How to cite: *Angew. Chem. Int. Ed.* **2023**, *62*, e202307092
doi.org/10.1002/anie.202307092

Macrocycles

Streamlined Protein-Protein Interface Loop Mimicry

*Tianxiong Mi, Siriwalee Siriwibool, and Kevin Burgess**

Abstract: Cyclic peptides comprising endocyclic organic fragments, *"cyclo*-organopeptides", can be probes for perturbing protein-protein interactions (PPIs). Finding loop mimics is difficult because of high conformational variability amongst targets. Backbone Matching (BM), introduced here, helps solve this problem in the illustrative cases by facilitating efficient evaluation of virtual *cyclo*-organopeptide core-structure libraries. Thus, 86 rigid organic fragments were selected to build a library of 602 *cyclo*-organopeptides comprising Ala and organic parts: *"cyclo*-{-$(Ala)_n$-organo-}". The central hypothesis is "hit" library members have accessible low energy conformers corresponding to backbone structures of target protein loops, while library members which *cannot* attain this conformation are probably unworthy of further evaluation. BM thereby prioritizes candidate loop mimics, so that less than 10 *cyclo*-organopeptides are needed to be prepared to find leads for two illustrative PPIs: iNOS·SPSB2, and uPA·uPAR.

Introduction

PPIs are important in cell signaling and as pharmaceutical targets.[1] Their interaction energies are dictated by peptidic hot segments,[2] hence peptides can provide sufficient pharmacophores to recapitulate key interface regions. Observable binding usually requires constrained peptides, so organic parts are often incorporated to make them less flexible, and reduce entropy loss on protein receptor binding. Thus "*cyclo*-organopeptides" (Figure 1a left, and 1b) are privileged, low molecular mass chemotypes to perturb PPIs.[3]

Loops occur approximately twice as frequently as helices or sheets at PPI interfaces.[4] However, unlike helices and sheets which can be mimicked by uniform structures, such as hydrocarbon staples[5] and hydrogen bond surrogates,[6] there are no generalized models for loop mimicry because their secondary structures are highly variable (Figure S1). Consequently, loop mimics "tailor made" for each target must be identified.

This paper outlines "Backbone Matching (BM)" to virtually identify loop mimics from user-defined libraries of *cyclo*-organopeptides. BM is designed to reduce wasted synthesis and testing resources during lead discovery for PPI targets. It is based on the following analysis of loop binding processes.

cyclo-Organopeptides can only bind the binding pocket of the targeted protein receptor if their backbones can access conformations similar to the natural protein ligand loop (Figure 1a right) in the PPI. Loop mimic backbones which cannot access these conformers in solution probably cannot fit the receptor pocket (Figure 1a left).

cyclo-Organopeptide conformations are determined by: (i) number of amino acids incorporated; (ii) physiochemical properties of the organic part; and, (iii) structures of the amino acids used. In conceiving BM we assumed conformational backbone components {*ie* (i) and (ii)} can be represented using all alanine systems, and simulated via conformational sampling in solution (Figure 1c). These *cyclo*-{-$(Ala)_n$-organo-} systems resemble backbone conformers because alanine is the simplest amino acid with a *Cα–Cβ* vector. Alanine residues induce backbone conformations unbiased by side-chain to side-chain interactions; this is important because BM is designed to exclude influences of side-chain functionalities until the end of the process when the *cyclo*-organopeptide is considered in the relevant environment: the protein receptor cavity. Temporarily applying that assumption means factor (iii) is initially held constant.

To act upon the assumption above, *cyclo*-{-$(Ala)_n$-organo-} fragment-, ring size- combinations are evaluated by automated overlays with the PPI solid state structure (Figure 1d). A few *cyclo*-{-$(Ala)_n$-organo-} systems may have conformers suitable for binding, *ie* similar to the protein ligand loop. They should be thermodynamically accessible in the receptor bound form, but not necessarily the global minima simulated in solution. These do not have appropriate side chain pharmacophores to strongly bind the receptor, so factor (iii) above is reintroduced. Thus overlay of hit *cyclo*-{-$(Ala)_n$-organo-} conformations on the bound protein ligand loop reveals which amino acid side chains might be added to the *cyclo*-{-$(Ala)_n$-organo-} backbone to enhance binding (Figure 1e). BM thereby evaluates backbone conformations of all-Ala systems using modest computational resources, then overlays them on the protein ligand loop to

[*] T. Mi, Prof. Dr. K. Burgess
Department of Chemistry, Texas A & M University
77842 College Station TX (USA)
E-mail: burgess@tamu.edu
S. Siriwibool
School of Chemistry, Institute of Science, Suranaree University of Technology
30000 Nakhon Ratchasima (Thailand)

