**Significance**

Peptide macrocycles are a promising class of drugs, but their weakness is conformational flexibility: target affinity can be limited by an unfavorable transition from a disordered unbound state to an ordered bound state. We introduce general computational methods for stabilizing peptide macrocycles in the bound state to an ordered bound state. We introduce general computational methods for stabilizing peptide macrocycles in the bound state to an ordered bound state. We introduce general computational methods for stabilizing peptide macrocycles in the bound state to an ordered bound state.

The current drug discovery process has shown exponentially decaying efficiency over the past several decades in terms of new drugs found per research dollar invested (9). Many factors contribute to this inefficiency, including the large numbers of lead compounds that show poor pharmacokinetic, pharmacodynamic, or toxicological properties in late-stage animal or clinical studies. A key early-stage bottleneck is the process of screening hundreds of thousands of candidate molecules in order to identify an initial hit. Rational structure-based drug design methods, which propose a small pool of candidate molecules for experimental screening that is likely to be enriched for hits, represent an attractive alternative to undirected screening-based approaches to address this bottleneck. Since these methods allow larger pools of initial hits to be identified at lower experimental cost, they...


Competing interest statement: Rosetta software has been licensed to numerous not-for-profit and for-profit organizations. Rosetta Licensing is managed by UW CoMotion, and royalty proceeds are managed by the RosettaCommons. Under institutional participation agreements between the University of Washington, acting on behalf of the RosettaCommons, their respective institutions may be entitled to a portion of revenue received on licensing Rosetta software including programs described here. R.B. and D.B. are unpaid board members of RosettaCommons. V.K.M. is a co-founder of Menten AI, in which he holds equity.

This article is a PNAS Direct Submission.

This open access article is distributed under Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 (CC BY-NC-ND).

To whom correspondence may be addressed. Email: vmulligan@flatironinstitute.org.

S.W. and T.S. contributed equally to this work.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1218900118/DCSupplemental.

Published March 15, 2021.
could also help to ease later-stage bottlenecks by providing more choice for lead identification and optimization, permitting candidates with higher probabilities of late-stage success to be carried forward.

High-affinity binding of a drug to its target depends on having a large free-energy gap between the bound and unbound states: the enthalpic favorability of the interactions between drug and target must outweigh the entropic cost of binding. Design methods generally focus on maximizing favorable interactions between a designed molecule and a target protein to maximize affinity and specificity. Unfortunately, as such methods append chemical groups to increase interactions with the target, the designed molecule becomes more flexible. This creates a mounting entropic cost associated with ordering the molecule so that it can bind, which reduces affinity and also introduces the possibility that the molecule may adopt alternative conformations that permit off-target interactions, which would hinder specificity (10). An ideal design method would maximize the favorability of intermolecular interactions between a drug and its target while simultaneously maximizing the rigidity of the drug in the unbound state, since both factors are critical for binding.

We previously reported computational methods, implemented within the Rosetta software suite (11), for designing and validating rigidly structured peptide macrocycles built from mixtures of natural and nonnatural amino acids (12–14). Rigidly structured peptide macrocycles should lose less conformational entropy on binding, and our working hypothesis is that this can address the problems hindering flexible meso-scale molecules and enable higher-affinity binding. Peptide macrocycles also combine many of the attractive properties of large-molecule (protein) therapeutics and of small-molecule drugs (15). Like protein therapeutics, peptide macrocycles present large surface areas for high-affinity, specific recognition of targets. This shared property of meso-scale and large-molecule therapeutics could account for their higher observed success rates when they reach clinical phases of testing (16). At the same time, macromolecule and incorporation of D-amino acids reduce recognition by the immune system and sensitivity to proteases, both of which are factors limiting the use of cellulyarily produced proteins as drugs (13, 17). Like small molecules, peptide macrocycles can be produced in large molar quantities, stored robustly, and administered relatively easily. Some natural peptide macrocycles, such as cyclosporine A, show oral bioavailability and cell permeability comparable to small-molecule drugs (18).

Starting with the X-ray crystal structure of NDM-1 bound to \( \text{L-} \)captopril, a weak small-molecule inhibitor of NDM-1 (19–21), we adapted our peptide macrocycle design methods to create inhibitors of NDM-1 that are simultaneously optimized for favorable interactions with the target and for rigidity in the binding-competent conformation. We promoted the latter by designing favorable internal interactions in this conformation and by strategic incorporation of rigidifying building blocks to render alternative conformations less favorable. Through enzyme inhibition assays and crystallographic studies, we show that our top design inhibits NDM-1 with 50-fold greater potency than the \( \text{L-} \)captopril control while binding to the active site in the designed binding mode and adopting the designed binding conformation. Unlike conventional drug screening approaches involving enormous compound libraries, our methods allowed us to shift most of the high-throughput exploration to in silico stages of the pipeline and to find hits from an initial experimental screen of only seven peptides. The computational methods developed here represent a general means of designing rigidly structured peptide macrocycles to bind to proteins of therapeutic interest, applicable to many targets beyond NDM-1.

