ACS central science

http://pubs.acs.org/journal/acscii

Research Article

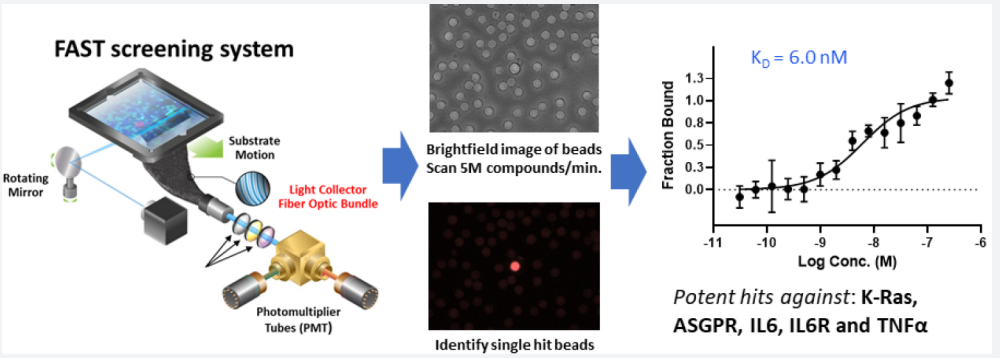
Mega-High-Throughput Screening Platform for the Discovery of Biologically Relevant Sequence-Defined Non-Natural Polymers

Michal Avital-Shmilovici, Xiaohe Liu, Thomas Shaler, Andrew Lowenthal, Pauline Bourbon, Janey Snider, Arlyn Tambo-Ong, Claire Repellin, Kenya Yniguez, Lidia Sambucetti, Peter B. Madrid,* and Nathan Collins*

Cite This: *ACS Cent. Sci.* 2022, 8, 86–101

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information



ABSTRACT: Combinatorial methods enable the synthesis of chemical libraries on scales of millions to billions of compounds, but the ability to efficiently screen and sequence such large libraries has remained a major bottleneck for molecular discovery. We developed a novel technology for screening and sequencing libraries of synthetic molecules of up to a billion compounds in size. This platform utilizes the fiber-optic array scanning technology (FAST) to screen bead-based libraries of synthetic compounds at a rate of 5 million compounds per minute (~83 000 Hz). This ultra-high-throughput screening platform has been used to screen libraries of synthetic "self-readable" non-natural polymers that can be sequenced at the femtomole scale by chemical fragmentation and high-resolution mass spectrometry. The versatility and throughput of the platform were demonstrated by screening two libraries of non-natural polyamide polymers with sizes of 1.77M and 1B compounds against the protein targets K-Ras, asialoglycoprotein receptor 1 (ASGPR), IL-6, IL-6 receptor (IL-6R), and TNFα. Hits with low nanomolar binding affinities were found against all targets, including competitive inhibitors of K-Ras binding to Raf and functionally active uptake ligands for ASGPR facilitating intracellular delivery of a nonglycan ligand.

INTRODUCTION

Natural biological polymers such as peptides, proteins, and nucleic acids have evolved molecular recognition functionalities to produce highly specific binding interactions and catalytic enzymatic functions. In an empirical approach to discover novel affinity agents, catalysts, and materials, biotechnologists have exploited these rich functional properties to create therapeutics, diagnostics, sensors, industrial reagents, and biomedical research probes by using powerful biological tools to screen large libraries of natural polymers of 10^{7-12} unique molecular sequences.[1−4] As data regarding the structures of biopolymers have amassed, so has the ability to rationally design *de novo* proteins and other biopolymers.[5−7]

There is a growing interest in expanding polymer discovery into the field of sequence-defined non-natural polymers or foldamers[8,9] using chemical rather than biological synthesis to further access molecular diversity in novel 3D folded structures with related functionality as affinity reagents, therapeutics, and catalysts.[10−16] At present, there is insufficient sequence and related structural data to design non-natural polymers *de novo*; as such, screening libraries to find polymers with desired properties are currently the most practical solution. An example using modified amino acid building blocks is the random nonstandard peptides integrated discovery (RaPID) system which integrates genetic code reprogramming with mRNA display technology to produce massive libraries (1012) of cyclic peptides.[17] While one of the most interesting methods of generating chemically diverse

Received: August 25, 2021
Published: January 11, 2022

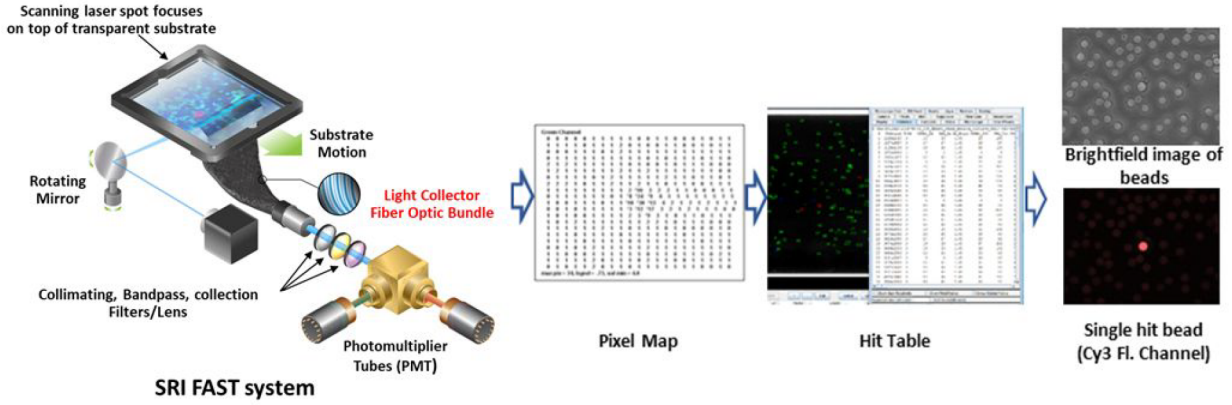


Figure 1. Diagram of the SRI fiber-optic array scanning technology (FAST) system. The FAST system uses rapid laser scanning with sensitive photomultiplier tube (PMT) fluorescence emission detection to rapidly generate a pixel map indicating the position of fluorescently labeled beads. An analysis of the pixel map generates a hit table with Cartesian coordinates and multiple calculated fluorescence metrics to detect hit beads with high sensitivity and specificity. The coordinates of the hits can then be transferred to other microscopy systems for an additional multiwavelength imaging analysis or bead extraction.

cyclic peptide libraries, it is still largely constrained to using L-a amino acid building blocks.

A chemical synthetic methodology for building libraries of peptide-based polymers is well-established on solid support beads via the "one-bead one-compound" (OBOC) method.[18] This combinatorial "mix and split" method has the potential to make chemical polymer libraries of similar size and complexity to biological methods; however, in practice, it has traditionally been used to produce relatively small libraries on the scale of only thousands to hundreds of thousands of compounds because of the two major hurdles of library screening and hit sequencing.[19] OBOC libraries can be synthesized with total chemical control, and the relatively harsh chemical reaction conditions of organic synthesis are well-tolerated compared to more biological library production methods.[20] To expand the application of OBOC libraries to the discovery of sequence-defined non-natural polymers, it is necessary to address the constraints of OBOC screening throughput for appropriately scaled molecular diversity and the requirement to handle low (ideally picomole or lower) amounts of hit compounds for sequence identification.

The current commercially available technology with the highest throughput for bead screening is fluorescence-activated cell sorting (FACS),[21,22] which has a theoretical throughput of $\sim 10^8$ in a 10 h period. In practice, however, hit enrichment from large libraries ($>10^6$) using pull-down methods is the only way to achieve time- and cost-effective FACS screening that reduces the number of beads to 10^{4-5}.[22] An alternative approach described by Carney et al.[23] used confocal laser scanning microscopy (CLSM) to screen beads immobilized on a 10 cm × 10 cm polystyrene surface. The autofluorescence of each bead (F_0) was measured first and compared with the fluorescence of labeled beads during screening (ΔF) to identify the brightest beads ($\Delta F/F_0$). By this approach, they could screen 200 000 beads in 20 min in this fluorometric assay and demonstrated its application in a multiplexed screen of 157 423 beads from a 9-mer peptide library on 90 μm beads. Setting the hit rate at 0.01%, they identified 22 hits of which the top 4 were sequenced and confirmed in subsequent assays. Recently, Quartararo et al.[24] demonstrated a synthesis of a 10^8 member peptide library and a screening strategy using affinity selection−mass spectrometry (AS−MS) methodology.[25−29] In-solution affinity selection was combined with nanoliquid chromatography−tandem mass spectrometry peptide sequencing to identify the highest-affinity binders. In this method, the target was immobilized onto magnetic beads, and potential binders from the library were pulled down and sequenced by liquid chromatography−tandem mass spectrometry. The peptides in the libraries were 10 amino acids long with nine diversity positions. The library was designed to have a theoretical diversity of 2×10^{11}, but since it was made on 30 μm TentaGel resin utilizing 2.9 g of resin, it contained only a fraction of the library—2×10^8 beads and peptides (with no compound redundancy in the library). The amino acids used for the library synthesis contained natural and non-natural amino acids, all with the standard scaffold of α-amino acids.

Sequencing of α-amino acid library hits on beads to identify peptides is routinely done by liquid chromatography with tandem mass spectrometry (LC−MS/MS)[30,31] or Edman degradation on a protein sequencer,[31,32] which works well for short oligomer peptides but does not translate well to novel non-α-amino acid backbone polymers. Many methods have been developed to improve the sequencing process such as introducing fluorescent dyes,[33] isotopic tags,[34] DNA encoding,[35] ladder-sequencing,[36−39] and chemical encoding methods.[40] However, none of these sequencing methods are sensitive enough for libraries synthesized on beads less than 20−30 μm in diameter[22] which contain ~4 picomoles of polymer/bead and for libraries with non-α-amino acid backbone polymers. Historically, OBOC libraries have typically used bead sizes of >90 μm containing >100 picomoles of polymer per bead to ensure that there is sufficient material for hit identification[20] (Table S1). Large libraries of beads with these diameters are prohibitively expensive in materials costs to synthesize and screen, particularly for high-molecular-weight polymers, which is another reason why such chemical libraries' sizes have been limited to date.

Here, we present a novel platform that enables the production of large libraries of synthetic, sequence-defined non-natural polymers (NNPs) on the scale of 10^7−10^9 members for megathroughput screening using a platform based on a fiber-optic array scanning technology (FAST) that screens up to ~5 million polymers a minute. Furthermore, we describe a method for sequencing single bead hits down to 10 μm in diameter with femtomole sensitivity. We demonstrate the platform's broad use

https://doi.org/10.1021/acscentsci.1c01041
ACS Cent. Sci. 2022, 8, 86−101

in screening against five targets of biomedical interest to identify biologically relevant non-natural polymers with affinities in the nanomolar to subnanomolar range that can inhibit protein−protein interactions (PPIs) and protein−glycan interactions and have exceptional biological activity and stability.

RESULTS

FAST Screening Approach. FAST was originally developed to identify rare circulating cancer cells in blood with high sensitivity and specificity (Figure 1).[41−46] In this application, cells preincubated with fluorescently labeled cell surface markers are plated as a monolayer on 108 × 76 mm glass slides, which are then scanned by excitation with a 488 nm laser. Emitted fluorescence is collected through a fiber-optic bundle, and the collected light is passed through a bandpass filter and analyzed by a photomultiplier tube to measure emission at 520 nm (green) and 580 nm (red/orange) to eliminate true negatives due to autofluorescence (see below). Cartesian coordinates of fluorescently labeled objects are located on a pixel map, along with fluorescent intensity measurements at the two emission wavelengths. In this well-free assay format, FAST can routinely identify the location of single rare cells in a milieu of 25 million white cells in a 1 min scan with an ∼8 μm resolution. In optimizing the FAST system for bead screening, the only major modification to the scanning process was the need to plate beads at a lower density than cells due to their propensity to aggregate and the need to extract them postanalysis for sequencing. Empirical optimization of bead plating density revealed that 10 μm diameter TentaGel beads plated with a density of 5 million beads per plate gave a relatively well-dispersed monolayer enabling automated analysis and bead picking for downstream processing. Similarly, 20 μm beads were optimally plated at a density of 2.5 million beads per plate. Detection sensitivity was assessed by spiking biotin-labeled beads into a pool of underivatized beads and incubating with Alexa Fluor 555-labeled streptavidin for 1 h before plating. The FAST screening process gave a detection sensitivity of over 99.99% (Table S2 and Figure S1).

