Mixed stereochemistry macrocycle acts as a helix-stabilizing peptide N-cap

Fabian Hink[1], Julen Aduriz-Arrizabalaga[2], Xabier Lopez[2], Hiroaki Suga[3], David De Sancho[2], Joseph M. Rogers[1]*

[1]Department of Drug Design and Pharmacology, University of Copenhagen, 2100, Copenhagen, Denmark.[2]Polimero eta Material Aurreratuak: Fisika, Kimika eta Teknologia, Kimika Fakultatea, UPV/EHU & Donostia International Physics Center (DIPC), PK 1072, 20018 Donostia-San Sebastian, Euskadi, Spain [3]Department of Chemistry, Graduate School of Science, The University of Tokyo, Bunkyo-ku, Tokyo 113-0033, Japan
*e-mail: joseph.rogers@sund.ku.dk

Keywords

Abstract

Interactions between proteins and α-helical peptides are abundant in the human cell. Many of these interactions are linked to disease and have been the focus of drug discovery campaigns. However, the large interfaces formed between multiple turns of α-helix and a binding protein represent a significant challenge to inhibitor discovery. Modified peptides featuring helix-stabilizing macrocycles have shown promise as inhibitors of these interactions. Here, we tested the ability of N-terminal to side-chain thioether-cyclized peptides to inhibit the α-helix binding protein Mcl-1, by screening a trillion-scale peptide library using RaPID technology. The top enriched peptides were lariats, featuring a small, four amino acid N-terminal macrocycle followed by a short linear sequence that resembled the natural α-helical Mcl-1 ligands. These 'Heliats' (*Hel*ical la*riats*) bound Mcl-1 with mid-nM affinity, and inhibited the interaction between Mcl-1 and a natural peptide ligand. Macrocyclization was found to stabilize α-helical structures, and significantly contribute to affinity and potency. Yet, the 2nd and 3rd positions within the macrocycle were permissible to sequence variation, so that a minimal macrocyclic motif, of an N-acetylated D-phenylalanine at the 1st position thioether connected a cysteine at the 4th, could be grafted into a natural peptide and stabilize helical conformations. We found that D-stereochemistry is more helix-stabilizing than L- at the 1st position in the motif, as the D-amino acid can utilize polyproline II torsional angles that allow for more optimal intrachain hydrogen bonding. This mixed stereochemistry macrocyclic N-cap is synthetically accessible, requiring only minor modifications to standard solid-phase peptide synthesis, and its compatibility with RaPID peptide screening can provide ready access to helix-focused peptide libraries for *de novo* inhibitor discovery.

Introduction

Protein-protein interactions are essential for cellular function[1]. A common type of interaction found in human cells is the binding of an intrinsically disordered region or protein (henceforth

referred to as IDPs) to a folded domain[2]. Upon binding, the IDP can fold into a distinct secondary structure, often an α-helix[2]. A number of these interactions are involved in disease processes, making these α-helix binding proteins attractive drug targets[3]. However, the large binding interfaces formed by multiple turns of α-helix and a binding protein makes the discovery of inhibitors a significant drug discovery challenge[4].

One modality showing promise as inhibitors of such interactions are peptides[5]. Peptides can mimic natural IDPs and inhibit otherwise intractable interactions[6]. However, short, linear peptides, much like natural IDPs, tend to be disordered in the absence of a binding partner and often only fold into a distinct secondary structure upon binding[7]. Their folding then comes with an entropic penalty, which negatively impacts binding strength[8]. To remedy this, macrocyclization, adding a ring of twelve or more atoms, has proved a successful strategy to 'pre-organize' a peptide for binding, by stabilizing the structures it adopts upon binding. This has included stabilization of specific secondary structures, and side-chain to side-chain 'staples' have been particularly successful in stabilizing α-helices[9].

A powerful alternative to modifying an existing peptide is to screen libraries for *de novo* macrocyclic peptides. One technology is the RaPID system (*Ra*ndom non-standard *P*eptides *I*ntegrated *D*iscovery)[10]. RaPID combines mRNA display, which uses ribosomal translation to generate trillions of DNA- or mRNA-tagged peptides[11], with flexizyme-based genetic code reprogramming, where a catalytically active RNA molecule loads non-canonical amino acids onto tRNA for use by ribosomal translation[12]. Typically this non-canonical amino acid is an N-chloroacetylated aromatic amino acid, which can spontaneously react with a C-terminal cysteine to form a thioether macrocycle[13]. mRNA coding for random amino acids between the N-terminal non-canonical and cysteine allows for the generation of macrocyclic peptide libraries. Affinity selection against an immobilized target protein and subsequent DNA sequencing allows for the identification of *de novo* cyclic peptide binders[14]. Hit cyclic peptides from RaPID selections form mainly sheet or coil structures, whereas only a few contain α-helix, typically one or two turns (Figure 1A)[15-22]. It is not clear how such short N-terminal to side-chain macrocyclic peptides can compete with IDPs that bind using long helices with several turns.

The Bcl-2 (B cell lymphoma 2) protein family contains prominent examples of helical peptides binding folded partners[23]. For instance, the folded domains of Mcl-1 (myeloid cell leukemia-1) can bind to proteins containing a BH3 (Bcl-2 homolog 3) and cause it to fold into a long α-helix (Figure 1B). Mcl-1 is an important regulator of apoptosis[24]. By binding to pro-apoptotic proteins containing a BH3 motif such as Bid (BH3-interacting domain death agonist), Mcl-1 prevents programmed cell death[25-27]. In cancer, apoptotic mechanisms are often abrogated by an overexpression of Mcl-1[24]. Higher Mcl-1 expression levels also result in decreased sensitivity to commonly applied anticancer therapies[28], making Mcl-1 an attractive cancer drug target. Several molecules targeting Mcl-1, including macrocycles, have entered clinical trials, highlighting the importance of Mcl-1 as a drug target[29].

Here, we sought to test the ability of RaPID to inhibit the protein-protein interaction between a long helical peptide and a folded protein target. We conducted RaPID against Mcl-1 and discovered an enrichment for peptides with a lariat structure, containing a small N-terminal macrocycle. We found that this 15-atom macrocycle, containing a D-stereocenter, could

facilitate binding to Mcl-1, improve the inhibitory capacity, and stabilize α-helical conformations in the discovered peptides.

Results

RaPID selection targeting Mcl-1 enriched for lariat peptides

We screened a library of cyclic peptides against Mcl-1, using RaPID. Through genetic code reprogramming, a cysteine-reactive ClAc-D-Phe was used as the initiator amino acid in each peptide (instead of methionine), this was followed by 6-15 random amino acids, then a cysteine for cyclization. Degenerate codons were used to encode the random amino acids, so that canonical amino acids (apart from methionine) would be sampled. mRNA display was used to link each cyclic peptide to its encoding mRNA-DNA duplex, and pull-down using immobilized Mcl-1 was used to select for binding peptides (Figure 1C). Eluted DNA was amplified by PCR and used to synthesize a new cyclic peptide library for additional rounds of selection. 7 rounds of increasingly stringent selection were performed, using higher temperatures and lower concentrations of Mcl-1 (Table S1). The fraction of mRNA displayed peptides recovered from the pull-down increased over the rounds (Figure S1), indicating enrichment of Mcl-1 binding peptides.

