Published in final edited form as:

Curr Opin Biotechnol. 2009 August ; 20(4): 390–397. doi:10.1016/j.copbio.2009.07.003.

Engineering Next Generation Proteases

Mark Pogson¹, George Georgiou^{1,2,3}, and Brent L. Iverson^{1,4}

¹Institute for Cellular and Molecular Biology

²Dept. of Chemical Engineering

³Dept. of Biomedical Engineering

⁴Dept. of Chemistry & Biochemistry

Abstract

The engineering of novel and precise sequence specificity into proteases will provide an important route to the development of exciting new tools for analytical, biotechnological, and therapeutic applications. Substantial progress has been made in reprogramming protease specificity, largely due to the development of high-throughput assay technologies that allow the isolation of novel protease variants from large libraries, a process referred to as directed evolution. Using directed evolution as well as other approaches, proteases have been reprogrammed to cleave specifically and with high activity peptide substrates containing a variety of different amino acids in the P1 and P1' positions. Protease variants have also been produced displaying activity towards post-translationally modified tyrosine, a specificity not yet identified in any naturally occurring protease. Together, these and other recent advances suggest that substantial progress is being made that could soon enable the widespread application of engineered proteases.

Introduction

Proteases make up the largest family of enzymes constituting an estimated 2% of all human genes [1,2]. As such, proteases are involved in a vast array of cell functions, and are intimately tied to cell survival at many different levels [3]. While some proteases, such as those involved with digestion and protein recycling in cells, are relatively non-selective, most proteases have evolved to recognize a limited set of peptide sequences. For example, the catalytic cascades associated with complex functions such as thrombosis, complement activation, and apoptosis are critically dependent on the exquisite substrate specificities of the participating proteases [4–6].

Proteases have proven to be indispensable tools for industrial applications and biomedical research, especially proteomics. For example, protease digestion is a routine sample preparation step for proteomic mass spectral analysis of complex protein mixtures. In addition, proteases have been utilized as therapeutics, with many clinical applications relegated to

© 2009 Elsevier Ltd. All rights reserved.

Correspondence: Brent L. Iverson: biverson@mail.utexas.edu, George Georgiou: gg@che.utexas.edu.

Addresses: Mark Pogson, The University of Texas at Austin, 2500 Speedway, Austin, TX 78712

Prof. George Georgiou, The University of Texas at Austin, 1 University Station C0400, Austin, TX 78712

Prof. Brent L. Iverson, The University of Texas at Austin, 1 University Station A5300, Austin, TX 78712

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

enzyme replacement in genetically deficient patients. One of the most visible protease therapeutics is probably tissue plasminogen activator (tPA) and its derivatives (Activase, TNKase, Retavase, among others), which are used to dissolve thrombi during myocardial infarction and stroke [7]. Of the six therapeutic proteases available in 2006, two have been engineered for enhanced properties such as serum half-life, but specificity and selectivity have remained unaltered [8,9].

The ability to precisely engineer protease substrate specificity advances the potential to use engineered proteases as therapeutics. As opposed to antibody-based therapies that function in a stoichiometric fashion, engineered proteases may be deployed for the catalytic and irreversible inactivation of target proteins. Thanks to catalytic turnover, a single protease molecule can inactivate numerous target proteins, thereby reducing dosing and treatment costs. The possibility of engineering novel therapeutic proteases has spawned at least two companies (Catalyst Biosciences and DIREVO Biotech), and several others have added enzyme engineering projects to their pipelines.

Earlier efforts to engineer the specificity of proteases have generally centered on structure-based approaches involving the transplantation of specificity between homologous enzymes, as in the conversion of trypsin to chymotrypsin, and the generation of subtilisin BPN' variants engineered to have furin-like specificity [10–14]. Although these and similar efforts pioneered the field of protease engineering, generating protease variants with truly novel substrate specificities has remained more elusive. Herein we review recent successes in the field of engineering protease specificity from the past two years and discuss some of the intriguing challenges that still remain.

