

Accurate de novo design of high-affinity protein-binding macrocycles using deep learning

Received: 31 December 2024

Accepted: 2 May 2025

Published online: 20 June 2025

Check for updates

Stephen A. Rettie ^{1,2,3,15}, David Juergens ^{2,4,15}, Victor Adebomi^{1,2,15}, Yensi Flores Bueso ^{1,2,5,6}, Qinqin Zhao⁷, Alexandria N. Leveille⁸, Andi Liu ⁷, Asim K. Bera ², Joana A. Wilms ^{9,10}, Alina Üffing ^{9,10,14}, Alex Kang ², Evans Brackenbrough², Mila Lamb², Stacey R. Gerben ², Analisa Murray ², Paul M. Levine ², Maika Schneider ^{1,2,11}, Vibha Vasireddy^{1,2}, Sergey Ovchinnikov¹², Oliver H. Weiergräber¹⁰, Dieter Willbold ^{9,10}, Joshua A. Kritzer ⁸, Joseph D. Mougous ^{7,13}, David Baker ^{2,5,13} , Frank DiMaio ^{2,5} & Gaurav Bhardwaj ^{1,2}

Developing macrocyclic binders to therapeutic proteins typically relies on large-scale screening methods that are resource intensive and provide little control over binding mode. Despite progress in protein design, there are currently no robust approaches for de novo design of protein-binding macrocycles. Here we introduce RFpeptides, a denoising diffusion-based pipeline for designing macrocyclic binders against protein targets of interest. We tested 20 or fewer designed macrocycles against each of four diverse proteins and obtained binders with medium to high affinity against all targets. For one of the targets, Rhombotarget A (RbtA), we designed a high-affinity binder ($K_d < 10$ nM) despite starting from the predicted target structure. X-ray structures for macrocycle-bound myeloid cell leukemia 1, γ -aminobutyric acid type A receptor-associated protein and RbtA complexes match closely with the computational models, with a C α root-mean-square deviation < 1.5 Å to the design models. RFpeptides provides a framework for rapid and custom design of macrocyclic peptides for diagnostic and therapeutic applications.

Macrocyclic peptides present a promising avenue for developing new therapeutics that bridge the gap between small-molecule drugs and large biologics^{1,2}. Biologics, while capable of binding diverse therapeutic targets with high affinity and selectivity, are usually unable to cross cell membranes because of their large size and high polarity, limiting them to extracellular targets. Conversely, small molecules can access intracellular targets but are not ideal for targeting proteins lacking deep hydrophobic pockets. In principle, macrocyclic peptides with sizes between small molecules and proteins can be developed to modulate molecular targets inaccessible to traditional therapeutic

modalities³. The ability to develop custom protein-binding macrocycles for diverse protein targets would have many diagnostic and therapeutic applications. Traditionally, the development of peptide therapeutics has relied on natural product discovery or high-throughput screening of trillions of random peptides for target binding using display-based techniques^{1,2}. However, natural product discovery has several challenges, particularly synthetic difficulties, marginal stability and low mutational tolerance of identified hits³. While powerful, the high-throughput screening methods are time-intensive, cost-intensive and labor-intensive and only span a small fraction of the rich chemical

A full list of affiliations appears at the end of the paper. e-mail: dabaker@uw.edu; dimaio@uw.edu; gauravb@uw.edu

and structural diversity accessible to macrocycles. Moreover, such approaches frequently fail to simultaneously optimize for multiple biophysical properties, such as target binding, selectivity and membrane permeability, because of the precise structural control required to achieve such functional properties⁵.

Structure-guided design methods offer a complementary approach to the library screening approaches, enabling rapid *in silico* exploration of a large chemical and structural diversity to design macrocycle binders for therapeutic targets. We previously developed physics-based methods for designing hyperstable constrained peptides, structured macrocycles and binders to protein targets by borrowing the motifs or interactions from previously described binding partners as anchors^{6–9}. However, despite the high accuracy observed in the design of monomeric macrocycles with these methods⁷, the design of protein-binding macrocycles has had limited success, achieving only modest binding affinities and, in many cases, with the experimentally determined structures not agreeing with the design models^{7,8,10}. The reliance on previously described binding partners for starting motifs also restricts such approaches to well-studied protein targets. In recent work, we described a pipeline for hallucinating and predicting the structures of macrocyclic peptide monomers by modifying AlphaFold2 (AF2) to include cyclic relative positional encoding (named 'AfCycDesign')¹¹. Other promising deep learning (DL) methods were described recently to predict the structures of macrocycles and macrocycle–target complexes^{12,13} and to design peptide binders to protein targets^{14–16}. However, these methods have not been extensively structurally validated to date or shown to robustly perform atomically accurate *de novo* design of macrocyclic peptide structures in complexes with diverse protein targets. Computational methods that can accurately design high-affinity macrocycle binders *de novo*, using just the information of target structure or sequence, are required for wider therapeutic applications.

We reasoned that recent breakthroughs in generative DL methods could be leveraged to develop a robust pipeline for the accurate and efficient design of macrocycle binders. Diffusion models for protein design, such as RFdiffusion¹⁷, are trained to generate diverse protein structures from randomly initialized residues as starting points and have demonstrated remarkable success in designing protein monomers, binders and symmetric oligomers of medium-sized to large-sized proteins. However, despite considerable recent progress in DL-based protein design methods, these methods are not readily applicable to designing macrocyclic peptides. Developing analogous methods for peptide design from scratch has been challenging because of the limited availability of experimental data for training such models. To address these challenges, we set out to extend the RoseTTAFold2 (RF2)¹⁸ structure prediction network and the RFdiffusion¹⁷ protein backbone generation framework to incorporate cyclic relative positional encoding and enable the generation of the macrocyclic peptide backbones.

Extending RF2 and RFdiffusion for macrocycles

We began by examining the ability of the RF2 (ref. 18) structure prediction network to model known macrocyclic peptide structures. We implemented a modified (Methods) cyclic relative position encoding for RF2 (Fig. 1a) and observed robust prediction of natural cyclic peptide structures (Supplementary Fig. 1). Given this success, we reasoned that the same relative positional encoding should enable RFdiffusion¹⁷ to generate macrocyclic peptide structures because of its similar network architecture. We added the cyclic positional encoding scheme to RFdiffusion and observed robust generation of diverse macrocyclic peptides (Fig. 1b,c and Supplementary Fig. 2). Similar to the previously described work on designing monomeric cyclic peptides with physics-based methods⁷ and AfCycDesign¹¹, we observed 9,045 and 8,913 structurally unique 10-residue and 12-residue backbones, respectively, when 48,000 macrocycle backbones were generated for each size (Supplementary Fig. 2). The distribution of phi and psi values in these

generated backbones is similar to the standard Ramachandran plot for protein structures (Supplementary Fig. 2), suggesting that generated backbones do not require extensive D-amino acids to stabilize the generated structures⁷. While we did not attempt to comprehensively enumerate the structural space of cyclic peptide monomers, RFpeptides can readily be scaled up to comprehensively cover the structural space accessible to macrocyclic peptides. Encouraged by the transferability of the cyclic positional encoding, we set out to use RFdiffusion for the *de novo* design of protein-binding macrocycles. We chose RFdiffusion for several reasons. Firstly, we expected the high experimental success rate of RFdiffusion^{17,19} for protein binder design to carry over to macrocycle binder design. Secondly, *de novo* binder design with AfCycDesign as it would be far more computationally expensive and has not been successfully implemented or experimentally validated. Thirdly, the method can still take advantage of the current built-in conditional generation functionalities of RFdiffusion, such as epitope-specific targeting and 'motif' scaffolding. Lastly, the method should be directly transferable to other current and future RoseTTAFold-based design networks, such as RFdiffusion All-Atom²⁰, for incorporating nonpeptidic molecules (nucleic acids, ions, etc.) during design calculations.

We modified the RFdiffusion protein binder design pipeline to use cyclic relative position encodings for the generated chain and standard positional encodings for the target and interbinder target indices (Fig. 1d). We then completed our design pipeline by using ProteinMPNN²¹ to design amino acid sequences compatible with the backbones generated by RFdiffusion (Fig. 1e). We chose ProteinMPNN for its improved performance in sequence design and ability to generate sequences with better solubility profiles over the sequences generated by traditional physics-based methods²². This pipeline readily generated macrocycles with diverse secondary structure content against target proteins (Fig. 1f) and the inclusion of standard RFdiffusion hotspot features clearly shifted the distribution of generated binders toward desired residues (Supplementary Fig. 3). We refer to this integrated pipeline as 'RFpeptides' throughout the remainder of the text.

De novo design of macrocyclic binders to myeloid cell leukemia 1 and MDM2

We selected myeloid cell leukemia 1 (MCL1) as our first target protein, given the availability of multiple high-resolution X-ray crystal structures available to initiate the design calculations. MCL1 is also a promising target for anticancer therapeutics because of its roles in autophagy, cell survival, DNA repair and cellular proliferation²³. For targeting MCL1, we used RFpeptides to generate 9,965 diverse cyclic peptide backbones, followed by four iterative rounds of ProteinMPNN and Rosetta Relax to design four amino acid sequences for each generated backbone. We expected the local changes to the generated backbone during the Rosetta Relax steps to allow for improved amino acid sequence diversity from the ProteinMPNN steps. While there are other ways to achieve increased sequence diversity, including generating multiple sequences per backbone from ProteinMPNN or adding noise during ProteinMPNN sequence generation, we did not explicitly try or compare them in this study. For downselecting the design candidates for experimental testing, we used AfCycDesign to repredict the designed macrocycle–target complexes from the macrocycle sequence and the target structure as a template. We selected the designs on the basis of the confidence metric (interface predicted aligned error (iPAE)) and the similarity between the original design model and the protein–macrocycle complex predicted by the AfCycDesign (Supplementary Fig. 4). For further stringency in the design selection process, we also used RF2 to repredict the complex structures, reasoning that the design models predicted identically by two orthogonal structure prediction networks (AfCycDesign and RF2) should have a higher likelihood of binding to the target as designed. However, the 1,984 selected designs at this stage were still more than the number of designs we could reasonably synthesize and test experimentally. Therefore, we next used Rosetta²⁴