Results and Discussion

Rationale and Approach for Structure-Guided Design. NDM-1 is competitively inhibited by both \( \text{L-} \) and \( \text{D-} \)isomers of captopril. Although the \( \text{D-} \)isomer is reported to be a 25-fold more potent inhibitor (21), only the \( \text{L-} \)isomer had an available X-ray crystal structure (Protein Data Bank [PDB] ID 4EXS) that we could use as a starting point when we began our peptide design work (Fig. 1A) (20). \( \text{L-} \)captopril occludes the NDM-1 active site cleft. Adjacent to this cleft are an ordered front loop (FL) consisting of amino acids 210 through 228 and a flexible hinge loop (HL) consisting of amino acids 64 through 73. The HL shows considerable conformational heterogeneity from one crystal structure to another or even in copies of the molecule in the asymmetric unit of a single crystal structure (Fig. 1B). The HL flexibility presents a major challenge for the design of larger inhibitors able to make more molecular contacts. For purposes of computational peptide design, we supposed that the observed HL conformations in available crystal structures represent relatively low-energy conformations of this loop. Since the conformation in chain B of PDB structure 4EXS presents Phe70 in a position likely to permit favorable hydrophobic interactions with an inhibitor, we chose this conformation for our in silico design work.

\( \text{L-} \)captopril resembles a \( \text{D-} \text{-cysteine-L-proline dipetide with a methyl group replacing the terminal amine. When it binds to the NDM-1 active site, the sulfur atom intercalates between and binds to the two catalytic zinc atoms, and the proline fills the space of the active site cleft (Fig. 1C). In silico, we converted the \( \text{L-} \)captopril methyl group in the 4EXS structure to an amine, yielding a \( \text{D-} \text{-cysteine-L-proline dipetide "stub" bound in the NDM-1 active site.} 

We then extended this stub, prepping a three-residue polyglycine chain by an amide bond to the \( \text{D-} \)cysteine, and similarly appending a three-residue polyglycine chain to the C terminus of the L-proline to yield an eight-residue peptide (Fig. 1D). Using the Rosetta generalized kinematic closure method (12, 13), we sampled conformations of this chain that were compatible with an amide bond linking the two termini and with favorable intramolecular backbone hydrogen bonding, keeping the \( \text{D-} \text{-cysteine-L-proline starting stub fixed.} 

For each conformation sampled, we designed sequences to maximize favorable interactions with the target while favoring the designed conformation (see below) using Rosetta side-chain packing methods, sampling \( \text{L-} \) and \( \text{D-} \)amino acids at positions able to accommodate each respectively (Fig. 1E). This was followed by a Monte Carlo-based refinement procedure in which we sampled small perturbations of the peptide conformation using generalized kinematic closure, reoptimizing side-chain identities and rotamers for each conformation sampled. We filtered this initial pool of several hundred designs based on the number of internal hydrogen bonds, shape complementarity to the target, and atomic clashes (Materials and Methods). To assess diversity of backbone conformations in the filtered population, we assigned each amino acid residue to one of four conformational bins, designated A, B, X, and Y, and representing left-handed \( \alpha \)-helical, left-handed \( \beta \)-strand, right-handed \( \alpha \)-helical, and right-handed \( \beta \)-strand conformations, respectively; these are described in greater detail in SI Appendix, section 2.1.6. We selected peptides with diverse backbone bin strings, and since we hypothesized that rigidity would be a key determinant of success, these were subjected to in silico conformational landscape analysis using the Rosetta simple cypep predict protocol (12, 13) to identify designs predicted to fold to the binding-competent conformation in the absence of the target. We used the \( P_{\text{Near}} \) metric described previously (12, 13, 15), which approximates the fractional occupancy (Boltzmann weight) of the designed conformation amid large sets of alternative conformations generated by extensive conformational sampling. \( P_{\text{Near}} \) values close to 0 indicate little predicted propensity to favor the
Inhibitory Activity of Designed Peptides. We chose seven designs for synthesis and experimental characterization, designated NDM1i-1A through NDM1i-1G, as shown in Fig. 2. These designs were selected for having favorable Rosetta peptide-target interaction energies, possessing diverse backbone conformations and intramolecular hydrogen bond patterns, and presenting hydrophobic side chains to interact with Leu65, Met67, Phe70, and Val73 on the inner hydrophobic face of the NDM-1 HL. The selected peptides were all optimized primarily for favorable interactions with the target during the design process, with folding propensity promoted by favoring conformationally constrained D- and l-proline residues. As such, they had $P_{near}$ values that ranged from 0.64 (NDM1i-1C) to 0.96 (NDM1i-1G).