Creating diversity to alter specificity

Protease specificity is denoted using the Schechter and Berger nomenclature [15] in which substrate residues are labeled PX-P1, where P1 is the new C-terminal residue following cleavage, and residues P1'-PX', where P1' is the new N-terminal residue following cleavage ($\text{H}_2\text{N-PX}\dots\text{P3-P2-P1}\downarrow\text{P1'-P2'-P3'}\dots\text{PX'-CO}_2\text{H}$). The corresponding subsites within the protease responsible for substrate recognition are labeled as SX-S1 and S1'-SX', respectively. Often, the functional consequences of amino acids that compromise the various protease subsites (SX-SX') are difficult or impossible to predict based on the crystal structure information alone. Presumably for this reason, rational mutagenesis approaches by themselves have been limited in their ability to generate highly active and selective protease variants displaying altered substrate specificity [16–18]. Nevertheless, site directed mutagenesis in and around a putative active site can be used to “zero in” efficiently on protease specificity determinants, thereby providing guidance for directed evolution approaches [19–20,21•,22•].

Technological advances in molecular biology have made gene assembly, modification and expression in microorganisms readily accessible [19,20,22•]. A common methodology for the somewhat random diversification of a gene sequence is the error-prone polymerase chain reaction (epPCR) [23], a technique that introduces mutations along the length of the gene being mutagenized. DNA shuffling and ITCHY-SCRATCHY [24,25] or their variations are employed for the homologous or non-homologous recombination of gene pools, often after some selection based on desirable phenotype has occurred. Targeted approaches such as saturation mutagenesis at selected codons are also used [26,27] when structural information is available.

The ability to interrogate the activities and substrate preferences of large numbers of enzyme variants is of particular importance in engineering proteases, because typically, a combination of amino acid changes is required to alter substrate specificity while maintaining high levels of catalytic activity. Although never truly predictable, choosing which residues to randomize

in a focused library can be informed by combining insights gleaned from structures as well as rational mutagenesis experiments that probe potential subsites and their vicinity [28•].

Assaying for Variants

The general scheme for isolating protease variants with desirable new attributes is easy to depict but technically very challenging to implement (Figure 1). Protease engineering has historically been most often accomplished using rational mutagenesis and low-throughput assays that preserve the essential link between genotype and phenotype. An example of two recently engineered proteases using methods such as this are the *E. coli* signal peptidase SPase I and rat brain neurolysin [29,30]. In both cases, enhanced activity for a desired new substrate was achieved using *in vitro* peptide hydrolysis assay technologies involving fluorescence resonance energy transfer (FRET) as indicators of activity. The specificity of SPaseI was reprogrammed to accept Arg, Gln, and Tyr residues at the P3 position instead of the wild-type preferred Ala or other small aliphatic amino acids. This change in specificity was accomplished by mutating two previously defined specificity-determining residues in SPase I [31]. The substrate specificity of rat brain neurolysin was altered at P2' by a single amino acid change to accommodate Glu in place of the wild-type preferred Ala, resulting in a 10-fold switch in specificity (Table 1).

The effective exploration of protease sequence space and the isolation of enzyme variants with desired catalytic properties from large libraries is predicated on the availability of quantitative high throughput assays for detecting catalytic activity. In simple terms, a higher throughput protease library screening assay increases the likelihood that rare variants can be found possessing extraordinary levels of a desired new activity.

However, the directed evolution of proteases poses unique complications compared to other enzyme systems. Expression of most proteases is toxic to a host cell due to adventitious cleavage of any number of essential host proteins. Also, in multi-round directed evolution studies, the conversion of a “specialist enzyme” with narrow substrate selectivity to a specialist enzyme variant with a different but equally narrow specificity often proceeds via “generalist” intermediates, present in the middle rounds, that display relaxed specificity [32]. Protease variants with relaxed substrate specificities are generally even more toxic to the host and therefore cannot be expressed effectively. In addition, there is always the possibility that a newly engineered protease will cleave and inactivate itself. Nonetheless, technologies that overcome these constraints and have led to the successful engineering of interesting and useful protease variants have been reported in recent years.

High-Throughput Assay Formats for Protease Engineering β -galactosidase Inactivation

In an attempt to counter enzyme evolution via promiscuous, or “generalist,” intermediates an *in vivo E. coli* screen based on β -galactosidase (β -gal) inactivation has been developed and applied to alter the specificity of the human immunodeficiency virus type I protease (HIV PR) [33•]. It has previously been demonstrated *in vivo* that HIV PR will cleave an insertion of a decapeptide HIV PR cleavage site into an accessible surface loop of β -gal when the two enzymes are co-expressed [34]. Furthermore, HIV PR cleavage is reduced and β -gal activity is increased when this decapeptide is altered to a different sequence found in the human cytokine TNF- α . Therefore, protease activity can be monitored as a decrease in β -gal activity (Figure 2A).