Next-generation sequencing revealed a progressive enrichment of particular cyclic peptide sequences (Figure 1D). Two features are notable: 1) the library became dominated by peptides with cysteine at the 4th position (Figure 1D, Figure S2). 2) the peptides had significant homology to natural Mcl-1 binding peptides i.e. they contained BH3 motifs (Figure 1E). The former is possible as cysteine is sampled by the degenerate codon library. Indeed, we calculate that 2.9% of the peptides in the initial library would feature a cysteine at the 4th position (Figure S2B). As the nearest cysteine in the sequence is most likely to undergo the irreversible reaction with the ClAc initiator[30], our peptides are likely to form 'lariats': small macrocycles connected to a linear chain. We note that this linear region contains significant homology to the BH3 motifs, having at least two of the four hydrophobic hotspot residues (Figure 1E). This linear region is likely mimicking the natural linear BH3 motifs and forming an α-helix upon binding Mcl-1.
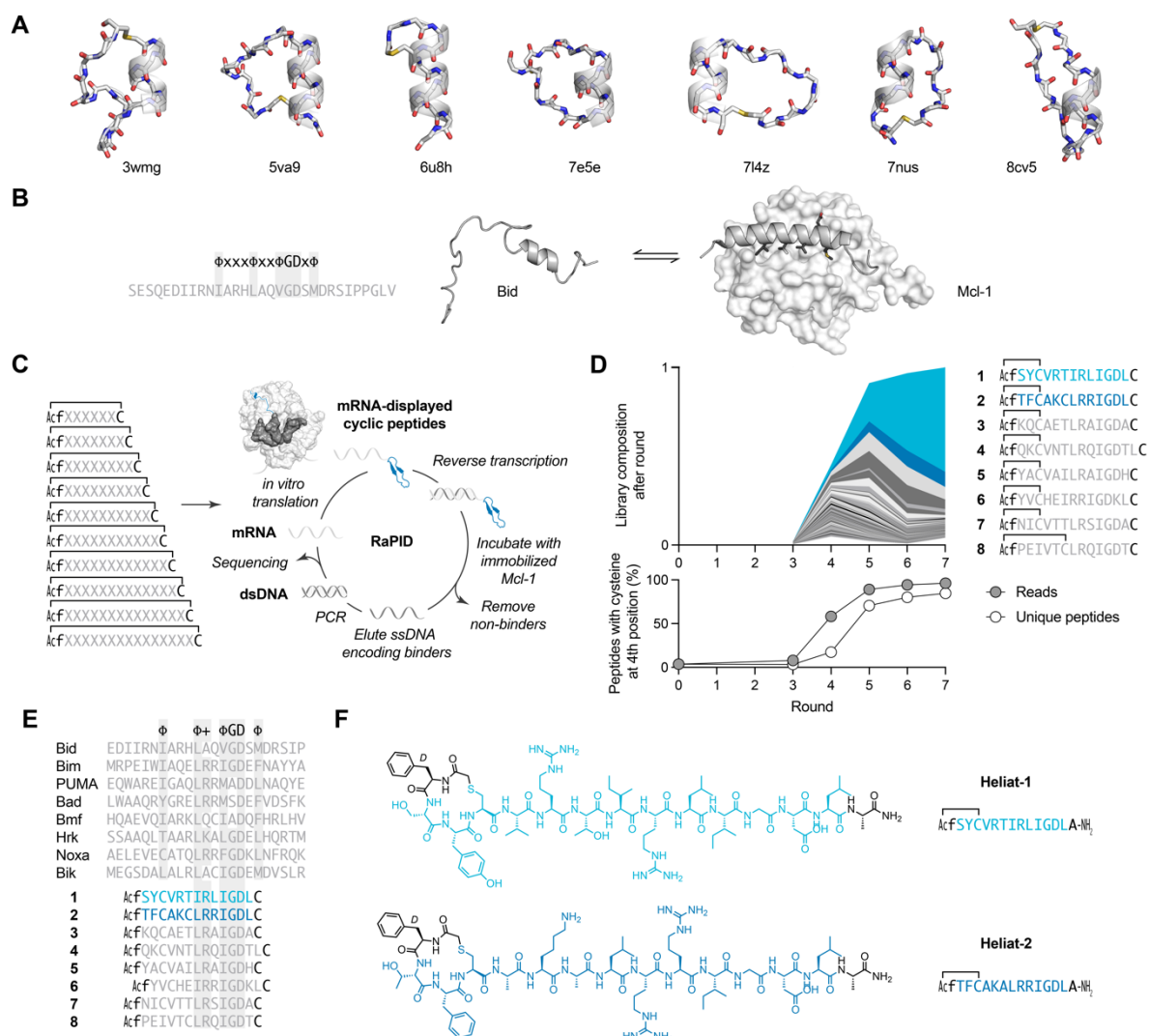
Figure 1. RaPID selection against Mcl-1 enriched for lariat peptides with a small 4-amino acid macrocycle.
A) Representative N-terminal to side-chain thioether-cyclized peptides that form an α-helix upon binding their target protein, PDB codes shown. **B**) The BH3 motif of Bid is a disordered peptide in isolation, but forms an α-helix upon biding Mcl-1. BH3 motifs have four hydrophobic hotspot residues (φ) with the third φ followed by a small amino acid (typically G) and an aspartate (D). **C**) RaPID was used to screen a large library (trillions) of peptides for binding to human Mcl-1, each thioether-cyclized and containing 6-15 randomized amino acids (X). **D**) Sequencing after 3-7 RaPID rounds reveals progressive enrichment of particular sequences. The majority of the sequencing reads, and unique peptides, contain a cysteine at the 4th position. **E**) Top peptides show high similarity to natural BH3 motifs, including two of the hydrophobic hotspots (φ). **F**) Cyclization through the 4th position cysteine will produce lariats: a small macrocycle connected to a short linear peptide. Derivates of RaPID hits **1** and **2**, **Heliat-1** and **Heliat-2**, were synthesized by SPPS for further characterization.

N-terminal macrocycle aids binding

We sought to understand the function, if any, of the small, 4 amino acid, N-terminal macrocycle. We synthesized the top two RaPID hits **1** and **2** using solid phase peptide synthesis (SPPS), which we named **Heliat-1** (*Heli*cal lar*iat*) and **Heliat-2**, respectively (Figure 1F). To guarantee the lariat structure, all non-cyclizing cysteines were mutated to alanine. We also synthesized linear variants lacking the thioether bond, **Linear-1** and **Linear-2** respectively (Figure 2A), and an aligned fragment (14 aa) of the BH3 motif from Bid, **Bid14**.

