

MIT Open Access Articles

Xenoprotein engineering via synthetic libraries

The MIT Faculty has made this article openly available. **Please share** how this access benefits you. Your story matters.

Citation: Gates, Zachary P. et al. "Xenoprotein Engineering via Synthetic Libraries." Proceedings of the National Academy of Sciences 115, 23 (May 2018): E5298–E5306 © 2018 National Academy of Sciences

As Published: <http://dx.doi.org/10.1073/PNAS.1722633115>

Publisher: National Academy of Sciences (U.S.)

Persistent URL: <http://hdl.handle.net/1721.1/120509>

Version: Final published version: final published article, as it appeared in a journal, conference proceedings, or other formally published context

Terms of Use: Article is made available in accordance with the publisher's policy and may be subject to US copyright law. Please refer to the publisher's site for terms of use.



Xenoprotein engineering via synthetic libraries

Zachary P. Gates^{a,1}, Alexander A. Vinogradov^a, Anthony J. Quartararo^a, Anupam Bandyopadhyay^a, Zi-Ning Choo^a, Ethan D. Evans^a, Kathryn H. Halloran^a, Alexander J. Mijalis^a, Surin K. Mong^a, Mark D. Simon^a, Eric A. Standley^a, Evan D. Styduhar^a, Sarah Z. Tasker^a, Faycal Touti^a, Jessica M. Weber^a, Jessica L. Wilson^a, Timothy F. Jamison^a, and Bradley L. Pentelute^{a,1}

^aDepartment of Chemistry, Massachusetts Institute of Technology, Cambridge, MA 02139

Edited by David Baker, University of Washington, Seattle, WA, and approved April 20, 2018 (received for review December 29, 2017)

Chemical methods have enabled the total synthesis of protein molecules of ever-increasing size and complexity. However, methods to engineer synthetic proteins comprising noncanonical amino acids have not kept pace, even though this capability would be a distinct advantage of the total synthesis approach to protein science. In this work, we report a platform for protein engineering based on the screening of synthetic one-bead one-compound protein libraries. Screening throughput approaching that of cell surface display was achieved by a combination of magnetic bead enrichment, flow cytometry analysis of on-bead screens, and high-throughput MS/MS-based sequencing of identified active compounds. Direct screening of a synthetic protein library by these methods resulted in the de novo discovery of mirror-image miniprotein-based binders to a ~150-kDa protein target, a task that would be difficult or impossible by other means.

xenoprotein | mirror-image miniprotein | D-protein | protein engineering | flow cytometry

Xenoproteins—protein molecules composed of noncanonical amino acids—might exhibit function not readily achieved by the use of proteogenic amino acids alone (1–9) and other favorable properties such as protease stability or altered immunogenicity (10). Chemical protein synthesis is a powerful approach for incorporating a virtually unlimited variety of noncanonical amino acids into a protein molecule for biophysical and structure–function studies (11–20). However, no experimental approach exists to differentiate beneficial versus deleterious mutations in a synthetic protein molecule with significant throughput. Such an approach would facilitate the discovery of functional xenoproteins, since incorporation of noncanonical amino acids into protein molecules is frequently deleterious (21–23).

Identification of functional variants from synthetic combinatorial libraries is a potential solution to the challenge of xenoprotein engineering (24–26). The synthesis of small proteins is straightforward with solid-phase peptide synthesis (SPPS); however, the screening of high-diversity synthetic libraries presents a formidable challenge. Generally, synthetic libraries have been limited to peptide and peptidomimetic structures comprising relatively short polymers with a limited number of varied positions. This is in contrast to molecular biology-based screening and selection strategies, which can routinely examine at least 10^7 – 10^8 variants of large protein molecules, such as antibody fragments, that contain point mutations across the entire polypeptide chain (27).

The one-bead one-compound (OBOC) approach (28) is amenable to the preparation of high-diversity synthetic libraries, provided that sufficiently small resin beads are employed. For example, 1 g of 10- μ m TentaGel resin contains $\sim 2 \times 10^9$ beads. The screening of at least $\sim 3 \times 10^6$ compounds by the OBOC approach has been reported (29); however, typical screens examine fewer than 10^6 compounds (30). The challenge of handling high-diversity OBOC libraries can be attributed to the labor associated with manually screening large numbers of beads (31) and to de novo sequencing of active compounds using the limited

amount of material present on 30- or 10- μ m beads (4 pmol or 100 fmol, respectively, for amine loadings of 0.2 mmol/g).

We set out to render practical the screening of high-diversity synthetic libraries and to explore the use of synthetic protein libraries for engineering de novo binding activity into a mirror-image miniprotein molecule (MIM). This task was chosen to illustrate the utility of chemical synthesis for xenoprotein engineering, since a MIM is composed entirely of noncanonical amino acids, with the exception of glycine. The overall approach is shown in Fig. 1. Key steps are the synthesis of folded protein variants bound to 30- μ m beads, high-throughput analysis of on-bead screens by flow cytometry, and de novo sequencing of identified active compound mixtures by a recently developed liquid chromatography/tandem mass spectrometry (LC-MS/MS) approach (32). These steps were individually optimized and then combined to discover MIM-based binders to a monoclonal antibody target. The results constitute important proof of concept for the synthetic library approach to xenoprotein engineering, which should in principle be applicable to the discovery of functional xenoproteins based on a variety of other small protein scaffolds.

Results and Discussion

EETI-II Is a Robust Scaffold for the Display of Chemical Diversity. As a molecular scaffold (33) for the generation of a mirror-image protein library we selected the trypsin inhibitor from *Ecballium elaterium*, EETI-II (Fig. 2). This 29-residue protein molecule is

Significance

Combinatorial protein libraries—prepared via molecular biology-based approaches—are invaluable tools for protein engineering. The inclusion of noncanonical amino acids in such libraries is of considerable interest. However, at present no approach competes with chemical synthesis in terms of the variety and number of noncanonical amino acids that can be simultaneously incorporated into a protein molecule. Here, we describe selection from synthetic libraries as a strategy for protein engineering. The approach enables identification of small (~30 aa), functional protein variants comprising a virtually unlimited variety of noncanonical amino acids. Increasing the throughput of synthetic library screening, which was achieved through this effort, is anticipated to improve the utility of synthetic libraries for identifying polypeptide-based ligands with de novo function.

Author contributions: Z.P.G., A.A.V., A.J.Q., A.B., E.D.E., M.D.S., T.F.J., and B.L.P. designed research; Z.P.G., A.A.V., A.J.Q., A.B., Z.-N.C., E.D.E., K.H.H., A.J.M., S.K.M., M.D.S., E.A.S., E.D.S., S.Z.T., F.T., J.M.W., J.L.W., and B.L.P. performed research; Z.P.G., A.A.V., A.J.Q., A.B., and B.L.P. analyzed data; and Z.P.G., A.A.V., and B.L.P. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Published under the PNAS license.

¹To whom correspondence may be addressed. Email: zgates@mit.edu or blp@mit.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1722633115/-DCSupplemental.

Published online May 21, 2018.

