Published in final edited form as:

Chembiochem. 2012 March 19; 13(5): 649–653. doi:10.1002/cbic.201100718.

Directed Evolution of Highly Selective Proteases Using a Novel FACS Based Screen that Capitalizes on the p53 Regulator MDM2

Tae Hyeon Yoo^{a,*}, Mark Pogson^{b,*}, Brent L. Iverson^{b,c}, and George Georgiou^{b,d,**}

^aDepartment of Molecular Science and Technology and Division of Applied Chemistry and Biological Engineering, Ajou University, Suwon, Korea

^bInstitute of Cellular and Molecular Biology, The University of Texas at Austin, 1 University Station A4800, Austin, TX 78712

^cDepartment of Chemistry and Biochemistry, The University of Texas at Austin, 1 University Station A5300, Austin, TX 78712

^dDepartment of Chemical Engineering, Department of Biomedical Engineering, and Section of Molecular Genetics and Microbiology, The University of Texas at Austin, 1 University Station C0400, Austin, TX 78712

Proteases could be very attractive as protein therapeutics that irreversibly deactivate proteins associated with disease states, provided that they could be engineered to cleave disease target proteins with specificities sufficient to avoid any unwanted side effects. Rational design and directed evolution approaches have been employed with various degrees of success for the engineering of protease specificity.^{[1] [2]} These studies have revealed two major bottlenecks in the pursuit of proteases engineered for therapeutic applications.

First, the generation of proteases exhibiting novel substrate specificity often amounts to relaxed selectivity rather than a true change in specificity.^[3] In order to combat this problem, our laboratory has developed directed evolution strategies for protease engineering using one or more counter-selection substrates.^[4] Using the counter-selection approach, we have been able to isolate variants of the *E. coli* outer membrane endopeptidase OmpT, which normally exhibits a strong preference for the hydrolysis of substrates that contain a basic amino acid (Lys or Arg) at the P1 and P1' positions (Schechter and Berger nomenclature),^[5] that instead, can selectively cleave peptides containing a small hydrophobic, polar, aromatic or even negatively charged residues in P1. We have even isolated OmpT variants that preferentially cleave a completely different sequence in both P1 and P1' such as Glu-Ala.^[4a, 4d] Additionally we were able to isolate OmpT variants engineered to cleave following post-translationally modified tyrosine residues, and one of them showed more than a 100-fold catalytic discrimination towards nitrotyrosine relative to peptides containing unmodified Tyr or other post-translational modifications (e.g. sulfotyrosine or phosphotyrosine).^[4b, 4c]

The second bottleneck in the engineering of therapeutic proteases is that enzymes designed or evolved to cleave peptides with high selectivities and catalytic activities may fail to recognize the same amino acid sequence in the context of a folded protein. Most proteases recognize their preferred sequences in an extended (β -sheet-like) conformation.^[6] As a result, a preferred amino acid sequence that as a free peptide is a suitable substrate for a protease, may not be cleaved in the context of an inaccessible or structurally constrained

^{**}Correspondence should be addressed to G.G. (gg@che.utexas.edu).

^{*}These authors contributed equally to this work.

portion of secondary structure within a protein.^[7] The inability of proteases to cleave proteins containing putative substrate peptide sequences within loops or sites that would be expected to be accessible based on crystallographic data, is well documented.^[8] Likewise we have found that OmpT variants engineered to cleave efficiently between Glu↓Ala in a peptide substrate^[4d] failed to digest the wild-type OmpT protein substrate T7 RNA polymerase,^[9] when the wild-type preferred dibasic OmpT cleavage site was substituted with Glu-Ala (MP, unpublished results). One way to overcome this problem is to devise high throughput screens that require the cleavage of a desired protein substrate. For example, in a pioneering study, Matsumura and coworkers^[2b] developed a screen for the selection of HIV protease variants expressed in *E. coli* that were both non-toxic to the host and had been evolved to cleave a peptide sequence introduced into an accessible loop of β -galactosidase.^[10] However, this strategy and other *in vivo* screens that exploit the cleavage of a substrate within the host^[2c] are applicable only to protein substrates that can be readily detected in colonies by virtue of their enzymatic activity. Furthermore, *in vivo* screens do not provide control over the kinetic properties of the selected enzymes, nor are they particularly effective in eliminating the cleavage of wild-type substrates or other undesired sequences.^[11]

