

The passive permeability landscape around geometrically diverse hexa- and heptapeptide macrocycles.

Chad E. Townsend, Mathew R. Naylor, Eva Jason, Cameron R. Pye†, Joshua A. Schwochert†, Quinn Edmondson, R. Scott Lokey*.

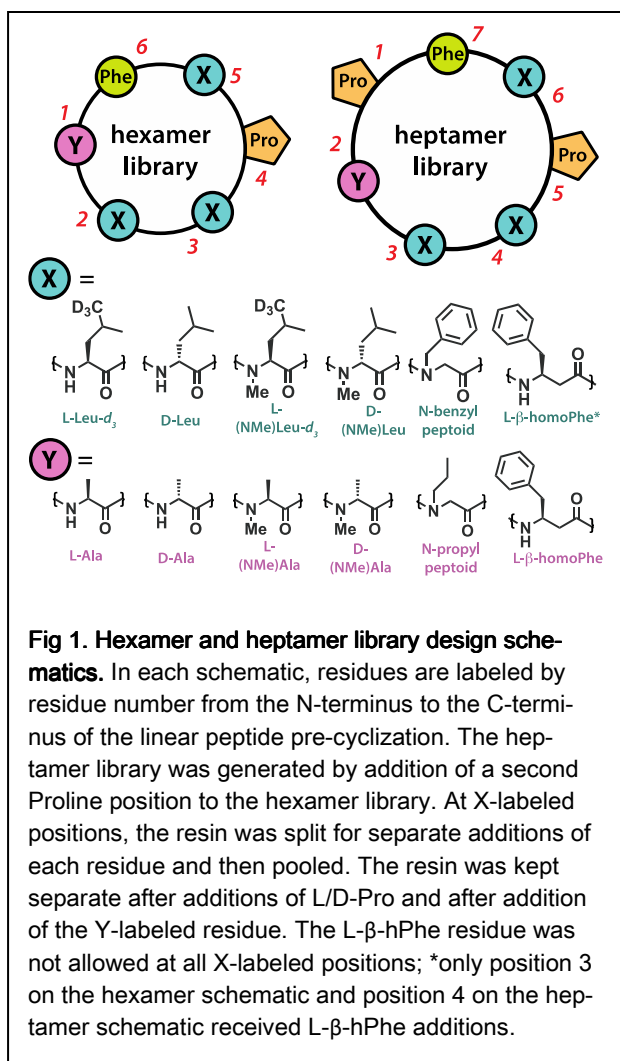
Department of Chemistry and Biochemistry, University of California Santa Cruz, 1156 High St., Santa Cruz, CA, 95060

†Unnatural Products, Inc., Santa Cruz, CA 95060, US

*Corresponding Author: slokey@ucsc.edu

ABSTRACT: Recent advances in DNA and mRNA encoding technologies have enabled the discovery of high-affinity macrocyclic peptides and peptide-like ligands against virtually any protein target of interest. However, even the most potent biochemical leads from these screening technologies often have only weak cellular activity due to poor membrane permeability. Biasing such libraries towards passive cell permeability in the design phase would facilitate development of leads against intracellular targets. We set out to empirically evaluate the intrinsic permeability of thousands of geometrically diverse hexa- and heptapeptide scaffolds by permuting backbone stereochemistry and N-methylation, and by including peptoid and β -amino acid residues at select positions, with the goals of providing a resource for biasing library-based screening efforts toward passive membrane permeability and studying the effects of the backbone elements introduced on a large number of compounds. Libraries were synthesized via standard split-pool solid phase peptide synthesis, and passive permeability was measured in pools of 150 compounds using a highly multiplexed version of the parallel artificial membrane permeability assay (PAMPA) under sink conditions. Compounds were identified using CycLS, a high-resolution mass spectrometry-based method that uses stable isotopes to encode stereochemistry and matches MSMS data to virtual fragment libraries based on the expected macrocyclic products. From the compounds that were identified with high confidence, 823 hexameric and 1330 heptameric scaffolds had PAMPA permeability coefficients greater than 1×10^{-6} cm/s. The prevalence of high permeability compounds in these two libraries suggests that passive permeability is achievable for hexa- and heptapeptides with highly diverse backbone geometries.

Introduction



Cyclic peptides have demonstrated the ability to inhibit a wide range of protein-protein interactions¹ and reach intracellular targets², with new cyclic peptide therapeutics reaching the market each year³. Powerful library construction and screening technologies such as mRNA display have made discovering potent cyclic peptide binders easier than ever^{4, 5}. Their ease of synthesis and diversification have made cyclic peptides an attractive platform for high throughput screening against protein-protein interactions and other targets that are difficult to inhibit with small molecule therapeutics. However, realizing that potential by achieving membrane permeability, which is a prerequisite for oral bioavailability and the ability to engage intracellular targets, remains a major barrier to the development of therapeutic cyclic peptides^{6, 7}.

The ligands produced by DNA/mRNA-encoded platforms are selected only for potency, which is often driven by charged and/or

highly polar residues that are generally incompatible with passive cell permeability. Yet, optimizing a lead compound for passive permeability in concert with target binding is often difficult even without these polar moieties. Biasing initial library designs toward passive cell permeability through inclusion of scaffolds with favorable ADME properties that can permit incorporation of sparse polar moieties may ease subsequent medicinal chemistry optimization.

Three-dimensional structure has been shown to be highly impactful on passive permeability in cyclic peptides⁸⁻¹⁰, with modifications to backbone geometry producing dramatic effects. In one study a single stereochemical inversion resulted in epimers with permeation rates that varied by nearly 100-fold¹¹. Sidechain modifications often modify permeability through simple lipophilicity rather than the conformational shifts implied by such permeability cliffs¹²⁻¹⁴. We therefore hypothesized that a large set of cyclic peptide backbone geometries of known permeability would both aid our understanding of passive permeation and bias the design library screens towards passively permeable hits through sidechain modification of permeable backbones.

We set out to chart the permeability landscape in geometrically diverse cyclic hexa- and heptapeptides using a one-bead one-compound (OBOC) based synthesis and highly parallel analytical approach that we have previously reported¹¹. Whereas previously we relied on resynthesis to decode stereochemical and sequence ambiguities when analyzing whole libraries, here we use stable isotopes to encode stereochemistry and an MSMS-based approach that we developed previously called CycLS¹⁵ to distinguish among isomers with different sequences. To cover the broadest possible conformational landscape in this size range, we varied stereochemistry and N-methylation, factors that have been shown to have a significant impact on passive permeability^{8, 9, 16, 17}. Proline residues were included to facilitate β -turn like conformations prone to forming internal hydrogen bonding networks and to increase cyclization efficiency. We also included β -amino acid and peptoid residues to demonstrate the utility of this method toward incorporating unnatural or potentially druglike moieties. In addition to increasing backbone flexibility, peptoids impart distinct conformational characteristics to a macrocycle by allowing both cis- and trans-rotamers at the N-substituted amide. They also provide potential sites for side chain diversification in future compounds or libraries based on these scaffolds, and we have previously shown that substitution of peptoids at specific positions in a cyclic peptide scaffold can preserve or even enhance membrane permeability^{12, 18}.

Although hexapeptide and peptide-peptoid hybrids have been extensively studied using similar approaches^{8, 9, 19-21}, few studies have explored the permeability landscape in heptapeptides. Herein we describe an extensive and unbiased survey of the permeability landscape in cyclic hexa- and hepta- peptides and peptide-peptoid hybrids. The results reveal not only the expected dependence of passive permeability on backbone geometry, but also show a remarkably high degree of passive permeability among all library members. Additionally, we observed strong correlations between measured mixture and pure compound PAMPA permeabilities, confirming the potential applicability of this split-pool/multiplexed approach to scaffold discovery toward the design of encoded libraries that are biased toward passive cell permeability.