The wild-type HIV PR is cytotoxic due to cleavage of essential *E. coli* proteins. Consequently, this toxicity was exploited to exclude promiscuous variants generated from an ePCR library.

Surviving variants were then assayed for a decrease in β -gal activity. Using this approach, an HIV PR variant containing two amino acid mutations, P9S/I50L, was recovered and further demonstrated to be specific for the new decapeptide sequence coupled with non-toxic expression (Table 1). In this respect, the evolution of HIV PR specificity was favored to bypass the “generalist” intermediates typically recovered during directed enzyme evolution [32].

The Yeast GASP system

A common theme in gene regulation is the localization of a transcriptional activator at a site that is physically separated from the nucleus. Interactions with signaling molecules or post-translational modifications, such as proteolytic cleavage of an anchoring motif, mediate the translocation of the activator to the nucleus resulting in transcription. This concept, similar in principle to one previously described [35], was exploited in a yeast-based transcriptional activation screen, named GASP (genetic assay for site-specific proteolysis), which was employed to alter the P2 specificity of the Hepatitis A Virus 3C protease (HAV 3CP) (Figure 2A) [36]. Briefly, in this approach, the LexA-b42 transcription factor is sequestered at the cytoplasmic face of the plasma membrane by fusing it to an integral membrane protein using a cleavable linker. Upon specific cleavage of the linker, the transcription factor diffuses to the nucleus and activates transcription of a reporter gene. An HAV 3CP variant recovered from a focused four member saturation mutagenesis library displayed a greater than 5-fold reduction in wild-type activity, Thr at P2, and a 160-fold switch in specificity towards a peptide substrate containing Gln at P2 (Table 1).

E. coli Surface Expression Combined with FACS Sorting

A versatile and powerful high throughput screen has also been developed for assaying protease activity in *E. coli* [37]. The technique relies on expression of the protease on the bacterial surface followed by fluorescence-activated cell sorting (FACS). Substrates that generate positively-charged fluorescent protease cleavage products are produced and electrostatically captured on the negatively-charged bacterial surface (Figure 2B). The capture of the fluorescent product by the cell results in a fluorescence signal that is proportional to the catalytic activity of the enzyme. Thus, variants displaying a desired level of activity can be isolated.

Advantages of using a FACS-based approach to library screening include the very high throughput nature of FACS, exceeding 10^6 clones per hour. In addition, multiple colors can be used in conjunction with multiple substrates so that a simultaneous selection/counterselection can be accomplished in the same screen. Because FACS is a quantitative assay, library quality as well as the number of positive clones present under the given screening conditions can be ascertained in real-time. Such real-time analysis provides for the rapid optimization of conditions to favor isolation of only the desired variants. This optimization can be as simple as adjusting substrate concentration and/or the FACS sort gates.

The *E. coli* surface expression/FACS assay system has proven to be remarkably successful for the engineering of the native *E. coli* protease outer membrane protein T (OmpT). OmpT is a surface membrane protease that represents the prototypical member of the OmpT family from Gram-negative bacteria and cleaves substrates between dibasic residues, preferring Arg in both the P1 and P1' positions [38,39]. In addition, OmpT cleaves protamine and other cationic peptides, an activity that confers protection against the action of these antibiotic agents [40, 41].

The engineering of OmpT variants with altered specificity at P1', changing the preference from a basic residue preferred by wild-type OmpT to Val, has been previously reported [37]. Using a single FRET selection substrate and FACS, a variant with a 60-fold enhanced preference for cleavage at Arg-Val was isolated. The implementation of multi-color FACS by using

fluorophores with different emission wavelengths enabled the simultaneous assay for activity with selection and counter-selection substrates. Using this approach, an OmpT variant with several different mutations was isolated from an epPCR library that displayed a 3 million-fold preference for cleavage of an Ala-Arg sequence, while maintaining high overall catalytic activity, which was equivalent to wild-type OmpT cleaving its preferred Arg-Arg substrate. This dramatic shift in specificity coupled with near wild-type activity was traced to a single mutation, Ser223Arg, which when introduced into OmpT alone, resulted in a variant that retained high levels of activity while completely abolishing all activity for cleavage of Arg-Arg substrates [42••].

