REVIEW ARTICLE

Cyclisation strategies for stabilising peptides with irregular conformations

Received 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

Quynh Ngoc Vu,[a]† Reginald Young,[a]† Haritha Krishna Sudhakar,[a]† Tianyi Gao,[a] Tiancheng Huang, [a] Yaw Sing Tan[b] and Yu Heng Lau*[a]

Cyclisation is a common synthetic strategy for enhancing the therapeutic potential of peptide-based molecules. While there are extensive studies on peptide cyclisation for reinforcing regular secondary structures such as α-helices and β-sheets, there are remarkably few reports of cyclising peptides which adopt irregular conformations in their bioactive target-bound state. In this review, we highlight examples where cyclisation techniques have been successful in stabilising irregular conformations, then discuss how the design of cyclic constraints for irregularly structured peptides can be informed by existing β-strand stabilisation approaches, new computational design techniques, and structural principles extracted from cyclic peptide library screening hits. Through this analysis, we demonstrate how existing peptide cyclisation techniques can be adapted to address the synthetic design challenge of stabilising irregularly structured binding motifs.

1. Introduction

Peptide-based therapeutics are currently experiencing a renaissance in drug discovery, offering a compelling alternative modality for inhibitor design that complements mainstream small molecule approaches.[1, 2] The size and functionality of peptides are well suited to forming intermolecular interactions with extended or shallow binding pockets, such as those found at protein-protein interaction (PPI) interfaces.[3-7] Many receptors and enzymes also bind peptides as their native substrate, serving as direct inspiration for rational inhibitor design.

Starting from a native peptide ligand, cyclisation is a common modification strategy used to constrain a peptide into its bioactive target-bound conformation and thus enhance its pharmacological properties. In addition to improving binding affinity and selectivity,[8-11] the cyclisation process can improve resistance to metabolic degradation and enable access to intracellular targets in some optimised cases.[12-15] To date, rational peptide cyclisation methodologies have primarily focused on reinforcing well-defined secondary structure elements, such as peptide stapling to inhibit helix-mediated PPIs.[16-19]

Analyses of the Protein Data Bank (PDB) have shown that approximately half of all PPI interfaces are in fact mediated by irregularly structured binding motifs.[20-22] Peptides with irregular conformations are especially common at the binding interfaces of heterodimeric complexes. In the case of enzyme–inhibitor complexes, both (23%) or at least one (48%) of the two partners predominantly display irregularly-structured motifs at the binding interface.[21] Given the widespread prevalence of irregular conformations at potentially druggable binding interfaces, there is a clear under-representation of cyclisation approaches that are designed to stabilise such peptide conformations. Existing cyclisation chemistries, which have been optimised for α-helices, β-strands, and β-turns, cannot be used directly without adaptation to stabilise irregular structures in a predictable manner. Hence, a deeper structural understanding of how different cyclic constraints impact the geometry of non-standard backbone conformations is needed to address this large set of potential therapeutic targets.

The purpose of this review is to highlight how existing cyclisation techniques can be adapted for conformational stabilisation of irregularly structured peptides found at the binding interfaces of target proteins. Firstly, we analyse important examples from the limited set of studies that have reported successful stabilisation of such conformations. We then draw comparisons to established cyclisation chemistries that are optimised for β-strands, outline several computational methods that have been used to identify promising starting points and guide structure-based design choices, and consider the implications of trends found in the irregular structures of cyclic peptides that have been discovered through screening and selection of cyclic peptide libraries. Through the examples chosen, we convey the wide range of potential strategies and opportunities available to medicinal chemists when confronting this unaddressed class of PPIs for drug development.

[a] School of Chemistry, Eastern Ave, The University of Sydney, NSW 2006, Australia
[b] Bioinformatics Institute, Agency for Science, Technology and Research (A*STAR), 30 Biopolis Street, #07-01, Matrix, Singapore 138671
† These authors contributed equally to this manuscript.

2. Rational stabilisation of irregular peptide conformations

In this review, we interpret the term *irregular conformation* as pertaining to peptides in their target-bound state that do not have a single well-defined secondary structure motif (ie. α-helix or related helices, β-strand/turn/hairpin) that dominates across the majority of residues in the sequence.

The primary challenge associated with stabilising irregular conformations is that there is inherently no single cyclic constraint that can be generally applied across a range of peptides and targets. Unlike helices, strands, and turns, the variability of intramolecular distances requires significant optimisation of ring size through experimental screening or structure-based design to achieve the desired stabilisation effect. In addition to ring size, other important considerations include determining the appropriate positions to anchor the cyclic linkage, the degree of flexibility needed for binding, interactions of the cyclic linkage with the target, and potential to form stabilising intramolecular interactions such as internal hydrogen bonding.

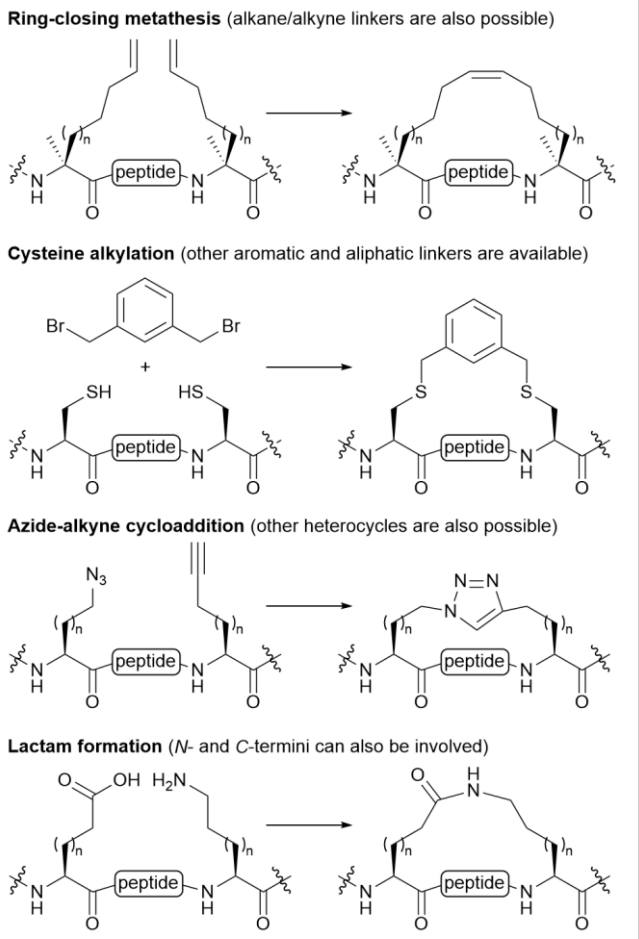

Figure 1. Selected examples of common peptide cyclisation chemistries used for conformational stabilisation. Comprehensive lists of examples can be found in reviews that focus on stabilisation of common secondary structures.[13,23,24]

While the extensive range of peptide cyclisation chemistries available has been the subject of numerous prior reviews (covering reactions such as cysteine alkylation and arylation, cycloaddition reactions, ring-closing metathesis, lactam formation *via* side-chains or *N*- and *C*-termini; see Figure 1 for examples),[13, 23, 24] the following discussion specifically showcases how existing peptide cyclisation chemistries can be tailored to fit irregularly structured peptides.

2.1 ExoS/14-3-3 – adapting hydrocarbon stapling chemistry

An early example that explicitly addressed the challenge of stabilising irregular peptide conformations was reported by Grossmann and co-workers, using peptides cyclised by ring-closing metathesis to inhibit the interaction between the human adaptor protein 14-3-3 and virulence factor exoenzyme S (ExoS) found in *Pseudomonas aeruginosa.*[25]

The key feature enabling rational design of a cyclic crosslink was the hybrid structure of the target-bound ExoS-derived peptide ESp, consisting of an *N*-terminal helical turn followed by an irregularly structured *C*-terminal tail (Figure 2a). The stabilisation strategy involved cyclising the helical *N*-terminus, relying on the well-established chemistry of hydrocarbon stapling[26] to propagate overall conformational stabilisation through the entire 11-residue peptide. The extremely hydrophobic nature of the ESp binding surface allowed for an unconventional choice of residues for stapling, where the binding residues were directly replaced with the hydrocarbon linker. This is a departure from the more common strategy of stapling at less critical residues that do not point into the binding pocket.

Despite the existence of design rules for hydrocarbon stapling within the context of α-helices, significant experimental optimisation was still necessary due to the dual functions of the hydrocarbon linker as both a conformational constraint and a hydrophobic binding motif. After varying the positioning, stereochemistry, and side-chain length of the olefinic amino acids, two optimal solutions were found. Both cyclised peptides used the same *i*,*i*+3 positioning of non-natural amino acids, but different absolute configurations (RS vs SS) and linker lengths (8 vs 12 atoms).