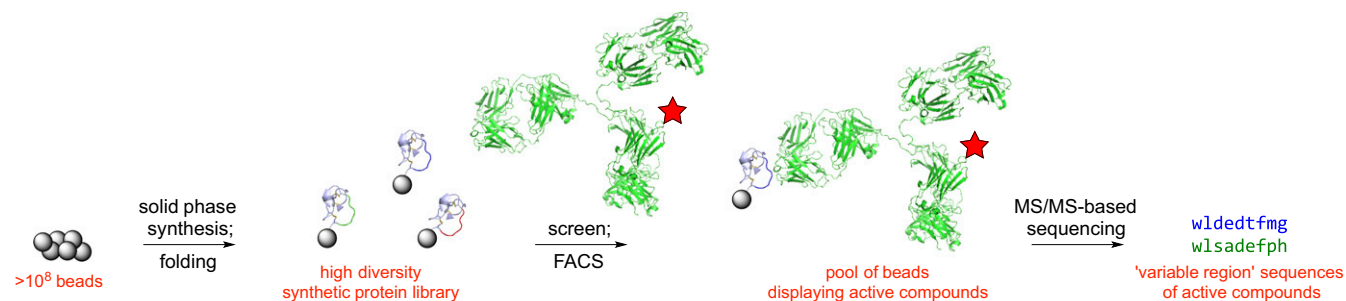


Fig. 1. An approach to identifying binders from synthetic protein libraries. Thirty-micrometer beads displaying xenoprotein variants are prepared by a combination of stepwise SPPS and on-bead folding. The resulting beads are incubated with a protein target bearing a fluorescent label (red star), and beads displaying functional xenoprotein variants are isolated by fluorescence-activated sorting (FACS). The sequences of xenoproteins contained on sorted beads are then determined by de novo MS/MS-based peptide sequencing.

amenable to chemical synthesis (34), is known to oxidatively fold spontaneously despite sequence variation of the trypsin-binding loop (23, 35–38), and has been successfully engineered to bind $\alpha_v\beta_3$ and $\alpha_v\beta_5$ integrins based on the Arg–Gly–Asp motif using yeast-surface display (39, 40). The broad tolerance of the EETI-II molecule to loop expansion and sequence variation (38) is typical of cystine knot proteins generally (41), and we anticipated that this feature could be leveraged to prepare synthetic libraries of folded variants for the selection of novel binding proteins. To confirm this important property of the EETI-II molecule, we tested the ability of a number of synthetic EETI-II variants, containing nonnative loop sequences, to fold spontaneously (Fig. 2C and *SI Appendix, section 3*). In all cases (20 shown), a single major product containing three disulfide bonds was formed. These data suggest that in a combinatorial library of EETI-II variants, many will be present as folded cystine knots, which likely retain the disulfide connectivity and general tertiary structure of the native EETI-II molecule (37, 42).

The susceptibility of L-polypeptides to protease digestion is a major drawback to their use *in vivo*, which has motivated the development of binding molecules based on D-polypeptides (43, 44) and D-proteins (10). Cystine knot proteins sometimes exhibit protease resistance, but this is not general (45). To demonstrate a significant advantage of D-protein molecules, we investigated the stability of several mirror-image EETI-II variants to proteinase K. Minimal degradation was observed over a period of 24 h, whereas the corresponding L-proteins survived less than 1 h under identical conditions (Fig. 2D and *SI Appendix, section 4*). Taken together, these data support the notion that mirror-image EETI-II is a robust scaffold for the generation of binding molecules, which may be of practical utility in biotechnology.

Protein-Based Fluorophores Enable Flow Cytometry Analysis of On-Bead Assays. Flow cytometry is a powerful technique for high-throughput analysis of cell-surface display libraries (46–48), and we sought to adopt its use for bead-based libraries, to increase screening throughput (49–51). An initial goal was to maximize the fluorescence contrast between beads displaying a known protein ligand versus beads displaying nonbinding ligands, after incubation with a fluorescently labeled target protein. We used SPPS to prepare 30- μ m TentaGel beads functionalized with the streptavidin (SA)-binding peptide StrepTag II (52) and assayed them for binding to a variety of commercially available SA conjugates alongside samples of beads displaying random peptides. SA conjugates were chosen based on compatibility with 633-nm excitation, to minimize bead autofluorescence, and included SA-Alexa Fluor 633 (SA-AF633), SA-allophycocyanin (SA-APC), and SA-quantum dot 655 (SA-QD).

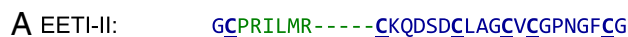
Of the fluorophores examined, only the protein-based APC gave acceptable contrast between StrepTag II beads and library

beads (Fig. 3A and *SI Appendix, section 6*). The comparatively poor performance of AF633 and QD fluorophores was due to a combination of 10- to 15-fold higher background fluorescence of the library beads, and approximately fivefold lower fluorescence of the bound StrepTag II beads. The propensity of highly charged, small-molecule fluorophores such as AF to bind non-specifically to library beads has been noted (53) and may be responsible for the high level of background binding observed with SA-AF633.

SA-APC enabled detection of protein-specific binding to beads functionalized with ligands to a variety of different protein targets. For example, incubation of beads functionalized with either the nine-residue linear epitope (HA epitope) of an anti-hemagglutinin monoclonal antibody (clone 12CA5; anti-HA mAb 12CA5) (Fig. 3B, *Middle*) or an IgG₁ F_c-binding peptide (Fig. 3B, *Right*) (54) with either SA-APC or a mixture of SA-APC and the appropriate biotinylated protein target resulted in target-dependent gains in fluorescence intensity. The ~100-fold difference in fluorescence intensity observed for bound versus unbound beads was comparable to that observed in cell-surface binding assays (55); therefore, we concluded that flow cytometry is suitable for the analysis of on-bead binding assays, and potentially for on-bead screens (discussed below).

On-Bead Binding Assays Are Highly Sensitive. For cell-surface binding assays, the fluorescence intensity of a cell displaying a bound ligand is directly proportional to ligand expression level (47, 55). To account for this effect, which could otherwise lead to the identification of weak binders displayed at high copy number, target binding is normalized based on binding to a coexpressed affinity tag, in a two-color experiment. We sought to determine whether a similar procedure would be required for beads, which reportedly exhibit substantial variation in amine loading (56) (moles of synthetic product per individual bead). To address this question, we prepared biotinylated TentaGel samples of varying biotin loading and assayed them by flow cytometry after incubation with SA-APC.

The fluorescence intensity of bound biotin beads varied only fivefold over a 1,000-fold range of biotin loadings (Fig. 3C), and a similar outcome was obtained for beads functionalized with varying levels of StrepTag II (*SI Appendix, Fig. S38*). These outcomes were not due to fluorescence quenching at high bead loadings: When incubated with mixtures of unlabeled SA and SA-APC, the fluorescence intensity of biotin-functionalized beads varied proportionally to the ratio of fluorescent and unlabeled SA (*SI Appendix, Fig. S39*). We concluded that a strategy to normalize bead fluorescence based on amine loading would not be required for success with bead-based screens, since the fluorescence intensity of a bound bead was largely invariant with bead loading. This result could be understood if the amount of protein target