As a first step towards the development of next generation screening systems for therapeutic protease engineering we have developed a novel single cell assay that exploits cell surface capture of exogenous protein substrates (Figure 1). This assay capitalizes on the p53 antagonist MDM2 as a detector of protease activity in addition to its utility as a counter-selection substrate. The tumor suppressor p53 functions to preserve genomic integrity in normally dividing cells.^[12] It's inactivation is implicated in >50% of cancers, including but not limited to, cancers of the colon, breast, lung, pancreas and brain.^[13] The p53 protein is negatively regulated by MDM2, which contains, among others, a p53 binding domain^[14] as well as an E3 ubiquitin ligase domain, responsible for marking p53 for proteosomal destruction.^[15] Similar to p53, the aberrant regulation of MDM2 has also been implicated in many cancers. The *N*-terminal domain of MDM2 binds directly to the *N*-terminal transactivation domain of p53.^[14a, 14c] More specifically, studies have shown that MDM2 recognizes an unstructured peptide with the sequence E₁₇TFS₂₉DLWKLLPEN₂₉ (**p53-TAD**, Figure 1E) located within the transactivating domain of p53 with an affinity in the 0.5–1 μ M range.^[16] Structural evidence suggests that p53-TAD binds to MDM2 in a helical conformation and that the *C*-terminus of the p53-TAD peptide is in close proximity to the *N*-terminus of MDM2 (Figure 1C).^[14a, 16] In addition to the p53-TAD peptide, several p53-TAD variants and other peptides capable of associating with MDM2 have been identified through peptide screens. One such peptide, designated **PMI**, exhibits a dissociation constant of 3nM (Figure 1E).^[16b]

As shown in Figure 1D, our protease screening platform, is built around MDM2 fused to a *C*-terminal poly-Arg tail and an *N*-terminal extension comprised of a putative proteolysis target cleavage sequence joined to a flexible linker and the p53-TAD sequence, designated construct **1**.^[17] Because of the facile intramolecular interaction between MDM2 and the p53-TAD peptide, the MDM2 binding site is primarily occupied. Thus, construct **1** displays a low effective binding affinity *in trans*, following incubation with an excess of **PMI** peptide conjugated to the fluorescent dye BODIPY (**PMI-BD**, Figure 1E). The *C*-terminal poly-Arginine sequence of construct **1** ensures strong association via electrostatic interactions with the negatively-charged surface of *E. coli* or of other microbial hosts, as desired. Cleavage within the protease substrate domain of construct **1** on the *E. coli* surface by expressed OmpT variants results in the release of the **p53-TAD** peptide, and the unliganded MDM2 on the cell surface can now bind **PMI-BD**, thus rendering cells fluorescent with an overall intensity that correlates with OmpT catalytic turnover.

Construct **1** and a control, construct **2**, composed of only MDM2 with a C-terminal poly-Arg fusion (Figure 1F) were expressed in *E. coli* and purified, and their apparent binding to **PMI-BD** on cell surfaces was determined. As expected, when *E. coli* BL21(DE3), an *ompT* deficient strain, is incubated with the control construct **2** and **PMI-BD**, the cells are rendered fluorescent (Figure 2A), but when incubated with the same concentration of construct **1** and **PMI-BD**, there is a >10-fold reduction in detectable fluorescence (Figure 2B).

Importantly, construct **1** includes a built-in counter-selection system. Any OmpT variants that display either the wild-type or greatly relaxed specificity are expected to give a significantly diminished fluorescent signal. For example, wild-type OmpT cleaves the poly-Arg tail, precluding association with the *E. coli* outer membrane (Figure S2). Thus, any variants retaining the wild-type OmpT specificity leading to cleavage of the polyArg tail or, alternatively, displaying relaxed selectivity leading to cleavage of the MDM2 domain are also expected to lead to weak fluorescence signal.