A crucial observation was the clear dependence of peptide conformation on linker structure (Figure 2a). The orientation of the shorter 8-atom linker (**β_{RS}8**, PDB 4N7Y) caused significant distortion in the *N*-terminus away from native helix of ESp with the concomitant loss of a key water molecule, while the longer 12-atom linker (**β_{SS}12**, PDB 4N84) preserved the native binding mode, leading to the most potent binding affinity of 150 nM as measured by isothermal calorimetry (ITC). Despite backbone distortion in the case of the shorter linker, the enthalpic losses were offset by the entropic gains from the conformational constraint and the liberation of water, resulting in a net increase in affinity over the wild-type ESp. Complementary evidence of the reduction in conformational flexibility was obtained by

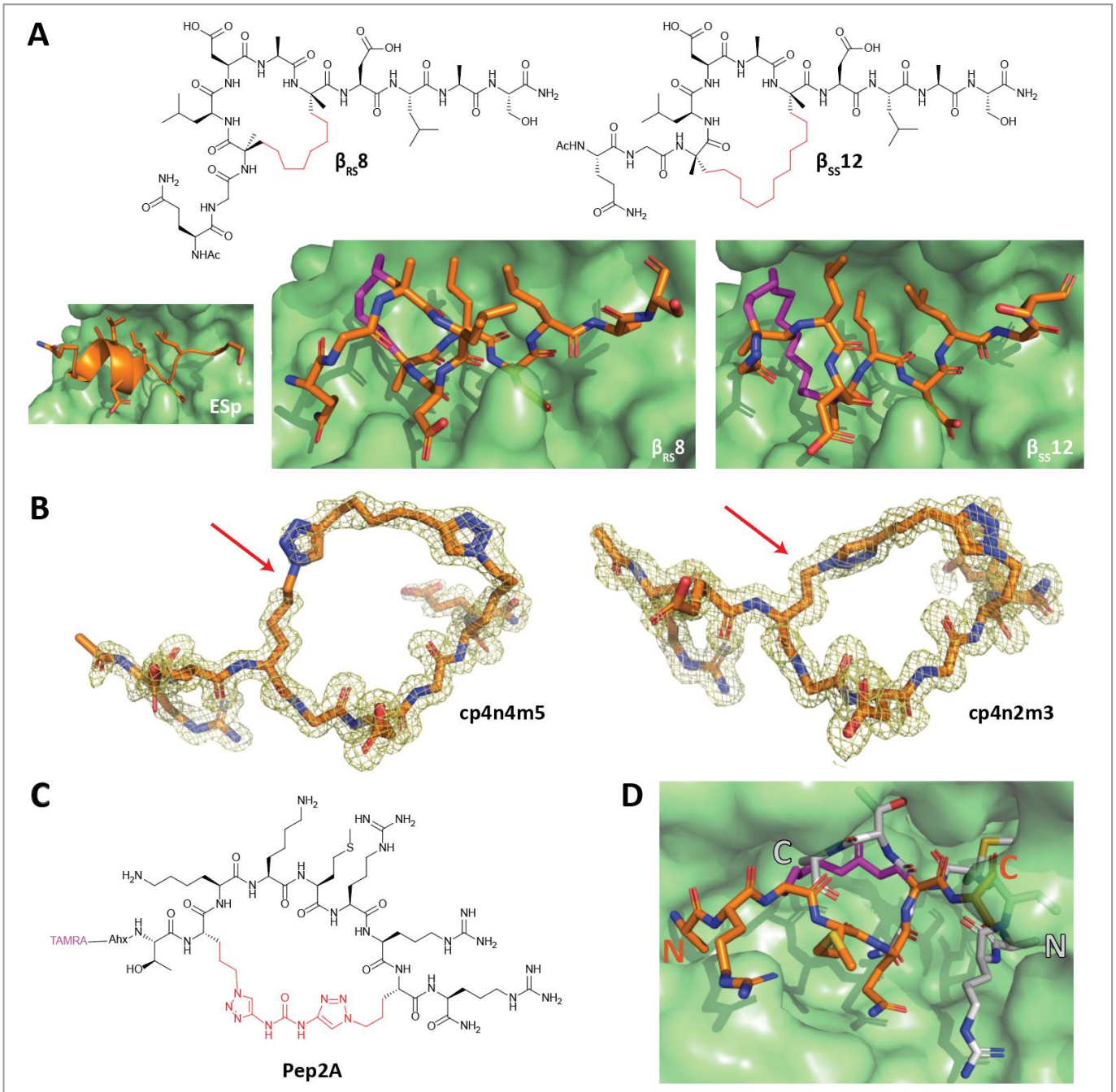

Figure 2. Examples of rational stabilisation of irregular peptide conformation. **a)** The impact of hydrocarbon-stapling on the conformation of the ESp peptide from virulence factor exoenzyme S (ExoS), which has a combination of helical and irregular structure.[25] Depending on linker length, cyclisation of the helical region can either preserve (**β_{SS}12**, PDB 4N84) or distort (**β_{RS}8**, PDB 4N7Y) the native conformation. **b)** The impact of linker length on the binding of stapled peptides targeting TNKS2.[30] Differences in the electron density maps ($2F_{obs}$-F_{calc}, shown as yellow mesh) of peptides **cp4n4m5** and **cp4n2m3** (PDB 5BXU/5BXO), especially near the staple at the region indicated by the red arrows, suggests differences in flexibility. **c)** Structure of **Pep2A** with a urea-bridged bis-triazole linker (red) that constrains the extended conformation of an intrinsically disordered region of the hepatocyte nuclear factor 1β transcription factor (HNF1β).[31] **d)** Comparison of binding poses for linear peptide **$H3_{1-21}$ K4M** (white) and lactam cyclised peptide, macrocycle **31** (orange).[32] Macrocycle **31** has an opposite backbone direction to the linear peptide and is partially protruding out from the binding pocket. The *N*- and *C*-terminus are labelled in grey and orange for **$H3_{1-21}$ K4M** and macrocycle **31** respectively (PDB 2V1D/6S35), with the lactam staple in magenta.

tracking changes in the circular dichroism (CD) spectra and chemical shifts of the relevant amide protons by NMR. The underlying thermodynamic factors may involve greater complexity than a simple conformational restriction however, as further experimental and molecular dynamics studies on ^{19}F-labelled analogues revealed that flexibility in the bound state may also play a role in driving the overall affinity.[27]

Together with a follow-up study using alkyne-containing staples,[28] this work from Grossmann and co-workers provides a structurally well-defined example in which peptide cyclisation constrains conformation, either reinforcing or distorting the amide backbone in a linker-dependent manner.

2.2 Tankyrase and importin-α – adapting double-click chemistry

Spring and co-workers adapted Cu-catalysed double-click stapling chemistry[29] to develop macrocyclic peptides against two cancer-related PPI targets involving completely irregular and extended binding conformations – the poly(ADP-ribose) polymerase tankyrase[30] and the nuclear transport protein importin α.[31] Similar to the previous example on ESp peptides, significant linker exploration was required to find the optimal chemical constraint to fit the irregular peptide structures.

Tankyrases bind a range of protein partners *via* their substrate recognition domain to mediate cellular processes such as control over the mitotic checkpoint and Wnt signalling. Various substrate-derived peptides are known to bind to this domain in an extended conformation. Starting with an 8-residue consensus sequence of the known peptide binders, Xu *et al.* synthesised a library of cyclised peptide variants by varying the staple position, linker length, and scaffold rigidity.[30]

In this study, crystallography played a crucial role in the optimisation process. Peptides from an initial cyclisation library were unable to improve upon the affinity of the linear consensus peptide, with the 1.35 Å resolution structure of tankyrase 2 bound by the highest affinity cyclic peptide (**cp4n4m5**, PDB 5BXU, Figure 2b) revealing that the screened ring sizes were too large for conformational restriction, as indicated by poorly resolved electron density for the linker component. Removing two methylene units from each of the non-natural azido amino acids resulted in cyclic peptides with improved binding affinity and increased rigidity, as reflected by more well-defined electron density in a corresponding crystal structure (**cp4n2m3**, PDB 5BXO). Furthermore, ITC confirmed a reduction in the entropic cost upon binding when compared to the parent consensus sequence.

Further shortening of the linker beyond the optimal length had a dramatic effect on peptide conformation and tankyrase binding. While the optimal linkage was constructed through double-click cycloaddition between two azidohomoalanine residues and a 1,3-diethynylbenzene linker, shortening the side-chains to azidoalanine resulted in over-restriction of the peptide into an α-helical conformation as observed by CD, accompanied by complete loss of detectable binding. Along with the studies in Section 2.1, these findings highlight the need for the linker to be in a 'goldilocks' zone to produce a high affinity peptide with the correct geometry.

With an optimal peptide in hand, Xu *et al.* showed that the cyclisation process improved proteolytically stability, and that the cyclised peptide was able to disrupt the tankyrase-axin interaction in an *in vitro* pulldown assay. To obtain effective inhibition of a Wnt pathway reporter in HEK 293T cells however, conjugation to the cell-penetrating peptide Antennapedia was required.

An example of a cell-permeable constrained peptide with an irregular structure was reported by Wiedmann *et al.*, who applied a similar double-click strategy to cyclise the nuclear localisation signal (NLS) region of the hepatocyte nuclear factor 1β transcription factor (HNF1β),[31] an intrinsically disordered peptide that adopts an extended conformation upon binding to importin-α for translocation across the nuclear membrane. After screening several linker lengths, a urea-bridged dialkynyl linker was found to have optimal binding (**Pep2A,** Figure 2c) while retaining the ability of the NLS to enter cells by virtue of its highly cationic sequence.

Both these studies highlight the opportunities and challenges when cyclising peptides with irregular extended conformations. In the absence of substantial enthalpic contributions from the linker itself, the entropic gains only arise with precise matching of the cyclic constraint to the required configuration. Furthermore, the ability to engage intracellular targets is somewhat contingent on the sequence itself, and not necessarily a property that arises from cyclisation alone.