To assess binding to Mcl-1, affinity was measured by Surface Plasmon Resonance (SPR). **Heliat-1** and **Heliat-2** bound with mid-nM affinity (K_D = 70 ± 30 nM and 180 ± 110 nM, respectively, Figure S3, Table S2 and S3), showing significantly tighter binding than their linear variants (K_D = 4,600 ± 1,600 nM and 23,000 ± 8,000 nM), resulting in a $\Delta\Delta G$ of 2.6 ± 0.2 kcal mol^{-1} and 2.91 ± 0.12 kcal mol^{-1}. In addition, the macrocyclic peptides bound with higher affinity than a fragment of Bid of similar length (14 aa), **Bid14** (K_D = 110,000 ± 40,000 nM), and bound with comparable affinity to a much longer Bid peptide (26 aa), **TAMRA-Bid26** (K_D = 105 ± 11 nM, Figure S4). All cyclic and linear peptides could inhibit binding of the fluorescently labelled **TAMRA-Bid26** to Mcl-1 (Figure 2B), suggesting these peptides bind at the BH3 motif binding site, consistent with the similarity of the RaPID hit sequences to these BH3 motifs. **Heliat-1** (IC$_{50}$ = 0.47 ± 0.07 μM) and **Heliat-2** (IC$_{50}$ = 0.87 ± 0.07 μM), showed an order of magnitude increase in potency relative to their linear counterparts (IC$_{50}$ = 5.4 ± 0.9 μM and IC$_{50}$ = 13.4 ± 1.7 μM, respectively). Both affinity and inhibition results suggest that macrocyclization aids peptide binding to Mcl-1.

Side-chains within the macrocycle are not all critical for binding

Macrocyclic peptides often bind with greater affinity than linear counterparts[15-22] as they are pre-organized to some degree, reducing the entropy cost for binding. One explanation for the higher affinity of **Heliat-1** and **Heliat-2** is that the side-chains within the macrocyclic regions are forming critical interactions with Mcl-1, and macrocyclization pre-organizes them for interaction. To discover which features of the peptides are important for binding, we conducted deep mutational scanning of RaPID hit **1** and **2**, providing an enrichment score (E) for all mutants which reports on the binding strength relative to the wild-type[31]. Many of the amino acids within the linear region were intolerant to mutation, suggesting they are critical for binding Mcl-1 (Figures S5 and S6). This is particularly true for the amino acids which aligned with the BH3 motif, e.g. D13 (Figure 2C, D). In both peptides **1** and **2**, mutations of C15 to hydrophobic amino acids were beneficial for binding (Figure S5 and S6), consistent with the C15 position aligning to a hydrophobic hotspot in the BH3 motif (Figure 1E). However, of the side-chains that are presented by the macrocycle of **1** (S2 and Y3), only Y3 is intolerant to mutation. None of the side-chains that are presented by the macrocycle of **2** (T2, F3) appear critical for binding. It is therefore difficult to conclude that the pre-organization of interactions between the side-chains of macrocyclic region and Mcl-1 is responsible for the increase in affinity observed relative to their linear variants, especially for **Heliat-2**.
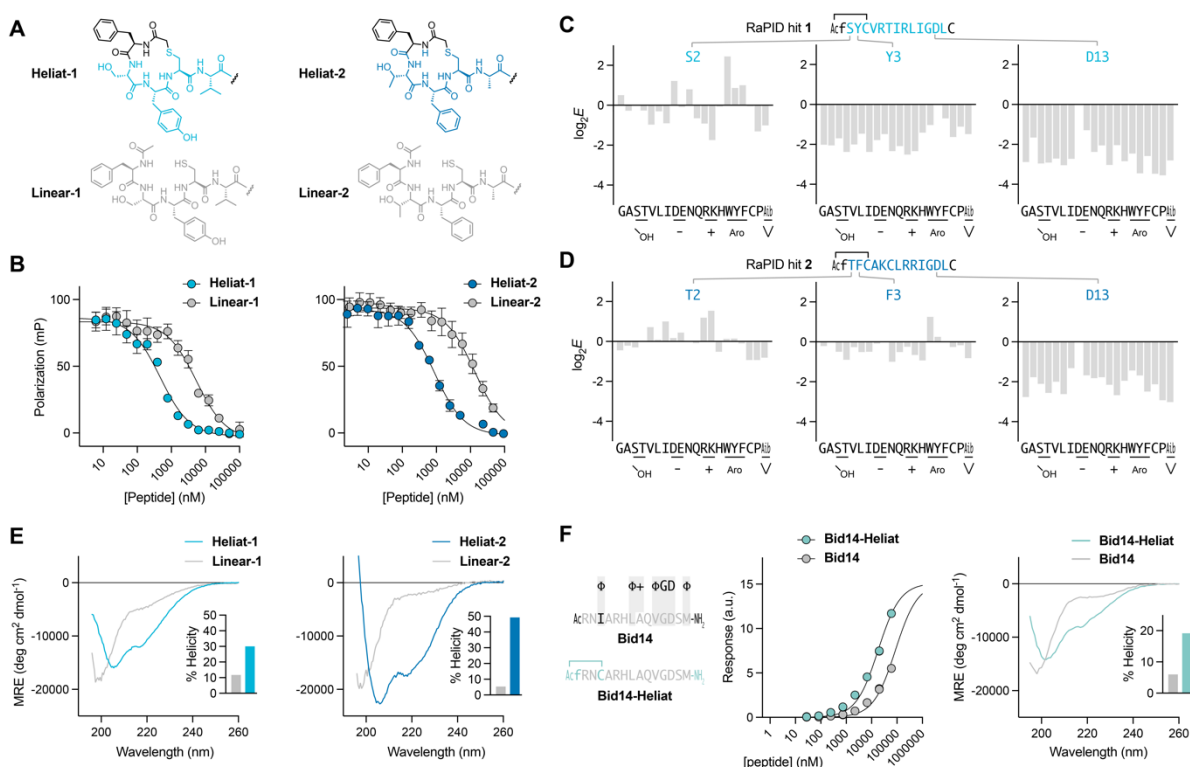
Figure 2. N-terminal macrocycle strengthens binding to Mcl-1, and stabilized an α-helical structure. A)
Linear variants of macrocyclic **Heliat-1** and **Heliat-2,** lacking the cyclizing thioether bond, were synthesized,
Linear-1 and **Linear-2** respectively. **B**) All peptides were able to outcompete 10 nM **TAMRA-Bid26** bound to 80
nM Mcl-1. Macrocyclic peptides showed lower IC_{50} than their linear variants, suggesting the macrocycle
contributes to affinity. **C**) Deep mutational scanning revealed that all mutations to Y3, but not S2, in RaPID hit **1**
resulted in reduced affinity for Mcl-1, producing lower enrichment scores (*E*). Suggesting the side-chain of Y3, but
not S2, is important for binding. Many amino acids that align with BH3 motifs in the linear region, D13 shown, are
largely intolerant to mutation. **D**) Many mutations to T2 and F3 in RaPID hit **2** do not significantly change binding
affinity, suggesting these side-chains are not critical for binding. **E**) Isolated macrocyclic peptides show a
significant negative circular dichroism (CD) signal at 222 nm relative to their linear variants, indicating an increase
in helical secondary structure. MRE = Mean Residue Ellipticity. **F**) A truncation of the human BH3 protein Bid (14
aa) was synthesized, **Bid14**, together with a variant containing an N-terminal 4 amino acid macrocycle, **Bid14-
Heliat**. SPR showed **Bid14-Heliat** bound Mcl-1 with higher affinity than **Bid14**. Isolated **Bid14-Heliat** showed
greater helicity than **Bid14**, suggesting the macrocycle can induce helicity in existing peptides that were not
selected in the presence of the macrocyclic structure.

