Title
Multiplexed Enzyme Activity-Based Probe Display via Hybridization

Journal
ACS Combinatorial Science, 22(11)

ISSN
2156-8952

Authors
Cavett, Valerie
Paegel, Brian M

Publication Date
2020-11-09

DOI
10.1021/acscombsci.0c00116

Peer reviewed

HHS Public Access

Author manuscript

ACS Comb Sci. Author manuscript; available in PMC 2021 November 09.

Published in final edited form as:

ACS Comb Sci. 2020 November 09; 22(11): 579–585. doi:10.1021/acscombsci.0c00116.

Multiplexed enzyme activity-based probe display via hybridization

Valerie Cavett[†], **Brian M. Paegel**[†,‡]

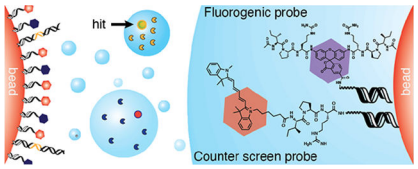
[†]Department of Pharmaceutical Sciences, University of California, Irvine

[‡]Departments of Chemistry & Biomedical Engineering, University of California, Irvine

Abstract

Emulsions offer the means to miniaturize and parallelize high-throughput screening, but require a robust method to localize activity-based fluorescent probes in each droplet. Multiplexing probes in droplets is impractical, though highly desirable for identifying library members that possess very specific activity. Here, we present multiplexed probe immobilization on library beads for emulsion screening. During library bead preparation, we quantitated ~10^6 primers per bead by fluorescence in situ hybridization, however emulsion PCR yielded only ~10^3 gene copies per bead. We leveraged the unextended bead-bound primers to hybridize complementary probe-oligonucleotide heteroconjugates to the library beads. The probe-hybridized bead libraries were then assayed in emulsion in vitro transcription/translation reactions and analyzed by FACS to perform multiplexed activity-based screening of trypsin and chymotrypsin mutant libraries for novel proteolytic specificity. The approach's modularity should permit a high degree of probe multiplexing and appears extensible to other enzyme classes and library types.

Graphical Abstract



Keywords

emulsion; screening; protease; beads; directed evolution

Emulsions offer an alternative to high-throughput screening in microplates wherein reactions are compartmentalized in water-in-oil droplets instead of individual wells.[1] As the number of reactions for screening is increased past the capacity of plates, emulsions offer advantages in scale, reagent requirements, and handling. The scaling advantages of emulsion-based

bpaegel@uci.edu.

Supporting Information Available

The supporting information is available free of charge on the ACS Publications website.

- Experimental detail and additional results.

biochemistry have been tremendously enabling for biotechnology, particularly for applications like next-generation DNA sequencing, single-cell transcriptomics, PCR-based diagnostics, enzyme evolution, and other methods involving highly diverse libraries of genes.[2–9] While many of these approaches are amenable to microfluidic emulsification, which generates monodisperse droplets and can be integrated with powerful droplet sorting circuitry,[5,10–12] the requisite specialized devices and flow controllers are not yet standard laboratory equipment. As an alternative, bulk emulsification typically relies on readily available equipment, such as stir plates or tissue homogenizers, but the resultant polydisperse aqueous droplets in oil are not sortable via commercial flow cytometry. The rising widespread availability of commercial flow cytometry instruments for sorting has prompted the development of secondary emulsification techniques, which replace the oil continuous phase of single-stage emulsions with an aqueous continuous phase that flow cytometers require for analysis.[13,14] The need to systematize double emulsification has led to further substantial microfluidic technology development.[15–17]

To avoid microfluidic droplet sorting or secondary emulsification steps, particles encapsulated in individual emulsion droplets[2] can be used to co-localize the gene and activity-based probe, and these particles can be sorted after all emulsion reactions are completed. Using this workflow, we screened for and identified a citrulline-dependent mutant of trypsin, trypsin^{+cit}, from a gene library of trypsin^{D189S} loop mutants.[18] Primers and a quenched fluorescent probe of proteolytic activity were chemically coupled to 2.8-μm diameter magnetic beads. The beads were emulsified with gene library templates in PCR mix. Emulsion PCR (emPCR) yielded a library of monoclonal beads that each displayed several thousand copies of a single gene mutant.[19,20] The beads were emulsified with in vitro transcription and translation (emIVTT) mix to generate the protein product of the gene library member, which then acted upon the bead-immobilized probes for subsequent analysis via FACS.

This initial emPCR/emIVTT activity-based screening workflow for identifying proteases with novel specificity had several limitations. While qPCR-based quantitation of post-emPCR bead-bound DNA templates was robust, confirmation of activity-based probe loading was difficult because the probe was coupled to the bead in a quenched state. Characterization of the probes prior to screening was not possible because covalent attachment to the bead precluded standard chemical analysis. Additionally, it was only feasible to immobilize a single type of probe due to coupling chemistry orthogonality; this eliminated the opportunity to counter-screen undesired off-target proteolytic activity. As a result, the activity profile of the trypsin^{+cit} mutant (the product of screening only for citrulline-dependent activity) featured unsurprisingly high levels of wild-type Arg-dependent activity. It was clear that both forward and counter screening probes were necessary to achieve further substrate selectivity, but we required a strategy for coupling multiple probe stoichiometrically and quantitatively.