We synthesized and purified the seven peptides and carried out NDM-1 inhibition assays using 1.5 μM nitrocefin as the substrate (Fig. 2, column v) at different designed inhibitor concentrations. IC$_{50}$ values were estimated as described in SI Appendix, section 3.3. As a positive control, we used the L-captopril isoform (a more potent inhibitor than the D-captopril starting point for design), which had an IC$_{50}$ value of 59.7 ± 6.3 μM (SI Appendix, Fig. S1). High-quality fits to the data for all peptides and controls were consistent with the expected 1:1 stoichiometry of binding (see SI Appendix, section 3.3 for details). All of the peptides but NDM1i-1C had IC$_{50}$ values lower than D-captopril, with the top peptide, NDM1i-1G, having an IC$_{50}$ value of 1.2 ± 0.1 μM, more than 50 times more potent than D-captopril.

Since a robust peptide therapeutic design pipeline would benefit considerably from computational metrics that could reliably rank designs to prioritize syntheses and experiments, we next examined which metrics best correlated with experimental success across our initial batch of designs. As noted above, the free energy of binding of a flexible molecule to a fixed target can be decomposed as the sum of two terms: \( \Delta G_{\text{binding}} \) the interaction free energy between the molecule and the target in the bound complex, and \( \Delta G_{\text{folding}} \), the free energy of ordering the flexible molecule into the conformation adopted in the complex. Rosetta estimates of \( \Delta G_{\text{binding}} \) using the difference in energy between the bound and unbound conformations (with limited conformational sampling of side chains across replicates) had little correlation with observed IC$_{50}$ values (Fig. 3A). Since \( \Delta G_{\text{folding}} = -RT \ln (\text{eq}) = -RT \ln \left( \frac{\text{bound}}{\text{unbound}} \right) \), where \( f \) is the fractional occupancy of the folded state at equilibrium, we can estimate folding free-energy changes using the $P_{near}$ metric described above as the approximate value of $f$ (SI Appendix, section 1.5.4). Such estimates of \( \Delta G_{\text{folding}} \) which are based on near-exhaustive sampling of the conformations of the peptide macrocycle in isolation, converge robustly and correlate well with the logarithm of the IC$_{50}$ value—so well that the rank order of computed \( \Delta G_{\text{folding}} \) values matches the rank order of experimental IC$_{50}$ values (Fig. 3B). Comparisons to conformational sampling simulations using earlier versions of the Rosetta energy function reveal that improvements to the energy function accuracy using small-molecule fluid simulations for parameter tuning (22, 23) have improved the correlation between estimated \( \Delta G_{\text{folding}} \) and observed IC$_{50}$ (SI Appendix, Fig. S2). There are several possible explanations for the lack of correlation between our $\Delta G_{\text{binding}}$ estimates and the observed IC$_{50}$ values. First, the differences in the interaction energies across these seven designs are likely to

