Engineering of protease variants exhibiting high catalytic activity and exquisite substrate selectivity

Navin Varadarajan*†‡, Jongsik Gam†‡, Mark J. Olsen*, George Georgiou*§¶, and Brent L. Iverson*†‖

*Institute for Cellular and Molecular Biology and Departments of †Chemistry and Biochemistry and §Chemical Engineering, University of Texas, Austin, TX 78712

Edited by Robert T. Sauer, Massachusetts Institute of Technology, Cambridge, MA, and approved April 1, 2005 (received for review January 4, 2005)

The exquisite selectivity and catalytic activity of enzymes have been shaped by the effects of positive and negative selection pressure during the course of evolution. In contrast, enzyme variants engineered by using *in vitro* screening techniques to accept novel substrates typically display a higher degree of catalytic promiscuity and lower total turnover in comparison with their natural counterparts. Using bacterial display and multiparameter flow cytometry, we have developed a novel methodology for emulating positive and negative selective pressure *in vitro* for the isolation of enzyme variants with reactivity for desired novel substrates, while simultaneously excluding those with reactivity toward undesired substrates. Screening of a large library of random mutants of the *Escherichia coli* endopeptidase OmpT led to the isolation of an enzyme variant, 1.3.19, that cleaved an Ala–Arg peptide bond instead of the Arg–Arg bond preferred by the WT enzyme. Variant 1.3.19 exhibited greater than three million-fold selectivity (-Ala–Arg-/-Arg–Arg-) and a catalytic efficiency for Ala–Arg cleavage that is the same as that displayed by the parent for the preferred substrate, Arg–Arg. A single amino acid Ser223Arg substitution was shown to recapitulate completely the unique catalytic properties of the 1.3.19 variant. These results can be explained by proposing that this mutation acts to ''swap'' the P_1 Arg side chain normally found in WT substrate peptides with the 223Arg side chain in the S_1 subsite of OmpT.

engineering | flow cytometry

The reprogramming of enzyme catalytic activity and selectivity is a central issue in protein biochemistry and biotechnology. Numerous structure-guided and directed evolution strategies have been used in search of enzyme variants that exhibit high catalytic rates with poor or inactive substrates of the parental enzyme (1–19). As impressive as these successes have been, the engineering of enzymes that exhibit turnover rates and selectivities with new substrates comparable to their natural counterparts has proven quite a challenge, especially when considering those enzymes for which a genetic selection strategy is not possible.

In particular, enzymes engineered through laboratory evolution involving *in vitro* catalytic assays have often been found lacking, either with respect to turnover rates or selectivity, relative to catalyst–substrate pairs isolated from natural sources. As a typical example, an extensive directed evolution program led to the isolation of *Escherichia coli* β-glucuronidase variants with significant β-galactosidase (10) or xylanosidase (11) activities, but nonetheless even the best clones exhibited k_{cat}/K_m values >1,000 times lower than those of naturally occurring enzymes such as the *E. coli* β-galactosidase or the *Thermoanaerobacterium saccharolyticum* β-xylosidase.

This trend appears to be general. In a recent comprehensive study, Aaron *et al.* (12) demonstrated that the evolution of higher activity toward poor substrates did not impair the parental catalytic activity, and, therefore, the evolved enzymes exhibited greater promiscuity. Enzymes evolved for higher substrate enantioselectivity often exhibit lower specific activities toward their new substrates relative to their respective parental enzymes (13–15). Similarly, the evolution of highly active variants of aspartate aminotransferases capable of accepting branched or aromatic amino acid substrates was accompanied by a relaxation of the substrate selectivity (16, 17).

In nature, the evolution of enzymes occurs as the result of positive selective pressure for turnover of physiological substrates, combined with simultaneous negative selective pressure to eliminate completely, or at least drastically suppress, deleterious activities. It follows that the engineering of enzymes exhibiting high catalytic activity and substrate selectivity for a particular desired substrate should be similarly accomplished by implementing selection and counterselection assay schemes in the laboratory. The recent directed evolution of novel tRNA synthetases and recombinases by capitalizing on *in vivo* selections support this expectation (18–19). Unfortunately, many desired enzyme activities are not amenable to *in vivo* selection strategies because cellular growth cannot be linked to the enzyme activity being sought. Consequently, laboratory-directed evolution approaches for the isolation of highly selective enzymes must rely on *in vitro* catalytic assays whereby the activity toward selection and counterselection substrates is determined sequentially for each library member (9–11, 13, 14). The successful implementation of the latter approach hinges on satisfying the following requirements. First, assays that afford the proper dynamic range must be designed because mutant proteins are likely to exhibit drastically different k_{cat} and K_m values for the various selection and counterselection substrates. Second, the isolation of rare, high-activity clones requires the screening of large libraries by assaying each clone toward multiple substrates, thus necessitating the availability of suitable high-throughput methods.

With these considerations in mind, we have extended our previous approach (20) and developed a new two-pronged strategy in which catalytic activities over a wide dynamic range for both a selection substrate and one or more counterselection substrates are quantified *simultaneously* at the single-cell level, enabling the rapid screening of mutant libraries. This approach relies on fluorescent substrates of different colors that label the surface of *E. coli* cells upon cleavage by a surface-anchored enzyme (20). Enzymes can be displayed on the surface of Gram-negative bacteria by established techniques (21, 22), and thus, access to the fluorescent substrate is assured. The net result is that the cell fluorescence profile accurately reflects the catalytic activity and selectivity of the surface-displayed enzyme. Multiparameter flow cytometry is then used to isolate clones

This paper was submitted directly (Track II) to the PNAS office.