In considering the application of FAST to bead-based screening, two major problems complicate the efficient fluorescence-based screening of TentaGel OBOC libraries: one problem is that the autofluorescence of TentaGel beads leads to low signal-to-noise ratios and complicates the identification of hits. The FAST screening approach uses several strategies to overcome problems due to bead autofluorescence based on optical properties of the TentaGel resin. Some key observations include the fact that TentaGel autofluorescence is highly significant in the FITC (fluorescein isothiocyanate) channel, and the fluorescence intensity diminishes as its wavelength shift increases;[47−49] autofluorescence intensifies with increasing bead size. In our hands, functionalized beads with different chemistries have various levels of autofluorescence which in general are slightly higher than the autofluorescence of unfunctionalized beads. As mentioned above, the size of beads we use for our library construction in this study is 10−20 μm in diameter (comparable to a mammalian-cell size), and the autofluorescence is significantly lower than those for the beads commonly used in other OBOC libraries (e.g., 90−300 μm). Second, a more favorable fluorophore (Alexa-fluor 555 or CF555, yellow/orange) for target probes is used in conjunction with a wavelength comparison technique engineered in the FAST system to eliminate the effects of fluorescence from autofluorescing particles.[50] The technique involves measuring emissions at two different wavelengths, one at the target emission wavelength (580 nm) and the other at 520 nm, a wavelength intermediate between the target emission wavelength and the laser excitation (488 nm). Because autofluorescence is typically more intense at wavelengths closer to the excitation, the ratio of the intermediate wavelength intensity to the target wavelength intensity is greater than one for unlabeled beads while for labeled beads the ratio is less than one. A software filter uses this ratio to eliminate the autofluorescing beads. The software filter also screens for and eliminates fluoresce-positive objects originating from dye aggregates and bead fragments by filtering for object size and relative brightness. We set negative controls in parallel for each screen. Negative controls include fluorescence cutoffs determined from unfunctionalized naked TentaGel beads that have gone through the assay staining process with labeled probe and a portion of the library beads taken through the assay staining protocol without the probe. As part of the comprehensive filter settings, the fluorescence intensity cutoff threshold is set to eliminate the selection of autofluorescent beads due to the TentaGel or library background (filter threshold setting details are described in Tables S6−S11). With this filter, 99.8% of the located objects on the glass slide can be eliminated as true negatives and are not selected for further analysis. With these filters, the typical hit number from a FAST scan is around 100−400 identified as corresponding coordinate locations on the glass slide from a sample containing 2.5 million 20 μm beads or 5 million 10 μm beads.

The hits identified by the FAST primary scan are then automatically imaged and analyzed by high-resolution automated digital microscopy (ADM) on a CellCelector instrument (ALS Automated Lab Solution GmbH) using bright-field, target Alexa Fluor 555 (AF555) or CF555 and counter target AF647/Cy5 channels. Hit beads are QC/QA reviewed based on morphology and fluorescence staining data. Damaged beads, beads with irregular shape, size, or staining pattern, and hit beads located within a large aggregate and impossible to exact are excluded. The mean fluorescence intensity (MFI) is then measured for all hits that pass initial QC/QA. All "true positive" (TP) hits are ranked based on MFI intensity and/or ratio of selected channels, and generally the top ∼50 beads from the initial 10−400 FAST hits are selected and isolated for sequencing and hit confirmation by resynthesis and K_D characterization.

Another problem with OBOC libraries is that during on-bead screening the signal strengths (e.g., fluorescence intensities) do not always correlate with the potency of the ligands on these beads. One of the contributing factors to this problem is that commercial resins typically used for library synthesis have high ligand loading (e.g., 90 μm TentaGel resin with a loading capacity of 0.3 mmol/g has a ligand density of ∼100 mM) which is necessary to provide a sufficient amount of material for subsequent hit identification but may cause false positives and screening biases due to the unintended multidentate interaction with high ligand density on the beads.[49] Chen et al. have showed that the decrease of ligand concentration on the beads leads to a significantly reduced number of false positives due to the reduction of nonspecific binding caused by avidity effects.[51] We are similarly able to minimize avidity effects by the use of smaller beads with less ligand loading. This also allows us to use lower probe concentrations in screening. For every screen, the probes are pretitrated to identify the minimum probe concentrations that achieve an optimal signal-to-noise ratio and hit numbers

https://doi.org/10.1021/acscentsci.1c01041
ACS Cent. Sci. 2022, 8, 86−101

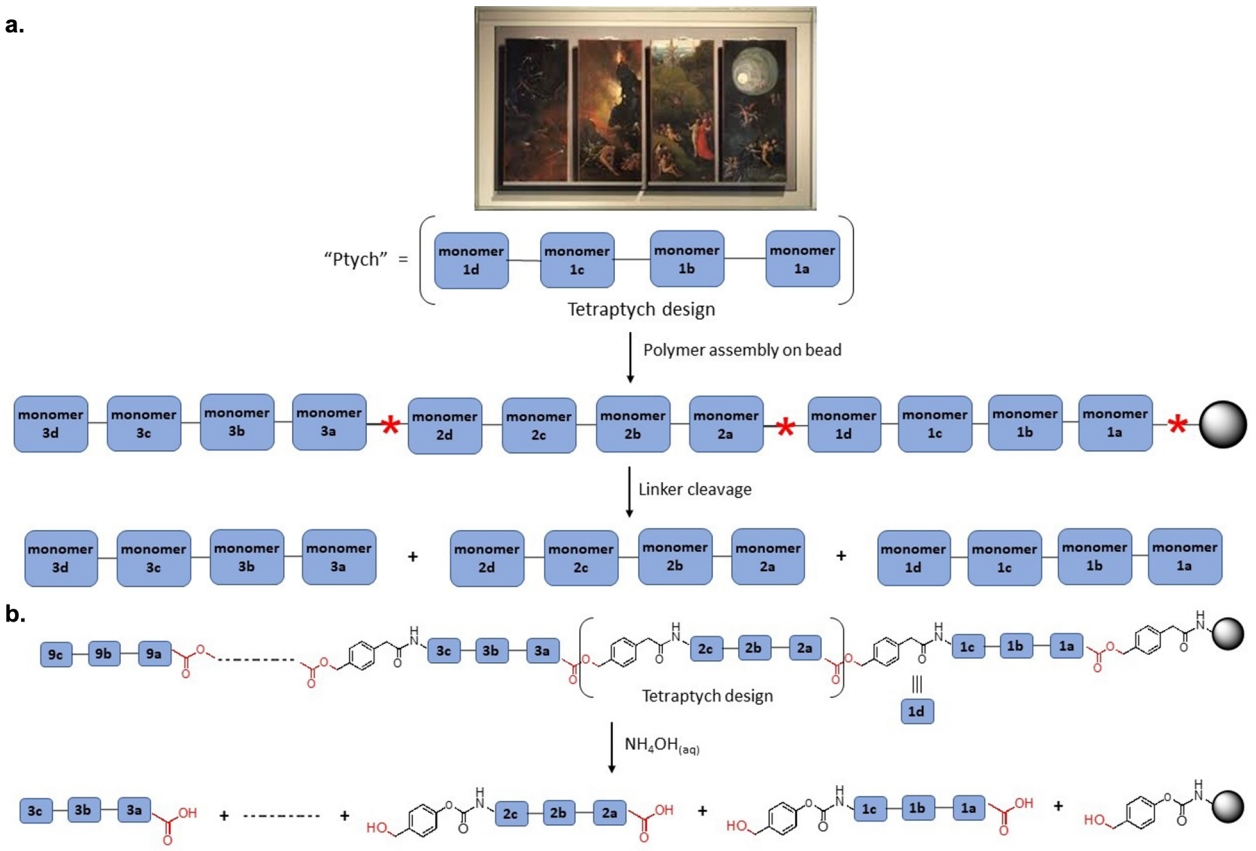


Figure 2. Description of the ptych design. (a) Example of a tetraptych painting (visioni dell'aldilà [Visions of the Hereafter] by Hieronymus Bosch) and the analogous ptych design of the polymers. Each polymer is constructed of a sequence of multiple tetraptychs each consisting of four diverse monomers and a cleavable linker (red ∗). During sequencing, each polymer is subjected to a chemical cleavage reaction in which the linkers are cleaved (at the red ∗ position) to generate a mixture of the ptych fragments. (b) Tetraptych-based polymer with PAM (phenyl-acetamido-methylene) as the cleavable linker (shown in black), and the cleavable ester linker shown in red. Once subjected to a solution of ammonium hydroxide, the ester linkers are cleaved to generate a mixture of all the tetraptychs constituting the polymer which can be each identified to reconstruct the full polymer sequence.

(titration and optimization details are described in Tables S6–S11). These strategies increase the probability of identifying the most active hit(s) while minimizing false positives.

Self-Readable Polymers and the "Ptych" Approach. We created a novel self-readable sequencing approach to polymer library design called the "ptych" (pronounced "tick") design. Figure 2a depicts a "tetraptych" (from Greek meaning "four-fold" from *tetra*, i.e., "four" and *ptysso*, i.e., "to fold"), which is a term used in the art world to describe a panel painting divided into four sections that can be folded to display a composite scene. In our application, a tetraptych is defined as a set of four monomers with folding properties that make up one diversity element in a longer sequence. The full sequence can then be formed by linking multiple ptychs together. Each tetraptych is selectively composed of monomers that enable diversity of physiochemical and structural properties at the individual ptych level. With this approach, polymer library scale and diversity can be built by choosing the number of ptychs in a linear sequence and the number of ptych variants at each position. The size of the ptych is also flexible. For example, two monomers define a diptych; four monomers tetraptychs; six monomers hexaptychs; and so forth.

By connecting ptychs via chemically cleavable linkers that can be cleaved under orthogonal conditions to those used in library synthesis and screening, the sequence of a ptych polymer can be directly read by mass spectrometry (MS). Figure 2a depicts a general polymer design with three tetraptychs, each consisting of four diversity building-block monomers and a cleavable linker. Cleavage of the linker monomers yields an equimolar mixture of the three tetraptychs. As a preliminary proof-of-concept, we used the phenyl-acetamido-methylene (PAM)[52] linker as a cleavable monomer building block. This monomer is normally used as a cleavable linker between a peptide and resin in Boc solid-phase peptide synthesis (SPPS).[52] In our application, it provides an ester bond between ptychs that is stable to the Fmoc/tBu/Alloc protection strategy[53,54] to build the polymers, but it can be readily cleaved using aqueous base such as ammonium hydroxide or sodium hydroxide. Figure 2b depicts a general polymer design with multiple tetraptychs, each consisting of a PAM linker and three diversity building-block monomers. Hydrolysis of the esters yields a mixture of the different tetraptychs.

The most important aspect of the ptych design is that building blocks can be selected so that each ptych diversity element in a sequence has a unique molecular weight. As a result, each of the ptychs present in the mixture after cleavage can be identified by its mass using high-resolution LC–MS and an electrospray source on an LTQ-Orbitrap XL mass spectrometer that can detect molecular ions at the femtomole level. This provides 3 orders of magnitude more sensitivity than the MS fragmentation methods used in peptide and protein sequencing.[55,56] The sequencing of ptych-designed libraries is independent of the nature of the building blocks, and virtually any chemical building block can be incorporated, whereas MS fragmentation

https://doi.org/10.1021/acscentsci.1c01041
ACS Cent. Sci. 2022, 8, 86–101

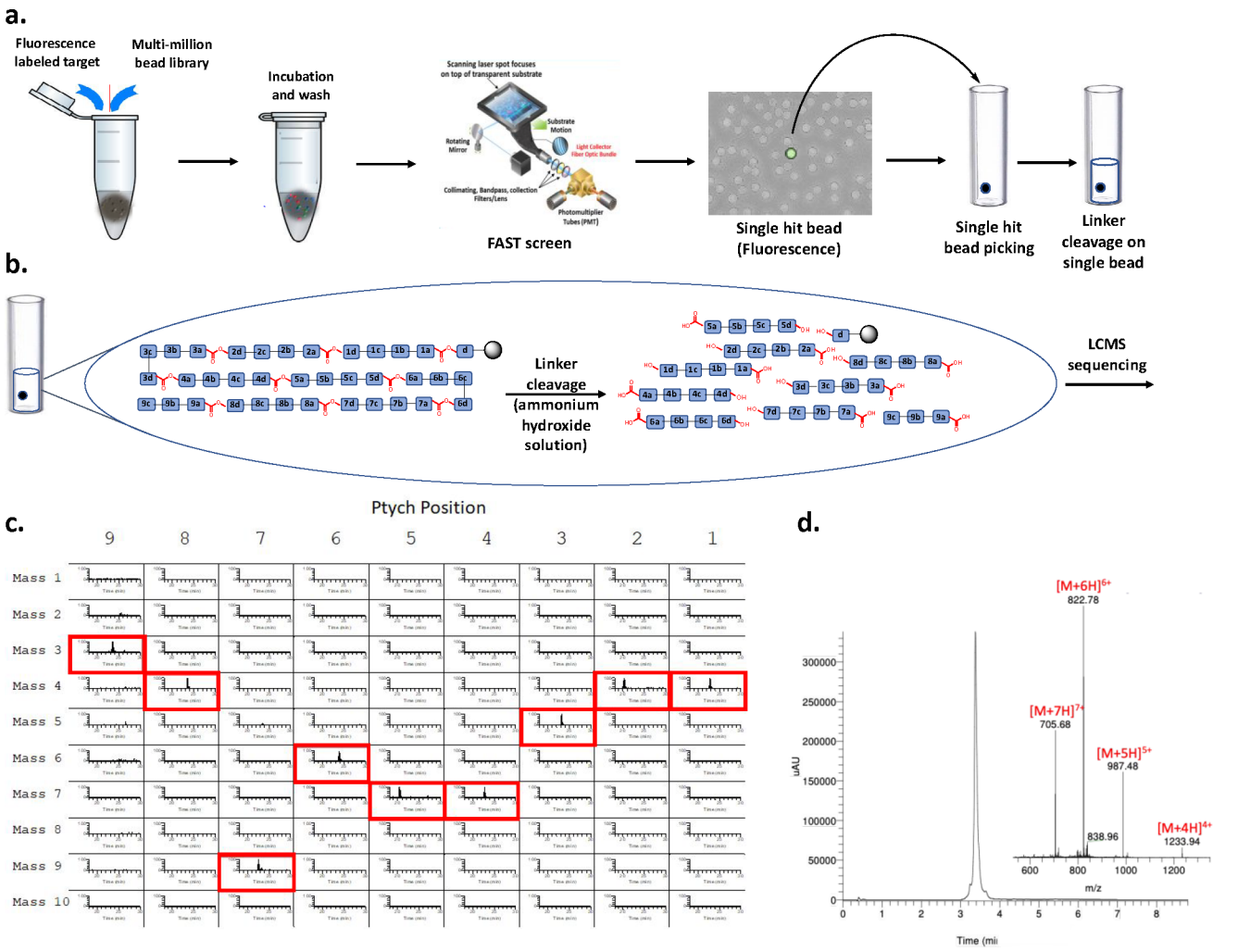


Figure 3. Summary of the screening process of a bead-based library. (a) The polymer-bead library is incubated with a fluorescently labeled target of interest, washed, and then screened by FAST to identify the location of positive hits. Hit beads are evaluated by bright-field and fluorescence microscopy, and confirmed bead hits are picked into separate LC−MS vials and subjected to the cleavage solution. (b) Upon treatment with ammonium hydroxide, all of the esters are hydrolyzed to yield a mixture of the different tetraptychs of a single sequence in each vial. (c) An analysis of ptych fragment masses allows reconstruction of the ptychs into a specific sequence based on the library design. (d) LC−MS chromatogram of a purified full-length 36-mer (9 tetraptych) polymer of which the sequence analysis is shown in panel c. The compound hit was characterized by LC−MS (ESI): calcd, 4932.5 Da; found, 4931.9 ± 0.9 Da. Fluorescein-(D)Asn-(D)Val-(L)Phe-PAM-(D)Tyr-(D)Ser-(L)Val-PAM-(D)Arg-(D)Ser-(L)Phe-PAM-(D)Phe-(D)Arg-(L)Ala-PAM-(D)Tyr-(D)Lys-(L)Ala-PAM-(D)Glu-(D)Arg-(L)Leu-PAM-(D)Ala-(D)Arg-(L)Leu-PAM-(D)Pro-(D)Arg-(L)Ala-PAM-(D)Phe-(D)Lys-Gly-PAM.