---

**Fig. 1.** Computational design approach for generating peptide macrocycle inhibitors of NDM-1. (A) Structure of NDM-1 (PDB ID 4EXS), chain B. The active site binds catalytic zinc atoms and is flanked by an ordered FL and a flexible HL. Hydrophobic amino acid residues on the inner face of the HL, and metal-coordinating residues, are labeled. (B) Comparison of a subset of NDM-1 crystal structures. PDB IDs 3RKJ, 3S0Z, 3ZR9, and 4HL1 are shown in gray. In lavender and green are PDB ID 4EXS, chains A and B, respectively. Where most of the structure, including the FL, is rigid, the HL shows extensive conformational flexibility, putting inner-face hydrophobic side chains (labeled) in diverse positions. (C) Crystal structure of NDM-1 active site (green) with L-captopril (purple) bound (PDB ID 4EXS, chain B). Active-site zinc atoms are shown beneath the surface as spheres. (D) In silico conversion of L-captopril to a D-proline, L-cysteine dipeptide (purple) flanked by polyglycine sequences (pink). (E) Rapid in silico sampling of closed conformations of a peptide macrocycle containing the o-cysteine, L-proline stub (purple), and flanking sequences (pink) in the context of the NDM-1 active site, using the generalized kinematic closure approach. For each closed conformation, Rosetta design heuristics were used to find side-chain identities and conformations (represented here by spheres).
Fig. 2. Designed eight-residue peptide macrocycle inhibitors of NDM-1, designated NDM1i-1A (A) through NDM1i-1G (G). (i) Amino acid sequences (AA) and backbone conformational bins (Bin) of designed peptides. In this and the following two columns, L-amino acids are shown in cyan and D-amino acids in orange. Backbone conformational bins are described in SI Appendix, section 2.1.6. (ii) Peptide design computer models shown as stick representations. Intramolecular backbone hydrogen bonds are shown as green lines. Sequence numbering is as shown in i. (iii) Space-filling computer models of designed peptides in the NDM-1 active site, with NDM-1 shown in gray. The HL, FL, and interacting residues Phe70 and Val73 are indicated. (iv) Conformational landscape analysis performed with the Rosetta simple_cycpep_predict application, showing computed energy of the peptide modeled in isolation plotted against rmsd to its designed binding conformation. Each point represents a separate conformational sampling attempt. Colors indicate the number of intramolecular backbone hydrogen bonds observed in the sampled conformation. \( P_{\text{norm}} \) values are indicated, with the mean and SE of three independent large-scale conformational sampling simulation replicates reported. (v) Experimentally measured activity of NDM-1 (vertical axis) in the presence of varying concentrations of peptide (horizontal axis). Points are mean of three independent replicates, and error bars represent the SEM. Red curves show fits to the Hill equation, with \( IC_{50} \) values and fit confidence indicated on each plot. (Insets) Fit residuals.
the fold propensity. We synthesized four point mutants that were predicted to preserve the fold and to interact favorably with the target: D-Arg1→D-Thr (r11), L-Leu3→L-Tyr (L3Y), L-Ile6→L-Leu (I6L), and L-Glu8→2-aminoisobutyric acid (E8AIB), along with seven combinations of these mutations (SI Appendix, Figs. S4 and S5). These peptides are designated NDM1i-2A through 2K. Several of these mutations increased IC50 values without reducing computed ΔGfolding values (SI Appendix, Fig. S6), suggesting that the manual introduction of these mutations to an optimized design weakened interactions with the target. A triple mutation with an IC50 value of 1.8 ± 0.1 μM (NDM1i-2I, bearing mutations L3Y/I6L/E8AIB) showed greater inhibition than any of the individual mutations or the L3Y/I6L double mutation (NDM1i-2H). The IC50 value was close to that of the NDM1i-1G (1.2 ± 0.1 μM) starting point, suggesting that there are multiple opportunities for finding variant inhibitors in the local sequence space of these peptides. These experiments are described in greater detail in SI Appendix, section 4.2.

Crystal Structures of Inhibitory Peptides Bound to NDM-1. To gain greater insight into the inhibition of NDM-1 by some of the top inhibitors, we crystallized the enzyme and solved structures by X-ray crystallography in complex with peptides NDM1i-1F and NDM1i-1G. Fig. 4 shows a comparison of the design and crystal structure of NDM1i-1G bound to NDM1. The binding mode observed in the crystal structure closely resembles that in the design, with the D-Cys-1-Pro stub coordinating active-site zinc residues as the D-captopeptir starting compound does. Peptide residues L-Leu3 and L-Ile6 pack against NDM-1 HL residues Met67 and Phe70, albeit with slightly different packing interactions than designed. This is due to flexibility of the HL, which moves in the crystal structure relative to the design structure (HL backbone heavyatom rmsd 3.1 Å), causing the peptide to rotate slightly about the stub residues in the opposite direction (peptide backbone heavyatom rmsd 1.8 Å) (Fig. 4C). Despite this, the internal conformation of the peptide remains rigid: when the peptide portion of the design model is aligned with the peptide portion of the crystal structure, the rmsd is 0.3 Å (Fig. 4D). Designed ionic interactions between NDM1i-1G residue d-Arg2 and NDM-1 residues Glu152 and Asp223 were blocked by the binding of a zinc ion to the anionic NDM-1 residues (Upper Insets in Fig. 4A and B). Despite these differences, the binding site and conformation are very close to the design model, demonstrating the power of the computational design methods used.