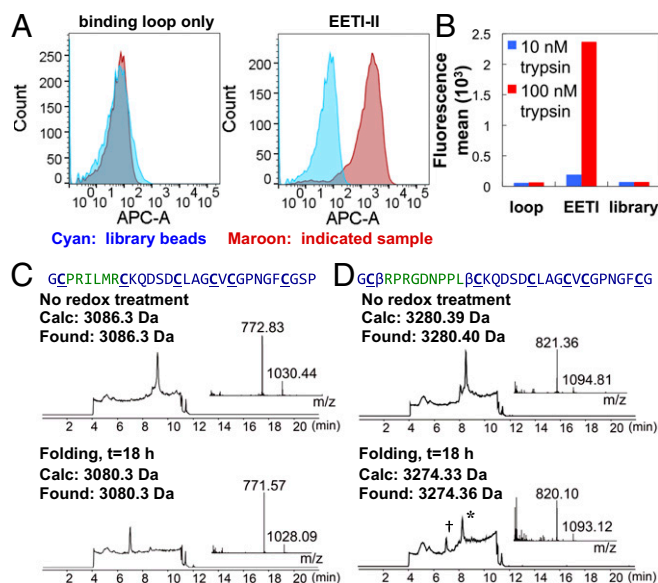


Fig. 4. Folded EETI-II can be prepared on beads. (A) Fluorescence histograms showing the contrast between library beads (cyan) and either EETI-II beads (maroon, *Right*) or trypsin-binding loop beads (maroon, *Left*) after incubation with a mixture of SA-APC (50 nM) and biotinylated trypsin (100 nM). (B) Fluorescence means obtained by treatment of the same samples with a fixed concentration of SA-APC (50 nM) and one of two different concentrations of biotinylated trypsin. Trypsin-dependent binding was observed for the EETI-II beads only. (C) LC-MS data showing the spontaneous oxidative folding of EETI-II while bound to beads. (D) LC-MS data showing the spontaneous oxidative folding of an analog of the engineered EETI-II variant 2.5F, while bound to beads; † denotes a three-disulfide-containing product and * denotes a two-disulfide-containing product. For C and D, the indicated mass values correspond to monoisotopic masses.

or after a folding treatment, we were able to evaluate the outcome of on-bead folding by LC-MS.

On-bead folding conditions effected the conversion of EETI-II to a single three-disulfide-linked product (64), as for the solution-based folding studies (Fig. 4C). To investigate the prospect of preparing folded EETI-II variants with expanded

loops—necessary to accommodate large libraries—we examined an analog of the engineered EETI-II variant 2.5F (39) in the same assay. In this case, the desired three-disulfide product formed in an approximately one-to-three ratio with a two-disulfide intermediate (Fig. 4D). These data support the notion that folded, synthetic EETI-II variants with randomized loop sequences can be prepared on beads. However, the possibility of synthetic coproducts participating in screens must be kept in mind, as for any bead or cell surface-based screen. In the case of libraries based on EETI-II, the two-disulfide folding intermediate may be a common and significant coproduct.

High-Diversity Mirror-Image EETI-II Libraries Were Prepared on Bilayer Beads. Having demonstrated the preparation of beads displaying functional EETI-II, we proceeded to construct a mirror-image EETI-II-based library, for use in screening (Fig. 5). Several considerations factored into the design of this strategy. First was the need to display a low density of synthetic protein in the accessible portion of the bead, while maintaining an adequate amount of material for sequencing. This need suggested the use of spatially segregated resin particles (59), which were prepared by an enzymatic shaving approach (65) (*SI Appendix, section 9*). Second was our inability to reliably assign *de novo* sequences to MS/MS spectra obtained from polypeptides longer than ~15 aa residues. This limitation suggested the use of a “coding structure” of amino acid sequence identical to the varied portion of the MIM, contained in the bead interior, and coupled to a cleavable linker for release postscreening. The coding structure was installed by removal of an Alloc group on the bead interior, after SPPS of the MIM “constant region” (Fig. 5).

Manual split-pool synthesis (28) of the MIM library was performed using 5 g of resin, which comprised more than 2×10^8 individual beads (*SI Appendix, section 10*). For each split, 15 protected D-amino acid monomers were employed (Cys, Arg, Ile, Asn, and Gln were excluded). The use of 15 possible monomers at each of nine varied positions equated to a theoretical diversity of 3.8×10^{10} compounds—roughly 200 times the number of beads used for synthesis. Thus, the totality of possible sequences was undersampled, as for many combinatorial protein libraries (27). We hypothesized that higher diversity would increase the odds of identifying consensus binding sequence motifs

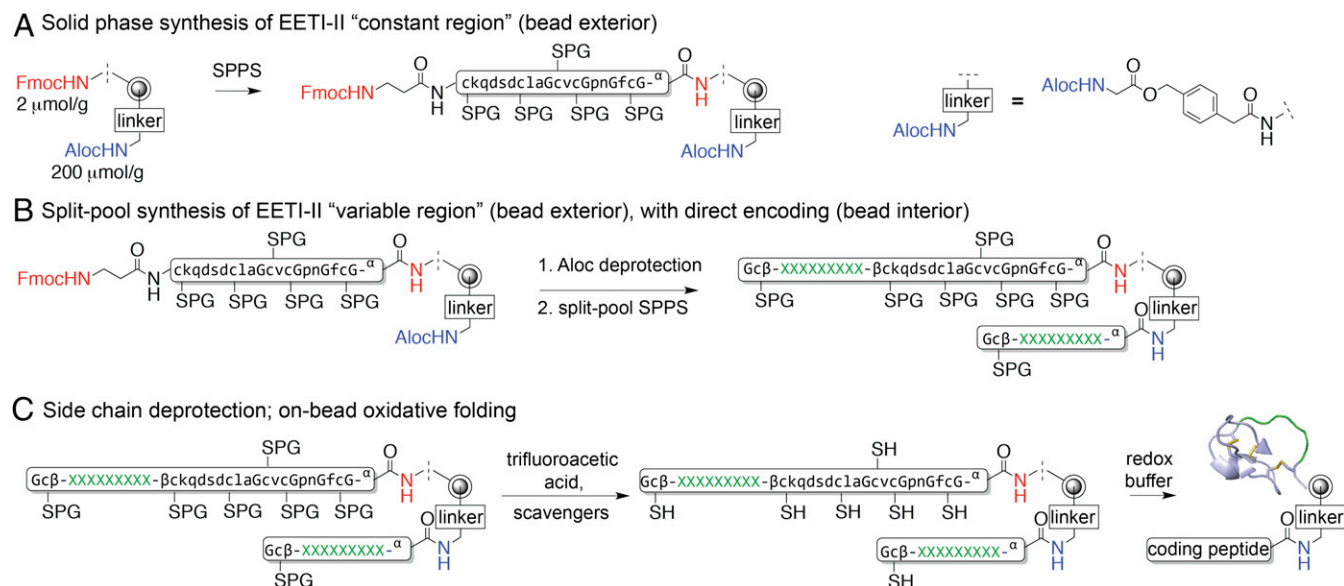


Fig. 5. (A–C) Strategy for the preparation of a mirror-image EETI-II-based library.

Table 1. Binding of protein targets to chemical functionality on beads limits the enrichment attainable by flow cytometry selection

Entry	Bead sample	Assay buffer	Protein target	Sort frequency	Maximum enrichment
1	Underivatized	BSA	—	9.8×10^{-7}	1.0×10^6
2	Library	BSA	—	1.0×10^{-5}	9.6×10^4
3	Underivatized	BSA	12CA5	3.2×10^{-3}	3.1×10^2
4	Underivatized	BSA	Thrombin	1.9×10^{-3}	5.3×10^2
5	Library	BSA	12CA5	6.6×10^{-2}	1.5×10^1
6	Library	BSA	Thrombin	1.9×10^{-1}	5.3
7	Underivatized	FBS	12CA5	3.5×10^{-4}	2.8×10^3
8	Underivatized	FBS	Thrombin	2.7×10^{-4}	3.7×10^3
9	Library	FBS	12CA5	5.7×10^{-4}	1.7×10^3
10	Library	FBS	Thrombin	4.2×10^{-4}	2.3×10^3

Underivatized beads and library beads were incubated with either of two SA-APC-conjugated protein targets, in either of two assay buffers. For each sample, at least 5×10^5 beads were analyzed. The fraction of beads comparably fluorescent to positive control ($>6,000$ counts; sort frequency) is noted for each sample. Entries 1 and 2 reflect averages of two independent experiments. Dash denotes assay buffer only.

without a major penalty, since nonredundant libraries could still be redundant with respect to such motifs.