Abbreviation: BODIPY, 4,4-difluoro-5,7-dimethyl-4-bora-3a,4a-diaza-*s*-indacene-3-propionic acid.

‡N.V. and J.G. contributed equally to this work.

¶To whom correspondence may be addressed at: Department of Chemical Engineering, University of Texas, Austin, TX 78712. E-mail: gg@che.utexas.edu.

‖To whom correspondence may be addressed at: Department of Chemistry and Biochemistry, University of Texas, Campus Mail Code A5300, Austin, TX 78712. E-mail: biverson@mail.utexas.edu.

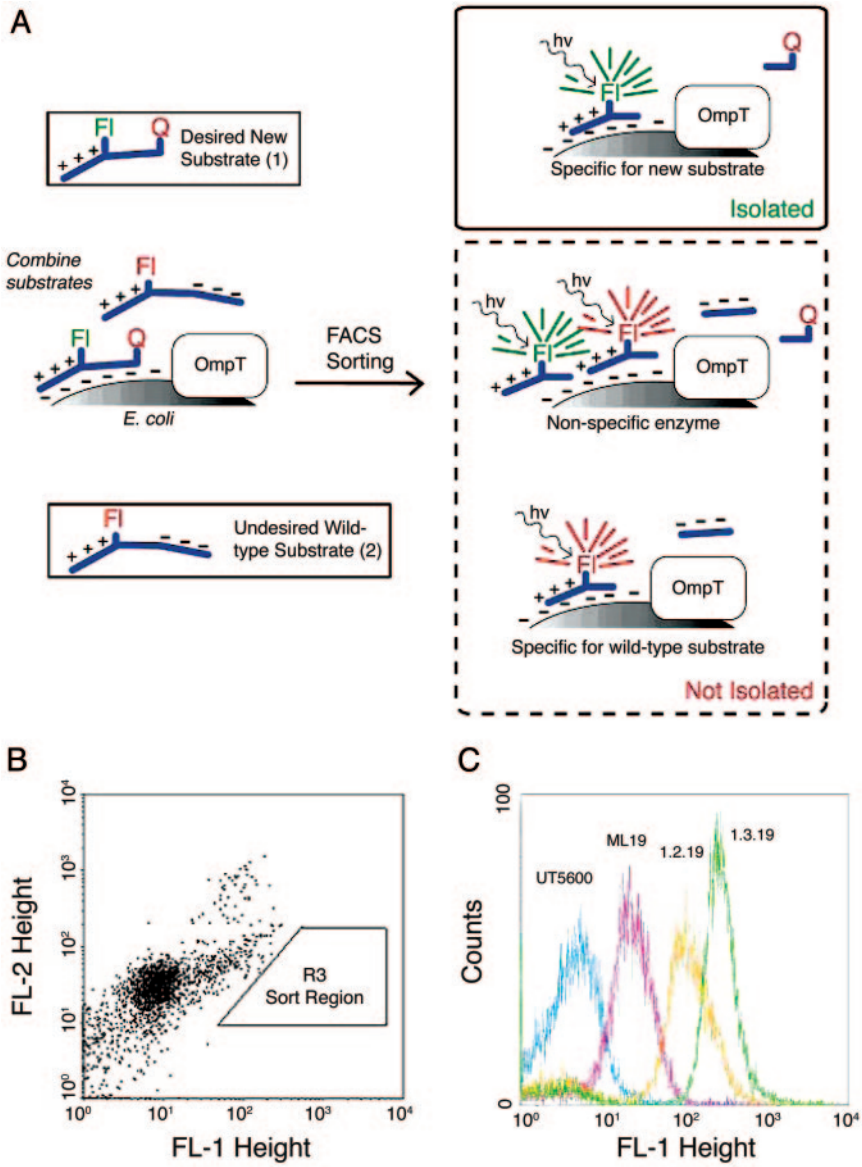

Fig. 1. Flow-cytometric schemes and assays described in this study. (*A*) Two-color discrimination and selection by using the FRET and electrostatic capture substrates. (*B*) Library sort gate R3, used to isolate positive clones displaying high FL-1 (from the hydrolysis of 1) and low FL-2 (lack of hydrolysis of 2) fluorescence. (*C*) Flow-cytometric discrimination of *E. coli* UT5600 (Δ*ompT*) transformed with pML19 (expressing WT OmpT), pML1.2.19, and pML1.3.19 by using FRET substrate **1**. Briefly, the cells were washed and resuspended in 1% sucrose and labeled with 50 nM (final concentration) of the AR-FRET substrate **1**. A 20-μl aliquot of the labeling reaction was transferred to 0.5 ml of 1% sucrose and analyzed on the flow cytometer.

expressing enzymes having a desired fluorescence profile from large libraries.

The *E. coli* endoprotease OmpT and its Omptin homologues play important roles in pathogenicity, are of significance in protein manufacturing, and have been exploited for biotechnology applications (23, 24). OmpT has a strong preference for cleavage between two basic residues (Lys and, especially, Arg) in the P_1 and P_1 positions of the substrate (25–27). We sought to isolate OmpT variants that exhibit (*i*) hydrolysis of a substrate that is cleaved poorly by the WT enzyme and (*ii*) a low rate of cleavage of dibasic sequences (preferred by the WT enzyme), thus conferring (*iii*) a high selectivity for the new cleavage over the one preferred by the parental enzyme (Fig. 1*A*), while maintaining (*iv*) a very high level of catalytic activity.