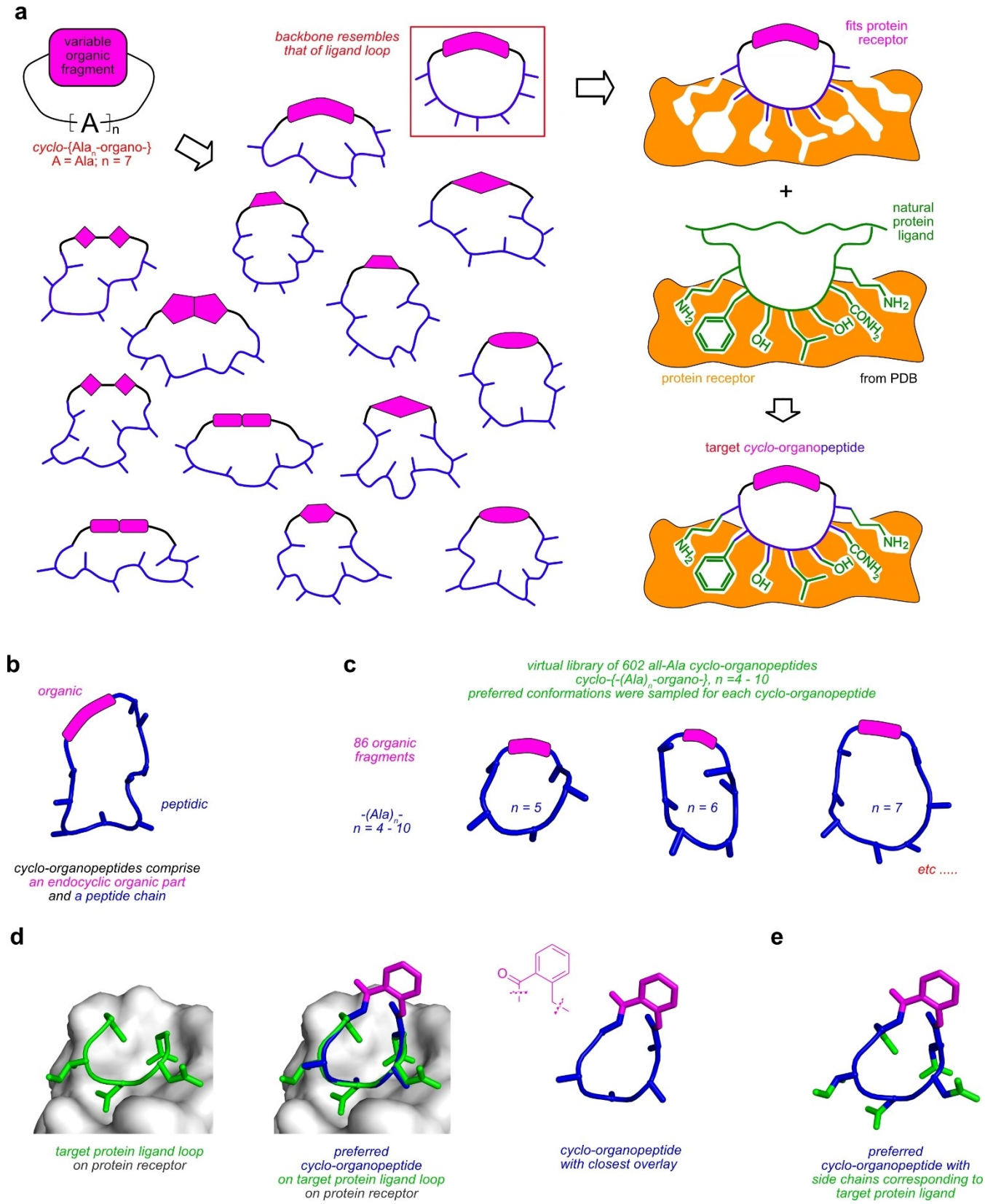

Figure 1. BM workflow. **a** A virtual process of *cyclo*-{-(Ala)$_n$-organo-} systems competing with target loops for receptor pockets. **b** The definition of *cyclo*-organopeptides. **c** A virtual library of preferred core conformations for *cyclo*-{-(Ala)$_n$-organo-} systems (n = 4–10). **d** Identification of the preferred conformer that best overlays protein ligand loop. **e** Preferred *cyclo*-organopeptides are synthesized with corresponding side chains.

deduce *cyclo*-{-(Ala)$_n$-organo-} hits and appropriate side chains to use for enhanced protein receptor binding. BM-generated leads can be computationally validated using high level docking simulations. Hits passing this test are assumed to be worthy of synthesis and testing.

Figure 1 illustrates BM featuring 86 organic parts and ring sizes incorporating 4–10 Ala's; this requires conformational simulations for a library of 86×7 = 602 *cyclo*-{-(Ala)$_n$-organo-} systems. Thus 602 simulations can identify accessible conformations (below a specified kcal mol^{-1} cut off from the global minimum) giving a *Core Conformation Database* which contains backbone orientations and projections of side chains represented by the Ala *Cα–Cβ* vectors. *Core Conformation Databases* only need to be assembled once for each library; they can be used repeatedly on any PPI.

Results and Discussion

Selection of Organic Fragments and Generation of A Virtual Library of *cyclo*-{-(Ala)$_n$-organo-}

An illustrative library of *cyclo*-{-(Ala)$_n$-organo-} was built using a set of rigid, distinct *non-canonical* amino acids as "organic fragments", mined from the ZINC database and filtered for diversity, availability, ease of incorporation into cyclic systems, and ability to impart rigidity (Figure 2a). Thus ZINC was filtered for commercial availability AND molecular weight < 300 Da, then for ONLY one carboxylic acid AND ONLY one primary or secondary amine, giving 27,006 amino acids. Flexible organic fragments were excluded via the following procedure. The collection was clustered into subsets of similar core structures (using the open source package RDkit); this drastically reduced the set to 712. Next, an in house *rigidity filter* script was applied wherein 50 random geometrically optimized conformers for each compound were generated. Only compounds with standard deviation of distances between the *OOC*- and the amine *N*- < 0.5 Å were retained. Analyses to test the rigidity filter are described in Supporting Information section E. Hits > $500/g (based on user SciFinder searching) were then removed, leaving 86 non-natural amino acids (Supporting Information section F).

Each organic fragment was virtually cyclized with an oligo-alanine sequence, -(Ala)$_n$- (n = 4–10), giving 602 (86*7) *cyclo*-{-(Ala)$_n$-organo-} peptides. Peptide strings of four or more amino acids were chosen because shorter ones are geometrically difficult to cyclize. Conversely, determination of all representative conformers for n > 10 is difficult because they tend to be too flexible. This is not a serious limitation because most hot loops feature 4–10 amino acids.[4,7]

Simulations of Preferred Core Conformations

Several conformational sampling methods were evaluated for the molecular dynamics (MD) studies and enhanced-

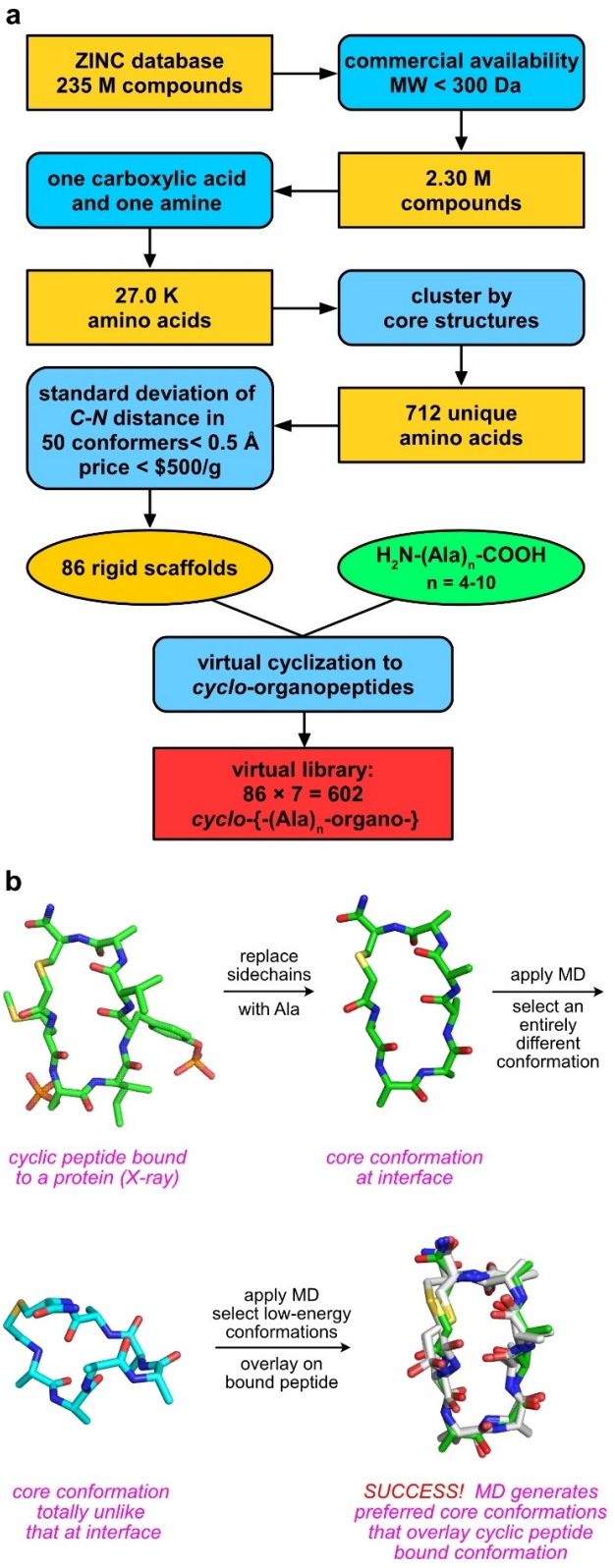

Figure 2. Generation of the *Core Conformation Database*. **a** Generation of a virtual library of *cyclo*-{-(Ala)$_n$-organo-}. **b** The conformational sampling method validated by regenerating known cyclic peptide receptor-bound conformations from the most different conformational starting point, is used to simulate populated conformers of each *cyclo*-{-(Ala)$_n$-organo-} system.

MCMM (Monte Carlo Multiple Minimum) appeared to be effective at relatively low computational costs.[8] Consequently, we designed a stringent test of enhanced-MCMM for several different cyclic peptides (Figure 2b).