2.3 LSD1 inhibitors – comparing different cyclisation chemistries

A recent example of stabilising an irregularly structured interaction between an enzyme and its protein substrate was the development of inhibitors for the epigenetic enzyme lysine-specific demethylase 1 (LSD1) by Danielson, Kihlberg and co-workers.[32] Similar to tankyrase, there are a numerous substrate-derived peptides that are known to bind the active site. Two different classes of irregular binding conformations had previously been observed in crystal structures of the bound state, involving an *N*-terminal helical turn consisting of either two β-turns of three γ-turns and a short *C*-terminal extension. The authors based their choice of cyclisation chemistries on this existing structural data, exploring ring-closing metathesis for *N*-terminal *i,i+2* cyclisation of the peptide conformation featuring β-turns, and both Cu-catalysed click triazole formation and lactamisation at an internal *i,i+3* position for the conformation with γ-turns. The study involved extensive exploration of multiple chemical parameters apart from cyclisation chemistry, including peptide truncation and point mutants, giving rise to complex structure-activity relationship (SAR) trends in both the binding affinity data obtained by surface plasmon resonance (SPR) and inhibition data from enzyme activity assays.

One of the most unusual findings arose from the target-bound crystal structure of their most potent lactam-cyclised peptide (macrocycle **31**, PDB 6S35, Figure 2d) in activity assays. Unlike the related linear and untruncated peptide (**H3$_{1-21}$K4M**, PDB 2V1D), the cyclised peptide adopted a vastly altered binding mode, flipping backbone direction and rotating away from the active site, thus partially protruding out from the binding pocket. The lactam bridge was observed to point into the pocket, forming hydrogen bonding interactions with nearby residues, suggesting that the structure of the pocket may have been unable to accommodate the lactam bridge while preserving the native peptide conformation.

The unexpected structural finding and complex SAR reveals the idiosyncratic nature of designing constraints for irregularly structured peptides. In comparison to small molecules, the highly dynamic nature of peptide binding and the sheer number of conformational possibilities can lead to unexpected discoveries and trends which are challenging to decipher and optimise in a rational manner. Nevertheless, the authors here were able to identify a promising cyclic peptide candidate for further lead development.

2.4 Summary of reported examples

While there is minimal structural similarity between the peptides and targets in each study discussed in this section, there are numerous recurring themes in the experimental approaches and results. Existing cyclisation techniques can be adapted for irregular peptides, but significant experimental optimisation is needed to identify the most suitable chemical constraint for a given conformation, with crystal structures of the bound complexes playing a crucial role in understanding the effects of linker SAR. Constraints that are too small lead to warping of the amide backbone, while overly large constraints can result in excess flexibility and poor binding. Furthermore, the observed affinity and activity gains due to entropic factors are modest, unless the linker itself also makes significant favourable interactions with the target. Despite these modest gains, there are also improvements in stability against degradation upon cyclisation. Cell permeability of irregularly structured peptides for intracellular targets has yet to be explored in depth, and is known to be idiosyncratic in many cases,[33] but may be influenced by factors such as the sequence identity and length, in addition to the conformation itself. Nevertheless, the success to date indicates that pursuing such challenging targets may be feasible, requiring sufficient throughput to explore chemical modifications, coupled with the capability for structural characterisation and robust binding and activity assays.

3. Lessons from cyclic β-strand peptides

Building on the successful studies discussed in Section 2 which adapted helix-stabilising chemistries (ring-closing metathesis, click triazole formation, lactamisation), we expect that existing cyclisation methods for stabilising other secondary structures should have sufficiently broad scope to be used for irregular conformations. In particular, strategies to constrain extended β-strand scaffolds may be well-suited to the intramolecular distances required for constraining irregular loops and extended strands, with optimisation to account for any relevant structural differences such as backbone dihedral angles or hydrogen bonding networks.

β-strand motifs were some of the earliest secondary structure elements to have been subjected to cyclisation, due to their important role in recognition events and enzymatic processing.[34] A large body of work has been dedicated to targeting proteolytic enzymes which universally recognise peptide substrates with a regular extended backbone conformation.[24, 35, 36] Illustrating this, β-strand mimetics have been reported for proteases involved in disease,[37-40] as well as other non-proteolytic enzymes that have protein substrates such as kinases and other transferases.[41-43]

Chemical approaches to β-strand mimicry have been the subject of comprehensive reviews in the past, covering a large number of possible peptide cyclisation motifs (eg. thioethers, biaryl ethers, lactams, lactones, hydrocarbon linkages)[36, 42, 44] as well as completely non-peptidic scaffolds as mimetics.[45] The primary focus here is on a small subset of instructive examples which retain most of the peptide framework and use cyclisation to restrict the endocyclic dihedral angles of the strand backbone.

3.1 Macrocyclic tethers

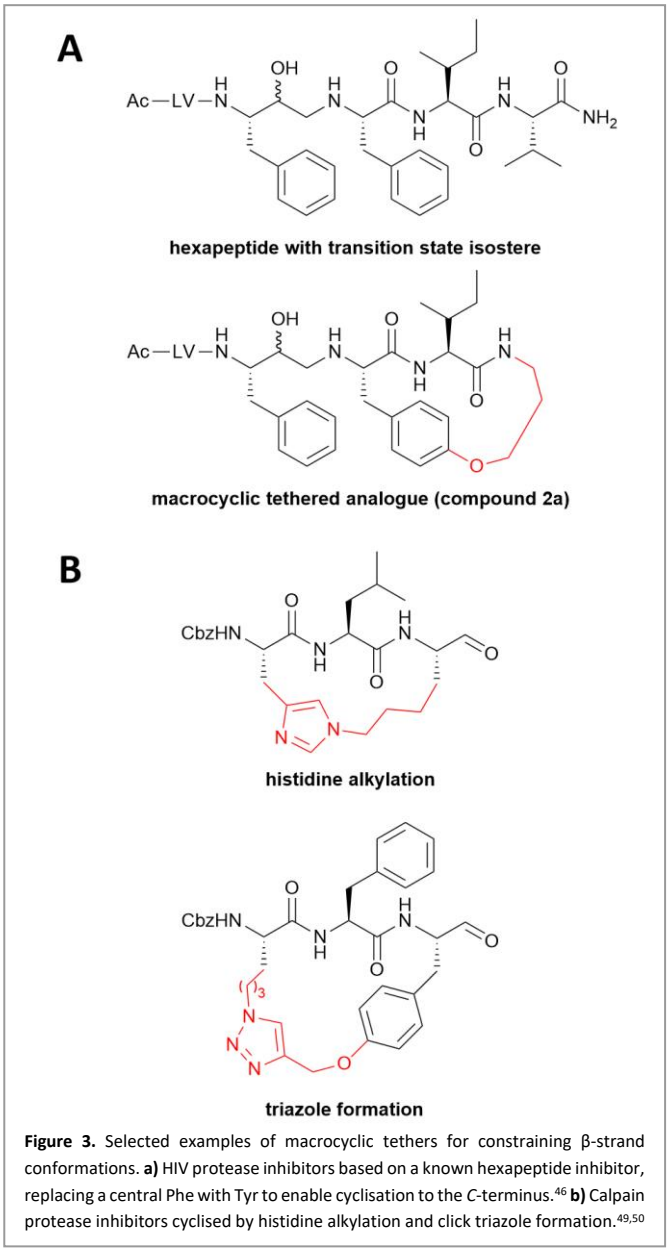

Figure 3. Selected examples of macrocyclic tethers for constraining β-strand conformations. **a)** HIV protease inhibitors based on a known hexapeptide inhibitor, replacing a central Phe with Tyr to enable cyclisation to the *C*-terminus.[46] **b)** Calpain protease inhibitors cyclised by histidine alkylation and click triazole formation.[49,50]

Macrocyclic tethers have been widely utilised to constrain bioactive peptides into β-strand conformations, typically by installing a covalent linkage between *i,i+2* side-chains residing on the same face of the strand. In the well-studied case of protease inhibitors, a wide variety of anchoring amino acids has been explored, including aromatic residues such as Tyr and His which have seen relatively less use in the context of helix stapling.

Early work on macrocyclic tethering for designing HIV protease inhibitors was conducted by Fairlie and co-workers,[46] inspired by previous cyclic peptide mimetics of other proteases such as renin.[47] Starting with a known transition-state analogue of a hexapeptide substrate (Figure 3a), the β-strand conformation was reinforced by replacement of a central Phe residue with Tyr and subsequent ether bridging to the *C*-terminus *via* an alkyl linkage (compound **2a**). Structures of the resulting cyclised peptides showed that native conformation and inhibitory activity were retained, while there was an increase in lipophilicity and resistance to degradation for the cyclic *C*-terminal portion of the inhibitors, although the acyclic *N*-terminal end remained vulnerable to hydrolysis.

More recently, Abell and co-workers reported a series of calpain protease inhibitors, expanding on the cyclisation chemistries available for β-strands. After initially exploring the use of the same tyrosine ether strategy as Fairlie and co-workers,[48] the authors explored alternative linker chemistries, including a click-derived triazole to replace the alkyl linker[49] and replacing Tyr with His for alkylation (Figure 3b).[50] Upon assaying the triazole series of cyclic peptides for inhibition of calpain, activity was correlated with the stabilisation of β-strand conformation as determined by NMR studies. However, conformational restriction was not the dominant factor affecting activity, as the presence of a *C*-terminal aldehyde functionality was required to drive potency by serving as a covalent warhead.