Peptide NDM1i-1F differs from NDM1i-1G by an 16V mutation, effectively replacing one methyl group by a hydrogen atom. This small change results in an approximately two-fold reduction in binding affinity. Differences in the crystal structures of NDM1i-1F and NDM1i-1G help to explain this. The C6 atom in NDM1i-1G L-Ile6 is buried between Met67 and Phe70 on the HL hydrophobic face (Fig. 4B). When this atom is removed, Met67 adopts an alternative conformation allowing a small (0.7 Å) shift of the HL to fill the void (SI Appendix, Fig. S7). This rearrangement may account for the change in binding affinity. Like NDM1i-1G, peptide NDM1i-1F binds in a binding mode that resembles the design, with the HL shifting by 3.7 Å, and the peptide rotating in the opposite direction by 1.3 Å (backbone heavyatom rmsds). The backbone heavyatom rmsds between the superimposed peptide portion of the design model and the crystal structure is 0.4 Å, again indicating atomic-resolution accuracy in computational design of the peptide conformation itself.

Incorporation of Noncanonical Side Chains. Our attempts to design NDM-1 inhibitors were carried out concurrently with development work to enhance the computational methods to expand the set of noncanonical amino acid building blocks available for computational design (SI Appendix, section 1). Past design

**Inhibitory Activity of Variants of NDM1i-1G.** We next carried out in silico mutagenesis of the top inhibitor NDM1i-1G, examining the effect on P_{Near} of mutations at every position to each of 46 possible amino acid types. As shown in SI Appendix, Fig. S3, the peptide is highly mutable, with many chirality-preserving mutations, as well as some chirality-inverting mutations, preserving
efforts involving exotic noncanonical building blocks used an energy function inspired by molecular dynamics force fields (25), but in the context of a target protein, this would lose the advantages of the Rosetta energy function, which has been highly optimized to reproduce conformational preferences of proteinogetic amino acids. We therefore opted to use the Rosetta ref2015 energy function with computed side-chain potentials produced by the MakeRotLib application, as described in SI Appendix, sections 1.1 and 1.2. We explored whether an expanded palette of amino acid building blocks could unlock new inhibition mechanisms.

Using the crystal structure of NDM1i-1G as our starting point, we sampled perturbations of the bound conformation of this peptide and designed sequences incorporating several noncanonical side chains (SI Appendix, Table S1). We found that many design trajectories converged to include L-norleucine (L-Nlu) at position 3, 2-aminomethyl-L-phenylalanine (L-A34) at position 6, and (4R)-4-hydroxy-L-proline (L-Hyp) at position 7. We synthesized and tested four designs with predicted fold propensities above 0.9 that incorporated these noncanonical amino acids. We therefore opted to use the Rosetta energy function inspired by molecular dynamics force fields (25), efforts involving exotic noncanonical building blocks used an energy function inspired by molecular dynamics force fields (25), but in the context of a target protein, this would lose the advantages of the Rosetta energy function, which has been highly optimized to reproduce conformational preferences of proteinogetic amino acids. We therefore opted to use the Rosetta ref2015 energy function with computed side-chain potentials produced by the MakeRotLib application, as described in SI Appendix, sections 1.1 and 1.2. We explored whether an expanded palette of amino acid building blocks could unlock new inhibition mechanisms.

Using the crystal structure of NDM1i-1G as our starting point, we sampled perturbations of the bound conformation of this peptide and designed sequences incorporating several noncanonical side chains (SI Appendix, Table S1). We found that many design trajectories converged to include L-norleucine (L-Nlu) at position 3, 2-aminomethyl-L-phenylalanine (L-A34) at position 6, and (4R)-4-hydroxy-L-proline (L-Hyp) at position 7. We synthesized and tested four designs with predicted fold propensities above 0.9 that incorporated these noncanonical amino acids, shown in SI Appendix, Fig. S8. One of these, NDM1i-3D, had an IC₅₀ value of 3.1 ± 0.3 μM, close to that of NDM1i-1G.