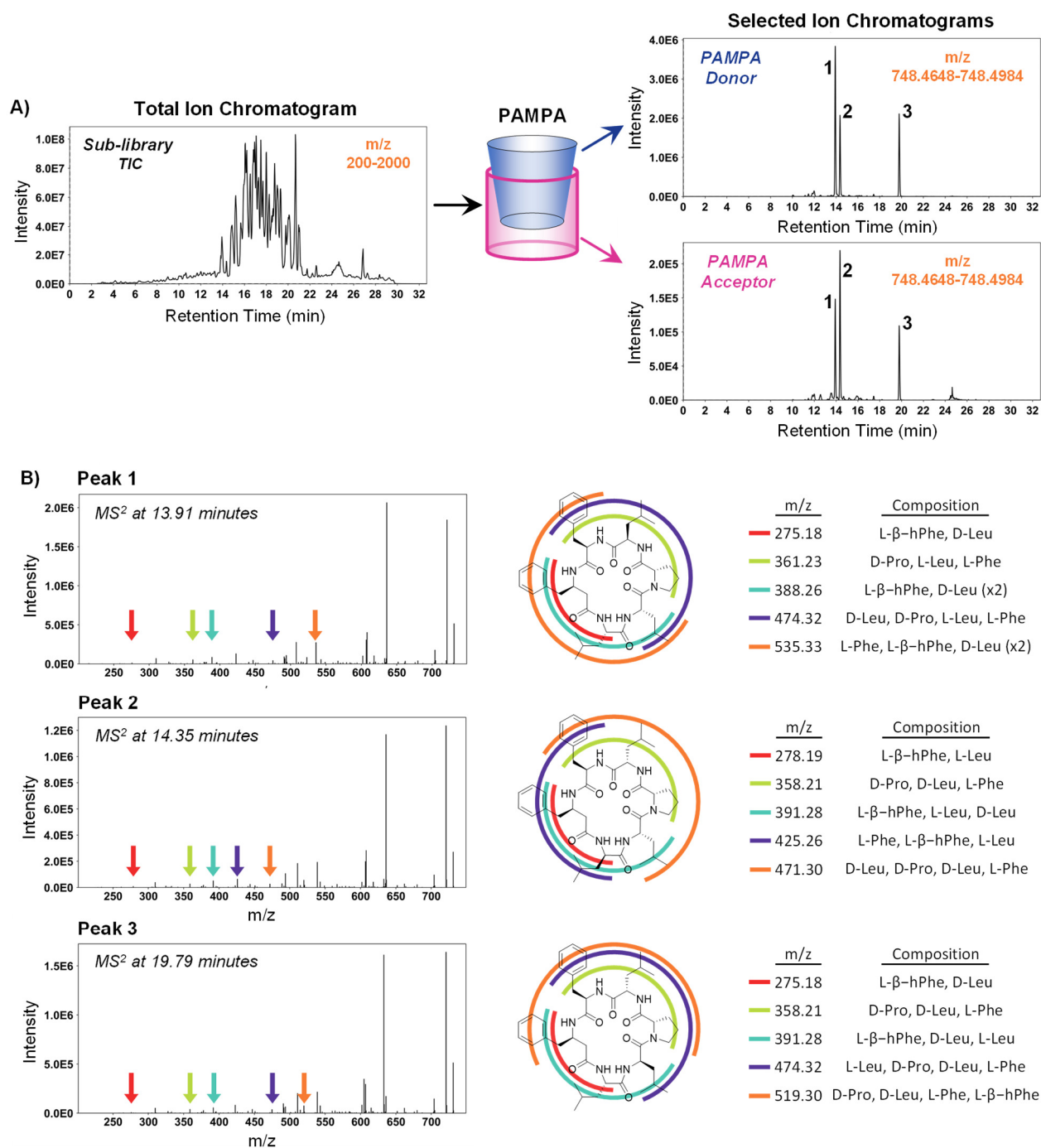
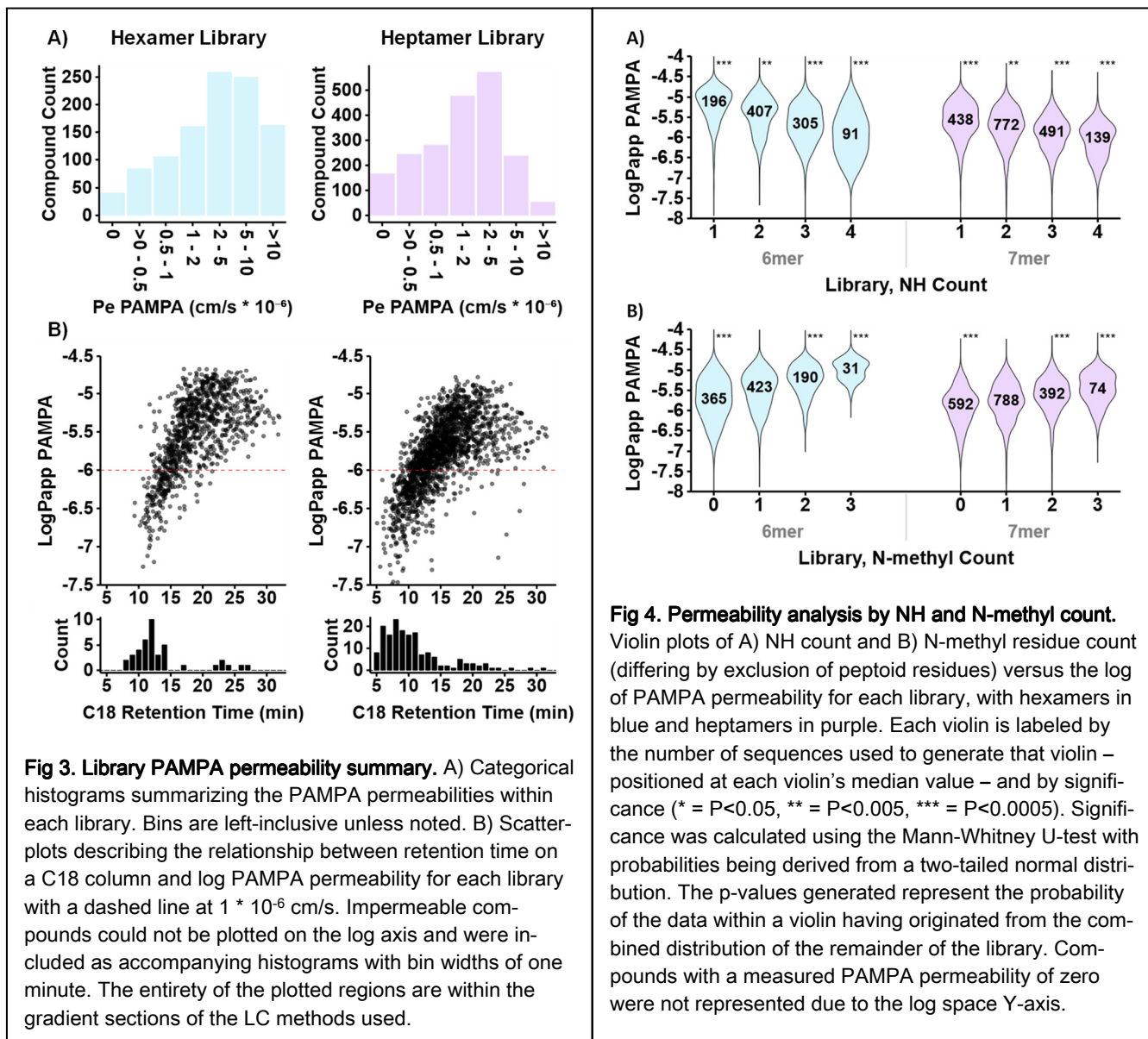


Fig 2. Analysis of mixtures workflow. A) A mixture of 150 cyclic hexa-peptides and peptide-peptoid hybrids had its passive permeation quantified by PAMPA. A total ion chromatogram of the mixture and extracted ion chromatograms of the donor and acceptor wells for a library mass with three expected compounds is shown. The MS² data of the three peaks were compared to virtual fragment libraries of each potential sequence to determine the position of an L-Leu- α_3 residue. B) Each MS² spectrum corresponding to a chromatographic peak in the extracted ion chromatograms is accompanied by the chemical structure of the library member that had the strongest match. Select fragments containing sequencing information key to distinguishing among the three candidates have been marked by colored arrows on the MS² spectrum and arcs denoting their position on the corresponding chemical structure. The table entry corresponding to each color notes the mass/charge ratio of each fragment and its composition. Individual cyclic peptide fragments do not generally provide sequence information directly, so information from multiple fragments must be combined to determine a sequence.



