2015

Peptide Drug Discovery: Innovative Technologies and Transformational Medicines

David J. Diller

Mark Jarosinski

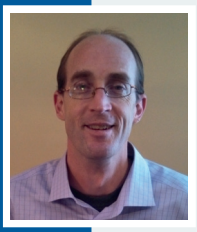
Tomi K. Sawyer

Joseph Audie
Sacred Heart University, audiej@sacredheart.edu

Recommended Citation

Diller, D.J., Jarosinski, M., Sawyer, T.K., & Audie, J. (2015). Peptide drug discovery: Innovative technologies and transformational medicines. In Kruger, H.G. & Albericio, F. (Eds). *Advances in the discovery and development of peptide therapeutics*(pp. 8-27). London: Future Medicine. doi: 10.4155/fseb2013.14.105

About the Authors



David J Diller

David Diller is the Director of Computational Chemistry at CMDBioscience. He has a PhD in mathematics from Northwestern University. He has worked in drug discovery groups at large pharma, medium-sized biotech and startup companies where he has applied many aspects of molecular modeling and informatics to problems in drug discovery including structure-based design, virtual screening and ADME/Tox modeling.

Mark Jarosinski

Mark Jarosinski is Senior Director of Peptide-Medicinal Chemistry at CMDBioscience. He studied peptide sciences under Victor Hruby. Mark worked for Novartis, Amgen, Midwest Bio-Tech prior to joining CMDBioscience. His scientific focus is lead identification and optimization of peptide therapeutics. His achievements span from molecular target recognition, rational design, synthesis and delivery of peptide therapeutics to the clinic.

Tomi K Sawyer

Tomi Sawyer is Emeritus Founding Chair of the SAB at CMDBioscience. He recently joined Merck & Company as Distinguished Scientist and Head, Peptide Drug Discovery & Innovative Technologies. His previous positions include Chief Scientific Officer at Aileron Therapeutics, Senior Vice-President of Drug Discovery at ARIAD Pharmaceuticals and Senior Director at Pfizer. Tomi is credited with >500 publications, patents and presentations.

Joseph Audie

Joseph Audie is co-founder and CEO of CMDBioscience. He has a PhD in biophysics from SUNY Stony Brook and is an Associate Professor of Chemistry at Sacred Heart University. His research interests include elucidating the physics of protein-ligand recognition and exploiting that knowledge to design peptide drugs.

Chapter 1

Peptide drug discovery: innovative technologies and transformational medicines

David J Diller, Mark Jarosinski, Tomi K Sawyer & Joseph Audie

doi:10.4155/fseb2013.14.105

Interest in peptide drug discovery is surging. In the past several years, numerous pharmaceutical and biotech companies have committed considerable resources to peptide-based drug discovery. In part, this is being fueled by an increasing recognition that peptide drugs combine many of the virtues of small molecules and proteins, while minimizing several of their drawbacks, and that peptides can potentially expand the druggable space to include intracellular, extracellular and membrane-associated protein–protein interactions. Moreover, powerful new *in vitro* and *in silico* technologies and breakthroughs in our understanding of natural peptides have emerged that provide peptide chemists with the tools and insights they need to solve the various pharmacokinetic problems that often plague peptide drug discovery efforts. From stapled peptides, to highly versatile macrocyclic peptides and disulfide-rich peptides, to other peptides with various nonstandard chemistries, peptides are poised to fulfill their promise of providing a drug class that straddles the chemical space between small molecules and proteins, ultimately resulting in transformational medicines and improved clinical outcomes.

Owing to the marginal return on investment in recent years, there is a pressing need to re-examine all aspects of drug discovery. One area that has received renewed focus and significant investment is that of innovative peptide technologies that may enhance transformational drug discovery. This renewed interest in therapeutic peptides may be attributed to two considerations. First, US FDA approval rates for peptides (23–26%) [1] are significantly greater than that of small-molecules (~13%) [2]. Although peptides currently account for a relatively small fraction of approved drugs, the approximately twofold increase in clinical success probability is an important metric, especially given that clinical trial attrition rates account for high overall R&D costs. The second consideration is the large physicochemical gap between small molecules (<500 Da) and biologics (>5000 Da). This gap has left a therapeutic gap that peptides and macrocycles have the potential to fill. Indeed, there is mounting evidence that peptides can be used to tackle challenging targets, such as protein–protein interactions (PPIs), where small-molecules have had limited success, and in modulating key intracellular PPI targets peptides may outperform antibodies that have limited cell-penetration properties.