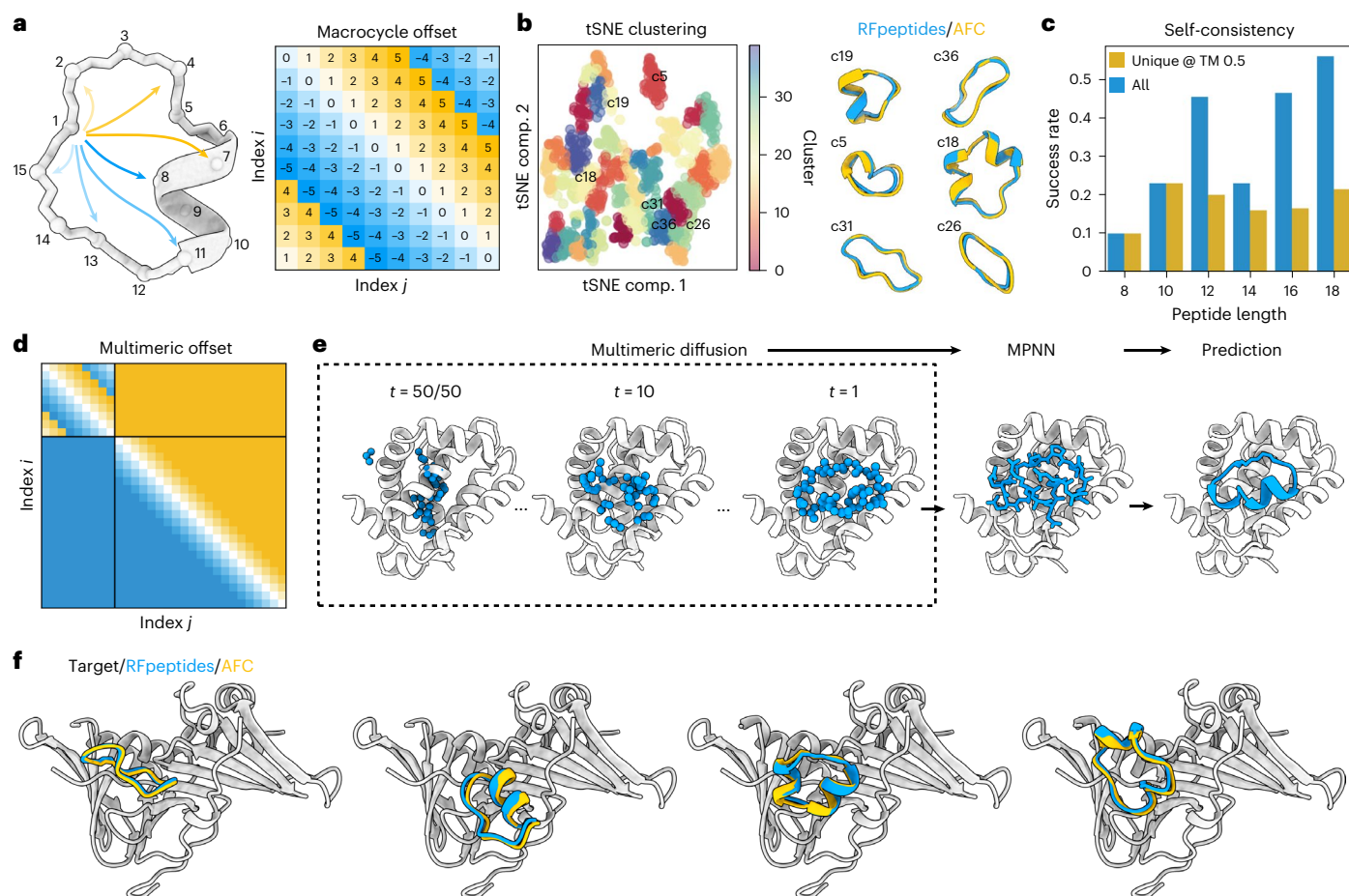


Fig. 1 RFpeptides is a diffusion-based pipeline for the de novo design of protein-binding macrocycles. **a**, Cyclically symmetric relative position encoding enables the generation of macrocyclic peptide backbones with N and C termini linked by a peptide bond. The relative position encodings are cyclized by switching from positive relative position encodings (that is, to the right) to negative encodings (that is, to the left) when index j is more than halfway around the peptide relative to index i . **b**, RFpeptides produces diverse and designable cyclic peptides. Left: structural clusters calculated using t -distributed stochastic neighbor embedding (tSNE)^{31,32} to reduce the dimensionality of an all-by-all TMscore matrix computed with TMalign³³ on an unfiltered set of 1,200 macrocycles generated using RFpeptides. Right: six RFpeptides outputs from differing structural clusters, all with $<1\text{ \AA}$ backbone r.m.s.d. between the design model (blue) and the structure predicted by AfCycDesign (gold). comp., component. **c**, Self-consistency of designed macrocycles of various lengths. For each peptide length, the fraction (with $n = 200$ per length) of backbones with at least one of the of eight LigandMPNN³⁴ sequences predicted by AfCycDesign

to refold with pLDDT > 0.8 and within 2.0 \AA backbone r.m.s.d. of the designed structure. Success rates for all sampled backbones are in blue and success rates only counting unique structural clusters (as calculated using MaxCluster^{35,36} at a TMscore threshold of 0.5) are in orange. **d**, For multichain diffusion trajectories (for example, macrocycle binder design), the relative positional encoding for the macrocycle chain is cyclized, whereas interchain and target chain relative positional encoding is kept as standard. **e**, Pipeline for the design of protein-binding macrocycles using RFpeptides. Macrocycle backbones are generated from randomly initialized atoms by a stepwise RFdiffusion-based denoising process, followed by amino acid sequence design using ProteinMPNN. Design models are downselected on the basis of the computational metrics from structure prediction using AfCycDesign and physics-based interface quality metrics using Rosetta. **f**, RFpeptides generates diverse macrocycles against selected targets. Four diverse cyclic peptide binders against the same target were generated using RFpeptides, with AfCycDesign iPAE < 0.3 and C α r.m.s.d. $< 1.5\text{ \AA}$ between the design model (blue) and AfCycDesign prediction (gold).

to calculate the ‘physics-based’ metrics of interface and macrocycle quality, such as calculated binding affinity (ddG), spatial aggregation propensity (SAP) of the designed macrocycle and the molecular surface area of the interface contacts (CMS) (Supplementary Fig. 4).

After strictly filtering the designed candidates on DL-based and physics-based metrics, we selected 27 designs for synthesis, biochemical and biophysical characterization. Despite specifying no hotspots to guide the generation process to a specific patch on the MCL1 structure, all selected designs bound to the functionally relevant MCL1-BH3 interaction site (Supplementary Fig. 5). While all selected designs include an α -helical segment, they feature different sequences, macrocycle placement and target interactions (Supplementary Fig. 5 and Supplementary Table 1a). In addition to the common helical motifs, the loop regions of the selected macrocycles also contribute extensive side-chain-mediated and backbone-mediated

interactions to the binding interface. During the chemical synthesis using Fmoc-based solid-phase synthesis (Methods), the yields for the correctly cyclized product for 13 designs were low and insufficient for further characterization. We tested the remaining 14 macrocycles for binding to biotinylated MCL1 using surface plasmon resonance (SPR) single-cycle kinetics experiments (Supplementary Fig. 6). Three macrocycles showed binding to the MCL1, with the best binder, MCB_D2 (MCL1 binding design 2) (Fig. 2a), demonstrating a binding affinity of $2\text{ }\mu\text{M}$ (Fig. 2b). To confirm whether the designed macrocycle adopts the designed structure and engages MCL1 in the designed binding mode, we determined the X-ray crystal structure of MCB_D2 bound to MCL1 at 2.1 \AA resolution. The crystal structure was nearly identical to the design model, with a root-mean-square deviation (r.m.s.d.) of 0.7 \AA over all of the C α atoms of the macrocycle with target chains aligned (Fig. 2c) and C α r.m.s.d. of 0.4 \AA within the macrocycles when aligned (Fig. 2d). The

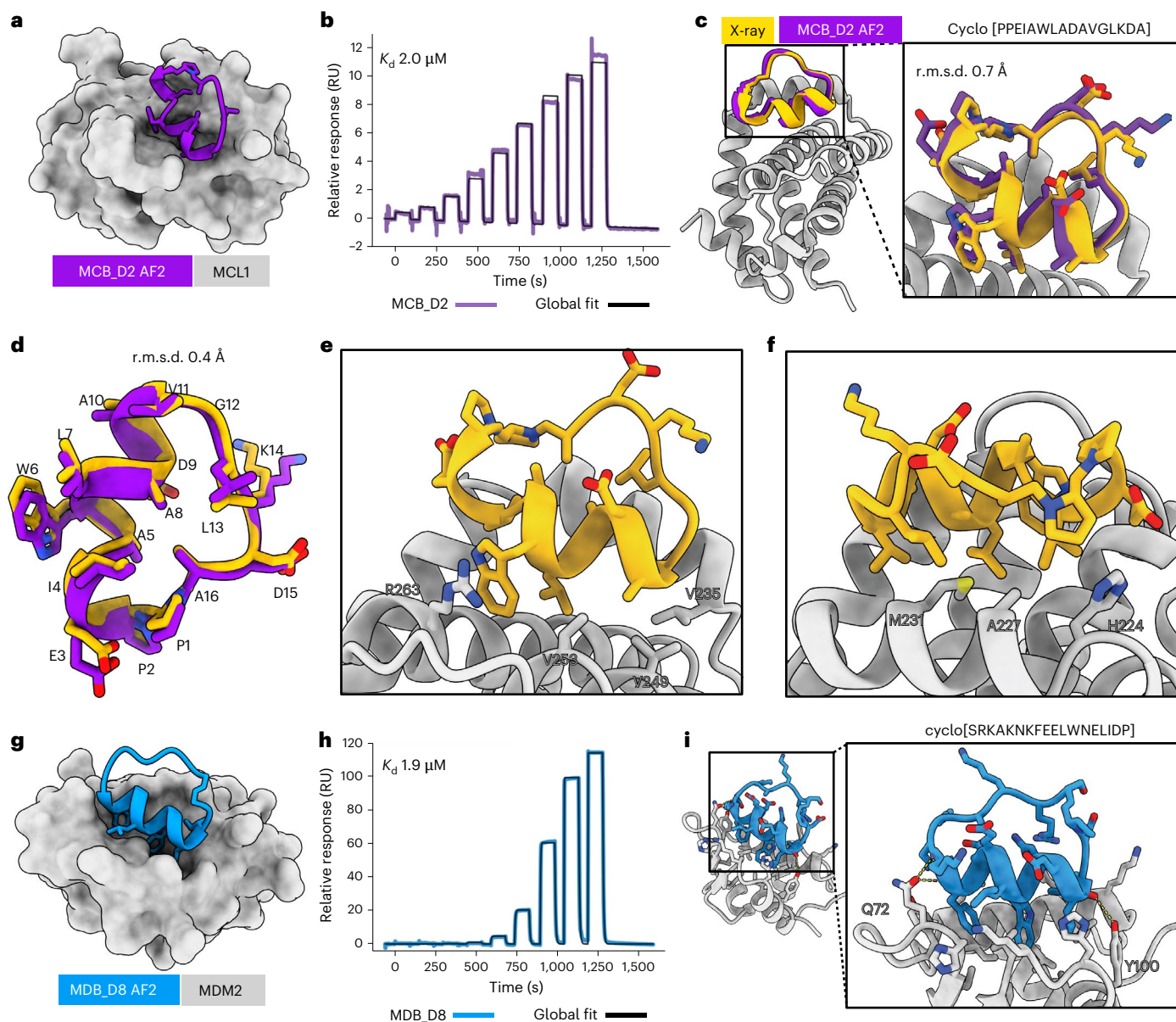


Fig. 2 | De novo design and characterization of macrocyclic binders to MCL1 and MDM2. **a**, AfCycDesign prediction of MCB_D2 (purple) bound to MCL1 (gray surface). MCB_D2 side chains are shown as sticks. **b**, Affinity determination of MCB_D2 using SPR. SPR sensorgram from a nine-point single-cycle kinetics experiment (twofold dilution, highest concentration: 20 μM). Experimental data are shown in purple and global fits are shown with black lines. The K_d is also shown on the plot. **c**, Experimentally determined complex structures closely match the design model. Overlap of the X-ray crystal structure (gold and gray) with the design model for MCB_D2 (purple). The $\text{C}\alpha$ r.m.s.d. for the macrocycle is 0.7 \AA when the experimental structure and design models are aligned by MCL1 residues. Close-up views demonstrate strong agreement between the side-chain rotamers of the design model and the X-ray structure. **d**, Overlay of the macrocycle model to the crystal structure shows a $\text{C}\alpha$ r.m.s.d. of 0.4 \AA with nearly