Results and Discussion

Using an approach that we reported previously¹⁵, we synthesized a library of 1800 cyclic hexapeptide and 3600 cyclic heptapeptide scaffolds using split-pool solid phase peptide synthesis (SPPS), starting with L-Phe attached to the solid support through a 2-chlorotrityl linker (in describing library synthesis residues will be addressed by their number in the hexamer library design schematic because the heptamer library was generated from a portion of the linear hexamer library)(Fig. 1). The resin was split into five pools for attachment of residue 5, which included L-Leu-*d*₃, D-Leu, L-(NMe)Leu-*d*₃, D-(NMe)Leu, and benzyl peptoid (peptoids were incorporated using the submonomer method²²; all other residues were incorporated using standard Fmoc SPPS conditions). After addition of residue 5, the resin was pooled and split once again for incorporation of either L-Pro or D-Pro; these subpools were kept separate for the remainder of

the synthesis. Each of the Pro pools was split into six sub-pools for the addition of residue 3, which included the same building blocks used at residue 5 with the addition of L- β -homoPhe (L- β -hPhe). Splitting and pooling was continued for incorporation of position 2, which was identical to position 5, and position 1, which included L-Ala, D-Ala, L-(NMe)Ala, D-(NMe)Ala, L- β -homo-Phe, and propyl peptoid, resulting in twelve mixtures defined by the known identities of residues 4 and 1. Each of the twelve mixtures of 150 compounds were split into two lots, with one lot cyclized (to generate the 1800 cyclic hexapeptides) and the other lot set aside for the addition of another L- or D-Pro to generate the 3600 cyclic heptapeptides. Cyclization was performed in dilute 1:1 ACN:THF with COMU, and each pool was purified by solid phase extraction on C18 media.

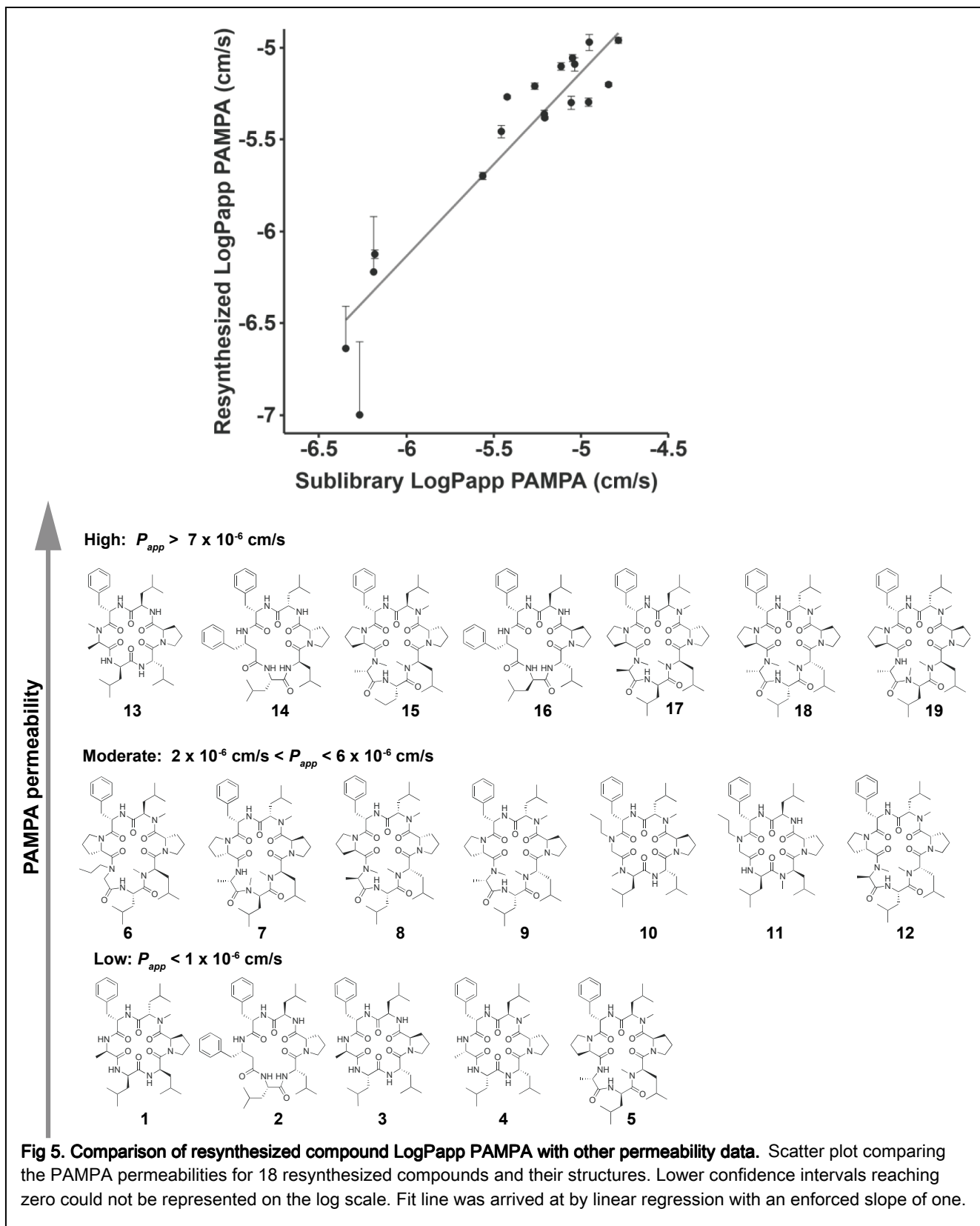
The resulting 12 hexamer and 24 heptamer sub-libraries were inspected by LCMS to assess synthetic success. Each sub-library was confirmed by LCMS to have approximately the expected number of major peaks at each relevant accurate mass through inspection of individual extracted ion chromatograms. Truncations were detected post-purification but most eluted before the start of the LCMS gradient and none shared a mass with an expected library compound. To measure synthetic success of the libraries on a peak-by-peak basis, we performed an analysis of all major peaks at each expected library mass of a single exemplar heptamer sub-library and were able to locate 146 of the 150 expected peaks. An analysis of an exemplar hexamer sub-library from our previous work on CycLS confirmed 139 of 150 expected peaks (SI Section 4.1).

Each sub-library of 150 compounds was tested for passive permeability using a version of the Parallel Artificial Membrane Permeability Assay (PAMPA) as previously reported¹³ with the addition of 0.2% polysorbate 80 in the donor well and 0.2% D- α -Tocopherol polyethylene glycol 1000 succinate (TPGS) to the acceptor well as “double-sink” conditions to mitigate against compound aggregation and adsorption to the apparatus²³. Tandem MS data were acquired for each sub-library, and the resulting spectra were analyzed using CycLS, with the peak identification process proceeding as in Figure 2. PAMPA and sequencing data were acquired with an identical LC method and were aligned by retention time to associate permeabilities and sequence identities. Various filters were applied to the merged data to ensure sequencing and PAMPA data quality. Where multiple peaks shared a top sequence – generally a result of a missing essential fragment – only the highest scoring peak was preserved, and the remainder were discarded. Finally, peaks with outlier permeabilities for their sub-library and retention time were manually curated by visual inspection of the appropriate extracted ion chromatogram to ensure correct automated integration and peak alignment between the donor and acceptor wells.