Protein Data Bank (PDB) coordinates from a cyclic peptide·protein receptor complex were collected for each cyclic peptide, then side chains were simplified to *-Me* to reveal all-Ala peptide core conformations. Enhanced-MCMM was applied, and a conformation most *dissimilar* to the bound conformation was selected. Using this as a starting point, enhanced-MCMM was applied again, and conformers within 15 kcal/mol from the lowest were clustered. Representative cluster members were examined to see if any overlaid with the original bound conformation. Identification of good overlays (RMSD, Root Mean Square Deviation, <1.0 Å) would prove the sampling method can regenerate bound core conformations of the native cyclic peptide from a drastically different starting point.

Enhanced-MCMM worked well on all but one case in the workflow above. The exception was a 14-mer cyclic peptide which was probably too flexible to be exhaustively sampled using the parameters provided. Given the similarity between regular cyclic peptides and *cyclo*-organopeptides, we believe enhanced-MCMM is a good choice for BM. Consequently, conformations of the virtual library of 602 *cyclo*-{-$(Ala)_n$-organo-} were simulated (Supporting Information section G).

Evaluation of cyclo-{-$(Ala)_n$-organo-} Core Conformations on Protein Ligand Loops

Two overlay procedures were devised to find *cyclo*-{-$(Ala)_n$-organo-} core conformations which overlaid protein interface loops with minimal RMSD devations. 'Precise alignment' only uses a sub-library where the length of *cyclo*-{-$(Ala)_n$-organo-} equals that of the target loop, and finds predicted best linkers among 86 fragments. For instance, a 6-mer loop will be screened by *cyclo*-{-$(Ala)_6$-organo-}; *n* is not varied (Figure 3a). Precise alignment works well for short loops which contain most of the protein ligand hot spots.[9]

The other algorithm, 'auto-slicing' (Figure S7) cuts loops into 4–10 residue fragments, then aligns them with *cyclo*-{-$(Ala)_n$-organo-} cores of same length. It is for larger loops with discontiguous hot spots. Auto-slicing searches for optimal *cyclo*-{-$(Ala)_n$-organo-} systems using the same number or less amino acids than those in the protein hot loop. Detailed descriptions of both algorithms are in Supporting Information section H, and source codes are on GitHub.

Kritzer and co-workers data mined the PDB to collect 1398 hot loops with 5 to 8 residues (Supporting Information section I);[4] these have diverse conformations (Figure S10 in Supporting Information for this paper, and Figure 2 of[4]). In our studies, each of those hot loops was overlaid using precise alignment (Figure 3b). Thus we found hits that overlaid conformers within RMSD 0.8 Å for 5 and 6-mer loops, and 1.0 for 7-, 8-mers, for each hot loop. Higher RMSD cut-offs relative to smaller loops are appropriate because larger ones can maintain suitable side chain orientations even as they flex.

BM found mimic candidates for 1245 of the 1398 hot loops screened indicating our library was sufficiently large and diverse to find hits for 89 % of those hot loops. Percentages of hot loops for which hits were identified as a function of each ring size are shown in Figure 3c. Interestingly, Kritzer discussed five promising hot loops which had not been studied before,[4] and BM found mimic candidates in our library for all of them (Figure 3d).

Two hot loops were selected to establish proof of principle for BM. One of these, iNOS·SPSB2, involves a short loop ideal for testing via the precise alignment algorithm. Overlaying via auto-slicing algorithm is more appropriate for large loops, such as the second: the uPA·uPAR interface loop.

Application of BM to iNOS·SPSB2

Inducible nitric oxide synthase (iNOS) produces nitric oxide and related reactive species which can kill intracellular pathogens. SPRY domain-containing SOCS box protein 2 (SPSB2) is a regulator protein which recruits an E3 ubiquitin ligase to polyubiquitinate iNOS, causing proteasomal degradation.[10] Inhibition of iNOS·SPSB2 can prolong iNOS's lifetime and enhance the host's immune response.

The short hot loop of iNOS (Figure 4a) is DINNN, and four of these residues are hot spots (Figure 4b).[10b] Hot spots concentrated in a short sequence are ideal for overlay via precise alignment, hence core conformers of *cyclo*-{-$(Ala)_5$-organo-} were overlaid on iNOS hot loop. Seven hits with RMSD ≤ 0.8 Å were selected (Figure S8). Candidates with alanines substituted by DINNN were synthesized.

Of molecules found to perturb iNOS·SPSB2[10a,11] wild type iNOS *N*-terminal fragment, a 13-mer peptide, is one of the best binders in the literature (K_d 7.1 nM by SPR[11b]); this was selected for further studies as the positive control. An unconstrained linear 5-mer derived from the hot loop was used for a negative control (Figure 4d).

Fluorescence polarization (FP) for the *N*-FITC modified positive control showed it bound SPSB2 with a K_d of 23 ± 9 nM (Figure S15a). Candidates **A1**–**A7** screened in competition with this probe rank as best to worst of the seven binders selected from BM based on overlay RMSDs (Figure 4c). Those RMSDs (≈ 0.5 for **A1**–**A3**, ≈ 0.6 for **A4**–**A5**, ≈ 0.7 for **A6** and ≈ 0.8 for **A7**; Å throughout) have the same trend as the binding data in Figure 4c suggesting these parameters tend to correlate.

Best binders from the FP screen, **A1**–**A3**, were analyzed via conventional modeling. All had good overlays of core conformers on the hot loop (Figure 4e). We further grafted natural side chains to these cores and flexibly docked them to receptor binding surface by induced-fit docking in *Schrödinger*. Core structures overlaid well with the hot loop and side chains accurately filled sub pockets at the surface (Figure S12), demonstrating synergistic stabilization of loop