N-terminal macrocycle is an α-helix inducer

An alternative explanation for how macrocyclization increases binding affinity, is that it pre-
organizes the peptide by stabilizing α-helical conformations, thereby reducing the entropic
cost of binding. To investigate, we measured the circular dichroism (CD) spectra of the
peptides in the absence of Mcl-1. **Heliat-1** and **Heliat-2** showed significantly higher α-helicity
than their linear variants (3- and 10-fold respectively, as judged by CD signal at 222 nm)
(Figure 2E), supporting the idea the macrocycle stabilizes α-helical structure in the peptide.
This CD signal was independent of peptide concentration, suggesting that this helicity was

not due to oligomerization into helical bundles, as has been previously observed for similar BH3 peptides[32] (Figure S7).

Heliat motif increases helicity in a natural BH3 peptide

We investigated whether the minimal 'heliat motif' (D-Phe at the 1st position, cysteine at 4th, and thioether connection through an acetyl group) is sufficient to stabilize helical conformations in a natural BH3 peptide. We synthesized a variant of **Bid14**, **Bid14-Heliat** which includes the heliat motif N-terminal macrocycle (Figure 2F). **Bid14-Heliat** bound Mcl-1 with higher affinity than **Bid14** (K_D = 20 ± 7 μM vs. K_D = 110 ± 40 μM, Figure 2F and S3, Table S2), even though the motif has replaced a hotspot hydrophobic leucine in the natural BH3 motif. **Bid14-Heliat** was also able to inhibit the interaction between **TAMRA-Bid26** and Mcl-1 with a modestly higher potency relative to **Bid14** (Figure S4). Importantly, the isolated **Bid14-Heliat** was more helical by CD than **Bid14** (Figure 2F) confirming that the heliat motif is a helix stabilizer, even for a sequence that has not been generated in the presence of the motif (as was the case for **Heliat-1** and **Heliat-2**).

Both the D-stereocenter and aromatic side-chain of D-Phe1 promote helicity

Both **Heliat-1** and **Heliat-2** have a non-canonical D-phenylalanine at the 1st position of the sequence (D-Phe1), which was a fixed feature during construction of the peptide library. To investigate the contribution of this amino acid to the helix stabilizing effect, we synthesized mutants without the aromatic ring (D-Ala), without the D-stereocenter (Gly), and with the opposite stereochemistry (L-Ala and L-Phe). Interestingly, progressive loss of helicity was observed from D-Phe, D-Ala, Gly, L-Ala to L-Phe, suggesting that both the D-stereocenter and the aromatic ring contribute to the helix stabilizing effect (Figure 3A).

High resolution structure studies suggest a H-bonding pattern consistent with α-helix stabilization

To understand how our heliat motif induces helical structure, we decided to examine the distribution of helical conformations across the peptide sequences. We used TOCSY, ROESY, and ^{1}H,^{13}Cα-HSQC NMR and compared ^{13}Cα chemical shifts of each amino acid to reference values for disordered peptides (Δ^{13}Cα). Unbound **Linear-1** was not sufficiently soluble for NMR. However, spectra could be obtained for unbound **Linear-2**, which showed only small Δ^{13}Cα values, suggesting it is random coil and lacking helical structure (Figure 3B). In contrast, **Heliat-1** and **Heliat-2** showed positive Δ^{13}Cα values characteristic for helical conformation (typically ~3 ppm if 100% helical), particularly within the macrocyclic region. Comparing **Heliat-2** with **Linear-2**, increased helicity was also observed immediately C-terminal of the macrocycle (amino acids 5-10), suggesting the macrocycle can, to some degree, stabilize helical conformations in the linear region of the lariat.
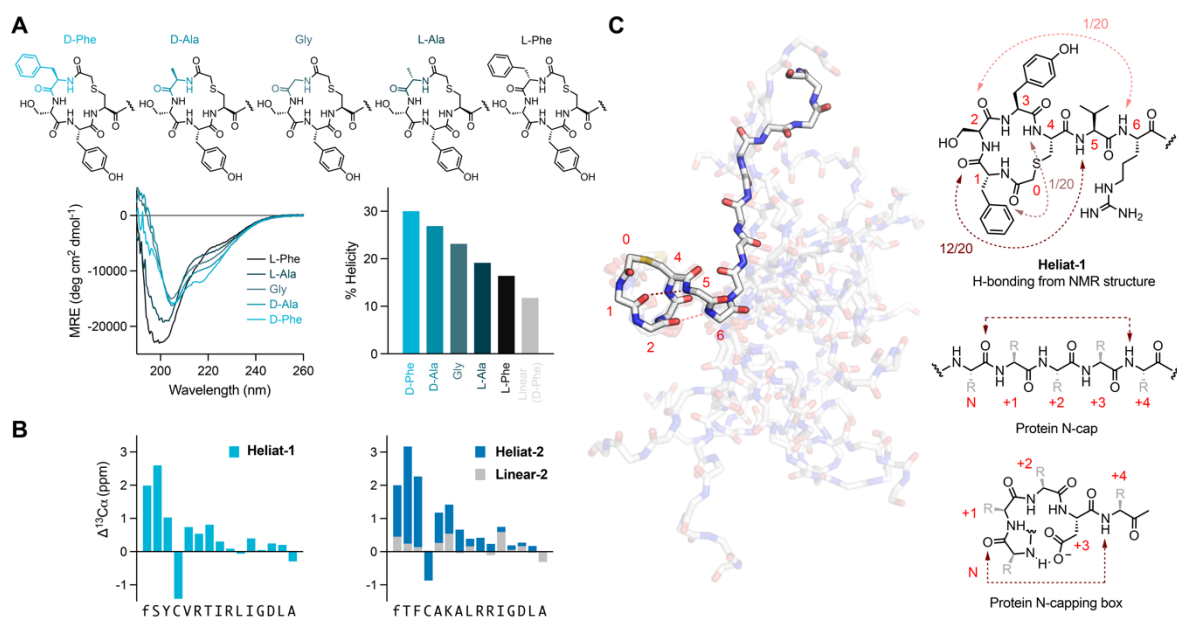
Figure 3. Heliat macrocycle occupies helical conformations in the unbound state and presents H-bond acceptors. A) Variants of **Heliat-1**, with D-Phe1 mutated to D-Ala, Gly, L-Ala, or L-Phe, showed less negative CD signal at 222 nm, indicating a loss of helical structure, suggesting both the aromatic ring and D-stereocenter D-Phe contribute to stabilizing the helical structure. **B) Heliat-1** and **Heliat-2** showed larger $^{13}C\alpha$ chemical shifts than reference values for random coil (positive $\Delta^{13}C\alpha$), with values expected for helical conformation (~3 ppm) in, and near to, the macrocycle. By comparison, **Linear-2** $\Delta^{13}C\alpha$ were small, as expected for a largely random coil. C4 produced negative $\Delta^{13}C\alpha$, this could be due to non-helical conformations, or could be because the thioether-bonded cysteine has a different intrinsic chemical shift to the free-thiol cysteines used for the reference value. **C)** The ROESY NMR structural ensemble of unbound **Heliat-1** reveals a backbone H-bonding pattern in the N-terminal macrocycle. In 12/20 models there is a H-bond between the carbonyl of D-Phe1 and amide hydrogen of V5. In 1/20 models, there is a H-bond between S2 and R6, if the H-bond geometry is slightly relaxed (see methods) this H-bond is present in 16/20 models.