Impermeable compounds, heptamers with low acceptor well intensity, and a limited set of random peaks were inspected similarly (SI section 3.3).

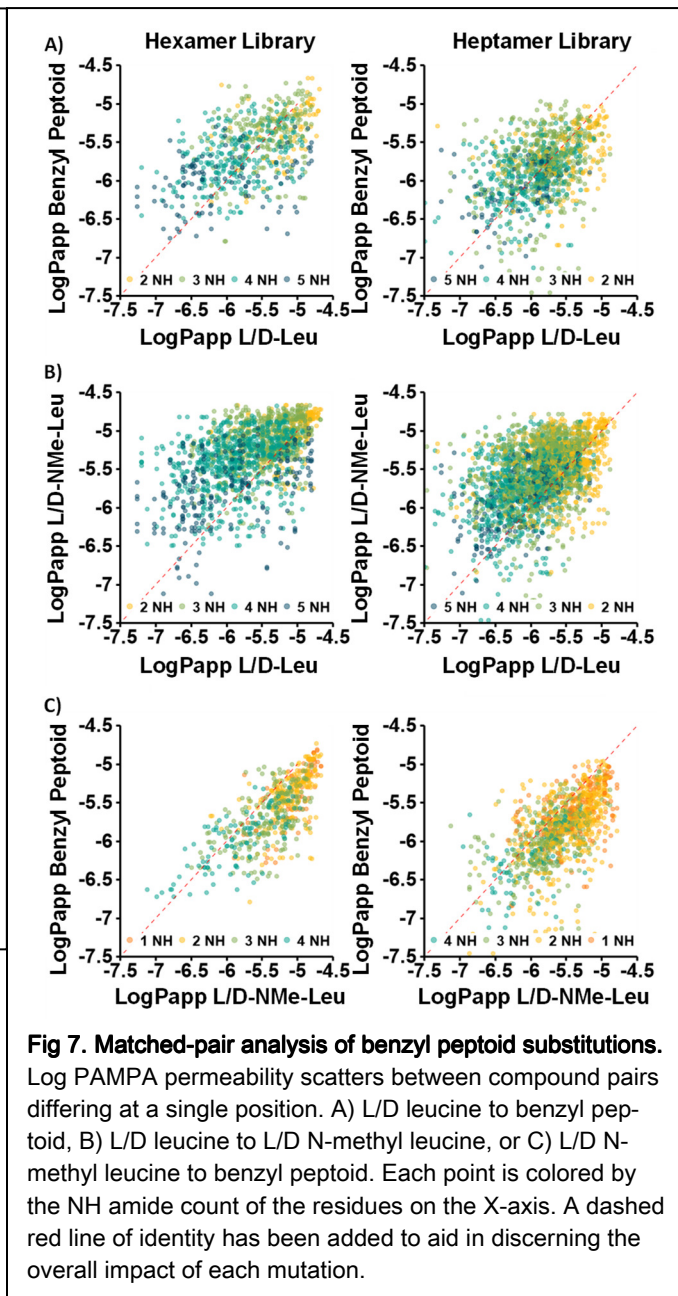
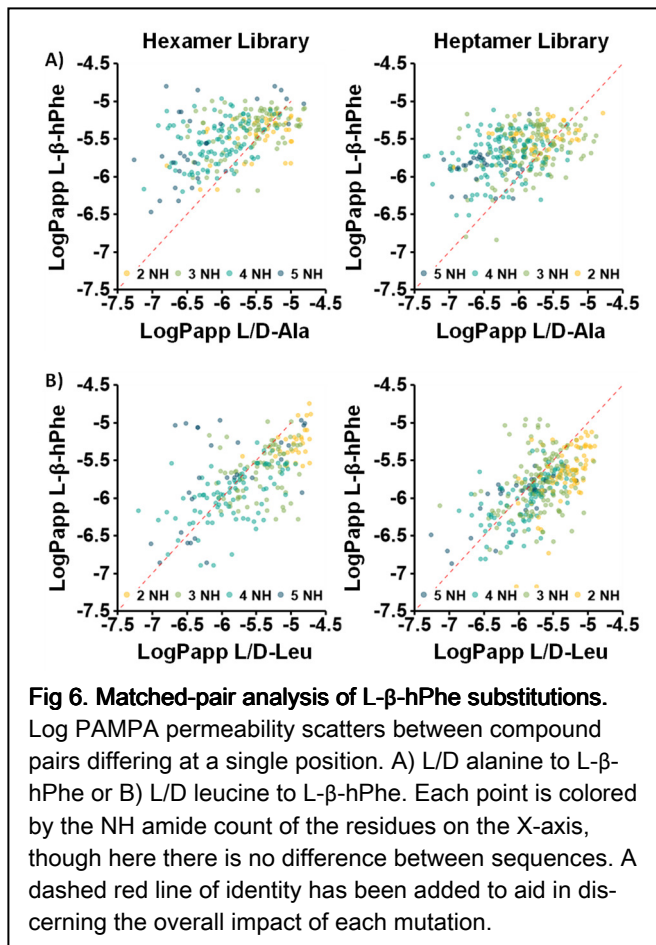
Unique sequences were obtained for 1063 of 1800 hexamers (59%) and 2023 of 3600 heptamers (56%), of which 823 hexamers and 1330 heptamers permeation rates greater than 1.0×10^{-6} cm/s, while 153 hexamers and 47 heptamers had permeation rates of 10×10^{-6} cm/s or greater (Fig. 3A). PAMPA permeability coefficients and HPLC retention times on a reverse-phase (C18) column (Fig. 3B) were correlated for both libraries as has been observed previously^{9, 11-13}. The majority of the 40 impermeable hexamers and 165 impermeable heptamers had relatively short HPLC retention times, in agreement with that correlation. Measured permeabilities ranged between 0.05×10^{-6} cm/s to 21.5×10^{-6} cm/s in the hexamer library and 0.01×10^{-6} cm/s to 16.4×10^{-6} cm/s in the heptamer library.

A simple compositional analysis was performed to demonstrate agreement with well-known trends in passive permeability. In general, removal of exposed hydrogen bond donors generally increases passive permeability in cyclic peptides^{8, 9, 17, 24} and other macrocycles¹⁴. We binned each library by amide NH count (Fig. 4A) and by N-methylated residue count (Fig. 4B) to obtain two different views of each library differing only by their treatment of peptoid residues. As expected, median log permeabilities decreased as amide NH count increased and increased as amide N-methyl count increased for both libraries. From the similar effect magnitudes of the two counts (NH and N-methyl), we can determine that the inclusion of peptoid residues in these libraries overall was neutral with respect to permeability. Additionally, the varying ability of individual compounds to form intramolecular hydrogen bonds can be observed through the range of permeabilities at each NH amide count, with a broad range of permeabilities observed for compounds with many NH amides and a more narrow range for those with few NH amides – though this effect was more subtle in the heptamer library.



PAMPA on complex mixtures has been noted to affect the permeability of the mixture components^{23, 25, 26}. To evaluate the severity of these differences and assess the utility of our data outside of bulk library analyses, we resynthesized and purified nine hexamers and ten heptamers and compared their pure PAMPA permeabilities with their mixture PAMPA permeabilities. Hexamer compounds 1, 2, 3, 4, 13, 14, and 16 (Fig. 5) were selected from the CycLS study for having sequenced correctly. Compounds 6, 10, and 11 were selected randomly among all compounds with high sequencing confidence and non-zero permeability from sub-libraries with a propyl-peptoid residue. The remaining compounds were chosen arbitrarily among all heptamer sub-libraries for high sequencing confidence with a bias toward high permeability. Synthesis was performed similarly to the library synthesis but without isotopic labeling of L-Leu. Resynthesized compound identities were confirmed through manual inspection of tandem MS data for a similar fragmentation profile and matching retention times when run on identical chromatographic methods (SI section 4.2).