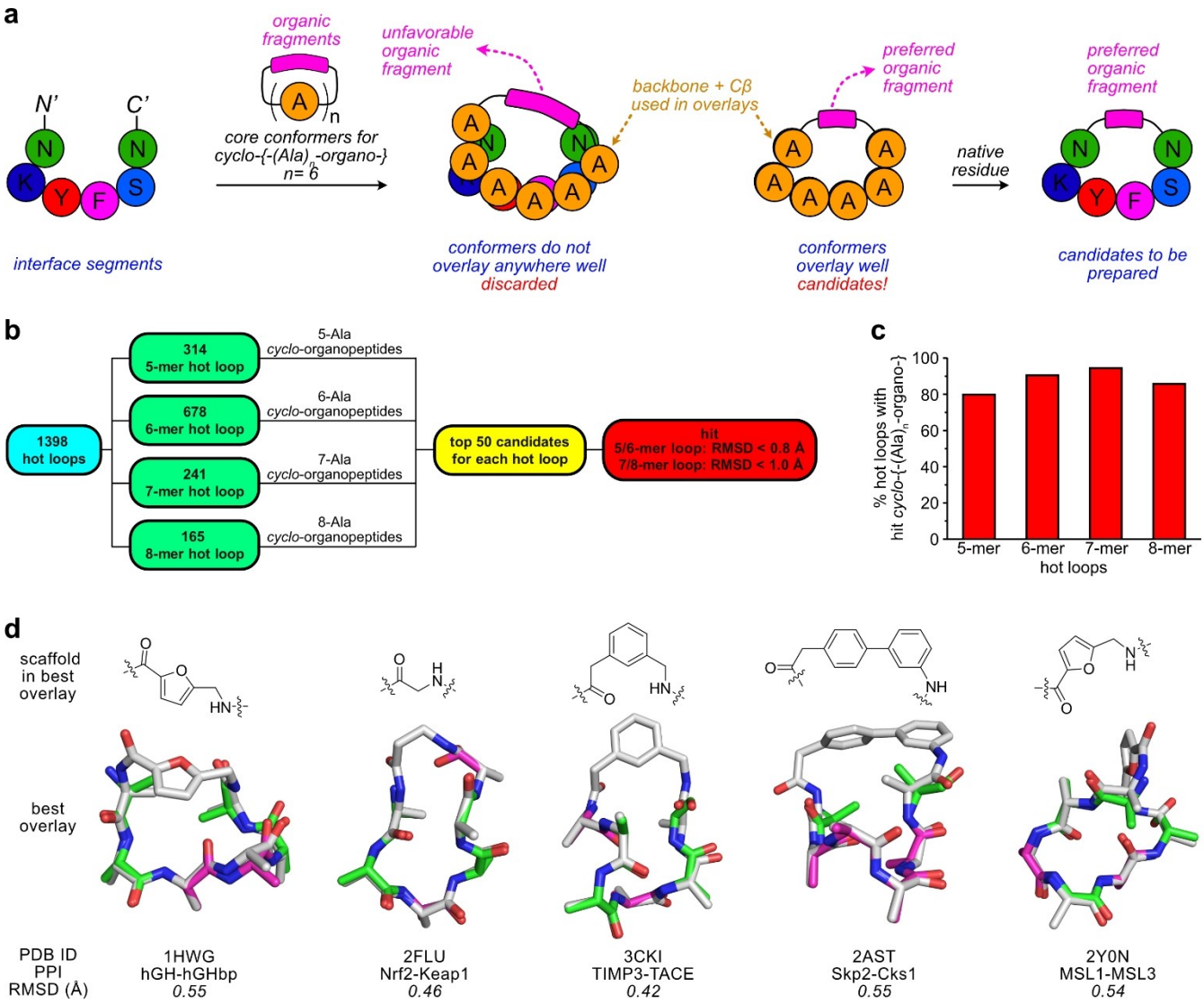

Figure 3. Alignment and virtual screening on hot loops. **a** 'Precise alignment' overlays a loop fragment with *cyclo*-{-(Ala)$_n$-organo-} of the same length using backbone and *Cβ* atoms ('auto-slicing alignment' is described in Supporting Information around Figure S7). **b** Workflow to screen hot loop dataset. **c** Percentage successful hit discovery after screening 5–8 mer hot loops by precise alignment.[4] **d** Top candidates from step **b** to mimic hot loops previously identified as interesting but not mimicked before; hot spots of these loops in magenta.

backbones and side chains interacting with the receptor pockets.

Affinities of **A1**–**A3** were further tested in dose-response competitive FP assays (Figure 4f). Their binding affinities were significantly better than the negative control (**A1** >100×; **A2**, 30×: **A3** 10×). The positive control and **A1** had a comparable K_i's.

Tryptophan fluorescence quenching (FQ) was also used to monitor binding (Figure 4g) via fluorescence of a solvent-exposed tryptophan in the receptor binding pocket (Figure 4a).[10b,11d] It showed the three hits had at least 20× more SBSP2 affinity than the negative control, and the best one, **A1** (K_d of 17±3 nM), was 2× better than the positive control.

In summary, *cyclo*-{-(Ala)$_5$-organo-} core conformations corresponding to ones expected for binding the SPSB2 receptor were selected from the virtual library by BM analyses. These overlaid well with the iNOS hot loop, and *cyclo*-organopeptides based on them bound SPSB2 well.

Application of BM at The uPA·uPAR Interface

Urokinase-type plasminogen activator (uPA)[12] overexpressed in cancers[13] facilitates degradation of interstitial tissues in tumor growth and metastasis. Proteolytic activity of this protease is upregulated by binding to its receptor (uPAR), leading to the production of the protease plasmin, through hydrolysis of plasminogen. Plasmin degrades extracellular matrix proteins such as fibronectin, vitronectin, and fibrin, clearing the way for epithelial to mesenchymal transitions to generate cancer stem cells which break out

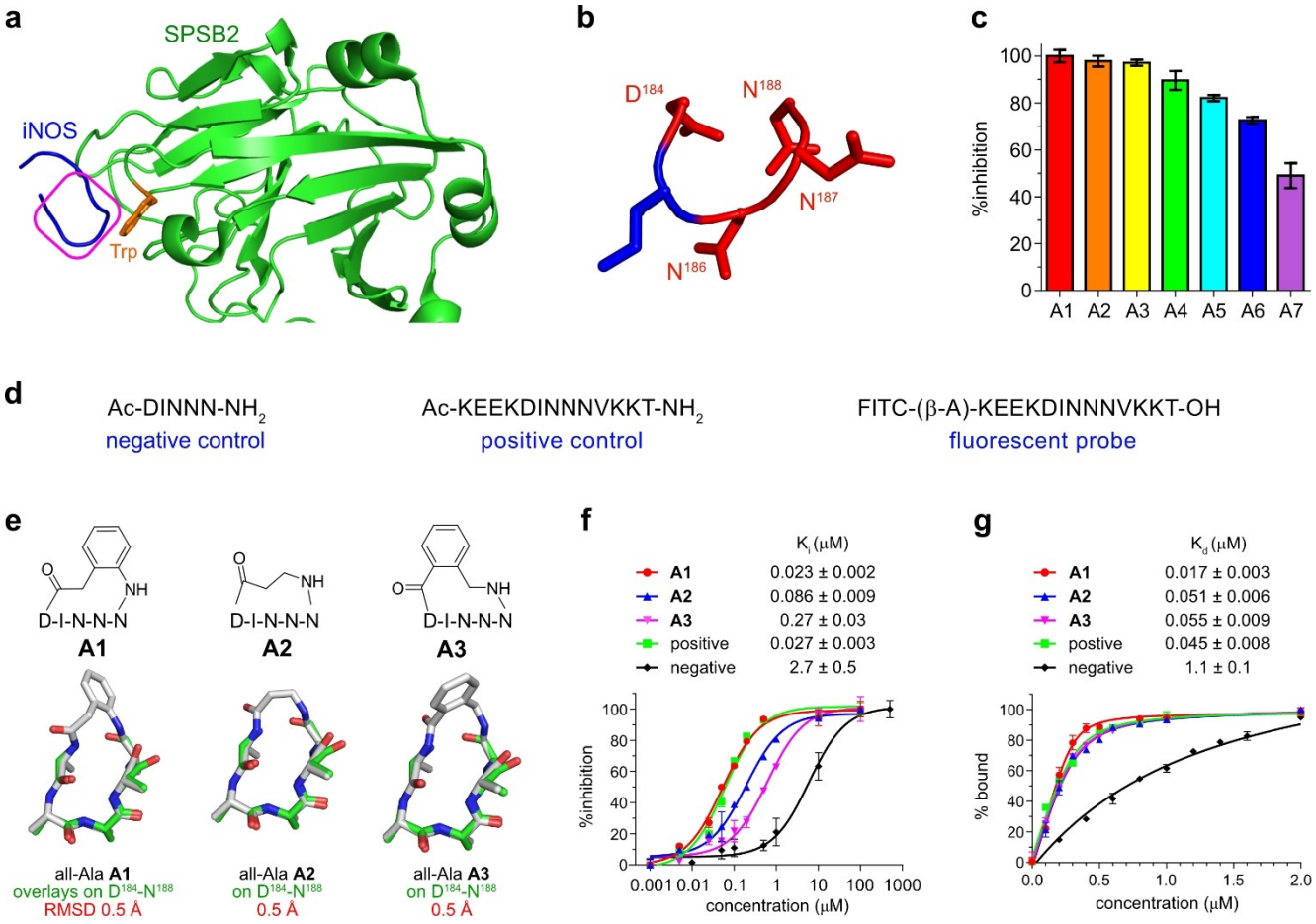

Figure 4. BM evaluation on iNOS·SPSB2. **a** Structure of iNOS·SPSB2 (PDB 6KEY) highlighting the hot loop and interface Trp. **b** Expanded structure of the hot loop from that PPI, sequence DINNN which encompasses the four hot spots shown in red. **c** Evaluation of seven hits (**A1** to **A7**) by competitive FP assay using a fixed ligand concentration of 50 μM. **d** Peptide-based controls: L-amino acids are capitalized, β-A is β-alanine. **e** Favored *cyclo*-{-(Ala)$_5$-organo-} and their hot loop overlays. **f** Data from dose-response, competitive FP, and **g** dose-response, direct tryptophan FQ assays.