We solved the X-ray crystal structure of peptide NDM1i-3D bound to NDM-1 and found that, although this peptide did occlude the NDM-1 active site, its binding mode was inverted relative to the design, with residue L-Glu8, rather than D-Cys4, coordinating the active-site zinc atoms (Fig. 5). When the peptide structure and design model of the complex were aligned, the peptide backbone heavy atom rmsd was 9.4 Å due to this inversion. The HL conformation was closer to the design than in the cases of NDM1i-1F and -1G, deviating by a backbone heavy atom rmsd of 1.2 Å. Despite this, the peptide was rigidly structured in the designed conformation: superposition of the peptide portion of the structure yielded a backbone heavy atom rmsd of 0.4 Å from design to crystal structure. A similar rotation by ~180° was observed previously in the X-ray crystal structure of a two-sided de novo-designed heterodimer interface between two protein scaffolds (26). Although the peptide macrocycle designed here is very different from the protein scaffold in this previous study, both have in common a certain rough symmetry: both present side chains in a manner that is roughly preserved on 180° rotation. In the case of the designed protein, 180° rotation roughly preserves the location of repeated helices. In the case of the designed macrocycle, the peptide superimposes on its own four-residue cyclic permutation, corresponding to a 180° rotation, with a backbone heavy atom rmsd of 2.0 Å. This cyclic permutation places each of the hydrophobic L-Nlu and L-A34 side chains in the space that the other formerly occupied, while permitting L-Glu8 to replace D-Cys4 at the metal-binding position (Fig. 5E). Future design...
Fig. 5. Comparison of design model and crystal structure (PDB ID 6XCI) of peptide NDM1i-3D bound to NDM-1. (A) Design model of NDM1i-3D (pink surface, with L- and D-amino acid residues shown in cyan and orange, respectively) in the NDM-1 active site (green). L-2-aminomethyl phenylalanine (L-A34) and L-norleucine (L-Nlu) residues make hydrophobic contacts with the inner face of the HL. (Inset) D-Cys at position 5 coordinates active-site zinc atoms. (B) X-ray crystal structure of NDM1i-3D bound to NDM-1. The peptide is rotated nearly 180° relative to the design model with L-Nlu and L-A34 residues in opposite positions. Water molecules are shown as sticks with blue surfaces. (Inset) L-Glu at position 8 coordinates the active-site zinc. Cadmium is observed in place of zinc at the adjacent site. (C) Overlay of X-ray crystal structure (darker colors) and design model (lighter colors). NDM-1 is shown in green; L- and D-amino acids in NDM1i-3D are shown in cyan and orange, respectively, and stub residues are shown in purple. The crystal structure’s positions are labeled in black, and the design model’s positions in white. As shown, the rotation of the design model puts l-Glu-8 (red arrows) where d-Cys-4 (orange arrows) would be. The motion of the peptide displaces it by an rmsd of 9.4 Å, while the HL moves by 1.2 Å. (D) Overlay of aligned peptide portions of the crystal structure (darker colors) and design model (lighter colors). Cyan and orange represent L- and D-amino acids, as before. Despite the change in binding orientation, the crystal structure peptide conformation matches the design to a backbone heavy-atom rmsd of 0.4 Å. (E) Overlay of crystal structure with peptide design circularly permuted by four residues. The rough symmetry of the backbone conformation allows d-Pro 1 to occupy the space that would be occupied by L-Pro 5 (green arrows), d-Cys 4 to occupy the space that would be occupied by L-Glu 8 blue arrows), and L-Nlu 3 to occupy the space that would be occupied by L-A34 6 (red arrows), possibly explaining why the peptide was able to bind to the same site in a very different binding mode.

Conclusions

We have introduced general computational methods for designing peptide macrocycles to bind to targets of therapeutic interest. Unlike screening-based approaches, computational design allows the creation of molecules able to bind to a desired site and in a desired binding mode. Of our seven NDM1i-1 designs, six were stronger inhibitors than the D-captopril control (itself a stronger inhibitor than the L-captopril starting point for design). By explicitly considering the propensity of a peptide macrocycle to favor a binding-competent conformation, we were able to predict the rank order of IC50 values, supporting our working hypothesis that the entropic cost of ordering larger molecules on binding must be minimized, while also providing a useful tool for in silico screening of future designs. X-ray crystallography confirmed that our top binders, NDM1i-1F and NDM1i-1G, bound to the active site in a binding mode very close to that designed, with flexibility of a flanking loop accounting for deviation from the design.

Accurately predicting the effect of point mutations on binding in silico can be difficult, as observed in our characterization of peptides NDM1i-2A through 2K, making experimental screens of variants of computationally designed starting points a useful complement. The fact that peptide NDM1i-2J, found in a very small experimental screen of variants, has inhibitory activity comparable to the NDM1i-1G starting point despite differing in...
sequence at three of eight positions suggests that scaffolds with high propensity to favor the binding-competent configuration provide good starting points for more extensive optimization of binding affinity through mutagenesis experiments.

Challenges for computational macrocycle design include the difficulty of considering the possible conformations of the peptide macrocycle backbone, the possible conformations of the loops flanking the target site, and the possible orientations of the peptide relative to the target. It can also be difficult to correctly model the conformational energetics of exotic chemical building blocks. Both of these may have contributed to the serendipitous discovery of an alternative binding mode of peptide NDM1i-3D, although the finding that this peptide adopted the designed backbone conformation raises the possibility of designing peptides with internal quasi-symmetry that have multiple possible binding modes for a target (27).