identical backbones and side-chain rotamers. **e**, Close-up view of the macrocycle-bound MCL1 structure showing the cation- π interaction at the interface. **f**, Close-up view of the macrocycle-bound MCL1 structure showing the hydrophobic contacts at the interface. **g**, AfCycDesign prediction of MDB_D8 design (blue) in complex with MDM2 (gray) shown as cartoons with interacting side chains shown as sticks, bound to MDM2 shown as surface. **h**, Affinity determination of MDB_D8 using SPR. SPR sensorgram from a nine-point single-cycle kinetics experiment (fivefold dilution, highest concentration: 50 μM). Experimental data are shown in blue and global fits are shown with black lines. The K_d is also shown on the plots. **i**, Overall and close-up views of the AfCycDesign prediction of the MDB_D8 design model, highlighting key interactions with the MDM2.

side-chain rotamers of the interacting residues in the crystal structure also closely matched the design model (Fig. 2d). The crystal structure also confirmed that the binding interactions are not restricted to the helix region of the designed macrocycle but are also contributed by the loop regions (Fig. 2e,f). While several hydrophobic interactions from the MCB_D2 helical segment are similar to those seen in natural MCL1 binders (for example, BH3 peptide), (Supplementary Fig. 7), the N-to-C orientation of the helix is flipped in the case of MCB_D2. The

loop region of MCB_D2 makes additional hydrophobic contacts and a cation- π interaction with MCL1 (Fig. 2e and Supplementary Fig. 7d) that we did not observe in previously reported natural MCL1 binders and their analogs. All three hits with an observable binding signal at 100 μM featured this cation- π interaction.

Encouraged by the experimental validation of the MCL1 binding macrocycles, we next sought to design binders to MDM2, an E3 ligase that interacts with tumor suppressor protein p53 and has multiple

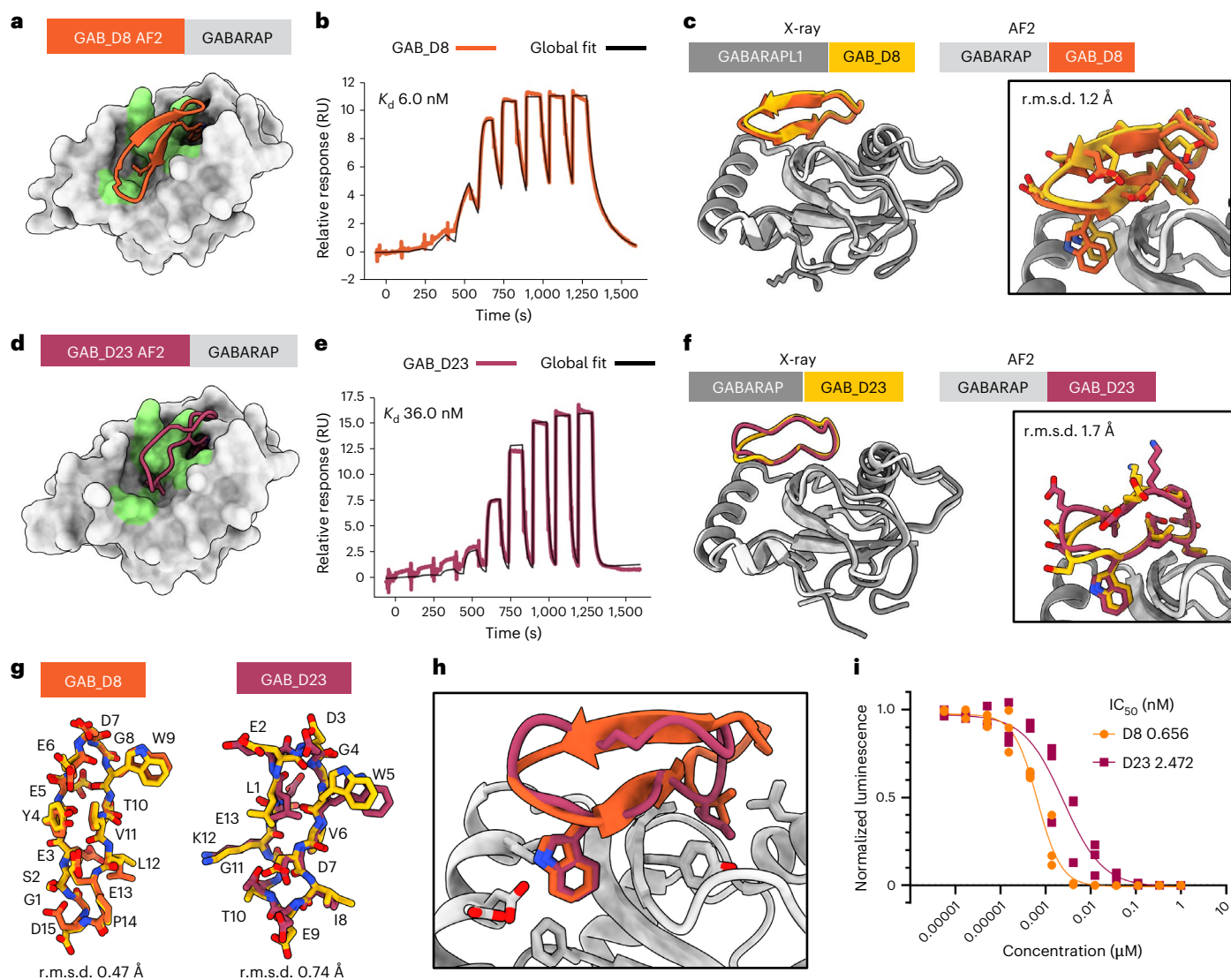


Fig. 3 | De novo design of high-affinity macrocycle binders to GABARAP.

a, AfCycDesign predicted model for design GAB_D8 bound to GABARAP shown as surface, with hotspot residues highlighted in green. **b**, Affinity determination of GAB_D8 using SPR. SPR sensorgram from a nine-point single-cycle kinetics experiment (fivefold dilution, highest concentration: 20 μ M). Experimental data are shown in orange and global fits are shown with black lines. The K_d is also shown on the plot. **c**, Superposition of chains E and F from the X-ray crystal structure of GAB_D8 bound to GABARAPL1 and the AfCycDesign model. **d**, AfCycDesign predicted model for design GAB_D23 bound to GABARAP shown as surface, with hotspot residues highlighted in green. **e**, Affinity determination

of GAB_D23 using SPR. SPR sensorgram from a nine-point single-cycle kinetics experiment (fivefold dilution, highest concentration: 20 μ M). Experimental data are shown in pink and global fits are shown with black lines. The K_d is also shown on the plot. **f**, Alignment of chains A and B from the X-ray crystal structure of GAB_D23 bound to GABARAP and the AfCycDesign model. **g**, Alignments of GAB_D8 and GAB_D23 macrocycle models to X-ray crystal structures show close matches. **h**, Comparison of GAB_D8 and GAB_D23 binding modes in the design models. **i**, Competitive AlphaScreen dose-response plot, IC_{50} from the average of three experiments. Donor and acceptor beads in the assay are bound to GABARAP and GABARAP-binding peptide K1, respectively.

critical roles in tumor growth and survival²⁵. We generated 10,000 macrocycle backbones spanning diverse lengths amenable to chemical synthesis (16–18 residues) and designed four amino acid sequences for each generated backbone using iterative rounds of ProteinMPNN and Rosetta Relax protocols (Methods). Design models were filtered on the basis of the confidence metrics and similarity of the AfCycDesign predictions to the designed complexes and the interface quality metrics calculated using Rosetta (Supplementary Fig. 4). AfCycDesign predicted 7,495 of the 40,000 design models to bind MDM2 with high confidence (normalized iPAE < 0.3) (Supplementary Fig. 4). In contrast to our approach for MCL1, we chose not to do any additional filtering with RF2 as the results between AfCycDesign and RF2 were fairly consistent. We also adjusted the filter thresholds for in silico filters as their overall distribution differed substantially from the distribution observed for

MCL1 (Supplementary Fig. 4). After filtering on interface metrics (Methods), we identified 17 designs with iPAE < 0.3, ddG < -50 kcal mol⁻¹, CMS > 300 Å² and SAP < 35. We selected 11 top-ranked designs by ddG for biochemical and biophysical characterization. The 11 selected designs had diverse sizes, shapes and sequences (Supplementary Fig. 8 and Supplementary Table 2); however, they were all predicted to bind the same site as the p53 transactivation domain (Supplementary Fig. 8). Three of the selected designs had poor yields during the cyclization step of the chemical synthesis, preventing further experimental characterization with them. We tested the remaining eight peptides for binding to the biotinylated MDM2 by SPR and identified three binders with observable binding signals at 100 μ M (Supplementary Fig. 9). The best design, MDB_D8 (Fig. 2g), demonstrated a binding affinity of 1.9 μ M in the SPR single-cycle kinetics experiment (Fig. 2h). The computational

model for this design makes several key contacts at the interface that are similar to interactions observed in native MDM2–p53 complex structures (Fig. 2i and Supplementary Fig. 10)²⁵. Despite different overall structures, all three hits from the SPR screen had a similar binding motif composed of phenylalanine, tryptophan and either leucine or methionine from the helical segment of the macrocycle. Together, these data highlight the promising accuracy of the RFpeptides pipeline to design diverse macrocyclic binders for selected targets of interest.