from primary tumors. Thus interaction of uPA with uPAR is central to neoplastic growth and metastasis via tissue remodeling, tumor cell invasion, adhesion, and proliferation.[13a] Consistent with this, elevated uPA is associated with poor prognosis for patient survival.[14]

uPA·uPAR has a large interface loop (Figure 5a and 5b) featuring a hot segment[4] S^{21}–I^{28} and a distal hot spot W^{30}.[15] The hot spots are discontiguous and it is unclear where to place a linker in this large loop, so auto-slicing was used. Auto-slicing alignment indicated **B1**–**B7** (RMSD $\leq$1.0 Å) were optimal in the library. Hot spots encompassed by **B1** and **B2** are K^{23}–W^{30}, but the smaller mimics **B3**–**B7** do not include W^{30} (Figure S9). AE105, a linear peptide discovered by phage display, has sub-nanomolar affinity to uPAR (by surface plasmon resonance, SPR),[16] and AE147 is an extended version of AE105 with improved water solubility.[17] AE147-FITC served as the fluorescent probe in the competitive FP assays. It had a K_d of 64 $\pm$ 15 nM with uPAR (Figure S16a). Linear fragments of the parent loop are used as negative controls (Figure 5d).

Data from FP screening of **B1**–**B7** *vs* AE147-FITC are shown in Figure 5c. These data indicate **B1** and **B2** (with W^{30}) have significantly higher affinities than **B3**–**B7** (no W^{30}), so Trp^{30} appears to be an important hot spot. The top three leads, **B1**–**B3**, are based on three different fragments of S^{21}–W^{30}, and linkers (Figure 5e); they were further docked into the uPAR cavity and settled in locations similar to the corresponding loop fragments (Figure 5f and S13). Consequently, **B1**–**B3**, were selected for dose responsive competitive FP assays to compare with controls (Figure 5g and S16). Poor binding was observed for the linear peptide controls (Table 1) proving cyclization via BM-identified linkers increased affinities. Positive control, AE105, gave a 4× better K_i than **B1**, but the comparison is uneven because the bound conformation of AE147 (and presumably AE105) does not resemble uPA in terms of residues or shape, and it changes the uPAR cavity geometry on binding.[17]

Affinities of **B1**–**B3** and the positive control were competed with uPA-HRP (horse radish peroxidase) in Enzyme-linked Immunosorbent Assay (ELISA). This is an even comparison insofar as the uPA mimics compete with

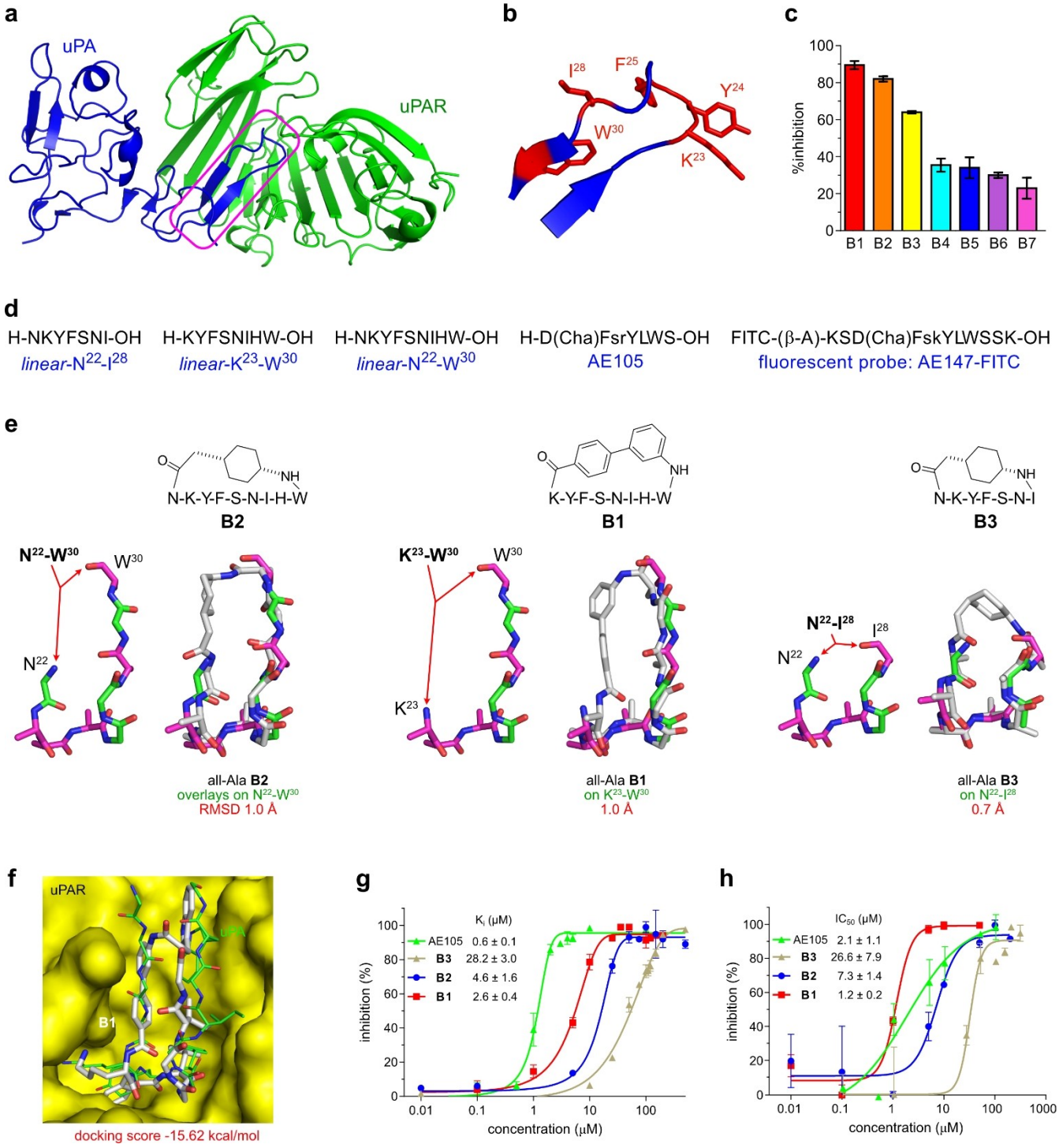

Figure 5. BM evaluation on uPA·uPAR using auto-slicing. **a** Structure of uPA·uPAR (3BT1). **b** Expansion highlighting the uPA hot loop having five potential hot spots shown in red. **c** Evaluation of hits **B1** to **B7** by competitive FP at fixed 50 μM ligand concentrations. **d** Peptide-based controls, where L-amino acids capitalized, D-amino acids in lower case, Cha is cyclohexyl alanine, β-A is β-alanine. **e** Favored *cyclo*-{-(Ala)$_n$-organo-} systems and their overlays with fragments of uPA loop. **f** Best pose of **B1** in uPAR cavity. **g** Dose-response competitive FP and **h** ELISA with immobilized uPAR *vs* uPA-HRP on the best hits **B1**–**B3** and AE105.

uPA itself and not with a system randomly generated in phage. In this assay, **B1** corresponds to a *smaller* IC_{50}, indicating better binding, compared with AE105 (Figure 5h).