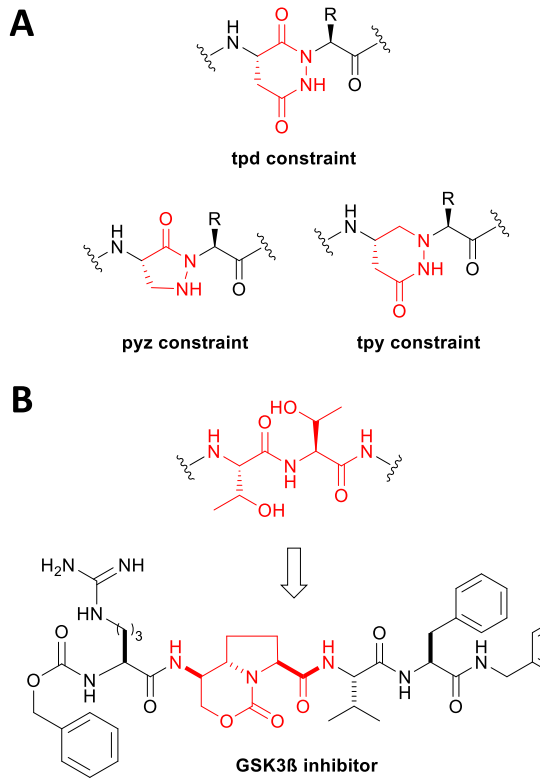

Figure 4. Amide backbone peptidomimetic isosteres based on heterocycles. **a)** Three different monocyclic surrogates for amides that reinforce β-strand conformations.[51] **b)** GSK3β inhibitor incorporating a bicyclic surrogate for a Thr-Thr dipeptide motif.[53]

3.2 Heterocyclic peptidomimetic amide isosteres

Smaller heterocyclic constraints for forming conformationally-locked amide isosteres offer an alternative strategy for β-strand stabilisation.[45] In 2014, Del Valle and co-workers reported a versatile methodology for peptidomimetic backbone rigidification using tetrahydropyridazine-3,6-dione scaffold (**tpd**) as a backbone amide surrogate (Figure 4a).[51] The tpd heterocycle was formed during solid-phase peptide synthesis by an acylation reaction between a hydrazino acid and a neighbouring aspartate side-chain. Tpd ring formation was found to constrain the ϕ and ψ angles of the peptide backbone to reinforce a strand-like conformation, while the backbone amination itself also served as a crucial hydrogen bond donor, interacting with the neighbouring amide carbonyl group to further restrict the conformation.[52]

Del Valle and co-workers also explored bicyclic motifs as surrogates for dipeptide segments within extended strand conformations, reporting a series of GSK3b mimetics as inhibitors of Akt protein.[53] The central Thr-Thr dipeptide of the native GSK3b peptide was replaced with an azobicycloalkane to promote a sawtooth peptide backbone arrangement, with a carbamate scaffold providing an oxygen atom to potentially mimic the alcohol group of a Thr sidechain (Figure 4b). The resulting compound was found to have a micromolar IC_{50} for inhibiting Akt1 kinase activity on the native crosstide substrate.

3.3 Implications of ß-strand stabilisation for irregular peptides

The examples in this section showcase a small number of the many cyclisation chemistries available for stabilising β-strand conformations. It is conceivable that similar constraints could be more generally applied to irregular structures after optimising linker length and stereochemistry.

For the macrocyclic tethers, many examples involve inhibitors of protease active sites that consist of relatively short peptide sequences, due to the small size of the catalytic pocket in comparison to the interface of a PPI target. Thus, such inhibitors are more amenable to traditional medicinal chemistry lead optimisation strategies that are routinely applied to small molecules. When addressing larger targets such as PPI interfaces which bind larger molecules, multiple short constraints or isosteric scaffold replacements may be required to fully constrain a peptide into the desired conformation.

For the cyclised *N*-amino peptide strategy and the broader field of heterocyclic surrogates for amide bonds, there is potential utility in stabilising local regions of strand-like or turn-like structures within larger irregularly structured peptides. For the Tpd scaffold, there are now numerous variants of the *N*-amino cyclic scaffold developed by Del Valle and co-workers featuring different heterocyclic ring sizes and structures (pyz and tpy, Figure 4a).[54] However, their application is expected to require screening and optimisation for any given target peptide, as detailed studies have revealed that ring formation does not necessarily lead to the desired stabilisation, with alternative intramolecular hydrogen-bonding possibilities and ring strain both influencing the final conformational outcome.[54]

4. Computational design tools

Computational tools are playing an increasingly important role in the identification and optimisation of peptide-based therapeutic candidates. As with traditional medicinal chemistry campaigns, computational approaches can rapidly narrow the pool of potential peptide candidates by screening of virtual libraries prior to experimental validation, and guide design and SAR through systematic variations such as computational alanine scanning. The major challenges associated with simulating peptide-based molecules are the increased degrees of freedom and importance of the initial conformation when compared to small molecule simulations.[55]

While a range of publicly available computational packages have been developed for predicting peptide conformation and scoring peptide-protein interactions,[56-60] they are often not ideally configured to accommodate chemical linkers found in many cyclised peptides, with disulfide and head-to-tail linkages as the exceptions.[61] Atomistic simulations are often required to sample the most probable conformations of a cyclic peptide, either in isolation or in the presence of the protein target.[62] The favoured conformations can then be compared to that of known native peptide ligands, or scored by the strength of interaction with the protein target.

Atomistic simulations feature in many of the studies discussed earlier in this review. Conformational search methods such as MacroModel's macrocyclic baseline search[63] from the Schrödinger software suite and LowModeMD[64] from the Molecular Operating Environment software platform have been used to sample low-energy conformations of peptide-based macrocycles (Sections 2.3 and 3.1).[32, 48] Molecular dynamics (MD) simulations followed by computational alanine scanning and binding free energy decomposition were used by Spring and co-workers (Section 2.2) to determine the best linker position such that important binding interactions between the peptide and protein target were not disrupted.[30, 31] Here, we highlight some other examples where computational approaches that have assisted rational design may be applicable to irregularly structured peptides.

4.1 Workflows involving molecular dynamics simulations

Voelz and co-workers developed a computational method to screen and characterise the conformations of β-hairpin peptide sequences upon cyclisation.[65] Starting from published structures of a β-hairpin peptide from a bacterial protein LapD involved in biofilm formation, cyclised designs using olefin cross-linkers were simulated using replica exchange molecular dynamics and the Folding@home distributed computing platform to identify favoured conformations, which were then analysed by constructing Markov state models. Olefin cross-linking (Figure 5a) was found to be superior in reinforcing the native conformation when compared to disulfide linked and linear peptide controls. In addition, there were many subtle factors that were found to influence peptide conformation including backbone *N*-methylation, choice of position to install the cross-linking residues, and both the length and geometry of the linker. The sensitive dependence of conformation on minor molecular changes such as methylation is consistent with the findings of experimental optimisation studies discussed in Section 2. This study further highlights the challenge of multi-parameter optimisation when rationally designing conformational constraints, and shows how computation can reduce the amount of laborious experimental optimisation, particularly in the case of secondary structures where the design rules are not well-defined.

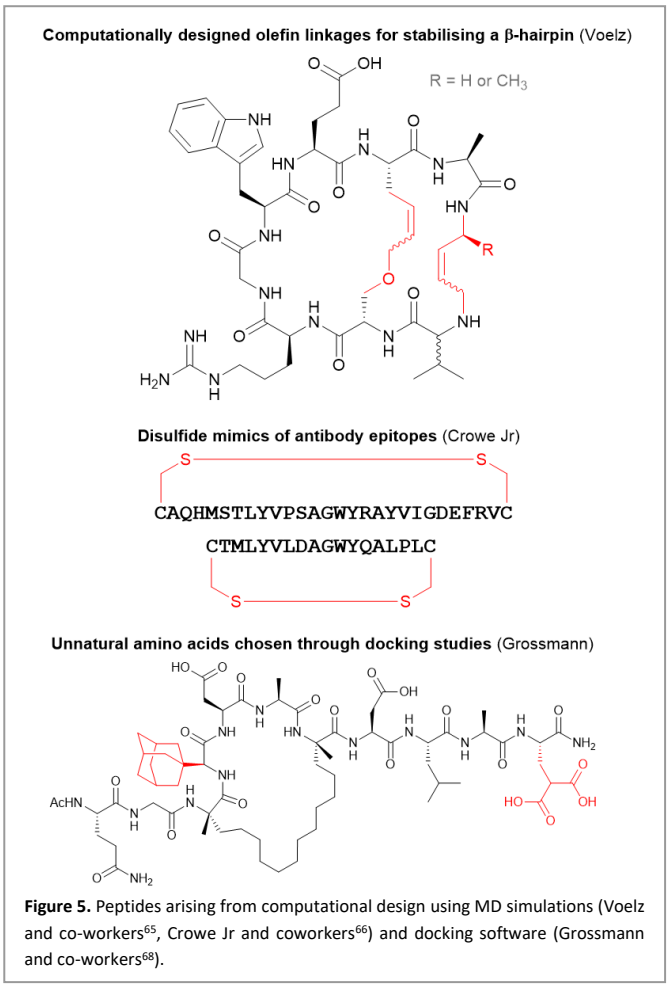

Figure 5. Peptides arising from computational design using MD simulations (Voelz and co-workers[65], Crowe Jr and coworkers[66]) and docking software (Grossmann and co-workers[68]).

Crowe Jr. *et al.* used computational methods to identify cyclic peptide sequences based off the CDRH3 loop region of the C05 influenza IgG antibody to capture the minimal epitope required.[66] Their methodology involved using the Rosetta software suite on the loop sequence with added cysteines at both termini for disulfide cyclisation, identifying the optimal fold for the peptide, then varying the sequence with the goal of stabilising the active conformation. After this Rosetta-based workflow, the best cyclised peptide candidates were subject to MD simulations to assess their conformational stability in the unbound state. Subsequent experimental validation revealed that two of the eight candidates (Figure 5b) were able to bind influenza hemagglutinin as measured by bilayer interferometry, with the disulfide constraint important for affinity. While inhibitory activity was not observed in functional hemagglutinin inhibition assays, this work nevertheless demonstrates another possible approach where computation can assist structure-based design of cyclic peptides which are truncated from irregular protein loops.

