

## MIT Open Access Articles

*Ultra-large chemical libraries for the  
discovery of high-affinity peptide binders*

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

**Citation:** Quartararo, Anthony J. et al. "Ultra-large chemical libraries for the discovery of high-affinity peptide binders." Nature Communications 11, 1 (June 2020): 3183 © 2020 The Author(s)

**As Published:** <http://dx.doi.org/10.1038/s41467-020-16920-3>

**Publisher:** Springer Science and Business Media LLC






**Persistent URL:** <https://hdl.handle.net/1721.1/128217>

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


**Terms of use:** Creative Commons Attribution 4.0 International license



# Ultra-large chemical libraries for the discovery of high-affinity peptide binders

Anthony J. Quartararo <sup>1</sup>, Zachary P. Gates<sup>1</sup>, Bente A. Somsen <sup>2</sup>, Nina Hartrampf <sup>1</sup>, Xiyun Ye<sup>1</sup>, Arisa Shimada<sup>3</sup>, Yasuhiro Kajihara<sup>3</sup>, Christian Ottmann<sup>2</sup> & Bradley L. Pentelute <sup>1,4,5,6</sup> 

High-diversity genetically-encoded combinatorial libraries ( $10^8$ – $10^{13}$  members) are a rich source of peptide-based binding molecules, identified by affinity selection. Synthetic libraries can access broader chemical space, but typically examine only  $\sim 10^6$  compounds by screening. Here we show that in-solution affinity selection can be interfaced with nano-liquid chromatography-tandem mass spectrometry peptide sequencing to identify binders from fully randomized synthetic libraries of  $10^8$  members—a 100-fold gain in diversity over standard practice. To validate this approach, we show that binders to a monoclonal antibody are identified in proportion to library diversity, as diversity is increased from  $10^6$ – $10^8$ . These results are then applied to the discovery of p53-like binders to MDM2, and to a family of 3–19 nM-affinity,  $\alpha/\beta$ -peptide-based binders to 14-3-3. An X-ray structure of one of these binders in complex with 14-3-3 $\sigma$  is determined, illustrating the role of  $\beta$ -amino acids in facilitating a key binding contact.

<sup>1</sup>Department of Chemistry, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. <sup>2</sup>Laboratory of Chemical Biology, Department of Biomedical Engineering and Institute for Complex Molecular Systems, Eindhoven University of Technology, PO Box 513, 5600 MB Eindhoven, Netherlands. <sup>3</sup>Department of Chemistry, Graduate School of Science, Osaka University, 1-1 Machikaneyama, Toyonaka, Osaka 560-0043, Japan. <sup>4</sup>The Koch Institute for Integrative Cancer Research, Massachusetts Institute of Technology, 500 Main Street, Cambridge, MA 02142, USA. <sup>5</sup>Center for Environmental Health Sciences, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139, USA. <sup>6</sup>Broad Institute of MIT and Harvard, 415 Main Street, Cambridge, MA 02142, USA. email: [blp@mit.edu](mailto:blp@mit.edu)

Drug discovery benefits from the ability to assay a large number of compounds for activity against a biomolecular target of interest. This target-based modality remains prominent in the pharmaceutical industry, and it is estimated that of all 113 first-in-class drugs approved by the FDA between 1999 and 2013, 71% were discovered using a target-based approach<sup>1</sup>. While a variety of techniques are used to this end, some of the most common strategies include high-throughput screening (HTS) and fragment-based drug discovery (FBDD)<sup>2</sup> for identification of small molecule binders, and protein engineering strategies<sup>3</sup> for discovery of novel biologics.

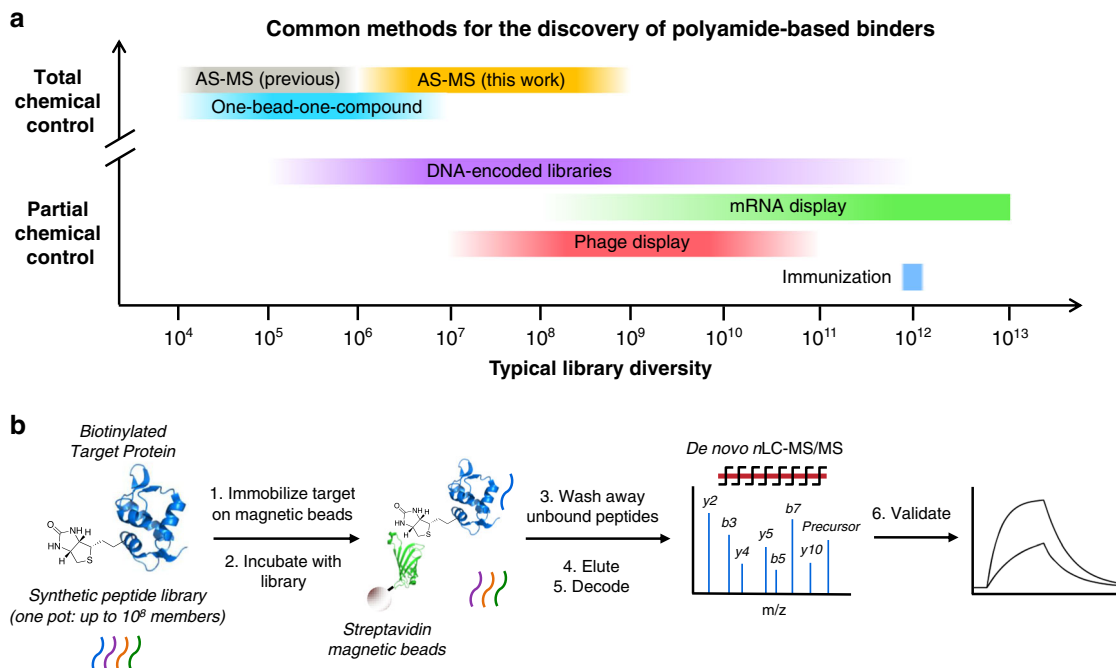
In the past decade, much research attention has been devoted to discovering and engineering peptide-based binders as an emerging therapeutic modality<sup>4,5</sup>. Given the unique niche of chemical space they occupy, possessing molecular weights in between those of small molecules (<500 Da) and biologics (up to ~150,000 Da), peptides offer a distinct profile of chemically attractive features. Peptides are synthetically accessible, amenable to chemical tailoring, and have the potential to bind the typically shallow surfaces seen in therapeutically relevant—and historically intractable—protein–protein interactions (PPIs)<sup>6–8</sup>. Importantly, chemical modifications, such as non-canonical amino acid incorporation, head-to-tail macrocyclization, and chemical stapling, can render peptides more proteolytically stable, more cell-penetrant, and even increase binding affinity relative to their natural, underivatized counterparts, which on their own tend to exhibit poor pharmacological properties<sup>9–13</sup>.

Molecular biology-based selection techniques, such as phage display<sup>14,15</sup> and mRNA display<sup>16</sup>, are powerful tools for target-based discovery of novel peptide binders, thanks in part to their ability to examine enormous libraries ( $10^8$ – $10^{13}$  members)<sup>17</sup>. However, incorporation of a large number and variety of non-canonical amino acids by these methods remains challenging (Fig. 1a)<sup>18,19</sup>. Chemical combinatorial methods, such as DNA-

encoded libraries (DELs)<sup>20,21</sup> and one-bead-one-compound (OBOC) libraries<sup>22,23</sup>, can easily overcome this hurdle; however, each of these techniques faces its own limitations. DELs are typically limited to three or four varied positions, due to inefficiencies in the chemistry used for their assembly, and as such are more often categorized as libraries of small molecules. OBOC libraries can easily incorporate more varied positions, but OBOC screening is technically challenging and typically examines only  $\sim 10^6$  compounds<sup>24</sup>.

Affinity selection-mass spectrometry (AS-MS)<sup>25–29</sup> represents an alternative strategy for target-based discovery of chemically accessed peptide binders. We recently leveraged LC-MS/MS for sequencing individual synthetic peptides present in complex mixtures<sup>30</sup>, and to increase the diversity of synthetic peptide libraries amenable to AS-MS<sup>12</sup>, from  $\sim 10$  to  $\sim 10^6$ . This advance was used to select improved variants of known binders from small focused libraries ( $10^3$  members)<sup>12</sup>. Discovery of binders from fully randomized libraries is a much greater challenge, which may require library diversities considerably higher than are typically examined by chemical screening. In principle, commercial mass spectrometers are sufficiently sensitive to detect and sequence peptides from mixtures as complex as  $\sim 10^9$  (Table 1)<sup>31,32</sup>. However, it is not obvious whether single-pass affinity selections from such libraries could provide sufficient enrichment (i.e., reduce the number of peptides sufficiently) to identify binders by LC-MS/MS-based sequencing, which is applicable to mixtures of up to several thousand peptides<sup>30</sup>.

Here, we show that magnetic bead capture-based AS-MS<sup>33,34</sup> is capable of identifying binders from libraries of up to  $\sim 10^8$  random synthetic peptides (Fig. 1b). Starting with an anti-hemagglutinin monoclonal antibody (anti-HA mAb) selection target, we demonstrate that high-affinity binders can be captured with near-quantitative recovery from relevant ligand concentrations. These results translate into a selection context, where high-



**Fig. 1 AS-MS enables discovery from randomized, high-diversity chemical libraries.** **a** Among existing techniques for peptide binder discovery, increased chemical control over library synthesis and screening typically comes at the cost of limited library diversity. Affinity selection-mass spectrometry (AS-MS; this work), which relies on chemically accessed libraries and direct identification of active members, can investigate library diversities on the order of  $10^8$ – $10^9$  while maintaining a high degree of chemical control over the synthesis and selection process. **b** A typical AS-MS workflow, which uses magnetic beads as a partitioning reagent to discriminate bound from unbound library members, and nano-liquid chromatography–tandem mass spectrometry to identify sequences of active peptides.

**Table 1 Mixtures of up to 10<sup>9</sup> peptides contain sufficient material for MS-based sequencing.**

Library size	Member concentration (pM)	Moles per member (fmol)
10 <sup>6</sup>	1000	1000
10 <sup>7</sup>	100	100
10 <sup>8</sup>	10	10
10 <sup>9</sup>	1	1

At a given scale (1 mL), maximum diversity is a tradeoff between the amount of individual peptide required for nLC-MS/MS sequencing, and the solubility limit of the total peptide mixture (1 mM). Even more peptides might be examined concurrently at a fixed total concentration by increasing the selection volume.

affinity binders containing the HA epitope are identified in proportion to library diversity over the range of 10<sup>6</sup>–10<sup>8</sup>. Similarly conducted selections for MDM2 binding identify p53-like peptides for 10<sup>8</sup>-member libraries only, illustrating the utility of high-diversity synthetic libraries for identifying binders previously accessible only to molecular biology-based approaches. The power of the synthetic library approach for identifying binders not otherwise accessible is illustrated by the discovery of mixed  $\alpha/\beta$ , phosphoserine-containing binders to the hub protein 14-3-3.