Protein Target-Dependent Binding Limits the Enrichment Attainable by Fluorescence-Activated Bead Sorting. With mirror-image EETI-II-based libraries in hand, we sought to determine the enrichment that could be achieved by flow cytometry selection for on-bead binding, where enrichment is defined as (positive beads/total beads)_{sorted}/(positive beads/total beads)_{analyzed} (46). Flow cytometry routinely achieves enrichments of $\sim 10,000$ (46, 66) for cell-surface display libraries. However, the sorting of beads is thought to be complicated by high background due to autofluorescence (67). To understand the severity of this issue, and other variables that might affect enrichment, we studied the frequency of sorted beads ($>6,000$ fluorescence counts) for two different bead samples—underivatized beads and library beads—under a variety of assay conditions (Table 1 and *SI Appendix, section 11.1*). Sort frequency sets an upper bound on enrichment, which can be expressed as (total beads_{analyzed}/total beads_{sorted}) \times (positive beads_{sorted}/positive beads_{analyzed}).

Enrichment was limited primarily by the binding of protein targets to chemical functionality on beads. The process of solid-phase synthesis did result in a 10-fold increase in the frequency of autofluorescent beads, as determined by comparison of library beads and underivatized beads in the absence of fluorescent protein (Table 1, entries 1 and 2); however, incubation of underivatized beads with either of two SA-APC conjugates resulted in more than 10^3 -fold gains in sort frequency (Table 1, entries 3 and 4). Library beads displayed a higher degree of protein binding compared with underivatized beads, with sort frequencies 10^3 - to 10^4 -fold above the autofluorescence level (Table 1, entries 5 and 6 vs. entry 2). These sort frequencies corresponded to maximum enrichments of 15 or 5 for anti-HA mAb 12CA5 or thrombin, respectively, which were sufficiently poor as to make sorting untenable.

To minimize nonspecific binding of protein targets, we conducted assays in complex media. For both underivatized beads and library beads, use of FBS buffer resulted in lower sort frequencies compared with BSA alone (Table 1, entries 7–10). In FBS, library beads exhibited only slightly higher sort frequencies compared with underivatized beads, corresponding to maximum enrichment factors of $\sim 2 \times 10^3$. Enrichment factors on this order are close to those obtained for the sorting of bacterial and yeast cells, suggesting that acceptable conditions for on-bead screening had been identified. However, since functional MIM variants were anticipated to be less frequent than one in several thousand, we concluded that a pre-enrichment step would be required for a successful screen. Magnetic bead enrichment was explored, since this procedure is used to reduce the initial size of cell surface display libraries before sorting (68) and is capable of retaining even modest-affinity ($K_d \sim$ micromolar) binders with high yield (69).

A Two-Stage Magnetic Bead Selection/On-Bead Screen Improves Enrichment for Active Beads. We investigated the utility of magnetic bead selection for improving enrichments achieved in screens of the mirror-image EETI-II library against anti-HA mAb 12CA5 and human α -thrombin, respectively. Library beads (400 mg, $\sim 2 \times 10^7$ beads) were incubated with SA-coated magnetic microparticles conjugated to the appropriate biotinylated target protein (*SI Appendix, section 11.3*). Retained library beads were isolated with the goal of maximizing recovery and incubated with the appropriate SA-APC conjugate in secondary screens. Two screens were carried out with anti-HA mAb, to assess reproducibility.

Magnetic bead selections for thrombin or 12CA5 binding retained 1–6% of library beads, corresponding to maximum enrichments of 16–90. Combined with the enrichment obtained by subsequent on-bead screens, overall maximum enrichments of up to 8.2×10^4 were obtained (Table 2). To evaluate whether the actual enrichments achieved were sufficient to identify functional

Table 2. Two-stage magnetic bead selection/on-bead screens can yield overall enrichments approaching 10^5

Entry	Protein target	Beads retained	Maximum enrichment (pull-down)	Sort frequency	Maximum enrichment (overall)
1	12CA5	221,066	90	1.1×10^{-3}	8.2×10^4
2	12CA5	1,049,851	19	1.2×10^{-3}	1.6×10^4
3	Thrombin	1,239,414	16	6.2×10^{-3}	2.6×10^3

For each experiment, 2×10^7 library beads were subjected to magnetic bead selection (pull-down); the number of retained beads was determined by flow cytometry, after on-bead screening with the appropriate SA-APC conjugate. Sort frequencies from the flow cytometry screens are noted, along with associated maximum enrichments for the combined pull-down and flow cytometry steps.

variants, the material present on sorted beads was sequenced as described (32) (*SI Appendix, section 11*).

For the first of two screens against 12CA5, 9 of 16 sequence assignments contained a motif, yp^*e^*d/e , where $*$ is any amino acid (*SI Appendix, Fig. S55*). This was in contrast to assignments from a single-pass screen of 1.5×10^6 beads, which lacked discernable similarity (*SI Appendix, Fig. S53*), suggesting that the magnetic bead selection preferentially retained positive beads. For the replicate 12CA5 screen, seven yp^*e^*d/e -containing sequences were obtained; however, these comprised a smaller fraction of the total (180), consistent with the lower overall enrichment obtained for the replicate screen (*SI Appendix, Figs. S57 and S58*). Likewise, no obvious similarity was discernable among sequences obtained for the thrombin screen, which had the lowest overall enrichment (*SI Appendix, Fig. S60*). These results highlight the importance of the magnetic bead selection and suggest that the protocol has yet to be optimized (70). Significantly higher enrichments can probably be achieved, since magnetic bead selection can enrich yeast cells 10^4 -fold with near-quantitative yield (69).

Sequence Motifs Are Predictive of Reproducible On-Bead Binding. To assess platform fidelity, we synthesized a select group of candidate 12CA5 binders for evaluation in an on-bead binding assay (*SI Appendix, section 12*). This experiment is analogous to the use of cell-surface binding assays to characterize active clones following selection by cell-surface display and would demonstrate the feasibility of our approach for identifying rare beads based on a reproducible screen criterion. Candidate sequences containing the yp^*e^*d motif were selected for evaluation, alongside several unrelated sequences.

With one exception, compounds containing the yp^*e^*d motif gave reproducible, 12CA5-dependent binding, similar to positive control beads functionalized with the HA epitope (Fig. 6, Table 3, and *SI Appendix, Fig. S61*). In contrast, candidate compounds lacking the yp^*e^*d motif were indistinguishable from negative control. These findings demonstrate the potential of our platform for isolating rare beads displaying active compounds and confirm the commonsense premise that emergence of a consensus binding sequence is indicative of a successful screen.

Orthogonal Binding Assays Confirm the Activity of Mirror-Image EETI-II-Based Binders. To confirm the activity of the identified MIM-based 12CA5 binders, select MIMs were synthesized for evaluation in off-bead assays (Fig. 7A and *SI Appendix, section 13*). In a biolayer interferometry assay (BLI), all MIMs tested exhibited a robust binding response to 12CA5 (*SI Appendix, section 14.1*). For three MIMs examined further, strong responses were also observed over a range of 12CA5 dilutions (*SI Appendix, section 14.2*). Individual kinetic traces were fit to a

Table 3. On-bead binding activity of putative 12CA5-binding MIMs

Entry	Sequence (variable region)	On-bead binding?
1	aldyplevd	Yes
2	ypyefdeph	Yes
3	Gypleydsv	Yes
4	ypeekdpys	Yes
5	GGypfevdd	Yes
6	Gypyeidwe	Yes
7	vlhypeeGd	Yes
8	Gypheedy	No
9	yphdpeed	No
10	vspfeypdm	No
11	lseypdwev	No
12	eyfydlkeG	No

Putative 12CA5-binding MIMs—identified from library screening—were synthesized on beads for evaluation in an on-bead binding assay. With one exception (entry 8), compounds that contained the yp^*e^*d motif exhibited 12CA5-dependent binding; compounds lacking the motif did not. For clarity, the “fixed” portions of the miniprotein sequences are omitted.