With the 29 peptides described here, including 6 with IC50 and KI values under 5 μM (SI Appendix, Table S3), we demonstrate an approach for the rapid identification of hits to inhibit antibiotic resistance factors. That these molecules are highly mutable provides a means of producing variants to continue the “arms race” as resistance mechanisms evolve. More broadly, these techniques could offer an alternative to costly high-throughput compound screens for a broad range of targets of therapeutic interest.

Materials and Methods

Enhancements of the Rosetta Software Suite. The computational work described here was carried out with the Rosetta software suite, a set of C++ libraries and applications for heteropolymer design, structure prediction, and modeling (11). Software enhancements needed to enable this work included improved support for energetic calculations involving noncanonical amino acids with the protein-centric ref2015 energy function (22, 23), the implementation of four new design-centric guidance scoring terms to control the construction of the creation of design-centric guidance terms (described in SI Appendix, section 2.1.3) to produce the NDM1i-4 designs. We also altered the protocol used for NDM1i-3 designs by adding sampling of small perturbations of the backbone and relaxing the restrictions on the number of exotic side chains that could be incorporated and including crystallographic water molecules during design as hydrogen bond donors and acceptors. See SI Appendix, section 2.1.3, for details on the design protocol, and SI Appendix, Table S1, for the energetic calculations, and SI Appendix, section 1.5.4, for details on both computational and experimental efforts.

Computational Protocols. NDM1-1 peptides were designed by starting with the structure of NDM-1 bound to l-captopril (PDB ID 4EXS). l-Captopril resembles a γ-cysteine-γ-proline dipeptide but for a methyl group that replaces the terminal amine. In silicon, we converted l-captoptil in the active-site pocket to a dipeptide and extended it with a polyglycine sequence to make an octapeptide, which we cycled using Rosetta’s generalized kine-
matic closure protocol. This approach is general and can be applied to starting stubs from experimentally solved complexes or from in silico docking. We discarded conformations with fewer than three internal backbone hydrogen bonds, those with oversaturated hydrogen bond acceptors, or those with egregious clashes between the macrocycle backbone and the target. We then used Rosetta’s rotamer optimization module (the Rosetta packer) to design the peptide sequence while simultaneously sampling alternative packings of nearby side chains on the NDM-1 target. We refined the initial design through a Monte Carlo search in which moves consisted of small perturbations of the macrocycle backbone (maintaining closure using generalized kine-
matic guidance) and side-chain reoptimization. Top confor-
mations encountered during the Monte Carlo trajectory were more rigorously redesigned using the Rosetta FastDesign protocol (12). To select designs for synthesis, metrics such as shape complementarity and overall Rosetta energy were considered. We also sought diversity in the design pool using backbone bin strings to classify conformations, as described in SI Appendix, section 2.1.6. In addition, top peptides were subjected to confor-
national landscape analysis using the Rosetta simple_cycep_predict approach, which computed the Pscore, metric and produced an estimate of the peptide ΔGguiding. See SI Appendix, section 1.5.4, for details on both calculations, and SI Appendix, sections 2.1.1 and 2.2.2, for details on the design protocol.

NDM1i-2 designs were variants on NDM1i-1G. Four point mutants were selected using in silico mutational scanning and Pscore analysis (SI Appendix, Fig. S3). These, and seven combinations of these mutations, were synthesized and tested.

NDM1i-3 designs were designed using a variant of the protocol used to produce NDM1i-1 designs. Starting with the X-ray crystal structure of NDM1i-1G bound to NDM-1 (Fig. 4), the macrocycle was subjected to small perturbations and redesigned using a much-expanded palette of amino acid building blocks. The step of performing extensive macrocycle backbone conformational sampling via a Monte Carlo search was omitted. To be as conservative as possible, we constrained the number of exotic noncanonical amino acids to one to two per design using the aa_composition design-centric guidance scoring term (13). We also made use of newly designed conformational libraries and applications for heteropolymer design, structure prediction, and modeling (11). Software enhancements needed to enable this work are also available from GitHub (https://github.com/vmullig/ndm1_design_scripts) (29), and from the RosettaCommons RossettaScripts scripts repository.

Enzymatic Assays and Data Analysis. NDM-1 was expressed in and purified from Escherichia coli BL21(DE3) cells. Inhibition of the hydrolysis of 1.5 μM nitrocefin was assayed as described in SI Appendix, section 3.3. C-Captopril was used as a positive control (SI Appendix, section 2.1.3). Activity was plotted as a function of inhibitor concentration, and data were fitted with Scipy using a modified Hill equation to extract IC50 values, as described in SI Appendix, section 3.3. Since all inhibition assays were performed with a constant initial concentration of substrate, IC50 values were proportional to KI values, allowing direct comparison across inhibitors; however, KI values for all peptides are also reported in SI Appendix, section 4.1.