The OmpT-S223R variant displays a dramatic 1.8×10^5 -fold change in catalytic selectivity (k_{cat}/K_m ratio) with Ala-Arg in the P1-P1' positions relative to Arg-Arg containing substrates, strongly preferred by the wild-type enzyme. [4d] Substrate phage analysis of the amino acid selectivity of OmpT-S223R also confirmed that it has a strong preference for Ala in the P1 position. [4d] However, surprisingly, when *E. coli* cells expressing the OmpT-S223R variant were incubated with either construct **2** (Figure 2C) or construct **1** (Figure S3) followed by **PMI-BD**, the cells exhibited low fluorescence, suggesting that OmpT-S223R cleaves either the MDM2 unit or the poly-Arg tail. SDS-PAGE analysis of construct **2** incubated with purified OmpT-S223R confirmed that OmpT-S223R indeed cleaves construct **2**, (Figure 2D, Lane 2) and mass spec analysis revealed cleavage at Leu₃₀↓Lys₃₁ in the MDM2 unit (Figure S4). Thus, the OmpT-S223R variant, which was isolated in a screen based upon the preferential cleavage of one selection peptide in the presence of a single counter selection peptide, can cleave MDM2 at a P1-P1' sequence not found in the counter selection peptide. In the MDM2 auto-inhibition assay described above the entire MDM2 domain as well as the poly-Arg tail serve effectively as counter-selection substrates, significantly increasing counter-selection stringency compared with screens using short peptides for counter selection.

To evaluate the utility of the auto-inhibited MDM2 system for highly specific protease engineering, an OmpT library, partially saturated across 21 positions, [4d] was incubated with construct **1** where the substrate sequence between the MDM2 and the autoinhibitory domains contains an Ala-Arg sequence (Figure 1D). In this protein the Ala-Arg dipeptide is found only within the linker region that constitutes the putative substrate, and is not present in MDM2. The cells were then labeled with **PMI-BD**, and sorted for high fluorescence. Following five rounds of sorting, the resulting cell pool exhibited around 20-fold higher fluorescence compared to background (Figure 3A). Colonies from the last round of sorting were grown individually, some of which exhibited high fluorescence as monoclonal populations (Figure 3B). SDS-PAGE analysis further demonstrated cleavage of construct **1** only at, or near, the expected Ala-Arg site (Figure 3C, marked in red). DNA sequencing revealed these clones encoded an identical OmpT variant (Figure S5), referred to as OmpT-AR1.

The exact cleavage site of construct **1** by the OmpT-AR1 protease was identified using mass spectrometry. Two major fragments were detected (Figure 4A), the masses of which corresponded with the expected fragments resulting from cleavage only at the desired Ala↓Arg site (Figure S6). Reaction kinetics were analyzed by HPLC, revealing a k_{cat}/K_m of $2.0 \times 10^2 \text{ M}^{-1} \text{ s}^{-1}$ (Figure 4B) using the entire fusion protein as substrate.

In summary we have developed a novel high-throughput method for the isolation of selective protease variants displayed on the surface of *E. coli*.^[18] Our method relies on an auto-inhibited protein fusion electrostatically anchored onto the cell surface using a poly-Arg tail, combined with external labeling with a fluorescently tagged peptide. Although we used the MDM2 **p53-TAD** pair to effect auto-inhibition, other such pairs might be used as well. The peptide or protein sequence to be cleaved is inserted between MDM2 and the **p53-TAD** sequences. Cleavage of the inserted site but not MDM2 or the poly-Arg tail results in high cell fluorescence, the intensity of which is dependent on protease catalytic turnover. Using this method, we succeeded in isolating an OmpT variant that cleaves between Ala-Arg, but nowhere else on the autoinhibited MDM2-substrate-fusion construct, unlike similar OmpT variants isolated using short peptide selection and counter selection substrates.