1:1 binding model, which for compound 1 yielded average values of $k_{on} = 4.5 (\pm 1.9) \times 10^4 \text{ M}^{-1} \text{ s}^{-1}$, $k_{off} = 2.2 (\pm 0.8) \times 10^{-3} \text{ s}^{-1}$, and $K_d = 50 \text{ nM}$ (22–120 nM) (Fig. 7B). Equilibrium binding responses were consistent with the K_d values obtained from kinetic analysis (50% of R_{max} achieved at 22 nM), which provided an elementary test of self-consistency for the BLI data (71).

Binding activity was investigated further using a fluorescence polarization-monitored competition assay (*SI Appendix, section 15*). MIM 1 was evaluated first and competed with fluorescent HA epitope for binding to 12CA5 (Fig. 7C, black trace). Based on the IC_{50} obtained from self-competition with unlabeled HA epitope (red trace), a K_d value of $460 \pm 100 \text{ nM}$ was obtained—~10-fold weaker than the value obtained from BLI. Based on reported discrepancies between the results of surface-based and other binding assays (72–74), and the ability of the fluorescence polarization assay to accurately measure the affinity of fluorescent HA epitope (*SI Appendix, Fig. S79*), we evaluated the remaining MIMs using the competition assay (Table 4 and *SI Appendix, section 15.3*). Affinities ranged from 460 nM to 6.1 μM ; binders containing aromatic residues in the $*$ positions of the sequence motif yp^*e^*d displayed the highest affinity.

“Hot-Spot” Residues Determine Binding Affinity of MIMs for 12CA5.

Identification of a sequence motif associated with 12CA5 binding implies that the identified residues are “hot spots,” since they are conserved among MIM-based binders. The interactions of L-peptide HA epitope Tyr98-Ala108 (YPYDVPDYALA) with anti-HA mAbs are also facilitated by a limited number of hot-spot residues: Asp101, Asp104, and Tyr105 (75–77). The D-polypeptide yp^*e^*d binding motif bears a striking resemblance to the L-peptide epitope, in terms of both sequence similarity (yp^*e vs. YP^*D) and overall side-chain composition. However, the hot-spot residues are different (yp^*e^*d vs. D^*DY). In crystal structures of anti-HA mAbs complexed with HA peptide Tyr100-Ala108, the side chain of Tyr105 is sandwiched between side chains of the mAb H3 loop (78, 79). Energetically, this is the most important interaction in the mAb–peptide complex. For the MIM-based binders, we speculate that the conserved D-tyrosine residue fulfills a similar interaction, and that the D-polypeptide traverses the mAb binding “cleft” formed by loops H2 and H3 in opposite direction of the L-peptide epitope.

To investigate a possible role of constraint in facilitating the binding interaction between 12CA5 and the MIM-based binders, select MIMs were exhaustively reduced/alkylated for evaluation in the fluorescence polarization (FP) assay (compounds 1 and 5;

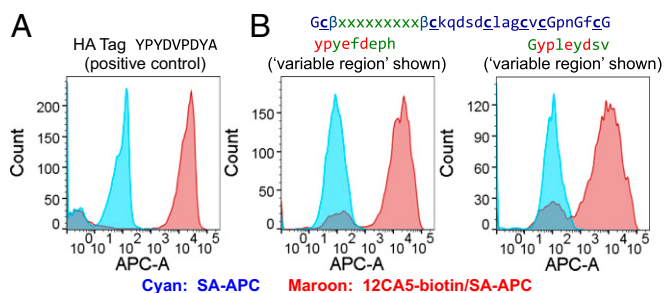


Fig. 6. Replicate preparations of identified active library beads exhibit reproducible on-bead binding. Comparable 12CA5-dependent binding was observed for beads functionalized with either (A) HA epitope or (B) putative 12CA5-binding MIMs.

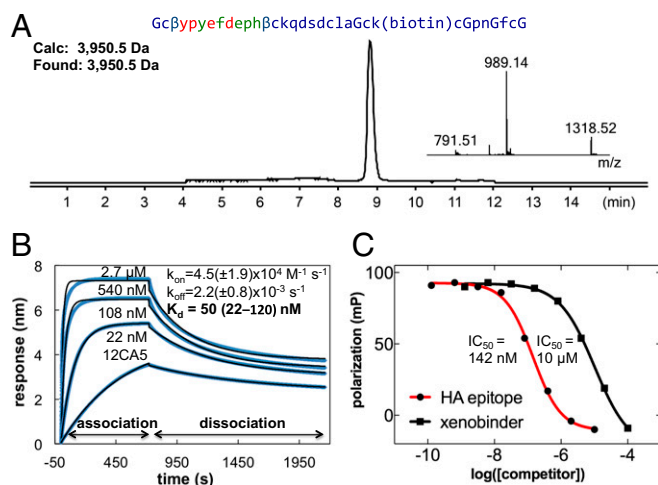


Fig. 7. Orthogonal binding assays confirm the activity of 12CA5-binding MIMs. (A) LC-MS analysis of representative biotinylated 12CA5-binding MIM 1. (B) BLI assay, showing association of exogenous 12CA5 to immobilized MIM 1. (C) FP competition binding assay, showing the displacement of fluorescent HA epitope by unlabeled HA epitope (red) and MIM 1 (black).

SI Appendix, section 16.1). When assayed side by side with the corresponding MIMs containing intact disulfide bonds, these “unfolded” variants exhibited approximately twofold weaker binding affinity (*SI Appendix, section 16.2*). The minimal role of constraint in stabilizing the anti-HA/engineered mirror-image EETI-II interface is in contrast to interactions between engineered cyclic peptide binders and targets such as integrin (80) or SA (60), and for the interaction of an engineered thioredoxin loop with an anti-hapten mAb (81), where constraint significantly improves binding affinity. Constraint may be comparatively unimportant in facilitating binding to mAbs that target linear epitopes, as for the mAb 12CA5 studied here (82). Potentially, the β -Ala residues play a role in offsetting a potential benefit of constraint. Further studies on a wide range of protein targets will be required to evaluate the general utility of constraint for increasing the binding activity of engineered knottins.

Significance. The methods described here form the basis of a platform for the discovery of xenoprotein molecules with de novo binding activity. As illustrated by the identification of mirror-image EETI-II-based “mimotopes” (83), discovery of MIM-based binding molecules is one application of this strategy. Mirror-image peptide and protein-based binding molecules are

of longstanding interest in biotechnology (84). However, to date their discovery by mirror-image phage or cell-surface display has required chemical synthesis of the pertinent target molecule (43, 85, 86), which has restricted the practical size limit of proteins that could realistically be targeted. Direct screening of synthetic MIM libraries is a potential solution to this limitation. Further work will be required to evaluate the general utility of engineered knottins for this application, and of other small protein scaffolds based on helical bundles (87), β -sheets (88), or de novo designs (89–91).

