C.

To assemble modules in pYESS2, each module (0.5 μL) was mixed with 25 ng of pYESS2 in a 5 μL reaction containing 0.5 μL of T4 ligase buffer, 0.25 μL of BsmBI, 0.25 μL of T7 ligase. Golden gate assembly program was as follows: (37 °C for 1 min, 16 °C for 2 min) × 29 cycles, 37 °C for 60 min, 65 °C for 20 min, 4 °C. Golden gate reactions were desalted by drop-dialysis (HVLP04700 Durapore, Millipore) and transformed in *E. coli* DH10B by electroporation, plated on LB + ampicillin (100 μg/mL) plates, and incubated at 37 °C overnight. White clones were picked, miniprepped, and sequenced.

For cloning of the enzyme into pYESS2, the choice of ERS strength was encoded on the reverse primer used to amplify the enzyme gene. The enzyme was added to pYESS2 or pYESS2 + substrate via a BsaI Golden Gate assembly. Routine transformations of pYESS2 plasmids in EBY100-Tune were performed in EZ-competent cells (Zymo Research, CA) and plated on YNB-CAA-glucose plates and incubated at 30 °C.

Screening of TEV-H at Various β-E Concentrations. Two plasmids encoding TEV-H + ERS (FEHDEL) and a TEV-H substrate cassette (Aga2-Flag-ENLYFHS-HA), with and without an FEHDEL ERS, respectively, were transformed in EBY-Tune and plated on YNB-CAA-Glucose plates. Individual clones were grown in YNB-CAA-glucose and induced at 30 °C in YNB-CAA-Galactose containing increasing concentrations of β-estradiol. After overnight induction, 10^7 cells were washed in cold PBS + 0.5% BSA, followed by labeling in ice-cold PBS + 0.5% BSA with anti-FLAG-PE (Biolegend, 0.10 μL/10^7 cells) and anti-HA-FITC (Genscript, 0.10 μL/10^7 cells) antibodies at a concentration of 10^5 cells/μL and incubated on ice, in the dark, for 1 h. Lastly, cells were washed in PBS + 0.5% BSA and scanned by flow cytometry.

TEV Protease Library Construction and Sorting. A TEV protease error-prone library (4.8×10^5) with no ERS with a 1.5% error rate was prepared using Taq DNA polymerase. A combinatorial NNS site saturation library (1×10^5) was prepared by overlap extension PCR. pYESS2 vectors were linearized by *Pst*I and SphI, and yeast transformation was performed as described previously.[32] Libraries were induced in YNB-CAA-Galactose supplemented with various concentrations of β-estradiol at a starting OD of 0.5. Typically, the number of induced cells corresponded to a ∼10-fold coverage of the initial library or the number of cells collected in the previous sorting round. For antibody labeling, cells were washed in cold PBS + 0.5% BSA, followed by labeling in ice-cold PBS + 0.5% BSA with anti-FLAG-PE (Biolegend, 0.25 μL/10^7 cells) and anti-HA-FITC (Genscript, 0.5 μL/10^7 cells) antibodies at a concentration of 10^5 cells/μL and incubated on ice, in the dark, for 1 h. Lastly, cells were washed in PBS + 0.5% BSA and scanned or sorted by FACS. After sorting, the cells were plated on selective medium plates, and individual colonies were reanalyzed and confirmed by flow cytometry. The DNA was extracted from the confirmed yeast single colonies using a Zymoprep kit (Zymo Research, CA) and transformed into *E. coli* for Sanger sequencing.

TEV Protease and MBP-ENLYFQS-GST Fusion Protein Expression and Purification. TEV protease variants were cloned downstream of a maltose-binding protein and transformed into *E. coli* BL21-RIL as described previously.[33] TEV proteases were purified using Ni-NTA chromatography, followed by dialysis overnight in storage buffer (50 mM Tris-HCl, pH 8.0, 0.1 mM NaCl, 1 mM DTT, 10% glycerol). For kinetic analysis, freshly purified TEV proteases were used. The MBP (maltose binding protein) and the GST (glutathione-S-transferase) protein were fused with a peptide linker containing the TEV substrate ENLYFQS, designated as MBP-ENLYFQS-GST. The MBP-ENLYFQS-GST fusion protein was expressed and purified on an amylose resin as described previously.[21]

Protease Characterization. TEV kinetic assays were carried out as described previously with minor modification.[33] Kinetic assays were carried out in 50 mM Tris-HCl, pH 8.0 containing 1 mM EDTA, and 2 mM freshly prepared DTT. In a 50 μL reaction, a total of 7.5 μM to 1 mM of substrate peptide (TENLYFQSGTRRW) was incubated with 0.1−0.5 μM of purified enzyme at 30 °C for 10−30 min. Subsequently, the reactions were quenched with 10 μL of 5% trifluoroacetic acid (TFA), followed by HPLC analysis on a C-18 column using an acetonitrile gradient from 2% to 60% over 10 min at a flow rate of 0.62 mL/min. The product amount was calculated upon the integration area at 280 nm and converted to concentration using a product response curve (SGTRRW). The MBP-ENLYFQS-GST fusion protein was used to monitor cleavage of proteins by TEV variants by SDS-PAGE analysis. The extent of cleavage of fusion proteins was calculated using the Image Lab Software (Bio Rad).

Structural Modeling of TEV Variants. The structure modeling was performed according to a previously published procedure.[34] The wild-type TEV protease structure (PDB: 1LVB) and peptide substrate, ENLYFQS, were derived from the protein data bank.[35] The structure of eTEV was derived from simulation using the online ITASSER program (https://zhanglab.ccmb.med.umich.edu/ITASSER/).[36] The modeling

https://dx.doi.org/10.1021/acssynbio.0c00452
ACS Synth. Biol. XXXX, XXX, XXX−XXX

studies were performed using the ZDOCK program,[37] followed by subsequent refinement with the RDOCK protocol.