As the excitement around peptide drug discovery builds, it is imperative to understand why in the 1980s and 1990s peptides first lost traction as development candidates. Two key reasons for this loss in momentum included their challenging pharmacokinetic (PK) properties and manufacturing costs. PK challenges for peptides can include poor intestinal absorption, low proteolytic stability and rapid renal clearance. However, it should be noted that for some disease indications apparent PK drawbacks can be beneficial. In addition, nature has solved numerous physiological and pharmacological peptide PK problems. For example, some natural product peptides (dietary-derived bioactive peptides) have been determined to be orally bioavailable and many peptide toxins have remarkable stability profiles. Such examples provide compelling evidence that peptides may be optimized into effective and broadly active pharmaceutical agents.

The therapeutic gap & the key opportunity for peptides in drug discovery: the example of class-B G-protein-coupled receptors & protein–protein interactions

Class-B G-protein-coupled receptors (GPCRs) are extracellular targets of natural peptides and are important drug targets. The peptide ligands for GPCRs are approximately 30 amino acids (AAs) long, with diffuse pharmacophoric domains (e.g., calcitonin, glucagon, parathyroid hormone). Class-B GPCRs are transmembrane proteins with an amino-terminal extracellular domain and an intracellular domain. The characteristic amino-terminal extracellular domain contains three conserved disulfide bonds and a conserved cleft for

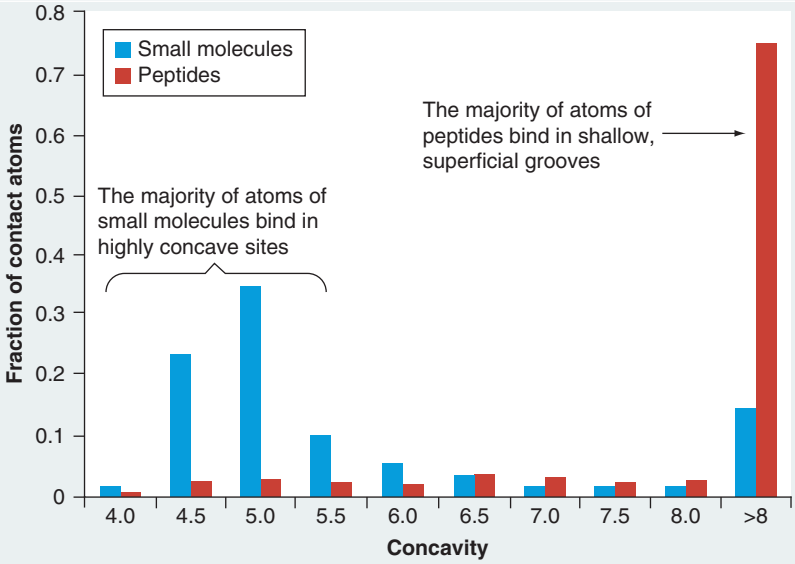
binding to the helical carboxyl-terminal region of many natural peptides. The extracellular domain orients the amino terminus of the peptides towards the transmembrane helical bundle, which is critical for receptor activation. Typically, to act as an antagonist a molecule needs to bind the extracellular domain. This is generally possible to do with a protein or peptide. However, given that most protein interactions involve large, relatively flat and diffuse interactions surfaces, small-molecule antagonist approaches would appear to be quite limited. To act as an agonist, a molecule must interact with the extracellular domain and the transmembrane helical bundle – a multidomain interaction that is difficult to mimic with small-molecules or proteins. Thus, the natural peptide ligands of class-B GCPRs have proven to be useful starting points for agonist lead optimization [3], as exemplified by the glucagon-like peptide-1 receptor (GLP1R) agonist peptides exenatide and liraglutide, which are used to treat Type II diabetes. Finally, more exceptional therapeutic strategies, such as targeting intracellular GPCR structures and interactions, have shown unique potential for peptide drug discovery versus small-molecule or protein strategies.

PPIs are challenging drug targets owing to the presence of large, diffuse and essentially flat binding surfaces that are not easily interrogated by small molecules. Several strategies have been used to design small molecules for protein interaction interfaces, with different levels of success. Discontinuous binding epitopes, often encountered in PPIs, have generally proven to be intractable with small molecules. By contrast, such binding epitopes can be fruitfully explored with peptides. And while proteins can modulate PPIs, they are limited to extracellular PPIs. Peptides, on the other hand, can be developed to modulate intracellular and extracellular PPIs. As demonstrated below, peptides can also bind to shallow and extended protein pockets. Thus, with peptides, PPI targets become more generally druggable; moreover, allosteric modulation of PPIs, where the interaction is further compromised by disease-related mutations, are likely more druggable with respect to peptide strategies.