PAMPA was performed separately on each pure, resynthesized compound in quadruplicate and the results compared to their library permeabilities in Figure 5. Compound 1 could not be displayed on the log scale because it was impermeable both in the library and as a pure compound, resulting in 18 contributing data points from 19 compounds. Library and resynthesized permeability values correlated well (R-squared of 0.88), but differed by up to 0.36 log units or 2.29-fold (average 0.14 or 1.39-fold) in either direction for data above 0.5×10^{-6} cm/s, below which the measurement deviation of the pure compounds increased. This mixture effect was not found to be concentration-dependent under non-sink conditions, which, lacking sinks in both wells, should be more vulnerable to such effects (SI section 5.1).



Matched-Pair Analysis

To best take advantage of the individual permeabilities of thousands of structurally diverse cyclic peptides, we performed an analysis of compound pairs differing only at a single residue. In most cases such comparisons can have their differences in permeability attributed partly to lipophilicity and partly to their impact on a peptide's conformational dynamics. While the intrinsic permeability of these scaffolds is not directly assessed by PAMPA, the “double-sink” conditions reduce the impact of water solubility on permeability, producing a closer estimate of intrinsic permeability than non-sink PAMPA. Differences in monomer lipophilicity were controlled by the library design except at position Y (Fig. 1), where the

N-propyl peptoid and beta-homo-Phe residues contribute much more to overall lipophilicity than the Ala and (NMe)Ala residues. Figures 6 and 7 are colored by the NH amide count of the X-axis to illuminate any differences in behavior originating from lipophilicity. In comparisons where the position of the matched pair of residues could vary, such as Figure 7, alternate scatterplots colored by pairing position were generated (SI section 5.2). Of particular interest were the effect of stereochemical inversions, peptoid residues, and beta residues on permeability.

Stereochemical Inversion

Single stereochemical inversions have been observed to have dramatic effects on the passive permeability of individual compounds, and we observe differences in permeability of over 70-fold in both libraries. However, we observed no systematic preference for either D- or L-stereochemistry greater than an average of 0.14 log units at any position, highlighting the context dependence of stereochemistry on conformation (and therefore, permeability). Stereochemical inversion had the greatest impact on the variable position located at $i - 1$ with respect to the static L-Phe residue, demonstrating preferences for D-Leu over L-Leu (0.14 log units hexa- and 0.12 log units heptamer) and L-(NMe)Leu over D-(NMe)Leu (0.11 log units hexa- and 0.09 log units heptamer). The adjacent proline ($i - 2$ with respect to the L-Phe) demonstrated no stereopreference in the hexamer library and only a small (0.08 log units) preference for D-Pro in the heptamer library, leading us to conclude that the static L-Phe templates the stereopreference of its $i - 1$ residue. No similar influence was observed $i + 1$ to the static L-Phe residue and all other positional trends averaged lower than 0.1 log units (SI section 5.2). This matched pair analysis revealed many "permeability cliffs" (represented by points that lie far off the diagonal) in which a single stereochemical inversion results in a large permeability increase or decrease. In addition, there are many points close to the diagonal representing positions which are relatively tolerant to stereoinversion.

Beta Residues

Having established that there is no systematic relationship between stereochemistry and permeability, we sought to understand the impact of β -residues on passive permeability by examining compound pairs in which L- β -hPhe replaces non-N-alkyl residues of either stereochemistry. L- β -hPhe appeared at only two positions per library and was therefore examined per position as a replacement for L/D-Ala at the Y position (Fig. 6A) and L/D-Leu at position X³ in the hexamer library and X⁴ in the heptamer library (Fig. 6B). One might expect a higher permeability from compounds including L- β -hPhe residues based on their greatly increased lipophilicity compared to Ala and Leu, especially under sink conditions where water solubility is less relevant. The large increase in permeability observed from the Ala to L- β -hPhe substitution

accorded with that expectation, however the Leu to L- β -hPhe substitution resulted in a decrease in permeability instead of the smaller, positive impact expected from a solely lipophilic perspective. Therefore, the L- β -hPhe residue must have a structural impact on passive permeability in one or both positions.

One major structural difference between Leu and L- β -hPhe is an increase in backbone flexibility. Given that the Leu to L- β -hPhe substitution is $i-1$ to a proline residue in all cases, the observed decrease in permeability may be a result of increased flexibility disrupting a proline-templated β -turn conformation hosting intramolecular hydrogen bonds. This is surely not true of all affected compounds, however. Another hypothesis is that the flexibility of L- β -hPhe may impair permeability at the Y position as well, but that the greatly increased lipophilicity of L- β -hPhe compared to Ala or the increased steric occlusion of the hPhe sidechain results in a net positive impact on permeability. Though our sampling of beta residues is limited, the increased flexibility of L- β -hPhe appears to be detrimental in at least the Leu to L- β -hPhe substitution.

Peptoid Residues

Peptoid residues have great utility for introducing chemical diversity not easily available in amino acid form into combinatorial libraries, and we therefore sought to explore their impact on passive permeability in detail. Here we focus on Leu or (NMe)Leu to benzyl peptoid substitutions, which were thoroughly sampled by these libraries. Benzyl peptoid was chosen for the X-residue position because of its lipophilic similarity to (NMe)Leu, simplifying this analysis. Removal of a hydrogen bond donor by substituting benzyl peptoid for L- or D-Leu (Fig. 7A) resulted in slightly increased permeability in both libraries. This substitution is overall positive only because the $i-1$ position to the static L-Phe has a permeability preference for N-alkylation; the other positions are neutral (SI section 5.2). By comparison, the Leu to (NMe)Leu substitution (Fig. 7B) increased permeability greatly at all positions in both libraries. Consistent with the previous comparisons, the (NMe)Leu to benzyl peptoid substitution (Fig. 7C) decreased permeability equally at all positions in both libraries. Neither the difference in lipophilicity between the two residues nor the position of the substitution accounts for this decrease, leading us to believe their difference in permeability originates from the structural impact of peptoid substitutions, perhaps arising from the comparative increase in backbone flexibility. Although previous work from our group has shown that strategic N-Me-to-peptoid substitutions can preserve and even enhance the permeability of individual cyclic peptide scaffolds^{12, 18, 25}, the present study provides a more comprehensive picture of the effect of peptoid substitutions across a wide variety of different backbone geometries. Thus, although substitution of N-Me amino acids for peptoids tends to

decrease permeability on average, numerous individual scaffolds still achieve high permeability, maintaining peptoids as an attractive option for easy functional diversification for scaffolds that tolerate them.

The negative effect on permeability of peptoid and beta residues suggests that additional flexibility is generally unfavorable in these libraries. The proline residue(s) in each library may contribute to this behavior by templating a double β -turn conformation for maximal intramolecular hydrogen bonding that increased flexibility disrupts. While rigidity can be a successful strategy for permeability in hexa- and heptamer cyclic peptides (e.g. 1NMe3¹⁰), large passively permeable scaffolds such as cyclosporin A²⁷ are required to leverage flexibility in a chameleonic behavior²⁸ to maintain both water solubility and passive permeability simultaneously. A requirement for flexibility, which we hypothesize to be detrimental to passive permeability for most scaffolds, may explain the rarity of large permeable scaffolds.