## Results

### AS-MS recovers high-affinity ligands from high dilution.

Identification of binders from high-diversity libraries would necessitate their efficient recovery at high dilution. To assess the utility of magnetic bead-based affinity capture for this purpose, we investigated the recovery of a number of peptide binders with varying affinity for anti-HA mAb clone 12ca5 (12ca5), which recognizes the linear motif DXXDYA (Supplementary Figs. 1–6; Supplementary Table 1). In these experiments, streptavidin-coated magnetic beads (1 mg; 0.13 nmol IgG binding capacity) functionalized with biotinylated 12ca5 were incubated with mixtures of 12ca5-binding peptides (either 1 nM/peptide, 100 pM/peptide, or 10 pM/peptide; 1 mL scale). The magnetic beads were isolated, washed with buffer, and treated with chemical denaturant to elute bound peptides. A portion of this eluate (~2%) was subjected to nLC-MS analysis, and recovery was quantified using normalized MS response. To calculate recovery, a separate dose-response curve was generated for each peptide, since their MS responses varied by ~5-fold (Supplementary Fig. 7).

Under the conditions examined, only the two highest affinity binders were significantly retained, with recoveries of 75% and 33% obtained for ~4 and ~25 nM-affinity binders, respectively (Fig. 2a; Supplementary Figs. 8 and 9). This result is consistent with the work of Sannino and coworkers, who observed low recoveries for micromolar affinity ligands in the context of DELs<sup>35</sup>. We tested the use of increased magnetic bead concentration to recover lower affinity ligands; however, this parameter had little effect on recovery, perhaps suggesting that these ligands are lost during the washing step (Supplementary Fig. 10). Consistent with this view, recovery of the 25 nM-affinity ligand (33%) is in quantitative agreement with that expected based on its measured dissociation rate, and the total wash time of ~6 min (34%, based on  $k_{\text{off}} = 3.0 \times 10^{-3} \text{ s}^{-1}$ ; Supplementary Fig. 6). Therefore, dissociation rate is an important factor in controlling the recovery of binders by AS-MS. For the retained binders, no loss in recovery was observed as their initial concentration was decreased from 1 nM to 10 pM, suggesting that selections from libraries as diverse as 10<sup>8</sup> would be feasible (Supplementary Figs. 8 and 9; Table 1). Below 10 pM, low MS detector response precluded convenient quantitation.

We examined whether recovery of high-affinity binders could be maintained at concentrations below those conveniently quantified. Affinity captures were performed at 10 and 1 pM ligand concentration (10 or 1 fmol/peptide, respectively), and the eluates were concentrated by solid phase extraction such that the majority of the bound fraction could be analyzed (~80%). Although recoveries could be maintained down to 1 pM ligand, the resulting MS signals were observed near the intensity threshold that would be used for MS/MS precursor isolation during an actual selection, to facilitate collection of high quality spectra ( $5 \times 10^4$  counts; Supplementary Fig. 11). The precursor selection threshold defines a limit of detection for sequenced peptides, which corresponds to ~50% or ~5% recovery of peptides initially present at 1 pM (1 fmol/mL) or 10 pM (10 fmol/mL), respectively. Therefore, 10<sup>9</sup> is an upper bound on the library diversity amenable to this AS-MS approach at the examined scale (1 mL), to retain sufficient material for nLC-MS/MS sequencing (Table 1).

**AS-MS enriches 12ca5 binders from a 10<sup>6</sup>-member library.** To assess whether magnetic bead-based pulldowns could recover binders from complex mixtures and reduce the mixture complexity sufficiently for nLC-MS/MS sequencing, we synthesized a library of design (X)<sub>9</sub>K, where X = all L-amino acids except for cysteine and isoleucine (theoretical diversity =  $2 \times 10^{11}$ ), by split-and-pool synthesis<sup>22,36</sup> on 30  $\mu\text{m}$  TentaGel resin (2.9 g;  $2 \times 10^8$  beads), for use in subsequent selections (Supplementary Fig. 12). Portions of this library comprising  $2 \times 10^7$  and  $2 \times 10^6$  members were taken prior to cleavage from resin to give three libraries of increasing size from the same batch of split-and-pool synthesis (Fig. 2b). Selections against 12ca5 (0.13 nmol) were then performed with the  $2 \times 10^6$ -member library, at 10 pM/member concentration (1 mL scale; 10 fmol/peptide, 20 nmol total library), with the goal of identifying sequences similar to the HA epitope (selections were performed in triplicate; Supplementary Note 1). Selection eluates were concentrated by solid phase extraction, and analyzed by nLC-MS/MS (theoretical loading of a retained peptide: 8 fmol). From these selections, just one peptide that matched the library design was identified with a sequencing score (average local confidence (ALC) score<sup>37</sup>)  $\geq 80$  (MNDLVYADK; Supplementary Table 2). This sequence had five residues in common with the HA epitope, DXVDYA, including the hot spot residues Asp4, Asp7, Tyr8, and Ala9 required for high-affinity 12ca5-binding (Supplementary Figs. 13–21)<sup>38</sup>.

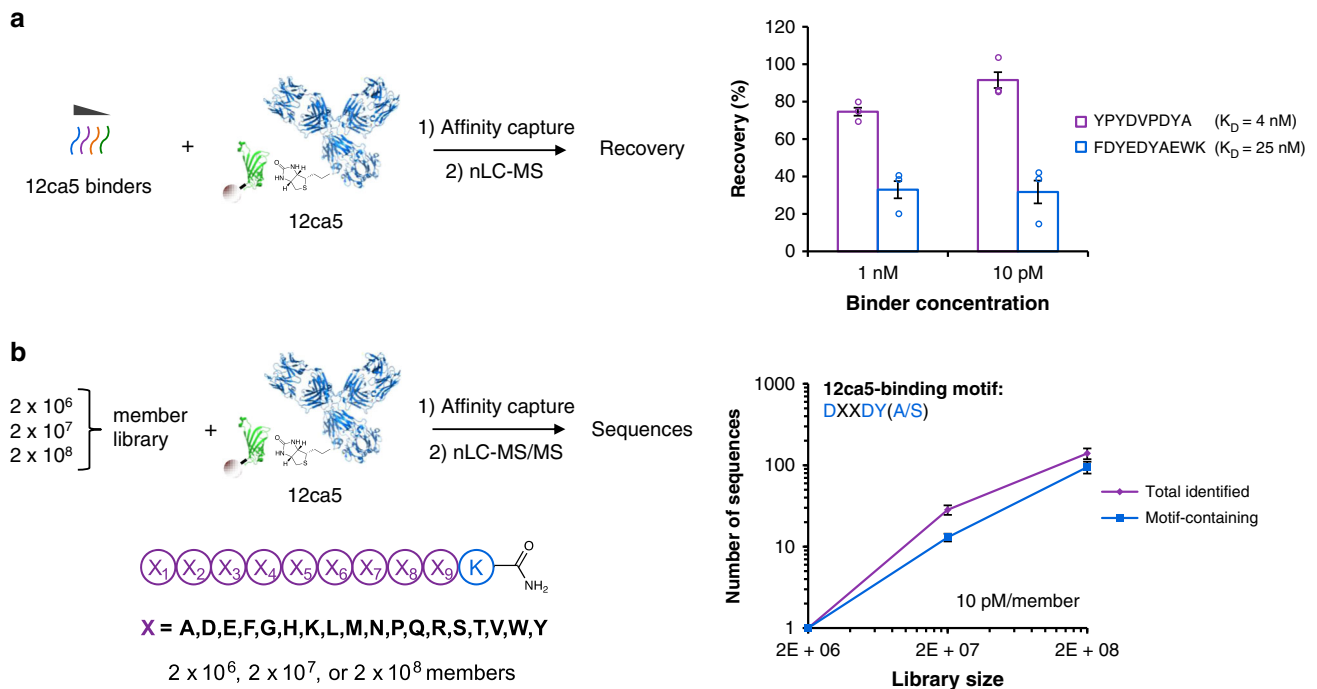
We endeavored to understand the enrichment achieved by the selection, where enrichment is defined as:

$$\frac{(\text{sequenced binders}/\text{total sequenced peptides})_{\text{isolated}}}{(\text{binders}/\text{total peptides})_{\text{assayed}}} \quad (1)$$

or equivalently

$$\frac{(\text{sequenced binders})_{\text{isolated}}/(\text{binders})_{\text{assayed}} * (\text{total peptides})_{\text{assayed}}}{(\text{total sequenced peptides})_{\text{isolated}}} \quad (2)$$

Achieving high enrichment in a single selection step would be essential for discovering rare binding molecules by AS-MS, since synthetic libraries cannot be propagated to facilitate sequential rounds of selection. Since only a single peptide was sequenced from the selection eluate, a maximum enrichment of  $2 \times 10^6$  was achieved. The actual enrichment may be significantly lower, if additional binders were present in the library, but not recovered and sequenced. For example, the expected number of DXXDY (A/S)-containing sequences in the  $2 \times 10^6$ -member library is 152 (although not all of these are necessarily high-affinity binders), suggesting the actual enrichment could be as low as  $\sim 1.3 \times 10^4$  (Supplementary Note 2).



**Fig. 2** Magnetic bead capture recovers high-affinity binders from both simple mixtures and random libraries. **a** High-affinity 12ca5-binding peptides were recovered effectively from a concentration of 10 pM, which corresponds to the concentration of an individual library member in a 10<sup>8</sup>-member library at 1 mM total peptide. Recoveries were determined by MS detector response, normalizing to an internal standard. **b** The number of identified sequences bearing the 12ca5-binding motif—DXXDY(A/S)—increased in proportion to library diversity, from two million to two hundred million-member libraries. On average, motif-containing sequences represented 71% of all sequences identified. Error bars correspond to one standard deviation among three technical replicates.

To check whether additional peptides were recovered by the selection at lower abundance, we performed a selection at 200 pM/member, such that less-abundant peptides could cross the precursor selection threshold (Supplementary Fig. 22). Increasing amounts of the selection eluate were analyzed by nLC-MS/MS, with sample loadings varied from 8 fmol/member (selection threshold: ~5% recovery) to 75 fmol/member (selection threshold: ~0.5% recovery). At the lowest sample loading, just one peptide was identified (MNDLV<sup>D</sup>YADK), identical to that observed in the original selection. As sample loading was increased to 75 fmol/member, two additional epitope-containing peptides were identified: PDVHDY<sup>T</sup>WGK and ENDWQDY<sup>S</sup>HK. However, recovery of these sequences came at the cost of identifying additional non-motif-containing peptides: a total of 32 from 3 replicate selections. These results suggest that with respect to recovery of binders, little benefit is conferred by higher library concentration, and that 8 fmol/member sample loading avoids detection of peptides present in the selection eluate at lower abundance. Accounting for these additional sequences yields an enrichment of  $1.1 \times 10^3$ , assuming 152 DXXDY(A/S)-containing peptides (Eq. (2)), which is on the order of enrichments reported for individual rounds of mRNA ( $\sim 10^3$ )<sup>39</sup>, phage ( $\sim 10^3$ )<sup>14</sup>, or cell surface display ( $\sim 10^4$ )<sup>40,41</sup>.