The auto-inhibited protein substrate design used in this study can be modified to contain different protease target sequences, including the potential for the insertion of entire protein domains. Additionally, the protein substrate has both a selection sequence and a wild-type counter-selection sequence, and can be engineered to include multiple counter-selection sequences that can be introduced between the MDM2 domain and the poly-Arg tail. The advantages of using multiple counter-selection sequences have been previously reported^[3a, 4c] and our own results support these prior conclusions. We are currently employing this method to isolate putative therapeutic protease candidates for the treatment of acute phase diseases. We envision that such enzymes could be deployed for a one-time treatment in which case adverse immune reactions due to the non-human nature of OmpT would not be expected to be an issue.

Methods

PMI-BD conjugate

The **PMI** peptide with an additional Glycine at the N-terminus (GTSFAEYWNLSP) was synthesized with >98% purity at ABGENT (San Diego, California). The peptide was conjugated to BODIPY FL-SE, from Invitrogen (Carlsbad, CA) using the N-terminal amine, and the peptide-dye conjugate was purified using reverse phase HPLC as previously described.^[4d]

Plasmid construction

An assembled gene for MDM2(25–125) plus the 10 Arg sequence was cloned into the pET-28a plasmid using the BamHI and XhoI sites. Then, a synthetic double strand oligonucleotide encoding the **p53-TAD** peptide (ETFSDLWKLLPEN) and the linker (GGGSGSARVGGGS) was further cloned into this pET-28 plasmid encoding the MDM2 gene using NcoI and BamHI sites (construct **1**). A synthetic double strand for the FLAG tag (DYKDDDDK) and assembled gene for MDM2(17–125) plus the 10 Arg sequence were cloned into pET28a using NcoI/BamHI and BamHI/XhoI respectively (construct **2**).

MDM2 expression and purification

Proteins with the C-terminal six Histidine tag were purified under native conditions by affinity chromatography using Ni-NTA resin (Qiagen) according to the manufacturer's protocols with some modifications. In order to prevent interactions between the positively charged MDM2 proteins and negatively charged molecules such as DNA and RNA, we increased the NaCl concentration of the lysis and washing buffers from 300mM to 1M. However, the elution buffer contains 300 mM NaCl. The purified proteins were purified further by gel filtration on Superdex G75 (GE Healthcare).

Library screening

E. coli BL21(DE3) cells harboring the 90% saturated-21 position NNS library^[4d] were inoculated into 2xYT supplemented with 100 mg/L ampicillin, and the cells were grown to an OD₆₀₀ of 2.0 at 37°C. A 1 mL aliquot of the culture was centrifuged at 12000 g for 3 min. The cell pellet was washed with 1 mL of 1% sucrose two times and resuspended in 1 mL of 1% sucrose. 50 uL of the resuspended cells and 5 uL of 10 uM construct **1** were added to 445 uL of 1% sucrose, and then the reaction mixture was incubated at 25 °C for 30 min. A 200 uL aliquot of the reaction was mixed with 4 mL of 1% sucrose including 40 nM **PMI-BD** conjugate. Library sorting was performed on the FACSaria using gates set based upon FSC/SSC and Green fluorescence channels. 5×10^7 cells were screened, and 1 % of the most fluorescent cells were collected and plated on 2xYT agar plates (100 mg/L Ampicillin). The enriched cells were subjected to the next round of sorting.

Enzyme purification and kinetic analysis

OmpT-AR1 was isolated as previously described.^[4d] For kinetic analyses, 125 nM of the purified enzyme was incubated with 1–20 uM construct **1** in 10 mM MES (pH 6.1) containing 10 mM EDTA and 150 mM NaCl at 25°C for 1 hour. The reaction mixtures were subjected to separation on a C18 reverse phase HPLC column. The amount of the **p53 peptide** (MGETFSDLWKLLPENGSGSGSA) in reaction solutions was determined using Tryptophan as an internal standard.

Acknowledgments

We thank Dr. Hsien-Hsiang Stony Lo in the University of Texas ICMB/CRED Protein and Metabolite Analysis Facility for the ESI and MALDI-TOF/TOF mass spectrometric analyses of protein and peptide samples, and Drs. E. Stone and M. Gebhard for suggestions regarding the manuscript. This research was conducted in whole or in part by Clayton Foundation for Research (BLI) and supported by US National Institutes of Health grants R01 GM065551 and R01 GM073089 (BLI and GG).