ROESY NMR ensemble of Heliat-1 shows helical hydrogen bonding between macrocycle and linear chain

We next sought atomic-resolution structural information to explain how the heliat motif is able to induce helical structure. ROESY NMR spectra were recorded for unbound **Heliat-1** and used to construct a structural ensemble (Figure 3C). The region around the N-terminal macrocycle was highly structured, showing good alignment between models, with ϕ/ψ angles consistent with an α-helical structure. In contrast, the linear C-terminus was largely random coil in this unbound state. In the majority of the models, backbone hydrogen bonding (H-bonding) could be seen between the carbonyl of D-Phe1 and amide of V5, together with additional *i* to *i+4* H-bonds bridging the macrocycle and the linear region, providing a plausible explanation for how the heliat motif is able to induce helicity and assist binding of the BH3-mimicking linear region.

Molecular dynamics simulations recapitulate the increase in helicity

In order to elucidate the interplay between peptide macrocyclization and helical propensity, we employed molecular dynamics (MD) simulations. We used the REST2 method, which ensures the efficient exploration of the conformational space of our peptides[33] (Figure S8), alongside the recently developed Amber99SB-disp force field and TIP4P-D water model[34], which was benchmarked against a large dataset of NMR data for both folded proteins and IDPs. We found that both **Heliat-1** and **Heliat-2** show a greater helicity than their linear counterparts (Figure 4A and D), particularly at or near the residues of the macrocycle. We have used two other force field-water model combinations and found consistent results (Figure S9). Additionally, back-calculated $\Delta^{13}C\alpha$ values are in overall very good agreement with experiments (Figure 4B). The only notable discrepancy is the chemical shift of C4, the residue whose sidechain is involved in cyclization. This may be due to the limited applicability of SPARTA+[35] to estimate chemical shifts for residues involved in chemical bonding via their sidechains, as is the case of disulfide bonds or our thioether bond[36]. In our simulations, we find that the increased helical propensity in the macrocyclic peptides is mediated by the formation of H-bonds that begin at the macrocycle, as evidenced by the distribution of distances between donor and acceptor atoms (respectively, amide hydrogens of amino acids 4, 5 and 6 and the carbonyl groups of the N-terminal acetyl thioether, D-Phe1 and amino acid 2, see Figure 4C). Comparing **Heliat-2** with **Linear-2**, we observe a greater tendency to have low distances between these pairs of atoms, which is compatible with H-bonding. We hypothesize that the macrocycle arranges the peptide in a favorable conformation to promote H-bonding. This induces a significant local α-helicity that partially propagates to the next residues in the sequence in the unbound state. These trends are also observed for **Heliat-1** and **Linear-1** (Figure S10).
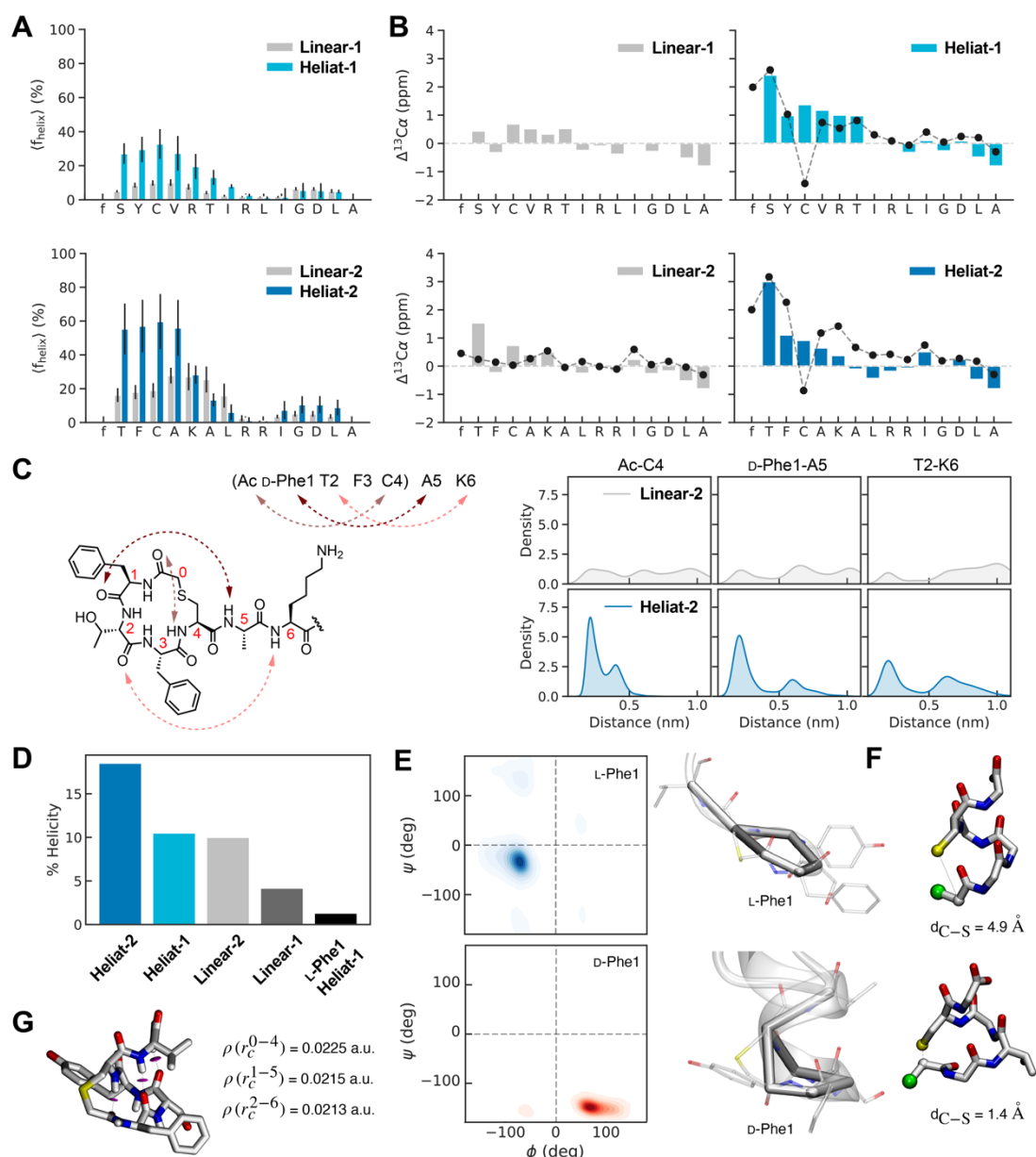
Figure 4. Computational investigations show that D-stereochemistry Phe1 in the macrocycle promotes helix conformations. A) Per-residue fraction of helicity **Heliat-1** (top) and **Heliat-2** (bottom), and their linear variants. **B**) Secondary NMR chemical shifts backcalculated from the MD simulation data (bars) and from experiment (black circles) for the linear and heliat peptides. **C**) Distance distributions for selected H-bond pairs for **Heliat-2** and **Linear-2**. The schematic illustrates the donor amide and acceptor carbonyls in the H-bonds. **D**) Global fraction helix for heliat and linear peptides. **E**) Torsional distribution of Ramachandran angles for D-Phe1 (bottom) and L-Phe1 (top) variant of **Heliat-1**. We show a trace representation of the macrocycle (i.e., including only the cycle Cα atoms and the acetyl carbon atom connected via the thioether bond) for the D and L variants, showing the planar (L-Phe1) and contorted geometries (D-Phe1) of the different diastereomers. **F**) Helical backbone models containing L-Phe1 (top) and D-Phe1 (bottom) illustrate the more optimal distance for D-Phe1 for macrocyclization. **G**) Optimized ROESY model with NCIPlot representation of H-bonds and corresponding electron density values calculated from QTAIM.