In subsequent studies, structural information regarding probable S1 residues led to the generation of a 3-member saturation mutagenesis library targeting the bottom of the S1 pocket of OmpT. Screening of this library yielded a variant specific for Glu-Arg [43••], a specificity that has not yet been observed in naturally occurring proteases found in the database. Importantly the active site mutations responsible for the switch in substrate selectivity of the Glu-Arg specific variant were shown to be epistatic (they did not provide benefit as individual mutations), thereby enhancing the importance of mutagenesis libraries targeting specific sets of residues.

The isolation of enzyme variants specific for aromatic or polar amino acids at P1 required full or partial saturation libraries encompassing up to 21 residues around the active site. These “expanding zone” libraries yielded enzyme variants specific for Tyr-Arg and Thr-Arg. Together these results validate the generality of using directed evolution principles coupled with a high-throughput FACS screen to isolate an entire family of protease variants reprogrammed to recognize a diverse array of peptide sequences. These newly accepted substrate sequences spanned the entire range of chemical properties such as charge, relative hydrophobicity, and size (Table 1). Importantly, every member of this variant family exhibited a level of overall catalytic activity that was similar to the wild-type enzyme reacting with its preferred Arg-Arg substrate.

Specificity for Post-Translationally Modified Amino Acids

Although natural protease specificity profiles exhibit a remarkable diversity of substrate preferences [44,45], no naturally occurring protease reported to date has been shown to recognize preferentially any post-translationally modified amino acids. Recently, however, two separate bacterial proteases, subtilisin BPN’ and OmpT, have been engineered to be specific at P1 for phosphorylated and sulfated tyrosine, respectively [46••,47••]. Structure guided mutagenesis of Subtilisin BPN’ at two positions lining the S1 pocket led to a variant (P129G, E156R) that displayed a 2500-fold switch in substrate selectivity, cleaving a phosphotyrosine containing substrate 5-fold more efficiently than the wild-type substrate. This same variant was found to be 2-fold more specific for phosphotyrosine over sulfotyrosine.

A sulfotyrosine-specific OmpT variant was isolated using the FACS-based assay from a 21-residue partial saturation mutagenesis library that was later backcrossed to wild-type OmpT using DNA shuffling. The sulfotyrosine specific variant displayed high overall catalytic efficiency (Table 1) and was over 200-fold more selective for sulfotyrosine over phosphotyrosine.

These two variants, specific for post-translationally modified Tyr, represent the crucial first step toward generating practical enzymes useful for the detection of post-translational modifications.

Current Limitations and Future Directions

As previously mentioned, proteases are difficult engineering targets for several reasons, including significant host toxicity problems and the potential for self-cleavage. Perhaps for these reasons, protease engineering is still an emerging field. The recent reports of high-throughput assays for screening protease libraries represent major milestones in the field, especially when used in conjunction with simultaneous selection/counterselection pressures using techniques such as FACS. Although these recent advancements are striking they are mainly centered on proteases of non-human origin. Protease utilization for human therapies will require the use of engineered human enzymes to reduce immunological responses. In addition, therapeutic proteases will require an extreme level of substrate specificity that will require expanded substrate recognition, extending well beyond P1 and P1', to minimize potentially lethal off-target cleavage events. In stark contrast, proteases intended for proteomic use, as in the detection of post-translationally modified amino acids, will need more relaxed substrate specificity displaying selectivity for the post-translationally modified amino acid only, regardless of the surrounding amino acid sequence.

A looming hurdle that must be confronted before engineered protease variants see widespread application, especially in the therapeutic realm, is their cleavage ability against fully folded target proteins. To date, most protease variants are characterized using the same peptide substrates as were used for screening, or a slight variation thereof. Consequently, with the exception of two OmpT variants that were shown to be active against biologically relevant targets [43••], it remains to be seen if activity against peptide substrates translates to activity against fully folded protein substrates. In other words, it is reasonable to expect that engineered proteases will require specific substrate conformations for activity, despite the presence of the same primary sequence used in a peptide selection substrate [48,49]. Thus, it might prove necessary to develop high throughput screens for library selection that utilize fully folded protein targets, not peptides.