De novo design of macrocyclic binders to γ -aminobutyric acid type A receptor-associated protein

We next set out to design binders against a target with a binding site that is structurally different from MCL1 and MDM2, formed by a mix of α -helices and β -strands (in contrast to all α -helical pockets of MCL1 and MDM2). We selected γ -aminobutyric acid type A receptor-associated (GABARAP) as the target, a protein responsible for mediating autophagy through its role in autophagosome biogenesis and recruitment of cargo, resulting in lysosomal degradation of damaged or surplus proteins and organelles²⁶. Peptide modulators against GABARAP could have therapeutic applications in the treatment of late-stage cancers²⁷ or as chimeric peptides for autophagy-mediated targeted protein degradation²⁸. Our target binding site for GABARAP, which is also the binding site for the native LC3-interacting region or Atg8-interacting motif²⁹, is formed by a mix of β -strand and α -helix secondary structures (Fig. 3 and Supplementary Fig. 13). For designing macrocyclic binders against the human GABARAP, we used a similar pipeline as described above for MCL1 and MDM2 (Methods) but we doubled the number of generated designs and defined six hotspot residues (Lys46, Lys48, Tyr49, Leu50, Phe60 and Leu63) to guide the macrocycle backbone generation to a specific site on the target (Fig. 3a,d). We generated 20,000 macrocycle backbones and designed the amino acid sequences using ProteinMPNN and Rosetta Relax protocols. Of the resulting 80,000 design models, we selected 335 macrocyclic designs on the basis of AfCycDesign (iPAE < 0.13) and Rosetta (ddG < -30 kcal mol⁻¹, SAP < 35 and CMS > 300 Å²) interface metrics (Supplementary Fig. 4). Instead of trying to synthesize and characterize all 335 cyclic peptides (which would have required substantial time and experimental resources), we clustered the 335 designs into 80 different clusters on the basis of their three-dimensional structures and selected representative designs from diverse clusters for further biochemical characterization. We selected 13 diverse macrocycles of 12–17 residues for synthesis and experimental validation (Supplementary Table 3 and Supplementary Fig. 11). Unlike the design candidates described above for MCL1 and MDM2, several of the selected macrocycles for GABARAP showed cyclic β -sheet structures with several edge-strand interactions with the target (Supplementary Fig. 11).

We successfully synthesized six designs with high purity (>90%) and tested them for binding to GABARAP using SPR (Supplementary Fig. 12). Two designs, GAB_D8 and GAB_D23, showed binding affinities of 6 nM and 36 nM, respectively (Fig. 3b,e). To further characterize the binding of GAB_D8 and GAB_D23, we tested the ability of these designs to disrupt the interaction of GABARAP with linear peptide K1 (a previously described binder to this site³⁰) in AlphaScreen assays. GAB_D8 and GAB_D23 demonstrated a half-maximal inhibitory concentration (IC₅₀) of 0.7 nM and 2.5 nM in the AlphaScreen assay, respectively (Fig. 3i). To our knowledge, GAB_D8 is the most potent macrocyclic GABARAP binder to date.

In crystallization trials, we did not obtain crystals of sufficiently high quality for GAB_D8 bound to GABARAP. We instead crystallized GAB_D8 bound to GABARAP1, a homolog of GABARAP with 86% overall sequence identity and 100% sequence identity for residues within 5 Å of GAB_D8 in the design model. The X-ray crystal structure for GAB_D8 bound to GABARAP1 matched very closely with the design model, with a C α r.m.s.d. of 1.2 Å over the macrocycle when aligned by the target

protein to the closest of the four copies in the asymmetric unit (Fig. 3c and Supplementary Fig. 13) and a C α r.m.s.d. of 0.47 Å when aligned by macrocycle alone (Fig. 3g). Notably, the X-ray structure of the GAB_D8–GABARAP1 complex showed two different bound conformations of GAB_D8, one that closely matched the design model and a second one that partially deviated from the design model (Supplementary Fig. 14), with a register shift nucleated by Thr10 from the macrocycle forming main-chain-mediated and side-chain-mediated hydrogen bonds with Lys48 on the target. GAB_D23 crystallized readily with GABARAP and also closely matched the design model with a C α r.m.s.d. of 1.7 Å when aligned by the target (Fig. 3f) and C α r.m.s.d. of 0.74 Å across the macrocycle alone (Fig. 3g). The X-ray crystal structure confirmed the key designed interactions, such as Trp5 and Ile8, with the main difference between the design model and the X-ray structure being the switch from a type I β -turn from Leu1 to Gly4 in the design model to a less regular conformation in the crystal structure, with a tendency for a type I' β -turn from Glu2 to Trp5. While our original design models were predicted with single sequences as inputs to AF2, we retrospectively predicted the GAB_D8–GABARAP1 and GAB_D23–GABARAP complex structures with multiple-sequence alignment (MSA) inputs. These MSA-based predictions of the designs matched even more closely with the X-ray crystal structures, with a C α r.m.s.d. of 0.5 Å and 0.9 Å for the GAB_D8–GABARAP1 and GAB_D23–GABARAP complexes, respectively, when aligned by the target structure (Supplementary Fig. 15). Overall, these data demonstrate the ability of our de novo design pipeline to identify high-affinity binders against targets with diverse pocket shapes and surfaces without requiring library-scale screening.

Design of macrocyclic binders to predicted structures

Given the high accuracy and binding affinity of macrocycles designed against selected targets, we next set out to design macrocyclic binders against targets without any experimentally determined structure. We reasoned that the high accuracy of RFpeptides could mitigate the inherent risk of designing against a predicted target structure. We designed macrocycles against Rhombotarget A (RbtA), a recently identified cell surface protein from the ESKAPE pathogen, *Acinetobacter baumannii*. There are no experimentally determined structures available for this protein and sequence-based searches against the Protein Data Bank (PDB) did not return notable matches to other protein structures. We predicted the structure of the 617-aa full-length protein using AF2 and RF2; both methods predicted similar overall structures (C α r.m.s.d. of 0.4 Å over 509 residues excluding the signal peptide and transmembrane domain) with high confidence (predicted local distance difference test (pLDDT) > 90) (Supplementary Fig. 16). AF2 and RF2 both predicted two distinct extracellular domains: an N-terminal β -helix domain and a C-terminal Ig-like domain (Supplementary Fig. 16). While there were some differences in the predicted structures from AF2 and RF2, we decided to focus our binder design calculations on regions that were predicted nearly identically and with high confidence by AF2 and RF2. On the basis of our preliminary design runs without hotspots to guide the diffusion, we identified a patch in the N-terminal domain to pursue in our large-scale design calculations against this target and defined hotspots Leu144, Phe202, Phe204, Tyr206, Val208, Leu231 and Ala269 for peptide backbone generation (Fig. 4a). In contrast to the concave pockets targeted for MDM2 and MCL1, this selected patch for RbtA is considerably flatter and difficult to target with conventional computational and experimental approaches (Supplementary Fig. 17). We generated 20,000 backbones for macrocycle binders and designed four amino acid sequences for each backbone using iterative rounds of ProteinMPNN and Rosetta Relax. Designs were filtered using AfCycDesign confidence metrics and Rosetta interface metrics, as described in earlier sections (Supplementary Fig. 4). On the basis of these in silico metrics, we selected 26 designs for biochemical and structural

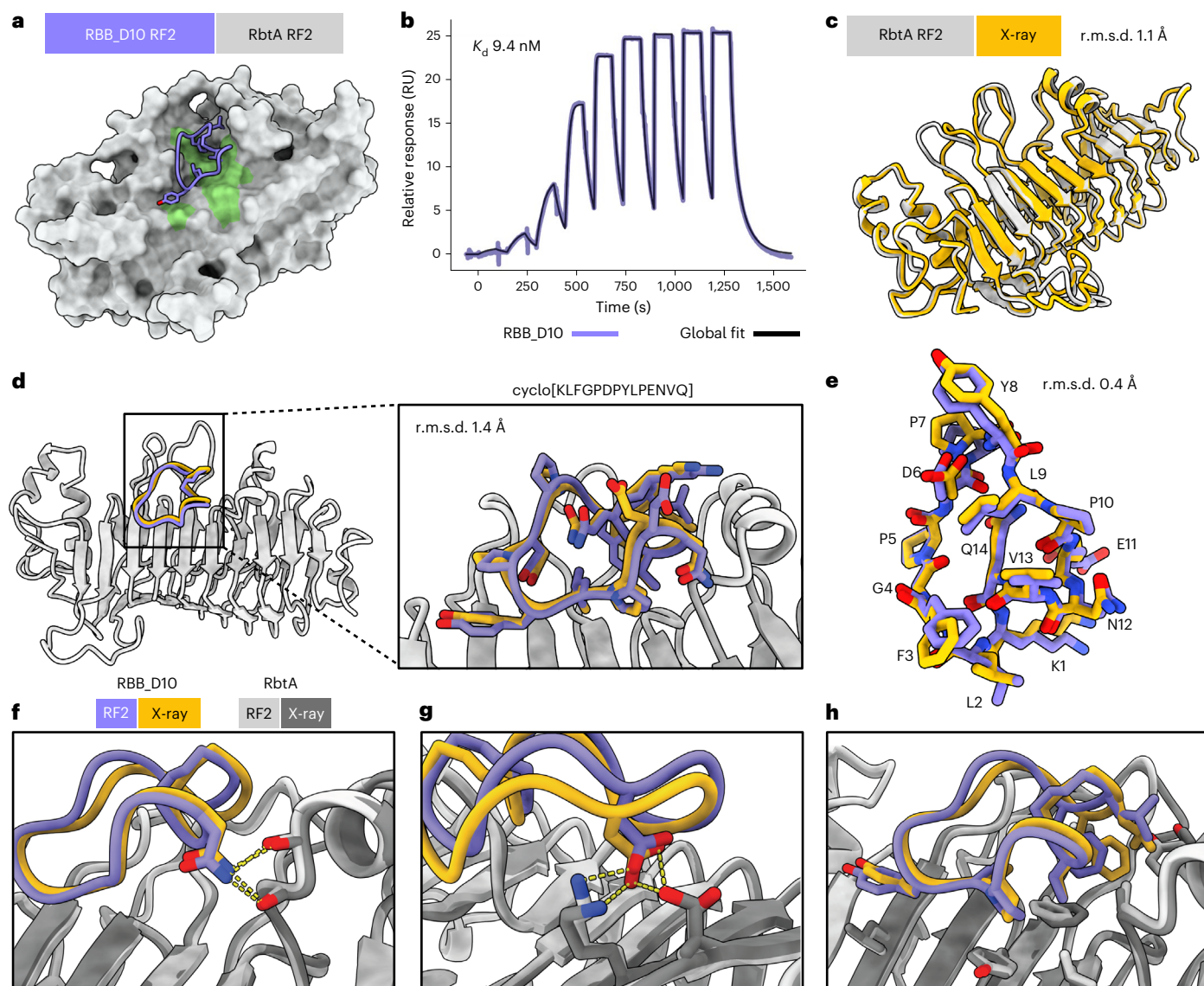


Fig. 4 | Accurate de novo design of a high-affinity cyclic peptide binder against the predicted structure of RbtA from *A. baumannii*. **a**, AfCycDesign prediction of design RBB_D10 (violet cartoon) bound to the AF2-predicted β -helix domain of RbtA shown as gray surface. Hotspot residues from RbtA used during the backbone design step are shown in green. **b**, SPR sensorgram from nine-point single-cycle kinetics experiment (fivefold dilution, highest concentration: 20 μ M). The K_d determined from the SPR experiment is also denoted on the plot. **c**, Close agreement of the RF2-predicted structure of RbtA (gray) with the X-ray structure (gold) of the RbtA N-terminal domain determined here confirms the predicted structure of the target used for the macrocycle design calculations. **d**, Alignment of the design model of RbtA-bound RBB_D10 (violet and gray) to the X-ray structure (gold) shows a close match between the design model and the