For de novo protein engineering, higher-diversity libraries yield generally superior outcomes. For example, the affinities of antibodies identified by selection from nonimmune libraries improve markedly with increasing library size (66, 92). Thus, extending the throughput by which synthetic libraries can be screened is expected to considerably extend their value. In this work, throughput was achieved by interfacing a multistage selection and screening procedure with a high-throughput LC-MS/MS-based sequencing approach, applicable to large numbers of small beads (32). Both of these steps were necessary to practically explore the utility of synthetic protein libraries for de novo engineering and should facilitate progress in combinatorial chemistry generally. Our work establishes significant parallels between flow cytometry analysis of bead-based (49) versus cell-surface display libraries (46, 47, 93) and highlights the importance of multistage, high-yield selection procedures in the case of beads, since synthetic libraries cannot be propagated and resorted to improve enrichment.

Other uses of synthetic protein libraries may be found. For example, screens for retention of binding could be used to identify positions of an engineered binding protein—identified through a molecular biology-based method—that are compatible with noncanonical amino acids. Such an approach could be useful for modifying the protease stability of biotherapeutics, which is currently achieved through rational design (94). In the future, synthetic libraries may be used to empirically evaluate the success of de novo xenoprotein designs, as has recently been achieved with yeast surface display and de novo designs based on proteogenic amino acids (91, 95). The combination of design and synthetic library screening could prove extremely powerful, with experimental feedback informing design strategy improvements.

Conclusion

The potential of synthetic protein libraries for identifying functional xenoproteins has been demonstrated. By using a combination of magnetic bead enrichment and flow cytometry, the screening of at least 2×10^7 synthetic protein variants was rendered feasible on the timescale of hours. This is a significant

Table 4. Affinity of mirror-image EETI-II-based binders (compounds 1–7) to 12CA5

Competitor	IC ₅₀ (FP)	K _d (FP)	K _d (BLI)
YPYDVDPYA (HA epitope)	174 ± 28 nM	8 nM	—
GcByppefdephβckqdsdclGckcGpnGfcG 1	10 ± 1 μM	460 ± 100 nM	50 nM (22–120 nM)
GcBgyppedydweβckqdsdclGckcGpnGfcG 2	19 ± 1 μM	870 ± 100 nM	—
GcBGGypfevdβckqdsdclGckcGpnGfcG 3	34 ± 1 μM	1.6 ± 0.3 μM	—
GcBgypleydsvβckqdsdclGckcGpnGfcG 4	36 ± 2 μM	1.7 ± 0.4 μM	45 nM (18–140 nM)
GcBvlhypeeGdβckqdsdclGckcGpnGfcG 5	50 ± 2 μM	2.3 ± 0.5 μM	85 nM (36–250 nM)
GcBaldyplevdβckqdsdclGckcGpnGfcG 6	98 ± 8 μM	4.5 ± 1 μM	—
GcByppeekdpysβckqdsdclGckcGpnGfcG 7	165 ± 11 μM	6.1 ± 2 μM	—

All cysteines were present as disulfide bonds. All compounds shown were characterized by FP competition binding assay; select compounds were characterized by both FP and BLI. Uncertainties in IC₅₀ values (FP) are from regression analysis, except for the HA epitope, where the uncertainty is the SD of three independent measurements. Relative K_d values (IC₅₀ values) were converted to K_d values using the experimentally determined K_d of the HA epitope (8 nM). Uncertainties in K_d values were determined by propagation of uncertainties in IC₅₀ values. For BLI measurements, ranges of K_d values are reported to account for scatter in the k_{on} and k_{off} values obtained from individual kinetic traces. Dash denotes not determined.

improvement over state-of-the-art bead-based screening methods, which examine ~100-fold fewer compounds—of smaller molecular weight (~1 kDa)—in a typical screen (51, 96). Throughput could be further improved by the use of even smaller beads, which in principle contain sufficient material for MS/MS-based sequencing (32). With continued development, and as synthetic methods continue to evolve, synthetic protein libraries may become a powerful tool for the discovery of folded, nonnatural polymers with biological function.

Methods

Peptide Synthesis. Thirty-micrometer TentaGel M NH₂ microspheres (M30352) (97) (Rapp Polymere GmbH) and Fmoc chemistry SPPS were employed throughout. For library synthesis, a protocol for Boc-chemistry SPPS (98) was adapted for use with manual Fmoc SPPS. Fmoc deprotection was achieved by treatment with 20% piperidine in dimethylformamide (30 s flow wash; 2 × 3-min batch wash). For the synthesis of defined peptidyl resins, a manual flow-based method for Fmoc SPPS (99) was employed. Methods for the preparation of spatially segregated and attenuated loading resin beads are described in *SI Appendix*.