RESULTS AND DISCUSSION

Generation of the YESS 2.0 System. Two key elements of YESS 2.0 are the newly designed and highly modular vector pYESS2 (Figure 1C) and an EBY100$^{kex2-20}$ strain with a chromosomally integrated synthetic transcription factor that enables β-estradiol (β-E) induction[20] (EBY100-Tune) (Figure 1D). The plasmid, pYESS2, contains four principal features. (1) The transcription of the substrate cassette and enzyme are decoupled as they are under the control of β-E and galactose-inducible promoters, respectively. (2) Induction by β-E enables tight control and titration of enzyme levels, addressing possible protease toxicity and enabling sorting in the absence of enzyme expression (i.e., allowing a negative sort by removing β-E and therefore enzyme expression). (3) Introduction of BsaI and BsmBI sites allow for rapid enzyme and substrate library incorporation via Golden Gate cloning.[38] (4) BsmBI sites on the substrate cassette side flank a constitutively expressed GFP that drops out upon correct substrate ligation, enabling a rapid initial green-white screen (Figure 1E).[39] The strain EBY100-Tune was generated by integrating the LexA-ER-haB112 synthetic chimeric transcription factor at the *leu2* marker of EBY100^{kex2-}. The EBY100^{kex2-} strain is a previously reported display strain that removes a major yeast endogenous proteolytic background, rendering this platform amenable to profiling proteases with trypsin-like activity.[20]

Rapid Assembly of Protease Activity Reporters. In YESS 1.0, the substrate cassettes were either assembled using traditional restriction digestion-ligation of a gBlock (aga2-substrate cassette fusion) or by overlap extension PCR. This is because the substrate cassettes downstream of *aga2* (substrates, epitope tags, and ERS) are normally as short as possible (<50 bp) to avoid having longer sequences that would increase the chances of containing an inadvertent substrate sequence for a given protease. These obstacles hinder the cloning of parts by traditional cloning or Gibson isothermal assembly. However, with any given enzyme, it is often necessary to (1) test several substrates, (2) permute epitope tags to eliminate adventitious protease substrate sequences and simultaneously optimize surface display, and (3) test various ERS sequences. A multivariable optimization using YESS 1.0, therefore, required many sequential cycles of laborious plasmid construction and analysis.

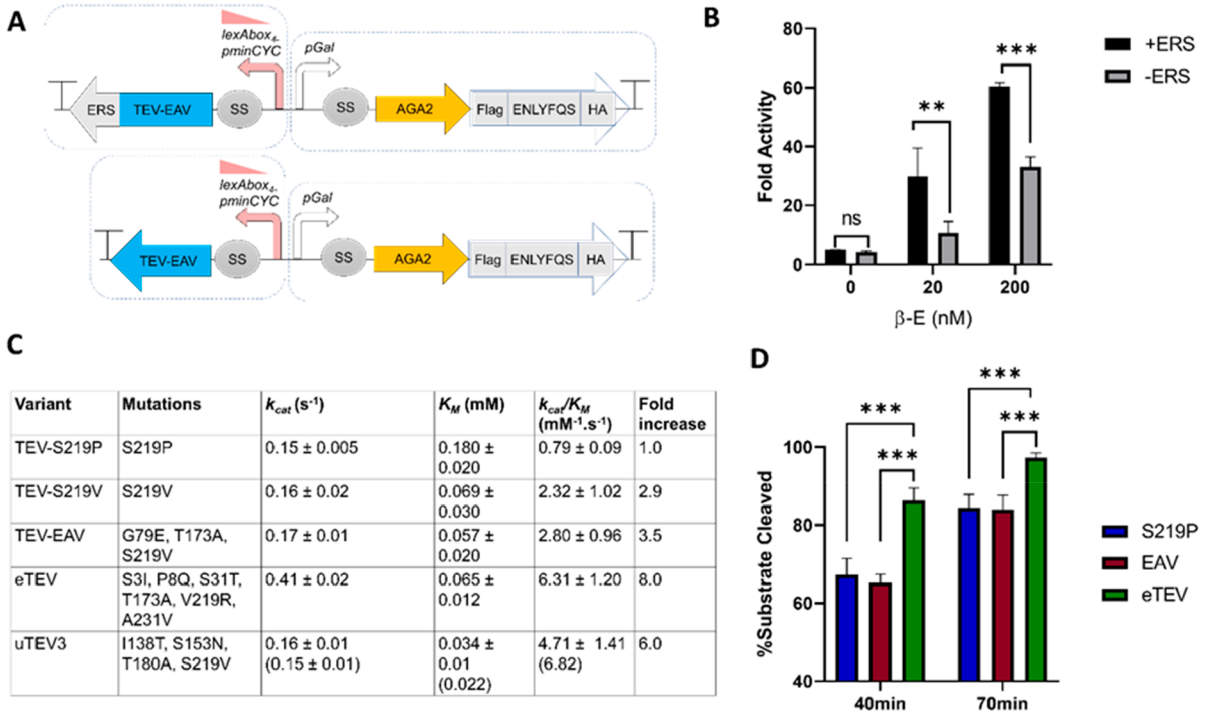
The second-generation vector pYESS2 was designed to enable the rapid assembly of substrate cassette parts in any desired order, number, and identity as needed, allowing for completion of cloning within 2−3 days rather than 2 weeks (Supplemental Figure 1). In particular, the substrate cassette sequence downstream of *aga2* was divided into short modules connected by a Gly-Ser linker (Figure 1C). Taking advantage of the degeneracy of the codons that translate into Gly-Ser, unique four base pair (bp) overhangs were designed to ensure that modules ligate in the correct order (Figure 1E). Therefore, epitope tag and ERS modules can be reused across multiple assemblies. Furthermore, modules are easily assembled by annealing and phosphorylating two DNA oligos in a way that generates the overhangs necessary for ligation. It needs to be pointed out that modules are only position-dependent because their overhangs are within the GS linker and do not interfere each module. The first and last module contain 4-bp overhangs that match the 5′- and 3′-backbone of pYESS2 substrate receiver module. Another strategy used in the pYESS2 plasmid is that a GFP gene was originally inserted in the substrate region, which could be replaced by the substrate cassette, thus enabling a white/green selection (Figure 1C). As an example, the BsmBI assembly efficiency of 6 TEV protease substrate modules in pYESS2 resulted in GFP dropout in over 90% of the colonies (Figure 1F and Supplemental Table 1). Moreover, all 10 white clones picked and sequenced were the expected construct. BsmBI assembly of 4, 5, and 6 modules are routinely performed with over 90% accuracy. Through a BsaI assembly, the protease can be added to pYESS2 before or after the substrate cassette construction (Figure 1C). However, the choice of ERS is encoded in the protease reverse primer sequence rather than as a separate module.

