# Supporting Information

## Yi et al. 10.1073/pnas.1215994110

### SI Methods

**Vector Construction.** The pESD (Yeast Epitope tagging vector for Suface Display) plasmid was constructed on the basis of the yeast surface display construct pCTCon2 (generously contributed by Dane K. Wittrup, Massachusetts Institute of Technology, Cambridge, MA) and the yeast epitope tagging vectors pESC-TRP vectors (generously contributed by Edward W. Marcotte, University of Texas at Austin) (Fig. 1). In the pESD plasmid (Fig. 1), the galactose (GAL) induced GAL1/10 bidirectional promoter (Agilent Technologies) in the pCTCon2 plasmid was replaced by the dual GAL1-GAL10 promoter that transcribes in both orientations. The promoter strength ratio of the GAL1:GAL10 in this bidirectional promoter is around 0.8:1.0 based on the product manual. The yeast adhesion receptor subunit Aga2 gene downstream of the GAL10 promoter was fused to a five-part cassette encoding (*i*) native substrate of Tobacco Etch Virus protease (TEV-P), ENLYFQS; (*ii*) the FLAG tag sequence, DYKDDDDK; (*iii*) the designed peptide substrate library, ENLYFXS (X can be any residue); (*iv*) the 6×His tag sequence; and (*v*) the endoplasmic reticulum (ER) retention signal sequence, FEHDEL. For the engineering of TEV-Fast, a similar construct (pESD-T) but without the 6×His tag sequence was generated (details in Fig. S9 and Table S1).

**Protease Library Construction.** The TEV-P gene used in this paper was originally extracted from Addgene plasmid 8830 (1). A saturation mutagenesis library of TEV-P, S1 pocket residues T146, D148, H167, and S170, was constructed (2) and then transformed into *Saccharomyces cerevisiae* strain EBY100 cells (*URA*+, *leu*−, *trp*−) (3). The TEV-P variant genes were subjected to random mutagenesis by error-prone PCR amplification (4) (see Table S3 for detailed information on primers).

The TEV-P library was inserted downstream of the GAL1 promoter in the pESD vector. The TEV-P contains a S219P mutation to increase its stability. For the construction of the S1 pocket library of TEV-P, PAGE-purified primers (primers 1–8, Table S3), which contained a randomized NNS codon (N = A, T, G, or C; S = G or C) in place of the wild-type codon at T146, D148, H167, and S170, were used to amplify the TEV-P gene by splicing overlap extension PCR (5). The PCR product was digested with KpnI and PstI, gel purified, and ligated into similarly digested E.coli cloning vector pTrc99A-MBP (pTrc99A-maltose binding protein). The ligation product was used to electroporate electrocompetent *Escherichia coli* MC1061 [$F^-$ Δ(*ara-leu*)7697 [*ara*D139]$_{B/r}$ Δ(*cod*B-*lac*I)3 *gal*K16 *gal*E15 $\lambda^-$ $e14^-$ *mcr*A0 *rel*A1 *rps*L150(strR) *spo*T1 *mcr*B1 *hsd*R2($r^-m^+$)] cells, and the library was plated on selective media. Plasmid DNA was isolated from the pooled clones, using a QIAprep Spin Miniprep kit (QIAGEN).

The library DNA was amplified by PCR, using primer 9 and primer 10 (Table S3), and a sequence encoding the ER retention peptide (FEHDEL) was fused in frame to the 3′ by overlap extension PCR. The library DNA was fused to a sequence encoding the Aga2 signal peptide QLLRCFSIFSVIASVLA. The DNA product was inserted downstream of the GAL1 promoter in pESD vector by homologous recombination, in the *S. cerevisiae* EBY100 strain (*URA*+, *leu*−, *trp*−) (3).

In addition, the random mutagenesis library of the TEV-P and its variants (TEV-PE10 and TEV-PH21) was generated by error-prone PCR amplification as described in ref. 4, using primer 11 and primer 12 (Table S3). The libraries were constructed in the E.coli cloning vector pMOPAC12 (6) and amplified using primer 9 and primer 13, which also removed the ER retention sequence (Table S3). The amplified library was then integrated into the pESD vector as described above.

**Substrate Library Construction.** The TEV-P substrate library was inserted downstream of the GAL10 promoter in the pESD vector. For the construction of the TEV-P substrate library, PAGE-purified primers (primer 14 and primer 15, Table S3), which contained a randomized NNS codon correspondent to the residue Q in the wild-type preferred substrate (ENLYFQS), were used to amplify the whole substrate fusion gene. The whole substrate fusion gene is composed of Aga2, selection and counterselection substrate sequences, multiple intervening epitope tag sequences, and a C-terminal ER retention sequence, The PCR products were then inserted downstream of the GAL10 promoter in pESD vector by homologous recombination, in the *S. cerevisiae* EBY100 strain (*URA*+, *leu*−, *trp*−) (3). The substrate library was labeled with anti-6×His-FITC antibody and presorted to remove any undesired mutations, including possible existing stop codons or shifted reading frames. After retransforming the recovered substrate constructs' DNA into *E. coli*, 96 randomly picked colonies were sequenced. Any substrate constructs not being recovered (ENLYFSS and ENLYFVS in our experiments) were remade separately and then added into the substrate construct mixture before later yeast transformation.

**Yeast Cell Screening.** Cells were grown to an $OD_{600}$ of 2.0–3.0 in 1 L YNB-CAA (yeast nitrogen base-casamino acid) + glucose medium, and then 2–5 × $10^8$ cells, around 10-fold larger than the library sizes, were induced with YNB-CAA + galactose medium at a final $OD_{600}$ of 0.5. Following media exchange, the cells were grown at 30 °C overnight, with shaking. A total of 2–5 × $10^8$ cells were washed and then labeled with fluorescently labeled antibodies: anti-FLAG-phycoerythrin (PE) antibody (ProZyme) and anti-6×His-FITC antibody (Genscript). During the antibody labeling steps, the cells were resuspended into 1× PBS solution containing 0.5% BSA with a final cell density of $10^5$ cells/μL. The amounts of antibody used for labeling are 0.02 μg/μL and 0.01 μg/μL for anti-FLAG-PE antibody and anti-6×His-FITC antibody, respectively. For the engineering of TEV-Fast, anti-6×His-FITC antibody was replaced by the anti-HA-FITC antibody (Genscript) with the same concentration and labeling process. The antibody-labeled cells were washed and resuspended in 1× PBS buffer and analyzed by a BD Biosciences FACSAria II flow cytometer. To avoid the signal interference, the sorting was performed using gates set on 575/30-nm as well as 510/20-nm emission filters in the flow cytometer. A total of ∼2 × $10^8$ cells were screened. To avoid the bacterial contamination, penicillin and streptomycin were added into the growth and inducing medium, with the final concentration of 100 units and 100 μg/mL, respectively. After four or five rounds of cell sorting and resorting, 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 and then transformed into *E. coli* and sequenced to obtain the mutated gene sequence information.