Our previous efforts to identify chemoselective bead functionalization chemistry for emPCR unexpectedly revealed a surplus of primers post-emPCR that hypothetically could be used to immobilize multiple different probes via hybridization.[21] We performed extensive characterization and optimization of the primer-functionalized beads used for emPCR and

routinely quantitated 2×10^6 primers per bead, while emPCR yielded only ~ 10^3 template copies per bead. The bead-bound primers that remained after emPCR-based extension were hybridized to a fluorescently labeled complementary oligonucleotide and analyzed by FACS to confirm that hybridized fluorescent probes would yield adequate signal for sorting. We used the statistical assay quality, Z', to quantify the signal to noise, which was excellent (0.71, Figure S1).[22] These data supported a proposed new strategy for selectively immobilizing activity-based probes on the bead surface by hybridization to the unextended bead-bound primer sites (Figure 1). In this "probe hybridization" workflow, single-molecule emPCR amplification of library members is performed on primer-functionalized beads to yield monoclonal gene library beads. The unextended remaining primers on the gene library beads are then hybridized to activity-based probes that are conjugated to a complementary oligonucleotide. Parallel expression of these probe-gene beads in emIVTT generates a library of mutant enzymes that each act on the encapsulated bead's hybridized probe set for subsequent screening by FACS.[18]

We next sought to demonstrate that the multiplexed probe hybridization approach was experimentally feasible. Modular probe-oligonucleotide heteroconjugates were designed and synthesized. A citrulline-dependent probe of proteolytic activity (forward probe, Figures 2A, S2) was prepared by acylating the pendant amines of the rhodamine 110 xanthene core with citrulline, quenching the rhodamine fluorescence.[23–25] The pendant citrulline residues were elaborated further using standard peptide synthesis to yield the requisite IPCit probe cleavage site sequence context. A PEG linker separated the rhodamine core from a propargylglycine that was used in CuAAC coupling to the 5′-azido oligonucleotide with sequence complementary to the bead-bound primer. A peptide sequence containing an Arg in the forward probe sequence context (IPRAA) and N-terminal Cy5 fluorescent dye was designed as a counter screening probe (counter probe, Figures 2B, S3). A C-terminal Ala spacer separated the Arg cleavage site from a propargylglycine that was used in CuAAC coupling to the bead-bound primer complement as described above. Cleavage of the counter screening probe at any of its amide bonds would liberate Cy5 from the bead, attenuating Cy5 fluorescence. Forward and counter screening probes were mixed and hybridized to primer functionalized beads, yielding particles with high red fluorescence (660 nm) and low green fluorescence (530 nm). Digestion with citrulline- and Arg-dependent trypsin^{+cit} mutant protease cleaved both forward and counter screening probes (Figure 2C). Rhodamine probe fluorescence (530 nm) was dequenched and Cy5 dye was liberated from the bead, decreasing red fluorescence (660 nm) (Figure 2C). Z'-factors were calculated for citrulline probe cleavage (0.59) and Arg cleavage (0.45), indicating good separation of positive and negative signal in both fluorescence channels.

The probe hybridization strategy offered several key characterization and assay design advantages. First, probe construction and characterization are decoupled from the bead surface functionalization. It is very difficult to conduct probe characterization in situ on surfaces, especially for probes that are quenched (e.g., R110 bisamide probes) and therefore exhibit negligible fluorescence or absorbance; the oligonucleotide provides a convenient spectroscopic handle for quantitating such quenched probes. The modular synthesis scheme is compatible with multiple different probe designs that transduce the desired enzymatic activity with increased fluorescence intensity, which is optimal for assay design. Here, the

forward screening probe increases in fluorescence signal with increasing desired (Cit-dependent) activity, whereas the counter screening probe fluorescence is highest when the off-target (Arg-dependent) activity is lowest. Second, hybridization is sensitive and mild compared to synthesis-based coupling (e.g., amide formation). Probe hybridization requires very little material (~ 30 pmol) owing to the high specificity and potency of hybridization. Finally, the data clearly demonstrate that at least two probes can be mixed and hybridized, obviating the need to design multiple orthogonal coupling chemistries that are also orthogonal with all functional groups on the probe.

The dual-probe assay was next tested by screening multiple libraries for Cit-specific proteases that exhibit low off-target Arg-dependent activity. Three different scaffolds, including wild-type (WT) trypsin, active site mutant trypsin^{D189S}, and chymotrypsin were mutagenized in two specificity-determining loops (loop 1 and loop 2).[26,27] Loop 1 (residues 184–188, chymotrypsin numbering) is located just N-terminal to the D189 active site. Loop 2 is more distally located (residues 217–220), and influences the specificity of ester hydrolysis in more indirect ways.[28] Each library was transferred to beads via emPCR, then hybridized with a mixture of $(IPCit)_2$R110 and Cy5-IPRAA probes, assayed for proteolytic activity in emIVTT, and analyzed by FACS (Figure 3A). Primer-functionalized beads hybridized to both probes were also analyzed by FACS to acquire a starting population (Q1 pane), defined in both channels as $<3\sigma$ from the mean (μ) of the starting population. Citrulline-directed activity is rare for either loop 1 or loop 2 mutants of WT trypsin, but more common for trypsin^{D189S} and even more so for chymotrypsin. The inverse is true for trypsin^{D189S} and chymotrypsin with respect to Arg-dependent activity. The loop 2 library contains more residual off-target Arg-dependent activity in the WT and more citrulline-dependent mutants in trypsin^{D189S} and chymotrypsin. Highly fluorescent beads (530 and 660 nm channels) were collected for each library, amplified, and the amplification products were analyzed in real-time IVTT (RT-IVTT) activity assays (Figure 3B). Trypsin^{D189S} and chymotrypsin scaffolds and unsorted libraries exhibit modest activity against the citrulline target. This activity is enhanced in the hit collections of these screens. WT trypsin, and the loop libraries derived from the trypsin scaffold exhibit high Arg-dependent activity, and no measurable activity against the citrulline target. In all hit collections, Cit-dependent activity is enhanced while Arg-dependent activity is suppressed.

The fraction of hits falling within each quadrant was quantitated for each scaffold and loop library. Functional densities, ρ, were calculated by dividing the number of hits in Q2 (forward screening density, ρ_{FWD}), Q3 (promiscuous density, ρ_{prom}), or Q4 (counter screening density, ρ_{CTR}) by the sum of events in Q2, Q3, and Q4 (Table 1). The events in Q1 were not considered because an unknown fraction of beads was not templated during emPCR. Of the scaffolds and loop libraries screened, the chymotrypsin loop 2 library was the most productive for citrulline-dependent activity. Its ρ_{FWD}:ρ_{CTR} bias was ~20 fold, followed by chymotrypsin loop 1 (~10-fold bias) and then trypsin^{D189S} loop 1 (4.5-fold).