experimentally determined structure (Ca r.m.s.d. for macrocycle: 1.4 Å). Close-up view of the RbtA-bound RBB_D10 with side chains shown as sticks. **e**, Overlay of RBB_D10 design model (after the AfCycDesign prediction step) aligned to the X-ray structure without RbtA demonstrates a nearly identical match for backbone coordinates and side-chain rotamers (Ca r.m.s.d.: 0.4 Å). The design model and X-ray structure are shown in violet and gold, respectively. **f**, Close-up view of the macrocycle-bound RbtA structure and design model showing polar side chain-to-backbone interactions mediated by RBB_D10 residue Asn12 at the interface. **g**, Close-up view of the polar side chain-to-side chain interactions mediated by RBB_D10 residue Asp6 at the interface. **h**, Close-up view of the hydrophobic interactions between RbtA and RBB_D10 at the binding interface.

characterization with AfCycDesign iPAE < 0.4, ddG < -30 kcal mol⁻¹, r.m.s.d. between the design model and AfCycDesign prediction < 1.5 Å and CMS > 300 Å² (Supplementary Fig. 18). The selected designs covered diverse sizes (13–18 aa), sequences, shapes and secondary structures (Supplementary Fig. 18 and Supplementary Table 4). We expressed the Avi-tagged version of the RbtA N-terminal domain (residues 20–458) and used it for binding screens using SPR. Four of 11 designs that were synthesized in sufficient quantity and purity showed a binding signal at 100 μ M in our screens (Supplementary Fig. 19). On the basis of further binding experiments with SPR, we determined the dissociation constant (K_d) of the best binder, RBB_D10, to be 9.4 nM

(Fig. 4b). The design model for RBB_D10 showed extensive contacts to the target with several side-chain-mediated polar contacts and hydrophobic interactions (Fig. 4f–h).

To confirm the structures of RbtA and RBB_D10 and the binding mode between them, we determined the high-resolution X-ray crystal structure of apo and macrocycle-bound RbtA using X-ray crystallography at 2 Å and 2.6 Å resolution, respectively. The apo structure of the RbtA N-terminal domain, which is also the first experimentally determined structure from this class of bacterial proteins, matched our AF2 and RF2 predictions for this target very closely, with an overall Ca r.m.s.d. of 1.2 Å and 1.1 Å between the X-ray structure of the RbtA

N-terminal domain and the AF2-predicted and RF2-predicted structures, respectively (Fig. 4c). The complex structure also confirmed the structure and binding mode of our designed macrocycle, RBB_D10, with the X-ray structure matching the design model with an r.m.s.d. of 1.4 Å (Fig. 4d). Notably, the conformation adopted by the macrocycle in the X-ray structure, including the side-chain rotamers involved in interactions with the target, was almost identical to the design model with an r.m.s.d. of 0.4 Å (Fig. 4e–h). Together, these data highlight the high accuracy and success rates provided by RFpeptides even while designing macrocycles against targets without deep pockets or targets with no known structures.

Overall, these data show that RFpeptides can sample extensive structural and chemical diversity of macrocycles during the backbone and sequence generation steps against selected targets and, finally, select the shapes and sequences ideally suited for binding the target surface or pockets. The highest-affinity binders against each target are also predicted to fold into the bound conformations even in the absence of the target (Supplementary Fig. 20), suggesting that macrocycles are designed to fold into binding-competent conformations. For all four design campaigns described here, selected designs demonstrated good solubility in aqueous buffers despite not imposing any particular sequence constraints related to solubility during the sequence design step using ProteinMPNN²¹. Notably, combining DL-based and physics-based in silico filters helps to select medium to high-affinity binders. However, we note that the distribution of such metrics varies substantially across the four selected targets and adjustments to filtering thresholds were required on the basis of the shape and chemical composition of the target pocket. While in silico metrics enrich well for binders, the relative ranking within the selected designs does not perfectly match the experimental binding affinities. The highest-affinity binders for MDM2 and RbtA had the best or second-best iPAE values among the designs chosen for those targets (Supplementary Tables 2b and 4b); however, the hit peptides against MCL1 and GABARAP were not among the top three ranked designs (Supplementary Tables 1b and 3b). Integration with high-throughput methods in the future should enable testing of more designs and inform absolute threshold values and filtering schemes for the single-shot design of peptide binders to any arbitrary target.

Discussion

Here, we describe RFpeptides, a generative DL pipeline for precise de novo design of macrocycle binders against a wide range of protein targets. The power of the approach is highlighted by the high affinities ($K_d < 10$ nM) of the designed macrocyclic binders to GABARAP and RbtA and the nearly identical X-ray crystal structures and design models of the macrocycle-bound MCL1, GABARAP and RbtA (C α r.m.s.d. of 0.7 Å, 1.2 Å and 1.4 Å, respectively). The RFpeptides approach offers several advantages over traditional methods. Firstly, the design approach should enable faster and more efficient discovery of macrocyclic binders. Despite testing fewer than 20 designed candidates per target (in contrast to trillions of peptides tested in traditional library-based approaches), we achieved high-affinity binders for two targets without requiring any further experimental optimization; to our knowledge, this is a considerably higher success rate than achieved with any previous method. Secondly, in contrast to the untargeted nature of the random library-based approaches, RFpeptides can be used for designing custom binders to specific patches and sites, as demonstrated for GABARAP and RbtA. Lastly, the atomically accurate nature of the design models enables structure-guided optimization for properties beyond target binding (as well as further increases in affinity), bypassing the bottleneck of complex structure determination, which has hindered the optimization of leads from library screening. Combined with the design principles for membrane traversal, RFpeptides could enable the design of peptides simultaneously optimized for target binding and cell permeability or oral bioavailability.

RFpeptides also has considerable advantages over previous computational peptide design methods. Information on known ligands and/or binding partners is not required to initiate design. RFpeptides can design macrocycles completely de novo from just the structure or sequence (as in the case of RbtA) of the target, enabling design against molecular targets intractable with previous methods. RFpeptides is not limited to generating macrocycles with particular motifs or topologies; the diffusion process generates macrocycles with diverse shapes and sizes and selects the topologies appropriate for the protein being targeted. Among the four targets tested here, binders for MCL1 and MDM2 have helical motifs, binders for GABARAP have a β -sheet topology and binders for RbtA sample looplike conformations that make extensive contacts with the flat surface of this target.

We anticipate that RFpeptides will enable the rapid design of custom macrocyclic binders against a wide range of molecular targets, accelerating efforts to develop peptides for diverse functional applications. With the rapid advances in DL methods and frameworks, including the recent development of all-atom diffusion models, we aim to extend the approach to generative design of macrocycles with noncanonical amino acids, crosslinkers and cyclization chemistries.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41589-025-01929-w>.

References

1. Vinogradov, A. A., Yin, Y. & Suga, H. Macrocyclic peptides as drug candidates: recent progress and remaining challenges. *J. Am. Chem. Soc.* **141**, 4167–4181 (2019).
2. Muttenthaler, M., King, G. F., Adams, D. J. & Alewood, P. F. Trends in peptide drug discovery. *Nat. Rev. Drug Discov.* **20**, 309–325 (2021).
3. Tsomaia, N. Peptide therapeutics: targeting the undruggable space. *Eur. J. Med. Chem.* **94**, 459–470 (2015).
4. Atanasov, A. G., Zotchev, S. B. & Dirsch, V. M. International Natural Product Sciences Taskforce & Supuran, C. T. Natural products in drug discovery: advances and opportunities. *Nat. Rev. Drug Discov.* **20**, 200–216 (2021).
5. Bhardwaj, G. et al. Accurate de novo design of membrane-traversing macrocycles. *Cell* **185**, 3520–3532 (2022).
6. Bhardwaj, G. et al. Accurate de novo design of hyperstable constrained peptides. *Nature* **538**, 329–335 (2016).
7. Hosseinzadeh, P. et al. Comprehensive computational design of ordered peptide macrocycles. *Science* **358**, 1461–1466 (2017).
8. Mulligan, V. K. et al. Computationally designed peptide macrocycle inhibitors of New Delhi metallo- β -lactamase 1. *Proc. Natl Acad. Sci. USA* **118**, e2012800118 (2021).
9. Hosseinzadeh, P. et al. Anchor extension: a structure-guided approach to design cyclic peptides targeting enzyme active sites. *Nat. Commun.* **12**, 3384 (2021).
10. Muratspahić, E. et al. Design and structural validation of peptide–drug conjugate ligands of the κ -opioid receptor. *Nat. Commun.* **14**, 8064 (2023).
11. Rettie, S. A. et al. Cyclic peptide structure prediction and design using AlphaFold2. *Nat. Commun.* **16**, 4730 (2025).
12. Grambow, C. A., Weir, H., Cunningham, C. N., Biancalani, T. & Chuang, K. V. CREMP: conformer–rotamer ensembles of macrocyclic peptides for machine learning. *Sci. Data* **11**, 859 (2024).
13. Zhang, C. et al. HighFold: accurately predicting structures of cyclic peptides and complexes with head-to-tail and disulfide bridge constraints. *Brief. Bioinform* **25**, bbae215 (2024).