Modulating Protease Activity in YESS 2.0. In YESS 1.0, the inability to finely control or turn off protease expression sometimes resulted in unwanted consequences. Since protease activity on a substrate within YESS is observed through a loss of fluorescence, we have observed that mutations or stop codons accumulated within the C-terminal epitope tag or the substrate itself can be enriched by FACS and overtake the sorted population. This phenomenon makes it challenging to assess whether a sorted library contains improved variants or false positives, as a significant fraction of the sorted library may contain epitope tag mutations and stop codons. False positives can only be removed by recloning a sorted protease library in a new plasmid. In contrast, the YESS 2.0 system overcomes these obstacles by shifting protease expression to a titratable, orthogonal β-E-inducible promoter and enabling one to switch off protease activity independently of substrate expression for use in negative screening rounds.

These YESS 2.0 features were first showcased by building and transforming two TEV-H activity reporters into EBY100-Tune. We had previously identified TEV-H as an orthogonal TEV variant, preferring a P1 His over the wild-type preferred P1 Gln. The two reporters constructed differed by the presence or absence of the FEHDEL ERS from the protease cassette, with no ERS present on the substrate end. When the FEHDEL ERS is present, TEV-H activity increases in a dose-dependent manner over nearly 3 orders of magnitude of β-E concentration (Figure 1G). In contrast, with the added stringency of an absence of any ERS, TEV-H shows observed activity only if the β-E concentration is above 125 nM (Figure 1H). This key result verified that, indeed, YESS 2.0 can control enzyme activity at both transcriptional and post-translational levels. Importantly, no protease activity is observed in the absence of β-E.

Directed Evolution of a Highly Active TEV Protease. Besides the need for TEV proteases with orthogonal substrate specificities,[9,12,40,41] it has long been a goal to improve the turnover rate (k_{cat}) of TEV protease on its native substrate ENLYFQS (0.15 s^{-1}) to leverage its exquisite substrate specificity fully. Importantly, synthetic systems that rely on TEV-mediated actuation[4,5,9,10,42] may benefit from a catalytically superior TEV protease, enabling faster proteolysis-based signaling and logic operations. In our previous effort, YESS 1.0 was used to evolve a triple mutant TEV variant (G79E, T173A, S219V) (referred to here as TEV-EAV) with 3.5-fold higher catalytic efficiency compared to the parent TEV-S219P when reacting with the native substrate ENLYFQS[21] and this was due to a decrease in K_M. To determine whether further enhanced TEV variants with increased k_{cat} could be isolated by

https://dx.doi.org/10.1021/acssynbio.0c00452
ACS Synth. Biol. XXXX, XXX, XXX−XXX



Variant	Mutations	k_{cat} (s^{-1})	K_M (mM)	k_{cat}/K_M ($mM^{-1}.s^{-1}$)	Fold increase
TEV-S219P	S219P	0.15 ± 0.005	0.180 ± 0.020	0.79 ± 0.09	1.0
TEV-S219V	S219V	0.16 ± 0.02	0.069 ± 0.030	2.32 ± 1.02	2.9
TEV-EAV	G79E, T173A, S219V	0.17 ± 0.01	0.057 ± 0.020	2.80 ± 0.96	3.5
eTEV	S3I, P8Q, S31T, T173A, V219R, A231V	0.41 ± 0.02	0.065 ± 0.012	6.31 ± 1.20	8.0
uTEV3	I138T, S153N, T180A, S219V	0.16 ± 0.01 (0.15 ± 0.01)	0.034 ± 0.01 (0.022)	4.71 ± 1.41 (6.82)	6.0

Figure 2. YESS 2.0 evolution of a highly active TEV variant. (A) Two TEV-EAV reporters were built to assess TEV-EAV activity in YESS 2.0. Experiments were in duplicates. (B) Activity of TEV-EAV can be tightly controlled using β-E in galactose medium. The legends "+ERS" and "−ERS" refer to the presence or absence of an ER retention sequence on the protease cassette. Both constructs are missing an ERS on the substrate cassette. (C) Kinetic characterization of evolved TEV variants. Kinetic characterization of TEV variants on the peptide TENLYFQSGGTRW. TEV variant eTEV was obtained by combining mutations from TEV-E2 and TEV-S7 (supplemental Table 2). Kinetic parameters for uTEV3, a variant published elsewhere,[7] were determined on a peptide substrate by HPLC and compared to published results (in parentheses). (D) Digestion of an MBP-TEVcs-GST by TEV variants. Ten μg of MBP-TEVcs-GST was mixed with 0.1 μg of TEV variant in a 10 μL reaction and incubated at 30 °C. Reactions were run in duplicates. Statistical two-tailed t test was performed ($p > 0.05$: ns; $p \leq 0.001$: ***; $p \leq 0.01$: **; $p \leq 0.05$: *).

leveraging the operational range of YESS 2.0, TEV-EAV reporter strains were induced at three concentrations of β-E (0, 20, and 200 nM). Induction of a TEV-EAV reporter devoid of ERS with 20 and 200 nM β-E showed a respective 2.5-fold and 7.5-fold increase in activity compared to cells induced with galactose only (Figure 2A). This assay indicated that inducing a TEV-EAV library with 20 nM β-E or lower would allow variants with higher catalytic efficiency to be selected. Since YESS 2.0 operation may approach enzyme saturation kinetics at low β-E concentrations in 2% galactose, faster variants may be obtained that show an increased k_{cat}.