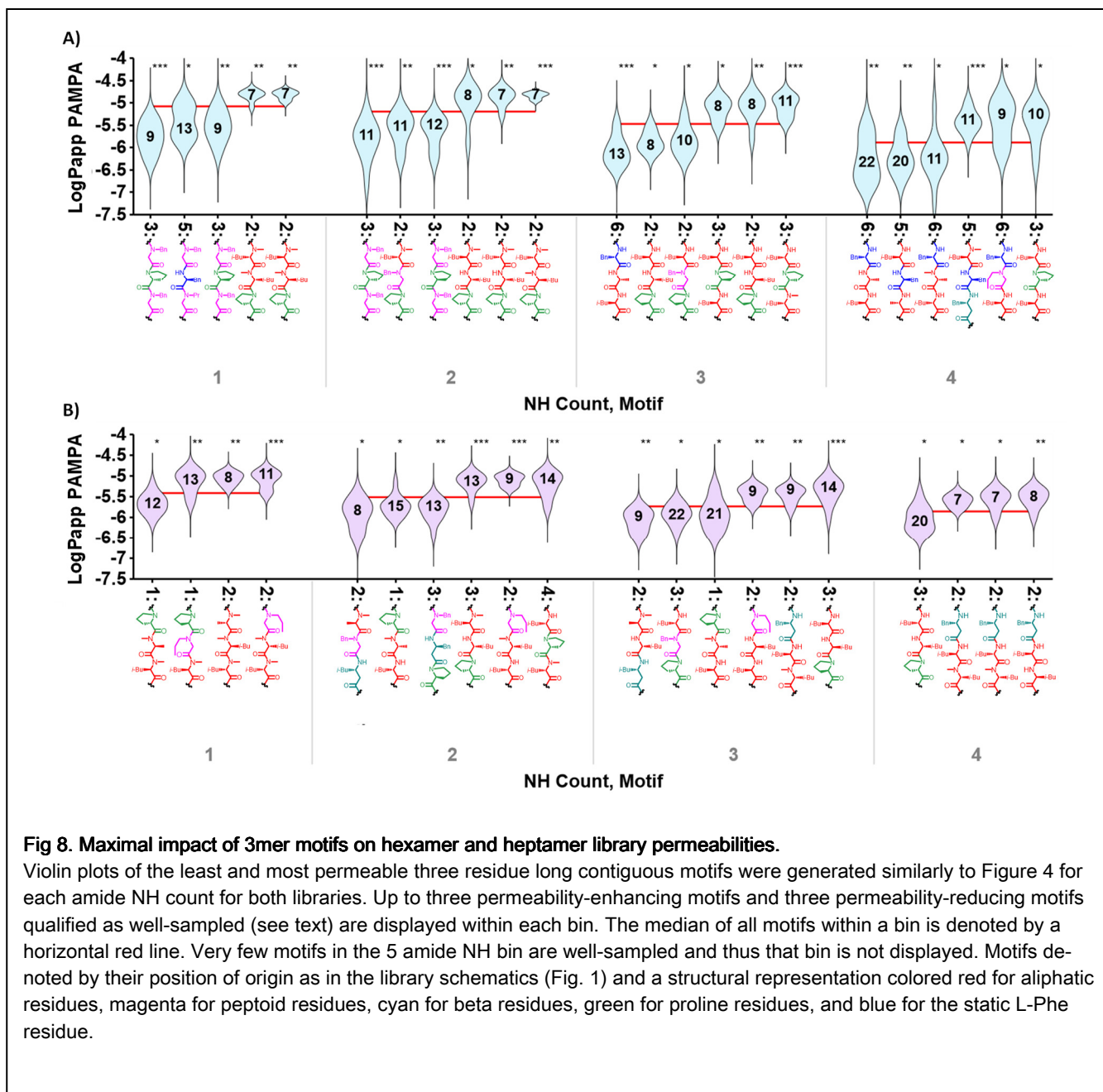


Fig 8. Maximal impact of 3mer motifs on hexamer and heptamer library permeabilities.

Violin plots of the least and most permeable three residue long contiguous motifs were generated similarly to Figure 4 for each amide NH count for both libraries. Up to three permeability-enhancing motifs and three permeability-reducing motifs qualified as well-sampled (see text) are displayed within each bin. The median of all motifs within a bin is denoted by a horizontal red line. Very few motifs in the 5 amide NH bin are well-sampled and thus that bin is not displayed. Motifs denoted by their position of origin as in the library schematics (Fig. 1) and a structural representation colored red for aliphatic residues, magenta for peptoid residues, cyan for beta residues, green for proline residues, and blue for the static L-Phe residue.

Passive Permeability Motifs

Technologies such as in-vitro translation⁴ and light-directed peptide array synthesis²⁹ can generate cyclic peptide libraries biased towards passive permeability by including only permeable backbone geometries, but other methods of library synthesis lack total synthetic control and cannot take the same approach without severely limiting their size. We therefore sought an alternative method of leveraging our permeability data for the benefit of such combinatorial library generation techniques by identifying permeability-enhancing substructures or “permeability motifs”. These permeability-enhancing motifs could be incorporated as static design elements while allowing the rest of the cycle to vary, freeing the

medicinal chemist to incorporate sidechain diversity while maintaining scaffolds statistically likely to maintain permeable geometries. Thus, we sought to address the prevalence of motifs with significant effects on permeability, their maximal impact on permeability, and their scope of applicability.

Although permeability motifs may be inferred from this data in a variety of patterns, we have chosen to evaluate contiguous 3-residue sequences generated from all possible such sequences at each position of a library's schematic in a position-specific manner. Each motif has many associated compound permeabilities, and each library compound is associated to a number of motifs equal to its length. A motif length of three was chosen as a useful length for both hexamers and heptamers while allowing significant variation in the remainder of the peptide. For all inter-motif comparisons, we found it necessary to segregate the compounds associated with each motif by NH amide count to ensure that simple lipophilicity did not drive comparisons of motif impact on permeability. This binning scheme introduces its own bias on compound membership in a given motif by restricting the N-alkylation of the remainder of the cycle, but nonetheless results in an approximation of intrinsic permeability (SI section 5.3).

Motif Prevalence

To determine the prevalence of motifs with a significant effect on permeability we used a Mann-Whitney U-test³⁰ between the log permeabilities of one motif vs. the population of all log permeabilities for all motifs within the same amide NH count. We found that 146 of 366 hexamer motifs detected and 197 of 391 heptamer motifs detected significantly affected permeability for one or more NH amide counts. As expected for conformational motifs derived from such a structurally diverse compound set, a roughly equal number of significant permeability-enhancing motifs were discovered (69 hexamers, 87 heptamers) as permeability-reducing motifs (75 hexamers, 106 heptamers). Statistical significance in a U-test alone, however, is not sufficient for a motif to be useful in practice, and we therefore implemented additional requirements for a motif to be considered well-characterized.

The ideal permeability-enhancing motif for biasing a combinatorically generated screening library contains many backbone geometries for better odds of target engagement and a greater potential library size. These factors in mind, we defined well-sampled motifs to be associated with a minimum of six compounds with known, non-zero permeabilities – which must be at least two-thirds of all theoretical library members containing that motif. These requirements resulted in a final total of 81 well-sampled permeability-enhancing motifs (25 hexamer, 56 heptamer) across all amide NH counts and 39 well-sampled permeability-reducing motifs (20 hexamer, 19 heptamer). Few permeability-reducing motifs were considered well-sampled due to the higher proportion of compounds with zero apparent permeability composing them

compared to permeability-enhancing motifs. Each of the 81 well-sampled permeability-enhancing motifs identified represents a set of backbone geometries potentially useful in biasing a combinatorically generated screening library towards passive permeability.

Motif Impact

To investigate the maximal effect of a well-sampled motif on passive permeability, up to three of the least and most permeable well-sampled motifs for each NH amide count were plotted in Figure 8. The same trends observed in the matched pair analysis were visible at some amide NH counts (low-permeability benzyl peptoid-containing motifs in hexamer bins 1 and 2, high-permeability L- β -hPhe-containing motifs where the L- β -hPhe would replace an Ala in heptamer bin 4) but did not dominate the motifs discovered. The most permeable motifs improved upon the population median permeability at that NH amide count by between 0.28 and 0.63 log units (1.9- and 4.3-fold) in the hexamer library and between 0.36 and 0.46 log units in the heptamer library (2.3- and 2.9-fold). While these benefits are smaller than picking only the most permeable backbone geometries, incorporating permeability-enhancing motifs into a library design also minimizes inclusion of backbone geometries with especially poor permeability. The least permeable motifs generally had a similar level of impact, ranging between 0.43 and 0.68 log units (2.7- and 4.8-fold) lower permeability than the population median in the hexamer library and 0.24 to 0.32 log units (1.7- and 2.1-fold) lower in the heptamer library. Reducing the requirements on well-sampled motifs had little effect on the maximal impact of permeability-enhancing motifs but increased the maximal impact of permeability-reducing motifs greatly (SI section 5.3). From this data we conclude that permeability motifs can have a sufficiently large impact to be of utility in biasing combinatorically generated libraries toward passive permeability even among this structurally diverse set of compounds.