References

1. a) Ballinger MD, Tom J, Wells JA. *Biochemistry*. 1996; 35:13579–13585. [PubMed: 8885837] b) Hedstrom L, Perona JJ, Rutter WJ. *Biochemistry*. 1994; 33:8757–8763. [PubMed: 8038165] c) Hedstrom L, Szilagyi L, Rutter WJ. *Science*. 1992; 255:1249–1253. [PubMed: 1546324] d) Kurth T, Ullmann D, Jakubke HD, Hedstrom L. *Biochemistry*. 1997; 36:10098–10104. [PubMed: 9254605] e) Lim EJ, Sampath S, Coll-Rodriguez J, Schmidt J, Ray K, Rodgers DW. *J Biol Chem*. 2007; 282:9722–9732. [PubMed: 17251185]
2. a) Olsen MJ, Stephens D, Griffiths D, Daugherty P, Georgiou G, Iverson BL. *Nat Biotechnol*. 2000; 18:1071–1074. [PubMed: 11017045] b) O'Loughlin TL, Greene DN, Matsumura I. *Mol Biol Evol*. 2006; 23:764–772. [PubMed: 16423863] c) Sellamuthu S, Shin B, Lee E, Rho S, Hwang W, Lee Y, Han H, Kim J, Park W. *Biochemical and Biophysical Research Communications*. 2008; 371:122–126. [PubMed: 18413229] d) Han H-E, Rho S-H, Lee YJ, Park WJ. *Biochemical and Biophysical Research Communications*. 2005; 337:1102–1106. [PubMed: 16229820] e) Kostallas G, Samuelson P. *Applied and Environmental Microbiology*. 2010; 76:7500–7508. [PubMed: 20851955]
3. a) Khersonsky O, Roodveldt C, Tawfik DS. *Current Opinion in Chemical Biology*. 2006; 10:498–508. [PubMed: 16939713] b) Aharoni A, Gaidukov L, Khersonsky O, McQ Gould S, Roodveldt C, Tawfik DS. *Nat Genet*. 2005; 37:73–76. [PubMed: 15568024] c) Matsumura I, Ellington AD. *Journal of Molecular Biology*. 2001; 305:331–339. [PubMed: 11124909]
4. a) Varadarajan JN, Gam J, Olsen MJ, Georgiou G, Iverson BL. *Proc Natl Acad Sci U S A*. 2005; 102:6855–6860. [PubMed: 15867160] b) Varadarajan N, Georgiou G, Iverson BL. *Angew Chem Int Ed Engl*. 2008; 47:7861–7863. [PubMed: 18780393] c) Varadarajan N, Pogson M, Georgiou G, Iverson BL. *J Am Chem Soc*. 2009; 131:18186–18190. [PubMed: 19924991] d) Varadarajan N, Rodriguez S, Hwang BY, Georgiou G, Iverson BL. *Nat Chem Biol*. 2008; 4:290–294. [PubMed: 18391948]