sequencing is extremely dependent on the nature of the fragmentation patterns of the backbone chemical bonds of the building blocks and has largely been limited to α-amino acid peptide polymers. The 10 μm diameter beads at ∼0.2 mmol/g loading typically carry ∼100 fmol of compound that is readily detectable by high-resolution LC−MS (Table S1). Using ptych design sequencing, polymer libraries can be synthesized on much smaller-diameter beads to create much larger libraries. Combined with the FAST platform, this enables screening and hit identification of much larger synthetic bead-based polymer libraries that has previously not been possible.

In a validation study to determine the reliability of sequences from individual beads, a set of 90 individual 10 μm beads, each containing one of four possible unique sequences, were mixed and then picked from a plate, cleaved, and sequenced. We were able to obtain the full correct sequence for 82 of the picked beads (91%) (Table S3a−d). There were no incorrect sequence assignments in any of the validation samples in which ptychs were detected. The samples that did not yield an identifiable sequence also did not yield any ptych assignments, indicating that the beads were likely not deposited correctly in the vial or were otherwise lost during automated sample processing, which is a factor in microscale handling efficiency and hit confirmation rates.

Figure 3 summarizes the typical screening process of a polymer library using the FAST screening and ptych library design. Assay development involves a preliminary titration screen using varying concentrations of targets against the library and naked control beads (Tables S6−S10) to minimize the effects of autofluorescence as described above while maximizing signal-to-noise. Based on these results, target screening concentrations and the background MFI threshold are selected for optimum hit fluorescent signals relative to background (Table S11). The library is incubated with a fluorescently labeled target in 50% Odyssey buffer and 0.5% CHAPS blocking buffer to screen out nonspecific binding and is followed by a

https://doi.org/10.1021/acscentsci.1c01041
ACS Cent. Sci. 2022, 8, 86−101

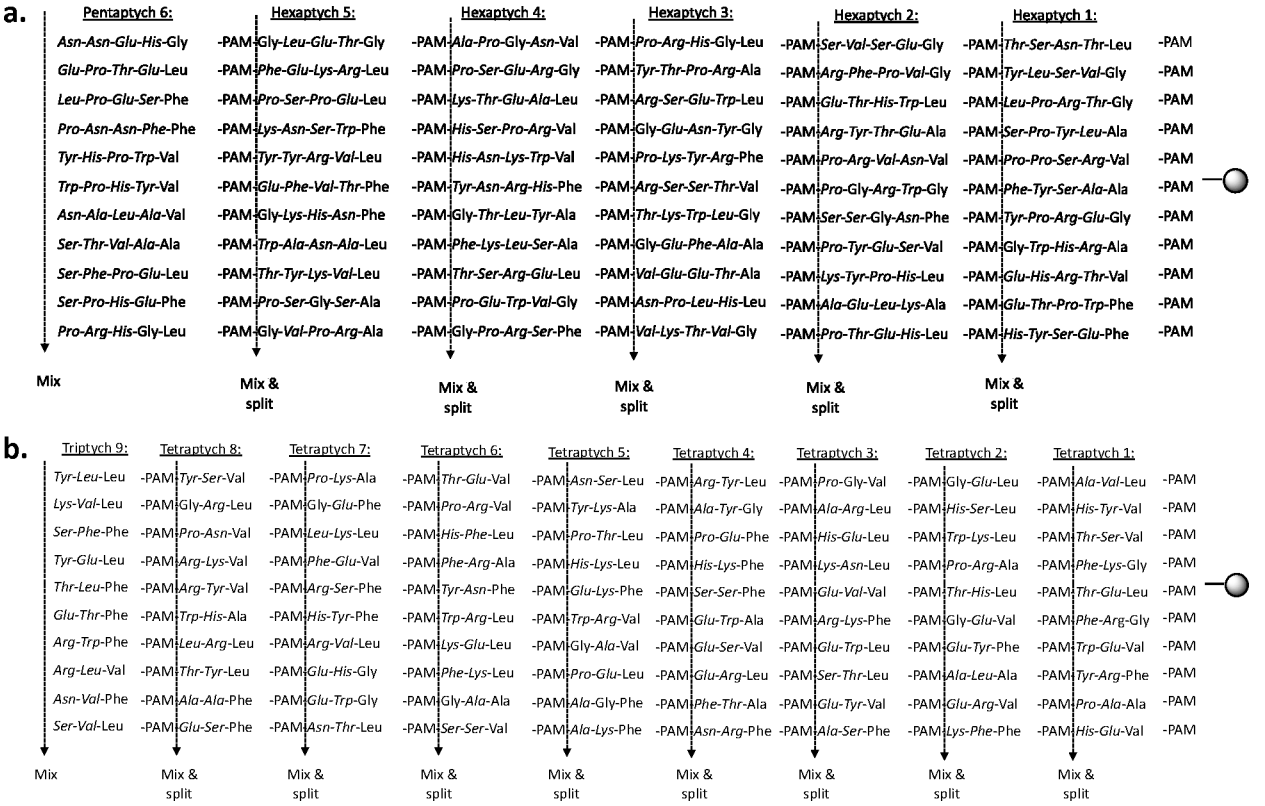


Figure 4. Library design and validation. (a) Hexaptych design for the ∼1.77 million compound library NNP1. (b) Tetraptych design for the 1 billion compound library NNP2. The sphere symbol represents resin beads. *Italics* indicates D-amino acid. In our design, there are D-amino acids and PAM linker, which are both non-natural building blocks. We chose to use L-amino acids at the position next to PAM as the corresponding Fmoc−L-amino acid−PAMs were commercially available, whereas the D-form was not.

sequence of washes (Table S11). The beads are then plated as a monolayer on glass slides and FAST screened to identify positive hits defined as fluorescently labeled beads that indicate binding to the target (Figure 3a). The plate and the hit location data from FAST are transferred to an automated fluorescence microscope and picking robot (ALS CellCelector, Video S1) for preliminary hit quality control. Confirmed hit beads are individually transferred into vials and treated with cleavage solution to hydrolyze the backbone esters yielding a mixture of, in this case, tetraptychs which are sequenced by LCMS (Figure 3b,c). The hit sequences are resynthesized and purified by preparative high-performance LC (HPLC) for hit confirmation and further testing (Figure 3d).

A preliminary study of hits showed that switching the backbone ester bonds to amides had only minor effects on measured binding affinities (Table S12), and as this greatly improves compound stability and simplifies hit resynthesis, all hits were prepared as the full backbone amide analogues. We measured resynthesized hit binding affinities for their respective targets using microscale thermophoresis (MST) (Figure 5c). Binding was confirmed in the majority of all backbone amide resynthesized hits with binding affinities in the nanomolar to subnanomolar range which indicated that switching out backbone esters for amide bonds generally has minimal effects on hit confirmation for these NNP designs.

Library Design and Screening Against Multiple Targets. We synthesized two large non-natural polymer libraries labeled NNP1 and NNP2. NNP1 (Figure 4a) consists of six hexaptychs as the diversity elements in which each ptych was composed of four D-amino acids (or glycine) and an L-amino acid (or glycine) ester linked to a PAM linker. This produced polymers of 36 monomers in length with an average molecular weight of ∼5 kDa. Each ptych was designed to have 1 of 11 possible hexaptychs per diversity position (listed under Hexaptych 1, Hexaptych 2, etc., in Figure 4a), making a 11^6 or an ∼1.77 million compound library. This corresponds to 66 hexaptychs, each of which was designed through a selection of monomers to give a range of physicochemical properties in each sequence position and a unique molecular weight for each hexaptych. Before synthesizing the library, we confirmed the synthetic feasibility of each individual hexaptych as a reference to determine the retention time by LC−MS and facilitate the sequencing of hits (Table S4). We made 75 copies of the library on 20 μm diameter monosized amino TentaGel microsphere resin beads with a loading of 0.27 mmol/g, which required only ∼550 mg of bead resin to produce.

Figure 4b shows the design for NNP2 that consists of nine tetraptychs constituting a total polymer length of 36 monomers. For each ptych in the sequence, there were 10 possible tetraptychs, constituting a total of 90 tetraptychs and creating a library of 10^9 or one billion compounds. We believe that this is the largest bead-based synthetic sequence-defined NNP library reported to date.[22] This library was constructed on 10 μm beads and required only 1.5 g of resin for the production of three copies (total of three billion beads). The individual ptychs in this library were also synthesized as controls and validated for sequencing (Table S5). After constructing the libraries, all side-chain protecting groups were removed, and the libraries were screened against multiple biological targets.

https://doi.org/10.1021/acscentsci.1c01041
ACS Cent. Sci. 2022, 8, 86−101

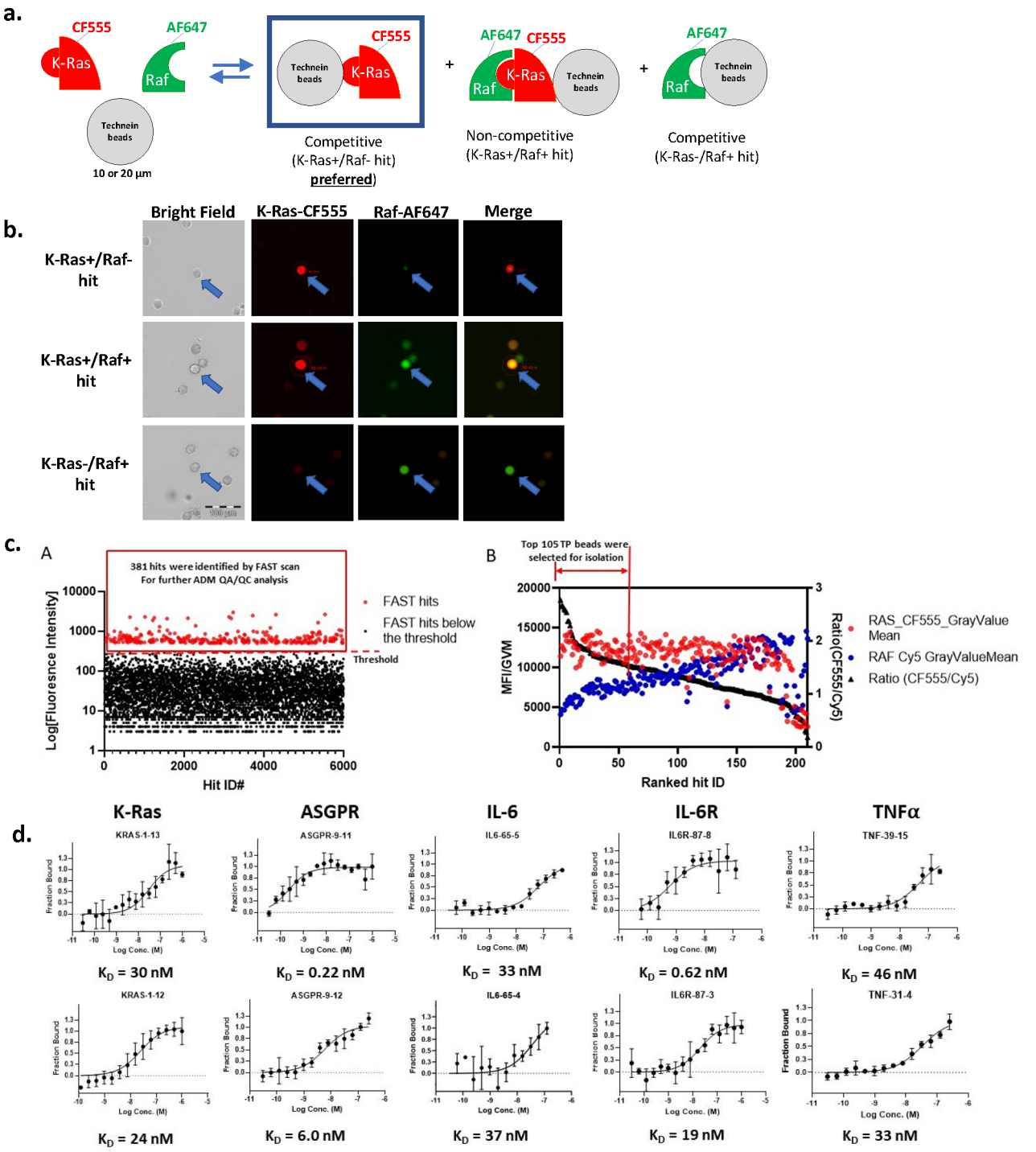


Figure 5. continued

https://doi.org/10.1021/acscentsci.1c01041
ACS Cent. Sci. 2022, 8, 86−101

e.