In similar steps, human Abelson tyrosine kinase (AblTK) expressed cells were grown, induced, and then labeled using Alexa Fluor 647 antiphosphotyrosine antibody (BioLegend) and anti-6×His-FITC antibody (Genscript) with the final concentration of 0.03 μg/μL and 0.01 μg/μL, respectively. The gates were set on a 660/20-nm emission filter for Alexa Fluor 647 as well as a 510/

20-nm emission filter for FITC in the BD Biosciences FACS Aria II flow cytometer.

**Protease Characterization.** TEV-P and select variants were expressed and purified as previously described (7). Kinetic assays were carried out as previously described with slight modifications (1). To monitor the cleavage of fusion proteins by TEV-P or its variants, the MBP (maltose binding protein) and the GST (glutathione s-transferase) protein were fused with a peptide linker containing ENLYFXS, where X can be Q, H, or E. The respective fusions were designated MBP-ENLYFQS-GST, MBP-ENLYFES-GST, and MBP-ENLYFHS-GST. For kinetic analysis, substrate peptides, TENLYFQSGTRRW, TENLYFESGTRRW, and TENLYFHSGTRRW, were purchased from Genscript (cleavage site after the underlined residue). All purified enzymes were >95% pure as determined by SDS/PAGE with Coomassie staining (7). Kinetic assays were carried out in 50 mM Tris·HCl buffer, pH 8.0, containing 1 mM EDTA and 2 μM freshly prepared DTT (1). A total of 5 μM to 6 mM of substrate peptide was incubated with 0.025–5 μM purified enzymes at 30 °C for 10–30 min. The reactions were quenched with freshly prepared 0.5% trifluoroacetic acid (TFA) (Sigma) followed by freezing at −80 °C. All of the enzymatic reactions were analyzed by HPLC on a Phenomenex $C_{18}$ reverse-phase column, using the acetonitrile gradient from 15% to 90% and a flow rate of 1 mL/min. The product amount was calculated upon the integration area at 280 nm and fitted to nonlinear regression of the Michaelis–Menten equation, using KaleidaGraph software (Synergy Software). LC-MS (ESI) of proteolysis products was performed on a Magic 2002 instrument (μM Bioresources).


1. Kapust RB, et al. (2001) Tobacco etch virus protease: Mechanism of autolysis and rational design of stable mutants with wild-type catalytic proficiency. *Protein Eng* 14(12):993–1000.
2. Varadarajan N, Cantor JR, Georgiou G, Iverson BL (2009) Construction and flow cytometric screening of targeted enzyme libraries. *Nat Protoc* 4(6):893–901.
3. Benatuil L, Perez JM, Belk J, Hsieh CM (2010) An improved yeast transformation method for the generation of very large human antibody libraries. *Protein Eng Des Sel* 23(4):155–159.
4. Drummond DA, Iverson BL, Georgiou G, Arnold FH (2005) Why high-error-rate random mutagenesis libraries are enriched in functional and improved proteins. *J Mol Biol* 350(4):806–816.
5. Cobaugh CW, Almagro JC, Pogson M, Iverson B, Georgiou G (2008) Synthetic antibody libraries focused towards peptide ligands. *J Mol Biol* 378(3):622–633.
6. Jung ST, et al. (2010) Aglycosylated IgG variants expressed in bacteria that selectively bind FcgammaRI potentiate tumor cell killing by monocyte-dendritic cells. *Proc Natl Acad Sci USA* 107(2):604–609.
7. Tropea JE, Cherry S, Waugh DS (2009) Expression and purification of soluble His(6)-tagged TEV protease. *Methods Mol Biol* 498:297–307.

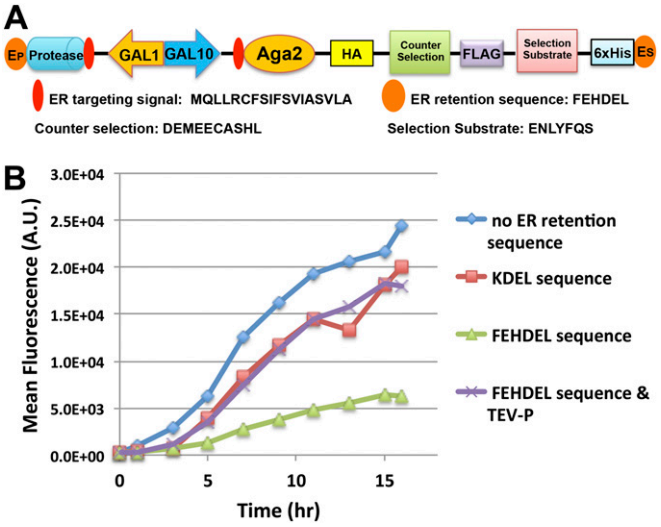




**Fig. S1.** ER retention effects in the YESS system. (*A*) Vector model of the constructs used in the time-course experiments to evaluate the ER retention effects of different ER retention sequences. In the experimental constructs, the counterselection gene encodes the substrate of hepatitis C virus (HCV) protease (DEMEECASHL), and the selection substrate gene encodes the substrate of TEV-P (ENLYFQS). (*B*) Time-course experiments. Cells containing different vectors were grown, induced, and analyzed using FACS (details *SI Methods*). Blue line: Vector contains no protease gene and no ER retention sequence at the C terminus of substrate (construct pESD-E, Table S1). Red line: Vector contains no protease gene but the KDEL ER retention sequence at the C terminus of substrate (construct pESD-F, Table S1). Green line: Vector contains no protease gene but the FEHDEL ER retention sequence at the C terminus of substrate (construct pESD-G, Table S1). Purple line: Vector contains the TEV-P gene with the FEHDEL ER retention sequence anchored at its C terminus and also the FEHDEL ER retention sequence at the C terminus of substrate (construct pESD-H, Table S1). The existence of the TEV-P will cause the cleavage at the selection substrate region, removing the ER retention sequence from the C terminus of the substrate. Ep, ER retention sequence at C terminal of protease; Es, ER retention sequence at C terminal of substrate; Aga1 and Aga2, subunits of the yeast adhesion receptor.