Conclusions

The engineering of novel specificity into protease scaffolds holds tremendous promise as a route to new tools for analytical, biotechnological, and therapeutic applications. Important successes have been reported over the years based upon the rational redesign of active sites guided by homologue sequence replacement or other structural insights. Nevertheless, directed evolution approaches have yielded the most promising recent results in the protease engineering field. Next generation high throughput, activity-based screens will enable even more extensive fine-tuning of specificity, without compromising overall protease catalytic activity.

Acknowledgment

We thank M. Jack Borrok, G. Skreatas, and T. Van Blarcom for helpful discussions and reading the manuscript.

References and recommended reading

1. Marnett AB, Craik C. Papa's got a brand new tag: advances in identification of proteases and their substrates. *Trends in Biotechnology* 2005;23:59–64. [PubMed: 15661339]
2. Schilling O, Overall C. Proteome-derived, database-searchable peptide libraries for identifying protease cleavage sites. *Nat Biotechnol* 2008;26:685–694. [PubMed: 18500335]
3. Doucet A, Overall C. Protease proteomics: revealing protease in vivo functions using systems biology approaches. *Mol Aspects Med* 2008;29:339–358. [PubMed: 18571712]
4. Chowdhury D, Lieberman J. Death by a thousand cuts: granzyme pathways of programmed cell death. *Annu Rev Immunol* 2008;26:389–420. [PubMed: 18304003]

5. Sjöberg AP, Trouw LA, Blom AM. Complement activation and inhibition: a delicate balance. *Trends in Immunology* 2009;30:83–90. [PubMed: 19144569]
6. Krem MM, Di Cera E. Evolution of enzyme cascades from embryonic development to blood coagulation. *Trends Biochem Sci* 2002;27:67–74. [PubMed: 11852243]
7. Marler JR, Brott T, Broderick J, Kothari R, Odonoghue M, Barsan W, Tomsick T, Spilker J, Miller R, Sauerbeck L, et al. TISSUE-PLASMINOGEN ACTIVATOR FOR ACUTE ISCHEMIC STROKE. *New Engl J Med* 1995;333:1581–1587. [PubMed: 7477192]
8. Walsh G. Biopharmaceutical benchmarks 2006. *Nat Biotechnol* 2006;24:769–776. [PubMed: 16841057]
9. Pipe SW. The promise and challenges of bioengineered recombinant clotting factors. *J Thromb Haemost* 2005;3:1692–1701. [PubMed: 16102035]
10. Hedstrom L, Szilagyi L, Rutter WJ. Converting trypsin to chymotrypsin: the role of surface loops. *Science* 1992;255:1249–1253. [PubMed: 1546324] ••The pioneering work describing the engineering of chymotrypsin specificity into trypsin
11. Ballinger MD, Tom J, Wells JA. Furilisin: a variant of subtilisin BPN' engineered for cleaving tribasic substrates. *Biochemistry* 1996;35:13579–13585. [PubMed: 8885837]
12. Hedstrom L, Perona JJ, Rutter WJ. Converting trypsin to chymotrypsin: residue 172 is a substrate specificity determinant. *Biochemistry* 1994;33:8757–8763. [PubMed: 8038165]
13. Lim EJ, Sampath S, Coll-Rodriguez J, Schmidt J, Ray K, Rodgers DW. Swapping the substrate specificities of the neuropeptidases neurolysin and thimet oligopeptidase. *J Biol Chem* 2007;282:9722–9732. [PubMed: 17251185]
14. Villa JP, Bertenshaw GP, Bond JS. Critical amino acids in the active site of meprin metalloproteinases for substrate and peptide bond specificity. *J Biol Chem* 2003;278:42545–42550. [PubMed: 12888571]
15. Schechter I, Berger A. On the size of the active site in proteases. I. Papain. *Biochemical and Biophysical Research Communications* 1967;27:157–162. [PubMed: 6035483]
16. Graf L, Craik CS, Patthy A, Rocznik S, Fletterick RJ, Rutter WJ. Selective alteration of substrate specificity by replacement of aspartic acid-189 with lysine in the binding pocket of trypsin. *Biochemistry* 1987;26:2616–2623. [PubMed: 3111531]
17. Caputo A, Parrish JC, James MN, Powers JC, Bleackley RC. Electrostatic reversal of serine proteinase substrate specificity. *Proteins* 1999;35:415–424. [PubMed: 10382669]
18. Harris JL, Peterson EP, Hudig D, Thornberry NA, Craik CS. Definition and redesign of the extended substrate specificity of granzyme B. *J Biol Chem* 1998;273:27364–27373. [PubMed: 9765264]
19. Jäckel C, Kast P, Hilvert D. Protein design by directed evolution. *Annual review of biophysics* 2008;37:153–173.
20. Reetz MT, Carballeira JD. Iterative saturation mutagenesis (ISM) for rapid directed evolution of functional enzymes. *Nat Protoc* 2007;2:891–903. [PubMed: 17446890]
21. Bershtein S, Tawfik DS. Advances in laboratory evolution of enzymes. *Current Opinion in Chemical Biology* 2008;12:151–158. [PubMed: 18284924] •A recent review of new methodologies and screening/selection techniques for directed evolution libraries with a focus on enzymes.
22. Kaur J, Sharma R. Directed Evolution: An Approach to Engineer Enzymes. *Critical Revs in Biotechnology* 2006;26:165–199. •A review of directed evolution principles and methods utilized in enzyme engineering.
23. Fromant M, Blanquet S, Plateau P. Direct random mutagenesis of gene-sized DNA fragments using polymerase chain reaction. *Anal Biochem* 1995;224:347–353. [PubMed: 7710092]
24. Lutz S, Ostermeier M, Benkovic SJ. Rapid generation of incremental truncation libraries for protein engineering using alpha-phosphothioate nucleotides. *Nucleic Acids Research* 2001;29:E16. [PubMed: 11160936]
25. Ostermeier M, Shim JH, Benkovic SJ. A combinatorial approach to hybrid enzymes independent of DNA homology. *Nat Biotechnol* 1999;17:1205–1209. [PubMed: 10585719]
26. Bloom JD, Meyer MM, Meinhold P, Otey CR, MacMillan D, Arnold FH. Evolving strategies for enzyme engineering. *Current Opinion in Structural Biology* 2005;15:447–452. [PubMed: 16006119]