Mimicry of large loops is more complicated than it is for shorter ones because their conformations are harder to recapitulate, and there are more possible positions to insert

Table 1: Comparison of FP and ELISA binding data for **B1**–**B3**, and controls.

compound	K_i (μM from FP)	IC_{50} (μM from ELISA)
B1	2.6 ± 0.4	1.2 ± 0.2
B2	4.6 ± 1.6	7.3 ± 1.4
B3	28.2 ± 3.0	26.6 ± 7.9
AE105	0.6 ± 0.1	2.1 ± 1.1
linear-N^{22}–W^{30}	28.1 ± 1.9	not tested
linear-K^{23}–W^{30}	343 ± 44	not tested
linear-N^{22}–I^{28}	> 500	not tested

a linker. This application of BM on uPA demonstrates its power to tackle this problem via auto-slicing alignment.

BM Augments Human Intuition in Loop Mimic Design

A routine criterion to select loop mimic linkers is matching their terminal distances (*ie* *NH*- to *CO*- in unnatural amino acids) with target loop fragments, but such reasoning simplifies reality for the following reason. In this work, the linkers connect via formation of two amide bonds that are not present in the loop fragment. Two new vectors corresponding to these could point in many directions depending on the docked *cyclo*-organopeptide conformation. In the illustrative case in Figure 6a, these distances vary from 2.3 to 6.1 Å depending on those vector orientations, whereas the brain probably only perceives the static *N*- to *C*- distance in the hot loop structure: 4.9 Å. Intuition cannot foresee the orientation of the two new bond vectors, or *cyclo*-organopeptide conformations accessible to each possible mimic.

The assertion is illustrated in Figure 6b. BM processes data from the loop fragment geometries, linker shapes, fragment *C*- to *N*-termini separations, and outputs RMSDs to evaluate overlays. Intuitive correlation of loop fragment *C*- to *N*-separations with accessible linker termini separations suggests the third best hit from BM, **B3**, and a system we developed intuitively, **B8**, are more likely to bind uPAR than **B1** and **B2**. However, the best conformer BM found for **B8** had a poor overlay, RMSD 1.3 Å (Figure S9b). Experimentally, and counter-intuitively, **B3** is a weaker binder than **B1** and **B2**, and **B8** had the weakest affinity in the series even though it spans a larger fragment of the hot loop than any of the other mimics (Figure 6b, right).

Extrapolating the discussion above, it is improbable intuition could be used to rank affinities of *cyclo*-organopeptides having structurally similar organic fragments. To test this, we intuitively selected three *cyclo*-organopeptides, **B1'**, **B1''** and **B2'**, comprising scaffolds with similar *C*- to *N*-separations and physiochemical properties to **B1** and **B2** (Figure 6c). The corresponding BM data was available, and we knew **B1** and **B2** core conformers should overlay better (from the rankings). In fact, the best overlay of **B1'** misaligns its W with W^{30}, so it is effectively similar to **B3** which does not have Trp. Two hot spots (I^{28}, W^{30}) are misaligned in best overlays of **B1''** and **B2'**. FP assays showed affinities correlating with the BM predicted overlays: **B1'**, 18.0 (*vs* **B3**, 28.2); **B1''**, 90.4; and, **B2'**, 124, respectively (μM throughout). Consequently, BM outperformed our intuition.

In summary, distance-based loop mimicry is useful for approximate evaluations. However, BM can process more data quantitatively, select better compounds to prepare and test, and is less vulnerable to false-positives.

Conclusion

Core conformation databases based on straightforward, fast, and computationally inexpensive conformational sampling of *cyclo*-{-$(Ala)_n$-organo-} systems are valuable. They facilitate evaluation of large libraries of *cyclo*-organopeptides and compare favorably with conventional high throughput virtual docking approaches which become decreasingly tractable as the library size increases. Another appealing attribute of core conformation databases is BM can evaluate them against *all* PPIs. Conversely, data from high throughput virtual docking to *particular* protein receptors does not extrapolate to other PPIs. Moreover, BM streamlines computational effort by initially ignoring amino acid side chain influences, then considering their effects after preferred backbone conformers are overlaid on protein ligand loops.

The alternative alignment options in BM are useful. Precise alignment on small hot loops is fast and nevertheless renders small, validated, *cyclo*-organopeptide hits. Alternatively, users should use auto-slicing to evaluate *cyclo*-organopeptide analogs of large loop fragments; this may indicate small mimics of large loops. This is useful because molecular size is often inversely correlated to factors which determine bioavailabilities.

Reliable, but computationally expensive and time-consuming docking procedures can be used to further validate BM-prioritized hits. These methods were implemented for the best BM hits in this study, and indicated all of them were credible binders (Supporting Information section J). We did not use high level computational docking to optimize hits by replacing natural amino acid side chains with unnatural ones, but that surely could be explored to find analogs with enhanced binding.

BM was more effective than intuition. Hit deduction requires more than crudely matching *N*- to *C*-distances in loops with linkers. An interesting illustration of BM's sophistication emerged when it was applied to uPA·uPAR. BM selected the *same* linker for **B2** and **B3** to span peptide strings of *different* lengths; we did not anticipate this and were surprised. Close inspection of the overlaying core conformers revealed *different* cyclohexane chair conformations; **B2** has $-CH_2CO-$ equatorial but **B3** has it axial. An equatorial $-CH_2CO-$ can accommodate the 9-mer peptide string in **B2**, while **B3** with the shorter string has the chair-flipped isomer corresponding to closer termini (Figure 7). BM selected different chair conformers of a linker demonstrating *more* predictive insight than us.

BM cannot, and was not designed to solve all problems in loop mimicry. Downstream issues like cell permeabilities, in vivo clearance, biological efficacies, and toxicities must be