Materials and Methods

Substrate Synthesis. The peptides **1a**, Ac-CARVGKGRGR-NH_2, and **2a**, Glu-EEGGRRIGRGGK-NH_2, were synthesized at the University of Texas peptide synthesis facility. Tetramethylrhodamine-5-iodoacetamide (TMRIA), 5-carboxytetramethylrhodamine, succinimidyl ester (5-TAMRA, SE), and BODIPY-FL-SE were purchased from Molecular Probes. The peptides **3**, WCARVGKGRGR-NH_2, and **4**, WEEGGRRIGRGGK-NH_2 (>95% purity), were purchased from Cell Essentials (Boston).

For the synthesis of the FRET substrate **1**, a 50-μl solution of peptide **1a** (5.4 mg, 4.9 μmol) in water was added to 100 μl of 1 M Na_2CO_3 and 50 μl of 0.5M $NaHCO_3$. A solution of 50 μl of TMRIA (4.0 mg, 4.8 μmol) in dimethylformamide was added to the reaction mixture and stirred at room temperature for 1h. The reaction mixture was quenched with 5 ml of 0.1% trifluoroacetic acid in water, and the product was purified by FPLC, using a 10–30% acetonitrile gradient. The purified product (2.6 mg, 35% yield) was freeze-dried and dissolved in 500 μl of water, and its identity was confirmed by both ^{1}H-NMR and electrospray ionization (ESI)-MS. For the conjugation of BODIPY, a 90-μl

solution of the purified product (1.9 mg, 1.2 μmol) in water was mixed with a 90-μl solution of BODIPY-FL-SE (0.5 mg, 1.3 μmol) in dimethylformamide. Twenty-five microliters of 1 M DMAP (25 μmol) was added to the reaction mixture and stirred at room temperature for 1 h. The reaction mixture was quenched with 5 ml of 0.1% trifluoroacetic acid in water and purified by FPLC with a 10–30% acetonitrile gradient. The purified product (0.7 mg, 32% yield) was freeze-dried, dissolved in 500 μl of water, and quantified by UV. The product identity was confirmed by both ^{1}H-NMR and ESI-MS.

The electrostatic capture substrate **2** was synthesized exactly as described above for BODIPY conjugation, except that 5-TAMRA, SE was used instead of BODIPY-FL-SE.

Flow-Cytometric Analysis. *E. coli* UT5600 [F$^-$ *ara-14 leuB6 secA6 lacY1 proC14 tsx-67* Δ(*ompT-fepC*)*266 entA403 trpE38 rfbD1 rpsL109*(*Str^R*)*xyl-5 mtl-1 thi-1*] was transformed with pML19 (28) encoding WT OmpT under the control of its native promoter. Overnight cultures of UT5600 and UT5600/pML19 were resuspended in 1% sucrose, diluted to 0.01 OD_{600}, labeled for 10 min with 50 nM **1** in 1% sucrose, diluted into 1 ml of 1% sucrose, and analyzed by using a Becton Dickinson FACSort.

Molecular Biology Methods. A library of random mutants was constructed by error-prone PCR using the Diversify mutagenesis kit (Clontech). The PCR product, pAMP1 vector DNA (Life Technologies), annealing buffer, and 2 units of uracil DNA glycosylase in a total volume of 20 μl were incubated for 40 min at 37°C, followed by 1 h at 4°C. The reaction mixture was used to electroporate electrocompetent *E. coli* DH10B [F$^-$ *mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) φ80*lac*ZΔM15 Δ*lac*X74 *rec*A1 *end*A1 *ara*D139 Δ(*ara*, *leu*)7697 *gal*U *gal*K λ-*rps*L *nup*G] cells, and the entire library was plated on selective media. The clones were pooled and frozen at −80°C in aliquots.

OmpT mutants in which Ser-223 was substituted with Arg, Gly, Leu, Lys, Phe, or Trp were constructed by using the QuikChange site-directed mutagenesis kit (Stratagene) with pML19 as the template and appropriate overlapping primer pairs.

Library Screening. Transformants were grown at 37°C in LB media, harvested after 16 h, washed once with 1% sucrose, and resuspended in 1% sucrose. A 50-μl aliquot of the cell suspension in sucrose was added to 949 μl of 1% sucrose and labeled by using 1 μl of each **1** and **2** (final concentration 100 nM). A 20-μl aliquot of this labeling reaction was diluted into 1% sucrose and analyzed on the flow cytometer. Library sorting was performed on a Becton Dickinson FACSCalibur instrument, using gates set based on forward scatter/side scatter and FL1/FL2 values. A total of ≈3 × 10^6 cells were sorted in 29 min, and 1,050 clones were isolated and then resorted. The collected solution was filtered, and the filters were placed on agar plates containing 200 μg/ml ampicillin. After 14 h, 99 colonies that grew were individually inoculated into 1 ml of LB media containing 200 μg/ml ampicillin. Individual colonies were screened by using either substrate **1** or **2** on the flow cytometer and the exact protocol described above.