An unconventional computational workflow for identifying cyclic peptide scaffolds was reported by Santini and Zacharias,[67] matching the backbone scaffold published structures of cyclic peptides to similar backbone motifs found at protein-protein interfaces. Their cPEPmatch approach then involved grafting the corresponding residues found in the PPI onto the matching cyclic peptide backbone. MD simulations and binding free energy calculations were then performed on the grafted cyclic peptides to predict their stability, optimal structure, and target binding affinity. An advantage of cPEPmatch is that it inherently screens for peptides that have already been cyclised with a variety of non-native chemistries, without requiring custom modification of the protocol to accommodate non-proteinogenic linkers.

4.2 Docking studies for exploring SAR

In an extension of work discussed in Section 2.1, Grossmann and co-workers screened almost 1500 single residue mutants of a previously designed macrocyclic inhibitor of the ExoS/14-3-3 interaction, covering a large range of non-proteinogenic amino acids. Each mutant was docked using AutoDock Vina software with filters to restrict the search to conformations similar to the known binding pose, thus reducing the computational burden.[68] Two different scoring functions, Astex Statistical Potential and ChemScore, were used to produce two different ranked lists for each varied residue. Of the twelve mutant candidates, two had increased binding affinity, and combining the changes into the corresponding double mutant (Figure 5c) led to an approximately three-fold increase in affinity compared to the starting macrocyclic inhibitor. This study shows how the process of determining SAR and optimising binding, which is especially laborious for irregularly structured peptides, can be accelerated by computation.

4.3 Computational tools for assessing cell permeability

Computational techniques in drug discovery can also assist in predicting pharmacokinetic properties. For intracellular targets, factors such as cell permeability can be as crucial as binding affinity. Lokey and co-workers have studied cyclic peptide natural products, which typically adopt uncommon secondary structures and are often able to permeate through cell membranes passively.[69] Simulations were carried out to study the effects of *N*-methylation of backbone amides on head-to-tail cyclised peptides. Partial N-methylation was found to allow cyclic peptides to form intramolecular H-bonds that favoured membrane permeability, often leading to greater cell penetration than their corresponding *N*-permethylated analogues.[70] The same authors also modelled permeability using a PAMPA-like computational model and correlated the results with experimentally determined solution-phase structures in the literature to identify common features. Backbone stereochemistry, amide *N*-methylation pattern, and incorporation of non-proteinogenic moieties all greatly affected conformation and permeability of the cyclic peptides.

4.4 The impact of computation on peptide design

Computational methods have had a significant impact on cyclic peptide inhibitor development, providing information that would be difficult or time consuming to obtain experimentally, such as calculating binding contribution of individual residues or linkers, predicting solution state peptide conformations, and identifying molecular factors that affect cellular permeability across large data sets. Ultimately however, computational findings still require experimental validation, and developing programs *ex novo* can require just as much specialised knowledge and training as creating new experimental workflows.

There are promising new technologies that may change the way that computation is used for peptide design. *De novo* protein design has become increasingly feasible for researchers with sufficient computational power and expertise.[71] As custom secondary structures for long peptides can be designed and constrained using disulfide linkages, similar methods may also be applicable to shorter peptides and irregular conformations. In the future, it is possible that other new technological trends such as machine learning[72] and distributed crowd-sourcing methods[73] will also boost the computational power available to researchers working on peptide design.

5. Lessons from structures of peptide screening hits

While the previous sections of this review outline rational and computational approaches to design appropriate cyclic constraints for peptides, there is a wide variety of unusual structural motifs discovered through library screening and selection approaches that have not been exploited in structure-based design studies.[74] The sheer size and diversity of these libraries can give rise to non-intuitive solutions to conformational stabilisation. By examining the how irregular

structures are stabilised in *de novo* cyclic peptide library screening hits, this information may serve as inspiration for novel design strategies.

There has been a growing interest in cyclic peptide screening and selection platforms over the last two decades.[75] Several strategies exist for selecting potent inhibitors from genetically encoded cyclic peptide libraries, including the use of biological

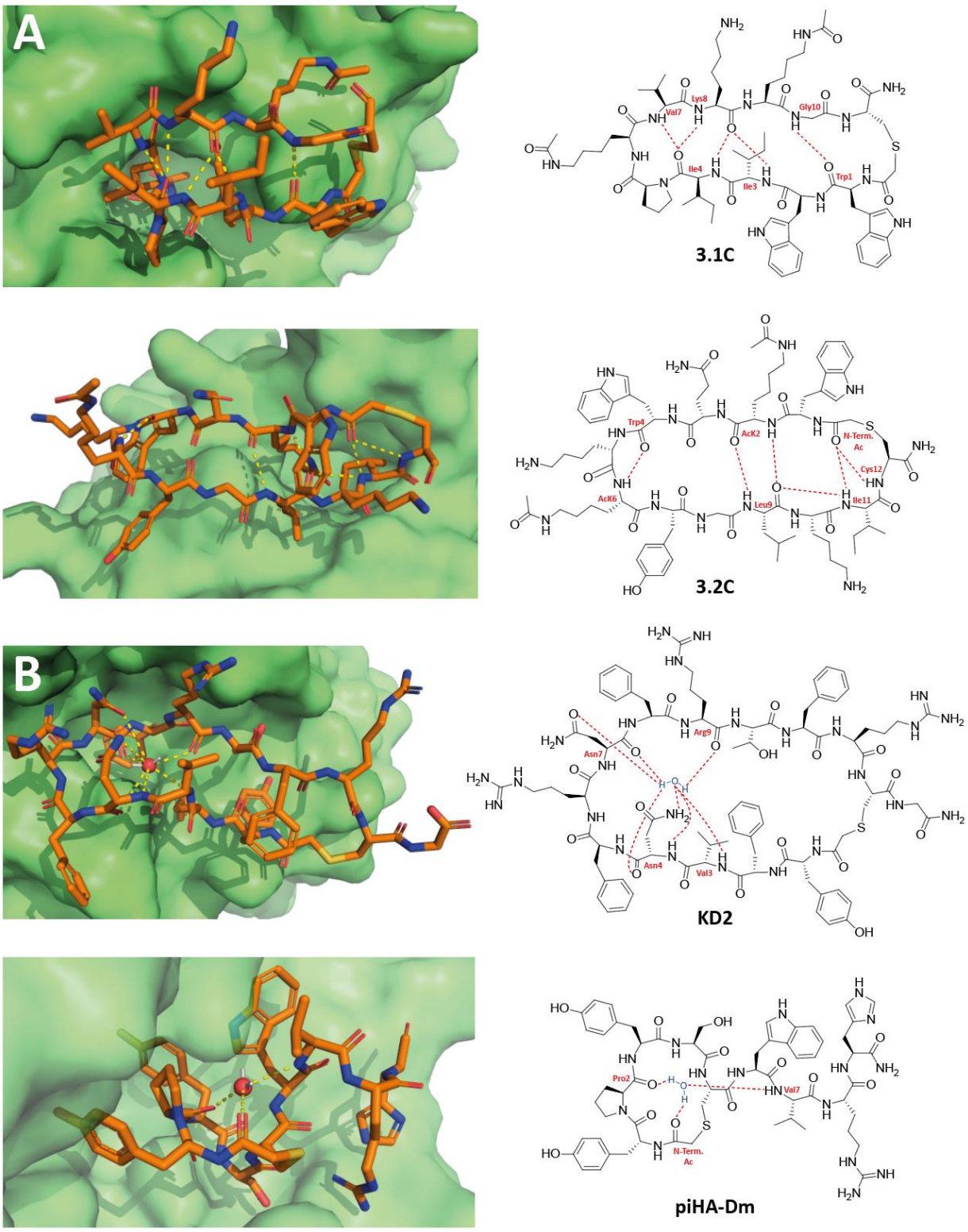

Figure 6. Bound structures and corresponding 2D representations of cyclic peptides selected through mRNA display. **a)** Structures of **3.1C** bound to BRD3-BD1 (PDB 6U4A) and **3.2C** bound to BRD3-BD2 (PDB 6ULP) derived through RaPID screens.[84] The key intramolecular hydrogen bonds are shown on the 2D representations in red. **b)** Structures of **KD2** bound to K-Ras (PDB 6WGN)[85] and **piHA-Dm** bound to α-amylase (PDB 5KEZ)[86] derived through RaPID screens mediated by a water molecule (shown as a red sphere). The intermolecular hydrogen bonds to water are show in red, while for clarity, the intramolecular hydrogen bonds are not drawn in this instance.

cyclisation methods such as intein splicing in SICLOPPS,[76, 77] and the introduction of exogenous chemoselective cross-linkers or disulfide bonds to cyclise libraries produced by phage display[78-80] or mRNA display,[81, 82] with the best candidates identified by DNA sequencing after multiple rounds of selection and sequence amplification. There are also purely synthetic techniques for generating and screening peptide libraries, such as one-bead-one-compound approaches which can be deconvoluted by mass spectrometry.[83]

In a recent survey of crystal structures for target-bound cyclic peptides derived from genetically encoded libraries up until 2019, Kawamura and co-workers reported that 7 out of 19 cyclic peptides contained no helical or sheet structure, with a further 3 containing less than 30% of either helical or sheet structure.[74] The success of these peptide libraries demonstrates that binding pockets that accept irregular peptide conformations are indeed tractable targets. In this section, we examine how peptide conformation is stabilised in a selection of these irregular structures.