Torsional preferences and the helicity of D-Phe and L-Phe variants

Experiments indicated that **Heliat-1** peptide with D-Phe at the 1st position had greater helicity than the L-Phe variant. We simulated the different diastereoisomers for both **Heliat-1** and **Heliat-2** and found L-Phe1 samples the right-handed helical α_R well, whereas, for D-Phe1, the most populated structure is polyproline II (PPII) (with positive values of the ψ torsion, Figure 4E). The choice of enantiomer for Phe1 is highly consequential for the overall arrangement of the torsion angles in the rest of the cycle (Figure S11). We observe that the L-Phe1 structure is more planar than that of D-Phe1, which is more contorted and favours helix-stabilizing interactions (Figure 4E). To rationalize the different structural propensities of the D- and L-Phe at the 1st position and their relation to secondary structure formation in the heliat motif, we have built ideal α-helical models of peptide chains with either ClAc-L-Phe and ClAc-D-Phe, which are the precursors for cyclization, setting all ϕ/ψ angles for amino acids 2-5 in an ideal α_R-helical configuration. For a D-Phe, if residue 1 populates the PPII region (with positive ϕ) the resulting thioether-bond distance between sulfur and acetyl Cα atom is 1.4 Å (Figure 4F), and hence is compatible with cyclization. On the other hand, the L enantiomer at the 1st position with typical α_R ϕ/ψ values results in a distance incompatible with thioether formation (4.9 Å, Figure 4F). Hence, the L-Phe variant requires a distortion of the helicity to form the macrocycle. As a result, the macrocycle adopts a planar conformation that is less ideal for helix-stabilization.

Another interesting detail is that, in the D-Phe1 macrocycle, the acetyl amide is rotated relative to the rest of the heliat motif, making its carbonyl group point in a different direction than the rest. This rotation enables H-bonding to the hydrogen of C4 amide, allowing it to form an *i* to *i+4* H-bond. We have confirmed the ring's conformation and formation of these H-bonds in models derived from ROESY spectra of **Heliat-1** that we have optimized using Quantum Mechanic calculations. Using the Quantum Theory of Atoms In Molecules (QTAIM) to the most stable structures, we have obtained a bond path connecting the atoms of these three H-bonds with well-defined bond critical points (r_c). The values of the electron density $\rho(r_c)$ at these points, which are a measure of the strength of a bond, are between 0.0213 and 0.0225 a.u, close to the value for an ideal backbone H-bond ($\rho(r_c) = 0.0226$ a.u., Figure 4G). Non-covalent interaction plots also show that H-bond interactions are present between these atoms (Figure S12). In this way, we independently support the results from the MD simulations, confirming the existence of H-bonds between the aforementioned donor-acceptor pairs.

Discussion

Cyclic peptides are emerging as a promising modality for drug discovery[37]. N-terminal to side-chain thioether-cyclized peptides are attractive for their *in vivo* stability, their ease of synthesis, and powerful methods for their *de novo* discovery, chiefly the RaPID system[10]. However, it was unclear how such peptides can target proteins that naturally interact with long α-helical peptides, such as the protein Mcl-1. Targeting Mcl-1 using RaPID, we discovered peptides with lariat topologies: small macrocycles at the N-terminus followed by a linear sequence which resembles the BH3 motifs of natural Mcl-1 ligands. The discovered peptides, coined 'Heliats' (*hel*ical lar*iats*), bound with Mcl-1 mid-nM affinity, comparable to other, longer, BH3 and BH3-mimicking peptides[38-41]. The small N-terminal, D-Phe containing

macrocycle aided binding, and structural investigation showed that this macrocycle was able to present H-bond acceptors to participate in *i* to *i+4* H-bonding. We show that the D-stereocenter amino acid at the 1st position, with a N-terminal acetyl thioether-bridge to a cysteine at the 4th, forms a minimal 'heliat motif' capable of stabilizing helical conformations.

The RaPID system has generated peptides for a number of disease-associated proteins, and these have exhibited diverse structures upon binding their targets[15]. This diversity is likely aided by the fact that RaPID can sample range of topologies, from lariats to essentially fully macrocyclic peptides. Helical structures in the bound state have been observed for peptides similar to our heliat motif, but with L-stereochemistry at the 1st position and the cyclizing cysteine at the 5th position[19]. One recent example had L-stereochemistry at 1st position and cysteine at the 4th position[42], a helix was observed in the bound state, with *i* to *i*+4 H-bonding between cycle and linear region, and there was a significant loss of affinity upon mutation of the cyclizing cysteine. This shows that even with non-ideal helical stabilizing of the L-stereochemistry macrocycle, peptides that bind as α-helices can nevertheless be discovered. Currently, there are no published structures for our D-stereochemistry containing motif. However, we note that such sequences have enriched in previous RaPID screens (Figure S13) and we would predict these to have helical structures.

We found that the L-amino acids within our N-terminal macrocycle were highly α-helical. This contrasts with most linear peptides, where helix 'fraying' means that helicity is less stable at the termini relative to the internal positions[43]. The N-terminal helicity in the macrocycle establishes H-bonding to propagate the α-helix to the L-amino acids beyond the cycle. However, the first canonical *i* to *i*+4 H-bond comes from the carbonyl of D-Phe (Figure 3C), making D-Phe the equivalent to the 'N-cap' amino acid found in α-helices of natural proteins. We argue that our heliat motif resembles the 'N-capping box'[44] motif found in natural proteins, where the N-cap H-bonds with an Asp (or Glu) side-chain at N+3 (Figure 3C); as both form a cycle containing 15 atoms, albeit with one non-covalent bond in the N-capping box. Our heliat motif also has some resemblance to the helix-stabilizing disulfide-bonded 'CXXC' motif[45] which occurs in certain natural proteins. While this motif uses 14 atoms instead of 15, it is a covalent macrocycle. Through an *in vitro* selection of a library containing a mixture of macrocyclic topologies, we have found a motif with aspects of natural helix-stabilizing motifs.

The stabilization of peptide helices has been a long-standing goal of peptide chemists, and a range of macrocyclic N-capping strategies have been explored to achieve this (Figure S14)[46]. Most similar to our heliat motif is a rationally designed ring closing metathesis strategy[47-48] to mimic the N-capping box, which also produced a helix-stabilizing macrocycle of 15 atoms. However, such rational N-capping methods require significant synthesis expertise, specialist reagents, often both. Whereas synthesis of our heliat motif requires only minor modification of standard SPPS procedures (coupling of a chloroacetyl group to the N-terminus), and the thioether cyclization is spontaneous later in aqueous solution.

Mixed stereochemistry peptides can provide access to structures and conformations not yet sampled by all L-peptides[49]. Here, we demonstrate that a D-stereocenter within our L-amino acid macrocycle was beneficial for helix stabilization. This adds to examples of other mixed stereocenter motifs that stabilize specific secondary structures, such D-Pro-L-Pro which can stabilize β-hairpins[50]. Yet, the α-helix is the most abundant secondary structure in biology,

and methods to stabilize, mimic and inhibit such structures offer many therapeutic possibilities. Here we report a synthetically facile, helix-stabilizing motif, that was discovered using, and is compatible with, *de novo* peptide discovery. We found that RaPID screening can discover valuable new chemical motifs, in addition to specific molecules for particular targets. We envision a minor adjustment to the RaPID system, fixing cysteine at the 4th position in the initial library, could reliably generate helical peptides for future drug discovery.