Dual-probe screens of the three scaffolds' loop libraries revealed profound differences in the density of proteolytic function. Qualitatively, trypsin^{D189S} and chymotrypsin share very similar distributions of citrulline- and Arg-dependent proteolysis (Table 1). That trypsin^{D189S} was chymotrypsin-like recapitulated seminal enzymological studies of this

variant[27,29] and agreed with our observation of chymotryptic activity in MS/MS experiments using trypsin^{+cit}, which introduced chymotryptic cleavages.[18] However, the single-point mutant played an outsized role in introducing novel Cit-dependent activity over either loop library, suggesting that systematic active site exploration[30] of this protease family may be fruitful, especially for discovering completely novel cleavage specificities. In fact, the screens clearly show that chymotryptic loop libraries are inverted from WT trypsin, which heavily favors charged Arg-directed proteolysis over the neutral citrulline. Overall, the analysis of proteolytic specificities (Arg vs. Cit) confirmed that they are completely decoupled in these scaffolds and libraries. Achieving clean Cit-dependent proteases appears very tractable, and might be even more easily attained in the chymotrypsin genetic background.

As a demonstration of the dual-probe hybridization approach's generality, we pivoted to exploration of chymotrypsin promiscuity using probes of canonical chymotryptic activity. R110 bisamide forward screening probes of chymotryptic activity, $(AAPF)_2R110$ and $(AAPY)_2R110$, and counter screening chymotryptic probe Cy5-AAPWAA were synthesized and coupled to the probe hybridization oligonucleotide (Figures S4, S5, S6). Probes were hybridized to loop 1 and loop 2 chymotrypsin gene bead libraries (Figure 4), assayed in emIVTT, and screened by FACS, reserving a portion of unreacted and probe-hybridized beads to set the 3σ gate (see above). The promiscuous activity quadrant, Q3, contained the highest functional density, and the populations formed a continuous band of active mutants bridging Q1 and Q3. Both Q2 and Q4 contained very little functional density. The distribution of proteolytic specificities in the population was qualitatively independent of forward screening probe or loop library identity.

The dual-probe chymotrypsin screen was designed to demonstrate the generality of both the probe design and screening approach. Chymotrypsin is a notoriously promiscuous protease, which limits its utility in protein mass spectrometry.[31] We speculated that narrowing its cleavage profile (e.g., by restricting activity to Phe and/or Tyr) might yield a more useful tool, though it was unclear whether removing activity from a promiscuous protease would be possible. With monoclonal gene bead loop 1 and loop 2 libraries of chymotrypsin already in hand, only synthesis of the chymotrypsin specific probes was required to perform screens of chymotryptic selectivity. Exploration of the chymotrypsin loop libraries revealed that Phe- and Tyr-specific activities were closely tied to Trp-specific activity, and that loop modifications were unlikely to result in a chymotrypsin variant with narrowed specificity. The majority of the population continuously bridged from Q1 (inactive) to Q3, in stark contrast with the results of the Cit/Arg screens. Given the importance of the active site mutant, trypsin^{D189S}, in accessing different activity space for Cit/Arg profiling, achieving a more desirable starting population of cleavage dependencies for chymotrypsin may require exploration of its active site as well.

Here we have shown that hybridization is a modular, mild, and selective method for immobilizing activity-based probes on emPCR bead libraries. Probe hybridization leverages the abundant unextended primer sites remaining after emPCR, linking emPCR amplicon as genotype to the post-emIVTT probe state phenotype via the bead substrate. As the approach is hybridization-based, probe association occurs under extremely mild conditions and

requires minute quantities of probe. Library screening using multiple probes is also straightforward since probes are simply mixed to the desired ratio in solution and added to the beads. In this study, we considered only two types of probes and two emission wavelengths, but recent advances in dye chemistry delivering an expanded palette of probe emission with superior stability would also be compatible with probe hybridization.[32] Finally, although probe immobilization was noncovalent, the probes remained hybridized through emIVTT and FACS analyses, as expected from an oligonucleotide with $T_m > 65°C$. In this experimental context, the high potency of association via hybridization may as well be considered covalent.

There are several notable limitations of probe hybridization. First, coupling of probe to the hybridization oligonucleotide must not compromise the probe's activity. There was good precedent for attaching activity-based probes to oligonucleotides for performing substrate selections.[33] And, while we did not observe reduced activity for any of the tryptic or chymotryptic R110 bisamide or dye-labeled peptide substrates, this may be problematic for other enzymes. For example, steric interference of DNA in screens of DNA-encoded libraries is well known.[34,35] Additionally, probe association is noncovalent, thus assay conditions must not disfavor duplex formation (e.g., high temperature, solvent), promote strand displacement, or degrade DNA.[36]

Although this study disclosed the use of probe hybridization for screening mutant protease libraries, there are numerous other possible applications. Many droplet- and emulsion-based directed evolution experiments targeting diverse protein classes could be enabled by probe hybridization.[8,10,37,38] Perhaps the most ideally suited would be enzymes that process nucleic acid substrates, such as polymerases and ligases, since their substrates are readily immobilized by hybridization. Hybridization is also known to be highly modular and predictable.[39] It should be possible to design more elaborate configurations of probe display by, for example, functionalizing the bead with multiple primers or by designing sequences that contain multiple (replicate or unique) hybridization regions. Finally, these concepts may extend to other library types, such as DNA-encoded solid-phase libraries,[40,41] or other substrates to which an oligonucleotide is readily coupled or, as in this case, repurposed to hybridize and thereby display an activity-based probe.

Bead functionalization and characterization.