5.1 Hits derived from mRNA display libraries

In 2020, Mackay and co-workers reported a series of structurally diverse cyclic peptide structures that bind the BET family of bromodomains, discovered using the flexizyme-augmented mRNA display technology known as RaPID (Random non-standard Peptide Integrated Discovery).[84] Using a head-to-sidechain cyclisation approach between an *N*-terminal *N*-chloroacetyltryptophan and a Cys thiol, multiple selections against related BET bromodomain (BD) proteins led to a range of cyclic peptide conformations as observed across 13 different target-bound structures. These conformations included well-defined β-hairpins and α-helices, but also instances of irregular structures.

Extensive internal hydrogen bonding was a common feature underpinning the stability of the diverse conformations observed. In the structure of a cyclic peptide bound to BRD3-BD1 (**3.1C**, PDB 6U4A, Figure 6a), the extended irregular hairpin structure was reinforced by five intramolecular hydrogen bonds bridging the backbones of the two strands. Similar hydrogen bonding networks were observed for another irregular hairpin cyclic peptide bound to BRD3-BD2 (**3.2C**, PDB 6ULP, Figure 6a). There was evidence of conformational preorganisation for unbound peptide 3.1C, as indicated by the well-dispersed signals of its ^{1}H NMR spectrum, suggesting the importance of cyclisation as an entropic constraint. Furthermore, there was a clear correlation between macrocyclic ring size and secondary structure. Cyclic peptides formed with a *C*-terminal Cys exclusively led to β-hairpin type structures, whereas lariat cyclisation to a Cys in the middle of the sequence in an *i,i+4* or *i,i+5* arrangement exclusively led to α-helices.

In some instances, water molecules appear to be critical for maintaining the bound secondary structure of cyclic peptides. Suga, Shokat and co-workers applied the same RaPID method to select for inhibitors against mutant K-Ras,[85] also discovering a cyclic peptide with a hairpin structure, with internal hydrogen bonding mediated in this instance by an ordered water molecule and an Asn side-chain amide (**KD2**, PDB 6WGN, Figure 6b). A similar phenomenon was also observed in a study by Jongkees *et al.* where a selected lariat inhibitor of α-amylase also featured a central water molecule between an *N*-terminal α-helical macrocycle and a *C*-terminal 3_{10}-helical tail (**piHA-Dm**, PDB 5KEZ, Figure 6b),[86] suggesting that water-mediated hydrogen bonds may be a common motif for stabilising cyclic peptide secondary structure.

The results from studies that use mRNA display suggest that the conformational stability and binding affinity of irregular extended structures containing loops and hairpins may potentially be improved by using larger ring sizes that span the entire peptide sequence, containing appropriately matched internal hydrogen bonding partners. Meanwhile, smaller lariat macrocycles appear to be more suitable for compact structures involving helices and loops. However, serendipity may well play a significant role in the attempted design of such structures, especially given the potential participation of water molecules in mediating intramolecular hydrogen bonding patterns.

5.2 Hits derived from phage display libraries

Numerous irregularly structured inhibitors have been reported against the serine protease urokinase-type plasminogen activator (uPA).[87-90] Different cyclisation approaches were explored across the various studies, including disulfide-bridged monocyclic and bicyclic peptides, as well as a 1,3,5-trisubstituted aromatic linker for crosslinking Cys residues (**UK18**, PDB 3QN7, Figure 7a).[90] While binding affinities vary in each case, all the structures involve irregular loops which lie across the protein surface, anchored by a key arginine residue buried into the substrate pocket.

A phage-derived cyclic peptide that displays a combination of well-defined secondary structure and an irregular loop was reported by Holliger and co-workers, in their study of bicyclic peptides for inhibiting proinflammatory cytokine tumour necrosis factor-alpha (TNFα).[91] The most potent molecule featured a 2,4,6-tris(bromomethyl)mesitylene core for Cys crosslinking, with a crystal structure bound to TNFα dimer revealing that a short irregular loop containing two Pro residues was responsible for interacting with one TNFα monomer, while a longer α-helical *i,i+6* loop interacted with the other monomer (**M21**, PDB 4TWT, Figure 7b). Remarkably, binding was abolished simply by changing the aromatic core to remove or add a single methylene unit at each of the 1,3,5-trimethyl positions, emphasising how the choice of cyclisation methods can have profound effects on peptide binding.

Although it is challenging to extract general implications for constraining irregular peptides from these limited examples, peptide length may play a role. It is known that phage display libraries have historically been designed with shorter peptide

sequences than in mRNA libraries, thus reducing the possibility of forming regular and well-defined secondary structure. As such, irregular peptide loops commonly appear in phage-derived cyclic peptide structures, though the chosen cyclisation chemistry and the structure of the binding pockets themselves may also be contributing factors.

5.3 Is rational design still necessary?

Despite the power of library-based techniques for identifying *de novo* cyclic peptide ligands with nanomolar affinity for binding pockets that natively accept peptides in an irregular conformation, rational design strategies still have an important role to play in the discovery and optimisation process. As in traditional medicinal chemistry pipelines, there are strengths and weaknesses of hit identification through either screening or structure-guided design. Although throughput enables wider exploration of chemical space, rational approaches are able to address specific binding pockets on a given target, while also circumventing target-specific incompatibility issues that render screens ineffective eg. RNA-binding activity that interferes with mRNA library screening. A hybrid approach can also be adopted, narrowing down a screening pool by rationally restricting variables such as peptide length, key residues, and linker positioning and chemistry, based on known properties of the target binding site. Beyond hit identification, rational exploration of SAR is still necessary to optimise binding affinity and other pharmacokinetic parameters. Emulating the common molecular features observed in screening hits may feed into the SAR exploration process, such as exploring different hydrogen bonding patterns and linker lengths to tailor the conformation towards the desired irregular structure.

6. Conclusions

Developing conformationally restricted analogues of peptides in irregular conformations is not a straightforward task. Given the prevalence of irregular peptide structures in nature however, there is strong motivation from a drug discovery perspective to develop robust workflows for their cyclisation, in search of highly potent and specific inhibitors with acceptable pharmacokinetics.[1, 92] Aside from direct application as therapeutic leads, cyclised peptide inhibitors can serve as important tool compounds in biological studies, and may assist small molecule design by defining possible binding interactions for a given target site.

The studies in this review demonstrate that the wealth of cyclisation techniques which have already been developed for regular secondary structures, such as α-helices, can be applied to the stabilisation of irregularly structured peptides. At the same time, a common theme is the need for thorough optimisation. Similar to lead optimisation processes in traditional medicinal chemistry, irregular peptides have multiple molecular parameters that can be varied, including

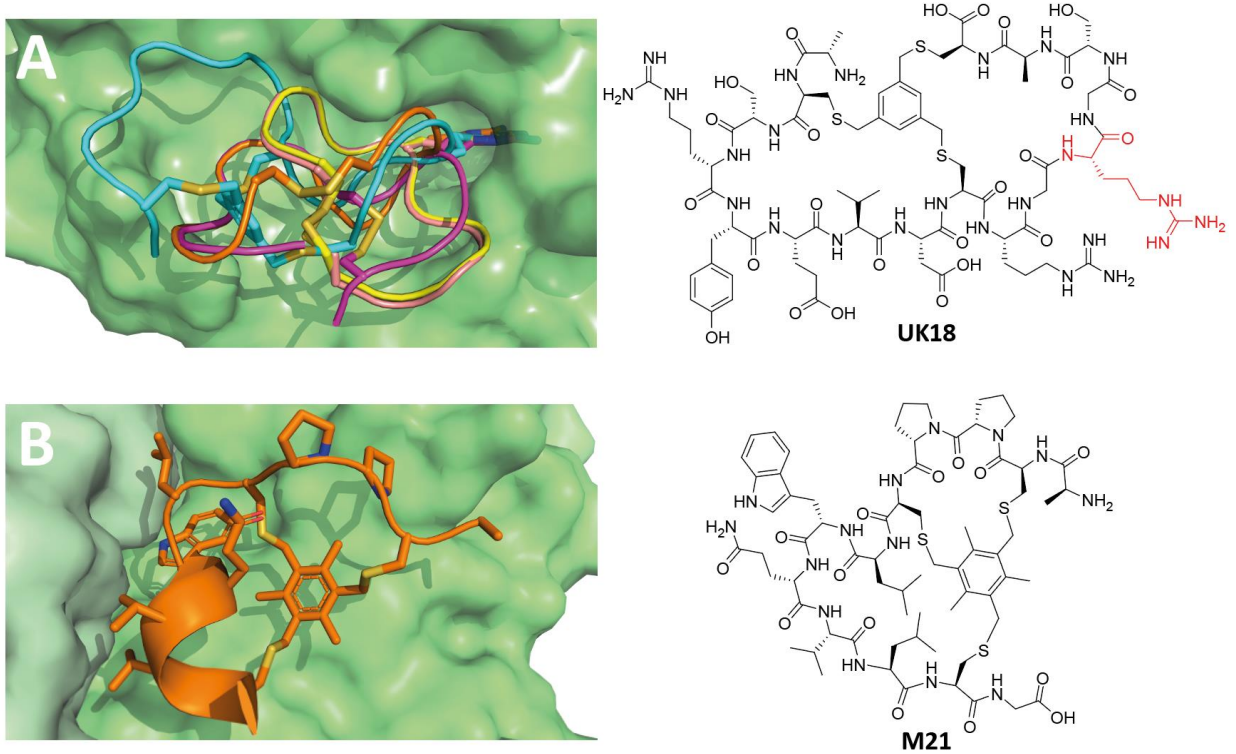

Figure 7. Bound structures and 2D representations of cyclic peptides selected through phage display. **a)** Structures of urokinase-type plasminogen activator (uPA) inhibitors derived from phage display libraries (PDB 2NWN/3QN7/4GLY/6A8G/6A8N, left)[87-90] and a 2D structure of one example **UK18** that was cyclised with a 1,3,5-trisubstituted aromatic linker (right).[90] All the structures feature irregular loops spread across the protein surface, anchored by an arginine residue in the substrate pocket (highlighted in red on the 2D structure). Figure adapted from McAllister *et al.*[74] **b)** Structures of peptide **M21** stabilised by a mesitylene core bound to proinflammatory cytokine tumour necrosis factor-alpha (TNFα) dimer (PDB 4TWT).[91] The α-helical loop interacts with one monomer while the short irregular loop with two proline residues interacts with the other monomer.