Acknowledgements

References

1. Westermarck, J.; Ivaska, J.; Corthals, G. L., Identification of protein interactions involved in cellular signaling. *Mol Cell Proteomics* **2013,** *12* (7), 1752-63.
2. Wright, P. E.; Dyson, H. J., Intrinsically disordered proteins in cellular signalling and regulation. *Nat Rev Mol Cell Biol* **2015,** *16* (1), 18-29.
3. Lu, H.; Zhou, Q.; He, J.; Jiang, Z.; Peng, C.; Tong, R.; Shi, J., Recent advances in the development of protein–protein interactions modulators: mechanisms and clinical trials. *Signal Transduct Target Ther* **2020,** *5* (1).
4. Shin, W. H.; Kumazawa, K.; Imai, K.; Hirokawa, T.; Kihara, D., Current Challenges and Opportunities in Designing Protein-Protein Interaction Targeted Drugs. *Adv Appl Bioinform Chem* **2020,** *13*, 11-25.
5. Wang, X.; Ni, D.; Liu, Y.; Lu, S., Rational Design of Peptide-Based Inhibitors Disrupting Protein-Protein Interactions. *Front Chem* **2021,** *9*.
6. Groß, A.; Hashimoto, C.; Sticht, H.; Eichler, J., Synthetic Peptides as Protein Mimics. *Front Bioeng Biotechnol* **2015,** *3*, 211.
7. Qian, H.; Schellman, J. A., Helix-coil theories: a comparative study for finite length polypeptides. *J Phys Chem* **1992,** *96* (10), 3987-3994.
8. Houk, K. N.; Leach, A. G.; Kim, S. P.; Zhang, X., Binding Affinities of Host–Guest, Protein–Ligand, and Protein–Transition-State Complexes. *Angew Chem Int Ed Engl* **2003,** *42* (40), 4872-4897.
9. Cheng, J.; Zhou, J.; Kong, L.; Wang, H.; Zhang, Y.; Wang, X.; Liu, G.; Chu, Q., Stabilized cyclic peptides as modulators of protein–protein interactions: promising strategies and biological evaluation. *RSC Medicinal Chemistry* **2023,** *14* (12), 2496-2508.

10. Yamagishi, Y.; Shoji, I.; Miyagawa, S.; Kawakami, T.; Katoh, T.; Goto, Y.; Suga, H., Natural product-like macrocyclic N-methyl-peptide inhibitors against a ubiquitin ligase uncovered from a ribosome-expressed de novo library. *Chem Biol* **2011,** *18* (12), 1562-70.

11. Roberts, R. W.; Szostak, J. W., RNA-peptide fusions for the in vitro selection of peptides and proteins. *Proc Natl Acad Sci U S A* **1997,** *94* (23), 12297-302.

12. Murakami, H.; Ohta, A.; Ashigai, H.; Suga, H., A highly flexible tRNA acylation method for non-natural polypeptide synthesis. *Nat Methods* **2006,** *3* (5), 357-359.

13. Goto, Y.; Ohta, A.; Sako, Y.; Yamagishi, Y.; Murakami, H.; Suga, H., Reprogramming the translation initiation for the synthesis of physiologically stable cyclic peptides. *ACS Chem Biol* **2008,** *3* (2), 120-9.

14. Huang, Y.; Wiedmann, M. M.; Suga, H., RNA Display Methods for the Discovery of Bioactive Macrocycles. *Chemical Reviews* **2019,** *119* (17), 10360-10391.

15. McAllister, T. E.; Coleman, O. D.; Roper, G.; Kawamura, A., Structural diversity in de novo cyclic peptide ligands from genetically encoded library technologies. *Pept Sci* **2021,** *113* (1), e24204.

16. Goldbach, L.; Vermeulen, B. J. A.; Caner, S.; Liu, M.; Tysoe, C.; van Gijzel, L.; Yoshisada, R.; Trellet, M.; van Ingen, H.; Brayer, G. D.; Bonvin, A.; Jongkees, S. A. K., Folding Then Binding vs Folding Through Binding in Macrocyclic Peptide Inhibitors of Human Pancreatic α-Amylase. *ACS Chem Biol* **2019,** *14* (8), 1751-1759.

17. Schneider, A. F. L.; Kallen, J.; Ottl, J.; Reid, P. C.; Ripoche, S.; Ruetz, S.; Stachyra, T. M.; Hintermann, S.; Dumelin, C. E.; Hackenberger, C. P. R.; Marzinzik, A. L., Discovery, X-ray structure and CPP-conjugation enabled uptake of p53/MDM2 macrocyclic peptide inhibitors. *RSC Chem Biol* **2021,** *2* (6), 1661-1668.

18. Kodan, A.; Yamaguchi, T.; Nakatsu, T.; Sakiyama, K.; Hipolito, C. J.; Fujioka, A.; Hirokane, R.; Ikeguchi, K.; Watanabe, B.; Hiratake, J.; Kimura, Y.; Suga, H.; Ueda, K.; Kato, H., Structural basis for gating mechanisms of a eukaryotic P-glycoprotein homolog. *Proc Natl Acad Sci U S A* **2014,** *111* (11), 4049-54.

19. Patel, K.; Walport, L. J.; Walshe, J. L.; Solomon, P. D.; Low, J. K. K.; Tran, D. H.; Mouradian, K. S.; Silva, A. P. G.; Wilkinson-White, L.; Norman, A.; Franck, C.; Matthews, J. M.; Guss, J. M.; Payne, R. J.; Passioura, T.; Suga, H.; Mackay, J. P., Cyclic peptides can engage a single binding pocket through highly divergent modes. *Proc Natl Acad Sci U S A* **2020,** *117* (43), 26728-26738.

20. Dai, S. A.; Hu, Q.; Gao, R.; Blythe, E. E.; Touhara, K. K.; Peacock, H.; Zhang, Z.; von Zastrow, M.; Suga, H.; Shokat, K. M., State-selective modulation of heterotrimeric Galphas signaling with macrocyclic peptides. *Cell* **2022,** *185* (21), 3950-3965 e25.

21. Norman, A.; Franck, C.; Christie, M.; Hawkins, P. M. E.; Patel, K.; Ashhurst, A. S.; Aggarwal, A.; Low, J. K. K.; Siddiquee, R.; Ashley, C. L.; Steain, M.; Triccas, J. A.; Turville, S.; Mackay, J. P.; Passioura, T.; Payne, R. J., Discovery of Cyclic Peptide Ligands to the SARS-CoV-2 Spike Protein Using mRNA Display. *ACS Cent Sci* **2021,** *7* (6), 1001-1008.

22. Franck, C.; Patel, K.; Walport, L. J.; Christie, M.; Norman, A.; Passioura, T.; Suga, H.; Payne, R. J.; Mackay, J. P., Discovery and characterization of cyclic peptides selective for the C-terminal bromodomains of BET family proteins. *Structure* **2023,** *31* (8), 912-923 e4.