Target	Library	Estimated number of compound screened	Number of hits after FAST screen and CellCelector confirmation	Number of hits isolated	Number of hits sequenced	Number of hits synthesized	Number of hits reconfirmed for binding by MST (K_D Range)
K-Ras / Raf	NNP1	5,000,000 (2 plates, 20 um beads)	381	105	85	14 (one-two of each cluster)	14 (18 – 180 nM)
ASGPR	NNP1	5,000,000 (2 plates, 20 um beads)	289	193	190	19 (one of each cluster)	8 (0.22 – 330 nM)
IL-6	NNP2	10,000,000 (2 plates, 10 um beads)	19	19	14	12	8 (25 – 500 nM)
IL-6R / IL6	NNP2	10,000,000 (2 plates, 10 um beads)	120	82	46	13 (one of each cluster)	8 (0.62 – 330 nM)
TNFα	NNP2	10,000,000 (2 plates, 10 um beads)	44	44	43	27	23 (0.31 – 270 nM)

f.

Hexaptych #6	Hexaptych #5	Hexaptych #4	Hexaptych #3	Hexaptych #2	Hexaptych #1	Hit ID	# of hits in the cluster	KD [nM]
*W*_P_H_Y_V_	_M_G_*V*_P_*R*_A_	_M_P_S_E_R_G_	_M_P_K_Y_***R***_F_	_M_S_S_G_N_F_	_M_G_*W*_H_R_A__M	KRAS-1-1	7	24
*W*_P_H_Y_V_	_M_G_*V*_P_*R*_A_	_M_P_S_E_R_G_	_M_P_K_Y_***R***_F_	_M_*P*_G_R_*W*_G_	_M_Y_P_R_E_G__M	KRAS-1-2	7	49
*L*_P_E_S_F_	_M_G_*V*_P_*R*_A_	_M_P_S_E_R_G_	_M_P_K_Y_***R***_F_	_M_*R*_Y_T_*E*_A_	_M_G_*W*_H_R_A__M	KRAS-1-3	7	18
*L*_P_E_S_F_	_M_G_*V*_P_*R*_A_	_M_P_S_E_R_G_	_M_P_K_Y_***R***_F_	_M_*R*_Y_T_*E*_A_	_M_*Y*_P_R_*E*_G__M	KRAS-1-4	7	36
*W*_P_H_Y_V_	_M_P_S_G_S_A_	_M_P_S_E_R_G_	_M_P_K_Y_***R***_F_	_M_*P*_G_R_*W*_G_	_M_*Y*_P_R_*E*_G__M	KRAS-1-5	7	75
*S*_P_H_E_F_	_M_P_S_G_S_A_	_M_*T*_S_R_*E*_L_	_M_P_K_Y_***R***_F_	_M_*R*_Y_T_*E*_A_	_M_G_*W*_H_R_A__M	KRAS-1-6	7	21
*W*_P_H_Y_V_	_M_G_*V*_P_*R*_A_	_M_G_P_R_S_F_	_M_*Y*_T_P_***R***_A_	_M_*P*_G_R_*W*_G_	_M_S_P_*Y*_L_A__M	KRAS-1-8	6	44
*S*_F_P_E_L_	_M_G_*V*_P_*R*_A_	_M_P_S_E_R_G_	_M_P_K_Y_***R***_F_	_M_*R*_F_P_*V*_G_	_M_G_*W*_H_R_A__M	KRAS-1-9	3	32
*E*_P_T_E_L_	_M_G_*V*_P_*R*_A_	_M_P_S_E_R_G_	_M_P_K_Y_***R***_F_	_M_*R*_Y_T_*E*_A_	_M_G_*W*_H_R_A__M	KRAS-1-10	4	60
*Y*_H_P_*W*_V_	_M_G_*K*_H_*N*_F_	_M_P_S_E_R_G_	_M_P_K_Y_***R***_F_	_M_*P*_G_R_*W*_G_	_M_*Y*_P_R_*E*_G__M	KRAS-1-11	5	60
*L*_P_E_S_F_	_M_F_E_*K*_R_L_	_M_P_S_E_R_G_	_M_P_K_Y_***R***_F_	_M_P_G_R_*W*_G_	_M_G_*W*_H_R_A__M	KRAS-1-12	2	24
*L*_P_E_S_F_	_M_G_*V*_P_*R*_A_	_M_*K*_T_E_A_L_	_M_*Y*_T_P_***R***_A_	_M_*P*_G_R_*W*_G_	_M_G_*W*_H_R_A__M	KRAS-1-13	4	30
*S*_P_H_E_F_	_M_G_*V*_P_*R*_A_	_M_P_S_E_R_G_	_M_P_K_Y_***R***_F_	_M_*A*_E_L_*K*_A_	_M_G_*W*_H_R_A__M	KRAS-1-14	2	155
*W*_P_H_Y_V_	_M_F_E_*K*_R_L_	_M_P_S_E_R_G_	_M_P_K_Y_***R***_F_	_M_*P*_G_R_*W*_G_	_M_*H*_Y_S_E_F__M	KRAS-1-15	3	180

Figure 5. Screening data and representative MST data. (a) Competition screen of library beads against the primary target K-Ras labeled with CF555 in the presence of Raf-RBD as a counter target labeled with Alexa Fluor 647 (AF647). Only the single positive beads with fluorescence at 555 nm were picked and sequenced. This strategy was performed to enrich for inhibitors that would bind to K-Ras while blocking its interaction with its downstream signaling partner Raf. (b) Automated digital microscopic images demonstrating three types of hits. Top: A single positive bead binds the primary target K-Ras but not the counter target Raf. Middle: A double single positive bead binds both the primary target K-Ras and the counter target Raf. Bottom: A single positive bead binds the counter target Raf but not the primary target K-Ras. (c) Hit identification by a rapid FAST screen for K-Ras binding NNPs. Left: After a FAST scan, the software filters out false positive hits including autofluorescence particles using the dual-wavelength comparison technology. ∼300 top hits with bright KRas-CF555 (in red) that, above the threshold, were identified for further ADM/CellCelector imaging and analysis from a 2.5 million 20 μm bead sample plate. Right: The FAST hits identified were imaged, reviewed, and analyzed on the CellCelector instrument at high resolution. The MFI was measured for each true positive (TP) bead, and the top ranked 55 TP hits based on high MFI of K-Ras-CF555 and ratio (CF555/Cy5) were isolated for sequencing, resynthesis, and characterization. (d) Representative of MST data showing the binding of resynthesized purified NNPs to their targets and the calculated K_D values. (e) Summary of all screens showing hit attrition going through the library screen to confirmed hits. (f) Confirmed hit sequences for K-Ras binders showing sequence homology. Hits were clustered according to sequence, and then, 1 or 2 sequences per cluster were synthesized and confirmed by MST (Table S16). Sequences are represented as single-letter amino acid codes. Italicized letters indicate D-amino acids; M represents PAM.

The two NNP libraries were constructed from two groups of building blocks. The first group included five premade Fmoc−L-amino acid (or Gly)−PAM esters: Fmoc−L-Phe−PAM ester, Fmoc−L-Ala−PAM ester, Fmoc−L-Val−PAM ester, Fmoc−L-Leu−PAM ester, and Fmoc−Gly−PAM ester. All five amino acid−PAM esters were commercially available with the boc protecting group, which was simply converted to the Fmoc form and was used in the library synthesis (see the Supporting Information for synthesis details). The second group of building blocks used in the library included 15 Fmoc-protected D-amino acids (or Gly): *Ala*, *Glu*, *Phe*, *Gly*, *His*, *Lys*, *Leu*, *Asn*, *Pro*, *Arg*, *Ser*, *Thr*, *Val*, *Trp*, and *Tyr* (see Scheme S3 for the structure of all of the building blocks). The NNP1 library was designed to have an amino acid distribution close to their average occurrence genome-wide (Scheme S1) (based on data from the UCSC Proteome Browser[57]). The library was designed in such a way that the amino acids will be distributed among the six hexaptychs as shown in Table S14. The design of the NNP2 library was more unique with less resemblance to the amino acid distribution in the genome (Scheme S2), and also, here, the amino acids were distributed among the nine tetraptychs in the library (Table S15).

To demonstrate the speed and efficiency of the screening and sequencing process, we screened five target proteins: K-Ras, ASGPR, IL6 and its receptor IL6R, and TNFα. K-Ras is an oncology drug target that is mutated in ∼30% of cancers and is associated with uncontrolled cell proliferation—particularly in pancreatic and lung cancers with poor prognoses.[58] ASGPR is the functional subunit of the asialoglycoprotein receptor (ASGPR) which is a C-type lectin glycan receptor predominantly found on the surface of liver hepatocytes and has been utilized as a mediator for liver-specific intracellular drug delivery of nucleic-acid-based therapeutics.[59] TNFα, IL6, and soluble IL6 receptor (IL6R) are, respectively, cytokines and a cytokine receptor involved in inflammatory signaling processes and are well-established immunotherapy targets. All are challenging

https://doi.org/10.1021/acscentsci.1c01041
ACS Cent. Sci. 2022, 8, 86−101

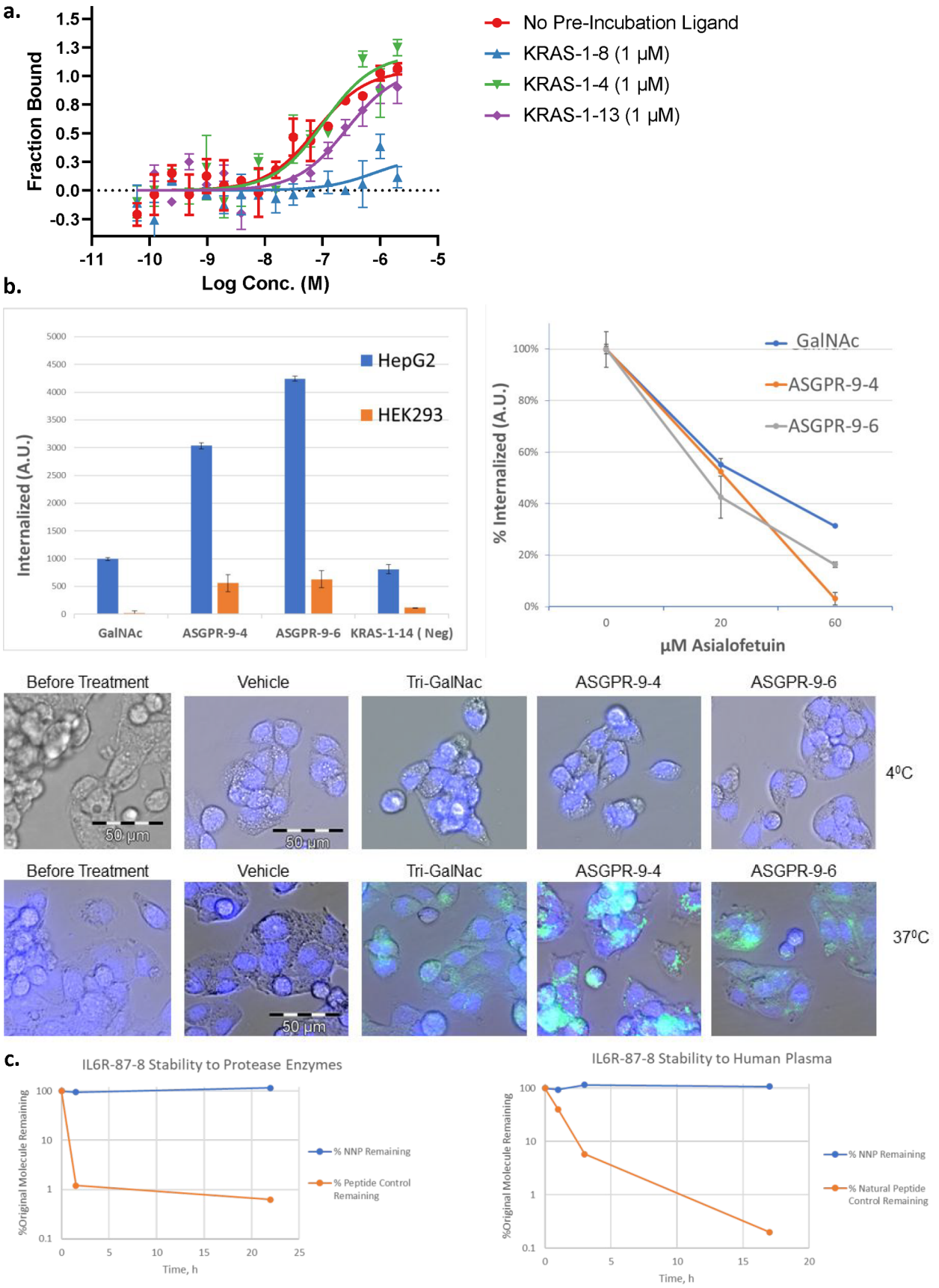


Figure 6. Biological relevance and stability of NNPs in biological matrices. (a) K-Ras−Raf interaction was measured by MST using a fixed concentration of K-Ras (5 nM) and a titration of Raf to give a 78 nM binding affinity. Competitive inhibition of the K-Ras−Raf protein−protein interaction was tested with a 15 min preincubation with NNP ligands (1 μM concentration) followed by titration with Raf to measure binding by MST. NNP KRAS-1−8 showed complete inhibition; NNP KRAS-1−13 caused a shift in the K_D (K_D = 260 nM), and NNP KRAS-1−4 showed no inhibition