cyclisation chemistry, linker position and length, stereochemistry, backbone modification, sequence identity, and internal hydrogen bonding patterns. Furthermore, the output that is being optimised involves a balance between *in vitro* binding affinity and effective target engagement *in vivo*. As this combination of variables represents a daunting task, the challenge for chemists is to identify methods that can accelerate throughput or enable predictable design.

Structural biology has a prominent role in the predictable design of cyclic constraints for irregular conformations. Crystallographic data can provide insight into binding poses, key interacting residues and even scaffold flexibility, enriching our capacity to interpret and understand data from SAR studies. Pairing experimentally determined structures with increasingly powerful computational techniques and library screening technologies should help to reduce the turnaround time for inhibitor optimisation.

Pharmacokinetics remains an ongoing challenge for the entire field of peptide therapeutics, including irregularly structured peptides. There are increasing efforts to understand how structure affects parameters such as cell permeability when addressing intracellular targets, including systematic experimental studies, high-throughput compatible methods, and computational studies. These efforts will be crucial for the future translational potential of peptide therapeutics to be realised, beyond well-established cases of peptide hormones such as insulin.[1]

In conclusion, rational cyclisation of irregularly structured peptides for the inhibition of PPIs and other drug targets is now achievable using the array of peptide discovery techniques that have been developed. As this class of stapled peptides matures and become more well-documented, we expect to see more studies achieving successful inhibition of a greater variety of PPI targets, opening new therapeutic windows for treating currently intractable diseases.

Author Contributions

All authors contributed to the planning and writing of this manuscript.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

YHL acknowledges funding from the National Health and Medical Research Council (APP2003250), Australian Research Council (DE190100624), and Westpac Scholars Trust (WRF2020). QV acknowledges funding from a Sydney Vital top-up scholarship.

References

1. J. L. Lau and M. K. Dunn, *Bioorg Med Chem*, 2018, **26**, 2700-2707.
2. E. Valeur, S. M. Gueret, H. Adihou, R. Gopalakrishnan, M. Lemurell, H. Waldmann, T. N. Grossmann and A. T. Plowright, *Angew Chem Int Ed Engl*, 2017, **56**, 10294-10323.
3. A. A. Kaspar and J. M. Reichert, *Drug Discov Today*, 2013, **18**, 807-817.
4. L. Nevola and E. Giralt, *Chem Commun (Camb)*, 2015, **51**, 3302-3315.
5. C. Yan, F. Wu, R. L. Jernigan, D. Dobbs and V. Honavar, *Protein J*, 2008, **27**, 59-70.
6. G. L. Verdine and G. J. Hilinski, *Methods Enzymol*, 2012, **503**, 3-33.
7. S. Perot, O. Sperandio, M. A. Miteva, A. C. Camproux and B. O. Villoutreix, *Drug Discov Today*, 2010, **15**, 656-667.
8. L. K. Henchey, A. L. Jochim and P. S. Arora, *Curr Opin Chem Biol*, 2008, **12**, 692-697.
9. C. E. Chang, W. Chen and M. K. Gilson, *Proc Natl Acad Sci U S A*, 2007, **104**, 1534-1539.
10. Z. Fang, Y. Song, P. Zhan, Q. Zhang and X. Liu, *Future Med Chem*, 2014, **6**, 885-901.
11. M. K. Gilson and H. X. Zhou, *Annu Rev Biophys Biomol Struct*, 2007, **36**, 21-42.
12. J. E. Bock, J. Gavenonis and J. A. Kritzer, *ACS Chem Biol*, 2013, **8**, 488-499.
13. Y. H. Lau, P. de Andrade, Y. Wu and D. R. Spring, *Chem Soc Rev*, 2015, **44**, 91-102.
14. T. Okamoto, D. Segal, K. Zobel, A. Fedorova, H. Yang, W. J. Fairbrother, D. C. Huang, B. J. Smith, K. Deshayes and P. E. Czabotar, *ACS Chem Biol*, 2014, **9**, 838-839.
15. G. H. Bird, A. Irimia, G. Ofek, P. D. Kwong, I. A. Wilson and L. D. Walensky, *Nat Struct Mol Biol*, 2014, **21**, 1058-1067.
16. V. Azzarito, K. Long, N. S. Murphy and A. J. Wilson, *Nat Chem*, 2013, **5**, 161-173.
17. A. L. Jochim and P. S. Arora, *ACS Chem Biol*, 2010, **5**, 919-923.
18. T. A. Cardote and A. Ciulli, *ChemMedChem*, 2016, **11**, 787-794.
19. M. T. J. Bluntzer, J. O'Connell, T. S. Baker, J. Michel and A. N. Hulme, *Peptide Science*, 2020, **113**.
20. J. A. Wells and C. L. McClendon, *Nature*, 2007, **450**, 1001-1009.
21. M. Guharoy and P. Chakrabarti, *Bioinformatics*, 2007, **23**, 1909-1918.
22. J. Gavenonis, B. A. Sheneman, T. R. Siegert, M. R. Eshelman and J. A. Kritzer, *Nat Chem Biol*, 2014, **10**, 716-722.
23. X. Li, S. Chen, W. D. Zhang and H. G. Hu, *Chem Rev*, 2020, **120**, 10079-10144.
24. A. K. Malde, T. A. Hill, A. Iyer and D. P. Fairlie, *Chem Rev*, 2019, **119**, 9861-9914.
25. A. Glas, D. Bier, G. Hahne, C. Rademacher, C. Ottmann and T. N. Grossmann, *Angew Chem Int Ed Engl*, 2014, **53**, 2489-2493.
26. L. D. Walensky and G. H. Bird, *J Med Chem*, 2014, **57**, 6275-6288.
27. A. Glas, E. C. Wamhoff, D. M. Kruger, C. Rademacher and T. N. Grossmann, *Chemistry*, 2017, **23**, 16157-16161.