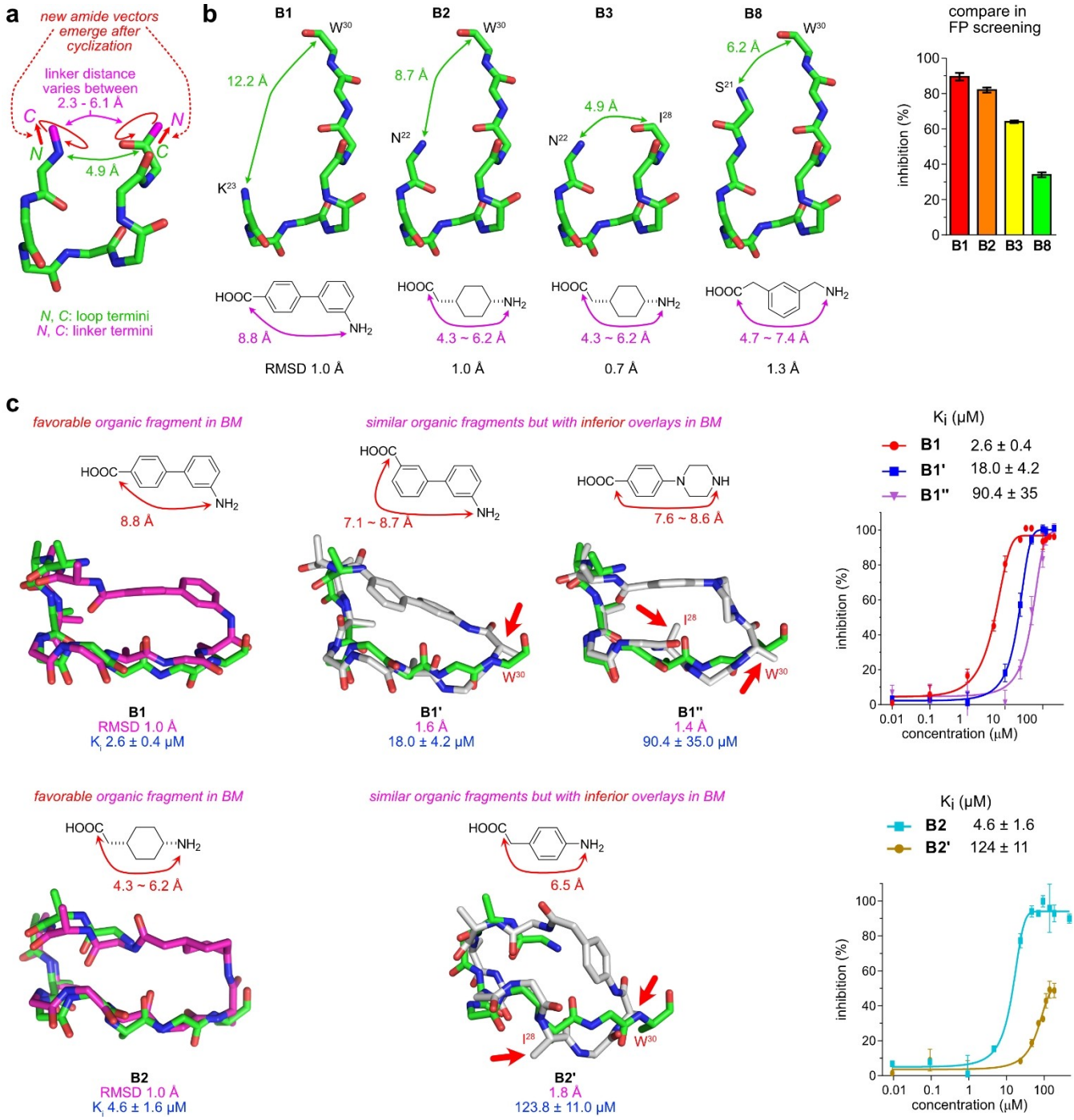

Figure 6. Comparison of BM with intuitive, termini-distance estimate methods. **a** A hot loop with a static *N*- to *C*-termini separation in the solid state can be mimicked using linkers of variable *C*- to *N*-termini distances because of rotations of the new amide bonds created on cyclization into *cyclo*-organopeptides. **b** Consistent with that conclusion, linkers selected by BM do not always have termini separations directly corresponding to those in the uPA loop (left and center), and linker separations *do not correlate* with the binding trends in competitive FP (right). **c** Goodness of fit in mimic/hot loop overlays (left and center) *do correlate* with *cyclo*-organopeptide affinities for uPAR (right); **B1** and **B2** selected by BM overlaid well and had higher affinities than **B1'**, **B1"**, and **B2'** which had similar organic fragments but overlaid inferiorly.

addressed separately. BM here was restricted to organic fragments with amine and acid functionalities, but we cannot foresee any major obstacles associated with modification for other organic fragments; ongoing studies in our laboratory feature generation of alternative libraries cyclized via reactions other than amide bond formation. Another limitation in this pilot study is interface loops containing Gly and Pro may be not adequately represented by all-Ala cores, but this would be easily resolved by expanding core libraries to include those residues for the PPI of interest.

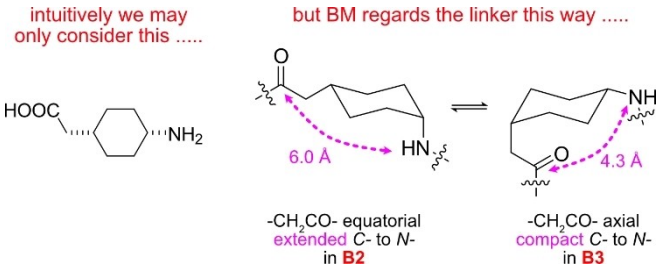

Figure 7. Intuition has less insight into conformational variability than BM using exhaustive conformational sampling.

BM is a useful tool. All the scripts described here are open access, so researchers could evaluate BM for design of probes to mimic or disrupt PPI targets of interest in their laboratories.

Acknowledgements

Financial support was provided by NIH R01EY029645, NSF CHE 1608009, and the Texas A&M University T3-Grants Program (246292-00000). We thank Dr. Lisa M. Perez and LMS (Laboratory for Molecular Simulation) for helpful discussions and computational resources during research.

Conflict of Interest

The authors declare no conflict of interest.

Data Availability Statement

The data that support the findings of this study are available in the supplementary material of this article.

Keywords: Loop Mimicry · Macrocycles · Protein-Protein Interactions · Virtual Screening

[1] a) A. Sehgal, *Protein-Protein Interactions as Drug Targets: Focusing on Cell Circuitry*, illustrated ed, D & MD Publications, **2005**; b) A. L. Hopkins, C. R. Groom, *Nat. Rev. Drug Discovery* **2002**, *1*, 727–730; c) A. L. Jochim, P. S. Arora, *ACS Chem. Biol.* **2010**, *5*, 919–923.

[2] a) N. London, B. Raveh, O. Schueler-Furman, *Curr. Opin. Chem. Biol.* **2013**, *17*, 952–959; b) J. A. Wells, C. L. McClendon, *Nature* **2007**, *450*, 1001–1009; c) W. L. DeLano, *Curr. Opin. Struct. Biol.* **2002**, *12*, 14–20.

[3] a) I. V. Smolyar, A. K. Yudin, V. G. Nenajdenko, *Chem. Rev.* **2019**, *119*, 10032–10240; b) A. K. Malde, T. A. Hill, A. Iyer, D. P. Fairlie, *Chem. Rev.* **2019**, *119*, 9861–9914; c) Y. Huang, M. M. Wiedmann, H. Suga, *Chem. Rev.* **2019**, *119*, 10360–10391; d) D. A. Guarracino, J. A. Riordan, G. M. Barreto, A. L. Oldfield, C. M. Kouba, D. Agrinsoni, *Chem. Rev.* **2019**, *119*, 9915–9949; e) Y. Chow Hoi, Y. Zhang, X. Li, E. Matheson, X. Li, *Chem. Rev.* **2019**, *119*, 9971–10001; f) S. D. Appavoo, S. Huh, D. B. Diaz, A. K. Yudin, *Chem. Rev.* **2019**, *119*, 9724–9752; g) P. Hosseinzadeh, G. Bhardwaj, V. K. Mulligan, M. D. Shortridge, T. W. Craven, F. Pardo-Avila, S. A. Rettie, D. E. Kim, D.-A. Silva, Y. M. Ibrahim, I. K. Webb, J. R. Cort, J. N. Adkins, G. Varani, D. Baker, *Science* **2017**, *358*, 1461–1466; h) E. Valeur, S. M. Gueret, H. Adihou, R. Gopalakrishnan, M. Lemurell, H. Waldmann, T. N. Grossmann, A. T. Plowright, *Angew. Chem. Int. Ed.* **2017**, *56*, 10294–10323; i) A. K. Yudin, *Chem. Sci.* **2015**, *6*, 30–49; j) T. A. Hill, N. E. Shepherd, F. Diness, D. P. Fairlie, *Angew. Chem. Int. Ed.* **2014**, *53*, 13020–13041; k) C. J. White, A. K. Yudin, *Nat. Chem.* **2011**, *3*, 509–524; l) E. Marsault, H. R. Hoveyda, M. L. Peterson, R. Gagnon, M. Vezina, J. F. Pinault, A. Landry, C. Saint-Louis, L. G. Ouellet, S. Beauchemin, K. Benakli, S. Beaubien, M. Brassard, Z. Wang, M. Champagne, F. Galaud, N. Fortin, D. Fortin, V. Plourde, M. Ramaseshan, S. Bhat, F. Bilodeau, D. Lonergan, R. Lan, S. Li, G. Berthiaume, L. Foucher, X. Peng, Y. Dory, P. Deslongchamps, *Adv. Exp. Med. Biol.* **2009**, *611*, 15–16; m) Y. S. Tan, D. P. Lane, C. S. Verma, *Drug Discovery Today* **2016**, *21*, 1642–1653; n) P. M. Cromm, J. Spiegel, T. N. Grossmann, *ACS Chem. Biol.* **2015**, *10*, 1362–1375; o) L. D. Walensky, G. H. Bird, *J. Med. Chem.* **2014**, *57*, 6275–6288.