https://doi.org/10.1021/acscentsci.1c01041
ACS Cent. Sci. 2022, 8, 86−101

Figure 6. continued

(K_D = 100 nM). (b) Left: cell uptake of tri-GalNAc (positive control), two NNP hits (ASGPR-9−4 and ASGPR-9−6), and a nonhit NNP (KRAS-1−14, negative control) in HepG2 (high ASGPR expressing) vs HEK293 (low ASGPR expressing) cells lines. Right: competition uptake assay of tri-GalNAc and two NNP hits (ASGPR-9−4 and ASGPR-9−6) in HepG2 cells after preincubation with different concentrations of asialofetuin (a naturally occurring serum protein ASGPR ligand). A decrease in cell uptake with increasing concentrations of asialofetuin indicates blocking of ASGPR-mediated uptake. Bottom: representative images of ASGPR NNP uptake by HepG2 cells (scale: 400× magnification). Nuclear dye Hoechst (blue) and either the tri-GalNAc or NNP molecules (green) show internalization of the ligands. HepG2 cells are shown before treatment and after 2 h of incubation with vehicle, tri-GalNAc, and the indicated fluorescein-labeled NNPs at both 4 and 37 °C to induce internalization. (c) Stability data for NNP IL6R-87-8 compared to its L variant in the presence of proteinase K (left) and for the same NNP in human plasma (right) compared with Angiotensin I.

targets for traditional small-molecule approaches and therefore represent interesting test cases for chemical NNP ligands.

In the case of the K-Ras, IL6, and IL6R, we wanted to screen for protein−protein interaction inhibitors. For K-Ras, we wanted to specifically block binding of the Ras binding domain to its downstream signaling partner Raf (Figure 5a,b), and for IL6, we wanted to block binding to its receptor IL6R and conversely in a separate screen find binders of IL6R that block IL6 binding. We labeled primary screening targets (Raf, IL6, and IL6R), with dyes maximally excited at ∼555 nm (AF555 or CF555) to identify binders in the FAST screen, and counter targets (Raf, IL6R, and IL6), with dyes maximally excited at a ∼647 nm wavelength (AF647 or CF647) which could be detected by ADM on the CellCelector instrument. After FAST screening hit detection, we measured the MFI for each dye on each bead to prioritize the hits using both the overall brightness of the bead as a qualitative measure of binding affinity and the MFI ratio of the target to the counter target (Figure 5b,c).

K-Ras and ASGPR were screened against library NNP1. The library size of NNP1 is 1.77M members on 20 mm beads, and it was screened at a 2.8-fold redundancy with 5M beads on two plates (2.5M beads per plate). After FAST screening, we identified a preliminary hit list of 381 K-Ras selective binding beads. Hit sequences in the K-Ras screen were pooled into 14 clusters, and the most prevalent sequences in each case were selected from each cluster. Similarly, 289 hit sequences from the ASGPR screen were grouped into 19 clusters, and individual hits from each cluster were selected for hit confirmation by resynthesis and measurement of K_D by MST (Figure 5e, Tables S16 and 17a,b). Equilibrium K_D binding affinities for K-Ras hits ranged from 18 to 180 nM, and from 0.22 to 330 nM for ASGPR (Figure 5d,e, Table S13).

With a library size of 1B members on 10 μm beads, library NNP2 would require 200 plates to screen the entire library at 5M beads per plate. With a custom industrial robotic high-throughput screening (HTS) suite, this would be fairly straightforward—the entire library could be FAST screened in <10 h. In this proof-of-concept study, we manually screened a 10 million compound portion of the library, corresponding to 2 screening plates against IL-6, IL6R, and TNFα. As with the K-Ras−Raf screen, we performed the IL-6R screen using counter labeled IL-6 to identify IL6R binding domain selective inhibitors. We selected and isolated the hits with the highest target to antitarget MFI ratios. For IL6 and TNFα, we were primarily interested in finding selective affinity agents and did not conduct competition screens. Binding affinities ranged from 25 to 500 nM for IL6, 0.6 to 330 nM for IL6R, and 0.3 to 270 nM for TNFα. The most potent hit in the NNP2 screen was a K_D of 310 pM against TNFα (Figure 5e, Tables S18−S20).

Figure 5e shows the hit rate broken down by screening and sequencing steps across the five targets. The average hit rate for beads identified by the FAST screen and passing QC/QA in ADM is 0.003% and ranged from 19 hit beads for IL6 to 381 hits for K-Ras. The bead hit rate is to a large extent determined by the threshold cut identified in assay development to eliminate the effects of autofluorescence producing false negatives. This is also a reasonable hit rate in terms of the downstream processing effort for hit confirmation. Hit bead selection for automated picking from the screening plates depends primarily on how isolated the beads are from neighboring beads. In a number of cases, hit beads are located in dense aggregates that make picking difficult to impossible without carrying over several other beads that will confound sequencing. In the five assays shown here, on average 72 ± 44% of the hit beads could be picked. Of these, on average 81 ± 25% were successfully sequenced by LCMS. The factors affecting sequencing were successful transfer of the beads to the cleavage vial where failure results in no observable ptychs' masses, or incomplete sequencing where individual ptychs failed to be identified in the LCMS. Sequencing however did not prove to be a major challenge, and the high sequencing rate enabled hit confirmation by resynthesis without major optimization. Where we identified a large number of redundant or similar sequence hits, such as for K-Ras, ASGPR, and IL6R, sequences were clustered, and the highest represented sequence in each cluster was selected for resynthesis. Over the five targets screened, hit confirmation of resynthesized hits was 71 ± 29% denoting a high true positive rate.

Significant sequence homology was observed between reconfirmed hits against each screening target. Tables S16−S20 list the hits and hit redundancy or hit clustering, as applicable, for each screen. Overlap in hit sequences based on ptych comparisons is observed in all screens and especially for ASGPR, K-Ras, and IL6R. For ASGPR, which was screened with the smaller library NNP1, we selected hits for resynthesis and confirmation on the basis of seeing hit redundancy of at least 2 or more identical sequences and selected 19 hits for follow-up on this basis (Table S17a,b). Although K-Ras was screened with the same library, we only saw a redundancy of 2 for one sequence, but we were able to cluster hits based on very similar sequences with typically 1−4 ptychs differing between sequences (Table S16). Interestingly, an *Arg* residue is conserved in position 2 of hexaptych 3 in all confirmed hits suggesting that this is a critical residue in binding or structural stability (Figure 5e). We see a similar result for IL6R even though this was screened with only a subset (10M compounds) of the much larger 1B compound library, and we were able to group similar hit sequences to identify 13 clusters (Table S18a,b).

The IL6R screen was a competition screen with IL6 to look for IL6−IL6R binding interaction inhibitors, and as such, it was tailored to look for a very specific binding site which may account for the sequence homology observed in the hits. By contrast, the IL6 and TNFα screens were conducted with the 1B

https://doi.org/10.1021/acscentsci.1c01041
ACS Cent. Sci. 2022, 8, 86−101

member library without competition, and as a result, a rather more diverse set of hits was observed in each case (Tables S19 and S20). However, even here, some sequence homology is evident. For example, in the case of TNFα residues *Arg*-Leu are largely conserved across tetraptych diversity in positions 2 and 1 for tetraptych position 8, and PAM-*Phe*-*Glu*-Val and PAM-*Pro*-Gly-Val are highly represented in tetraptychs 7 and 3, respectively (Table S19b). The sequence homology between hits seen in each screen suggests that the screening process is identifying specific binders with defined molecular contacts in each case.

Discovery of PPI Inhibitors and Receptor-Mediated Intercellular Delivery Agents. To demonstrate the biological significance of these NNP hits and their target selectivity, we focused on K-Ras and ASGPR functional biological activities. For K-Ras, we investigated the specific inhibition of K-Ras and Raf binding for a range of confirmed K-Ras hits. We tested this in an MST competition binding assay (Figure 6a). As a control, we measured the K-Ras−Raf interaction alone for which an average K_D value of 78 nM was obtained from two technical runs. Then, we measured the same interaction with three different K-Ras lead hits (KRAS-1−4, K_D = 36 nM; KRAS-1−8, K_D = 44 nM; KRAS-1−13, K_D = 30 nM; at 1 μM each), each preincubated with K-Ras at room temperature for 15 min. The three NNPs showed a range of inhibition activity from complete inhibition (KRAS-1−8) to partial inhibition (KRAS-1−13) and no inhibition (KRAS-1−4). The MFI ratios of the target (K-Ras-CF555) to the counter target (Raf-AF647) for the hit beads corresponding to these hits were higher for KRAS-1−8 than the other two hit beads (MFI ratios: KRAS-1−8 = 2.55; KRAS-1−4 = 1.89; and KRAS-1−13 = 1.53), suggesting that this could be a useful metric for functional inhibitors of PPIs from the primary competition screen.

ASGPR is a glycoprotein receptor, and all of the published ligands are glycans that to date mimic the native substrates.[60−62] To determine if NNP hits could specifically internalize into liver cells with a high expression of ASGPR but not cells lacking ASGPR expression, we compared the ASGPR-mediated uptake of two lead NNP hits, ASGPR-9−4 (K_D = 230 nM) and ASGPR-9−6 (K_D = 34 nM), in the HepG2 human hepatocarcinoma (high expressing) and HEK293 (nonexpressing) cell lines (Figure 6b). As a positive control, we utilized the ASGPR trivalent ligand *N*-acetylgalactosamine (tri-GalNAc)[60] and a nonhit NNP from the same NNP1 library (KRAS-1−14) as a negative control. All compounds were labeled with fluorescein and analyzed by flow cytometry (see the Experimental Section for detailed procedures and the Supporting Information for synthesis details). The internalization of the two NNP hits was significantly higher in HepG2 cells compared to that in HEK293 cells and significantly higher than uptake of the positive control tri-GalNAc. The nonhit NNP negative control showed minimal uptake in either cell line. To further demonstrate ASGPR-mediated cellular uptake, competitive cell uptake assays were performed utilizing the two NNP hits and the positive control tri-GalNAc in HepG2 cells in the presence of asialofetuin, a naturally occurring serum protein ligand for ASGPR[63] (Figure 6b). Cells were preincubated with two concentrations of asialofetuin (20 and 60 μM), representing 67- and 200-fold excess compared with the test compound (0.3 μM). The NNP hits and positive control's cellular uptake decreased with increasing concentrations of asialofetuin and is mostly abolished with 60 μM asialofetuin. These results indicate that these ligands compete for the same receptor and that uptake is ASGPR-mediated. Importantly, these results for K-Ras and ASGPR demonstrate that the NNP hits found in screening are not nonspecific binders but are capable of selectively binding to their target proteins in the presence of other selective protein-binding partners, plasma media, and cell membranes and are capable of eliciting functional biological responses.

Biological Stability of NNP Hits. To confirm the superiority in stability of these largely D-amino acid NNPs over peptides, we investigated the stability of an IL-6R hit from NNP2 to proteinase K and in human plasma. For the proteinase K stability assay, we compared the original hit with its fully L-amino acid variant. As expected, the L-amino acid variant was completely degraded within less than 2 h in the presence of proteinase K (Figure 6c), whereas minimal degradation was observed for the NNP hit after an overnight incubation. Similar stability was observed in human plasma where the stability of the hit was compared to the natural peptide Angiotensin I. Angiotensin I was completely degraded within 4 h in human plasma (Figure 6c), whereas NNP2 hits stayed largely intact even after overnight incubation.

CONCLUSIONS

Using the FAST screening platform and ptych design, we have demonstrated a megathroughput screening and sequencing strategy for the discovery of potent and functional NNPs. The novel ability to screen at the femtomole scale on 10 μm beads enables time- and cost-effective screening with much larger chemical diversity than has previously been reported. In this proof-of-concept study, we used commercially available amino acid building blocks and well-established solid-phase chemistry to construct these first NNP libraries to validate the screening and sequencing methodology. Using the same approach, it is relatively straightforward to move into increasingly novel synthetic building blocks and coupling chemistry[64] as well as different cleavable linkers enabling an unlimited access to polymer diversity through library synthesis and empirical screening. We have shown here that we can find low nanomolar to picomolar hits from primary screening and have used this to validate biological selectivity and activity in a range of molecular targets. We have shown biological functionality of the hits of two targets as representative use cases, which are the ability to disrupt PPI by inhibition of the K-Ras−Raf interaction and protein−glycan interaction (PGI) in ASGPR-mediated cellular uptake and internalization. While we do not anticipate α-amino acid NNPs will be passively permeable to cell membranes, our interest in screening for inhibitors of K-Ras which is an intracellular target was driven by recent breakthroughs in cell selective receptor-mediated intracellular delivery of biologic molecules[65] which could feasibly be used for delivery of NNP-like payloads. Utilizing a similar approach, we have also identified NNPs that could potentially be intracellular delivery agents by targeting ASGPR to identify receptor selective NNPs that not only bind ASGPR but also are actively transported across cell membranes in a selective receptor-mediated manner. The NNP hits identified here, without optimization, are more efficiently intracellularly transported than the previously reported molecular transport ligand tri-GalNAc which is being used commercially for the delivery of nucleic acid drugs.[59] This is particularly noteworthy as tri-GalNAc and other reported ligands for ASGPR are glycans, and we have demonstrated here that ASGPR can also bind and transport non-natural peptide-like ligands. Lastly, as observed by others,[22] the primarily D-

https://doi.org/10.1021/acscentsci.1c01041
ACS Cent. Sci. 2022, 8, 86−101

amino acid NNPs show unique stability against biological degradation.