27. Varadarajan N, Cantor JR, Georgiou G, Iverson B. Construction and flow cytometric screening of targeted enzyme libraries. *Nature Protocols* 2009;4:893–901.
28. Reetz M, Kahakeaw D, Lohmer R. Addressing the numbers problem in directed evolution. *Chembiochem* 2008;9:1797–1804. [PubMed: 18567049] •The authors describe tools that can be used in the generation of more focused directed evolution libraries, helping to reduce overall library size.
29. Kadonosono T, Kato-Murai M, Ueda M. Alteration of substrate specificity of rat neurolysin from matrix metalloproteinase-2/9-type to -3-type specificity by comprehensive mutation. *Protein Eng Des Sel* 2008;21:507–513. [PubMed: 18499680]
30. Ekici OD, Karla A, Paetzel M, Lively MO, Pei D, Dalbey RE. Altered -3 substrate specificity of *Escherichia coli* signal peptidase 1 mutants as revealed by screening a combinatorial peptide library. *J Biol Chem* 2007;282:417–425. [PubMed: 17077081]
31. Karla A, Lively MO, Paetzel M, Dalbey R. The identification of residues that control signal peptidase cleavage fidelity and substrate specificity. *J Biol Chem* 2005;280:6731–6741. [PubMed: 15598653]
32. Khersonsky O, Roodveldt C, Tawfik DS. Enzyme promiscuity: evolutionary and mechanistic aspects. *Current Opinion in Chemical Biology* 2006;10:498–508. [PubMed: 16939713]
33. O'Loughlin TL, Greene DN, Matsumura I. Diversification and specialization of HIV protease function during in vitro evolution. *Mol Biol Evol* 2006;23:764–772. [PubMed: 16423863] •HIV PR protease specificity altered using a β -galactosidase inactivation assay. Inducing protease expression prior to plating initially eliminated promiscuous enzyme variants. Therefore, the resulting enzyme variants displayed reduced cytotoxicity and represented "specialist" enzymes.
34. Baum EZ, Bebernitz GA, Gluzman Y. beta-Galactosidase containing a human immunodeficiency virus protease cleavage site is cleaved and inactivated by human immunodeficiency virus protease. *Proc Natl Acad Sci USA* 1990;87:10023–10027. [PubMed: 2124694]
35. Han HE, Rho SH, Lee YJ, Park WJ. Engineering of Kex2 variants exhibiting altered substrate specificity. *Biochemical and Biophysical Research Communications* 2005;337:1102–1106. [PubMed: 16229820]
36. Sellamuthu S, Shin B, Lee E, Rho S, Hwang W, Lee Y, Han H, Kim J, Park W. Engineering of protease variants exhibiting altered substrate specificity. *Biochemical and Biophysical Research Communications* 2008;371:122–126. [PubMed: 18413229] •Yeast genetic assay for site-specific proteolysis (GASP) is used to engineer the specificity of HIV 3C protease.
37. Olsen MJ, Stephens D, Griffiths D, Daugherty P, Georgiou G, Iverson BL. Function-based isolation of novel enzymes from a large library. *Nat Biotechnol* 2000;18:1071–1074. [PubMed: 11017045]
38. McCarter JD, Stephens D, Shoemaker K, Rosenberg S, Kirsch JF, Georgiou G. Substrate specificity of the *Escherichia coli* outer membrane protease OmpT. *J Bacteriol* 2004;186:5919–5925. [PubMed: 15317797]
39. Stathopoulos C. Structural features, physiological roles, and biotechnological applications of the membrane proteases of the OmpT bacterial endopeptidase family: a micro-review. *Membrane & cell biology* 1998;12:1–8. [PubMed: 9829254]
40. Stumpe S, Schmid R, Stephens DL, Georgiou G, Bakker EP. Identification of OmpT as the protease that hydrolyzes the antimicrobial peptide protamine before it enters growing cells of *Escherichia coli*. *J Bacteriol* 1998;180:4002–4006. [PubMed: 9683502]
41. Guina T, Yi EC, Wang H, Hackett M, Miller SI. A PhoP-regulated outer membrane protease of *Salmonella enterica* serovar typhimurium promotes resistance to alpha-helical antimicrobial peptides. *J Bacteriol* 2000;182:4077–4086. [PubMed: 10869088]
42. Varadarajan N, Gam J, Olsen MJ, Georgiou G, Iverson B. Engineering of protease variants exhibiting high catalytic activity and exquisite substrate selectivity. *Proc Natl Acad Sci USA* 2005;102:6855–6860. [PubMed: 15867160] ••Introduction of a novel method employing simultaneous selection, positive pressure for a target substrate, and counter-selection, negative pressure to reduce parentallike activity, in the engineering of a protease specificity.
43. Varadarajan N, Rodriguez S, Hwang BY, Georgiou G, Iverson B. Highly active and selective endopeptidases with programmed substrate specificities. *Nat Chem Biol* 2008;4:290–294. [PubMed: 18391948] ••A high-throughput FACS based assay employing simultaneous selection and counter-selection was used in the engineering of OmpT protease specificity. A number of highly specific and