On the basis of this result, an error-prone (1.5% error rate) and a site-saturation mutagenesis library (at residues E79, A173, and V219) with no ERS on either substrate or protease ends were prepared on the TEV-EAV template and screened against the AGA2-FLAG-ENLYFQS-HA substrate cassette. After being transformed into EBY100-Tune,[32] 2×10^8 cells were induced in YNB-CAA-galactose supplemented with 20 nM β-E and subsequently labeled with fluorescently labeled anti-FLAG-PE and anti-HA-FITC antibodies. Cells that displayed high PE and low FITC fluorescence (0.2% of the population) were collected. Post-R1 cells were induced with 20 nM and 8 nM β-E in YNB-CAA-galactose, respectively. After analyzing the activity of the library, cells induced with 8 nM β-E that displayed high PE and low FITC fluorescence were collected (1% of the population). Finally, a negative sorting round (in the absence of β-E) was performed, and cells displaying both high PE and FITC signals were collected (Supplemental Figure 2). This sorting strategy appraised the functionality of the sorted library after each round and can inform the choice of sorting stringency or whether to perform a negative round of sorting. A negative sort allows one to collect cells that display both FLAG and HA epitope tags, effectively removing stop codons and mutations from the library. After four rounds of sorting, all screened clones showed improved cleavage activity in YESS 2.0 compared to parent TEV-EAV (Supplemental Figure 3). These phenotypes were retained *in vitro* as six selected clones showed higher activity on the peptide TENLYFQSGTRRW compared to TEV-EAV (Supplemental Table 2, Supplemental Figure 4).

Next, we combined the two best variants, E2 (S3I, P8Q, S31T, A231V) and S7 (E79G, V219R), from the error-prone and saturation mutagenesis libraries, respectively, to generate the hexamutant enhanced TEV "eTEV" (S3I, P8Q, S31T, E79G, T173A, V219R, A231V). This enhanced TEV shows a 2-fold improvement in catalytic efficiency (k_{cat}/K_M) compared with TEV-EAV, corresponding to an ∼8-fold overall increase in catalytic efficiency relative to the parent TEV-S219P. Notably, while K_M of eTEV remained almost unchanged at 65 μM with that of TEV-EAV, the turnover rate (k_{cat}) was improved 2.4-fold from 0.17 s^{-1} to 0.41 s^{-1} (Figure 2C, Supplemental Figure 5).

Finally, eTEV was tested in MBP-ENLYFQS-GST fusion protein cleavage assays using conditions recommended for commercial TEV protease. eTEV routinely outperformed TEV-S219P and TEV-EAV at various ratios of enzyme to substrate (Figure 2D and Supplemental Figure 6). For instance, eTEV was able to digest 85% of the MBP-ENLYFQS-GST fusion protein in 40 min at 30 °C at a 1:100 ratio, compared with ∼68% for TEV-EAV and TEV-

https://dx.doi.org/10.1021/acssynbio.0c00452
ACS Synth. Biol. XXXX, XXX, XXX−XXX

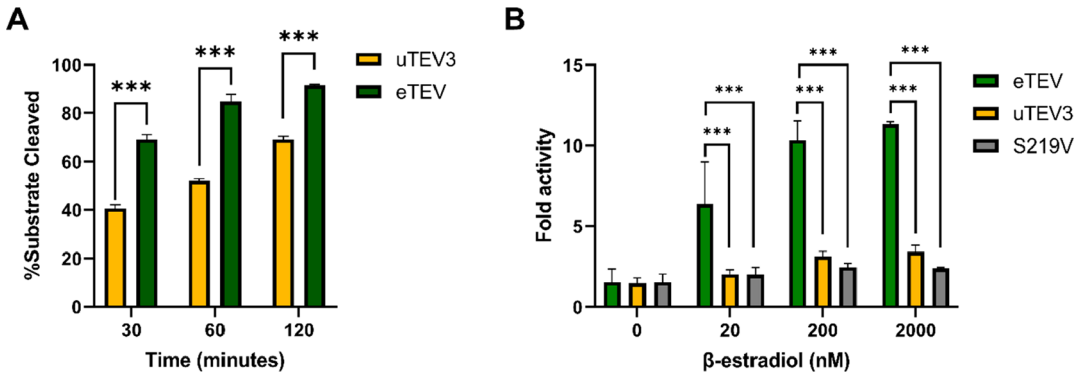


Figure 3. eTEV outperforms uTEV3 on protein substrates and in YESS 2.0. (A) MBP-TEVcs-GST cleavage assay comparing eTEV and uTEV3. Statistical two-tailed *t* test was performed ($p > 0.05$: ns; $p \leq 0.001$: ***; $p \leq 0.01$: **; $p \leq 0.05$: *). (B) Comparing the activity of uTEV3, eTEV, and TEV-S219V in YESS 2.0. In these constructs, the protease contains an ERS, while the substrate cassette does not. Statistical two-tailed *t* test was performed ($p > 0.05$: ns; $p \leq 0.001$: ***; $p \leq 0.01$: **; $p \leq 0.05$: *).