An aliquot of carboxylic acid-functionalized magnetic beads (M-270 carboxylic acid Dynabeads, 1 mg, Life Technologies, Carlsbad, CA) was washed (DMF, 2 × 500 μL). A solution of HOAt (500 mM), propargylamine (500 mM), and DIC (700 mM) was prepared in DMF (400 μL total volume), added to the beads, and the beads were incubated (3 h, 50 °C). The alkyne-functionalized beads were washed (DMF, 5 × 500 μL) and resuspended (DMF, 1 mL). Solvent was removed from an aliquot of alkyne-functionalized beads (0.5 mg) and the beads were combined with a solution 5′-azidopentanoic acid-AGACCGAGATAGGGTTGAGTGTTG-3′ (144 μM), $CuSO_4$ (1.6 mM), L(+)-ascorbic acid (8 mM), TBTA (1.9 mM), Tween-20 (0.022% w/v), and TEAA (86 mM, pH 7) prepared in 50% v/v DMSO (62.5 μL total volume), the reaction was incubated (3 h, RT), and the beads were washed (breaking buffer, 3 × 1 mL), and resuspended (bead buffer, 1 mL). Buffer was

removed from an aliquot of primer-functionalized beads (10^6 beads), and the beads were combined with fluorescein-labeled probe oligonucleotide hybridization solution (1 μM 5′-FA M-CACTCAACCCTATCTC-3′ in 2× SSC, 20 μL) and incubated (RT 5 min). The beads were washed (2×SSC, 3 × 20 μL; 0.1 N NaOH, 2 × 20 μL, 60 °C) and all washes saved. Standard solutions were prepared (10 – 400 nM, 5′-FAM-CACTCAACCCTATCTC-3′ in 2×SSC) and analyzed together with SSC washes and NaOH eluants using a real-time PCR instrument (channel 1, CFX96, Bio-Rad, Hercules, CA) to determine the unknown [FAM] in the washes and eluants. The quantitated number of fluorophores was divided by the quantitated number of beads to yield the average oligonucleotide primer loading per bead.

Emulsion PCR (emPCR) bead library preparation.

Primer-functionalized beads (1.3×10^7) were washed twice (1×PCR MM). Each emPCR sample is prepared in a mi-crocentrifuge tube (2 mL, SafeLock, Eppendorf) containing a steel ball (6 mm dia) and amplification reaction mixture. Amplification reaction mixture contained dNTPs (0.2 mM each dNTP), 8 μM forward primer 5′-TGCGTCCGGCGTAGAGGATC-3′, 2 pg/μL library template, *Taq* polymerase (90U, New England Biolabs) and KF-6012 (0.02%) in 1×PCR MM. Oil (1200 μL, 4/20/76, KF-6038/mineral oil/DMF-A-6CS, w/w/w) was added to the top of each aqueous reaction mix. The reaction as emulsified (10 s, 15 Hz; 10 s, 17.1 Hz) using a bead mill homogenizer (TissueLyzer, Qiagen). Using a wide bore pipet tip, aliquots (50 μL) were transferred and thermally cycled ([95 °C 20 s, 68 °C 90 s] × 30 cycles, 68 °C 5 min).

Emulsion clean up.

After amplification, breaking buffer (50 μL) was added to each PCR well to pool aliquots to a single tube. Each well was rinsed with an additional aliquot of breaking buffer (50 μL). Beads were harvested by centrifugation (3000 × g, 5 min). Supernatant was removed and beads washed with breaking buffer (3 × 1 mL). Beads were transferred to a clean centrifuge tube, rinsed with breaking buffer (1 mL), resuspened in bead buffer (1 mL) and counted on a hemacytometer.

Library characterization.

Beads were diluted (average 0.05 bead/μL) and quantitated via Taqman qPCR assay (20 μL assay volume). A standard curve was prepared using serial dilutions of the template, adding a constant volume (1 μL) to each standard reaction (100 pg - 0.1 fg in logs). The number of templates per bead was calculated from the standard curve for all wells that amplify. The number of wells without amplification signal was counted; excess wells over the expected number of empty wells based on the Poisson distribution ($\lambda = 1$ bead/well) were assumed to contain untemplated beads. The fraction of untemplated beads was then used to calculate an estimated λ for the beads in the emPCR mix, providing an estimate of monoclonality.

emIVTT.

A mixture of R110-forward screening probe heteroconjugate and Cy5-counter-screening probe heteroconjugate (1 μM each in 2× SSC) was prepared. $5 \times 10^6 - 1 \times 10^7$ beads are

incubated in the probe mix (30 μL, 5 min, RT). Supernatant was removed on a magnet and the beads washed (2 × 100 μL 2× SSC, followed by 2 × 100 μL 1× MM).

100 μL IVTT reactions (PURExpress with Protein Disulfide Bond Enhancers, NEB) were prepared in tubes (2 mL SafeLock, Eppendorf) on ice with a steel bead (6 mm). Beads (5×10^6 — 1×10^7) were resuspended in reagent A (40 μL), vortexed, combined with bait (2 pmol), DSBE 1 (4 μL), DSBE 2 (4 μL), enterokinase light chain (0.5 μL) and nanopure water (to 70 μL). Immediately before emulsification, reagent B (30 μL) was added to the aqueous mix, then the reaction covered with oil mix (1200 μL, 4/20/76, KF-6038/mineral oil/DMF-A-6CS, w/w/w). Samples were emulsified on a TissueLyzer bead mill homogenizer (15.1 Hz 10s, 17.1 Hz 10s). Emulsions were immediately transferred to heat block and incubated (17 h, 37 °C) in the dark.

Bulk oil was removed from top of emulsions after settling overnight. Breaking buffer (500 μL) was added and samples transferred to a clean tube (1.5 mL). Incubation tubes were rinsed with additional breaking buffer (500 μL) and samples centrifuged (5 min, 3000 × g). Supernatant was removed on a magnet and beads washed (breaking buffer, 2 × 500 μL) before transfer to clean tubes. Supernatant was removed and beads rinsed (100 μL 1× PCR MM). Beads were resuspended in (50 μL, 1× PCR MM) and treated with RNaseA (1 μL, 30 min, 37 °C). Supernatant was removed on magnet and beads washed with breaking buffer (200 μL) before resuspension (PBS).