14. Brixi, G. et al. SaLT&PepPr is an interface-predicting language model for designing peptide-guided protein degraders. *Commun. Biol.* **6**, 1081 (2023).
15. Xie, X., Valiente, P. A., Kim, J. & Kim, P. M. HelixDiff, a score-based diffusion model for generating all-atom α -helical structures. *ACS Cent. Sci.* **10**, 1001–1011 (2024).
16. Li, Q., Vlachos, E. N. & Bryant, P. Design of linear and cyclic peptide binders of different lengths from protein sequence information. Preprint at *bioRxiv* <https://doi.org/10.1101/2024.06.20.599739> (2024).
17. Watson, J. L. et al. De novo design of protein structure and function with RFDiffusion. *Nature* **620**, 1089–1100 (2023).
18. Baek, M. et al. Accurate prediction of protein structures and interactions using a three-track neural network. *Science* **373**, 871–876 (2021).
19. Vázquez Torres, S. et al. De novo design of high-affinity binders of bioactive helical peptides. *Nature* **626**, 435–442 (2024).
20. Krishna, R. et al. Generalized biomolecular modeling and design with RoseTTAFold All-Atom. *Science* **384**, ead12528 (2024).
21. Dauparas, J. et al. Robust deep learning-based protein sequence design using ProteinMPNN. *Science* **378**, 49–56 (2022).
22. Wicky, B. I. M. et al. Hallucinating symmetric protein assemblies. *Science* **378**, 56–61 (2022).
23. Widden, H. & Placzek, W. J. The multiple mechanisms of MCL1 in the regulation of cell fate. *Commun. Biol.* **4**, 1029 (2021).
24. Leman, J. K. et al. Macromolecular modeling and design in Rosetta: recent methods and frameworks. *Nat. Methods* **17**, 665–680 (2020).
25. Kussie, P. H. et al. Structure of the MDM2 oncoprotein bound to the p53 tumor suppressor transactivation domain. *Science* **274**, 948–953 (1996).
26. Szalai, P. et al. Autophagic bulk sequestration of cytosolic cargo is independent of LC3, but requires GABARAPs. *Exp. Cell Res.* **333**, 21–38 (2015).
27. Brown, H. et al. Structure-based design of stapled peptides that bind GABARAP and inhibit autophagy. *J. Am. Chem. Soc.* **144**, 14687–14697 (2022).
28. Ji, C. H. et al. The AUTOTAC chemical biology platform for targeted protein degradation via the autophagy–lysosome system. *Nat. Commun.* **13**, 904 (2022).
29. Popelka, H. & Klionsky, D. J. Analysis of the native conformation of the LIR/AIM motif in the Atg8/LC3/GABARAP-binding proteins. *Autophagy* **11**, 2153–2159 (2015).
30. Weiergräber, O. H. et al. Ligand binding mode of GABA_A receptor-associated protein. *J. Mol. Biol.* **381**, 1320–1331 (2008).
31. van der Maaten, L. & Hinton, G. Visualizing data using t-SNE. *J. Mach. Learn. Res.* **9**, 2579–2605 (2008).
32. Pedregosa, F. et al. Scikit-learn: machine learning in Python. *J. Mach. Learn. Res.* **12**, 2825–2830 (2011).
33. Zhang, Y. & Skolnick, J. TM-align: a protein structure alignment algorithm based on the TM-score. *Nucleic Acids Res* **33**, 2302–2309 (2005).
34. Dauparas, J. et al. Atomic context-conditioned protein sequence design using LigandMPNN. *Nat. Methods* **22**, 717–723 (2025).
35. Sternberg, A. MaxCluster: a tool for protein structure comparison and clustering. <http://www.sbg.bio.ic.ac.uk/maxcluster/index.html> (accessed 1 November 2024).
36. Siew, N., Elofsson, A., Rychlewski, L. & Fischer, D. MaxSub: an automated measure for the assessment of protein structure prediction quality. *Bioinformatics* **16**, 776–785 (2000).

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

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

© The Author(s) 2025

¹Department of Medicinal Chemistry, University of Washington, Seattle, WA, USA. ²Institute for Protein Design, University of Washington, Seattle, WA, USA. ³Molecular and Cellular Biology Program, University of Washington, Seattle, WA, USA. ⁴Graduate Program in Molecular Engineering, University of Washington, Seattle, WA, USA. ⁵Department of Biochemistry, University of Washington, Seattle, WA, USA. ⁶Cancer Research @UCC, University College Cork, Cork, Ireland. ⁷Department of Microbiology, University of Washington, Seattle, WA, USA. ⁸Department of Chemistry, Tufts University, 62 Talbot Avenue, Medford, MA, USA. ⁹Heinrich–Heine–Universität Düsseldorf, Institut für Physikalische Biologie, Düsseldorf, Germany. ¹⁰Forschungszentrum Jülich, Institute of Biological Information Processing, Structural Biochemistry (IBI–7), Jülich, Germany. ¹¹Department of Chemistry, University of Washington, Seattle, WA, USA. ¹²Department of Biology, Massachusetts Institute of Technology, Cambridge, MA, USA. ¹³Howard Hughes Medical Institute, University of Washington, Seattle, WA, USA. ¹⁴Present address: Molecular Cell Biology of Autophagy Laboratory, The Francis Crick Institute, London, UK. ¹⁵These authors contributed equally: Stephen A. Rettie, David Juergens, Victor Adebomi. ✉ e-mail: dabaker@uw.edu; dimaio@uw.edu; gauravb@uw.edu

Methods

Computational methods for cyclic peptide binder design

Macrocytic peptide monomers and binders were designed with RFpeptides using a three-stage pipeline: backbone generation using RFdiffusion with the cyclic offset applied to the peptide chains, followed by sequence design using ProteinMPNN and, finally, structure prediction of the designed peptide–target complexes using either AfCycDesign and/or RoseTTAFold with the cyclic offset applied to the peptide. Designs were further filtered and downselected using Rosetta metrics and, in some cases, clustered on the basis of C α r.m.s.d. Detailed computational methods, including example scripts, can be found in Supplementary Section 2.2.

Peptide synthesis

Macrocytic peptides described here were either purchased from Wuxi AppTec at greater than 90% purity or synthesized in-house using Fmoc-based solid-phase peptide synthesis. Peptides were typically synthesized on preloaded CTC resin. The resin was swollen in DCM followed by iterative deprotection with 20% piperidine in DMF and coupling with either HBTU (Sigma) or PyAOP (Novabiochem) and DIEA (Sigma). The linear peptides were cleaved from the resin using either 2% TFA in DCM or 20% HFIP (Oakwood Chemical) in DCM. The solvent was removed by rotary evaporation and linear protected peptides were cyclized in either DCM, DMF or a mixture of both depending on the solubility of the peptide, using two equivalents of PyAOP and five equivalents of DIEA overnight. The protecting groups were removed using a cocktail of 95:2.5:2.5, TFA, water and TIPS for 2.5 h. The crude peptides were precipitated using cold diethyl ether. The precipitate containing the crude cyclization reaction was dissolved in a mixture of water and acetonitrile for purification using reverse-phase high-performance liquid chromatography (LC). Peptide identities were confirmed by mass spectrometry (MS). Purities for all synthesized and tested macrocytic peptides are also summarized in Supplementary Tables 7–10. The mass spectrograms and analytical LC chromatograms for all purified peptides are shown in Supplementary Section 4.

Protein expression and purification

MDM2 and MCL1. The amino acid sequences of MCL1 (PDB 2PQK)³⁷ and MDM2 (PDB 4HFZ)³⁸ were retrieved from the PDB. The optimized genes were then cloned into a Novogen pRSF-DUET plasmid (Sigma, 71341-3), incorporating a 6xHis-tag at the N terminus, followed by an Avi-tag and a tobacco etch virus (TEV) protease cleavage site. The resulting constructs were codon-optimized for *Escherichia coli* expression and synthesized by GenScript. For propagation, the plasmids were transformed into *E. coli* NEB α cells (New England Biolabs, C2987); for protein expression, the plasmids were transformed into *E. coli* BL21(DE3) cells (New England Biolabs, C2527). A single sequence-verified colony was cultured in 50 ml of kanamycin (50 μ g ml⁻¹) selective Luria Broth (LB) medium. This culture was incubated at 37 °C with shaking at 200 rpm for 16 h overnight. Subsequently, 50 units of optical density at 600 nm (OD₆₀₀) of the overnight culture were transferred to 1 L of fresh kanamycin (50 μ g ml⁻¹) selective LB medium. The culture was grown at 37 °C with shaking at 200 rpm for 2 h (until it reached an OD₆₀₀ of 0.4–0.5), at which point the temperature was decreased to 20 °C. The culture was grown until an OD₆₀₀ of 0.7–0.8; protein expression was induced by adding 1 mM IPTG and the culture was left to grow overnight for 14 h.

Cells were harvested by centrifugation at 5,000g for 10 min at 4 °C, resulting in a cell pellet with a density of 5 g L⁻¹. The pellet was immediately flash-frozen and stored at –20 °C for later use. For lysis, the pellet was thawed on ice and resuspended in 5 ml of lysis buffer per gram of pellet. This lysis buffer contained 50 mM Tris-HCl, 300 mM NaCl and 10 mM imidazole and was supplemented with 1 \times BugBuster protein extraction reagent (Sigma-Aldrich, 70921), 200 μ g ml⁻¹ lysozyme (Sigma-Aldrich, L6876), 25 U per ml benzonase nuclease (Sigma-Aldrich, E8263) and 1 \times cOmplete EDTA-free protease inhibitor

cocktail (Sigma-Aldrich, I1836170001). The buffer was filter-sterilized using a 0.2 μ m filter before the addition of benzonase, mixed by inversion and kept on ice until use. Cells were completely resuspended in the lysis buffer using a homogenizer at low speed and incubated for 30 min at room temperature (22–25 °C). Following incubation, the suspension was sonicated using a Q500 Sonicator equipped with a four-tip probe. Sonication was conducted for 2–3 min using pulses of 10–15 s on followed by 10–15 s off at 70% amplitude. The lysate was clarified by centrifugation at 16,000g for 20 min.

Ni-NTA agarose resin (Qiagen, 30210) was equilibrated with 20 column volumes (CV) of ultrapure water, followed by 20 CV of equilibration buffer (50 mM Tris-HCl, 300 mM NaCl and 10 mM imidazole). Then, 4 ml of 50% resin suspended in equilibration buffer was used to bind His-tagged proteins from 25 ml of clarified lysate. All immobilized metal affinity chromatography (IMAC) steps were conducted at 4 °C. The lysate–resin mixture was incubated for 60 min on a rotary shaker set to a slow speed. After incubation, the resin was transferred to a 20-ml gravity column and allowed to completely settle. The resin was first washed with 20 CV of wash buffer 1 (20 mM Tris-HCl, 250 mM NaCl, 10 mM imidazole and 5 mM β -mercaptoethanol), followed by another 20 CV of wash buffer 2 (20 mM Tris-HCl, 500 mM NaCl and 35 mM imidazole). The bound proteins were then eluted with 8 ml of elution buffer (20 mM Tris-HCl, 250 mM NaCl, 350 mM imidazole and 2 mM DTT). Aliquots of the eluate were collected and analyzed using SDS–PAGE gels.