Taken together, these results suggest that magnetic bead affinity capture can facilitate selections from complex mixtures of synthetic peptides at a concentration of 10 pM/member. In conjunction with nLC-MS/MS peptide sequencing, very high enrichment for sequenced peptides is enabled by the use of a precursor selection threshold, to detect only those peptides that are significantly recovered by the selection (here, a DXXDY(A/S)-containing peptide—present in highest abundance). Peptides recovered in lower abundance (here, mostly non-motif-containing peptides) are present in the selection eluate, but not

sequenced (Supplementary Table 4). As described below (‘Parallel selections distinguish non-specific binders’), the majority of these peptides are non-specific binders.

**Enrichment is maintained as diversity increases from 10<sup>6</sup>-10<sup>8</sup>.** To investigate whether comparable selection performance could be achieved using higher-diversity libraries, we assayed 2 × 10<sup>7</sup>- and 2 × 10<sup>8</sup>-member libraries in selections for 12ca5 binding. In theory, these libraries should contain 10- and 100-fold more DXXDY(A/S) peptides compared to the 2 × 10<sup>6</sup>-member library. If selections from the higher-diversity libraries performed comparably, then they should recover all of these additional DXXDY(A/S)-containing peptides, with no increase in the proportion of non-binding peptides. For each library, selections were performed in triplicate, with library concentration maintained at 10 pM/member and using 0.13 nmol of selection target, as above.

From the 2 × 10<sup>7</sup>-member library, a representative selection identified 23 peptides that matched the library design (ALC ≥ 80), 14 of which (61%) contained DXXDYA or DXXDYS (Fig. 2b; Supplementary Table 2). An additional two sequences closely resembled 12ca5 binders—KVL<sup>D</sup>YDYAWK and YDDRYAD<sup>T</sup>FK—and may correspond to inaccurate sequence assignments. Selections from the 2 × 10<sup>8</sup>-member library identified 156 total peptides (ALC ≥ 80), 109 of which (70%) contained either DXXDYA or DXXDYS (Fig. 2b; Supplementary Tables 2-3). These results illustrate that selections identified binders from the higher-diversity libraries without loss in recovery, since the expected ~10 and ~100 DXXDY(A/S)-containing peptides were identified, and without loss in enrichment, since DXXDY(A/S)-containing peptides comprised the majority of selected peptides in each case (100%, 61%, and 70% for 2 × 10<sup>6</sup>, 2 × 10<sup>7</sup>, and 2 × 10<sup>8</sup>-member libraries, respectively). Therefore, we concluded that

single-pass AS-MS could be applied to libraries of at least  $10^8$  random peptides without loss in performance.

### Recovery of binders drops from libraries beyond $10^8$ members.

We next set out to understand whether libraries of diversity beyond  $10^8$  would also be amenable to AS-MS. To access  $10^9$  random synthetic peptides on a convenient lab scale, we used 20  $\mu\text{m}$  TentaGel resin beads (vs. 30  $\mu\text{m}$  beads, above) to prepare a library of design  $(X)_9K$  ( $X$  = all L-amino acids except for cysteine and isoleucine; 5.4 g;  $1.3 \times 10^9$  beads) (Supplementary Fig. 23). As above, a portion of the beads was set aside prior to cleavage from resin to give a  $10^8$ -member library for side-by-side comparison.

The  $10^8$ -member library performed comparably to the  $10^8$ -member library prepared on 30  $\mu\text{m}$  TentaGel beads (above), yielding 257 total peptides ( $\text{ALC} \geq 80$ ), 183 of which (71%) contained either DXXDYA or DXXDYS. In contrast, selections from the  $10^9$ -member library identified only 34 peptides ( $\text{ALC} \geq 80$ ), 21 of which (62%) contained DXXDYA or DXXDYS (Supplementary Table 5). For the  $10^9$ -member library, selections were performed at 2 pM/member to maintain solubility; however, the lower sequence identification rate cannot be attributed to material losses alone, since selection from a  $10^8$ -member library performed at 2 pM/member identified 131 DXXDY(A/S) sequences (vs. 183 from the selection at 10 pM/member; Supplementary Table 5). Analysis of pooled eluates from replicate selections from the  $10^9$ -member library yielded no additional sequences, providing further evidence that material losses were not responsible for the reduction in sequence identification rate (since similar populations of DXXDY(A/S)-containing sequences are recovered by replicate selections) (Supplementary Tables 3 and 6).

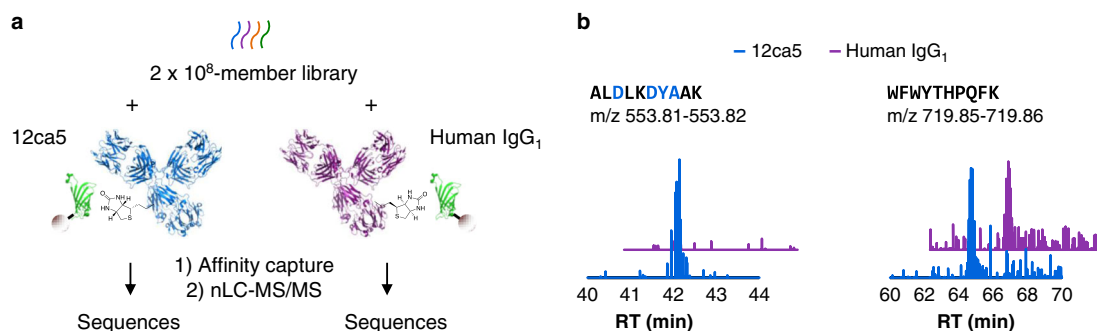
Since affinity selections involve a large number of potential binders competing for a limited number of binding sites, it is possible that many weaker binders present in the library reduce the individual recoveries of all ligands. Two experiments were performed to address whether competition was responsible for the lower recovery of DXXDY(A/S) peptides from the  $10^9$ -member library. First, we performed a selection using 10-fold higher stoichiometry of 12ca5 (1.3 nmol) relative to library (2 fmol/member). If competition were limiting recovery of binders, this experiment should have recovered additional DXXDY(A/S)-containing peptides. Instead, their recovery was abrogated entirely (Supplementary Table 7). Second, to determine the frequency of binders that would be required for competition to become significant, we studied the effect of exogenous competitors on the recovery of DXXDY(A/S) peptides from the  $10^8$ -

member library. In all, 4 nM- or 3  $\mu\text{M}$ -affinity competitors required concentrations of 100 nM or 100  $\mu\text{M}$ , respectively, to attenuate the recovery of DXXDY(A/S) peptides (Supplementary Fig. 24; Supplementary Table 8). These concentrations correspond to  $5 \times 10^4$  or  $5 \times 10^7$  peptides (present at 2 pM/peptide), suggesting that: (1) the expected  $\sim 10^3$  DXXDY(A/S) peptides in the  $10^9$ -member library would not compete for selection target; and (2) in order for  $\mu\text{M}$ -affinity binders to compete with DXXDY(A/S) peptides, they would need to comprise 5% of the  $10^9$ -member library.

The combined results are consistent with the interpretation that selections from the  $10^9$ -member library yield more peptides than are compatible with nLC-MS/MS sequencing. As the number of peptides in the selection eluate (the sample complexity) increases, the proportion of sequenced peptides decreases<sup>30</sup>. This drop in sequencing coverage was observed previously to be particularly significant beyond  $10^3$  peptides<sup>30</sup>, which is consistent with the approximate number of total peptides (binders and non-binders) likely isolated from the  $10^9$ -member library ( $>3.5 \times 10^3$ , based on a maximum of 35 peptides isolated from the  $10^6$ -member library). Additionally, the competition experiments provide strong evidence that selections took place under conditions of equilibrium binding, since the relative strength of the two competitors was in proportion to their binding constants. Therefore, predictable competition effects should be expected when AS-MS is conducted in the presence of added competitor, or from focused libraries containing a higher proportion of binders.

**Parallel selections distinguish non-specific binders.** Selections from identical portions of synthetic peptide mixtures obtained by split-and-pool synthesis are readily conducted in parallel. For example, a  $2 \times 10^8$ -member library synthesized on 30  $\mu\text{m}$  TentaGel ( $\sim 3.7$  pmol/bead) provides sufficient material for  $\sim 370$  selections performed at 10 fmol/member scale. We leveraged this capability to understand what proportion of non-HA epitope-containing peptides were common to an unrelated selection target, and therefore attributable to non-specific binding<sup>42</sup>. Side-by-side selections from a  $2 \times 10^8$ -member library were conducted in triplicate against both 12ca5 (mouse IgG<sub>2b</sub> $\kappa$ ) and a polyclonal human IgG<sub>1</sub>, with the goal of quantifying the degree of overlap among selected peptides (Fig. 3a).

Three replicate 12ca5 selections yielded a total of 133 DXXDY(A/S)-containing peptides, all of which were specific for 12ca5, along with 65 non-motif containing peptides (Fig. 3b; Supplementary Table 9). Of the non-motif-containing peptides, 18



**Fig. 3** Parallel selections enable differentiation of specific and non-specific binders. **a** Using AS-MS, the same population of compounds can be used in selections against more than one target in parallel, enabling assessment of target-dependent enrichment of peptides. Here, selections using a  $2 \times 10^8$ -member library were performed against 12ca5 and a polyclonal human IgG<sub>1</sub> in parallel, followed by sequential nLC-MS/MS analysis. **b** Extracted ion chromatograms of a DXXDY(A/S)-containing peptide (left) and a non-motif-containing peptide (right) from a selection against 12ca5 (blue) and polyclonal human IgG<sub>1</sub> (purple). The peptide containing the 12ca5-binding motif is selectively enriched towards 12ca5, while the non-motif-containing peptide is promiscuously pulled down in both conditions.

contained motifs that differed from the HA epitope by one position (for example, EXXDYA). Of the remaining 47 peptides, closer inspection revealed that 13 may contain mis-sequenced HA epitopes. For example, VFQDWEDFSK and YMDTVDFSEK contain the FS dipeptide fragment, which is isobaric to YA (Supplementary Fig. 25). A total of 34 peptides had no obvious sequence similarity to the DXXDY(A/S) motif. 17 of these were also sequenced from the IgG<sub>1</sub> selections. Examination of the LC-MS data revealed that many or all of the remaining 17 peptides were also present in the IgG<sub>1</sub> selection eluates, but not sequenced (9 of 9 selected peptides examined; Supplementary Fig. 26). Therefore, essentially all of the non-motif-containing peptides sequenced from 12ca5 selection can be attributed to non-specific binding. Presumably, many additional non-motif-containing peptides were recovered at lower abundance and not sequenced, as for the 10<sup>6</sup>-member library.