Enzyme Purification and Kinetic Analysis. Proteins were isolated as previously described, with minor modifications (20), to a final purity >90% as determined by SDS/PAGE. For kinetic analyses, 10–20 nM of the purified enzymes were incubated with 20 μM to 1 mM of the appropriate substrate in 0.1 M Tris/10 mM EDTA (pH 8.0) at room temperature (25°C), and the reaction was monitored by HPLC on a Phenomenex C18 reverse-phase column, using the following gradient: 5% AcN/95% H_2O for 1 min, increasing to 95% AcN/5% H_2O over a period of 29 min and returning to 5% AcN/95% H_2O over 5 min. The product

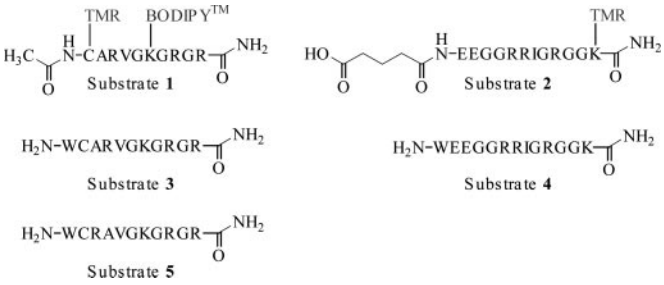

Fig. 2. Substrates used for the detection of catalytic activity and for the kinetic analysis.

concentration was determined by using the integration areas at 280 nm, and the apparent rates were fitted to a Michaelis–Menten kinetic equation by nonlinear regression. The cleavage products were determined by LC-MS (ESI) on a Magic 2002 instrument (Micron Bioresources, Auburn, CA).

Results

Two-Color Flow Cytometric Screening Strategy. Electrostatic interactions between the negatively charged bacterial surface (29) and substrates **1** and **2** (Fig. 2) were exploited to capture the fluorescent and positively charged products of the enzymatic cleavage reaction on the surface of the bacteria. The selection substrate **1** is a FRET peptide with a net +3 charge, the minimum charge required for cell-surface capture (30). Substrate **1** is electrostatically adsorbed on the cell surface but upon cleavage, the N-terminal moiety consisting of Ac-NH_2-CA bound to the tetramethylrhodamine dye is released from the cell, resulting in accumulation of the C-terminal peptide, which is positively charged and contains the BODIPY (4,4-difluoro-5,7-dimethyl-4-bora-3a,4a-diaza-*s*-indacene-3-propionic acid) fluorophore. Substrate **1** was used as the selection substrate because it accumulates on the cell surface before cleavage, resulting in a high effective molarity, which in turn, enables the detection of even very weak catalytic activity. The counterselection substrate **2** was a zwitterionic peptide containing a single fluorophore and at least three positive charges on one side of the scissile bond (Arg–Arg), as well as an equal number of negatively charged groups on the other (Fig. 1*A*). The intact substrate has no net charge, but enzymatic cleavage generates a positively charged moiety carrying the fluorescent dye that is deposited on the cell surface. Upon incubation with **2**, *ompT*$^+$ cells exhibited a >10-fold higher fluorescence relative to the *E. coli ompT*$^-$ deletion strain UT5600 (data not shown). The combination of a surface-bound selection FRET substrate **1** and zwitterionic counterselection substrate **2** that is free in solution afforded the proper dynamic range required for isolation of highly active and selective enzymes.

Library Screening. C5 is an OmpT variant exhibiting a 60-fold increase in activity toward cleavage at a nonpreferred Arg–Val site (20). However, as is typically the case for engineered enzymes, C5 exhibits relaxed overall specificity and is not selective for the cleavage of Arg–Val sites. For example, the C5 variant also cleaved peptide substrates at Ala–Arg sequences with a catalytic efficiency even higher than that of the WT OmpT (Table 1), as well as other sequences (20). The gene encoding C5 downstream from the *ompT* promoter was subjected to random mutagenesis by error-prone PCR (30), and a library of ≈1 × 10^6 transformants was generated. DNA sequencing of 10 randomly selected clones revealed a mutation rate of 1.1% nucleotides per gene. The *E. coli* library was incubated in a solution of 1% wt/vol sucrose (to maintain proper osmotic balance) together with 100 nM of the electrostatic capture substrate **2** and the FRET

BIOCHEMISTRY

Table 1. Kinetic parameters for the cleavage of substrate 3 and substrate 4 by OmpT, C5, and the four mutants

	–A–R–*			–R–R–†			
Enzyme	k_{cat}, s^{-1}	K_M, μM	k_{cat}/K_M, $s^{-1}\cdot M^{-1}$	k_{cat}, s^{-1}	K_M, μM	k_{cat}/K_M, $s^{-1}\cdot M^{-1}$	Specificity –AR–/–RR–
OmpT	$3.1 \pm 0.5 \times 10^{-2}$	16 ± 5	$2.2 \pm 0.9 \times 10^{3}$	8.8 ± 0.7	55 ± 9	$1.7 \pm 0.4 \times 10^{5}$	1.3×10^{-2}
C5	$1.7 \pm 0.5 \times 10^{-2}$	1.5 ± 0.6	$1.4 \pm 0.8 \times 10^{4}$	0.7 ± 0.1	2.2 ± 0.4	$3 \pm 1.0 \times 10^{5}$	3.6×10^{-2}
1.2.19	1.7 ± 0.3	9 ± 1	$2.1 \pm 0.6 \times 10^{5}$	0.4 ± 0.1	260 ± 90	$2 \pm 1 \times 10^{3}$	1.0×10^{2}
1.3.19	2.2 ± 0.4	15 ± 3	$1.5 \pm 0.1 \times 10^{5}$	$4.9 \pm 0.6 \times 10^{-4}$	160 ± 30	3 ± 1	4.7×10^{4}
S223R	2.3 ± 0.1	9 ± 2	$2.6 \pm 0.8 \times 10^{5}$	n.d.	n.d.	n.d.	$>1.8 \times 10^{5}$
D208G	1.6 ± 0.1	7.3 ± 0.8	$2.4 \pm 0.7 \times 10^{5}$	0.3 ± 0.1	240 ± 100	$2 \pm 1 \times 10^{3}$	1.4×10^{2}