FACS analysis.

Samples were sorted (FACSJazz, BD BioSciences, San Jose, CA) after calibration with rainbow and Accudrop beads (BD Biosciences). An initial gate was set using forward and side scatter to collect only the single bead population. A sub-gate was then set in the FAM channel to collect the high fluorescence population. A sub-gate was then set from this population to collect high fluorescence in the Cy5 channel. Beads in this final gate were sorted into aliquots of qPCR mix (20 μL) in a 96-well plate at either 1 or 10 beads per well.

After bead collection, plates were centrifuged (1 min, 1000 × g) and thermally cycled (CFX96, [95 °C 15 s, 68 °C 80 s] × 40 cycles). Sample wells with observable signal were processed on MinElute columns (Qiagen) to isolate the sorted DNA populations.

RT-IVTT assays.

Activity of the starting scaffold, libraries and selected populations were assessed using in vitro expression of the proteins with solution phase R110 probes. Reactions were prepared using the PURExpress in vitro expression kit supplemented with disulfide bonds enhancers, EK light chain (67 fg/μL) and rhodamine probe (2 μM). PCR product template (1 ng) was used per reaction (5 μL). Samples were incubated (37 °C) in a real-time qPCR instrument (CFX96, BioRad), and channel 1 monitored (1 min intervals).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgement

This work was funded by the National Institutes of Health (R01GM120491).

References

(1). Tawfik DS; Griffiths AD Man-made cell-like compartments for molecular evolution. Nature Biotechnology 1998, 16, 652–656.

(2). Sepp A; Tawfik DS; Griffiths AD Microbead display by in vitro compartmenta-lisation: selection for binding using flow cytometry. FEBS Letters 2002, 532, 455–458. [PubMed: 12482612]

(3). Agresti JJ; Kelly BT; Jäschke A; Griffiths AD Selection of ribozymes that catalyse multiple-turnover Diels-Alder cycloadditions by using in vitro compartmentalization. Proceedings of the National Academy of Sciences of the United States of America-Biological Sciences 2005, 102, 16170–16175.

(4). Huebner A; Sharma S; Srisa-Art M; Hollfelder F; Edel JB; deMello AJ Microdroplets: a sea of applications? Lab on a Chip 2008, 8, 1244–1254. [PubMed: 18651063]

(5). Baret J-C; Miller OJ; Taly V; Ryckelynck M; El Harrak A; Frenz L; Rick C; Samuels ML; Hutchison JB; Agresti JJ; Link DR; Weitz DA; Grif-fiths AD Fluorescence-activated droplet sorting (FADS): efficient microfluidic cell sorting based on enzymatic activity. Lab on a Chip 2009, 9, 1850–1858. [PubMed: 19532959]

(6). Paegel BM; Joyce GF Microfluidic compartmentalized directed evolution. *Chemistry &* Biology 2010, 17, 717–724.

(7). Mazutis L; Gilbert J; Ung WL; Weitz DA; Griffiths AD; Heyman JA Single-cell analysis and sorting using droplet-based microfluidics. Nature Protocols 2013, 8, 870–891. [PubMed: 23558786]

(8). Larsen AC; Dunn MR; Hatch A; Sau SP; Youngbull C; Chaput JC A general strategy for expanding polymerase function by droplet microfluidics. Nature Communications 2016, 7, 11235.

(9). Kim SC; Clark IC; Shahi P; Abate AR Single-cell RT-PCR in microfluidic droplets with integrated chemical lysis. Analytical Chemistry 2018, 90, 1273–1279. [PubMed: 29256243]

(10). Agresti JJ; Antipov E; Abate AR; Ahn K; Rowat AC; Baret J-C; Marquez M; Klibanov AM; Griffiths AD; Weitz DA Ultrahigh-throughput screening in drop-based microfluidics for directed evolution. Proceedings of the National Academy of Sciences of the United States of America 2010, 107, 4004–4009. [PubMed: 20142500]

(11). Miller OJ; Harrak AE; Mangeat T; Baret J-C; Frenz L; Debs BE; Mayot E; Samuels ML; Rooney EK; Dieu P; Galvan M; Link DR; Griffiths AD High-resolution dose–response screening using droplet-based microfluidics. Proceedings of the National Academy of Sciences of the United States of America 2012, 109, 378–383. [PubMed: 22203966]

(12). Vallejo D; Nikoomanzar A; Paegel BM; Chaput JC Fluorescence-activated droplet sorting for single-cell directed evolution. ACS Synthetic Biology 2019, 1–11.

(13). Bernath K; Hai M; Mastrobattista E; Griffiths AD; Magdassi S; Tawfik DS In vitro compartmentalization by double emulsions: Sorting and gene enrichment by fluorescence activated cell sorting. Analytical Biochemistry 2004, 325, 151–157. [PubMed: 14715296]

(14). Mastrobattista E; Taly V; Chanudet E; Treacy P; Kelly BT; Griffiths AD High-throughput screening of enzyme libraries: in vitro evolution of a β-Galactosidase by fluorescence-activated sorting of double emulsions. Chemistry & Biology 2005, 12, 1291–1300. [PubMed: 16356846]

(15). Romanowsky MB; Abate AR; Rotem A; Holtze C; Weitz DA High throughput production of single core double emulsions in a parallelized microfluidic device. Lab on a Chip 2012, 12, 802. [PubMed: 22222423]

(16). Sukovich DJ; Kim SC; Ahmed N; Abate AR Bulk double emulsification for flow cytometric analysis of microfluidic droplets. Analyst 2017,

(17). Brower KK; Carswell-Crumpton C; Klemm S; Cruz B; Kim G; Calhoun SGK; Nichols L; Fordyce PM Double emulsion flow cytometry with high-throughput single droplet isolation and nucleic acid recovery. Lab Chip 2020, 46, 114004–13.