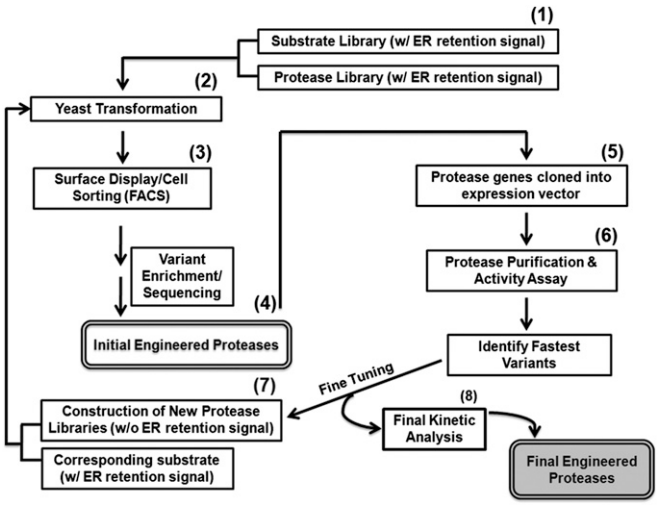



**Fig. S2.** Scheme of the general strategy for screening the evolved protease. Shown are the general steps of the protease library sorting using the yeast ER sequestration screening (YESS) system. Step 1: Generating the substrate and protease gene libraries. Step 2: Performing the yeast transformation to incorporate the substrate and protease gene libraries. Step 3: Yeast cells were sorted via FACS analysis. Step 4: Obtaining the initial evolved proteases. Step 5: Evolved proteases obtained in step 4 were subcloned into expression vector. Step 6: Initial characterization of the evolved proteases obtained in step 4 to identify the desired variants. Step 7: On the basis of the results of step 6, the desired variants were chosen for the second round of library sorting. New protease gene libraries based on these variants were generated and incorporated into the new construct, in which the ER retention sequence was removed from the C terminus of the protease to generate a more stringent proteolytic condition. Steps 2–7 were repeated until the desired variants were obtained. Step 8: Detailed kinetic analysis was performed for the obtained protease variants.

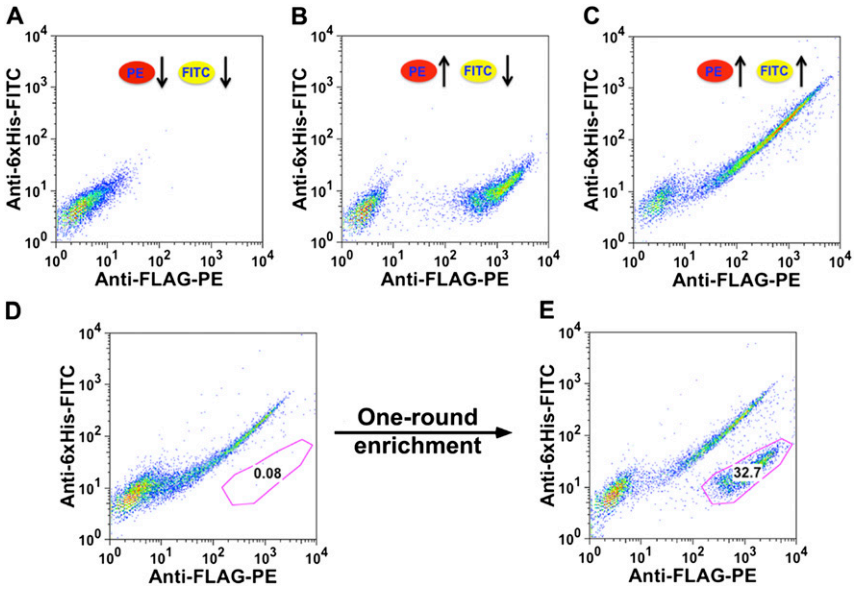



**Fig. S3.** Validation of the YESS system, using the TEV-P. The YESS system was validated using the TEV-P with its canonical substrate ENLYFQS and mutated substrate ENLYFKS. All of the constructs were generated on the basis of the pESD vector model (detailed vector information in Table S1). After induction with galactose, all of the cells were labeled with anti-FLAG-PE and anti-6×His-FITC antibodies followed by FACS analysis. (*A*) Cells containing the construct pESD-I. (*B*) Cells containing the construct pESD-J. (*C*) Cells containing the construct pESD-K. (*D*) Cells containing pESD-I, pESD-J, or pESD-K were mixed with a cell density ratio of 500:1:500, respectively. The sorting gate was drawn on the basis of the signals presented by the cells containing construct pESD-J. (*E*) Cells after one round of enrichment of the mixture in *D*. Cells falling in the gate were sorted, grown on an agar plate, and randomly picked for sequencing. Six of 10 sequences were identified as the target protease vector (pESD-J). The enrichment is from 1:1,000 to 6:10, which gives a one-round enrichment factor of 600.