Reactions were carried out at room temperature. n.d., not detected.
*Substrate sequence: WCARVGKGRGR-NH_2.
†Substrate sequence: WEEGGRRIGRGGK-NH_2.

substrate **1** for 10 min. Bacterial cells displaying increased BODIPY fluorescence (green, emission at 530 nm, FL-1) and reduced tetramethylrhodamine fluorescence (red, emission at 560 nm, FL-2) corresponding to high activity with substrate **1** and reduced activity with the counterselection peptide **2**, respectively, were isolated (Fig. 1*B*). A total of 1,050 clones were obtained, and after resorting, 99 clones were isolated, grown individually in microtiter well plates. Two clones, 1.3.19 and 1.2.19, were selected for further study because they both exhibited high green fluorescence (Fig. 1*C*) and reduced red fluorescence consistent with the sorting criteria used for their isolation. Sequencing determined that 1.2.19 contained Asp208Gly and Asp214Val mutations, whereas 1.3.19 contained Ser17Gly and Ser223Arg mutations in addition to those found in C5, the parent enzyme used in this study.

The 1.2.19 and 1.3.19 proteins were extracted in *n*-octyl-β-glucoside and purified (25). Kinetic analysis of the purified OmpT 1.2.19 and 1.3.19 (Table 1) proteins by using unlabeled peptide substrates **3** and **4** (Fig. 2) revealed that the selection yielded the anticipated outcome, namely, highly active enzymes that are specific for the hydrolysis of **3** (in particular, cleavage of the Ala–Arg peptide bond) but impaired in their ability to attack Arg–Arg. It is interesting to note that although cleavage between any two amino acids in the linker region between the fluorophore and the quencher would lead to increased FL-1 fluorescence, the two best variants 1.2.19 and 1.3.19 hydrolyzed substrate **3** between Ala and Arg. In addition, 1.3.19 did not exhibit the secondary cleavage activities displayed by its parental enzyme C5. Importantly, compared with the WT OmpT, 1.3.19 displayed a $>3 \times 10^6$ reversal in selectivity (Table 1) for Ala–Arg over Arg–Arg cleavage, but its enzymatic activity for cleavage of **3** ($k_{cat}/K_M = 1.5 \times 10^5$ $s^{-1}\cdot M^{-1}$) was nearly identical to that of the WT OmpT for its own preferred substrate **4** ($k_{cat}/K_M = 1.7 \times 10^5$ $s^{-1}\cdot M^{-1}$). The isolation of both 1.2.19 and 1.3.19 suggests that such highly selective, highly active catalysts, as opposed to enzymes with expanded specificity or low catalytic activity toward new substrates, can be easily isolated by multiparameter, quantitative screening of large libraries.

Site-Specific OmpT Mutants. With the exception of the Ser17Gly mutation in 1.3.19, the other three amino acid substitutions in the two isolated clones resulted in mutations that involved charged residues. Examination of the recently reported structure of OmpT (31) pointed to a significant role for Ser223Arg and Asp208Gly in modulating peptide substrate specificity, because these residues are located deep in the active-site cleft.

Ser223Arg and Asp208Gly were constructed by using site-directed mutagenesis and purified as described above. LC-MS of the substrate cleavage products revealed that the OmpT Ser223Arg and Asp208Gly enzymes cleaved **3** only between Ala and Arg. Neither variant produced the secondary Lys–Gly cleavage that is generated by the WT OmpT ($k_{cat}/K_M = 4 \times 10^3$ $s^{-1}\cdot M^{-1}$) as well as the C5 enzyme. In addition, the single amino acid variants recapitulated the high selectivities and catalytic activities displayed by 1.2.19 and 1.3.19, respectively (Table 1). In particular, Ser223Arg completely abolished the ability of OmpT to cleave Arg–Arg sites (no cleavage of **4** could be detected after >24 h incubation) yet cleaved **3** at Ala–Arg with a k_{cat}/K_M value slightly higher than that of 1.3.19. In addition, the Ser223Arg mutant did not cleave substrate **5** (Fig. 2), further underscoring its specificity for Ala–Arg. The effect of the Ser223Arg mutation appeared to be unique, because insertion of Trp, Lys, Leu, Gly, or Phe at the 223 position by using site-directed mutagenesis resulted in variants that did not exhibit cleavage of substrate **1** comparable to even WT OmpT when monitored by FACS (Fig. 5).

Discussion

We have developed a high-throughput strategy for the quantitative screening of enzyme libraries by using simultaneous selection and counterselection criteria. The methodology presented depends on (*i*) the display of enzyme libraries on the surface of microorganisms, (*ii*) retention of fluorescent reaction products on the cell surface, and (*iii*) multicolor flow cytometry for the isolation of clones that can selectively turn over one or more substrates. The combination of electrostatic and FRET substrates afforded the proper dynamic range required to screen a wide array of enzymatic activities, an important consideration when screening large enzyme libraries.

Display of enzymes on the surface of microorganisms such as *E. coli* or yeast can be accomplished in a variety of ways (32, 33). In this manner, the enzyme can react with exogenous synthetic substrates, circumventing the substrate transport limitations associated with methods that use intracellularly expressed enzymes. In our approach, the enzymatic reaction generates fluorescent products that become associated with the cell surface, resulting in a fluorescence profile representative of the catalytic selectivity of the displayed enzyme. Although we have capitalized on electrostatic interactions for product capture, a number of other methods for cell-surface modification (34, 35) may be exploited for the capture of reaction products. In addition to the directed evolution of protease selectivity reported here, we believe that our methodology can be extended to other enzymes including various hydrolases and ligases.