The eluate was loaded onto a pre-equilibrated Superdex 75 10/300 GL column (25 mM Tris-HCl, 250 mM NaCl and 2 mM DTT) and run at a flow rate of 0.6 ml min⁻¹ using an ÄKTA pure system for size-exclusion chromatography (SEC). Then, 1 ml fractions were collected from the elution volume of 8–16 ml and those corresponding to peaks in the absorbance at 280 nm between an elution volume of 10 and 13 ml were assessed with SDS–PAGE gels. Fractions confirming the expected molecular weight were pooled and concentrated by centrifugation at 4,000g for 30 min at 4 °C using Amicon Ultra-4 concentrators with a 3 kDa cutoff (Millipore Sigma, UFC800308) to a final volume of 500 μ l. The identity of the eluted proteins were confirmed by MS using an Agilent 6230 LC–MS time-of-flight system.

Verified protein samples were processed for further applications: biotinylation for SPR analysis or tag removal by TEV protease cleavage for crystallography. Biotinylation was performed using the BirA biotin–protein ligase standard reaction kit (Avidity, BirA-500) according to the manufacturer's recommended conditions. The reaction was carried out at 4 °C overnight on a slowly shaking platform. For TEV protease cleavage, the proteins were treated with a 25:1 protein to TEV enzyme ratio³⁹. Similarly, the mixture was incubated at 4 °C overnight on a slowly shaking platform. Following these treatments, samples underwent a cleanup step using 1 ml of Ni-NTA resin per 20 mg of protein. The resin was pre-equilibrated with 10 CV of ultrapure water and 10 CV of a buffer containing 25 mM Tris-HCl, 250 mM NaCl and 10 mM imidazole. The pre-equilibrated resin was added to the protein mixture and incubated for 30 min on a rolling platform at 4 °C. Subsequently, the mixtures were filtered through a 0.45 μ m PVDF centrifugal filtering unit to remove the Ni-NTA-bound substrates. The eluate was collected and dialyzed in 2 L of 25 mM Tris-HCl, 250 mM NaCl and 2 mM DTT using a Slide-A-Lyzer G3 dialysis cassettes with a 3.5 kDa molecular weight cutoff (Thermo Scientific, A52966) overnight for 18 h at 4 °C stirring. The dialyzed protein was concentrated to 0.2–0.5 ml (as required for downstream assays), using the Amicon ultra concentrators (as above), aliquoted and flash-frozen. Fractions were analyzed by mass spectroscopy for the efficacy of the biotinylation and TEV protease cleavage treatments, as previously described.

GABARAP for SPR. A synthetic complementary DNA was designed on the basis of the amino acid sequence of GABARAP (UniProt O95166) and optimized for expression in *E. coli* using Benchling software.

Diffraction-quality crystals for the complex appeared in 0.2 M sodium chloride, 0.1 M Bis-Tris pH 6.5 and 25% (w/v) polyethylene glycol 3350 (Hampton Research) in 2 weeks.

GABARAP and GABARAPL1 with cyclic peptides. Cyclic peptides GAB_D8 and GAB_D23 were dissolved in 10 mM Tris-HCl and 150 mM NaCl (pH 7.0) and each mixed with both GABARAP and GABARAPL1, targeting a peptide-to-protein molar ratio of 3:2. After incubation for 10 min at room temperature, any insoluble components were removed by centrifugation (10 min at 20,000g and 4 °C). The protein-peptide complexes were concentrated using Amicon Ultra-0.5 centrifugal filter units with a 3 kDa cutoff (Merck) until a final protein concentration of 6–8 mg ml⁻¹ (GABARAPL1-GAB_D8) or 13–15 mg ml⁻¹ (GABARAP-GAB_D23) was reached. Samples were once again cleared of particles by centrifugation (30 min at 20,000g and 4 °C) before application in crystallization experiments. Search for crystallization conditions was performed by the sitting-drop vapor diffusion method using robotic systems Freedom Evo (Tecan) and Mosquito LCP (SPT Labtech) with commercially available screening sets. Experiments were set up by combining 200 nl of protein-peptide complex with 100 nl (for GABARAPL1-GAB_D8) or 200 nl (for GABARAP-GAB_D23) of reservoir solution and plates were incubated at 20 °C. Crystals appeared for a number of conditions, which were subjected to optimization as appropriate. Diffraction-quality samples used for X-ray structure determination developed with reservoir solutions containing 0.17 M ammonium sulfate, 25.5% (w/v) PEG 4000 and 15% (v/v) glycerol for GABARAPL1-GAB_D8 and 0.1 M MES pH 5.0 and 30% (w/v) PEG 6000 in the case of GABARAP-GAB_D23. Diffraction data (<https://doi.esrf.fr/10.15151/ESRF-DC-1966164200> and <https://doi.esrf.fr/10.15151/ESRF-DC-1979522808> for GABARAPL1-GAB_D8 and GABARAP-GAB_D23, respectively) were collected at 100 K on beamline BM07 of the European Synchrotron Radiation Facility (ESRF) tuned to an X-ray wavelength of 0.9795 Å, using a Pilatus 6M detector (DECTRIS). Data processing was carried out with XDS and XSCALE⁴³ and included reflections up to a diffraction limit of 1.5 Å for GABARAP-GAB_D23 and 2.5 Å for GABARAPL1-GAB_D8. The GABARAP-GAB_D23 structure featuring space group C2 was determined by molecular replacement (MR) using MOLREP⁴⁴ with the structure of GABARAP from its K1 peptide complex (PDB 3D32)³⁰ as a template. For the GABARAPL1-GAB_D8 complex, initial evaluation suggested tetragonal symmetry but with strong indications of twinning. Data integration in maximal translationengleiche subgroups followed by MR search using MoRDa⁴⁵ revealed P2₁2₁2₁ as the true space group, with near-perfect pseudomero-hedral twinning accounting for apparent Laue group 4/mmm. To avoid bias in cross-validation, this pseudosymmetry of the data was explicitly accounted for in flag assignment. The solution obtained for GABARAPL1-GAB_D8 was subjected to a round of automated rebuilding in phenix.autobuild⁴⁶. In either case, model refinement was performed with phenix.refine⁴⁷, alternating with interactive rebuilding in Coot⁴⁸, which included stepwise introduction of cyclic peptides GAB_D8 and GAB_D23. According to validation using MolProbity⁴⁹ and the wwPDB validation system (<https://validate-rcsb-2.wwpdb.org/>), both models featured good geometry. Detailed statistics of data collection and refinement can be found in Supplementary Table 6.

RbtA with cyclic peptide and apo RbtA. RbtA (10 mg ml⁻¹) and RBB_D10 were mixed in a 1:5 molar ratio and incubated for 30 min at room temperature. Initial crystallization trials were set up in 200 nl drops using 96-well crystallization plates and the experiments were conducted by the sitting-drop vapor diffusion method. Crystal drops were imaged using the UVEX crystal plate hotel system by JANSi. Diffraction-quality crystals for the RbtA-RBB_D10 complex appeared in 0.2 M lithium sulfate, 0.1 M Tris pH 8.5 and 40% (v/v) PEG 400 (JCSG Plus, Hampton Research). Additionally, we soaked the crystals in 22.32 mg ml⁻¹ RBB_D10 for 5 min before flash-freezing. Crystals for

RbtA alone (18.7 mg ml⁻¹) were grown in 0.1 M Bis-Tris pH 6.5 and 20% (v/v) PEG 5,000 MME (SG1, Molecular Dimensions). All crystals were flash-cooled in liquid nitrogen before shipping to the synchrotron for data collection.

Diffraction data were collected at the NSLS2 beamline AMX/FMX (17-ID-1/17-ID-2). X-ray intensities and data reduction were evaluated and integrated by XDS⁴³ and merged and scaled by Pointless and Aimless in the CCP4i2 program suite⁵⁰. The X-ray crystal structure was determined by MR using the designed model for phasing by Phaser⁵¹. Next, the structure obtained from the MR was improved and refined by Phenix⁴⁷. Model building was performed by Coot⁴⁸ in between the refinement cycles. The final model was evaluated by MolProbity⁴⁹. Data collection and refinement statistics are reported in Supplementary Table 5.

SPR

SPR experiments were performed using a Cytiva Biacore 8K in HBS-EP+ buffer from Cytiva. Measurements were obtained by immobilization of biotinylated target protein using the biotin capture kit from Cytiva. Binding screens were performed by single-cycle kinetics experiments using the standard protocol in the Biacore 8K control software at 30 μl min⁻¹ with serial injections of 10 nM, 100 nM, 1 μM, 10 μM and 100 μM, an association time of 60 s and a dissociation time of 120 s. For MCL1 designs, a dissociation time of 150 s was used. To evaluate the affinity of successful designs, a nine-point single-cycle kinetics experiment was performed with an association time of 90 s and dissociation time of 300 s. The dilution series for MCB_D2 was twofold starting at 20 μM, that for MDB_D8 was fivefold starting at 50 μM, and those for GAB_D8, GAB_D23 and RBB_D10 were fivefold starting at 20 μM. Reported measurements were analyzed using Biacore Insight evaluation software; sensorgrams were double-referenced and fit with a 1:1 binding kinetics fit model.

AlphaScreen assay

We used the AlphaScreen assay as described by Leveille et al.⁵² to measure inhibition of the GABARAP-K1 interaction by the computationally designed macrocycles. K1 is a previously described GABARAP binder with a K_d of 10 nM (ref. 27). Biotin-labeled peptide K1 was used at a final concentration of 10 nM and incubated with 10 nM (final concentration) of 6xHis-GABARAP in a final reaction volume of 50 μl. Computationally designed inhibitor peptides were serially diluted with 1:3 dilutions using the highest final concentration of 50 μM and added to the reaction mixture. The buffer used was 25 mM HEPES pH 7.3, 150 mM NaCl, 0.01% Tween, 1 mg ml⁻¹ BSA and 0.5% DMSO. The plate was covered in foil, centrifuged at 1,500 rpm for 2 min and incubated for 150 min at room temperature with shaking. Then, 20 μg ml⁻¹ (final concentration) of the streptavidin donor beads and nickel chelate acceptor beads were added in the dark before incubating for another 45 min. Data were collected on a Tecan plate reader using excitation at 680 nm and emission at 520–620 nm. Data were normalized to 0% (buffer only) and 100% (protein and tracer peptide, no inhibitor) controls. IC₅₀ values were obtained from curve fits using GraphPad Prism 9 software, using the equation $Y = \frac{100}{1 + (\frac{X}{IC_{50}})^n}$, where X is the concentration of inhibitor and h

is the Hill coefficient. At least three independent replicates were used to calculate the average IC₅₀ and the s.e.m.

Statistics and reproducibility

No statistical method was used to predetermine sample size. One trial from the AlphaScreen that was used to determine the IC₅₀ of GAB_D8 was repeated and the repeated value is what was used. All data are included in the Source Data. The experiments were not randomized. The investigators were not blinded to allocation during experiments and outcome assessment.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The design models and sequences are available in Supplementary Information. Crystal structures of MCB_D2 bound to MCL1, GAB_D8 bound to GABARAPL1, GAB_D23 bound to GABARAP, RBB_D10 bound to RbtA and apo RbtA were deposited to the PDB under accession codes [9CDT](#), [9HGC](#), [9HGD](#), [9CDU](#) and [9CDV](#), respectively. Source data are provided with this paper.