23. Rautureau, G. J.; Day, C. L.; Hinds, M. G., Intrinsically disordered proteins in bcl-2 regulated apoptosis. *Int J Mol Sci* **2010,** *11* (4), 1808-24.

24. Widden, H.; Placzek, W. J., The multiple mechanisms of MCL1 in the regulation of cell fate. *Commun Biol* **2021,** *4* (1), 1029.

25. Clohessy, J. G.; Zhuang, J.; de Boer, J.; Gil-Gómez, G.; Brady, H. J., Mcl-1 interacts with truncated Bid and inhibits its induction of cytochrome c release and its role in receptor-mediated apoptosis. *J Biol Chem* **2006,** *281* (9), 5750-9.

26. Singh, R.; Letai, A.; Sarosiek, K., Regulation of apoptosis in health and disease: the balancing act of BCL-2 family proteins. *Nat Rev Mol Cell Biol* **2019,** *20* (3), 175-193.

27. Adams, J. M., Ways of dying: multiple pathways to apoptosis. *Genes & Development* **2003,** *17* (20), 2481-2495.

28. Negi, A.; Murphy, P. V., Development of Mcl-1 inhibitors for cancer therapy. *Eur J Med Chem* **2021,** *210*, 113038.

29. Wang, H.; Guo, M.; Wei, H.; Chen, Y., Targeting MCL-1 in cancer: current status and perspectives. *J Hematol Oncol* **2021,** *14* (1), 67.

30. Iwasaki, K.; Goto, Y.; Katoh, T.; Suga, H., Selective thioether macrocyclization of peptides having the N-terminal 2-chloroacetyl group and competing two or three cysteine residues in translation. *Org Biomol Chem* **2012,** *10* (30), 5783.

31. Rogers, J. M.; Passioura, T.; Suga, H., Nonproteinogenic deep mutational scanning of linear and cyclic peptides. *Proc Natl Acad Sci U S A* **2018,** *115* (43), 10959-10964.

32. Assafa, T. E.; Nandi, S.; Śmiłowicz, D.; Galazzo, L.; Teucher, M.; Elsner, C.; Pütz, S.; Bleicken, S.; Robin, A. Y.; Westphal, D.; Uson, I.; Stoll, R.; Czabotar, P. E.; Metzler-Nolte, N.; Bordignon, E., Biophysical Characterization of Pro-apoptotic BimBH3 Peptides Reveals an Unexpected Capacity for Self-Association. *Structure* **2021,** *29* (2), 114-124.e3.

33. Wang, L.; Friesner, R. A.; Berne, B. J., Replica Exchange with Solute Scaling: A More Efficient Version of Replica Exchange with Solute Tempering (REST2). *J Phys Chem B* **2011,** *115* (30), 9431-9438.

34. Robustelli, P.; Piana, S.; Shaw, D. E., Developing a molecular dynamics force field for both folded and disordered protein states. *Proc Natl Acad Sci U S A* **2018,** *115* (21), E4758-E4766.

35. Shen, Y.; Bax, A., SPARTA+: a modest improvement in empirical NMR chemical shift prediction by means of an artificial neural network. *J Biomol NMR* **2010,** *48* (1), 13-22.

36. Paissoni, C.; Nardelli, F.; Zanella, S.; Curnis, F.; Belvisi, L.; Musco, G.; Ghitti, M., A critical assessment of force field accuracy against NMR data for cyclic peptides containing β-amino acids. *Phys Chem Chem Phys* **2018,** *20* (23), 15807-15816.

37. Ji, X.; Nielsen, A. L.; Heinis, C., Cyclic Peptides for Drug Development. *Angew Chem Int Ed Engl* **2024,** *63* (3), e202308251.

38. Ku, B.; Liang, C.; Jung, J. U.; Oh, B. H., Evidence that inhibition of BAX activation by BCL-2 involves its tight and preferential interaction with the BH3 domain of BAX. *Cell Res* **2011,** *21* (4), 627-41.

39. Muppidi, A.; Doi, K.; Edwardraja, S.; Drake, E. J.; Gulick, A. M.; Wang, H. G.; Lin, Q., Rational design of proteolytically stable, cell-permeable peptide-based selective Mcl-1 inhibitors. *J Am Chem Soc* **2012,** *134* (36), 14734-7.

40. Foight, G. W.; Ryan, J. A.; Gullá, S. V.; Letai, A.; Keating, A. E., Designed BH3 peptides with high affinity and specificity for targeting Mcl-1 in cells. *ACS Chem Biol* **2014,** *9* (9), 1962-8.

41. Rezaei Araghi, R.; Ryan, J. A.; Letai, A.; Keating, A. E., Rapid Optimization of Mcl-1 Inhibitors using Stapled Peptide Libraries Including Non-Natural Side Chains. *ACS Chem Biol* **2016,** *11* (5), 1238-1244.

42. Bertran, M. T.; Walmsley, R.; Cummings, T.; Valle Aramburu, I.; Benton, D. J.; Assalaarachchi, J.; Chasampalioti, M.; Swanton, T.; Joshi, D.; Federico, S.; Okkenhaug, H.; Yu, L.; Oxley, D.; Walker, S.; Papayannopoulos, V.; Suga, H.; Christophorou, M. A.; Walport, L. J., A cyclic peptide toolkit reveals mechanistic principles of peptidylarginine deiminase IV (PADI4) regulation. *BioRxiv* **2023**.

43. Baldwin, R. L., Alpha-helix formation by peptides of defined sequence. *Biophys Chem* **1995,** *55* (1-2), 127-35.

44. Harper, E. T.; Rose, G. D., Helix stop signals in proteins and peptides: the capping box. *Biochemistry* **1993,** *32* (30), 7605-9.

45. Iqbalsyah, T. M.; Moutevelis, E.; Warwicker, J.; Errington, N.; Doig, A. J., The CXXC motif at the N terminus of an alpha-helical peptide. *Protein Sci* **2006,** *15* (8), 1945-50.

46. Whisenant, J.; Burgess, K., Synthetic helical peptide capping strategies. *Chem Soc Rev* **2022,** *51* (14), 5795-5804.

47. Pham, T. K.; Kim, Y. W., Helix stabilization by stapled N-capping box. *Bioorg Chem* **2020,** *101*, 104024.

48. Nguyen, L. T.; Luong, H. X.; Kim, Y.-W., Helix Nucleation via Hydrocarbon Cross-link Mimicking N-capping Box. *Bull Korean Chem Soc* **2016,** *37* (4), 566-570.

49. Bhardwaj, G.; Mulligan, V. K.; Bahl, C. D.; Gilmore, J. M.; Harvey, P. J.; Cheneval, O.; Buchko, G. W.; Pulavarti, S. V.; Kaas, Q.; Eletsky, A.; Huang, P. S.; Johnsen, W. A.; Greisen, P. J.; Rocklin, G. J.; Song, Y.; Linsky, T. W.; Watkins, A.; Rettie, S. A.; Xu, X.; Carter, L. P.; Bonneau, R.; Olson, J. M.; Coutsias, E.; Correnti, C. E.; Szyperski, T.; Craik, D. J.; Baker, D., Accurate de novo design of hyperstable constrained peptides. *Nature* **2016,** *538* (7625), 329-335.

50. Robinson, J. A., Beta-hairpin peptidomimetics: design, structures and biological activities. *Acc Chem Res* **2008,** *41* (10), 1278-88.