We screened an OmpT error-prone library to select for variants that preferred a new peptide substrate at the expense of activity with the WT preferred substrate. After three rounds of sorting, we obtained 10 clones that displayed the desired specificity profile, the best of which were 1.2.19 and 1.3.19. Although the targeted error-rate for the library was 1.1% (11 bases), both 1.2.19 (Asp208Gly and Asp214Val) and 1.3.19 (Ser17Gly and Ser223Arg) had just two extra mutations each, relative to C5

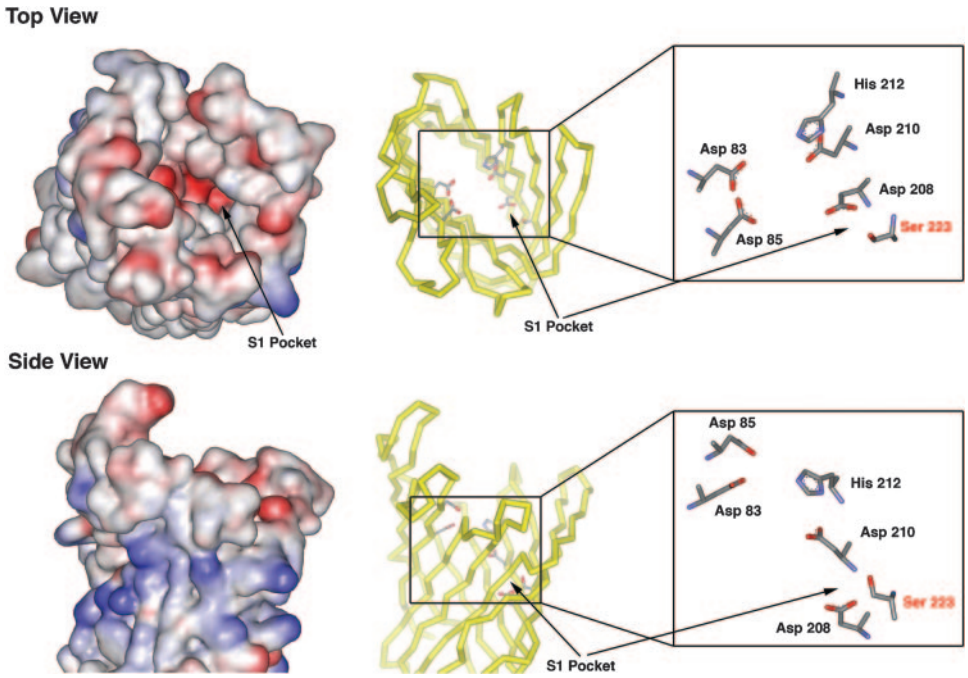

Fig. 3. The crystal structure of OmpT. (*A*) The electrostatic potential surface of OmpT (31) generated by using WEBLAB VIEWERLITE (Accelrys, San Diego). (*B*) The side chains corresponding to Asp-208 and Ser-223 are shown in addition to the proposed catalytic residues, Asp-83, Asp-85, and Asp-210 and His-212.

(20), the starting enzyme construct. A detailed kinetic analysis of the isolated enzymes proved both to be consistent with the flow-cytometric screening criteria. In other words, both of these enzymes have altered, as opposed to relaxed, substrate specificity yet maintain a native level of catalytic activity.

Examination of the crystal structure of OmpT (31) indicated that the Asp208Gly mutation in 1.2.19 and the Ser223Arg in 1.3.19 might be the primary determinants of the altered specificity profiles of these enzymes. Construction of these single point mutants of OmpT confirmed that Asp208Gly was similar to 1.2.19 in terms of specificity and overall activity. Unexpectedly, Ser223Arg was more specific than 1.3.19, especially in eliminating cross-reactivity with the WT preferred Arg–Arg-containing substrate **4**. Importantly, the Ser223Arg variant was as active with substrate **3** as WT OmpT with its preferred substrate **4**. Thus, for the Ser223Arg single-mutation variant of OmpT, altered specificity has not come at the cost of overall catalytic activity.

It is worth mentioning that although OmpT is a trypsin-like protease (36), the Ser223Arg substitution conferred specificity more analogous to that of chymotrypsin. The effect of this single amino acid substitution should be contrasted with the conversion of trypsin to chymotrypsin that required a monumental effort and hinged on major reorganization of the substrate-binding surface (37, 38).

A crystal structure of OmpT has been published (31), but no structural information is available for OmpT containing a bound substrate or substrate analog. Nevertheless, inspection of the OmpT structure leads to a reasonable prediction for the location of an unusually deep S_1 binding pocket (Fig. 3), with the Ser-223 residue located near the bottom of this pocket. Consistent with our proposed role of the 223 residue in substrate recognition, a recent molecular dynamics calculation using WT OmpT and an

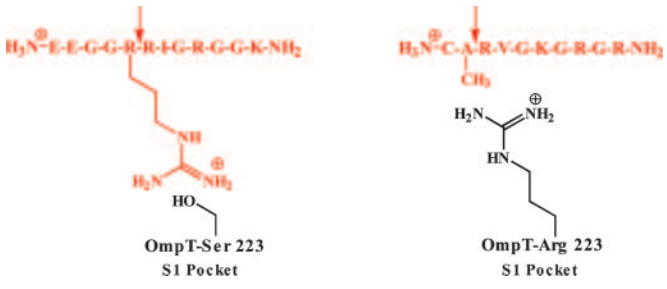

Fig. 4. Arg side-chain swapping. The Arg occupying the P_1 subsite is accommodated by Ser-223 in S_1. In the 1.3.19 mutant, the Ser223Arg only allows occupancy by Ala in the P_1 site of the substrate.