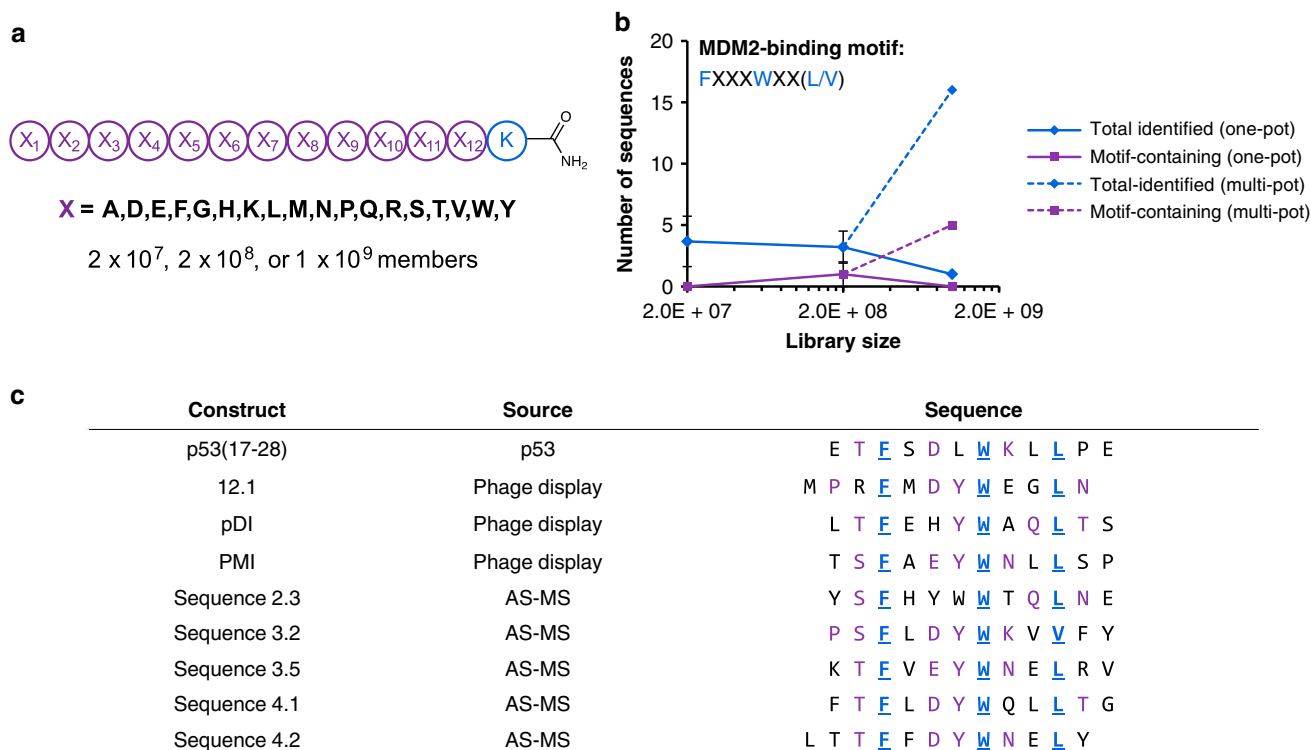
### High diversity enables discovery of MDM2-binding peptides.

Having established magnetic bead-based AS-MS as a selection protocol applicable to high-diversity libraries of random synthetic peptides, we sought to benchmark its performance relative to affinity selection from genetically-encoded libraries. As a model target for this purpose, we selected MDM2, an oncogenic ubiquitin ligase that binds its substrate p53 through a FXXXWXXL motif. Phage display has identified a number of well-characterized, high-affinity MDM2 binders based on this motif<sup>15,43–45</sup>, and we sought to determine whether AS-MS could recapitulate these

results by identifying similar sequences from synthetic libraries of comparable size and design.

To mimic previous phage display libraries, a library of design (X)<sub>12</sub>K, where X = all L-amino acids except cysteine and isoleucine (theoretical diversity =  $1.2 \times 10^{15}$ ), was synthesized on 20 μm TentaGel resin (4.2 g;  $1.3 \times 10^9$  beads). Prior to cleavage from resin, this library was split to yield five distinct  $2 \times 10^8$ -member libraries, as well as a  $2 \times 10^7$ -member library, to investigate the importance of library size in the context of selections for MDM2 binding (Fig. 4a; Supplementary Fig. 27). The N-terminal domain of MDM2 (residues 25–109) was accessed synthetically in biotinylated form, to enable its use as a selection target in conjunction with streptavidin-coated magnetic beads (Supplementary Fig. 28).

Selections from three of the five  $2 \times 10^8$ -member libraries against (25–109)MDM2 (0.13 nmol) yielded sequences containing the FXXXWXX(L/V) motif characteristic of MDM2-binding (Fig. 4b). In total, 16 sequences from these selections were identified (ALC ≥ 80), five of which (31%) contained FXXXWXX(L/V) (Supplementary Table 10). An additional two sequences appeared to be MDM2-binding peptides, containing the minimal FXXXW motif, but were potentially mis-sequenced due to poor-quality fragmentation spectra (Supplementary Fig. 29). Selections from the  $2 \times 10^7$ -member library did not yield motif-containing sequences, consistent with the frequency of binders identified from  $2 \times 10^8$ -member libraries. Selections were also performed from a 10<sup>9</sup>-member library, obtained by pooling the individual  $2 \times 10^8$ -member libraries. These selections failed to identify



**Fig. 4 High-diversity libraries facilitate discovery of p53-like peptides.** **a** Design of libraries used in selections against MDM2. Each library was derived from the same batch of split-and-pool SPPS. The readout for selections was the number of sequences bearing the MDM2-binding motif, FXXXWXX(L/V). **b** Motif-containing sequences could be identified by selection from individual  $2 \times 10^8$ -member libraries (one-pot), but not from  $2 \times 10^7$ -member or  $1 \times 10^9$ -member libraries (obtained by pooling the five individual  $2 \times 10^8$ -member libraries; one-pot). Selections from the five individual  $2 \times 10^8$ -member libraries (multi-pot;  $1 \times 10^9$  members total) identified five motif-containing sequences, suggesting that library size was enabling for identification of p53-like peptides as long as the mixture was not overly complex. Error bars correspond to one standard deviation among three technical replicates. **c** Sequences of known MDM2-binding peptides, including the native MDM2-binding epitope of p53 (residues 17–28) and three peptides identified from phage display, aligned with sequences identified from **b**. Known hot spot residues (F, W, L/V) are indicated in blue; aligned hot spot positions are indicated in bold and underline; and additional conserved residues are indicated in purple.

motif-containing sequences, consistent with the poor recovery of motif-containing sequences from  $10^9$ -member libraries in selections against 12ca5. In summary, library diversity correlates with number of binders identified, provided the diversity is within the technical limits of AS-MS for decoding single-pass selections.

A closer examination of the FXXXWXX(L/V)-containing sequences identified here alongside known MDM2-binding peptides revealed sequence similarity outside of the FWL triad (Fig. 4c). Specifically, a 6-residue motif was observed among the majority of sequences: (S/T)FX(D/E)YWXXL. Each of the conserved positions corresponds to a hot spot of binding energy, as determined by mutational analysis<sup>46</sup>. The ability to identify not only the FWL triad but also other significant determinants of binding affinity supports our interpretation that AS-MS is capable of matching the performance of phage display, in the context of selections against MDM2. While others have demonstrated the utility of synthetic libraries for identifying MDM2 binders<sup>42,47</sup>, our results illustrate their utility for mapping the determinants of a protein-protein interaction.

**AS-MS identifies non-canonical 14-3-3 $\gamma$ -binding peptides.** To illustrate a key advantage of the synthetic library approach, we investigated whether AS-MS could achieve comparable selection performance from a library based on non-canonical amino acids. As a selection target, we chose the hub protein 14-3-3, which interacts with a range of disease-relevant proteins including p53, Raf kinases, and estrogen receptor  $\alpha$ . Considerable effort has been devoted to developing modulators of these interactions<sup>48</sup>, which are generally mediated by phosphorylation<sup>49</sup>.

For use in selections against 14-3-3, we designed a library based on a fixed phosphoserine, flanked by eight varied positions (Fig. 5a). At each varied position, we incorporated one of 18 non-canonical amino acids—including  $\beta$ - and D-amino acids—encompassing a variety of polar, non-polar, charged, and aromatic side-chain functionalities (Fig. 5a) (theoretical diversity =  $1.1 \times 10^{10}$ ). This library was synthesized on 30  $\mu$ m TentaGel resin (2.9 g;  $2 \times 10^8$  beads), yielding  $2 \times 10^8$  members (Supplementary Fig. 30).

Side-by-side selections were performed against the 14-3-3 $\gamma$  isoform, and 12ca5 (negative control, to identify non-specific binders; 0.13 nmol each). These selections yielded a total of 19 sequences that matched the library design (ALC  $\geq$  80), 17 of which were unique to 14-3-3 $\gamma$ , and 2 of which were unique to 12ca5 (Supplementary Table 11; Supplementary Fig. 31). Extracted ion chromatograms were used to verify the absence of 14-3-3 $\gamma$ -unique sequences from the 12ca5 selections (Supplementary Figs. 32 and 33). Among the 14-3-3 $\gamma$ -unique sequences, seven contained a C-terminal motif: ( $\beta$ -homoserine)-( $\beta$ -alanine/ $\beta$ -homoserine)-(4-nitrophenylalanine). In general, D-amino acids were not present within the sequences identified.  $\beta$ -amino acids were enriched at positions 3, 7, and 8, but were otherwise largely absent. A positional frequency analysis revealed additional preferences for cyclohexylalanine at the N-terminus, and a  $\beta$ -homoserine at position three (Supplementary Fig. 31).

#### Discovered peptides bind 14-3-3 $\gamma$ with low nanomolar affinity.

To test the binding affinity of putative 14-3-3 $\gamma$ -binding peptides that contained the C-terminal ( $\beta$ -homoserine)-( $\beta$ -alanine/ $\beta$ -homoserine)-(4-nitrophenylalanine) motif, we synthesized fluorophore-labeled forms of three selected peptides, for use in a fluorescence anisotropy binding assay (Fig. 5b; Supplementary Figs. 34–37). As a negative control, we employed a sequence identified from a 12ca5 selection. Each of the putative 14-3-3 $\gamma$ -binding peptides examined exhibited low nanomolar affinity for 14-3-3 $\gamma$ , with  $K_D$  values ranging from 3 to 19 nM (Fig. 5c). By

contrast, the negative control peptide exhibited  $\sim 10,000$ -fold weaker binding, suggesting that the specific amino acid sequences of the 14-3-3 $\gamma$ -binding peptides—not the phosphoserine alone—were required for their identification.

Unlabeled forms of the non-canonical 14-3-3 $\gamma$  binders were assayed for their ability to compete with BiExoS, a peptide ligand derived from the *Pseudomonas aeruginosa* cytotoxin Exoenzyme S<sup>50</sup>, in a fluorescence anisotropy competition assay (Supplementary Figs. 38–41). This experiment would test whether the peptides bind to the amphipathic 14-3-3 $\gamma$  binding groove, or elsewhere on the protein. The non-canonical 14-3-3 $\gamma$  binders were found to compete off BiExoS with IC<sub>50</sub> values ranging from 78 to 530 nM, suggesting that they indeed bind in the canonical, phosphopeptide-accepting binding channel on 14-3-3 $\gamma$  (Fig. 5d)<sup>48</sup>. By contrast, the negative control peptide showed no inhibitory activity.