Transition melt temperatures across all hits ranged from 39 to 65 °C (data not shown), which is within the range of folded proteins of similar lengths (for example, see ref 66) and indicates that a tertiary structure is probably important for the molecular interactions of these hits. As the diversity of synthetic polymers expands, 3D structures of hits by crystallography or nuclear magnetic resonance (NMR) will most likely identify templates for novel secondary and tertiary structural motifs that can be rapidly refined by building focused libraries for secondary screening. As structural motifs become better understood, individual ptychs can be engineered to promote intra- and intermolecular recognition to stabilize structure and maximize affinity. We used a regular repeating ptych design for libraries described here, but more elaborate designs that use different numbers and types of monomers in ptych positions are possible. These will provide further chemical diversity for primary screening and strategies that allow the optimization of hits.

In summary, we have demonstrated a method for stepping outside of the bounds of natural polymers and moving into a new field of designer polymers with completely new structures and functions through empirical screening. The application area of this platform is vastly broad and includes therapeutics for drug discovery, affinity reagents for sensors and diagnostics, and reagents for catalysis.

EXPERIMENTAL SECTION

Synthesis of Libraries NNP1 and NNP2. All libraries were synthesized using "one-bead−one-compound" and "mix-and-split" methods of solid-phase synthesis on TentaGel amine 10 or 20 μm resin. Library NNP1 was synthesized on 554 mg of 20 μm TentaGel M NH_2 (0.27 mmol/g amine loading) with a theoretical diversity of 1.77×10^6 and 75 copies (i.e., 1.33×10^8 beads). Library NNP2 was synthesized on 1.5 g of 10 μm TentaGel M NH_2 (0.25 mmol/g amine loading) with a theoretical diversity of 1×10^9 and 3 copies (i.e., 3×10^9 beads).

For the synthesis of library NNP1, the beads were swollen in DCM for 1 h. Then, the DCM was drained, and the beads were suspended in DMF and divided evenly by pipet between 11 plastic fritted syringes placed on a manifold. Then, 11 different hexaptychs were constructed on the beads, a different hexaptych in each fritted syringe, by coupling first an L-amino acid−PAM ester followed by the coupling of four more D-amino acids, according to the library design in Figure 4a. The beads were then mixed and split evenly again between the 11 plastic fritted syringes, and the synthesis was carried out in the same manner with the next hexaptychs, until all six hexaptychs were constructed.

For the synthesis of library NNP2, after swelling the beads in DCM for 1 h, the DCM was drained, and the beads were suspended in DMF. Then, the beads were divided evenly by pipet between 10 plastic fritted syringes placed on a manifold. Then, 10 different tetraptychs were constructed on the beads, a different tetraptych in each fritted syringe, by coupling first an L-amino acid−PAM ester followed by the coupling of two more D-amino acids, according to the library design in Figure 4b. The beads were then mixed and split evenly again between the 10 plastic fritted syringes, and the synthesis was carried out in the same manner with the next tetraptychs, until all nine tetraptychs were constructed.

Coupling Conditions for Fmoc−L-Amino Acid−PAM Esters in the Library Synthesis. 3.5 equiv of Fmoc−L-amino acid−PAM ester was dissolved in a solution of 0.5 M HATU in NMP (3.18 equiv of HATU). Then, DIEA (10 equiv) was added to this mixture to activate the amino acid for 30 s, and the solution was added to the resin and reacted for 30 min. After completion of the coupling reaction (confirmed by a ninhydrin test), the resin was drained and washed with DMF (3 × 5 mL).

Coupling Conditions for Fmoc−D-Amino Acids. 5.5 equiv of Fmoc−D-amino acid was dissolved in a solution of 0.5 M HATU in NMP (5 equiv of HATU). Then, DIEA (10 equiv) was added to this mixture to activate the amino acid for 30 s, and the solution was added to the resin and reacted for 30 min. After completion of the coupling reaction (confirmed by a ninhydrin test), the resin was drained and washed with DMF (3 × 5 mL).

Fmoc Deprotection. Fmoc deprotection was performed by the addition of 25% 4-methylpiperidine in DMF (5 mL) to the resin (1 × 5 min + 1 × 10 min), followed by draining and washing the resin with DMF (5 × 5 mL).

Side-Chain Deprotection. At the end of the library's construction, after the last Fmoc deprotection, all of the library beads were mixed into one fritted syringe, and the side-chain protecting groups were removed with a solution of 95% (v/v) TFA, 2.5% (v/v) water, and 2.5% (v/v) triisopropylsilane (1 mL of cleavage solution per 10 mg of resin) for 2 h. Then, the TFA cocktail was drained, and the resin was thoroughly washed with DCM, DMF, DCM, and MeOH (3 × 10 mL of each solvent) and was ready for the screening process.

Activation of K-Ras by GTP Loading for the Screen and Binding Assays. To activate K-Ras for binding NNP or Raf, the K-Ras protein had to be loaded with GTP. Loading was performed according to the following protocol: The 200 μM stock solution of the target protein was diluted to 10 μM in 20 mM HEPES pH 8.0, 150 mM NaCl, 10 mM $MgCl_2$, 1 mM TCEP, and 0.05% Tween-20 (total volume 110 μL). A portion of 10 μL was set aside for later labeling quality control. EDTA pH 8.0 (stock concentration 10 mM) was added to the protein solution to a final concentration of 80 μM. GTP (stock concentration 50 mM) was added to the protein solution to a final concentration of 750 μM. The solution was incubated at 30 °C for 2 h (PCR tube) and then placed on ice for 2 min. $MgCl_2$ was added to the protein solution to a final concentration of 100 mM. The resulting protein solution was buffer exchanged into the buffer required in the labeling kit for labeling. This procedure was used before the screen, the MST analysis, and the K-Ras/Raf inhibition assay. All of the other target molecules (TNFα, IL-6, IL-6R, and ASGPR) were used as received without any additional treatment.

Screening of OBOC Libraries on FAST. FAST screening assays were specifically optimized for each target in terms of probe concentration, blocking and washing stringency, etc. The probe binding to the NNP1 and NNP2 library beads was performed in tubes. Typically, the library or control beads were hydrated in the buffer (1% PEG, 50 mM Tris, pH 7.5, 25% Odyssey blocking buffer PBS) for 30 min at room temperature (RT) with vortex followed by 1 min of sonication to break apart the large bead clumps. Beads were then centrifuged down, and the bead pellets were washed 2× with Odyssey/PBS buffer; the bead suspension was further filtered through a 30 μm cell strainer to remove bead aggregates. The concentration of the hydrated beads was determined based on bead counting using a hemocytometer. Aliquots of the bead suspension with the required number of beads then were centrifuged down, resuspended in blocking buffer (100% Odyssey, 0.5% Chaps, 200 mM NaCl in PBS), and incubated overnight at RT with

https://doi.org/10.1021/acscentsci.1c01041
ACS Cent. Sci. 2022, 8, 86−101

gentle rotating. After blocking, the beads were pelleted, resuspended in 100% Odyssey buffer, and then mixed at a 1:1 volume ratio with the CF555 or AF555 conjugated probe that was diluted in the prebinding buffer (1% Chaps, 400 mM NaCl, 2 mM TCEP, in PBS) to 2× the final working concentration. The probe/library bead mixtures were incubated for 1 h at RT with gentle rotation to allow probe to bind to the library beads. After incubation, the beads were pelleted, and the unbound probes were aspirated, followed by 3 washes with 10 mL of wash buffer (0.5% Chaps, 200 mM NaCl, 1 mM TCEP in PBS), 5 min/time, and an additional 2 washes with 10 mL of 0.5% Chaps/PBS. After the last centrifugation, the buffer was aspirated, with the exception of the final ∼500 μL. This volume was sonicated for 30 s to dissociate newly formed bead clumps. Then, 1.5 mL of prepared 0.3% low-melting agarose (LMT) that was kept in a 37 °C water bath before use was added to the resonicated beads to make the bead/soft agar suspension.

Beads in the LMT suspension were then transferred and evenly plated onto the FAST slide (the screening plates), and then, the slides were placed on a cold tray to accelerate the curing and immobilization of the beads. Following the gel formation, a layer of mounting medium (e.g., 500 μL of Live-Cell medium) was gently placed on top of the gel to keep the beads from rapid drying or photoquenching of the fluorescence. The sample slides (plates) were scanned and analyzed using the FAST system. The FAST analysis generates a bead hit list, where each bead is quantified by an MFI measurement.

Bead Analysis and Picking Using an ALS CellCelector instrument. The beads with MFI values above a threshold determined by the "no probe" control condition were identified, and then, a coordinate list of the hits were transferred to the CellCelector instrument for automated digital microscopy (ADM). This imaging analyzes the hits with multiple channels at higher resolution. Images of the hit beads were then QC/QA reviewed based on the morphology and fluorescence staining, and the fluorescence of selected channels was quantified to rank the top hits for isolation. Then, each selected single hit bead was isolated with the CellCelector instrument individually into the HPLC vials in ddH_2O for MS-based sequencing.

Processing and Sequence Analysis of Picked Beads. Beads were deposited directly into glass autosampler vials containing deionized water. The vials were inserted into deep-well 96-well plates and dried in a vacuum centrifugal concentrator (GeneVac II Plus) at 40 °C. To hydrolyze the interptych ester linkages, 50 μL of 7% aqueous ammonium hydroxide or 150 mM NaOH was added, and the samples were incubated at 37 °C for 6 h and then evaporated under vacuum in the centrifugal concentrator. The samples were then prepared for analysis by adding 50 μL of 5% acetonitrile in water with 0.1% formic acid and analyzed by capillary reversed-phase gradient LC−MS/MS using an Agilent capillary HPLC pump and CTC Analytics autosampler coupled to an LTQ-Orbitrap mass spectrometry system. Expected masses of hydrolysis products were loaded into an inclusion list for targeted MS/MS when detected above threshold in a high-resolution Orbitrap scan. Data analysis used both MS and MS/MS data to assign high-confidence hits for assembling sequences for the hits.

Hit Resynthesis. *Solid-Phase NNP Synthesis.* Hits were synthesized on ChemMatrix Rink amide resin (loading 0.5 mmol/g, typical scale: 30 mg, 0.015 mmol) by an automated peptide synthesizer (Biotage Syro I) using standard Fmoc-based amide coupling conditions with DIC/Oxyma as the coupling reagents. Fmoc-protected L-amino acid−PAM esters used in the library synthesis were replaced here by two separate residues: Fmoc−L-amino acid and Fmoc−PAM. This was in order to avoid having ester linkage (but rather a standard amide linkage) in the synthesized hits, for stability purposes. The synthesis was performed using the following protocol: ChemMatrix Rink amide resin was swollen in DCM for 1 h, drained, washed with DMF, and placed on the peptide synthesizer for constructing the full sequence. Fmoc deprotection was performed by the addition of 25% 4-methylpiperidine in DMF (1.2 mL) to the resin (1 × 5 min + 1 × 10 min), followed by draining and washing the resin with DMF (5 × 1.2 mL). Couplings were performed by adding 250 μL of NMP to the resin followed by 90 μL of 0.5 M Fmoc-protected amino acids (or Fmoc−PAM) in DMF (3 equiv, 0.045 mmol), 90 μL of 0.5 M Oxyma in DMF (3 equiv, 0.045 mmol), and 90 μL of 0.5 M DIC in DMF (3 equiv, 0.045 mmol). The resin mixture was allowed to react for 15 min at 60 °C and was then drained, washed with DMF (3 × 1.2 mL), and treated again with the same coupling conditions for double coupling. At the end of the double coupling, the Fmoc was deprotected, and these synthesis cycles were repeated on the peptide synthesizer until all of the residues were constructed onto the resin. After the last Fmoc-deprotection, the resin-NNP was taken out of the peptide synthesizer for manual fluorescein incorporation.

Incorporation of Fluorescein. fluorescein was incorporated on the N-terminus of all of the resynthesized hits. 21.3 mg of NHS-fluorescein (3 equiv, 0.045 mmol) was dissolved in 300 μL of DMF and was added to the resin-NNP. The resin mixture was allowed to react for 3 h and was monitored by a ninhydrin test. Upon completion, the resin was drained, washed thoroughly with DMF (3 × 5 mL) and DCM (3 × 5 mL), and dried before cleavage.

NNP Cleavage. Cleavage from the solid support and side-chain deprotection were performed by the treatment of resin-NNP with a solution of 95% (v/v) TFA, 2.5% (v/v) water, and 2.5% (v/v) triisopropylsilane (3 mL of cleavage solution per 30 mg of resin) for 2 h. TFA was then evaporated on the SpeedVac concentrator (Thermo Scientific Savant SpeedVac concentrator) until the solution volume reached ∼1 mL. The crude NNP was then precipitated, triturated with chilled diethyl ether (×3), and then purified by preparative LC−MS as described above (see the Supporting Information for the LC−MS analysis of representative crude NNP hits).