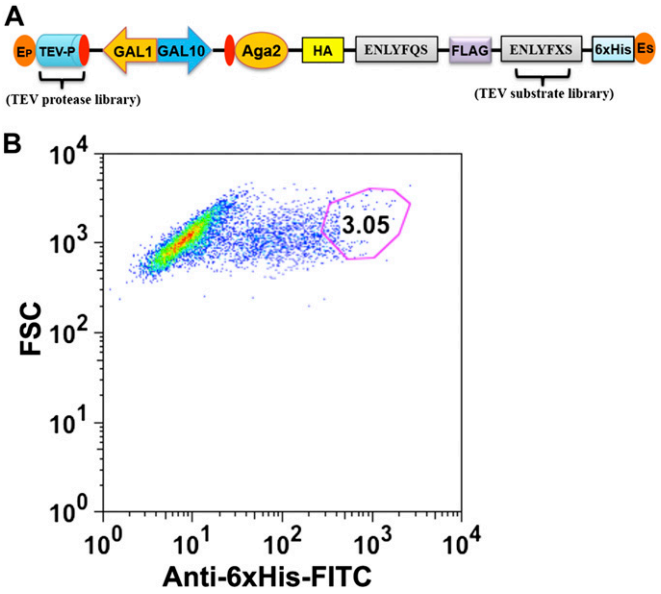



**Fig. S4.** Preparation of the TEV-P substrate P1 position library. (*A*) Vector model of the constructs used in the prescreening of the TEV-P substrate library (pESD-L, Table S1, without the TEV-P S1 pocket library) and the screening of the TEV-P S1 pocket library (pESD-M, Table S1, with the TEV-P S1 pocket library). (*B*) FACS data of the prescreening of the TEV-P substrate library based on the vector pESD-L. Cells were labeled with the anti-6×His-FITC antibody, and the plasmid DNA of the top 3.0% of the cells presenting the highest fluorophore signals was collected.

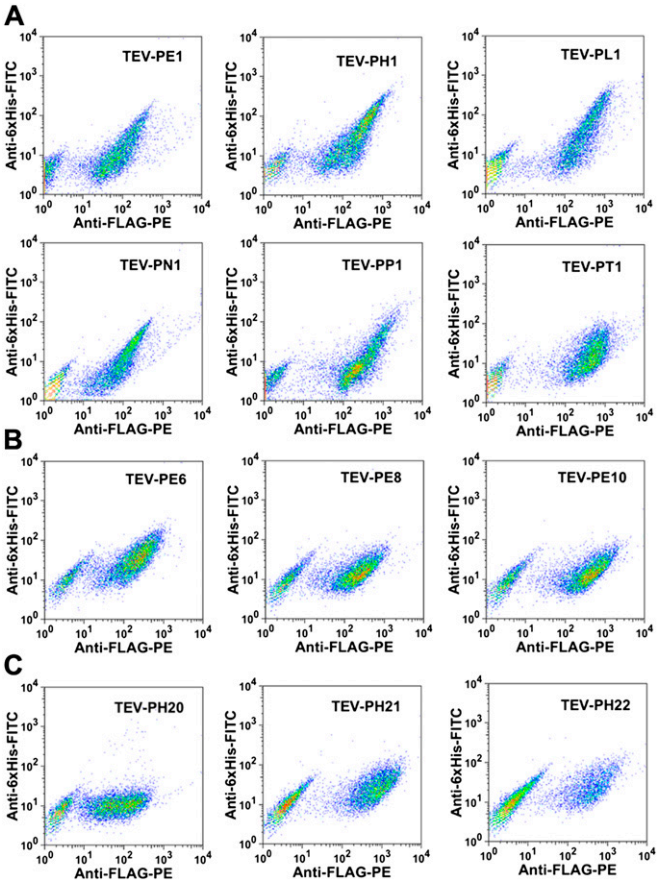



**Fig. S5.** FACS data of the selected TEV-P variants after the cell sorting of the S1 pocket library and the error-prone PCR libraries. (*A*) Representative single colonies obtained from the S1 pocket library sorting were sequenced, grown, induced, and analyzed using FACS (details in *SI Methods*). (*B*) Representative single colonies obtained for recognizing the ENLYFES substrate from the sorting of the TEV-PE3–based error-prone PCR library. The ER retention sequence was removed from the C terminus of the protease in the construct. (*C*) Representative single colonies obtained for recognizing the ENLYFHS substrate from the sorting of the TEV-PH7 based error-prone PCR library. The ER retention sequence was removed from the C terminus of the protease in the construct. See detailed variant information in Table S2.

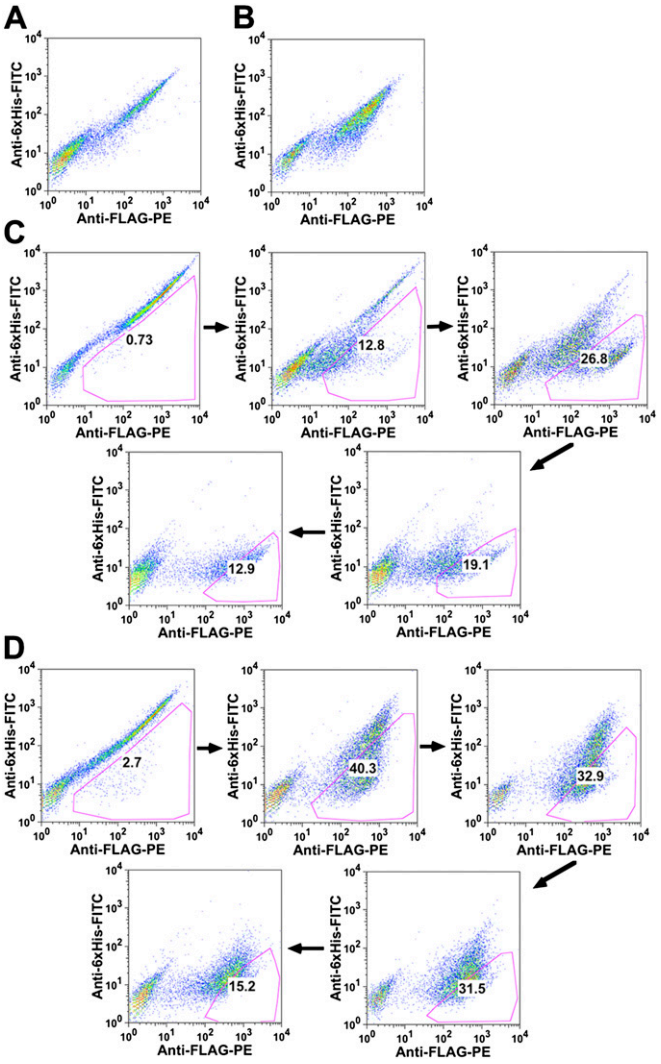



**Fig. S6.** FACS data of the sorting process of the error-prone PCR libraries. Error-prone PCR libraries were constructed using the genes encoding TEV-PE3 or TEV-PH7 as the template, and the cells were sorted as stated in *SI Methods*. (*A*) Cells expressing the mutated TEV-PE3 in which the ER-retention signal peptide was removed from the C terminus of the protease. (*B*) Cells expressing the mutated TEV-PH7 in which the ER-retention signal peptide was removed from the C terminus of the protease. (*C*) FACS data of the cell sorting of the TEV-PE3–based error-prone PCR library. Data were recorded after the first-, second-, third-, and fourth-round enrichments. (*D*) FACS data of the cell sorting of the TEV-PH7–based error-prone PCR library. Data were recorded after the first-, second-, third-, and fourth-round enrichments.