S219P (Figure 2D, Supplemental Figure 6). Therefore, the low amount of eTEV and reduced time needed to reach complete cleavage of a protein fusion (~70 min) makes eTEV a valuable reagent for protein purification.

Comparison of eTEV and uTEV3. During our eTEV engineering, an evolved TEV variant, named uTEV3, was reported. uTEV3 was generated using a yeast platform based on TEV-mediated transcription factor release via induction of protein−protein interactions.[7] Based on our characterization, eTEV and uTEV3 possess similar catalytic efficiencies under the same conditions (Figure 2C), except that the improved catalytic efficiency in uTEV3 appears to be mainly due to a significant K_M decrease. In the MBP-ENLYFQS-GST fusion protein cleavage assay, eTEV outperformed uTEV3 at every point along the assay time course (Figure 3A). After 30 min, at a 1:200 protease to substrate ratio, eTEV cleaved 70% of the MBP-TEVcs-GST fusion, compared to only 30% cleavage with uTEV3. Furthermore, at a 1:200 ratio, eTEV reached ~90% digestion of the fusion protein after 120 min (Supplemental Figure 7).

As evidence that the TEV engineering strategy in YESS 2.0 was operating at saturating substrate levels, uTEV3 was cloned in a pYESS2 construct with no ERS similar to TEV-EAV and tested by FACS. Under these conditions, uTEV3 showed comparable activity to TEV-S219P, highlighting that the YESS 2.0 approach used to make eTEV did not select for TEV variants with stronger substrate−binding interactions.

Structural Modeling of TEV Variants. Although eTEV contains 6 mutations, none of them exist in the S1 binding pocket or catalytic core residues. To help understand the better catalytic properties of eTEV, structure modeling of TEV-S219P and eTEV was performed with the peptide substrate, ENLYFQS. Our results revealed a clear substrate channel and compared with TEV-S219P, the S1 binding pocket in eTEV was rearranged and further buried inside the protease (Figure 4A,B). Moreover, detailed analysis of the S1 binding pocket indicated that although the main residues interacting with substrate P1 residue, Gln, in the S1 binding pocket remained the same, the residue interaction in eTEV is tighter, which might be due to the more buried nature of its S1 binding pocket (Figure 4A). Considering that the eTEV has an approximately 2.8-fold lower K_M and 2.7-fold higher k_{cat}, it is reasonable to assume that the more buried S1 binding pocket in eTEV leads to a stronger substrate interaction, which apparently simultaneously enhances substrate proteolysis.

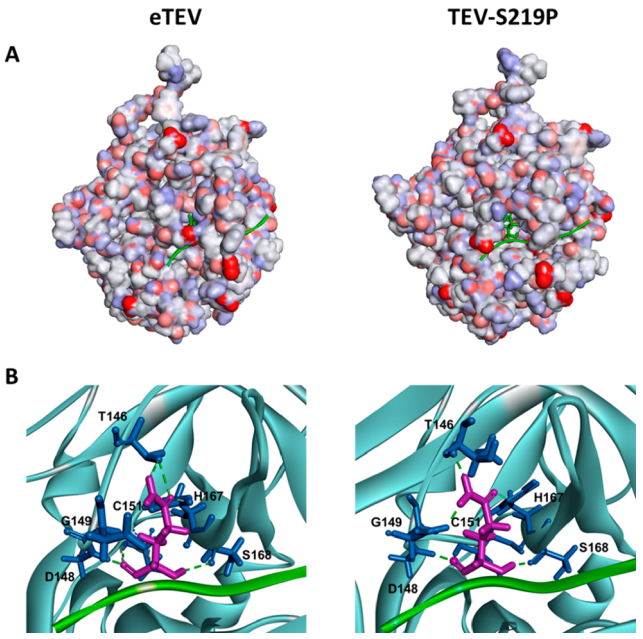


Figure 4. Structural modeling of TEV variants. (A) Compared with TEV-S219P, the S1 binding pocket in eTEV is further buried inside the protease. (B) Overview of the S1 binding pocket of eTEV and TEV-S219P.

Another interesting finding concerns the role residue 219 (Supplemental Table 2). Residue 219 is part of the flexible C-terminal loop positioned below the catalytic core and plays a critical role in regulating TEV autolysis and substrate binding.[43,44] Variant S6 reverted to the wild-type sequence and showed autolysis after purification as expected (Supplemental Figure 5B). Importantly, the respective mutations of S219 to Arg and Lys in variants S7 and S8 are shown for the first time to increase the catalytic efficiency of TEV while preventing autolysis (Supplemental Figure 5A). In particular, the V219R mutation in variant S7 suggests that it may be solely responsible for the observed improvements in catalytic efficiency. Our report adds to the known effects of mutating residue 219. On the one hand, S219P and S219E prevent TEV autolysis but occur at the detriment of k_{cat} and K_M, respectively.[43] On the other hand, mutating S219 to Val, Asn, Lys, or Arg also prevents autolysis, yet, in addition, they result in increased catalytic activity[43,44] (Figure 2B). Whether

https://dx.doi.org/10.1021/acssynbio.0c00452
ACS Synth. Biol. XXXX, XXX, XXX−XXX

residue 219 has greater impacts on K_M or k_{cat} may not be straightforward and may depend on the context of other mutations present in TEV.