#### $\beta$ -amino acids facilitate a key binding contact with 14-3-3.

As an additional means of characterizing the binding interaction of non-canonical peptides with 14-3-3, we crystallized 14-3-3.12 in complex with 14-3-3 $\sigma$ . 14-3-3 $\sigma$  was used in place of 14-3-3 $\gamma$  to facilitate crystallization, and retained most of the binding activity for 14-3-3.12 (Supplementary Fig. 42). Diffraction data were collected to a resolution of 1.8 Å, and the structure was solved by molecular replacement (Supplementary Table 12).

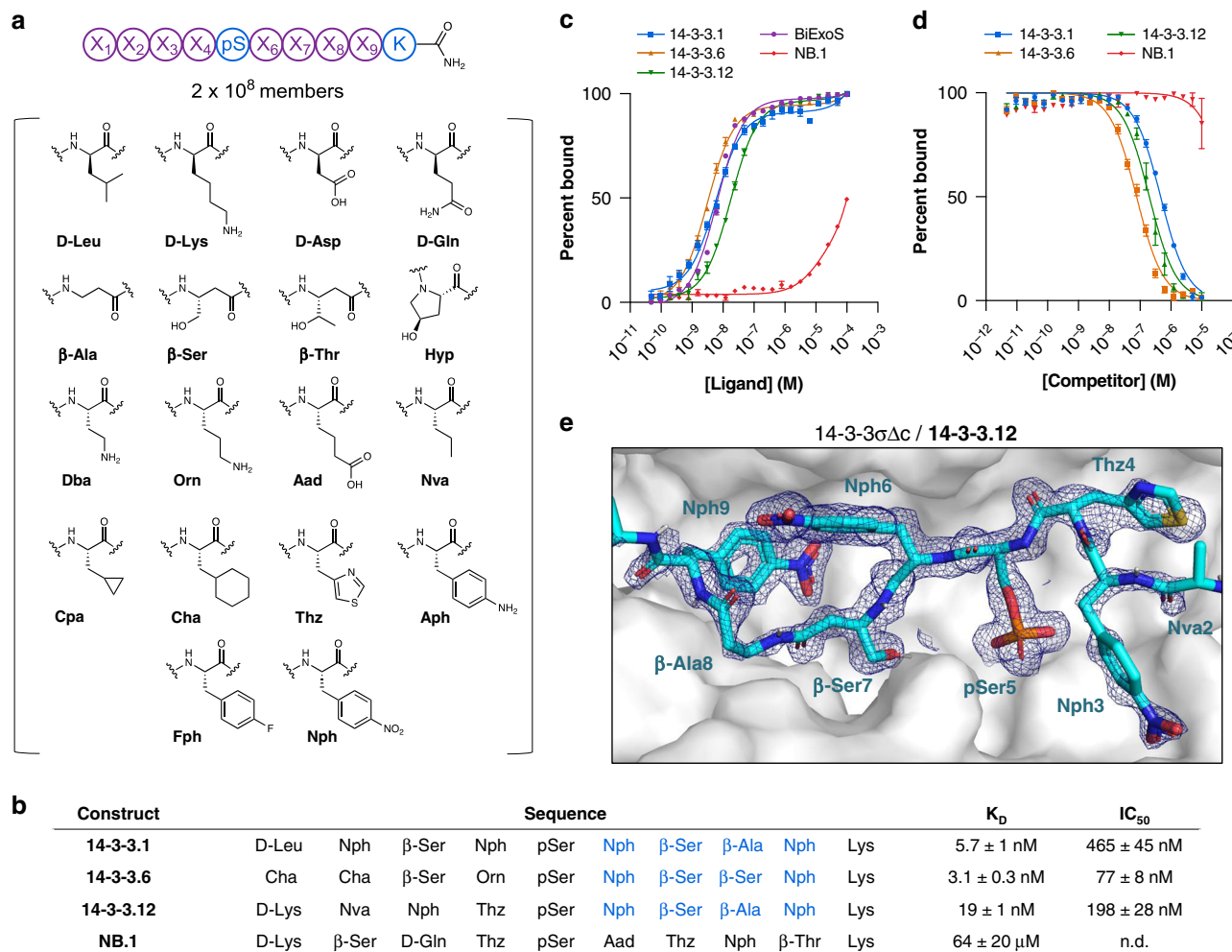
The 14-3-3.12 backbone adopts an extended conformation in the 14-3-3 $\gamma$  binding groove, flanked by two half-turns<sup>51</sup> defined by thiazolylalanine4 and  $\beta$ -alanine8 (Fig. 5e). 4-Nitrophenylalanine6, which was selected along with thiazolylalanine and 4-fluorophenylalanine at this position, makes hydrophobic contacts with Leu218, Ile219, and Leu222 of 14-3-3 $\sigma$  (Supplementary Fig. 43). 4-Nitrophenylalanine9—the residue most conserved by the selection—participates in an electrostatic interaction and/or H-bond with the  $\epsilon$ NH<sub>3</sub> group of Lys122 (N–O distance = 3.2 Å), and makes a hydrophobic contact with Ile168 (Supplementary Fig. 43). We speculate that the  $\beta$ -residues conserved at positions 7 and 8 of 14-3-3.12 provide the backbone flexibility necessary to accommodate these energetically-important interactions, which were not identified by selection from peptide libraries based on canonical amino acids<sup>49</sup>.

#### Discussion

In this work, we demonstrate that affinity selection-mass spectrometry, using magnetic bead reagents, provides sufficient enrichment to identify high-affinity binders from randomized libraries of  $10^8$  synthetic peptides. With respect to accessible diversity, this advance brings synthetic libraries up to the level of molecular biology-based combinatorial libraries. Diversity is a key determinant of selection outcome, as illustrated here for the discovery of p53-like binders to MDM2, and in the field of antibody engineering<sup>41,52</sup>. Therefore, the results described here can be expected to considerably extend the utility of synthetic libraries for discovering novel binding molecules.

The practical limit to library diversity amenable to single-pass AS-MS is  $10^8$ , beyond which the number of binders identified decreases. Our combined results are consistent with non-specific binding as the origin of this limit, which results in the recovery of more peptides from  $>10^8$ -member libraries than can be sequenced by nLC-MS/MS with high coverage. Since diversity is limited by selection performance and nLC-MS/MS sequencing coverage, rather than peptide solubility, future work should focus on these areas. For example, multi-stage selections might be employed to improve enrichment further, and to reduce the number of peptides in a  $>10^8$ -member library sufficiently for nLC-MS/MS sequencing. Sequencing coverage might also be





**Fig. 5 Synthetic libraries identify a 14-3-3 $\gamma$ -binding consensus based on  $\beta$ -amino acids.** **a** Design of a non-canonical library used in selections against 14-3-3 $\gamma$ . A phosphoserine was fixed in the middle, flanked by four varied positions on either side. A suite of non-canonical amino acids was incorporated at each varied position. **b** Sequences of putative 14-3-3 $\gamma$ -binding peptides (**14-3-3.1**, **14-3-3.6**, and **14-3-3.12**) identified from affinity selections, as well as a negative control peptide (non-binder or **NB.1**, identified as an artifact in selections against 12ca5) chosen for resynthesis and binding validation studies. Residues comprising the conserved C-term motif among 14-3-3 $\gamma$  binders—Nph- $\beta$ -Ser-( $\beta$ -Ala/ $\beta$ -Ser)-Nph—are indicated in blue. **c** Identified binders exhibited nanomolar affinities for 14-3-3 $\gamma$ , as determined by fluorescence anisotropy of FITC-labeled 14-3-3 $\gamma$ -binding peptides, NB.1, and a known 14-3-3 $\gamma$ -binding peptide (BiExoS; positive control). These affinities were approximately 10,000-fold higher relative to the negative control.  $K_D$  values are given in **b**. Uncertainties correspond to 95% confidence intervals derived from nonlinear regression. **d** Identified 14-3-3 $\gamma$ -binders could compete off bound BiExoS in a fluorescence anisotropy competition assay, suggesting they bind 14-3-3 $\gamma$  in the canonical, amphipathic binding groove.  $IC_{50}$  values are given in **b**. **e** Molecular structure of 14-3-3 $\sigma\Delta C$  (white surface) in complex with **14-3-3.12** (cyan sticks), based on a 1.80 Å crystal structure. The 2Fo-Fc electron density map corresponding to **14-3-3.12** is shown (blue mesh), contoured at  $1\sigma$ . Abbreviations:  $\beta$ -Ala  $\beta$ -alanine;  $\beta$ -Ser  $\beta$ -homoserine;  $\beta$ -Thr  $\beta$ -homothreonine; Aad amino adipic acid; Aph 4-aminophenylalanine; Cha cyclohexylalanine; Cpa cyclopropylalanine; Dbal diaminobutyric acid; Fph 4-fluorophenylalanine; Hyp hydroxyproline; Nph 4-nitrophenylalanine; Nva norvaline; Orn ornithine; pSer phosphoserine; Thz thiazolylalanine. Error bars correspond to standard error among three technical replicates.

improved by the use of specialized nLC columns and extended analysis times<sup>53</sup>.

The primary benefit of synthetic peptide libraries is the chemical control gained over the library design. Here, the combination of large library diversity and non-canonical amino acids led to the discovery of high-affinity 14-3-3 $\gamma$ -binding peptides that utilize  $\beta$ -amino acids to facilitate a binding contact. Given the comparatively low diversities examined by AS-MS relative to the upper bounds of genetically-encoded techniques ( $10^8$  vs.  $10^{13}$ ), we anticipate that taking advantage of the chemical capabilities AS-MS affords—such as straight-forward non-canonical amino acid incorporation—may prove critical for more intractable targets. For example, AS-MS may be

particularly suited to engineering peptide and peptoid foldamers<sup>54,55</sup>. Interfacing non-canonical amino acid incorporation with the macrocyclic architectures that have been rendered accessible to phage<sup>56</sup> and mRNA display<sup>57</sup> would further expand the breadth of chemical space amenable to exploration by AS-MS. In our case, performing selections on libraries of macrocycles would require an additional, post-enrichment linearization step for reliable MS/MS-based sequencing<sup>58,59</sup>. We envision that progress in these areas, along with improved mass spectral methods to enable investigation of libraries of even greater diversities, may ultimately facilitate discovery of fully non-natural peptide binders to historically undruggable targets.