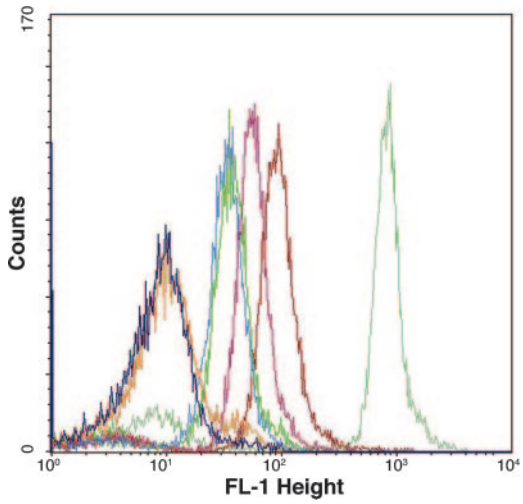

Fig. 5. Flow-cytometric analysis of the OmpTSer223 mutants. Briefly, the cells were washed and resuspended in 1% sucrose and labeled with 50 nM (final concentration) of the AR-FRET substrate **1**. A 20-μl aliquot of the labeling reaction was transferred to 0.5 ml of 1% sucrose and analyzed on the flow cytometer. Dark blue, S223W; orange, S223L; light blue, S223F; dark green, S223K; pink, S223G; red, OmpT; light green, S223R.

BIOCHEMISTRY

Arg–Arg-containing substrate placed the P_1 Arg side chain deep in the same S_1 pocket proposed here and adjacent to Ser-223 (39).

The high activity seen with Ser223Arg likely reflects a simple swapping of an important residue from the peptide substrate to the enzyme active site. Because enzymes evolve to be catalytically active in the presence of substrate, it makes sense that some functional groups attached to the substrate could modulate active site structure and thus catalytic activity to a significant degree. This is the basis of Koshland's induced fit theory of enzyme catalysis (40) in which substrate binding is required to convert an enzyme active site into the proper catalytic arrangement. A logical interpretation of our data is that the Ser223Arg substitution replaces important induced fit interactions of the P_1 Arg side chain of a bound substrate (Fig. 4) with the 223Arg side chain that is now attached to the enzyme in S_1. A significant structural role for the restored guanidium group would explain why the Ser223Arg variant has WT levels of catalytic activity. Interestingly, the decreased rate of catalysis seen for WT OmpT reacting with the nonpreferred substrate **3** (that lacks the P_1 Arg residue) is primarily a reflection of a substantially lower k_{cat} value, possibly consistent with the induced fit hypothesis that a guanidium group in the P_1-S_1 site is important for high levels of catalytic activity. To investigate this hypothesis, we constructed five OmpTSer223 mutants, OmpTSer223Phe, OmpTSer223Gly, OmpTSer223Lys, OmpTSer223Leu, and OmpTSer223Trp. These five mutants were designed to test whether the guanidium group is essential in the P_1-S_1 site for high catalytic activity or if steric (Phe, Trp), hydrophobic (Leu), or positively charged (Lys) residues can help achieve the same goal. Flow-cytometric analysis (Fig. 5) demonstrates that none of these mutants had even a WT level of activity with substrate **1**, confirming the need for a guanidium group.

Swapping important residues at protein–protein and protease–substrate interfaces has been accomplished by rational design (41, 42). For 1.3.19, this apparent swapping occurred as a result of combined selection and counterselection. In that regard, the isolation of the OmpT Ser223Arg variant supports the notion that residue swapping at protein–peptide interfaces represents a facile mechanism for substrate diversification as the result of evolutionary pressure without any preconceived notions regarding the protein structure.

Impressive as our findings with OmpT Ser223Arg might be, in general, the engineering of large changes in substrate specificity is expected to require multiple amino acid substitutions leading to extensive remodeling of the active site. Thus, whereas in this case the Ser223Arg mutant could have been identified by screening a small library using manual techniques, the evolution of mutants of OmpT or other enzymes that can react selectively with nonnative substrates will likely require the screening of very large libraries. Therefore, a high overall assay throughput/dynamic range and the ability to place selection and counterselection evolutionary pressure on an enzyme to be evolved are key to the isolation of rare, change-of-function clones. Selection and counterselection assays can be carried out in a tiered approach, wherein the library is screened first for one substrate and then the active clones are screened for activity with the counterselection substrate. As shown here, however, selection and counterselection can also be applied simultaneously, analogous to living systems. Simultaneous screening by FACS accommodates high throughput and is particularly attractive for carrying out evolutionary experiments using three or more substrates, as required to explore "substrate space."

We thank Daren Stephens and Marsha Demers for technical assistance; Drs. Andrew Ellington, Jack Kirsch, and Frances Arnold for constructive comments on the manuscript; and the University of Texas Peptide Synthesis and Mass Spectrometer facilities. This work was supported by National Institutes of Health Grant R01 GM065551 (to B.L.I. and G.G.).