- Wang L, Schultz PG (2004) Expanding the genetic code. *Angew Chem Int Ed Engl* 44: 34–66.
- Jackson JC, Duffy SP, Hess KR, Mehl RA (2006) Improving nature's enzyme active site with genetically encoded unnatural amino acids. *J Am Chem Soc* 128:11124–11127.
- Ugwumba IN, et al. (2011) Improving a natural enzyme activity through incorporation of unnatural amino acids. *J Am Chem Soc* 133:326–333.
- Liu CC, et al. (2008) Protein evolution with an expanded genetic code. *Proc Natl Acad Sci USA* 105:17688–17693.
- Liu CC, et al. (2009) Evolution of proteins with genetically encoded "chemical warheads". *J Am Chem Soc* 131:9616–9617.
- Young TS, et al. (2011) Evolution of cyclic peptide protease inhibitors. *Proc Natl Acad Sci USA* 108:11052–11056.
- Hammerling MJ, et al. (2014) Bacteriophages use an expanded genetic code on evolutionary paths to higher fitness. *Nat Chem Biol* 10:178–180.
- Kang M, et al. (2014) Evolution of iron(II)-finger peptides by using a bipyridyl amino acid. *ChemBioChem* 15:822–825.
- Xiao H, et al. (2015) Exploring the potential impact of an expanded genetic code on protein function. *Proc Natl Acad Sci USA* 112:6961–6966.
- Uppalapati M, et al. (2016) A potent D-protein antagonist of VEGF-A is non-immunogenic, metabolically stable, and longer-circulating in vivo. *ACS Chem Biol* 11: 1058–1065.
- Baca M, Kent SBH (1993) Catalytic contribution of flap-substrate hydrogen bonds in "HIV-1 protease" explored by chemical synthesis. *Proc Natl Acad Sci USA* 90: 11638–11642.
- Smith R, Brereton IM, Chai RY, Kent SBH (1996) Ionization states of the catalytic residues in HIV-1 protease. *Nat Struct Biol* 3:946–950.
- Romanelli A, Shekhtman A, Cowburn D, Muir TW (2004) Semisynthesis of a segmental isotopically labeled protein splicing precursor: NMR evidence for an unusual peptide bond at the N-extein-intein junction. *Proc Natl Acad Sci USA* 101:6397–6402.
- Wu Z, et al. (2004) Total chemical synthesis of N-myristoylated HIV-1 matrix protein p17: Structural and mechanistic implications of p17 myristoylation. *Proc Natl Acad Sci USA* 101:11587–11592.
- Shogren-Knaak M, et al. (2006) Histone H4-K16 acetylation controls chromatin structure and protein interactions. *Science* 311:844–847.
- Valiyaveetil FI, Leonetti M, Muir TW, Mackinnon R (2006) Ion selectivity in a semi-synthetic K⁺ channel locked in the conductive conformation. *Science* 314:1004–1007.
- Torbeev VY, et al. (2011) Protein conformational dynamics in the mechanism of HIV-1 protease catalysis. *Proc Natl Acad Sci USA* 108:20982–20987.
- Liu Z, et al. (2014) Structure of the branched intermediate in protein splicing. *Proc Natl Acad Sci USA* 111:8422–8427.
- Murakami M, et al. (2016) Chemical synthesis of erythropoietin glycoforms for insights into the relationship between glycosylation pattern and bioactivity. *Sci Adv* 2: e1500678.
- Thompson RE, et al. (2017) Tyrosine sulfation modulates activity of tick-derived thrombin inhibitors. *Nat Chem* 9:909–917.
- Montclare JK, Tirrell DA (2006) Evolving proteins of novel composition. *Angew Chem Int Ed Engl* 45:4518–4521.
- Reinert ZE, Lengyel GA, Horne WS (2013) Protein-like tertiary folding behavior from heterogeneous backbones. *J Am Chem Soc* 135:12528–12531.
- Simon MD, et al. (2016) D-Amino acid scan of two small proteins. *J Am Chem Soc* 138: 12099–12111.
- DeGrado WF, Sosnick TR (1996) Protein minimization: Downsizing through mutation. *Proc Natl Acad Sci USA* 93:5680–5681.
- Bianchi E, et al. (1995) A conformationally homogeneous combinatorial peptide library. *J Mol Biol* 247:154–160.
- Lowe G, Quarrell R (1994) Targeted libraries of CMT-I mutants and the identification of potential serine proteinase inhibitors. *Innovations and Perspectives in Solid Phase Synthesis*, ed Epton R (Mayflower Worldwide Limited, Birmingham, UK), pp 377–380.
- Packer MS, Liu DR (2015) Methods for the directed evolution of proteins. *Nat Rev Genet* 16:379–394.
- Lam KS, et al. (1991) A new type of synthetic peptide library for identifying ligand-binding activity. *Nature* 354:82–84.
- Wu X, Upadhyaya P, Villalona-Calero MA, Briesewitz R, Pei D (2013) Inhibition of Ras-effector interactions by cyclic peptides. *MedChemComm* 4:378–382.
- Obexer R, Walport LJ, Suga H (2017) Exploring sequence space: Harnessing chemical and biological diversity towards new peptide leads. *Curr Opin Chem Biol* 38:52–61.
- Das S, et al. (2015) A general synthetic approach for designing epitope targeted macrocyclic peptide ligands. *Angew Chem Int Ed Engl* 54:13219–13224.
- Vinogradov AA, et al. (2017) Library design-facilitated high-throughput sequencing of synthetic peptide libraries. *ACS Comb Sci* 19:694–701.
- Binz HK, Amstutz P, Plückthun A (2005) Engineering novel binding proteins from nonimmunoglobulin domains. *Nat Biotechnol* 23:1257–1268.
- Le-Nguyen D, Nalis D, Castro B (1989) Solid phase synthesis of a trypsin inhibitor isolated from the Cucurbitaceae *Ecballium elaterium*. *Int J Pept Protein Res* 34: 492–497.
- Le Nguyen D, et al. (1990) Molecular recognition between serine proteases and new bioactive microproteins with a knotted structure. *Biochimie* 72:431–435.
- Christmann A, Walter K, Wentzel A, Krätzner R, Kolmar H (1999) The cystine knot of a squash-type protease inhibitor as a structural scaffold for *Escherichia coli* cell surface display of conformationally constrained peptides. *Protein Eng* 12:797–806.
- Mong SK, et al. (2017) Heterochiral knottin protein: Folding and solution structure. *Biochemistry* 56:5720–5725.
- Lahti JL, Silverman AP, Cochran JR (2009) Interrogating and predicting tolerated sequence diversity in protein folds: Application to E. *Elaterium* trypsin inhibitor-II cystine-knot miniprotein. *PLOS Comput Biol* 5:e1000499.
- Kimura RH, Levin AM, Cochran FV, Cochran JR (2009) Engineered cystine knot peptides that bind α v β 3, α v β 5, and α 5 β 1 integrins with low-nanomolar affinity. *Proteins* 77:359–369.
- Kimura RH, Cheng Z, Gambhir SS, Cochran JR (2009) Engineered knottin peptides: A new class of agents for imaging integrin expression in living subjects. *Cancer Res* 69: 2435–2442.
- Craik DJ, Du J (2017) Cyclotides as drug design scaffolds. *Curr Opin Chem Biol* 38:8–16.
- Kryshtafovich A, et al. (2014) Challenging the state of the art in protein structure prediction: Highlights of experimental target structures for the 10th critical assessment of techniques for protein structure prediction experiment CASP10. *Proteins* 82: 26–42.
- Schumacher TN, et al. (1996) Identification of D-peptide ligands through mirror-image phage display. *Science* 271:1854–1857.
- Garton M, et al. (2018) Method to generate highly stable D-amino acid analogs of bioactive helical peptides using a mirror image of the entire PDB. *Proc Natl Acad Sci USA* 115:1505–1510.
- Werle M, et al. (2006) The potential of cystine-knot microproteins as novel pharmacophoric scaffolds in oral peptide drug delivery. *J Drug Target* 14:137–146.
- Georgiou G, et al. (1997) Display of heterologous proteins on the surface of microorganisms: From the screening of combinatorial libraries to live recombinant vaccines. *Nat Biotechnol* 15:29–34.
- Boder ET, Wittrup KD (1997) Yeast surface display for screening combinatorial polypeptide libraries. *Nat Biotechnol* 15:553–557.
- Chao G, et al. (2006) Isolating and engineering human antibodies using yeast surface display. *Nat Protoc* 1:755–768.
- Needels MC, et al. (1993) Generation and screening of an oligonucleotide-encoded synthetic peptide library. *Proc Natl Acad Sci USA* 90:10700–10704.
- Müller K, et al. (1996) Rapid identification of phosphopeptide ligands for SH2 domains. Screening of peptide libraries by fluorescence-activated bead sorting. *J Biol Chem* 271:16500–16505.