## Methods

**Materials.** H-Rink Amide-ChemMatrix resin was purchased from PCAS BioMatrix Inc. (St-Jean-sur-Richelieu, Quebec, Canada). In all, 30  $\mu\text{m}$  TentaGel M NH<sub>2</sub> microspheres (M30352; 0.20–0.25 mmol/g amine loading) were purchased from Rapp Polymere (Tübingen, Germany). In all, 20  $\mu\text{m}$  TentaGel S NH<sub>2</sub> microspheres (TMN-9909-PI; 0.2–0.3 mmol/g amine loading) was purchased from Peptides International (Louisville, KY). Fmoc-Ala-OH, Fmoc-Arg(Pbf)-OH, Fmoc-Asn(Trt)-OH, Fmoc-Asp(tBu)-OH, Fmoc-Gln(Trt)-OH, Fmoc-Glu(tBu)-OH, Fmoc-Gly-OH, Fmoc-His(Trt)-OH, Fmoc-Leu-OH, Fmoc-Lys(Boc)-OH, Fmoc-Met-OH, Fmoc-Phe-OH, Fmoc-Pro-OH, Fmoc-Ser(tBu)-OH, Fmoc-Thr(tBu)-OH, Fmoc-Trp(Boc)-OH, Fmoc-Tyr(tBu)-OH, and Fmoc-Val-OH were purchased from Advanced ChemTech (Louisville, KY). Fmoc-D-Asp(tBu)-OH, Fmoc-D-Gln(Trt)-OH, Fmoc-D-Leu-OH, and Fmoc-D-Lys(Boc)-OH were also purchased from Advanced ChemTech (Louisville, KY). 1-[Bis(dimethylamino)methylene]-1H-1,2,3-triazolo [4,5-b]pyridinium-3-oxid-hexafluorophosphate (HATU) was purchased from P3 BioSystems (Louisville, KY). 4-[(R,S)- $\alpha$ -[1-(9H-Fluoren-9-yl)-methoxyformamido]-2,4-dimethoxybenzyl]-phenoxyacetic acid (Fmoc-Rink amide linker), Fmoc-L-His(Boc)-OH, Fmoc- $\beta$ -Ala-OH, and Fmoc-L-Lys(Alloc)-OH, and di-tert-butyl dicarbonate were purchased from Chem-Impex International (Wood Dale, IL). Fmoc- $\beta$ -cyclopropyl-L-alanine, Fmoc- $\beta$ -cyclohexyl-L-alanine, Fmoc-L-norvaline, Fmoc-O-tert-butyl-L- $\beta$ -homoserine, Fmoc-O-tert-butyl-L- $\beta$ -homothreonine, Fmoc-L- $\alpha$ -aminoadipic acid  $\delta$ -tert-butyl ester, N $\alpha$ -Fmoc-Ny-Boc-L-2,4-diaminobutyric acid, N $\alpha$ -Fmoc-N $\delta$ -Boc-L-ornithine, Fmoc-3-(4-thiazolyl)-L-alanine, Fmoc-O-tert-butyl-L-trans-4-hydroxyproline, Fmoc-4-(Boc-amino)-L-phenylalanine, Fmoc-4-fluoro-L-phenylalanine, Fmoc-4-nitro-L-phenylalanine, and fluorescein isothiocyanate isomer I were also purchased from Chem-Impex International (Wood Dale, IL). Biotin-(PEG)4-NHS ester and biotin-(PEG)4-propionic acid were purchased from ChemPep Inc. (Wellington, FL). Peptide synthesis-grade N,N-dimethylformamide (DMF), dichloromethane (DCM), diethyl ether, HPLC-grade acetonitrile (MeCN), and HPLC-grade methanol (MeOH) were purchased from VWR International (Philadelphia, PA). Trifluoroacetic acid (TFA; for HPLC,  $\geq 99\%$ ), piperidine (ReagentPlus; 99%), triisopropylsilane (98%), 1,2-ethanedithiol ( $\geq 98\%$ ), phenylsilane (97%), tetrakis(triphenylphosphine)palladium(0) (99%), and N- $\alpha$ -Fmoc-O-benzyl-L-phosphoserine were purchased from MilliporeSigma (St. Louis, MO). Diisopropylethylamine (99.5%; biotech. grade; DIEA) was also purchased from MilliporeSigma, and purified by passage through an activated alumina column (Pure Process Technology solvent purification system; Nashua, NH). Water was deionized using a Milli-Q Reference water purification system (Millipore).

Peptone from casein, granulated yeast extract, glycerol, and imidazole were purchased from Merck. LB Broth (Miller) powder, ampicillin sodium salt, potassium phosphate dibasic ( $\geq 99.0\%$ ), potassium phosphate monobasic, magnesium chloride hexahydrate ( $\geq 99.0\%$ ),  $\beta$ -mercaptoethanol (BME;  $\geq 99.0\%$ ), and Trizma base ( $\geq 99.0\%$ ) were purchased from MilliporeSigma. Isopropyl  $\beta$ -D-Thiogalactopyranoside (IPTG) was purchased from PanReac AppliChem.

Mouse anti-hemagglutinin (HA) monoclonal antibody clone 12ca5 (anti-HA mAb 12ca5) and polyclonal human IgG<sub>1</sub> were purchased from Columbia Biosciences (Frederick, MD). HyClone™ Fetal Bovine Serum (SH30071.03HI, heat inactivated) was purchased from GE Healthcare Life Sciences (Logan, UT). Bovine serum albumin (BSA; RIA grade) and Tween 20 (reagent grade) were purchased from Amresco (Solon, OH). Dynabeads MyOne Streptavidin T1 magnetic microparticles were purchased from Invitrogen (Carlsbad, CA).

Phosphate buffered saline (10x, Molecular biology grade) was purchased from Corning. Tris(hydroxymethyl)aminomethane (Tris) was purchased from J.T. Baker. 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES;  $\geq 99.5\%$ ), sodium bicarbonate (ACS grade,  $\geq 99.7\%$ ), and magnesium chloride ( $\geq 98\%$ ) were purchased from MilliporeSigma. Tris(2-carboxyethyl)phosphine hydrochloride (TCEP) was purchased from Hampton Research (Aliso Viejo, CA). Sodium chloride (ACS grade) was purchased from Avantor. Guanidine hydrochloride (Technical grade) and sodium phosphate monobasic monohydrate (ACS grade) were purchased from Amresco.

**Manual fast-flow peptide synthesis.** Manual fast-flow synthesis<sup>60</sup> of peptide- $\alpha$ -carboxamides was carried out on a 0.1 mmol scale, using H-Rink amide-ChemMatrix resin (0.45 mmol/g). Reagents and solvents were delivered to a stainless steel reactor containing the resin by either an HPLC pump (DMF or 20% (v/v) piperidine in DMF) or a syringe pump (active esters of Fmoc-amino acids). The reactor was submerged in a water bath for the duration of the synthesis and the temperature was maintained at 70 °C. The procedure for each coupling cycle included: a 30 s coupling with a mixture of Fmoc-protected amino acid (1 mmol), HBTU (0.95 mmol), and diisopropylethyl amine (DIEA; 2.9 mmol, 500  $\mu\text{L}$ ) in 2.5 mL of DMF, at a flow rate of 6 mL/min (for the coupling of tryptophan and histidine, 190  $\mu\text{L}$  of DIEA was used to minimize racemization); 1 min DMF wash, at a flow rate of 20 mL/min; 20 sec deprotection with 20% (v/v) piperidine in DMF, at a flow rate of 20 mL/min; and 1 min DMF wash, at a flow rate of 20 mL/min. After each synthesis was complete, resins were washed with DCM (5x) and dried under reduced pressure.

**Automated fast-flow peptide synthesis.** Automated fast-flow synthesis<sup>61</sup> of peptide- $\alpha$ -carboxamides was carried out using H-Rink amide-ChemMatrix resin (0.18 mmol/g, 0.03 mmol scale for N-terminal domain of MDM2; 0.45 mmol/g, 0.1

mmol scale for all other peptides). Syntheses were carried out at 90 °C; amide bond formation was effected in 8 s; and Fmoc removal was carried out in 8 s with 20% (v/v) piperidine in DMF. Individual cycle times were each  $\sim 40$  s. After each synthesis was complete, resins were washed with DCM (5x) and dried under reduced pressure.

**Manual solid-phase synthesis of non-canonical peptides.** Manual, batch synthesis of peptide- $\alpha$ -carboxamides was carried out on a 0.05 mmol scale, using H-Rink amide-ChemMatrix resin (0.45 mmol/g). Resin was weighed into disposable fritted plastic syringes (Torviq), washed 3x with DMF, then swollen in DMF for  $\sim 1$  h. Fmoc-protected amino acids (5 eq) in HATU solution (0.38M, 4.5 eq) were activated with DIEA (15 eq) and added to the resin bed. Reaction scales were variable and are described in the Supplementary Information. Couplings were allowed to proceed for 20 min. At this time, reaction mixtures were drained and resins were washed 3x with DMF. Fmoc removal was carried out by treatment of the resin with 20% piperidine in DMF (1x flow wash; 2  $\times$  5 min batch treatments). Following removal of N-terminal Fmoc group, resins were washed 3  $\times$  5 mL with DMF, then 3  $\times$  5 mL with DCM.

**Split-and-pool synthesis of peptide libraries.** Split-and-pool synthesis was carried out on 30  $\mu\text{m}$  TentaGel resin (0.26 mmol/g) for 10<sup>8</sup>-member libraries, and 20  $\mu\text{m}$  TentaGel resin (0.26 mmol/g) for 10<sup>9</sup>-member libraries. Splits were performed by suspending the resin in DMF and dividing it evenly (via pipet) among 18 plastic fritted syringes on a manifold. Couplings were carried out as follows: solutions of Fmoc-protected amino acids, HATU (0.38M in DMF; 0.9 eq relative to amino acid), and DIEA (1.1 eq for histidine; 3 eq for all other amino acids) were each added to individual portions of resin. Couplings were allowed to proceed for 20 min. Resin portions were recombined, and washed with DMF. Fmoc removal was carried out by treatment of the resin with 20% piperidine in DMF (1x flow wash; 2x 5 min batch treatments). Resin was washed again with DMF.

**Peptide cleavage and global deprotection.** Global side chain deprotection and cleavage from solid support were carried out by treatment of dry resin with a solution of 94% (v/v) TFA, 2.5% (v/v) ethanedithiol, 2.5% (v/v) water, and 1.0% (v/v) triisopropylsilane, for 2 h at ambient temperature ( $\sim 1.5$  mL of deprotection solution/50 mg of resin). TFA was then evaporated under a stream of nitrogen, and crude peptide was precipitated by addition of cold diethyl ether. Precipitated peptide was triturated (3x) with cold diethyl ether, dissolved in 50/50 water/acetoneitrile (0.1% TFA), passed through a 0.2  $\mu\text{m}$  PTFE syringe filter, and lyophilized.

**Magnetic bead preparation.** In all, 100  $\mu\text{L}$  portions of MyOne Streptavidin T1 Dynabeads (10 mg/mL; 1 mg; 0.13 nmol IgG binding capacity) were transferred to 1.7 mL plastic centrifuge tubes, and placed in a magnetic separation rack (New England Biolabs, cat# S1506S). The beads were washed three times with blocking buffer (1 mg/mL BSA or 10% FBS, 0.02% Tween 20, 1x PBS), and then treated with portions of biotinylated target protein ( $\sim 1.2$  eq). The resulting suspensions were transferred to a rotating vertical mixer, and kept for 15 min at ambient temperature or 1 h at 4 °C. After this time, the beads were returned to the separating rack, the supernatant was removed, and the beads were washed 3x with blocking buffer.