28. P. M. Cromm, K. Wallraven, A. Glas, D. Bier, A. Furstner, C. Ottmann and T. N. Grossmann, *Chembiochem*, 2016, **17**, 1915-1919.
29. Y. H. Lau, Y. Wu, P. de Andrade, W. R. Galloway and D. R. Spring, *Nat Protoc*, 2015, **10**, 585-594.
30. W. Xu, Y. H. Lau, G. Fischer, Y. S. Tan, A. Chattopadhyay, M. de la Roche, M. Hyvonen, C. Verma, D. R. Spring and L. S. Itzhaki, *J Am Chem Soc*, 2017, **139**, 2245-2256.
31. M. M. Wiedmann, Y. S. Tan, Y. Wu, S. Aibara, W. Xu, H. F. Sore, C. S. Verma, L. Itzhaki, M. Stewart, J. D. Brenton and D. R. Spring, *Angew Chem Int Ed Engl*, 2017, **56**, 524-529.
32. J. Yang, V. O. Talibov, S. Peintner, C. Rhee, V. Poongavanam, M. Geitmann, M. R. Sebastiano, B. Simon, J. Hennig, D. Dobritzsch, U. H. Danielson and J. Kihlberg, *ACS Omega*, 2020, **5**, 3979-3995.
33. P. G. Dougherty, A. Sahni and D. Pei, *Chem Rev*, 2019, **119**, 10241-10287.
34. Y. Dou, P. F. Baisnee, G. Pollastri, Y. Pecout, J. Nowick and P. Baldi, *Bioinformatics*, 2004, **20**, 2767-2777.
35. Y. M. Wong, H. Masunaga, J. A. Chuah, K. Sudesh and K. Numata, *Biomacromolecules*, 2016, **17**, 3375-3385.
36. J. R. Del Valle, in *Peptidomimetics II*, 2015, DOI: 10.1007/7081_2015_163, ch. Chapter 163, pp. 25-49.
37. D. Leung, G. Abbenante and D. P. Fairlie, *J Med Chem*, 2000, **43**, 305-341.
38. R. E. Babine and S. L. Bender, *Chem Rev*, 1997, **97**, 1359-1472.
39. M. L. West and D. P. Fairlie, *Trends in Pharmacological Sciences*, 1995, **16**, 67-75.
40. A. D. Pehere and A. D. Abell, in *Proteases in Health and Disease*, 2013, DOI: 10.1007/978-1-4614-9233-7_11, ch. Chapter 11, pp. 181-192.
41. T. K. Sawyer, R. S. Bohacek, D. C. Dalgarno, C. J. Eyermann, N. Kawahata, C. A. Metcalf, 3rd, W. C. Shakespeare, R. Sundaramoorthi, Y. Wang and M. G. Yang, *Mini Rev Med Chem*, 2002, **2**, 475-488.
42. W. A. Loughlin, J. D. A. Tyndall, M. P. Glenn and D. P. Fairlie, *Chemical Reviews*, 2004, **104**, 6085-6118.
43. J. D. Tyndall, T. Nall and D. P. Fairlie, *Chem Rev*, 2005, **105**, 973-999.
44. T. A. Hill, N. E. Shepherd, F. Diness and D. P. Fairlie, *Angew Chem Int Ed Engl*, 2014, **53**, 13020-13041.
45. S. Hanessian, G. McNaughton-Smith, H.-G. Lombart and W. D. Lubell, *Tetrahedron*, 1997, **53**, 12789-12854.
46. G. Abbenante, D. R. March, D. A. Bergman, P. A. Hunt, B. Garnham, R. J. Dancer, J. L. Martin and D. P. Fairlie, *Journal of the American Chemical Society*, 1995, **117**, 10220-10226.
47. A. E. Weber, T. A. Halgren, J. J. Doyle, R. J. Lynch, P. K. Siegl, W. H. Parsons, W. J. Greenlee and A. A. Patchett, *J Med Chem*, 1991, **34**, 2692-2701.
48. A. D. Abell, M. A. Jones, J. M. Coxon, J. D. Morton, S. G. Aitken, S. B. McNabb, H. Y. Lee, J. M. Mehrtens, N. A. Alexander, B. G. Stuart, A. T. Neffe and R. Bickerstaffe, *Angew Chem Int Ed Engl*, 2009, **48**, 1455-1458.
49. A. D. Pehere and A. D. Abell, *Org Lett*, 2012, **14**, 1330-1333.
50. H. Chen, W. Jiao, M. A. Jones, J. M. Coxon, J. D. Morton, R. Bickerstaffe, A. D. Pehere, O. Zvarec and A. D. Abell, *Chem Biodivers*, 2012, **9**, 2473-2484.
51. C. W. Kang, S. Ranatunga, M. P. Sarnowski and J. R. Del Valle, *Org Lett*, 2014, **16**, 5434-5437.
52. M. P. Sarnowski, C. W. Kang, Y. M. Elbatrawi, L. Wojtas and J. R. Del Valle, *Angew Chem Int Ed Engl*, 2017, **56**, 2083-2086.
53. S. Ranatunga and J. R. Del Valle, *Bioorg Med Chem Lett*, 2011, **21**, 7166-7169.
54. M. P. Sarnowski, K. P. Pedretty, N. Giddings, H. L. Woodcock and J. R. Del Valle, *Bioorg Med Chem*, 2018, **26**, 1162-1166.
55. S. E. Allen, N. V. Dokholyan and A. A. Bowers, *ACS Chem Biol*, 2016, **11**, 10-24.
56. M. Trellet, A. S. Melquiond and A. M. Bonvin, *PLoS One*, 2013, **8**, e58769.
57. B. Raveh, N. London, L. Zimmerman and O. Schueler-Furman, *PLoS One*, 2011, **6**, e18934.
58. B. Raveh, N. London and O. Schueler-Furman, *Proteins*, 2010, **78**, 2029-2040.
59. H. Lee, L. Heo, M. S. Lee and C. Seok, *Nucleic Acids Res*, 2015, **43**, W431-435.
60. I. Antes, *Proteins*, 2010, **78**, 1084-1104.
61. Y. Zhang and M. F. Sanner, *J Chem Theory Comput*, 2019, **15**, 5161-5168.
62. Y. S. Tan, D. P. Lane and C. S. Verma, *Drug Discov Today*, 2016, **21**, 1642-1653.
63. K. S. Watts, P. Dalal, A. J. Tebben, D. L. Cheney and J. C. Shelley, *J Chem Inf Model*, 2014, **54**, 2680-2696.
64. P. Labute, *J Chem Inf Model*, 2010, **50**, 792-800.
65. A. M. Razavi, W. M. Wuest and V. A. Voelz, *J Chem Inf Model*, 2014, **54**, 1425-1432.
66. A. M. Sevy, I. M. Gilchuk, B. P. Brown, N. G. Bozhanova, R. Nargi, M. Jensen, J. Meiler and J. E. Crowe, Jr., *Structure*, 2020, **28**, 1114-1123 e1114.
67. B. L. Santini and M. Zacharias, *Front Chem*, 2020, **8**, 573259.
68. D. M. Kruger, A. Glas, D. Bier, N. Pospiech, K. Wallraven, L. Dietrich, C. Ottmann, O. Koch, S. Hennig and T. N. Grossmann, *J Med Chem*, 2017, **60**, 8982-8988.
69. T. R. White, C. M. Renzelman, A. C. Rand, T. Rezai, C. M. McEwen, V. M. Gelev, R. A. Turner, R. G. Linington, S. S. Leung, A. S. Kalgutkar, J. N. Bauman, Y. Zhang, S. Liras, D. A. Price, A. M. Mathiowetz, M. P. Jacobson and R. S. Lokey, *Nat Chem Biol*, 2011, **7**, 810-817.
70. C. L. Ahlbach, K. W. Lexa, A. T. Bockus, V. Chen, P. Crews, M. P. Jacobson and R. S. Lokey, *Future Med Chem*, 2015, **7**, 2121-2130.
71. G. Bhardwaj, V. K. Mulligan, C. D. Bahl, J. M. Gilmore, P. J. Harvey, O. Cheneval, G. W. Buchko, S. V. Pulavarti, Q. Kaas, A. Eletsky, P. S. Huang, W. A. Johnsen, P. J. Greisen, G. J. Rocklin, Y. Song, T. W. Linsky, A. Watkins, S. A. Rettie, X. Xu, L. P. Carter, R. Bonneau, J. M. Olson, E. Coutsias, C. E. Correnti, T. Szyperski, D. J. Craik and D. Baker, *Nature*, 2016, **538**, 329-335.
72. J. M. Wolfe, C. M. Fadzen, Z. N. Choo, R. L. Holden, M. Yao, G. J. Hanson and B. L. Pentelute, *ACS Cent Sci*, 2018, **4**, 512-520.
73. M. Shirts and V. S. Pande, *Science*, 2000, **290**, 1903-1904.
74. T. E. McAllister, O. D. Coleman, G. Roper and A. Kawamura, *Peptide Science*, 2020, **113**.
75. C. Sohrabi, A. Foster and A. Tavassoli, *Nature Reviews Chemistry*, 2020, **4**, 90-101.
76. K. R. Lennard and A. Tavassoli, *Chemistry*, 2014, **20**, 10608-10614.
77. C. P. Scott, E. Abel-Santos, M. Wall, D. C. Wahnon and S. J. Benkovic, *Proc Natl Acad Sci U S A*, 1999, **96**, 13638-13643.

78. K. Deyle, X. D. Kong and C. Heinis, *Acc Chem Res*, 2017, **50**, 1866-1874.
79. C. Heinis, T. Rutherford, S. Freund and G. Winter, *Nat Chem Biol*, 2009, **5**, 502-507.
80. S. Ng, K. F. Tjhung, B. M. Paschal, C. J. Noren and R. Derda, *Methods Mol Biol*, 2015, **1248**, 155-172.
81. T. Passioura and H. Suga, *Chem Commun (Camb)*, 2017, **53**, 1931-1940.
82. K. Josephson, A. Ricardo and J. W. Szostak, *Drug Discov Today*, 2014, **19**, 388-399.
83. Z. Qian, P. Upadhyaya and D. Pei, *Methods Mol Biol*, 2015, **1248**, 39-53.
84. K. Patel, L. J. Walport, J. L. Walshe, P. D. Solomon, J. K. K. Low, D. H. Tran, K. S. Mouradian, A. P. G. Silva, L. Wilkinson-White, A. Norman, C. Franck, J. M. Matthews, J. M. Guss, R. J. Payne, T. Passioura, H. Suga and J. P. Mackay, *Proc Natl Acad Sci U S A*, 2020, **117**, 26728-26738.
85. Z. Zhang, R. Gao, Q. Hu, H. Peacock, D. M. Peacock, S. Dai, K. M. Shokat and H. Suga, *ACS Cent Sci*, 2020, **6**, 1753-1761.
86. S. A. K. Jongkees, S. Caner, C. Tysoe, G. D. Brayer, S. G. Withers and H. Suga, *Cell Chem Biol*, 2017, **24**, 381-390.
87. G. Zhao, C. Yuan, T. Wind, Z. Huang, P. A. Andreasen and M. Huang, *J Struct Biol*, 2007, **160**, 1-10.
88. S. Chen, I. Rentero Rebollo, S. A. Buth, J. Morales-Sanfrutos, J. Touati, P. G. Leiman and C. Heinis, *J Am Chem Soc*, 2013, **135**, 6562-6569.
89. D. Wang, Y. Yang, L. Jiang, Y. Wang, J. Li, P. A. Andreasen, Z. Chen, M. Huang and P. Xu, *J Med Chem*, 2019, **62**, 2172-2183.
90. A. Angelini, L. Cendron, S. Chen, J. Touati, G. Winter, G. Zanotti and C. Heinis, *ACS Chem Biol*, 2012, **7**, 817-821.
91. S. Luzi, Y. Kondo, E. Bernard, L. K. Stadler, M. Vaysburd, G. Winter and P. Holliger, *Protein Eng Des Sel*, 2015, **28**, 45-52.
92. C. Morrison, *Nat Rev Drug Discov*, 2018, **17**, 531-533.