Motif Scope of Applicability

To determine the potential generality of these motifs, we investigated the correlation between their effects on permeability in the hexamer vs. heptamer libraries. The previous positional definition for motifs made a direct comparison between the hexamer and heptamer libraries impossible, and so we created a position-independent motif naming scheme. This new motif naming scheme categorized residues as L-NH, L-N-Me, D-NH, D-N-Me, L-proline, D-proline, peptoid residue, or beta residue rather than use exact residue identities. While this analysis resulted in some cases of composite motifs with less extreme impacts on permeability, few motifs were impacted (SI section 5.3).

A linear regression of the permeability comparison between hexamer and heptamer libraries of all identically named motifs without any filtering resulted in an R-squared of 0.52. Although there was not a strong overall correlation between

the permeabilities of individual motifs in the hexamer and heptamer libraries, some permeability-determining motifs were common to both ring sizes. For example, the D-Leu – L-Leu – D-Pro motif was highly favorable among compounds containing 3 backbone NH amides across both the hexamer and heptamer libraries while all stereochemical variants of the Leu – Bnz – Pro motif were among the least permeable. Similarly, the D-NMe-Leu – D-Leu – D-Pro motif was highly favorable among compounds containing 2 NH amides in both libraries. While instances of agreement do exist between libraries, most permeability-affecting motifs did not correlate, leading us to conclude that their utility is limited to libraries of the same ring size and possibly the same proline placement. Despite these restrictions of scope, the conformational nature of a permeability-enhancing motif should allow these motifs to maintain utility in the presence of a variety of alternative sidechains.

Conclusions

To increase our understanding of passive permeation in cyclic peptides and to aid in biasing screening libraries composed of cyclic peptides towards passively permeable hits, we synthesized hexamer and heptamer libraries with diverse backbone geometries and quantified their PAMPA permeabilities as mixtures by MSMS. Agreement between the pure PAMPA permeabilities of 19 resynthesized compounds with the corresponding mixture-based library peaks validated that individual compounds can be extracted from our library analyses. Our results confirm our previous understanding of the impact of free amide NHs and N-methylation across thousands of compounds, elucidate the effects of single backbone modifications, and introduce the concept of permeability-enhancing motifs.

Our matched-pair analysis of backbone modifications showed that although single stereochemical inversions do not generally control the passive permeability of a set of geometrically diverse cyclic peptides, dramatic “permeability cliffs” were not uncommon in individual cases. Although the matched pair analysis of L- β -hPhe was complicated by differences in lipophilicity, sidechain volume, and proximity to proline residues, we hypothesize that its increased flexibility had a generally negative impact on permeability. The results of our matched pair analysis of benzyl peptoid led us to a similar conclusion, with the negative effect of replacing an N-methyl residue especially telling. Our data suggests that for hexa- and heptapeptides, in most cases, the entropic cost of populating a non-polar conformer for membrane transit is higher than the enthalpic gains over a more rigid scaffold. Although beta and peptoid residues are attractive for the structural and chemical diversity they provide, these findings suggest that their inclusion in libraries biased towards passive

permeability should be limited where possible or strategically incorporated via a permeability-enhancing motif that accommodates their presence.

Our investigation of passive permeation in relation to 3-mer motifs revealed that “permeability motifs” and “anti-permeability motifs” exist even among compound sets with diverse backbone geometries and few shared structural elements. We found dozens of permeability-enhancing motifs sampled well enough to be of practical use in biasing combinatorically generated libraries towards passive permeability, the best of which increased permeability 2-fold to 4-fold compared to all other compounds with the same number of hydrogen bond donors. We find this effect size remarkable for libraries with such a high structural diversity. Although few motifs were shared between the hexa- and heptamer libraries, permeability-enhancing motifs likely exist for larger ring sizes as well. These observations demonstrate the presence and potential impact of permeability-enhancing motifs on passive permeability, yielding insights into the structural scope of their applicability. The permeability-enhancing motifs discovered expand the utility of our library data to combinatorically generated screening libraries in addition to those technologies with full synthetic control by offering sets of interrelated backbone geometries with a net positive impact on passive permeability.

We have demonstrated that individual library members can be extracted from geometrically diverse libraries with a good correlation to their permeability in mixtures. Scaffolds with particularly high permeabilities may be useful to bias screening technologies with “privileged” backbones expected to accommodate diverse sidechains in the search for activity. With the workflow herein performed as a scaffold permeability screen prior to designing an activity screening library, subsequent hits would, in principle, require less optimization to achieve passive permeability and move towards oral bioavailability. The permeability-enhancing motifs that we identified extend the utility of our permeability data to library generation technologies without complete synthetic control and to those seeking to incorporate backbone diversity. Though we present only hexa- and heptapeptide/peptoid scaffolds herein, this approach can identify privileged backbone geometries in any modular, synthetically accessible amide-bonded system, and we are currently using it to explore larger macrocycles.

Associated Content

Supporting information includes synthetic and analytical methods, a summary of in-house python scripts used and links to their GitHub pages, characterization of sub-libraries and resynthesized compounds, and additional discussion.

Spreadsheets have been provided containing all the curated library data used for this analysis and a summary of all motifs identified.

Author Information

Corresponding Author: R. Scott Lokey, slokey@ucsc.edu. Chad Townsend, Eva Jason, and Quinn Edmondson synthesized all libraries. Chad Townsend and Eva Jason synthesized all individual compounds. Chad Townsend, Matthew Naylor, Cameron Pye, and Joshua Schwochert developed PAMPA and tandem mass spectrometry LCMS methods. Chad Townsend wrote in-house python scripts assisted by Eva Jason and with advice from Cameron Pye and Joshua Schwochert. Chad Townsend performed all data analysis. Chad Townsend and Scott Lokey wrote the manuscript.

Acknowledgements

We would like to acknowledge Colin Kelly and Victoria G. Klein for providing valuable feedback throughout this project. We gratefully acknowledge NIH grant GM131135 for funding.

References

1. Villar, E. A.; Beglov, D.; Chennamadhavuni, S.; Porco, J. A., Jr.; Kozakov, D.; Vajda, S.; Whitty, A., How proteins bind macrocycles. *Nat Chem Biol* **2014**, *10* (9), 723-31.
2. Nielsen, D. S.; Shepherd, N. E.; Xu, W.; Lucke, A. J.; Stoermer, M. J.; Fairlie, D. P., Orally Absorbed Cyclic Peptides. *Chem Rev* **2017**, *117* (12), 8094-8128.
3. Zorzi, A.; Deyle, K.; Heinis, C., Cyclic peptide therapeutics: past, present and future. *Curr Opin Chem Biol* **2017**, *38*, 24-29.
4. Passioura, T.; Suga, H., A RaPID way to discover nonstandard macrocyclic peptide modulators of drug targets. *Chem Commun (Camb)* **2017**, *53* (12), 1931-1940.
5. Franzini, R. M.; Neri, D.; Scheuermann, J., DNA-Encoded Chemical Libraries: Advancing beyond Conventional Small-Molecule Libraries. *Acc Chem Res* **2014**, *47* (4), 1247-1255.
6. Liras, S.; McClure, K. F., Permeability of Cyclic Peptide Macrocyces and Cyclotides and Their Potential as Therapeutics. *ACS Med Chem Lett* **2019**, *10* (7), 1026-1032.
7. White, A. M.; Craik, D. J., Discovery and optimization of peptide macrocycles. *Expert Opin Drug Discov* **2016**, *11* (12), 1151-1163.
8. Wang, C. K.; Northfield, S. E.; Colless, B.; Chaousis, S.; Hamernig, I.; Lohman, R. J.; Nielsen, D. S.; Schroeder, C. I.; Liras, S.; Price, D. A.; Fairlie, D. P.; Craik, D. J., Rational design and synthesis of an orally bioavailable peptide guided by NMR amide temperature coefficients. *Proc Natl Acad Sci U S A* **2014**, *111* (49), 17504-9.
9. Wang, C. K.; Northfield, S. E.; Swedberg, J. E.; Colless, B.; Chaousis, S.; Price, D. A.; Liras, S.; Craik, D. J., Exploring experimental and computational markers of cyclic peptides: Charting islands of permeability. *Eur J Med Chem* **2015**, *97*, 202-13.
10. Rezai, 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* **2006**, *128* (8), 2510-1.
11. Hewitt, W. M.; Leung, S. S.; Pye, C. R.; Ponkey, A. R.; Bednarek, M.; Jacobson, M. P.; Lokey, R. S., Cell-permeable cyclic peptides from synthetic libraries inspired by natural products. *J Am Chem Soc* **2015**, *137* (2), 715-21.