(18). Tran DT; Cavett VJ; Dang VQ; Torres HL; Paegel BM Evolution of a mass spectrometry-grade protease with PTM-directed specificity. Proceedings of the National Academy of Sciences of the United States of America 2016, 113, 14686–14691. [PubMed: 27940920]

(19). Diehl F; Li M; He Y; Kinzler KW; Vogelstein B; Dressman D BEAMing: single-molecule PCR on microparticles in water-in-oil emulsions. Nature Methods 2006, 3, 551–559. [PubMed: 16791214]

(20). Dressman D; Yan H; Traverso G; Kinzler KW; Vogelstein B Transforming single DNA molecules into fluorescent magnetic particles for detection and enumeration of genetic variations. Proceedings of the National Academy of Sciences of the United States of America 2003, 100, 8817–8822. [PubMed: 12857956]

(21). Malone ML; Cavett VJ; Paegel BM Chemoselective coupling preserves the substrate integrity of surface-immobilized oligonucleotides for emulsion PCR-based gene library construction. ACS Combinatorial Science 2017, 19, 9–14. [PubMed: 28064476]

(22). Zhang J; Chung T; Oldenburg K A simple statistical parameter for use in evaluation and validation of high throughput screening assays. Journal of Biomolecular Screening 1999, 4, 67–73. [PubMed: 10838414]

(23). Leytus SP; Patterson WL; Mangel WF New Class of sensitive and selective fluorogenic substrates for serine proteinases. Biochemical Journal 1983, 215, 253–260. [PubMed: 6228222]

(24). Leytus SP; Melhado LL; Mangel WF Rhodamine-based compounds as fluorogenic substrates for serine proteinases. Biochemical Journal 1983, 209, 299–307. [PubMed: 6342611]

(25). Zerfas BL; Trader DJ Monitoring the immunoproteasome in live cells using an activity-based peptide–peptoid hybrid probe. Journal of the American Chemical Society 2019, 141, 5252–5260. [PubMed: 30862160]

(26). Evnin LB; Vásquez JR; Craik CS Substrate specificity of trypsin investigated by using a genetic selection. Proceedings of the National Academy of Sciences of the United States of America 1990, 87, 6659–6663. [PubMed: 2204062]

(27). Craik CS; LARGMAN C; Fletcher T; ROCZNIAK S; Barr PJ Redesigning trypsin: Alteration of substrate specificity. Science 1985,

(28). Hedstrom L; Szilagyi L; Rutter WJ Converting trypsin to chymotrypsin - the role of surface loops. Science 1992, 255, 1249–1253. [PubMed: 1546324]

(29). Craik CS; Roczniak S; Largman C; Rutter WJ The catalytic role of the active-site Aspartic-acid in serine proteases. Science 1987, 237, 909–913. [PubMed: 3303334]

(30). Toscano MD; Woycechowsky KJ; Hilvert D Minimalist active-site redesign: teaching old enzymes new tricks. Angewandte Chemie International Edition 2007, 46, 3212–3236. [PubMed: 17450624]

(31). Tsiatsiani L; Heck AJR Proteomics beyond trypsin. The FEBS journal 2015, 282, 2612–2626. [PubMed: 25823410]

(32). Lavis LD; Chao T-Y; Raines RT Fluorogenic label for biomolecular imaging. ACS Chem. Biol 2006, 1, 252–260. [PubMed: 17163679]

(33). Jetson RR; Krusemark CJ Sensing enzymatic activity by exposure and selection of DNA-encoded probes. Angew. Chem 2016, 128, 9714–9718.

(34). Clark MA et al. Design, synthesis and selection of DNA-encoded small-molecule libraries. Nature Chemical Biology 2009, 5, 647–654. [PubMed: 19648931]

(35). Goodnow RA; Dumelin CE; Keefe AD DNA-encoded chemistry: enabling the deeper sampling of chemical space. Nature Reviews Drug Discovery 2017, 16, 131–147. [PubMed: 27932801]

(36). Malone ML; Paegel BM What is a “DNA-compatible” reaction? ACS Combinatorial Science 2016, 18, 182–187. [PubMed: 26971959]

(37). Ghadessy FJ; Ong JL; Holliger P Directed evolution of polymerase function by compartmentalized self-replication. Proceedings of the National Academy of Sciences of the United States of America 2001, 98, 4552–4557. [PubMed: 11274352]

(38). Laos R; Shaw R; Leal NA; Gaucher E; Benner S Directed evolution of polymerases to accept nucleotides with nonstandard hydrogen bond patterns. Biochemistry 2013, 52, 5288–5294. [PubMed: 23815560]

(39). Seeman NC DNA in a material world. Nature 2003, 421, 427–431. [PubMed: 12540916]

(40). MacConnell AB; McEnaney PJ; Cavett VJ; Paegel BM DNA-Encoded solidphase synthesis: Encoding language design and complex oligomer library synthesis. ACS Combinatorial Science 2015, 17, 518–534. [PubMed: 26290177]

(41). Mendes KR; Malone ML; Ndungu JM; Suponitsky-Kroyter I; Cavett VJ; McEnaney PJ; MacConnell AB; Doran TM; Ronacher K; Stanley K; Utset O; Walzl G; Paegel BM; Kodadek T High-throughput identification of DNA-encoded IgG ligands that distinguish active and latent Mycobacterium tuberculosis infections. ACS Chemical Biology 2017, 12, 234–243. [PubMed: 27957856]