Hit Characterization. The hit binding affinities to various targets were determined using microscale thermophoresis (MST). MST experiments were performed on a Monolith NT.115pico (NanoTemper Technologies GmbH, Munich, Germany). Measurements were performed at room temperature, in triplicate, with incubation periods of 15, 30, and 45 min. Binding affinities were obtained from a 16 point, 2-fold dilution series with the ligand starting concentration at 1 μM and target concentration at 5 nM. Targets were labeled using Nanotemper Monolith second-generation protein labeling kits. A RED-MALEIMIDE (Maleimide-647-dye) labeling kit was used for K-Ras, and a RED-NHS (NHS-647-dye) labeling kit was used for IL-6, IL-6R, TNFα, and ASGPR. The buffer for the ASGPR contained 20 mM HEPES pH 7.4, 150 mM NaCl, 10 mM $MgCl_2$, 2 mM $CaCl_2$, 0.05% Pluronic F-127, and 1 mM DTT; that for IL-6 20 mM HEPES pH 7.4, 150 mM KCl, 10 mM $MgCl_2$, and 0.1% Pluronic F-127; that for the soluble IL-6 receptor 20 mM HEPES pH 7.4, 150 mM NaCl, 10 mM $MgCl_2$, and 0.05% Tween-20; that for K-Ras 20 mM HEPES pH 7.4, 150 mM NaCl, 10 mM $MgCl_2$, 0.05% Tween-20, and 1 mM

https://doi.org/10.1021/acscentsci.1c01041
ACS Cent. Sci. 2022, 8, 86−101

DTT; and that for TNFα 10 mM HEPES pH 7.4, 150 mM NaCl, 10 mM $MgCl_2$, 0.05% Polysorbate-20. Triplicate data was analyzed using MO.AffinityAnalysis software (NanoTemper Technologies GmbH).

K-Ras/Raf Inhibition Assay. The interaction between the target protein K-Ras and the ligand protein Raf was investigated using a microscale thermophoresis (HTS-MST) assay in the absence and presence of three NNP hits: KRAS-1−4, KRAS-1−8, and KRAS-1−13. The loading of K-Ras with the GTP was performed according to the protocol detailed above (Activation of K-Ras by GTP Loading for the Screen and Binding Assays section). After the GTP loading, the resulting protein solution was buffer exchanged into 100 mM HEPES pH 6.5, 5 mM $MgCl_2$, 50 mM NaCl, 1 mM TCEP. The resulting concentration of the target protein was 8.9 μM, which was used for Maleimide-647-dye labeling. For the interaction between K-Ras and Raf with no NNP present, two technical runs with the same samples were performed between GTP-loaded K-Ras and Raf in the same buffer conditions that were used to test the interaction between K-Ras and the NNP hits: 20 mM HEPES pH 7.4, 150 mM NaCl, 10 mM $MgCl_2$, 1 mM DTT, 0.05% Tween-20. For the interaction between K-Ras and Raf in the presence of NNPs, the labeled target protein K-Ras was diluted to 10 nM in assay buffer containing 2 μM NNP and incubated at room temperature for 15 min. This solution was then mixed with the ligand protein Raf serial dilution 1:1 to yield the final assay samples with 5 nM target protein and 1 μM NNP.

Cell Culture for the ASGPR Uptake Assay. The HEK293T (human embryonic kidney cells) and human hepatoma HepG2 cells were grown according to the protocols provided by the American Type Culture Collection (ATCC). Cells were seeded at ~1.5×10^5 cells/well in a 24-well culture plate for the uptake assay. After at least 16 h of culture to allow cells to attach and equilibrate, the compound treatment was set up for the uptake assay.

ASGPR Uptake Assay. The fluorescein-labeled NNPs or fluorescein-labeled trivalent ligand *N*-acetylgalactosamine (tri-GalNAc) recognizing ASGRPR was added to wells at indicated concentrations and incubated for 2 h. Two plates were prepared for each treatment condition, one serving as the 4 °C no internalization control that was kept on ice during incubation, while the second plate was incubated at 37 °C to allow for energy-dependent internalization. Following the incubation period, all plates were placed on ice and washed three times with ice-cold PBS/3% BSA/2 mM EDTA and then lifted with trypsin. Cells were transferred to 96-well round-bottom plates in FACS buffer. Cells were then analyzed by flow cytometry using LSR-II with an HTS sampler (BD Biosciences, San Jose, CA). Data (mean fluorescence intensities) was further analyzed using Flowjo software (BD Biosciences, San Jose, CA). The internalized fraction was expressed as the difference between the corresponding 4 and 37 °C MFIs as previously described.[59]

For the competitive uptake assay with the natural ligand of ASGPR (asialofetuin), cells were preincubated with 20 and 60 μM asialofetuin on ice or at 37 °C for 1.5 h, followed by the treatment with compounds for an additional 2 h before the flow cytometric analysis as described above. The reduction of the uptake under asialofetuin competition was expressed by the percentage against the same treatment condition without asialofetuin.[67]

Stability Assays. For proteinase K stability, solutions of each tested compound (200 μM) in 10% DMSO and 20 mM Tris HCl at pH 8 were prepared. Proteinase K was added to a final concentration of 100 μg/mL and 100 μM of the tested compound in 5% DMSO and 10 mM Tris HCl at pH 8. The solutions were incubated at 37 °C, and aliquots after 0, 1.5, and 16 h were analyzed by LC−MS.

For human plasma stability, lyophilized human plasma was reconstituted in sterile water for injection, aliquoted into 200 μL aliquots and stored frozen at −80 °C prior to the stability studies. For the stability studies, three aliquots per NNP were thawed at room temperature. An additional set of three aliquots for a positive control peptide (Angiotensin I) were also thawed. The incubations for the plasma stability were initiated by mixing 2 μL of a 2 mM stock solution of NNP IL6R-87-8 or positive control in DMSO with the 200 μL thawed plasma aliquot. After briefly vortex-mixing, 50 μL zero-time-point samples were removed, mixed with 50 μL of water and 400 μL of acetonitrile, and frozen on dry ice until all time point samples were collected. The samples were incubated at 37 °C with samples removed and water/acetonitrile added at 1, 3, and 17 h. Proteins were precipitated by centrifuging the samples at 17 000*g*, 4 °C, for 1 h. The supernatants were removed and concentrated in a centrifugal vacuum concentrator (GeneVac Genie II) at 45 °C until the volume had been reduced to ~60 μL. The samples were then diluted with 95% water, 5% acetonitrile, and 0.1% formic acid to a volume of 200 μL; , 2 μL of an internal standard peptide (Val5-Angiotensin I, Sigma) was added prior to sample analysis by LC−MS on the LTQ-Orbitrap XL system described above.

Safety Statement. All chemical synthesis procedures were performed in appropriate fume hoods using standard chemistry best practices. No unexpected or unusually high safety hazards were encountered.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acscentsci.1c01041.

Supplemental tables and figures, general experimental procedures, detailed synthetic methods, and analytical data on all compounds (PDF)

Video S1: Plate and hit location data from FAST are transferred to an automated fluorescence microscope and picking robot. The robot uses a glass capillary on a robotic arm to extract the bead from the plate and place it in a tube for sequencing (MP4)

■ AUTHOR INFORMATION

Corresponding Authors

Peter B. Madrid − *SRI Biosciences Division, SRI International, Menlo Park, California 94025, United States;* orcid.org/0000-0002-4895-5610; Email: peter.madrid@sri.com

Nathan Collins − *SRI Biosciences Division, SRI International, Menlo Park, California 94025, United States;* orcid.org/0000-0002-5793-4870; Email: nathan.collins@sri.com

Authors

Michal Avital-Shmilovici − *SRI Biosciences Division, SRI International, Menlo Park, California 94025, United States*

Xiaohe Liu − *SRI Biosciences Division, SRI International, Menlo Park, California 94025, United States*

Thomas Shaler − *SRI Biosciences Division, SRI International, Menlo Park, California 94025, United States*

Andrew Lowenthal − *SRI Biosciences Division, SRI International, Menlo Park, California 94025, United States*

https://doi.org/10.1021/acscentsci.1c01041
ACS Cent. Sci. 2022, 8, 86−101

Pauline Bourbon − *SRI Biosciences Division, SRI International, Menlo Park, California 94025, United States*
Janey Snider − *SRI Biosciences Division, SRI International, Menlo Park, California 94025, United States*
Arlyn Tambo-Ong − *SRI Biosciences Division, SRI International, Menlo Park, California 94025, United States*
Claire Repellin − *SRI Biosciences Division, SRI International, Menlo Park, California 94025, United States*
Kenya Yniguez − *SRI Biosciences Division, SRI International, Menlo Park, California 94025, United States*
Lidia Sambucetti − *SRI Biosciences Division, SRI International, Menlo Park, California 94025, United States*

Complete contact information is available at: https://pubs.acs.org/10.1021/acscentsci.1c01041

Author Contributions

N.C. conceived of the concepts presented here with support from P.B.M. and M.A.-S. M.A.-S., T.S., X.L., P.B., J.S., C.R., A.L., A.T.-O., K.Y., and L.S. ran the experiments. M.A.-S., P.B.M., and N.C. wrote the manuscript.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by the DARPA Fold F(X) program N66001-14-C-4059. (Distribution Statement "A"—Approved for Public Release, Distribution Unlimited. The views, opinions, and/or findings expressed are those of the authors and should not be interpreted as representing the official views or policies of the Department of Defense or the U.S. Government.) We also thank SRI for internal R&D support as well as experimental guidance from our colleagues Kathlynn Brown and Michael McGuire. The reference tri-GalNAc ligand for ASGPR was kindly provided by Ionis Pharmaceuticals, Inc. The K-Ras−Raf competition assay was performed by 2bind GmbH (Germany).

■ REFERENCES

(1) Barbas, C. F., 3rd; Kang, A. S.; Lerner, R. A.; Benkovic, S. J. Assembly of combinatorial antibody libraries on phage surfaces: the gene III site. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88* (18), 7978−7982.
(2) Nixon, A. E.; Sexton, D. J.; Ladner, R. C. Drugs derived from phage display: from candidate identification to clinical practice. *mAbs* **2014**, *6* (1), 73−85.
(3) Maeda, Y.; Javid, N.; Duncan, K.; Birchall, L.; Gibson, K. F.; Cannon, D.; Kanetsuki, Y.; Knapp, C.; Tuttle, T.; Ulijn, R. V.; et al. Discovery of catalytic phages by biocatalytic self-assembly. *J. Am. Chem. Soc.* **2014**, *136* (45), 15893−15896.
(4) Seker, U. O.; Demir, H. V. Material binding peptides for nanotechnology. *Molecules* **2011**, *16* (2), 1426−1451.
(5) Langan, R. A.; Boyken, S. E.; Ng, A. H.; Samson, J. A.; Dods, G.; Westbrook, A. M.; Nguyen, T. H.; Lajoie, M. J.; Chen, Z.; Berger, S.; et al. De novo design of bioactive protein switches. *Nature* **2019**, *572* (7768), 205−210.
(6) Huang, P. S.; Boyken, S. E.; Baker, D. The coming of age of de novo protein design. *Nature* **2016**, *537* (7620), 320−327.
(7) Wachsmuth, M.; Findeiß, S.; Weissheimer, N.; Stadler, P. F.; Mörl, M. De novo design of a synthetic riboswitch that regulates transcription termination. *Nucleic Acids Res.* **2013**, *41* (4), 2541−2551.
(8) Hill, D. J.; Mio, M. J.; Prince, R. B.; Hughes, T. S.; Moore, J. S. A field guide to foldamers. *Chem. Rev.* **2001**, *101* (12), 3893−4012.
(9) Zhao, Y.; Moore, J. S. Foldamers Based on Solvophobic Effects. *Foldamers* **2007**, 75−108.
(10) Gellman, S. H. Foldamers: A Manifesto. *Acc. Chem. Res.* **1998**, *31* (4), 173−180.
(11) Chaput, J. C.; Herdewijn, P. What Is XNA? *Angewandte Chemie (International ed. in English)* **2019**, *58* (34), 11570−11572.
(12) Al Ouahabi, A.; Charles, L.; Lutz, J.-F. Synthesis of Non-Natural Sequence-Encoded Polymers Using Phosphoramidite Chemistry. *J. Am. Chem. Soc.* **2015**, *137* (16), 5629−5635.
(13) Chen, Z.; Lichtor, P. A.; Berliner, A. P.; Chen, J. C.; Liu, D. R. Evolution of sequence-defined highly functionalized nucleic acid polymers. *Nat. Chem.* **2018**, *10* (4), 420−427.
(14) Kodadek, T. Synthetic receptors with antibody-like binding affinities. *Curr. Opin. Chem. Biol.* **2010**, *14* (6), 713−720.
(15) Badi, N.; Lutz, J.-F. Sequence control in polymer synthesis. *Chem. Soc. Rev.* **2009**, *38* (12), 3383−3390. 10.1039/B806413J
(16) Chan-Seng, D.; Lutz, J.-F. Solid-Phase Synthesis as a Tool for the Preparation of Sequence-Defined Oligomers Based on Natural Amino Acids and Synthetic Building Blocks. In *Sequence-Controlled Polymers: Synthesis, Self-Assembly, and Properties*; ACS Symposium Series; American Chemical Society, 2014; Vol. *1170*, pp 103−116.
(17) Goto, Y.; Suga, H. The RaPID Platform for the Discovery of Pseudo-Natural Macrocyclic Peptides. *Acc. Chem. Res.* **2021**, *54* (18), 3604−3617.
(18) Lam, K. S.; Salmon, S. E.; Hersh, E. M.; Hruby, V. J.; Kazmierski, W. M.; Knapp, R. J. A new type of synthetic peptide library for identifying ligand-binding activity. *Nature* **1991**, *354* (6348), 82−84.
(19) Komnatnyy, V. V.; Nielsen, T. E.; Qvortrup, K. Bead-based screening in chemical biology and drug discovery. *Chem. Commun.* **2018**, *54* (50), 6759−6771.
(20) Lam, K. S.; Lebl, M.; Krchňák, V. The "One-Bead-One-Compound" Combinatorial Library Method. *Chem. Rev.* **1997**, *97* (2), 411−448.
(21) Sklar, L. A.; Carter, M. B.; Edwards, B. S. Flow cytometry for drug discovery, receptor pharmacology and high-throughput screening. *Current opinion in pharmacology* **2007**, *7* (5), 527−534.
(22) Gates, Z. P.; Vinogradov, A. A.; Quartararo, A. J.; Bandyopadhyay, A.; Choo, Z. N.; Evans, E. D.; Halloran, K. H.; Mijalis, A. J.; Mong, S. K.; Simon, M. D.; et al. Xenoprotein engineering via synthetic libraries. *Proc. Natl. Acad. Sci. U.S.A.* **2018**, *115* (23), E5298−e5306.
(23) Carney, R. P.; Thillier, Y.; Kiss, Z.; Sahabi, A.; Heleno Campos, J. C.; Knudson, A.; Liu, R.; Olivos, D.; Saunders, M.; Tian, L.; et al. Combinatorial Library Screening with Liposomes for Discovery of Membrane Active Peptides. *ACS combinatorial science* **2017**, *19* (5), 299−307.
(24) Quartararo, A. J.; Gates, Z. P.; Somsen, B. A.; Hartrampf, N.; Ye, X.; Shimada, A.; Kajihara, Y.; Ottmann, C.; Pentelute, B. L. Ultra-large chemical libraries for the discovery of high-affinity peptide binders. *Nat. Commun.* **2020**, *11* (1), 3183.
(25) 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. U. S. A.* **1992**, *89* (10), 4505−4509.
(26) 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.* **1997**, *11* (11), 1178−1184.
(27) 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.* **1997**, *16* (5), 505−511.
(28) van Breemen, R. B.; Huang, C.-R.; Nikolic, D.; Woodbury, C. P.; Zhao, Y.-Z.; Venton, D. L. Pulsed Ultrafiltration Mass Spectrometry: A New Method for Screening Combinatorial Libraries. *Anal. Chem.* **1997**, *69* (11), 2159−2164.
(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.* **2016**, *138* (4), 1335−1343.
(30) Aebersold, R.; Mann, M. Mass spectrometry-based proteomics. *Nature* **2003**, *422* (6928), 198−207.