51. Mendes KR, et al. (2017) High-throughput identification of DNA-encoded IgG ligands that distinguish active and latent *Mycobacterium tuberculosis* infections. *ACS Chem Biol* 12:234–243.
52. Schmidt TGM, Koepke J, Frank R, Skerra A (1996) Molecular interaction between the strep-tag affinity peptide and its cognate target, streptavidin. *J Mol Biol* 255:753–766.
53. Jee J-E, et al. (2014) Investigating fluorescent dyes in fluorescence-assisted screenings. *Chem Commun (Camb)* 50:15220–15223.
54. DeLano WL, Ultsch MH, de Vos AM, Wells JA (2000) Convergent solutions to binding at a protein-protein interface. *Science* 287:1279–1283.
55. VanAntwerp JJ, Wittrup KD (2000) Fine affinity discrimination by yeast surface display and flow cytometry. *Biotechnol Prog* 16:31–37.
56. Doran TM, et al. (2014) Utility of redundant combinatorial libraries in distinguishing high and low quality screening hits. *ACS Comb Sci* 16:259–270.
57. Hintersteiner M, Buehler C, Auer M (2012) On-bead screens sample narrower affinity ranges of protein-ligand interactions compared to equivalent solution assays. *ChemPhysChem* 13:3472–3480.
58. Weber PC, Pantoliano MW, Thompson LD (1992) Crystal structure and ligand-binding studies of a screened peptide complexed with streptavidin. *Biochemistry* 31:9350–9354.
59. Wang X, Peng L, Liu R, Xu B, Lam KS (2005) Applications of topologically segregated bilayer beads in 'one-bead one-compound' combinatorial libraries. *J Pept Res* 65:130–138.
60. Giebel LB, et al. (1995) Screening of cyclic peptide phage libraries identifies ligands that bind streptavidin with high affinities. *Biochemistry* 34:15430–15435.
61. Munson MC, Barany G (1993) Synthesis of α -Conotoxin SI, a bicyclic tridecapeptide amide with two disulfide bridges: Illustration of novel protection schemes and oxidation strategies. *J Am Chem Soc* 115:10203–10210.
62. Wentzel A, Christmann A, Krätzner R, Kolmar H (1999) Sequence requirements of the GPNG β -turn of the Ecballium elaterium trypsin inhibitor II explored by combinatorial library screening. *J Biol Chem* 274:21037–21043.
63. Alexander M, Kent SBH, Engelhard M, Merrifield RB (1978) A new synthetic route to tert-butyloxycarbonylaminoacyl-4-(oxymethyl)phenylacetamidomethyl-resin, an improved support for solid-phase peptide synthesis. *J Org Chem* 43:2845–2852.
64. Johnson ECB, Durek T, Kent SBH (2006) Total chemical synthesis, folding, and assay of a small protein on a water-compatible solid support. *Angew Chem Int Ed Engl* 45:3283–3287.
65. Vágner J, et al. (1996) Enzyme-mediated spatial segregation on individual polymeric support beads: Application to generation and screening of encoded combinatorial libraries. *Proc Natl Acad Sci USA* 93:8194–8199.
66. Feldhaus MJ, et al. (2003) Flow-cytometric isolation of human antibodies from a nonimmune *Saccharomyces cerevisiae* surface display library. *Nat Biotechnol* 21:163–170.
67. Townsend J, et al. (2010) 3-nitro-tyrosine as an internal quencher of autofluorescence enhances the compatibility of fluorescence based screening of OBOC combinatorial libraries. *Comb Chem High Throughput Screen* 13:422–429.
68. Yeung YA, Wittrup KD (2002) Quantitative screening of yeast surface-displayed polypeptide libraries by magnetic bead capture. *Biotechnol Prog* 18:212–220.
69. Ackerman M, et al. (2009) Highly avid magnetic bead capture: An efficient selection method for de novo protein engineering utilizing yeast surface display. *Biotechnol Prog* 25:774–783.
70. Mendes K, Ndungu JM, Clark LF, Kodadek T (2015) Optimization of the magnetic recovery of hits from one-bead one-compound library screens. *ACS Comb Sci* 17:506–517.
71. Schuck P, Minton AP (1996) Kinetic analysis of biosensor data: Elementary tests for self-consistency. *Trends Biochem Sci* 21:458–460.
72. Bondeson K, Frostell-Karlsson A, Fägerstam L, Magnusson G (1993) Lactose repressor-operator DNA interactions: Kinetic analysis by a surface plasmon resonance biosensor. *Anal Biochem* 214:245–251.
73. Kelley RF, O'Connell MP (1993) Thermodynamic analysis of an antibody functional epitope. *Biochemistry* 32:6828–6835.
74. Schuck P (1997) Use of surface plasmon resonance to probe the equilibrium and dynamic aspects of interactions between biological macromolecules. *Annu Rev Biophys Biomol Struct* 26:541–566.
75. Houghten RA (1985) General method for the rapid solid-phase synthesis of large numbers of peptides: Specificity of antigen-antibody interaction at the level of individual amino acids. *Proc Natl Acad Sci USA* 82:5131–5135.
76. Pinilla C, Appel JR, McPherson SE, Houghten RA (1992) Comparison of structural and functional approaches for the study of peptide-mAb interactions. *Peptides: Chemistry and Biology*, eds Smith JA, Rivier JE (ESCOM, Leiden, The Netherlands), pp 867–868.
77. Churchill MEA, et al. (1994) Crystal structure of a peptide complex of anti-influenza peptide antibody Fab 26/9. Comparison of two different antibodies bound to the same peptide antigen. *J Mol Biol* 241:534–556.
78. Rini JM, Schulze-Gahmen U, Wilson IA (1992) Structural evidence for induced fit as a mechanism for antibody-antigen recognition. *Science* 255:959–965.
79. Schulze-Gahmen U, Rini JM, Wilson IA (1993) Detailed analysis of the free and bound conformations of an antibody. X-ray structures of Fab 17/9 and three different Fab-peptide complexes. *J Mol Biol* 234:1098–1118.
80. O'Neil KT, et al. (1992) Identification of novel peptide antagonists for GP1b/IIIa from a conformationally constrained phage peptide library. *Proteins* 14:509–515.
81. James LC, Roversi P, Tawfik DS (2003) Antibody multispecificity mediated by conformational diversity. *Science* 299:1362–1367.
82. Souriau C, Chiche L, Irving R, Hudson P (2005) New binding specificities derived from Min-23, a small cysteine-stabilized peptidic scaffold. *Biochemistry* 44:7143–7155.
83. Geysen HM, Rodda SJ, Mason TJ (1986) A priori delineation of a peptide which mimics a discontinuous antigenic determinant. *Mol Immunol* 23:709–715.
84. Robson B (1996) Doppelgänger proteins as drug leads. *Nat Biotechnol* 14:892–893.
85. Mandal K, et al. (2012) Chemical synthesis and X-ray structure of a heterochiral D-protein antagonist plus vascular endothelial growth factor protein complex by racemic crystallography. *Proc Natl Acad Sci USA* 109:14779–14784.
86. Levinson AM, et al. (2017) Total chemical synthesis and folding of all-L and all-D variants of oncogenic KRas(G12V). *J Am Chem Soc* 139:7632–7639.
87. Braisted AC, Wells JA (1996) Minimizing a binding domain from protein A. *Proc Natl Acad Sci USA* 93:5688–5692.
88. Patel S, Mathonet P, Jaulent AM, Ullman CG (2013) Selection of a high-affinity WW domain against the extracellular region of VEGF receptor isoform-2 from a combinatorial library using CIS display. *Protein Eng Des Sel* 26:307–315.
89. Bhardwaj G, et al. (2016) Accurate de novo design of hyperstable constrained peptides. *Nature* 538:329–335.
90. Baker EG, Bartlett GJ, Porter Goff KL, Woolfson DN (2017) Miniprotein design: Past, present, and prospects. *Acc Chem Res* 50:2085–2092.
91. Chevalier A, et al. (2017) Massively parallel de novo protein design for targeted therapeutics. *Nature* 550:74–79.
92. Griffiths AD, Duncan AR (1998) Strategies for selection of antibodies by phage display. *Curr Opin Biotechnol* 9:102–108.
93. Francisco JA, Campbell R, Iverson BL, Georgiou G (1993) Production and fluorescence-activated cell sorting of *Escherichia coli* expressing a functional antibody fragment on the external surface. *Proc Natl Acad Sci USA* 90:10444–10448.
94. Checco JW, et al. (2015) Targeting diverse protein-protein interaction interfaces with α/β -peptides derived from the Z-domain scaffold. *Proc Natl Acad Sci USA* 112:4552–4557.
95. Rocklin GJ, et al. (2017) Global analysis of protein folding using massively parallel design, synthesis, and testing. *Science* 357:168–175.
96. Lian W, Upadhyaya P, Rhodes CA, Liu Y, Pei D (2013) Screening bicyclic peptide libraries for protein-protein interaction inhibitors: Discovery of a tumor necrosis factor- α antagonist. *J Am Chem Soc* 135:11990–11995.
97. Bayer E (1991) Towards the chemical synthesis of proteins. *Angew Chem Int Ed Engl* 30:113–129.
98. Schnölzer M, Alewood P, Jones A, Alewood D, Kent SBH (1992) In situ neutralization in Boc-chemistry solid phase peptide synthesis. Rapid, high yield assembly of difficult sequences. *Int J Pept Protein Res* 40:180–193.
99. Simon MD, et al. (2014) Rapid flow-based peptide synthesis. *ChemBioChem* 15:713–720.