5. Schechter I, Berger A. Biochemical and Biophysical Research Communications. 1967; 27:157–162. [PubMed: 6035483]
6. Tyndall JDA, Nall T, Fairlie DP. Chem Rev. 2005; 105:973–999. [PubMed: 15755082]
7. Timmer J, Zhu W, Pop C, Regan T, Snipas S, Eroshkin A, Riedl S, Salvesen G. Nat Struct Mol Biol. 2009
8. Mahrus S, Trinidad J, Barkan D, Sali A, Burlingame A, Wells J. Cell. 2008
9. Grodberg J, Dunn JJ. J Bacteriol. 1988; 170:1245–1253. [PubMed: 3277950]
10. Baum EZ, Beberitz GA, Gluzman Y. Proc Natl Acad Sci USA. 1990; 87:10023–10027. [PubMed: 2124694]
11. Rowe LA, Geddie ML, Alexander OB, Matsumura I. Journal of Molecular Biology. 2003; 332:851–860. [PubMed: 12972256]
12. Vogelstein B, Lane D, Levine AJ. Nature. 2000; 408:307–310. [PubMed: 11099028]
13. Sherr CJ. Cell. 2004; 116:235–246. [PubMed: 14744434]
14. a) Kussie PH, Gorina S, Marechal V, Elenbaas B, Moreau J, Levine AJ, Pavletich NP. Science. 1996; 274:948–953. [PubMed: 8875929] b) Oliner JD, Kinzler KW, Meltzer PS, George DL, Vogelstein B. Nature. 1992; 358:80–83. [PubMed: 1614537] c) Oliner JD, Pietenpol JA, Thiagalingam S, Gyuris J, Kinzler KW, Vogelstein B. Nature. 1993; 362:857–860. [PubMed: 8479525]
15. a) Haupt Y, Maya R, Kazaz A, Oren M. Nature. 1997; 387:296–299. [PubMed: 9153395] b) Honda R, Tanaka H, Yasuda H. FEBS Lett. 1997; 420:25–27. [PubMed: 9450543]
16. a) Chi SW, Lee SH, Kim DH, Ahn MJ, Kim JS, Woo JY, Torizawa T, Kainosho M, Han KH. J Biol Chem. 2005; 280:38795–38802. [PubMed: 16159876] b) Pazgier M, Liu M, Zou G, Yuan W, Li C, Li J, Monbo J, Zella D, Tarasov SG, Lu W. Proc Natl Acad Sci U S A. 2009; 106:4665–4670. [PubMed: 19255450]
17. Liu Z, Olejniczak ET, Fesik SW. Protein Expr Purif. 2004; 37:493–498. [PubMed: 15358376]
18. a) Francisco JA, Earhart CF, Georgiou G. Proc Natl Acad Sci USA. 1992; 89:2713–2717. [PubMed: 1557377] b) Georgiou G, Stathopoulos C, Daugherty PS, Nayak AR, Iverson BL, Curtiss R. Nat Biotechnol. 1997; 15:29–34. [PubMed: 9035102]

Figure 1.

Cyto metric analysis of OmpT catalytic activity. The activity of OmpT can be analyzed using two components: autoinhibited protein and peptide-dye conjugate. **A)** The positively charged autoinhibited protein associates with negatively charged bacterial cell surfaces. However, the protein has very low binding affinity to the peptide-dye conjugate, and the cells are not fluorescent. **B)** When cells express a specific OmpT variant, the autoinhibition is removed, and thus the cells become fluorescent. **C)** Structure of the N-terminal domain of MDM2 (green) and the p53 peptide (pink). **D)** Engineering of an autoinhibited MDM2 that has a polyarginine tail (construct **1**). **E)** Sequences of peptides and substrates used for autoinhibition and fluorescent labeling. **F)** Schematic representation of autoinhibited-substrate-MDM2-polyArg (construct **1**) and MDM2-polyArg (construct **2**).

Figure 2.

Flow cytometric analysis of *E. coli* BL21(DE3) incubated with **A)** PMI-BD and construct **2** or **B)** construct **1**. **C)** Flow cytometric analysis of *E. coli* BL21(DE3) expressing OmpT-S223R incubated with PMI-BD and construct **2**. **D)** Cleavage of construct **2** (Lane 1) by OmpT-S223R (Lane 2) or OmpT-AR1 (Lane 3): 10 μM construct **2** was incubated with 250 nM OmpT-S223R or OmpT-AR1 for two hours at room temperature.

Figure 3.

Screening of OmpT for cleavage of the AR site in construct 1. **A)** Fluorescence histogram of cells after 5 rounds of sorting. **B)** Fluorescence of cells expressing OmpT-AR1 incubated with PMI-BD and construct 1. **C)** SDS-PAGE analysis of individual clones expressing OmpT variants after incubating whole cells with construct 1.

Figure 4. Characterization of OmpT-AR1. **A)** Mass spectrometric analysis of cleaved products of construct **1** by the purified OmpT-AR1. **B)** Kinetic trace for the hydrolysis of construct **1** between Arginine and Alanine.