CONCLUSIONS

Extensive rescreening of individual clones to confirm activity is an underreported bottleneck of directed evolution experiments. It is therefore noteworthy that 14 out of 16 TEV variants isolated from our screen showed higher activity when analyzed by FACS as an individual clone in YESS 2.0 as well as when the same clone was analyzed using *in vitro* cleavage assays with a peptide substrate (Supplemental Figure 4), thereby circumventing the rescreening bottleneck. Furthermore, by targeting the protease to the ER, the YESS 2.0 system promotes protein folding and could potentially synergize with ER engineering approaches that enhance protein folding and secretion.[45−47] Such tools might be necessary to facilitate rapid activity tests for uncharacterized and difficult-to-express proteases. Nevertheless, in soon to be reported studies, several different mammalian proteases have now been shown to be active by ER sequestration, including serine[48] and metalloproteases.[23]

Overall, the YESS 2.0 system brings us closer to unlocking the full potential of a YSD approach for understanding and redesigning the specificity of protein-modifying enzymes such as proteases. First, the speed and versatility with which protease activity reporters are constructed and verified in YESS 2.0 compared with YESS 1.0 were much-needed improvements that significantly improve experimental workflow (Supplemental Figure 1). Second, the transcriptional and translational control nodes introduced in YESS 2.0 provide for substantially higher sensitivity and dynamic range compared to YESS 1.0. These features allowed the system to operate at, or near, saturation kinetics, so that TEV variants with higher k_{cat} were selected. This TEV variant exhibits 2.25-fold higher catalytic efficiency compared to the parent enzyme TEV-EAV and this activity increase was almost entirely attributable to an increase in catalytic turnover (k_{cat}). To the best of our knowledge, while TEV protease has been extensively engineered, eTEV is the first TEV variant with increased k_{cat}. While we have shown that eTEV is faster at tag cleavage experiments under substrate saturating conditions, it remains to be tested whether eTEV properties could improve other applications, including synthetic circuits based on TEV-mediated actuation.

The modifications introduced in YESS 2.0 will also benefit substrate profiling[20,23,49] and drug screening[24,50] applications. The decoupling of enzyme and substrate transcriptions will ensure that sorted libraries contain very few false positives, resulting in populations highly enriched for true positives. Furthermore, the ability to finely modulate stringency via enzyme to substrate ratios allows one to stratify substrates directly based on their activities, generating high-quality data for next-generation sequencing and data-driven applications.[49,51]

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acssynbio.0c00452.

Sequence of parts used to assemble the 6-piece substrate cassette (Aga2-His-ENLYFQS-FLAG-ENLYFES-HA-ERS), description of substrate cloning strategy used in YESS 1.0, sorting the error-prone library of TEV-EAV, scan of individual clones induced at 20 nM β-E, detailed list of mutations found in selected variants, initial characterization of TEV variants *in vitro* on synthetic peptide TENLYFQSTRGGW, SDS-Page gels of purified TEV variants, biochemical characterization of TEV variants on synthetic peptide TENLYFQSTRGGW, representative protein gels of MBP-ENLYFQS-GST digestion by TEV variants, sequence of pYESS 2.0 plasmid (PDF)

AUTHOR INFORMATION

Corresponding Authors

Carl A. Denard − *Department of Chemistry, University of Texas at Austin, Austin, Texas 78712, United States;* orcid.org/0000-0002-2804-9426; Email: cdenard@ufl.edu

Brent L. Iverson − *Department of Chemistry, University of Texas at Austin, Austin, Texas 78712, United States;* orcid.org/0000-0001-7974-3605; Email: iversonb@austin.utexas.edu

Authors

Chelsea Paresi − *Department of Chemistry, University of Texas at Austin, Austin, Texas 78712, United States*

Rasha Yaghi − *Department of Chemistry, University of Texas at Austin, Austin, Texas 78712, United States*

Natalie McGinnis − *Institute for Cellular and Molecular Biology, University of Texas at Austin, Austin, Texas 78712, United States*

Zachary Bennett − *Department of Biomedical Engineering, University of Texas at Austin, Austin, Texas 78712, United States*

Li Yi − *Department of Chemistry, University of Texas at Austin, Austin, Texas 78712, United States*

George Georgiou − *Department of Chemical Engineering, University of Texas at Austin, Austin, Texas 78712, United States*

Complete contact information is available at: https://pubs.acs.org/10.1021/acssynbio.0c00452

Funding

This work was supported by the Clayton Foundation (B.L.I.) for Research.

Notes

The authors declare the following competing financial interest(s): The authors have filed a provisional patent application on components of this manuscript and the mutations in eTEV.

REFERENCES

(1) Craik, C. S., Page, M. J., and Madison, E. L. (2011) Proteases as therapeutics. *Biochem. J. 435*, 1−16.

(2) Arnau, J., Lauritzen, C., Petersen, G. E., and Pedersen, J. (2006) Current strategies for the use of affinity tags and tag removal for the purification of recombinant proteins. *Protein Expression Purif. 48*, 1−13.

(3) Giansanti, P., Tsiatsiani, L., Low, T. Y., and Heck, A. J. (2016) Six alternative proteases for mass spectrometry-based proteomics beyond trypsin. *Nat. Protoc. 11*, 993−1006.

(4) Kim, C. K., Cho, K. F., Kim, M. W., and Ting, A. Y. (2019) Luciferase-LOV BRET enables versatile and specific transcriptional readout of cellular protein-protein interactions. *eLife 8*, No. e43826.