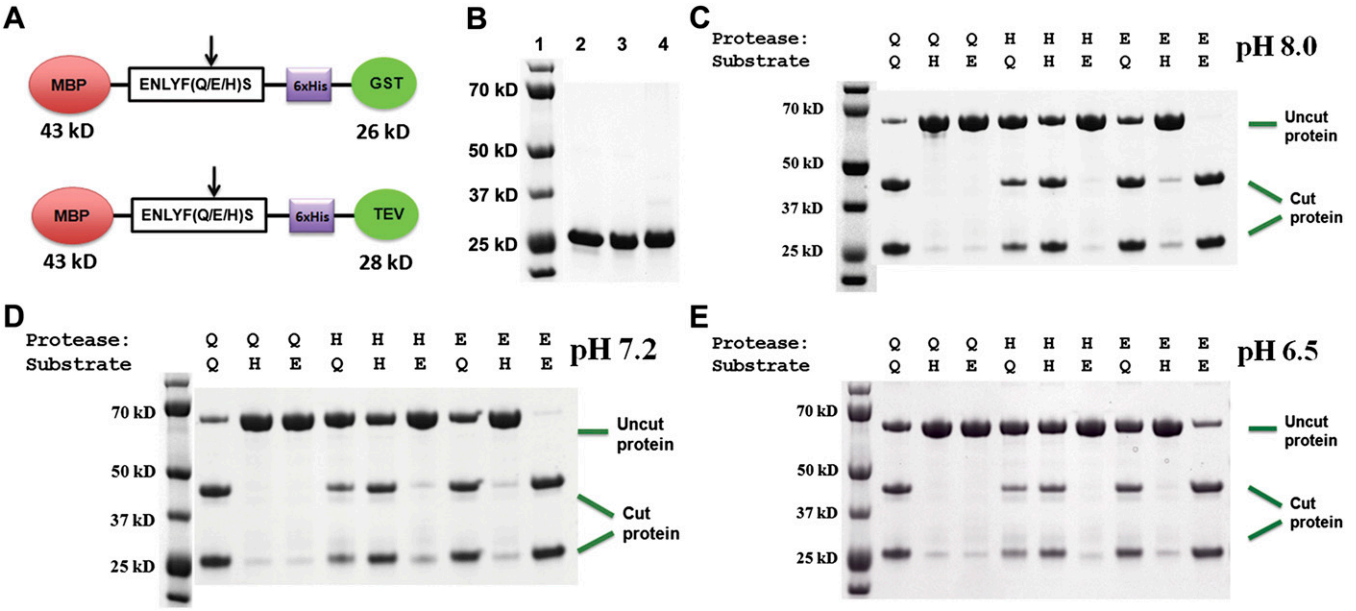



**Fig. S7.** Protein substrate digestion by TEV-P, TEV-PE10, and TEV-PH21 under different pH conditions. (*A*) Diagrams of the protein substrate purification constructs (MBP-ENLYFXS-6×His-GST) and the protease purification constructs (MBP-ENLYFXS-6×His-TEV-P). X can be Q, E, and H, corresponding to TEV-P, TEV-PE10, and TEV-PH21, respectively. (*B*) SDS/PAGE analysis of the purified TEV-P, TEV-PE10, and TEV-PH21. Lane 1, molecular mass ladders; lane 2, 10 μg purified TEV-P; lane 3, 10 μg purified TEV-PE10; lane 4, 10 μg purified TEV-PH21. (*C*) Protein substrate digestion reactions were performed at pH 8.0; all reactions were incubated at 30 °C for 1 h with 5 μg protein substrates mixed with 0.1 μg proteases in 20 μL reaction buffer. (*D*) Protein substrate digestion reactions performed at pH 7.2; same reaction conditions as in *C*. (*E*) Protein substrate digestion reactions performed at pH 6.5; same reaction conditions as in *C*.