selective variants were isolated with altered specificities spanning the entire range of amino acid chemical properties.

44. Gosalia DN, Salisbury CM, Ellman JA, Diamond SL. High throughput substrate specificity profiling of serine and cysteine proteases using solution-phase fluorogenic peptide microarrays. *Mol Cell Proteomics* 2005;4:626–636. [PubMed: 15705970]
45. Harris JL, Backes BJ, Leonetti F, Mahrus S, Ellman JA, Craik CS. Rapid and general profiling of protease specificity by using combinatorial fluorogenic substrate libraries. *Proc Natl Acad Sci USA* 2000;97:7754–7759. [PubMed: 10869434]
46. Knight ZA, Garrison JL, Chan K, King DS, Shokat KM. A remodelled protease that cleaves phosphotyrosine substrates. *J Am Chem Soc* 2007;129:11672–11673. [PubMed: 17803306] •
•Rational mutagenesis of Subtilisin BPN' resulted in the first example of a protease engineered to be specific for a post-translationally modified tyrosine.
47. Varadarajan N, Georgiou G, Iverson BL. An Engineered Protease that Cleaves Specifically after Sulfated Tyrosine. *Angew Chem Int Ed Engl.* 2008 • •An OmpT variant specific for sulfotyrosine was isolated from a large saturation mutagenesis library using a FACS assay, representing the second protease to be engineered for a post-translationally modified amino acid, and the first protease specific for sulfotyrosine.
48. Coombs GS, Bergstrom RC, Madison EL, Corey DR. Directing sequence-specific proteolysis to new targets. The influence of loop size and target sequence on selective proteolysis by tissue-type plasminogen activator and urokinase-type plasminogen activator. *J Biol Chem* 1998;273:4323–4328. [PubMed: 9468480]
49. Tyndall JD, Nall T, Fairlie DP. Proteases universally recognize beta strands in their active sites. *Chem Rev* 2005;105:973–999. [PubMed: 15755082]