X-Ray Crystallography. For crystallization, NDM-1 was expressed in and purified from E. coli BL21(DE3) cells as described in SI Appendix, section 3.4.1. NDM1i-1F, NDM1i-1G, or NDM1i-3D peptide was added to the protein, and crystals were grown by the hanging-drop method; full details are provided in SI Appendix, section 3.4.2. Following cryoprotection with 25% glycerol and immersion in liquid nitrogen, diffraction data were collected on beamline 5.0.1 at the Advanced Light Source, at 100 K using a wavelength of 0.979 Å. Data for NDM1i-3D were collected on beamline 5.0.1 at the Advanced Light Source, at 100 K using a wavelength of 0.977 Å. The structure of NDM-1 with no peptide bound (PDB ID 35PU) was used for molecular replacement phasing. Model validation was carried out with MolProbity (30) with the NDM1i-1F complex having 99.02 and 0.22%, the NDM1i-1G complex having 98.69 and 0%, and NDM1i-3D having 98.88 and 0% Ramachandran-favored and outliers, respectively. Full details of analysis and refinement are provided in SI Appendix, section 3.4.2, and data processing, refinement, and model statistics are shown in SI Appendix, Table S3.

Data Availability. RosettaScripts XML scripts for peptide macrocycle inhibitor design data have been deposited in Github (https://github.com/vmullig/ndm1_design_scripts), and are also available in the SI Appendix. Structure factors and coordinates for the NDM1i-1F, NDM1i-1G, and NDM1i-3D complexes have been deposited in the Protein Data Bank (PDB IDs 6XBE, 6XBF, and 6XCI, respectively).

Acknowledgments. V.K.M., P.D.R., and R.B. were supported by the Simulation Initiative. S.W. was supported by an Alexander Graham Bell Canada Graduate Scholarship from the Natural Sciences and Engineering Research Council. T.S. was supported by a Michael Smith Foundation for
Health Research Postdoctoral Fellowship. D.T.K. was supported by a doctoral award from the Canadian Institutes of Health Research (CIHR). P.H. and T.W.C. were supported by Washington Research Foundation Innovation postdoctoral fellowships. P.H. was also supported by NIH Ruth Kirschstein F32 award no. F32GM120791-02. A.M.W. was supported by NIH Grant R21 CA219847. J.W.L. was supported by the NIH Grants 1F32-CA189246 and R01- GM127578 and by the RosettaCommons. R.M. was supported by the RosettaCommons. N.C.J.S. was supported by operating funds awarded from the CIHR and Tier 1 Canada Research Chair program. D.B. was supported by the Howard Hughes Medical Institute. Crystallographic data were collected using beamline 08ID-1 at the Canadian Light Source, which is supported by the Howard Hughes Medical Institute. Crystallographic data were collected using beamline 5.0.2 at the Advanced Light Source, which is supported by National Laboratory, the University of Washington Hyak cluster, and the supported by the Canada Foundation for Innovation, Natural Sciences and Engineering Research Council of Canada, the University of Saskatchewan, the Government of Saskatchewan, Western Economic Diversification Canada, the National Research Council Canada, and the CIHR. Crystallographic data were also accrued at the Advanced Light Source, a Department of Energy (DOE) Office of Science User Facility under contract no. DE-AC02-05CH11231. The authors thank the staff of beamline 5.0.2 for their assistance and Fred Rosell for advice on the nitrocefin assay. An award of computer time was provided to D.B. and V.K.M. by the Innovative and Novel Computational Impact on Theory and Experiment program. This research used resources of the Argonne Leadership Computing Facility, which is a DOE Office of Science User Facility supported under Contract DE-AC02-06CH11357. Computations were carried out on the Mira and Theta supercomputers at Argonne National Laboratory, the University of Washington Hyak cluster, and is supported by the Simons Foundation Iron, Gordon, and Popeye clusters. We thank Andrew Leaver-Fay and Sergey Lyskov for Rosetta support and Yuri Alexeev for support on Mira and Theta.

9. J. W. Scannell, A. Blanckley, H. Boldon, B. Warrington, Diagnosing the decline in β-lactamase 1
27. V. K. Mulligan et al., Computational design of mixed chirality peptide macrocycles with internal symmetry. Protein Sci. 29, 2433–2445 (2020).