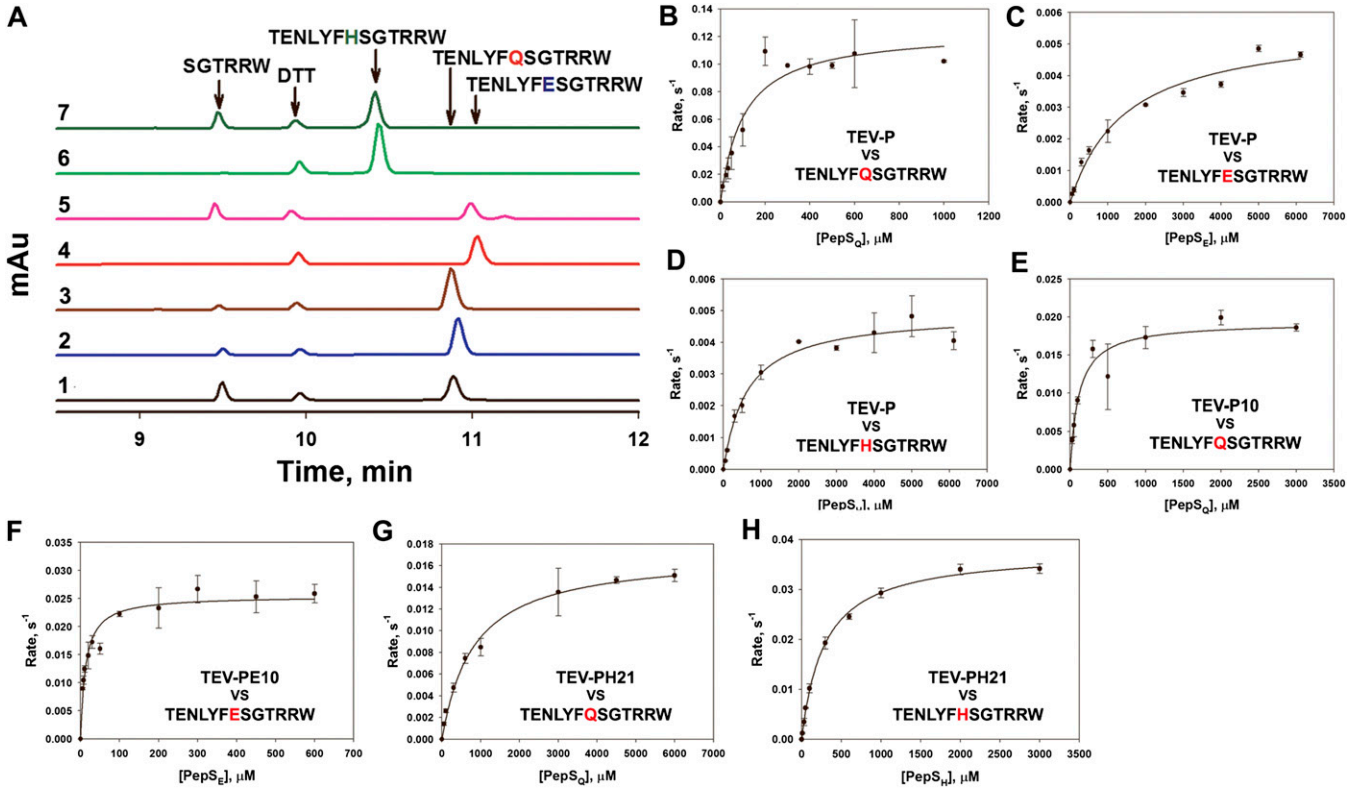



**Fig. S8.** Kinetic studies of TEV-P, TEV-PE10, and TEV-PH21 against different peptide substrates through HPLC analysis. All reactions of purified protease and synthesized peptides were incubated at 30 °C for 1 h with 100 μM peptide substrates. The total reaction volume is 100 μL. The different peptides were eluted at different time points according to the acetonitrile gradients. Data were plotted with the peptide–substrate concentrations against the rates ($s^{-1}$). (*A*) HPLC analysis of the digestion of peptide substrates by the TEV-P, TEVPE10, and TEV-PH21. 1, 0.1 μM TEV-P incubated with 100 μM TENLYFQSGTRRW; 2, 0.1 μM TEV-PE10 incubated with 100 μM TENLYFQSGTRRW; 3, 0.5 μM TEV-PH21 incubated with 100 μM TENLYFQSGTRRW; 4, 1 μM TEV-P incubated with 100 μM TENLYFESGTRRW; 5, 0.1 μM TEV-PE10 incubated with 100 μM TENLYFESGTRRW; 6, 1 μM TEV-P protease incubated with 100 μM TENLYFHSGTRRW; 7, 0.5 μM TEV-PH21 incubated with 100 μM TENLYFHSGTRRW. The corresponding products were confirmed using mass spectrometry. (*B*–*H*) Plots of substrate concentration vs. rate of (*B*) the TEV-P against the TENLYFQSGTRRW, (*C*) the TEV-P against the TENLYFESGTRRW, (*D*) the TEV-P against the TENLYFHSGTRRW, (*E*) the TEV-PE10 against the TENLYFQSGTRRW, (*F*) the TEV-PE10 against the TENLYFESGTRRW, (*G*) the TEV-PH21 against the TENLYFQSGTRRW, and (*H*) the TEV-PH21 against the TENLYFHSGTRRW.

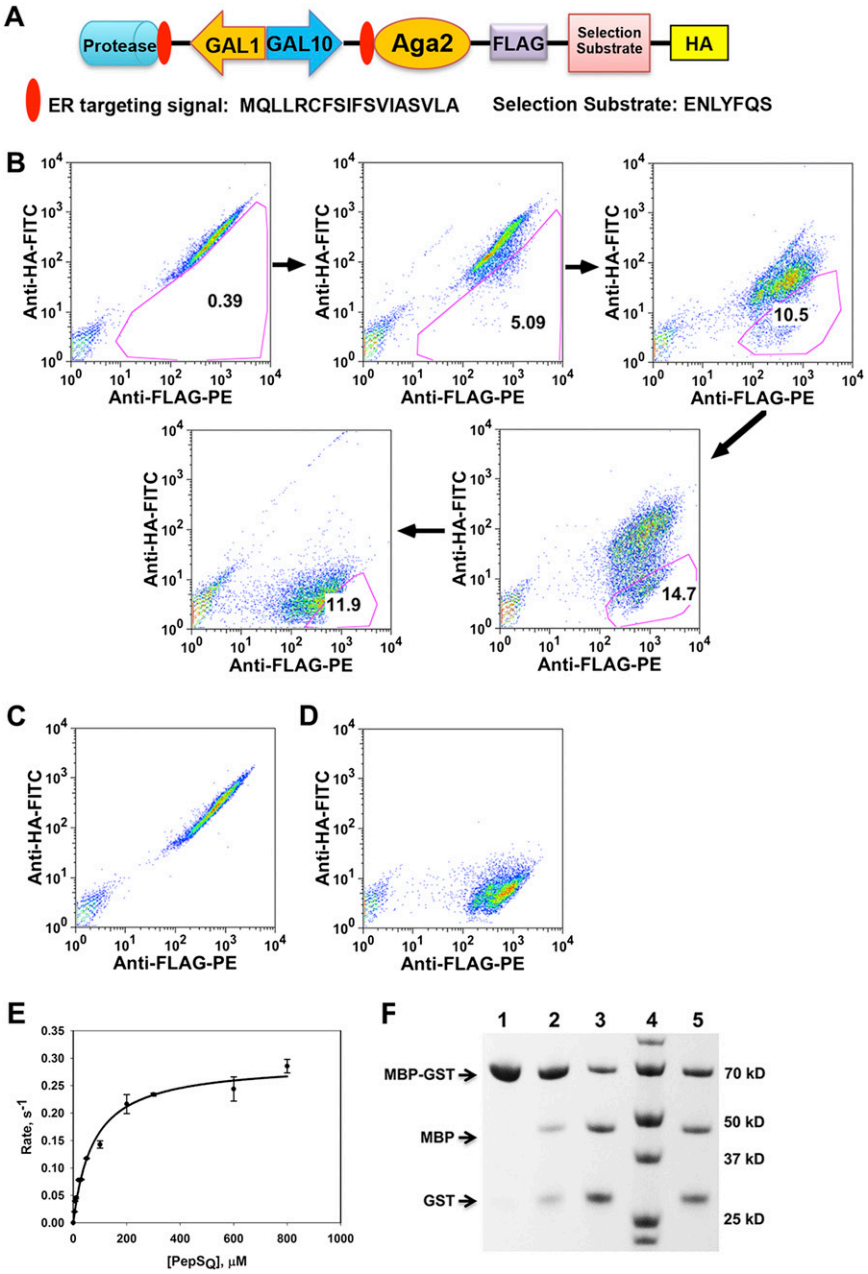
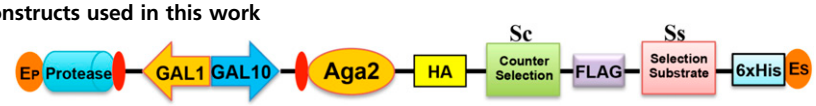