Figure 1. Schematic summary of protease engineering. Including, but not limited to, some processes and techniques used to perform laboratory-directed evolution.

Figure 2.

Overview of the recently described *in vivo* assays utilized to screen large protease libraries for variants possessing altered substrate specificity. (A) β -galactosidase inactivation assay utilized in engineering HIV PR [33]. (B) Schematic principal of GASP, a cytoplasmic based assay used in the engineering of HAV 3CP protease specificity [36]. (C) Screening methodology of the two-color FACS assay used for the isolation of OmpT variants with altered specificity [42, 43,47].

Table 1

Enzyme	Mutations	Cleavage Site	$k_{cat}/K_M (M^{-1}s^{-1})$	Fold switch in specificity	Ref
Trypsin Tr Ch[S1-L1+L2+Y172W]	— H18T, Y172W, L185A, E186S, D189S, Q192M, Y217S, 219T, A221S, P222T, D223S, N224T	LvsL ₁ AMC Phe/Tvs/TrpL ₁ AMC	$1.9 \pm 0.1 \times 10^5$ 10^5	10^8	[12]
Rat Brain Neurolysin	— Y610L	RPKP↓YAN RPKP↓VEN	2.9×10^4 1.9×10^3	— 10	[29]
SPase	— I86T, I144C	AXA↓AT XXA↓AT	Not Reported Not Reported	— *	[30]
HIV PR	— P9S, I50L	VSFNF↓PQITL NRPDYLLFAE**	Not Reported Not Reported	— ≥ 2	[33]
HAV 3CP	— M29V, H145C, K146W, K147G, L155W	ELRTQ↓SFSN ELRQQ↓SFSN	Not Reported Not Reported	— 160	[36]
Var7 OmpT	— S223R	Arg↓Arg Ala↓Arg	$1.7 \pm 0.4 \times 10^5$ $> 1.8 \times 10^5$	— $> 10^7$	[47] [42]
AR-OmpT	— E27L, D208R, S223G, L183F	Glu↓Arg	$3 \pm 2 \times 10^5$	$\geq 10^6$	
ER-OmpT	— Y221A, L265V	Tyr↓Arg	$8 \pm 0.5 \times 10^4$	$\geq 10^6$	[43]
YR-OmpT	— E27H, V29S, D208L, D214N, S223D, P243S, W253G, N270Y, S276G	Thr↓Arg	$2 \pm 1 \times 10^4$	$> 10^6$	
TR2-OmpT	— D97H, S223D, Q63R	Arg↓Val Glu↓Ala	$5 \pm 2 \times 10^5$ $1 \pm 0.3 \times 10^5$	Not Reported Not Reported	
RV-OmpT	— E27L, N48D, L183F, D208R, D214N, G216E, S223G, N244I, D274G, A280E	sTyr↓Arg	$1.1 \pm 0.2 \times 10^5$	$> 10^6$	[47]
EA-OmpT	— E27F, V29A, M87R, Y126C, E153D, I170V, D208R, Y221A, I282H	Phe/Tyr↓Gly pTyr↓Gly	3×10^6 1.41×10^7	— 2500	[46]
sT4-OmpT	— E156R, P129G				
Subtilisin in BPN					

* In the case of SPase engineering, the specificity was relaxed at P3 and not altered for a specific substrate.

** The precise cleavage site of HIV PR PSS/I50L was not identified