(31) Thakkar, A.; Cohen, A. S.; Connolly, M. D.; Zuckermann, R. N.; Pei, D. High-throughput sequencing of peptoids and peptide-peptoid hybrids by partial edman degradation and mass spectrometry. *J. Comb. Chem.* **2009**, *11* (2), 294−302.
(32) Lebl, M.; Krchnák, V.; Sepetov, N. F.; Seligmann, B.; Strop, P.; Felder, S.; Lam, K. S. One-bead-one-structure combinatorial libraries. *Biopolymers* **1995**, *37* (3), 177−198.
(33) Gao, Y.; Amar, S.; Pahwa, S.; Fields, G.; Kodadek, T. Rapid Lead Discovery Through Iterative Screening of One Bead One Compound Libraries. *ACS Comb. Sci.* **2015**, *17* (1), 49−59.
(34) Paulick, M. G.; Hart, K. M.; Brinner, K. M.; Tjandra, M.; Charych, D. H.; Zuckermann, R. N. Cleavable hydrophilic linker for one-bead-one-compound sequencing of oligomer libraries by tandem mass spectrometry. *J. Comb. Chem.* **2006**, *8* (3), 417−426.
(35) Kodadek, T.; McEnaney, P. J. Towards vast libraries of scaffold-diverse, conformationally constrained oligomers. *Chemical communications (Cambridge, England)* **2016**, *52* (36), 6038−6059.
(36) Youngquist, R. S.; Fuentes, G. R.; Lacey, M. P.; Keough, T. Generation and screening of combinatorial peptide libraries designed for rapid sequencing by mass spectrometry. *J. Am. Chem. Soc.* **1995**, *117* (14), 3900−3906.
(37) Wang, X.; Peng, L.; Liu, R.; Gill, S. S.; Lam, K. S. Partial Alloc-Deprotection Approach for Ladder Synthesis of "One-Bead One-Compound" Combinatorial Libraries. *J. Comb. Chem.* **2005**, *7* (2), 197−209.
(38) Chait, B. T.; Wang, R.; Beavis, R. C.; Kent, S. B. Protein ladder sequencing. *Science (Washington, DC, U. S.)* **1993**, *262* (5130), 89−92.
(39) Wang, P.; Arabaci, G.; Pei, D. Rapid Sequencing of Library-Derived Peptides by Partial Edman Degradation and Mass Spectrometry. *J. Comb. Chem.* **2001**, *3* (3), 251−254.
(40) Liu, R.; Marik, J.; Lam, K. S. A Novel Peptide-Based Encoding System for "One-Bead One-Compound" Peptidomimetic and Small Molecule Combinatorial Libraries. *J. Am. Chem. Soc.* **2002**, *124* (26), 7678−7680.
(41) Liu, X.; Hsieh, H. B.; Campana, D.; Bruce, R. H. A new method for high speed, sensitive detection of minimal residual disease. *Cytometry, Part A* **2012**, *81A* (2), 169−175.
(42) Krivacic, R. T.; Ladanyi, A.; Curry, D. N.; Hsieh, H. B.; Kuhn, P.; Bergsrud, D. E.; Kepros, J. F.; Barbera, T.; Ho, M. Y.; Chen, L. B.; et al. A rare-cell detector for cancer. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101* (29), 10501−10504.
(43) Ao, Z.; Liu, X. Fiber-Optic Array Scanning Technology (FAST) for Detection and Molecular Characterization of Circulating Tumor Cells. *Methods Mol. Biol.* **2017**, *1634*, 235−246.
(44) Curry, D. N.; Krivacic, R. T.; Hsieh, H. B.; Ladanyi, A.; Bergsrud, D. E.; Ho, M. Y.; Chen, L. B.; Kuhn, P.; Bruce, R. H. High-speed detection of occult tumor cells in peripheral blood. *Conf Proc. IEEE Eng. Med. Biol. Soc.* **2004**, *2004*, 1267−1270.
(45) DeMaster, L. K.; Liu, X.; VanBelzen, D. J.; Trinité, B.; Zheng, L.; Agosto, L. M.; Migueles, S. A.; Connors, M.; Sambucetti, L.; Levy, D. N.; et al. A Subset of CD4/CD8 Double-Negative T Cells Expresses HIV Proteins in Patients on Antiretroviral Therapy. *J. Virol* **2016**, *90* (5), 2165−2179.
(46) Wang, D.; Wu, L.; Liu, X. Glycan Markers as Potential Immunological Targets in Circulating Tumor Cells. *Adv. Exp. Med. Biol.* **2017**, *994*, 275−284.
(47) 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.* **2003**, *125* (46), 13995−14004.
(48) Olivos, H. J.; Bachhawat-Sikder, K.; Kodadek, T. Quantum dots as a visual aid for screening bead-bound combinatorial libraries. *ChemBioChem* **2003**, *4* (11), 1242−1245.
(49) Pei, D. On-bead library screening made easier. *Chem. Biol.* **2010**, *17* (1), 3−4.
(50) Liu, X.; Hsieh, H. B.; Campana, D.; Bruce, R. H. A new method for high speed, sensitive detection of minimal residual disease. *Cytometry A* **2012**, *81* (2), 169−175.
(51) Chen, X.; Tan, P. H.; Zhang, Y.; Pei, D. On-bead screening of combinatorial libraries: reduction of nonspecific binding by decreasing surface ligand density. *J. Comb. Chem.* **2009**, *11* (4), 604−611.
(52) Mitchell, A. R.; Erickson, B. W.; Ryabtsev, M. N.; Hodges, R. S.; Merrifield, R. B. Tert-butoxycarbonylaminoacyl-4-(oxymethyl)-phenylacetamidomethyl-resin, a more acid-resistant support for solid-phase peptide synthesis. *J. Am. Chem. Soc.* **1976**, *98* (23), 7357−7362.
(53) Albericio, F., Ed. *Solid-Phase Synthesis*; CRC Press: Boca Raton, FL, 2000; DOI: 10.1201/9781482270303.
(54) Thieriet, N.; Alsina, J.; Giralt, E.; Guibé, F.; Albericio, F. Use of Alloc-amino acids in solid-phase peptide synthesis. Tandem deprotection-coupling reactions using neutral conditions. *Tetrahedron Lett.* **1997**, *38* (41), 7275−7278.
(55) Zubarev, R. A.; Makarov, A. Orbitrap Mass Spectrometry. *Anal. Chem.* **2013**, *85* (11), 5288−5296.
(56) Hecht, E. S.; Scigelova, M.; Eliuk, S.; Makarov, A.; Meyers, R. A. Fundamentals and Advances of Orbitrap Mass Spectrometry. *Encyclopedia of Analytical Chemistry* **2019**, 1−40.
(57) Hsu, F.; Pringle, T. H.; Kuhn, R. M.; Karolchik, D.; Diekhans, M.; Haussler, D.; Kent, W. J. The UCSC Proteome Browser. *Nucleic Acids Res.* **2004**, *33*, D454−D458.
(58) Stephen, A. G.; Esposito, D.; Bagni, R. K.; McCormick, F. Dragging ras back in the ring. *Cancer Cell* **2014**, *25* (3), 272−281.
(59) Tanowitz, M.; Hettrick, L.; Revenko, A.; Kinberger, G. A.; Prakash, T. P.; Seth, P. P. Asialoglycoprotein receptor 1 mediates productive uptake of N-acetylgalactosamine-conjugated and unconjugated phosphorothioate antisense oligonucleotides into liver hepatocytes. *Nucleic Acids Res.* **2017**, *45* (21), 12388−12400.
(60) Nair, J. K.; Willoughby, J. L.; Chan, A.; Charisse, K.; Alam, M. R.; Wang, Q.; Hoekstra, M.; Kandasamy, P.; Kel'in, A. V.; Milstein, S.; et al. Multivalent N-acetylgalactosamine-conjugated siRNA localizes in hepatocytes and elicits robust RNAi-mediated gene silencing. *J. Am. Chem. Soc.* **2014**, *136* (49), 16958−16961.
(61) Sanhueza, C. A.; Baksh, M. M.; Thuma, B.; Roy, M. D.; Dutta, S.; Préville, C.; Chrunyk, B. A.; Beaumont, K.; Dullea, R.; Ammirati, M.; et al. Efficient Liver Targeting by Polyvalent Display of a Compact Ligand for the Asialoglycoprotein Receptor. *J. Am. Chem. Soc.* **2017**, *139* (9), 3528−3536.
(62) Huang, X.; Leroux, J. C.; Castagner, B. Well-Defined Multivalent Ligands for Hepatocytes Targeting via Asialoglycoprotein Receptor. *Bioconjugate Chem.* **2017**, *28* (2), 283−295.
(63) Westerlind, U.; Westman, J.; Törnquist, E.; Smith, C. I.; Oscarson, S.; Lahmann, M.; Norberg, T. Ligands of the asialoglycoprotein receptor for targeted gene delivery, part 1: Synthesis of and binding studies with biotinylated cluster glycosides containing N-acetylgalactosamine. *Glycoconj. J.* **2004**, *21* (5), 227−241.
(64) Nanjan, P.; Porel, M. Sequence-defined non-natural polymers: synthesis and applications. *Polym. Chem.* **2019**, *10* (40), 5406−5424.
(65) McGuire, M. J.; Gray, B. P.; Li, S.; Cupka, D.; Byers, L. A.; Wu, L.; Rezaie, S.; Liu, Y. H.; Pattisapu, N.; Issac, J.; et al. Identification and characterization of a suite of tumor targeting peptides for non-small cell lung cancer. *Sci. Rep.* **2015**, *4*, 4480.
(66) Liu, F.; Du, D.; Fuller, A. A.; Davoren, J. E.; Wipf, P.; Kelly, J. W.; Gruebele, M. An experimental survey of the transition between two-state and downhill protein folding scenarios. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105* (7), 2369−2374.
(67) Westerlind, U.; Westman, J.; Törnquist, E.; Smith, C. I. E.; Oscarson, S.; Lahmann, M.; Norberg, T. Ligands of the asialoglycoprotein receptor for targeted gene delivery, part 1: Synthesis of and binding studies with biotinylated cluster glycosides containing N-acetylgalactosamine. *Glycoconjugate J.* **2004**, *21* (5), 227−241.

NOTE ADDED AFTER ASAP PUBLICATION

This paper was published ASAP on January 11, 2022, with errors in Figure 2. The corrected version was reposted on January 13, 2022.

https://doi.org/10.1021/acscentsci.1c01041
ACS Cent. Sci. 2022, 8, 86−101