**Fig. S9.** Engineering TEV-P to display increased catalytic activity. (*A*) Relevant portion of the DNA vector (pESD-T) used in the screening of TEV-P variants displaying increased catalytic activity. (*B*) FACS histograms of an error-prone PCR library of TEV-P and the cell populations isolated following five consecutive rounds of enrichments, using the YESS system. The error-prone PCR library was constructed using the genes encoding TEV-P as the template, and the cells were sorted as described in *SI Methods*. (*C*) FACS histogram of cells expressing TEV-P from a DNA construct as in *A*. (*D*) FACS data of cells expressing the isolated variant (TEV-Fast) from DNA construct as in *A*. (*E*) Michaelis–Menten kinetics of TEV-Fast with the peptide TENLYFQSGTRRW as the substrate. (*F*) Digestion of fusion proteins by TEV-P or TEV-Fast. Reactions were incubated at 30 °C, pH 8.0, for 20 min with 5 μg protein fusion substrate with or without 0.1 μg protease in a 20-μL reaction buffer. Lane 1, MBP-ENLYFQS-GST substrate only; lane 2, MBP-ENLYFQS-GST substrate incubated with TEV-P; lane 3, MBP-ENLYFQS-GST substrate incubated with TEV-Fast; lane 4, molecular mass ladders; lane 5, MBP-ENLYFQS-GST substrate incubated with S219V variant.

**Table S1. pESD constructs used in this work**




| Construct | Ep | Enzyme | Sc | Ss | Es | Cutting position | PE signal | FITC signal |
|---|---|---|---|---|---|---|---|---|
| pESD-E | ✗ | ✗ | DEMEECASHL | ENLYFQS | ✗ | None | ↓↓ | ↓↓ |
| pESD-F | ✗ | ✗ | DEMEECASHL | ENLYFQS | ✓ (KDEL) | None | ↑↑ | ↓↓ |
| pESD-G | ✗ | ✗ | DEMEECASHL | ENLYFQS | ✓ | None | ↑↑ | ↑↑ |
| pESD-H | ✓ | TEV-P | DEMEECASHL | ENLYFQS | ✓ | Ss | ↑↑ | ↑↑ |
| pESD-I | ✓ | TEV-P | ENLYFQS | ENLYFKS | ✓ | Sc | ↓↓ | ↓↓ |
| pESD-J | ✓ | TEV-P | ENLYFKS | ENLYFQS | ✓ | Ss | ↑↑ | ↓↓ |
| pESD-K | ✓ | TEV-P | DEMEECASHL | ENLYFKS | ✓ | None | ↑↑ | ↑↑ |
| pESD-L | ✗ | ✗ | ENLYFQS | ENLYFXS | ✓ | — | ↑↑ | ↑↑ |
| pESD-M | ✓ | TEV-P S1 library | ENLYFQS | ENLYFXS | ✓ | — | | |
| pESD-N | ✓ | ✗ | ENLYFQS | DEMEECASHL | ✓ | None | ↑↑ | ↑↑ |
| pESD-O | ✓ | HCV-P | ENLYFQS | DEMEECASHL | ✓ | Ss | ↑↑ | ↓↓ |
| pESD-P | ✓ | ✗ | ENLYFQS | CGYGPKKKRKVGG | ✓ | None | ↑↑ | ↑↑ |
| pESD-Q | ✓ | GrK | ENLYFQS | CGYGPKKKRKVGG | ✓ | Ss | ↑↑ | ↓↓ |
| pESD-R* | ✓ | AblTK | ✗ | FKGSTAENAEYLRVAPQSSEF | ✓ | Y[†] | ↑↑[‡] | ↑↑ |
| pESD-S* | ✗ | ✗ | ✗ | FKGSTAENAEYLRVAPQSSEF | ✓ | N[†] | ↓↓[‡] | ↑↑ |
| pESD-T[§] | ✗ | ✗ | ✗ | ENLYFQS | ✗ | Ss | ↑↑[§] | ↓↓[§] |

Ep, ER retention sequence at C terminus of TEV-P (FEHDEL); Es, ER retention sequence at C terminus of substrates (FEHDEL if not other annotated); CGYGPKKKRKVGG, human Grk substrate sequence; DEMEECASHL, HCV-P substrate sequence; ENLYFQS, TEV-P canonical substrate sequence; ENLYFKS, mutated TEV-P substrate sequence; AblTK, human Abelson tyrosine kinase; Sc, counter selection substrate; Ss, selection substrate.

*pESD-R and pESD-S do not contain the fragments of HA-Sc-FLAG in the substrate fusion polypeptide.

[†]Tyrosine phosphorylation.

[‡]Fluorescent signal of Alexa Fluor647.

[§]Construct pESD-T was generated in a different model with no counterselection substrate; it has the fragments of Aga2-FLAG-Ss-HA. Details can be seen in Fig. S9.