12. Furukawa, A.; Townsend, C. E.; Schwochert, J.; Pye, C. R.; Bednarek, M. A.; Lokey, R. S., Passive Membrane Permeability in Cyclic Peptomer Scaffolds Is Robust to Extensive Variation in Side Chain Functionality and Backbone Geometry. *J Med Chem* **2016**, *59* (20), 9503-9512.
13. Naylor, M. R.; Ly, A. M.; Handford, M. J.; Ramos, D. P.; Pye, C. R.; Furukawa, A.; Klein, V. G.; Noland, R. P.; Edmondson, Q.; Turmon, A. C.; Hewitt, W. M.; Schwochert, J.; Townsend, C. E.; Kelly, C. N.; Blanco, M. J.; Lokey, R. S., Lipophilic Permeability Efficiency Reconciles the Opposing Roles of Lipophilicity in Membrane Permeability and Aqueous Solubility. *J Med Chem* **2018**, *61* (24), 11169-11182.
14. Over, B.; Matsson, P.; Tyrchan, C.; Artursson, P.; Doak, B. C.; Foley, M. A.; Hilgendorf, C.; Johnston, S. E.; Lee, M. D. t.; Lewis, R. J.; McCarren, P.; Muncipinto, G.; Norinder, U.; Perry, M. W.; Duvall, J. R.; Kihlberg, J., Structural and conformational determinants of macrocycle cell permeability. *Nat Chem Biol* **2016**, *12* (12), 1065-1074.
15. Townsend, C.; Furukawa, A.; Schwochert, J.; Pye, C. R.; Edmondson, Q.; Lokey, R. S., CycLS: Accurate, whole-library sequencing of cyclic peptides using tandem mass spectrometry. *Bioorg Med Chem* **2018**, *26* (6), 1232-1238.
16. Rader, A. F. B.; Reichart, F.; Weinmuller, M.; Kessler, H., Improving oral bioavailability of cyclic peptides by N-methylation. *Bioorg Med Chem* **2018**, *26* (10), 2766-2773.
17. White, T. R.; Renzelman, C. M.; Rand, A. C.; Rezai, T.; McEwen, C. M.; Gelev, V. M.; Turner, R. A.; Linington, R. G.; Leung, S. S.; Kalgutkar, A. S.; Bauman, J. N.; Zhang, Y.; Liras, S.; Price, D. A.; Mathiowetz, A. M.; Jacobson, M. P.; Lokey, R. S., On-resin N-methylation of cyclic peptides for discovery of orally bioavailable scaffolds. *Nat Chem Biol* **2011**, *7* (11), 810-7.
18. Schwochert, J.; Turner, R.; Thang, M.; Berkeley, R. F.; Ponkey, A. R.; Rodriguez, K. M.; Leung, S. S.; Khunte, B.; Goetz, G.; Limberakis, C.; Kalgutkar, A. S.; Eng, H.; Shapiro, M. J.; Mathiowetz, A. M.; Price, D. A.; Liras, S.; Jacobson, M. P.; Lokey, R. S., Peptide to Peptoid Substitutions Increase Cell Permeability in Cyclic Hexapeptides. *Org Lett* **2015**, *17* (12), 2928-31.
19. Goetz, G. H.; Philippe, L.; Shapiro, M. J., EPSA: A Novel Supercritical Fluid Chromatography Technique Enabling the Design of Permeable Cyclic Peptides. *ACS Med Chem Lett* **2014**, *5* (10), 1167-72.
20. Guixer, B.; Arroyo, X.; Belda, I.; Sabido, E.; Teixido, M.; Giral, E., Chemically synthesized peptide libraries as a new source of BBB shuttles. Use of mass spectrometry for peptide identification. *J Pept Sci* **2016**, *22* (9), 577-91.
21. Ciudad, S.; Bayó-Puxán, N.; Varese, M.; Seco, J.; Teixidó, M.; García, J.; Giral, E., 'À La Carte' Cyclic Hexapeptides: Fine Tuning Conformational Diversity while Preserving the Peptide Scaffold. *ChemistrySelect* **2018**, *3* (8), 2343-2351.
22. Zuckermann, R. N.; Kerr, J. M.; Kent, S. B. H.; Moos, W. H., Efficient method for the preparation of peptoids [oligo(N-substituted glycines)] by submonomer solid-phase synthesis. *J Am Chem Soc* **1992**, *114* (26), 10646-10647.
23. Oh, M. H.; Lee, H. J.; Jo, S. H.; Park, B. B.; Park, S. B.; Kim, E. Y.; Zhou, Y.; Jeon, Y. H.; Lee, K., Development of Cassette PAMPA for Permeability Screening. *Biol Pharm Bull* **2017**, *40* (4), 419-424.
24. Fouche, M.; Schafer, M.; Berghausen, J.; Desrayaud, S.; Blatter, M.; Piechon, P.; Dix, I.; Martin Garcia, A.; Roth, H. J., Design and Development of a Cyclic Decapeptide Scaffold with Suitable Properties for Bioavailability and Oral Exposure. *ChemMedChem* **2016**, *11* (10), 1048-59.
25. Furukawa, A.; Schwochert, J.; Pye, C. R.; Asano, D.; Edmondson, Q. D.; Turmon, A.; Klein, V.; Ono, S.; Okada, O.; Lokey, R. S., Drug-like properties in macrocycles above MW 1000: Backbone rigidity vs. side-chain lipophilicity. *Angew Chem Int Ed Engl* **2020**.
26. Zhang, J.; Maloney, J.; Drexler, D. M.; Cai, X.; Stewart, J.; Mayer, C.; Herbst, J.; Weller, H.; Shou, W. Z., Cassette incubation followed by bioanalysis using high-resolution MS for in vitro ADME screening assays. *Bioanalysis* **2012**, *4* (5), 581-93.
27. el Tayar, N.; Mark, A. E.; Vallat, P.; Brunne, R. M.; Testa, B.; van Gunsteren, W. F., Solvent-dependent conformation and hydrogen-bonding capacity of cyclosporin A: evidence from partition coefficients and molecular dynamics simulations. *J Med Chem* **1993**, *36* (24), 3757-64.

28. Whitty, A.; Zhong, M.; Viarengo, L.; Beglov, D.; Hall, D. R.; Vajda, S., Quantifying the chameleonic properties of macrocycles and other high-molecular-weight drugs. *Drug Discov Today* **2016**, *21* (5), 712-7.
29. Bhushan, K. R., Light-directed maskless synthesis of peptide arrays using photolabile amino acid monomers. *Org Biomol Chem* **2006**, *4* (10), 1857-9.
30. Mann, H. B.; Whitney, D. R., On a Test of Whether one of Two Random Variables is Stochastically Larger than the Other. *Ann. Math. Statist.* **1947**, *18* (1), 50-60.
-