[4] J. Gavenonis, B. A. Sheneman, T. R. Siegert, M. R. Eshelman, J. A. Kritzer, *Nat. Chem. Biol.* **2014**, *10*, 716–722.

[5] A. M. Ali, J. Atmaj, N. Van Oosterwijk, M. R. Groves, A. Domling, *Comput. Struct. Biotechnol. J.* **2019**, *17*, 263–281.

[6] A. Nazzaro, B. Lu, N. Sawyer, A. M. Watkins, P. S. Arora, *Angew. Chem. Int. Ed.* **2023**, *62*, e202303943.

[7] a) T. R. Siegert, M. Bird, J. A. Kritzer, *Methods Mol. Biol.* **2017**, *1561*, 255–277; b) T. R. Siegert, M. J. Bird, K. M. Makwana, J. A. Kritzer, *J. Am. Chem. Soc.* **2016**, *138*, 12876–12884.

[8] G. Olanders, H. Alogheli, P. Brandt, A. Karlén, *J. Comput.-Aided Mol. Des.* **2020**, *34*, 231–252.

[9] a) T. Kortemme, D. Baker, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 14116–14121; b) T. Kortemme, E. Kim David, D. Baker, *Sci. STKE* **2004**, *2004*, pl2; c) A. A. Ibarra, G. J. Bartlett, Z. Hegedüs, S. Dutt, F. Hobor, K. A. Horner, K. Hetherington, K. Spence, A. Nelson, T. A. Edwards, D. N. Woolfson, R. B. Sessions, A. J. Wilson, *ACS Chem. Biol.* **2019**, *14*, 2252–2263; d) I. S. Moreira, *Curr. Top. Med. Chem.* **2015**, *15*, 2068–2079; e) J. M. Martins, R. M. Ramos, A. C. Pimenta, I. S. Moreira, *Proteins* **2014**, *82*, 479–490; f) N. Tuncbag, A. Gursoy, O. Keskin, *Bioinformatics* **2009**, *25*, 1513–1520.

[10] a) K. Li, T. You, P. Zhao, Y. Luo, D. Zhang, H. Wei, Y. Wang, J. Yang, X. Guan, Z. Kuang, *Nitric Oxide* **2021**, *113–114*, 1–6; b) Z. Kuang, R. S. Lewis, J. M. Curtis, Y. Zhan, B. M. Saunders, J. J. Babon, T. B. Kolesnik, A. Low, S. L. Masters, T. A. Willson, L. Kedzierski, S. Yao, E. Handman, R. S. Norton, S. E. Nicholson, *J. Cell Biol.* **2010**, *190*, 129–141.

[11] a) P. Filippakopoulos, A. Low, T. D. Sharpe, J. Uppenberg, S. Yao, Z. Kuang, P. Savitsky, R. S. Lewis, S. E. Nicholson, R. S. Norton, A. N. Bullock, *J. Mol. Biol.* **2010**, *401*, 389–402; b) S. M. Gueret, S. Thavam, R. J. Carbajo, M. Potowski, N. Larsson, G. Dahl, A. Dellsen, T. N. Grossmann, A. T. Plowright, E. Valeur, M. Lemurell, H. Waldmann, *J. Am. Chem. Soc.* **2020**, *142*, 4904–4915; c) M. M. Sadek, N. Barlow, E. W. W. Leung, B. J. Williams-Noonan, B. K. Yap, F. M. Shariff, T. T. Caradoc-Davies, S. E. Nicholson, D. K. Chalmers, P. E. Thompson, R. H. P. Law, R. S. Norton, *ACS Chem. Biol.* **2018**, *13*, 2930–2938; d) B. K. Yap, E. W. Leung, H. Yagi, C. A. Galea, S. Chhabra, D. K. Chalmers, S. E. Nicholson, P. E. Thompson, R. S. Norton, *J. Med. Chem.* **2014**, *57*, 7006–7015.

[12] a) K. Dass, A. Ahmad, A. S. Azmi, S. H. Sarkar, F. H. Sarkar, *Cancer Treat. Rev.* **2008**, *34*, 122–136; b) M. C. Kriegbaum, M. Persson, L. Haldager, W. Alpizar-Alpizar, B. Jacobsen, H. Gaardsvoll, A. Kjaer, M. Ploug, *Curr. Drug Targets* **2011**, *12*, 1711–1728.

[13] a) M. V. Carriero, P. Franco, G. Votta, I. Longanesi-Cattani, M. T. Vento, M. T. Masucci, A. Mancini, M. Caputi, I. Iaccarino, M. P. Stoppelli, *Curr. Drug Targets* **2011**, *12*, 1761–1771; b) A. H. Mekkawy, D. L. Morris, M. H. Pourgholami, *Future Oncol.* **2009**, *5*, 1487–1499; c) M. Jo, R. D. Lester, V. Montel, B. Eastman, S. Takimoto, S. L. Gonias, *J. Biol. Chem.* **2009**, *284*, 22825–22833; d) X.-w. Hu, H.-f. Duan, L.-H. Gao, S.-y. Pan, Y.-m. Li, Y. Xi, S.-R. Zhao, L. Yin, J.-f. Li, H.-P. Chen, C.-T. Wu, *Cancer Biol. Ther.* **2008**, *7*, 651–659; e) J. Romer, B. S. Nielsen, M. Ploug, *Curr. Pharm. Des.* **2004**, *10*, 2359–2376.
[14] L. Tang, X. Han, *Biomed. Pharmacother.* **2013**, *67*, 179–182.
[15] P. A. Morais, F. F. Maia, C. Solis-Calero, E. W. S. Caetano, V. N. Freire, H. F. Carvalho, *Phys. Chem. Chem. Phys.* **2020**, *22*, 3570–3583.
[16] M. Ploug, S. Ostergaard, H. Gaardsvoll, K. Kovalski, C. Holst-Hansen, A. Holm, L. Ossowski, K. Dano, *Biochemistry* **2001**, *40*, 12157–12168.
[17] P. Llinas, M. H. Le Du, H. Gardsvoll, K. Dano, M. Ploug, B. Gilquin, E. A. Stura, A. Menez, *EMBO J.* **2005**, *24*, 1655–1663.

Manuscript received: June 1, 2023
Accepted manuscript online: October 17, 2023
Version of record online: November 3, 2023