**Table S2. Evolved TEV-P variants obtained after the cell sorting of the S1 pocket library and the error-prone PCR libraries**

| Number | Variant, P1 | T146 | D148 | H167 | S170 | Other mutations |
|---|---|---|---|---|---|---|
| TEV-PE1 | E | T | K | H | S | |
| TEV-PE2 | E | T | P | H | R | |
| TEV-PE3 | E | T | R | H | S | |
| TEV-PE4 | E | T | R | H | A | T22A, L56W, F172L, T173A, Q197R |
| TEV-PE5 | E | T | R | H | A | V125A, T173A, N174H |
| TEV-PE6 | E | T | R | H | A | K65R, T173A, N177K, F186L, M218I |
| TEV-PE7 | E | T | R | H | A | P92L, R108H, M121L, S122P, T173A, N177K |
| TEV-PE8 | E | T | R | H | S | N171S |
| TEV-PE9 | E | T | R | H | S | N177K |
| TEV-PE10 | E | T | R | H | S | S120R, T173A, N177K, M218I |
| TEV-PE11 | E | A | P | H | A | T173A, N177K |
| TEV-PE12 | E | A | P | H | A | T173A, N177K, V199D |
| TEV-PE13 | E | A | P | H | A | T173A, N177K, Q196R |
| TEV-PE14 | E | A | P | H | A | K65E, T173A, N177K, F179L |
| TEV-PE15 | E | A | P | H | A | N12D, I163V, T173A, N177K, M218I |
| TEV-PE16 | E | A | P | H | A | M124I, T173A, N177K, K184R, P221S |
| TEV-PE17 | E | A | P | H | A | C110R, R159G, T173A, N177K, N192D, E223G |
| TEV-PH1 | H | A | A | H | S | V228A |
| TEV-PH2 | H | A | A | H | T | |
| TEV-PH3 | H | A | A | H | T | R203Q |
| TEV-PH4 | H | A | S | H | T | |
| TEV-PH5 | H | A | P | H | A | |
| TEV-PH6 | H | A | P | H | A | E106G, T173A |
| TEV-PH7 | H | A | P | H | A | T173A, M218I |
| TEV-PH8 | H | A | P | H | S | |
| TEV-PH9 | H | A | P | H | S | K89R, T173A |
| TEV-PH10 | H | A | P | H | S | Q96R |
| TEV-PH11 | H | A | P | H | T | |
| TEV-PH12 | H | C | P | H | T | |
| TEV-PH13 | H | C | Q | H | S | N171D |
| TEV-PH14 | H | V | A | H | S | |
| TEV-PH15 | H | V | P | H | A | N171S |
| TEV-PH16 | H | V | P | H | S | T128S, D136G |
| TEV-PH17 | H | V | P | H | S | D136G |
| TEV-PH18 | H | V | P | H | T | K147T |
| TEV-PH19 | H | V | R | H | S | |
| TEV-PH20 | H | A | P | H | A | T173A |
| TEV-PH21 | H | A | P | H | A | T17A, S153C, S168T, T173A |
| TEV-PH22 | H | A | P | H | A | Y11F, C110R, I144T, T173A, F186L, M218I |
| TEV-PL1 | L | C | A | V | T | T173N |
| TEV-PL2 | L | C | A | V | T | T173N, N192S |
| TEV-PL3 | L | C | P | V | T | T17A, T173N, K184R |
| TEV-PL4 | L | C | R | V | T | T70M, T173N |
| TEV-PL5 | L | Q | R | V | T | Q58K, K99E, T173N |
| TEV-PN1 | N | A | S | H | T | |
| TEV-PN2 | N | V | E | H | T | Q104R, T173A |
| TEV-PN3 | N | V | P | H | S | D136G |
| TEV-PN4 | N | V | P | H | S | P39H, D136G |
| TEV-PN5 | N | V | P | H | A | N171S |
| TEV-PN6 | N | V | R | H | S | |
| TEV-PP1 | P | T | D | Y | L | G213C |
| TEV-PT1 | T | V | R | Q | A | T113A, T173A |

Variants all contain the S219P mutation. Variants TEV-PE4 to TEV-PE17 and TEV-PH20 to TEV-PH21 were obtained from the error-prone PCR libraries.

**Table S3. List of primers**

| Primers | Sequences |
|---|---|
| 1 | 5′-GGAAGCATTGGATTCAANNSAAGNNSGGGCAGTGTGGCAGTCC-3′ |
| 2 | 5′-ATTAGTATCAACTAGAGATGGGTTCATTGTTGGTATA-3′ |
| 3 | 5′-ATCTCTAGTTGATACTAATGGACTGCCACACTGCCC-3′ |
| 4 | 5′-TTTGTGTTGGTGAAATTSNNTGCTGASNNTATACCAACAATGAACCC-3′ |
| 5 | 5′-TTGAATCCAATGCTTCCAGAA-3′ |
| 6 | 5′-AATTTCACCAACACAAACAA-3′ |
| 7 | 5′-TACCATCTGCAGAGCGACGGCGACGACGATTCATGAG-3′ |
| 8 | 5 -ATGGTTGGTACCGAAAATCTTTATTTTAGCGGTCATCATCATC-3′ |
| 9 | 5′-CGTCAAGGAGAAAAAACCCCGGATCCGTAATACGACTCACTATAGGGCCCGGGCGTCGACATGC AACTTTTGAGATGCTTCAGTATTTTCAGCGTCATCGCCAGTGTGCTGGCCAGCTTGTTTAAGGG GCCGCGTG-3′ |
| 10 | 5′-GTACAGTGGGAACAAAGTCGATTTTGTTACATCTACACTGTTGTTATCAGATCTCGAGCGGTAC CTTACTCATTACAATTCGTCGTGTTCGAAACTACCCAAGTCCTCTTCAGAAATAAGCTTTTGTT CGGATCCATTCATGAGTTGAGTCGCTTCC-3′ |
| 11 | 5′-ATGGCTGGCCCAGCCGGCCAGCTTGTTTAAGGGGCCGCG-3′ |
| 12 | 5′-GTCCATGGCCCCCGAGGCCTTAATTCATGAGTTGAGTCGCTTCCTTAAC-3′ |
| 13 | 5′-GTACAGTGGGAACAAAGTCGATTTTGTTACATCTACACTGTTGTTATCAGATCTCGAGCGGTA CCTTACTCATTAATTCATGAGTTGAGTCGCTTCC-3′ |
| 14 | 5′-GAGCTCACAATTCGTCGTGTTCGAAACTACCATGATGATGATGATGACTGCCAGASNNGAA ATACAAATTTTCACTGCCTTTATCGTCGTCATCTTTATAATC-3′ |
| 15 | 5′-CGAATTCAACCCTCACTAAAGGGCGGCCGCACTAGTATCGATG-3′ |