(5) Wang, W., Wildes, C. P., Pattarabanjird, T., Sanchez, M. I., Glober, G. F., Matthews, G. A., Tye, K. M., and Ting, A. Y. (2017) A light- and calcium-gated transcription factor for imaging and manipulating activated neurons. *Nat. Biotechnol. 35*, 864−871.
(6) Lin, M. Z., Glenn, J. S., and Tsien, R. Y. (2008) A drug-controllable tag for visualizing newly synthesized proteins in cells and whole animals. *Proc. Natl. Acad. Sci. U. S. A. 105*, 7744−7749.
(7) Sanchez, M. I., and Ting, A. Y. (2020) Directed evolution improves the catalytic efficiency of TEV protease. *Nat. Methods 17*, 167−174.
(8) Copeland, M. F., Politz, M. C., Johnson, C. B., Markley, A. L., and Pfleger, B. F. (2016) A transcription activator-like effector (TALE) induction system mediated by proteolysis. *Nat. Chem. Biol. 12*, 254−260.
(9) Gao, X. J., Chong, L. S., Kim, M. S., and Elowitz, M. B. (2018) Programmable protein circuits in living cells. *Science 361*, 1252−1258.
(10) Fink, T., Lonzaric, J., Praznik, A., Plaper, T., Merljak, E., Leben, K., Jerala, N., Lebar, T., Strmsek, Z., Lapenta, F., Bencina, M., and Jerala, R. (2019) Design of fast proteolysis-based signaling and logic circuits in mammalian cells. *Nat. Chem. Biol. 15*, 115−122.
(11) Stein, V., and Alexandrov, K. (2014) Protease-based synthetic sensing and signal amplification. *Proc. Natl. Acad. Sci. U. S. A. 111*, 15934−15939.
(12) Packer, M. S., Rees, H. A., and Liu, D. R. (2017) Phage-assisted continuous evolution of proteases with altered substrate specificity. *Nat. Commun. 8*, 956.
(13) Pogson, M., Georgiou, G., and Iverson, B. L. (2009) Engineering next generation proteases. *Curr. Opin. Biotechnol. 20*, 390−397.
(14) Sellamuthu, S., Shin, B. H., Han, H. E., Park, S. M., Oh, H. J., Rho, S. H., Lee, Y. J., and Park, W. J. (2011) An engineered viral protease exhibiting substrate specificity for a polyglutamine stretch prevents polyglutamine-induced neuronal cell death. *PLoS One 6*, No. e22554.
(15) Sandersjoo, L., Kostallas, G., Lofblom, J., and Samuelson, P. (2014) A protease substrate profiling method that links site-specific proteolysis with antibiotic resistance. *Biotechnol. J. 9*, 155−162.
(16) Li, Q., Yi, L., Marek, P., and Iverson, B. L. (2013) Commercial proteases: present and future. *FEBS Lett. 587*, 1155−1163.
(17) Varadarajan, N., Gam, J., Olsen, M. J., Georgiou, G., and Iverson, B. L. (2005) Engineering of protease variants exhibiting high catalytic activity and exquisite substrate selectivity. *Proc. Natl. Acad. Sci. U. S. A. 102*, 6855−6860.
(18) Varadarajan, N., Georgiou, G., and Iverson, B. L. (2008) An engineered protease that cleaves specifically after sulfated tyrosine. *Angew. Chem., Int. Ed. 47*, 7861−7863.
(19) Varadarajan, N., Rodriguez, S., Hwang, B. Y., Georgiou, G., and Iverson, B. L. (2008) Highly active and selective endopeptidases with programmed substrate specificities. *Nat. Chem. Biol. 4*, 290−294.
(20) Li, Q., Yi, L., Hoi, K. H., Marek, P., Georgiou, G., and Iverson, B. L. (2017) Profiling protease specificity: combining yeast ER sequestration screening (YESS) with next-generation Sequencing. *ACS Chem. Biol. 12*, 510−518.
(21) Yi, L., Gebhard, M. C., Li, Q., Taft, J. M., Georgiou, G., and Iverson, B. L. (2013) Engineering of TEV protease variants by yeast ER sequestration screening (YESS) of combinatorial libraries. *Proc. Natl. Acad. Sci. U. S. A. 110*, 7229−7234.
(22) Yi, L., Taft, J. M., Li, Q., Gebhard, M. C., Georgiou, G., and Iverson, B. L. (2015) Yeast endoplasmic reticulum sequestration screening for the engineering of proteases from libraries expressed in yeast. *Methods Mol. Biol. 1319*, 81−93.
(23) Mei, M., Zhai, C., Li, X., Zhou, Y., Peng, W., Ma, L., Wang, Q., Iverson, B. L., Zhang, G., and Yi, L. (2017) Characterization of aromatic residue-controlled protein retention in the endoplasmic reticulum of Saccharomyces cerevisiae. *J. Biol. Chem. 292*, 20707−20719.
(24) Taft, J. M., Georgeon, S., Allen, C., Reckel, S., DeSautelle, J., Hantschel, O., Georgiou, G., and Iverson, B. L. (2019) Rapid screen for tyrosine kinase inhibitor resistance mutations and substrate specificity. *ACS Chem. Biol. 14*, 1888−1895.
(25) Gai, S. A., and Wittrup, K. D. (2007) Yeast surface display for protein engineering and characterization. *Curr. Opin. Struct. Biol. 17*, 467−473.
(26) Cherf, G. M., and Cochran, J. R. (2015) Applications of yeast surface display for protein engineering. *Methods Mol. Biol. 1319*, 155−175.
(27) Konning, D., and Kolmar, H. (2018) Beyond antibody engineering: directed evolution of alternative binding scaffolds and enzymes using yeast surface display. *Microb. Cell Fact. 17*, 32.
(28) Lim, S., Chen, B., Kariolis, M. S., Dimov, I. K., Baer, T. M., and Cochran, J. R. (2017) Engineering high affinity protein-protein interactions using a high-throughput microcapillary array platform. *ACS Chem. Biol. 12*, 336−341.
(29) Lim, S., Glasgow, J. E., Filsinger Interrante, M., Storm, E. M., and Cochran, J. R. (2017) Dual display of proteins on the yeast cell surface simplifies quantification of binding interactions and enzymatic bioconjugation reactions. *Biotechnol. J. 12*, 1600696.
(30) Mei, M., Zhou, Y., Peng, W., Yu, C., Ma, L., Zhang, G., and Yi, L. (2017) Application of modified yeast surface display technologies for non-Antibody protein engineering. *Microbiol. Res. 196*, 118−128.
(31) Ottoz, D. S., Rudolf, F., and Stelling, J. (2014) Inducible, tightly regulated and growth condition-independent transcription factor in Saccharomyces cerevisiae. *Nucleic Acids Res. 42*, No. e130.
(32) Benatuil, L., Perez, J. M., Belk, J., and Hsieh, C. M. (2010) An improved yeast transformation method for the generation of very large human antibody libraries. *Protein Eng., Des. Sel. 23*, 155−159.
(33) Tropea, J. E., Cherry, S., and Waugh, D. S. (2009) Expression and purification of soluble His(6)-tagged TEV protease. *Methods Mol. Biol. 498*, 297−307.
(34) Fan, X., Li, X., Zhou, Y., Mei, M., Liu, P., Zhao, J., Peng, W., Jiang, Z. B., Yang, S., Iverson, B. L., Zhang, G., and Yi, L. (2020) Quantitative analysis of the substrate specificity of human rhinovirus 3C protease and exploration of its substrate recognition mechanisms. *ACS Chem. Biol. 15*, 63−73.
(35) Phan, J., Zdanov, A., Evdokimov, A. G., Tropea, J. E., Peters, H. K., 3rd, Kapust, R. B., Li, M., Wlodawer, A., and Waugh, D. S. (2002) Structural basis for the substrate specificity of tobacco etch virus protease. *J. Biol. Chem. 277*, 50564−50572.
(36) Yang, J., Yan, R., Roy, A., Xu, D., Poisson, J., and Zhang, Y. (2015) The I-TASSER Suite: protein structure and function prediction. *Nat. Methods 12*, 7−8.
(37) Chen, R., Li, L., and Weng, Z. (2003) ZDOCK: an initial-stage protein-docking algorithm. *Proteins: Struct., Funct., Genet. 52*, 80−87.
(38) Engler, C., and Marillonnet, S. (2014) Golden Gate cloning. *Methods Mol. Biol. 1116*, 119−131.
(39) Lee, M. E., DeLoache, W. C., Cervantes, B., and Dueber, J. E. (2015) A Highly characterized yeast toolkit for modular, multipart assembly. *ACS Synth. Biol. 4*, 975−986.
(40) Fernandez-Rodriguez, J., and Voigt, C. A. (2016) Post-translational control of genetic circuits using Potyvirus proteases. *Nucleic Acids Res. 44*, 6493−6502.
(41) Renicke, C., Spadaccini, R., and Taxis, C. (2013) A tobacco etch Virus protease with increased substrate tolerance at the P1′ position. *PLoS One 8*, No. e67915.
(42) Barnea, G., Strapps, W., Herrada, G., Berman, Y., Ong, J., Kloss, B., Axel, R., and Lee, K. J. (2008) The genetic design of signaling cascades to record receptor activation. *Proc. Natl. Acad. Sci. U. S. A. 105*, 64−69.
(43) Kapust, R. B., Tozser, J., Fox, J. D., Anderson, D. E., Cherry, S., Copeland, T. D., and Waugh, D. S. (2001) Tobacco etch virus protease: mechanism of autolysis and rational design of stable mutants with wild-type catalytic proficiency. *Protein Eng., Des. Sel. 14*, 993−1000.
(44) Nam, H., Hwang, B. J., Choi, D. Y., Shin, S., and Choi, M. (2020) Tobacco etch virus (TEV) protease with multiple mutations to improve solubility and reduce self-cleavage exhibits enhanced enzymatic activity. *FEBS Open Bio 10*, 619−626.