Code availability

The code and scripts for running RFpeptides are available from Zenodo (<https://doi.org/10.5281/zenodo.15264344>)⁵³. The code and scripts for RFpeptides are also available from RFDiffusion GitHub repository (<https://github.com/RosettaCommons/RFDiffusion>).

References

- Fire, E., Gullá, S. V., Grant, R. A. & Keating, A. E. Mcl-1–Bim complexes accommodate surprising point mutations via minor structural changes. *Protein Sci.* **19**, 507–519 (2010).
- Anil, B., Riedinger, C., Endicott, J. A. & Noble, M. E. M. The structure of an MDM2–Nutlin-3a complex solved by the use of a validated MDM2 surface-entropy reduction mutant. *Acta Crystallogr. D* **69**, 1358–1366 (2013).
- Sumida, K. H. et al. Improving protein expression stability and function with ProteinMPNN. *J. Am. Chem. Soc.* **146**, 2054–2061 (2024).
- Bhandari, B. K., Gardner, P. P. & Lim, C. S. Solubility-Weighted Index: fast and accurate prediction of protein solubility. *Bioinformatics* **36**, 4691–4698 (2020).
- Davis, G. D., Elisee, C., Newham, D. M. & Harrison, R. G. New fusion protein systems designed to give soluble expression in *Escherichia coli*. *Biotechnol. Bioeng* **65**, 382–388 (1999).
- Dang, B. et al. SNAC-tag for sequence-specific chemical protein cleavage. *Nat. Methods* **16**, 319–322 (2019).
- Kabsch, W. XDS. *Acta Crystallogr. D* **66**, 125–132 (2010).
- Vagin, A. & Teplyakov, A. Molecular replacement with MOLREP. *Acta Crystallogr. D* **66**, 22–25 (2010).
- Vagin, A. & Lebedev, A. MoRDa, an automatic molecular replacement pipeline. *Acta Crystallogr. A Found. Adv.* **71**, s19 (2015).
- Terwilliger, T. C. et al. Iterative model building structure refinement and density modification with the PHENIX AutoBuild wizard. *Acta Crystallogr.* **64**, 61–69 (2008).
- Liebschner, D. et al. Macromolecular structure determination using X-rays neutrons and electrons: recent developments in Phenix. *Acta Crystallogr. D Struct. Biol.* **75**, 861–877 (2019).
- Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D* **66**, 486–501 (2010).
- Williams, C. J. et al. MolProbity: more and better reference data for improved all-atom structure validation. *Protein Sci.* **27**, 293–315 (2018).
- Winn, M. D. et al. Overview of the CCP 4 suite and current developments. *Acta Crystallogr. D* **67**, 235–242 (2011).
- McCoy, A. J. et al. Read Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
- Leveille, A. N. et al. Exploring arylidene–indolinone ligands of autophagy proteins LC3B and GABARAP. *ACS Med. Chem. Lett.* **16**, 271–277 (2025).
- Rettie, S. et al. Accurate de novo design of high-affinity protein-binding macrocycles using deep learning. *Zenodo* <https://doi.org/10.5281/zenodo.15264344> (2025).
- Hunter, J. D. Matplotlib: a 2D graphics environment. *Comput. Sci. Eng* **9**, 90–95 (2007).
- Waskom, M. seaborn: statistical data visualization. *J. Open Source Softw* **6**, 3021 (2021).
- Pettersen, E. F. et al. UCSF ChimeraX: structure visualization for researchers, educators, and developers. *Protein Sci* **30**, 70–82 (2021).

Acknowledgements

We thank L. Stuart, L. Stewart, K. van Wormer, L. Goldschmidt, M. Kennedy, I. Haydon, J. Woodward, X. Li, Z. Taylor, H. Osterstock, G. Zhou, G. Gökçe, J. Palmer, K. Lindenauer, K. Campbell, M. Gloegl, R. Ragotte and M. Sadilek for their helpful feedback, guidance and support. We also thank the IPD Peptide, Crystallography and Biologics, Vaccines and Process Development core labs and the University of Washington Chemistry MS Facility for providing instrumentation support and expertise. This work was supported by funds from the DARPA Harnessing Enzymatic Activity for Lifesaving Remedies program HRO01120S0052 contract HRO011-21-2-0012 (to G.B., D.B., J.D.M. and M.L.), the Defense Threat Reduction Agency HDTRA1-19-1-0003 (to D.B., G.B. and S.A.R.), the National Institutes of Health (NIH) 5R21AI178088-02 (to G.B. and S.A.R.), the Howard Hughes Medical Institute (HHMI) Emerging Pathogens Initiative (to J.D.M., G.B. and V.A.), startup funds from the University of Washington's Department of Medicinal Chemistry and Institute for Protein Design (to G.B.), the Audacious Project (to G.B., A.K.B. and A.K.), the C19 HHMI Initiative grant (to A.K.B. and A.K.), NIH R35-GM148407 (to J.A.K. and A.N.L.), the European Union's Horizon Europe research and innovation program under the Marie Skłodowska-Curie 101059124 (to Y.F.B.), NIH R01-ROAI160052 (to A.K.B. and A.K.), Deutsche Forschungsgemeinschaft project ID 267205415-SFB 1208 (to D.W.) and the Bill and Melinda Gates Foundation GR047983 (to D.J.). All plots in this paper were generated using matplotlib or seaborn^{54,55}. Peptide structures were rendered using ChimeraX 1.9 (ref. 56) or PyMOL 2.5.4. Data were analyzed and plotted with Pandas versions 1.4.3 and plotted using matplotlib version 3.7.0 and seaborn version 0.12.2. All figures were created with [BioRender.com](#). This research used resources (FMX/AMX) of the National Synchrotron Light Source II, a US Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Brookhaven National Laboratory under contract No. DE-SC0012704. The Center for BioMolecular Structure is primarily supported by the NIH National Institute of General Medical Sciences through a Center Core P30 Grant (P30GM133893) and by the DOE Office of Biological and Environmental Research (KP1607011). This publication resulted from data collected using the beamtime obtained through Northeastern Collaborative Access Team BAG proposal no. 311950. Lastly, we would like to thank the staff of the ESRF and European Molecular Biology Laboratory for assistance and support in using beamline BM07 under proposal number MX-2587.

Author contributions

S.A.R., D.J., V.A., F.D. and G.B. conceptualized the study. D.J. and F.D. implemented the cyclic relative positional offsets into RF2 and RFDiffusion. S.A.R., D.J., V.A. and G.B. developed the protocol for generating and filtering designs. A.L. and J.D.M. identified RbtA as a surface-exposed target in *A. baumannii*. S.A.R., V.A., M.L., P.M.L., M.S. and V.V. synthesized the designs. Y.F.B., Q.Z., M.L., S.R.G., A.L., J.A.W., A.U. and A.M. expressed and purified the target proteins. S.A.R., V.A. and A.N.L. biophysically characterized the designed macrocyclic peptides. A.K.B., J.A.W., A.U., A.K., E.B. and O.H.W., determined the X-ray crystal structures of the designed macrocycle peptides bound to their targets. S.O., O.H.W., D.W., J.A.K., J.D.M., D.B., F.D. and G.B. offered supervision throughout the project. S.A.R., D.J., V.A., D.B. and G.B. wrote the paper. All authors read and contributed to the paper. S.A.R., D.J. and V.A. agree that the order of their respective names may be changed for personal pursuits to best suit their interests.

Competing interests

D.W. is a cofounder of Priavoid and Attyloid. D.B. and G.B. are cofounders, advisors and shareholders of Vilya. The other authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41589-025-01929-w>.

Correspondence and requests for materials should be addressed to David Baker, Frank DiMaio or Gaurav Bhardwaj.

Peer review information *Nature Chemical Biology* thanks Hyun Ho Park, Francesca Peccati and the other, anonymous reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a | Confirmed

- The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided
Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
Give P values as exact values whenever suitable.
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection

Data analysis

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

The design models and sequences are available in the Supplementary Information. Crystal structures of MCB_D2 bound to MCL1, RBB_D10 bound to RbtA and apo RbtA have been deposited in the PDB under accession codes 9CDT, 9CDU, and 9CDV and will be released publically upon publication. GABARAP structures with bound macrocycles are available under PDB accession codes 9HGC and 9HGD.

Research involving human participants, their data, or biological material

Policy information about studies with [human participants or human data](#). See also policy information about [sex, gender \(identity/presentation\), and sexual orientation](#) and [race, ethnicity and racism](#).

Reporting on sex and gender	N/A
Reporting on race, ethnicity, or other socially relevant groupings	N/A
Population characteristics	N/A
Recruitment	N/A
Ethics oversight	N/A

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences Behavioural & social sciences Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Sample size for binder screening was determined by synthesis throughput available. No statistical method was used to pre-determine sample sizes for computational sampling for binder design, sample sizes were determined by compute availability.
Data exclusions	One independent replicate for GABARAP AlphaScreen assay was repeated as an outlier. One technical replicate during that repeat was dropped due to space on the plate.
Replication	RFpeptides was applied to 4 diverse design problems. We observed more than one peptide binder with signal at at least 100 μ M against each target during screening by SPR for each target and all replicates for hits from the screens were successful.
Randomization	N/A (all analysis was automated, data points were generated under controlled and uniform settings)
Blinding	N/A (all analysis was automated, data points were generated under controlled and uniform settings)

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> Antibodies
<input checked="" type="checkbox"/>	<input type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology and archaeology
<input checked="" type="checkbox"/>	<input type="checkbox"/> Animals and other organisms
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data
<input checked="" type="checkbox"/>	<input type="checkbox"/> Dual use research of concern
<input checked="" type="checkbox"/>	<input type="checkbox"/> Plants

Methods

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging

Plants

Seed stocks

Report on the source of all seed stocks or other plant material used. If applicable, state the seed stock centre and catalogue number. If plant specimens were collected from the field, describe the collection location, date and sampling procedures.

Novel plant genotypes

Describe the methods by which all novel plant genotypes were produced. This includes those generated by transgenic approaches, gene editing, chemical/radiation-based mutagenesis and hybridization. For transgenic lines, describe the transformation method, the number of independent lines analyzed and the generation upon which experiments were performed. For gene-edited lines, describe the editor used, the endogenous sequence targeted for editing, the targeting guide RNA sequence (if applicable) and how the editor was applied.

Authentication

Describe any authentication procedures for each seed stock used or novel genotype generated. Describe any experiments used to assess the effect of a mutation and, where applicable, how potential secondary effects (e.g. second site T-DNA insertions, mosaicism, off-target gene editing) were examined.