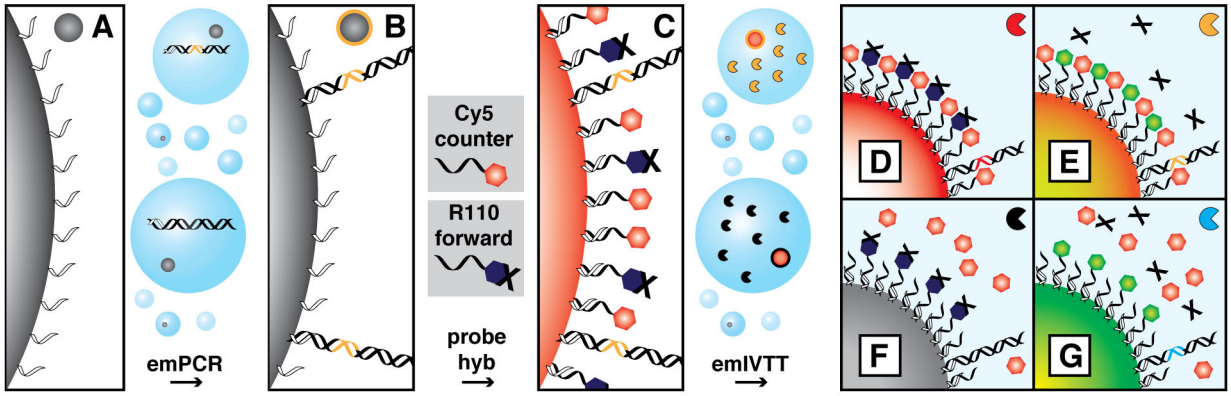


Figure 1:

Multiplexed activity-based screening workflow. **(A)** Magnetic beads (2.8 μm) are chemoselectively functionalized with DNA oligonucleotide PCR primer (white, ~ 2×10^6 primers immobilized per bead) and compartmentalized in emPCR with a limiting dilution of mutant library. **(B)** After emPCR, the bead clonally displays ~1k mutant gene copies. The majority of primers remain unextended. Different colors of DNA sequence represent different mutant library members, shown schematically as a color halo. Probes are conjugated to the oligonucleotide complement (black) of the bead-bound PCR primer. **(C)** The heteroconjugate is hybridized to the bead. Forward screening probes report the target proteolytic activity (e.g., a quenched R110 probe, purple hexagon). Counter screening probes (e.g., a linear peptide with terminal Cy5 fluorophore, bright red hexagon) report off-target activity. The probe-gene beads are emulsified in IVTT for high-throughput expression and cleavage challenge, and harvested for FACS. Four types of two-color fluorescence profiles result from the population of protease mutants. **(D)** Inactive proteases (red gene) do not cleave the R110 or counter screening probes, maintaining the red-fluorescent bead produced upon hybridization to the quenched R110 forward screening probe and fluorescent Cy5 counter screening probe. **(E)** Active proteases with only the target activity (yellow gene) cleave and dequench the R110 probe, but not the Cy5 counter screening probe, producing a green-and red-fluorescent bead. **(F)** Promiscuous proteases lacking the target activity (black) do not cleave the R110 probe; cleavage at any residue in the Cy5 counter-screening peptide results in non-fluorescent beads. **(G)** Promiscuous proteases that also possess the target activity cleave and dequench the R110 probe and cleave Cy5 from the bead, producing a green-fluorescent bead.

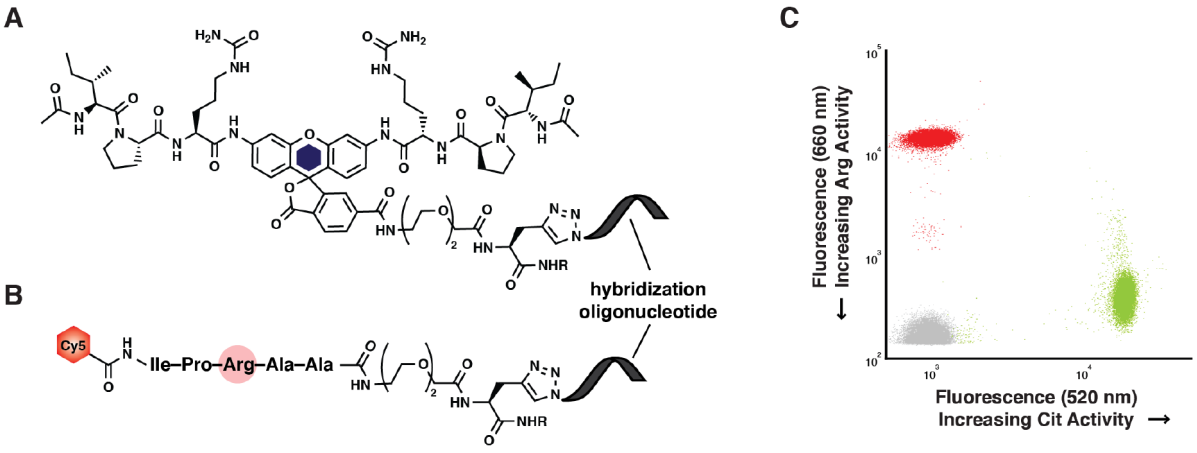


Figure 2:

Multiplexed probe hybridization-based on-bead assay development. **(A)** The $(IPCit)_2R110$ bisamide forward screening probe contains citrulline at the P1 position adjacent the R110 core. The probe is conjugated via CuAAC to an oligonucleotide for immobilization via hybridization to the bead surface. Cleavage C-terminal to citrulline dequenches the R110 core, increasing bead fluorescence emission at 520 nm. **(B)** A counter-screening probe is constructed with the same peptide sequence context as the forward screening $(IPCit)_2R110$ probe, substituting Arg for Cit (red highlighting) and appending an N-terminal Cy5 fluorophore. Cleavage at Arg or any other off-target amino acid releases the Cy5 fluorophore from the probe, attenuating bead fluorescence emission at 660 nm. **(C)** Flow cytometry analysis was performed for beads without probes (gray), beads hybridized with forward and counter screening probes (red), and beads digested with Arg- and Cit-dependent $Trypsin^{+cit}$ protease (green).