**Magnetic bead capture of control binders.** In all, 1 mL solutions containing blocking buffer (1 mg/mL BSA, 0.02% Tween 20, 1x PBS), and varying concentrations and amounts of 12ca5 control binders were prepared in 1.7 mL plastic centrifuge tubes, and chilled on ice for 10 min (the 12ca5 binders were added from mixtures containing 1  $\mu\text{M}$ /peptide or 10 nM/peptide in 6 M guanidine hydrochloride, 200 mM phosphate, pH 7 buffer). The resulting chilled solutions were then added to 1 mg portions of 12ca5-functionalized magnetic beads, and the resulting suspensions were kept on a rotating vertical mixer (1 h, in 4 °C cold room). The centrifuge tubes containing the bead suspensions were then transferred to the magnetic separation rack. The beads were isolated, and washed three times with 1 mL each of chilled 1x PBS (beads were exposed to buffer for a total of  $\sim 6$  min). Then, each drained bead aliquot was treated with 2  $\times$  150  $\mu\text{L}$  of elution buffer (6 M guanidine hydrochloride, 200 mM phosphate, pH 7.0 buffer containing 1 fmol/ $\mu\text{L}$  Peptide Retention Time Calibration Standard (PRTC; Pierce, cat# 88320, for use as an internal reference in MS-based quantitation)). For some conditions featuring lower binder concentration, eluates were concentrated via C18 ZipTip® pipette tips prior to nLC-MS analysis.

**Affinity selection from random libraries.** Library (typically 10 fmol/member) was incubated with 100  $\mu\text{L}$  (1 mg) portions of protein-immobilized magnetic beads in the presence of 10% FBS, 1x PBS (final volume: 1 mL) on a rotating mixer for 1 h at 4 °C. Typical final conditions: 1 mg/mL magnetic beads, 10 pM/member library. The centrifuge tubes containing the bead suspensions were transferred to the magnetic separation rack. The beads were washed 3  $\times$  1 mL w/ 1x PBS. Bound peptides were eluted with 2  $\times$  100  $\mu\text{L}$  6M guanidine hydrochloride, 200 mM phosphate, pH 6.8. Eluates were concentrated via C18 ZipTip® pipette tips and lyophilized. Powders were typically resuspended in 6  $\mu\text{L}$  water (0.1% formic acid), and 5  $\mu\text{L}$  were submitted for nLC-MS/MS analysis.

**nLC-MS/MS analysis.** Analysis was performed on an EASY-nLC 1200 (Thermo Fisher Scientific) nano-liquid chromatography handling system connected to an Orbitrap Fusion Lumos Tribrid Mass Spectrometer (Thermo Fisher Scientific). Samples were run on a PepMap RSLC C18 column (2  $\mu\text{m}$  particle size, 15  $\text{cm} \times 50 \mu\text{m}$  ID; Thermo Fisher Scientific, P/N ES801). A nanoViper Trap Column (C18, 3  $\mu\text{m}$  particle size, 100  $\text{\AA}$  pore size, 20  $\text{mm} \times 75 \mu\text{m}$  ID; Thermo Fisher Scientific, P/N 164946) was used for desalting. The standard nano-LC method was run at 40  $^{\circ}\text{C}$  and a flow rate of 300  $\text{nL}/\text{min}$  with the following gradient: 1% solvent B in solvent A ramping linearly to 61% B in A over 40 or 60 min, where solvent A = water (0.1% FA), and solvent B = 80% acetonitrile, 20% water (0.1% FA). Positive ion spray voltage was set to 2200 V. Orbitrap detection was used for primary MS, with the following parameters: resolution = 120,000; quadrupole isolation; scan range = 200–1400  $\text{m}/\text{z}$ ; RF lens = 30%; AGC target =  $1 \times 10^6$ ; maximum injection time = 100 ms; 1 microscan. Acquisition of secondary MS spectra was done in a data-dependent manner: dynamic exclusion was employed such that a precursor was excluded for 30 s if it was detected four or more times within 30 s (mass tolerance: 10.00 ppm); monoisotopic precursor selection used to select for peptides; intensity threshold was set to  $5 \times 10^4$ ; charge states 2–10 were selected; and precursor selection range was set to 200–1400  $\text{m}/\text{z}$ . The top 15 most intense precursors that met the preceding criteria were subjected to subsequent fragmentation. Three fragmentation modes—collision-induced dissociation (CID), higher-energy collisional dissociation (HCD), and electron-transfer/higher-energy collisional dissociation (ETHCd)—were used for acquisition of secondary MS spectra. Only precursors with charge states 3 and above were subjected to all three fragmentation modes; precursors with charge states of 2 were subjected to CID and HCD only. For all three modes, detection was performed in the Orbitrap (resolution = 30,000; quadrupole isolation; isolation window = 1.3  $\text{m}/\text{z}$ ; AGC target =  $2 \times 10^4$ ; maximum injection time = 100 ms; 1 microscan). For CID, a collision energy of 30% was used. For HCD, a collision energy of 25% was used. For ETHCd, a supplemental activation collision energy of 25% was used.

**Direct fluorescence anisotropy.** All direct fluorescence anisotropy affinity measurements were conducted in FA buffer (10 mM HEPES pH 7.4, 150 mM NaCl, 0.1% (v/v) Tween 20, 0.1% (w/v) BSA). A 1:1 dilution series of 14-3-3 $\gamma$  (starting at 100  $\mu\text{M}$ ) was performed in polystyrene non-binding low-volume Corning Black Round Bottom 384-well plates (Corning 4514) containing a fixed concentration FITC-labeled peptide (10 nM or 50 nM). Measurements were performed at ambient temperature using a Tecan Infinite F500 plate reader with the following parameters:  $\lambda_{\text{ex}}$ : 485 (20) nm;  $\lambda_{\text{em}}$ : 535 (25) nm; mirror: Dichroic 510; flashes: 20; integration time: 50  $\mu\text{s}$ ; settle time: 0  $\mu\text{s}$ ; gain: manual 90; Z-position: calculated from well. The G-factor was set at 35 mP based on wells containing only the FITC-labeled peptide.

**Fluorescence anisotropy competition.** All fluorescence anisotropy competition experiments were conducted in FA buffer (10 mM HEPES pH 7.4, 150 mM NaCl, 0.1% (v/v) Tween-20, 0.1% (w/v) BSA). A 1:1 dilution series of unlabeled peptides (starting at 10  $\mu\text{M}$ ) were performed in wells containing a fixed concentration FITC-labeled biExoS (10 nM) and 14-3-3 $\gamma$  (20 nM). Measurements were performed at ambient temperature using a Tecan Infinite F500 plate reader with the following parameters:  $\lambda_{\text{ex}}$  = 485 (20) nm;  $\lambda_{\text{em}}$  = 535 (25) nm, mirror: Dichroic = 510, flashes: 20; integration time: 50  $\mu\text{s}$ ; settle time: 0  $\mu\text{s}$ ; gain: manual 90; and Z-position: calculated from well. The G-factor was set at 35 mP based on wells containing only the FITC-labeled peptide.

**X-ray structure determination.** Unlabeled 14-3-3.12 was soaked into preformed crystals of 14-3-3 $\sigma\Delta\text{C}$ , which grew in 25% PEG400, 5% Glycerol, 0.2 M CaCl<sub>2</sub>, 0.1 M HEPES pH 7.5 plus 2 mM BME within two weeks. The soaked crystal was fished after 15 days of incubation and flash-frozen in liquid nitrogen. Diffraction data was collected at 100K on an in-house Rigaku Micromax-003 (Rigaku Europe, Kemsing Sevenoaks, UK) sealed tube X-ray source and Dectris Pilatus 200K detector (DECTRIS Ltd Baden-Daettwil, Switzerland). Integration, scaling and merging of data was done using DIALS (CCP4i2) after which molecular replacement is done with MOLREP (CCP4i2) using PDB 4JC3 as search model. A three-dimensional structure of 14-3-3.12 was generated using eLBOW (Phenix) after which it was built within this structure based on visual inspection of Fo-Fc and 2FoFc electron density maps in Coot. Several rounds of model building and refinement (based on isotropic b-factors and standard set of stereo-chemical restraints: covalent bonds, angles, dihedrals, planarities, chiralities, non-bonded) were performed using Coot and Phenix.refine. See Supplementary Table 12 for data collection and refinement statistics. Structural coordinates have been deposited in the PDB (6TCH).

## Data availability

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request. The source data underlying Figs. 2b, 3b, and 4b, and Supplementary Figs. 12, 22–24, 27, and 30 are provided as Source data file. Source data are provided with this paper.

Received: 2 December 2019; Accepted: 27 May 2020;