1. Carter, P. & Wells, J. A. (1987) *Science* **237,** 394–399.
2. Bone, R., Fujishige, A., Kettner, C. A. & Agard, D. D. (1991) *Biochemistry* **30,** 10388–10398.
3. Palzkill, T. & Botstein, D. (1992) *J. Bacteriol.* **174,** 5237–5243.
4. Graham, L. A., Brandt, U., Sargent, J. S. & Trumpower, B. L. (1993) *J. Bioenerg. Biomembr.* **25,** 245–257.
5. Harris, J. L. & Craik, C. S. (1998) *Curr. Opin. Chem. Biol.* **2,** 127–132.
6. Farinas, E. T., Butler, T. & Arnold, F. H. (2001) *Curr. Opin. Biotechnol.* **12,** 545–551.
7. Arnold, F. H. (2001) *Nature* **409,** 253–257.
8. Hedstrom, L. (2002) *Chem. Rev.* **102,** 4501–4524.
9. Tao, H. & Cornish, V. W. (2002) *Curr. Opin. Chem. Biol.* **6,** 858–864.
10. Matsumura, I. & Ellington, A. D. (2001) *J. Mol. Biol.* **305,** 331–339.
11. Geddie, M. L. & Matsumura, I. (2004) *J. Biol. Chem.* **279,** 26462–26468.
12. Aharoni, A., Gaidukov, L., Khersonsky, O., McQ. Gould, S., Roodveldt, C. & Tawfik, D. S. (2005) *Nat. Genet.* **37,** 73–76.
13. Reetz, M. T., Wilensek, S., Zha, D. & Jaeger, K.-E. (2001) *Angew. Chem. Int. Ed.* **40,** 3589–3591.
14. DeSantis, G., Wong, K., Farwell, B., Chatman, K., Zhu, Z., Tomlinson, G., Huang, H., Tan, X., Bibbs, L., Chen, P., *et al.* (2003) *J. Am. Chem. Soc.* **125,** 11476–11477.
15. Van Loo, B., Spelberg, J. H., Kingma, J., Sonke, T., Wubbolts, M. G. & Janssen, D. B. (2004) *Chem. Biol.* **11,** 981–990.
16. Rothman, S. C. & Kirsch, J. F. (2003) *J. Mol. Biol.* **327,** 593–608.
17. Yano, T. & Kagamiyama, H. (2001) *Proc. Natl. Acad. Sci. USA* **98,** 903–907.
18. Santoro, S. W. & Schultz, P. G. (2002) *Proc. Natl. Acad. Sci. USA* **99,** 4185–4190.
19. Bucholz, F. & Stewart, A. F. (2001) *Nat. Biotechnol.* **19,** 1047–1052.
20. Olsen, M. J., Stephens, D., Griffiths, D., Daugherty, P., Georgiou, G. & Iverson, B. L. (2000) *Nat. Biotechnol.* **18,** 1071–1074.
21. Francisco, J. A., Campbell, R., Iverson, B. L. & Georgiou, G. (1993) *Proc. Natl. Acad. Sci. USA* **90,** 10444–10448.
22. Georgiou, G., Stathopoulos, C., Daugherty, P. S., Nayak, A. R., Iverson, B. L. & Curtiss, R., III (1997) *Nat. Biotechnol.* **15,** 29–34.
23. Kukkonen, M. & Korhonen, T. K. (2004) *Int. J. Med. Microbiol.* **294,** 7–14.
24. Okuno, K., Yabuta, M., Ooi, T. & Kinoshita, S. (2004) *Appl. Environ. Microbiol.* **70,** 76–86.
25. McCarter, J. D., Stephens, D., Shoemaker, K., Rosenberg, S., Kirsch, J. F. & Georgiou, G. (2004) *J. Bacteriol.* **186,** 5919–5925.
26. Dekker, N., Cox, R. C., Kramer, R. A. & Egmond, M. R. (2001) *Biochemistry* **40,** 1694–1701.
27. Okuno, K., Yabuta, M., Kawanishi, K., Ohsuye, K., Ooi, T. & Kinoshita, S. (2002) *Biosci. Biotechnol. Biochem.* **66,** 127–134.
28. Matthews, D. J. & Wells, J. A. (1993) *Science* **260,** 1113–1117.
29. Razatos, A., Ong, Y. L., Sharma, M. M. & Georgiou, G. (1998) *Proc. Natl. Acad. Sci. USA* **95,** 11059–11064.
30. Daugherty, P. S., Iverson, B. L. & Georgiou, G. (2000) *J. Immunol. Methods* **243,** 211–227.
31. Vandeputte-Rutten, L., Kramer, R. A., Kroon, J., Dekker, N., Egmond, M. R. & Gros, P. (2001) *EMBO J.* **20,** 5033–5039.
32. Lee, S. Y., Choi, J. H. & Xu, Z. (2003) *Trends Biotechnol.* **21,** 45–52.
33. Chen, W. & Georgiou, G. (2002) *Biotechnol. Bioeng.* **79,** 496–503.
34. Saxon, E. & Bertozzi, C. R. (2001) *Annu. Rev. Cell Dev. Biol.* **17,** 1–23.
35. Link, A. J., Vink, M. K. & Tirrell, D. A. (2004) *J. Am. Chem. Soc.* **126,** 10598–10602.
36. Leytus, S. P., Bowles, L. K., Konisky, J. & Mangel, W. F. (1981) *Proc. Natl. Acad. Sci. USA* **78,** 1485–1489.
37. Hedstrom, L., Szilagyi, L. & Rutter, W. J. (1992) *Science* **255,** 1249–1253.
38. Kurth, T., Ullmann, D., Jakubke, H. D. & Hedstrom, L. (1997) *Biochemistry* **36,** 10098–10104.
39. Baaden, M. & Sansom, M. S., *Biophys. J.*, in press.
40. Koshland, D. E. (1958) *Proc. Natl. Acad. Sci. USA* **44,** 98–104.
41. Atwell, S., Ulstch, M., De Vos, A. M. & Wells, J. A. (1997) *Science* **278,** 1125–1128.
42. Caputo, A., Parrish, J. C., James, M. N., Powers, J. C. & Bleackley, R. C. (1999) *Proteins* **35,** 415–424.