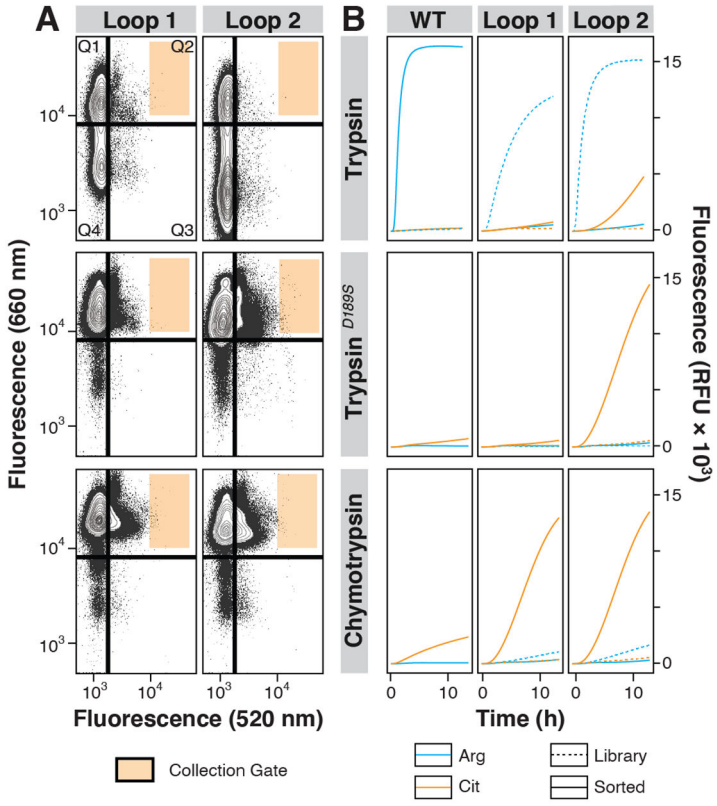


Figure 3:

Library activity profiles and screening hit pool analysis. **(A)** Library activity profiles of trypsin, trypsin^{D189S}, or chymotrypsin (loop 1 or loop 2 site saturation mutagenesis) revealed each library's functional density of Cit-dependent and Arg-dependent proteolytic activity. Increasing 520 nm fluorescence indicates an increase of citrulline-dependent activity. Decreasing 660 nm fluorescence indicates increasing Arg-dependent activity. Each FACS scatterplot contains two bisecting lines that are drawn 3σ from the mean of the probe-hybridized libraries analyzed prior to emIVTT, and a hit collection gate (orange) that is drawn based on the mean signal of positive control beads hybridized with fluorescent complementary oligonucleotide. The trypsin loop 1 plot displays labels for the four quadrants: Q1 (inactive proteases and empty beads), Q2 (target-dependent proteases), Q3 (promiscuous proteases), and Q4 (off-target proteases). **(B)** RT-IVTT activity assay traces of the wild type (WT), loop 1, or loop 2 mutant libraries of trypsin, trypsin^{D189S}, or chymotrypsin display probe proteolysis reaction progress. Assays used either $(acIPR)_2R110$ (cyan) or $(acIPCit)_2R110$ (orange) probes for beads obtained either from the unsorted library (dashed) or gated hit collection in FACS (solid).

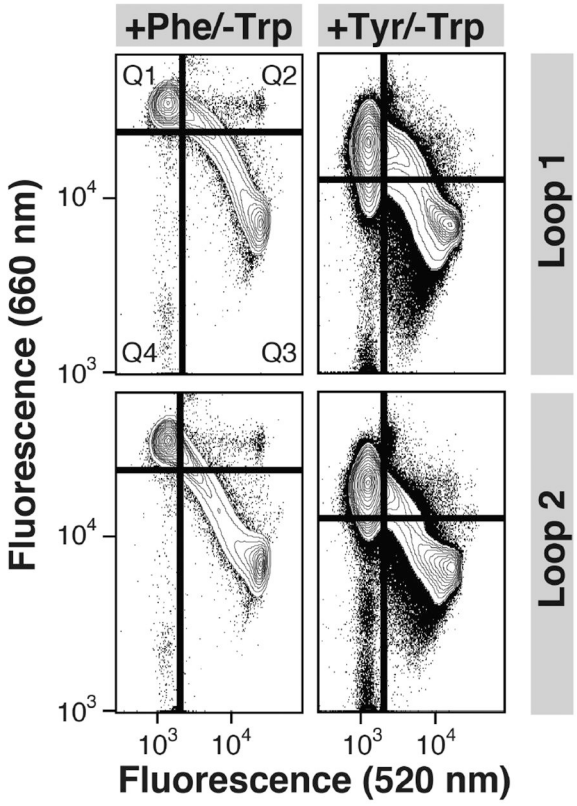


Figure 4:

Chymotrypsin proteolytic specificity sculpting. Library activity profiles of chymotrypsin (loop 1 or loop 2 site saturation mutagenesis), reveal each library's functional density of Phe- or Tyr-specific activity (Q2) in comparison to Trp-specific activity (Q4). Increasing 520 nm fluorescence indicates an increase of either Phe- or Tyr-specific activity. Decreasing 660 nm fluorescence indicates increasing Trp-specific activity. Bisecting lines are drawn 3σ from the mean of the probe-hybridized libraries analyzed prior to emIVTT.

Table 1:

Forward (Cit-dependent) functional density, ρ_{FWD}, is Q2/(Q2+Q3+Q4), promiscuous (Cit- and Arg-dependent) functional density, ρ_{prom}, is Q3/(Q2+Q3+Q4), and the counter (Arg-dependent) functional density, ρ_{CTR}, is Q4/(Q2+Q3+Q4).

Scaffold	Loop Library	ρ_{FWD} (%)	ρ_{prom} (%)	ρ_{CTR} (%)
Trypsin	1	5.1	2.1	93
Trypsin	2	0.60	1.2	98
Trypsin^{D189S}	1	82	0.38	18
Trypsin^{D189S}	2	67	2.3	31
Chymotrypsin	1	91	0.47	8.5
Chymotrypsin	2	94	0.69	4.6