Published online: 23 June 2020

## References

- Mignani, S., Huber, S., Tomás, H., Rodrigues, J. & Majoral, J.-P. Why and how have drug discovery strategies in pharma changed? What are the new mindsets? *Drug Discov. Today* **21**, 239–249 (2016).
- Erlanson, D. A., McDowell, R. S. & O'Brien, T. Fragment-based drug discovery. *J. Med. Chem.* **47**, 3463–3482 (2004).
- Gebauer, M. & Skerra, A. Engineered protein scaffolds as next-generation antibody therapeutics. *Curr. Opin. Chem. Biol.* **13**, 245–255 (2009).
- Tsomaia, N. Peptide therapeutics: targeting the undruggable space. *Eur. J. Medicinal Chem.* **94**, 459–470 (2015).
- Vinogradov, A. A., Yin, Y. & Suga, H. Macrocyclic peptides as drug candidates: recent progress and remaining challenges. *J. Am. Chem. Soc.* **141**, 4167–4181 (2019).
- Grossmann, T. N. et al. Inhibition of oncogenic Wnt signaling through direct targeting of  $\beta$ -catenin. *Proc. Natl Acad. Sci. USA* **109**, 17942–17947 (2012).
- Chang, Y. S. et al. Stapled  $\alpha$ -helical peptide drug development: a potent dual inhibitor of MDM2 and MDMX for p53-dependent cancer therapy. *Proc. Natl Acad. Sci. USA* **110**, E3445–E3454 (2013).
- Leshchiner, E. S. et al. Direct inhibition of oncogenic KRAS by hydrocarbon-stapled SOS1 helices. *Proc. Natl Acad. Sci. USA* **112**, 1761–1766 (2015).
- Rezaei, T., Yu, B., Millhauser, G. L., Jacobson, M. P. & Lokey, R. S. Testing the conformational hypothesis of passive membrane permeability using synthetic cyclic peptide diastereomers. *J. Am. Chem. Soc.* **128**, 2510–2511 (2006).
- Walensky, L. D. & Bird, G. H. Hydrocarbon-stapled peptides: principles, practice, and progress. *J. Med. Chem.* **57**, 6275–6288 (2014).
- Bird, G. H. et al. Biophysical determinants for cellular uptake of hydrocarbon-stapled peptide helices. *Nat. Chem. Biol.* **12**, 845–852 (2016).
- Touti, F., Gates, Z. P., Bandyopadhyay, A., Lautrette, G. & Pentelute, B. L. In-solution enrichment identifies peptide inhibitors of protein–protein interactions. *Nat. Chem. Biol.* **15**, 410–418 (2019).
- Rogers, J. M., Passioura, T. & Suga, H. Nonproteinogenic deep mutational scanning of linear and cyclic peptides. *Proc. Natl Acad. Sci. USA* **115**, 10959–10964 (2018).
- Clackson, T. & Wells, J. A. In vitro selection from protein and peptide libraries. *Trends Biotechnol.* **12**, 173–184 (1994).
- Kay, B. K., Kurakin, A. V. & Hyde-DeRuyscher, R. From peptides to drugs via phage display. *Drug Discov. Today* **3**, 370–378 (1998).
- Wilson, D. S., Keefe, A. D. & Szostak, J. W. The use of mRNA display to select high-affinity protein-binding peptides. *Proc. Natl Acad. Sci. USA* **98**, 3750–3755 (2001).
- Obexer, R., Walport, L. J. & Suga, H. Exploring sequence space: harnessing chemical and biological diversity towards new peptide leads. *Curr. Opin. Chem. Biol.* **38**, 52–61 (2017).
- Katoh, T., Tajima, K. & Suga, H. Consecutive elongation of D-Amino acids in translation. *Cell Chem. Biol.* **24**, 46–54 (2017).
- Katoh, T. & Suga, H. Ribosomal incorporation of consecutive  $\beta$ -amino acids. *J. Am. Chem. Soc.* **140**, 12159–12167 (2018).
- Eidam, O. & L. Satz, A. Analysis of the productivity of DNA encoded libraries. *MedChemComm* **7**, 1323–1331 (2016).
- Zhao, G., Huang, Y., Zhou, Y., Li, Y. & Li, X. Future challenges with DNA-encoded chemical libraries in the drug discovery domain. *Expert Opin. Drug Discov.* **14**, 735–753 (2019).
- Lam, K. S. et al. A new type of synthetic peptide library for identifying ligand-binding activity. *Nature* **354**, 82–84 (1991).
- Lam, K. S., Lebl, M. & Krchňák, V. The “One-Bead-One-Compound” Combinatorial Library Method. *Chem. Rev.* **97**, 411–448 (1997).
- Gates, Z. P. et al. Xenoprotein engineering via synthetic libraries. *Proc. Natl Acad. Sci. USA* **115**, E5298–E5306 (2018).
- Zuckermann, R. N., Kerr, J. M., Siani, M. A., Banville, S. C. & Santi, D. V. Identification of highest-affinity ligands by affinity selection from equimolar peptide mixtures generated by robotic synthesis. *Proc. Natl Acad. Sci. USA* **89**, 4505–4509 (1992).
- Dunayevskiy, Y. M., Lai, J.-J., Quinn, C., Talley, F. & Vouros, P. Mass spectrometric identification of ligands selected from combinatorial libraries using gel filtration. *Rapid Commun. Mass Spectrom.* **11**, 1178–1184 (1997).
- Kaur, S., McGuire, L., Tang, D., Dollinger, G. & Huebner, V. Affinity selection and mass spectrometry-based strategies to identify lead compounds in combinatorial libraries. *J. Protein Chem.* **16**, 505–511 (1997).
- van Breemen, R. B. et al. Pulsed ultrafiltration mass spectrometry: a new method for screening combinatorial libraries. *Anal. Chem.* **69**, 2159–2164 (1997).

29. Maaty, W. S. & Weis, D. D. Label-free, in-solution screening of peptide libraries for binding to protein targets using hydrogen exchange mass spectrometry. *J. Am. Chem. Soc.* **138**, 1335–1343 (2016).
30. Vinogradov, A. A. et al. Library design-facilitated high-throughput sequencing of synthetic peptide libraries. *ACS Comb. Sci.* **19**, 694–701 (2017).
31. Jiang, J. et al. Development of an immuno tandem mass spectrometry (iMALDI) assay for EGFR diagnosis. *Proteom. Clin. Appl.* **1**, 1651–1659 (2007).
32. Li, H., Popp, R. & Borchers, C. H. Affinity-mass spectrometric technologies for quantitative proteomics in biological fluids. *TrAC Trends Anal. Chem.* **90**, 80–88 (2017).
33. Choi, Y. & van Breemen, R. B. Development of a screening assay for ligands to the estrogen receptor based on magnetic microparticles and LC-MS. *Comb. Chem. High. Throughput Screen.* **11**, 1–6 (2008).
34. Rush, M. D., Walker, E. M., Prehna, G., Burton, T. & van Breemen, R. B. Development of a magnetic microbead affinity selection screen (MagMASS) using mass spectrometry for ligands to the retinoid X receptor- $\alpha$ . *J. Am. Soc. Mass Spectrom.* **28**, 479–485 (2017).
35. Sannino, A. et al. Quantitative assessment of affinity selection performance by using DNA-encoded chemical libraries. *ChemBioChem* **20**, 955–962 (2019).
36. Furka, A., Sebastyén, F., Asgedom, M. & Dibó, G. General method for rapid synthesis of multicomponent peptide mixtures. *Int. J. Pept. Protein Res.* **37**, 487–493 (1991).
37. Ma, B. et al. PEAKS: powerful software for peptidome novo sequencing by tandem mass spectrometry. *Rapid Commun. Mass Spectrom.* **17**, 2337–2342 (2003).
38. Churchill, M. E. et al. 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 (1994).
39. Olson, C. A. et al. Single-round, multiplexed antibody mimetic design through mRNA display. *Angew. Chem. Int. Ed. Engl.* **51**, 12449–12453 (2012).
40. Georgiou, G. et al. Display of heterologous proteins on the surface of microorganisms: From the screening of combinatorial libraries to live recombinant vaccines. *Nat. Biotechnol.* **15**, 29–34 (1997).
41. Feldhaus, M. J. et al. Flow-cytometric isolation of human antibodies from a nonimmune *Saccharomyces cerevisiae* surface display library. *Nat. Biotechnol.* **21**, 163–170 (2003).
42. Alluri, P. G., Reddy, M. M., Bachhawat-Sikder, K., Olivos, H. J. & Kodadek, T. Isolation of protein ligands from large peptoid libraries. *J. Am. Chem. Soc.* **125**, 13995–14004 (2003).
43. Böttger, V. et al. Identification of novel mdm2 binding peptides by phage display. *Oncogene* **13**, 2141–2147 (1996).
44. Hu, B., Gilkes, D. M. & Chen, J. Efficient p53 activation and apoptosis by simultaneous disruption of binding to MDM2 and MDMX. *Cancer Res.* **67**, 8810–8817 (2007).
45. Pazgier, M. et al. Structural basis for high-affinity peptide inhibition of p53 interactions with MDM2 and MDMX. *Proc. Natl Acad. Sci. USA* **106**, 4665–4670 (2009).
46. Li, C. et al. Systematic mutational analysis of peptide inhibition of the p53-MDM2/MDMX interactions. *J. Mol. Biol.* **398**, 200–213 (2010).
47. Kritzer, J. A., Luedtke, N. W., Harker, E. A. & Schepartz, A. A rapid library screen for tailoring  $\beta$ -peptide structure and function. *J. Am. Chem. Soc.* **127**, 14584–14585 (2005).
48. Stevers, L. M. et al. Modulators of 14-3-3 protein–protein interactions. *J. Med. Chem.* **61**, 3755–3778 (2018).
49. Yaffe, M. B. et al. The structural basis for 14-3-3:phosphopeptide binding specificity. *Cell* **91**, 961–971 (1997).
50. de Vink, P. J. et al. A binary bivalent supramolecular assembly platform based on Cucurbit[8]uril and dimeric adapter protein 14-3-3. *Angew. Chem. Int. Ed. Engl.* **56**, 8998–9002 (2017).
51. Efimov, A. V. Standard structures in proteins. *Prog. Biophysics Mol. Biol.* **60**, 201–239 (1993).
52. Griffiths, A. D. & Duncan, A. R. Strategies for selection of antibodies by phage display. *Curr. Opin. Biotechnol.* **9**, 102–108 (1998).
53. Zhang, Y., Fonslow, B. R., Shan, B., Baek, M.-C. & Yates, J. R. Protein analysis by shotgun/bottom-up proteomics. *Chem. Rev.* **113**, 2343–2394 (2013).
54. Mándity, I. M. & Fülöp, F. An overview of peptide and peptoid foldamers in medicinal chemistry. *Expert Opin. Drug Discov.* **10**, 1163–1177 (2015).
55. Nizami, B. et al. FoldamerDB: a database of peptidic foldamers. *Nucleic Acids Res* **48**, D1122–D1128 (2020).
56. Heinis, C., Rutherford, T., Freund, S. & Winter, G. Phage-encoded combinatorial chemical libraries based on bicyclic peptides. *Nat. Chem. Biol.* **5**, 502–507 (2009).
57. Goto, Y. et al. Reprogramming the translation initiation for the synthesis of physiologically stable cyclic peptides. *ACS Chem. Biol.* **3**, 120–129 (2008).
58. Lee, J. H., Meyer, A. M. & Lim, H.-S. A simple strategy for the construction of combinatorial cyclic peptoid libraries. *Chem. Commun.* **46**, 8615–8617 (2010).
59. Liang, X., Vézina-Dawod, S., Bédard, F., Porte, K. & Biron, E. One-pot photochemical ring-opening/cleavage approach for the synthesis and decoding of cyclic peptide libraries. *Org. Lett.* **18**, 1174–1177 (2016).
60. Simon, M. D. et al. Rapid flow-based peptide synthesis. *ChemBioChem* **15**, 713–720 (2014).
61. Mijalis, A. J. et al. A fully automated flow-based approach for accelerated peptide synthesis. *Nat. Chem. Biol.* **13**, 464–466 (2017).

## Acknowledgements

This work was supported by the NIH/NIGMS Interdepartmental Biotechnology Training Program (T32 GM008334 to A.J.Q.), the Defense Advanced Research Projects Agency (DARPA; Award 023504-001 to B.L.P.), and Calico (to B.L.P.). We gratefully acknowledge Anne Fischer and Tyler Stukenbroeker (DARPA) for their support and guidance; Earl Moore, Mark Paul, Louis Abruzzese, and David Sarracino for their technical assistance with nanoLC and Orbitrap mass spectrometry; and Eric Spooner, Marko Jovanovic, and Dan Maloney for their discussions regarding MS-based analysis and sequencing. We thank Joseph Brown for assistance with library synthesis, Suan Tuang for assistance with automated peptide synthesis, and Faycal Touti, Ethan Evans, and Alex Vinogradov for many fruitful scientific discussions.

## Author contributions

A.J.Q., Z.P.G., B.A.S., and B.L.P. designed research; A.J.Q., Z.P.G., B.A.S., N.H., X.Y., and A.S. performed research, with support from B.L.P., C.O., and Y.K.; A.J.Q., Z.P.G., B.A.S., C.O., and B.L.P. analyzed data; A.J.Q., Z.P.G., and B.A.S. generated figures; and A.J.Q. and Z.P.G. wrote the paper with support from B.L.P.

## Competing interests

B.L.P. is a co-founder of Amide Technologies and Resolute Bio. Both companies focus on the development of protein and peptide therapeutics.

## Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41467-020-16920-3>.

**Correspondence** and requests for materials should be addressed to B.L.P.

**Peer review information** *Nature Communications* thanks Richard Breemen, Ronald Zuckermann and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020