https://dx.doi.org/10.1021/acssynbio.0c00452
ACS Synth. Biol. XXXX, XXX, XXX−XXX

(45) Mei, M., Li, J., Wang, S., Lee, K. B., Iverson, B. L., Zhang, G., Ge, X., and Yi, L. (2019) Prompting Fab yeast surface display efficiency by ER retention and molecular chaperon co-expression. *Front. Bioeng. Biotechnol. 7*, 362.
(46) de Ruijter, J. C., Koskela, E. V., and Frey, A. D. (2016) Enhancing antibody folding and secretion by tailoring the Saccharomyces cerevisiae endoplasmic reticulum. *Microb. Cell Fact. 15*, 87.
(47) Huang, M., Wang, G., Qin, J., Petranovic, D., and Nielsen, J. (2018) Engineering the protein secretory pathway of Saccharomyces cerevisiae enables improved protein production. *Proc. Natl. Acad. Sci. U. S. A. 115*, E11025−E11032.
(48) Guerrero, J. L., O'Malley, M. A., and Daugherty, P. S. (2016) Intracellular FRET-based screen for redesigning the specificity of secreted proteases. *ACS Chem. Biol. 11*, 961−970.
(49) Pethe, M. A., Rubenstein, A. B., and Khare, S. D. (2019) Data-driven supervised learning of a viral protease specificity landscape from deep sequencing and molecular simulations. *Proc. Natl. Acad. Sci. U. S. A. 116*, 168.
(50) Dickinson, B. C., Packer, M. S., Badran, A. H., and Liu, D. R. (2014) A system for the continuous directed evolution of proteases rapidly reveals drug-resistance mutations. *Nat. Commun. 5*, 5352.
(51) Pethe, M. A., Rubenstein, A. B., and Khare, S. D. (2017) Large-scale structure-based prediction and identification of novel protease substrates using computational protein design. *J. Mol. Biol. 429*, 220−236.

https://dx.doi.org/10.1021/acssynbio.0c00452
ACS Synth. Biol. XXXX, XXX, XXX−XXX