Increasing the viable drug space: the characteristics of peptide binding sites

To quantify the extent to which peptides broaden the viable druggable space, it is worth highlighting a recent analysis of peptide–protein co-crystal structures [4], which included 103 co-crystal structures of peptides (5–15 AAs) bound to nonredundant proteins. Importantly, this analysis uncovered several aspects of peptide–protein interactions that differed from those generally known for PPIs. First, peptide–protein interactions were observed to pack more effectively than PPIs. This is not surprising given that peptide–protein interfaces have significantly less buried surface area than PPIs. The

Figure 1.1 A comparison of peptide and small-molecule binding sites.

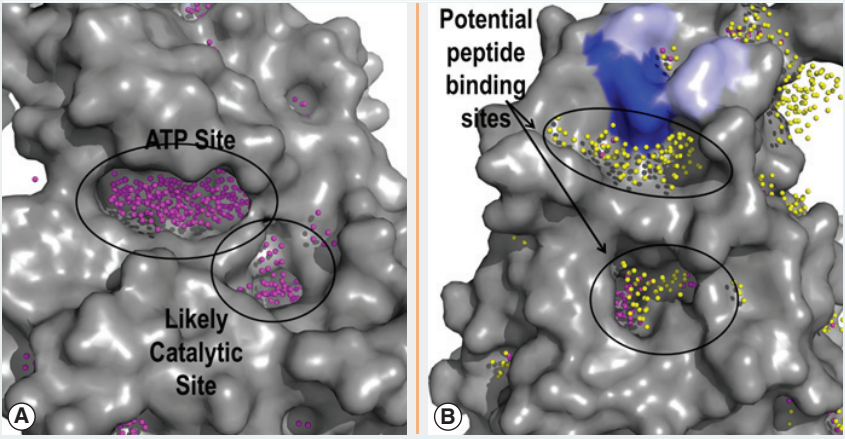


The blue histogram represents the contact atoms of small molecules. The red histogram represents the contact atoms of peptides. A heavy atom of the ligand was considered a contact atom if it was within 4.5 Å of a heavy atom of the protein. The metric used to define the depth of the pocket is concavity. Low values of concavity indicate deep pockets, whereas high values for concavity indicate shallow superficial pockets.

second, more surprising, difference was that the peptide binding induced little change in the target protein's binding site. This can be attributed to the fact that peptides are disordered in solution and lose a significant amount of entropy upon binding. This entropy loss is partly compensated for by pre-organized protein-binding partners. Hence, there are important differences between PPIs and protein–peptide interactions that may be exploited to increase the viable drug space.

Peptide–protein binding sites are also very different from small-molecule protein binding sites. To support this point, we performed a survey of co-crystal structures found in the protein data bank (PDB): 150 protein:small-molecule complexes and 150 protein:peptide complexes. We next calculated the binding site concavity at each atom of the ligand in contact with the protein. As **Figure 1.1** shows, nearly all the contact atoms of small molecules are bound in deep pockets, whereas nearly all the contact atoms of peptides bind in superficial grooves. As a concrete example, we show in **Figure 1.2**

Figure 1.2 Potential binding sites on the pseudo-kinase domain of Jak2.



The purple spheres are points of low curvature – consistent with small-molecule binding sites. The yellow spheres are points of high curvature – consistent with peptide binding sites. **(A)** The ATP binding site. **(B)** Other potential peptide binding sites. This is the face opposite the ATP binding site. The portion of the surface colored in dark blue highlights a mutation, V617F, common to many hematological disorders.

the pseudo-kinase domain of Jak2. The ATP-binding site is clearly consistent with small-molecule binding. There are, however, several additional sites on this domain that are consistent with peptide binding. These sites could potentially be sites for biologically relevant PPIs or sites for allosteric modulation by peptides. Particularly for targets such as kinases, where the primary functional site is highly conserved over a large family of proteins, allosteric sites offer the potential to improve upon selectivity and ultimately safety. Hence, peptides offer excellent molecular probes for exploring novel as well as conventional protein surfaces and binding sites.

Peptide binding site identification

The above observations regarding peptide–protein binding sites open exciting opportunities for peptide drug design. Given that more than 25% of the human genome either has a publically available crystal structure or could be readily modeled, there is significant interest in finding and annotating potential peptide binding sites.

Numerous computational methods have been proposed to identify potential peptide binding sites and residue hot spots on proteins; here, a few methods are highlighted. At one extreme are purely sequence-based approaches. For example, Yip and coworkers proposed a method that uses covariance

within multiple sequence alignments (MSAs) to identify potential functional sites [5]. The key to this strategy is to identify correlated residue changes within a MSA or pair of MSAs that suggest structural or functional relationships and potential interaction sites. PepSite offers an approximate structure-based approach for locating potential peptide binding sites [6]. Specifically, this approach is based on two key principles: the recognition that peptide binding sites are often enriched in localized hot spots; and the calculation of spatial position-specific AA scoring matrices that are derived from a training set of protein–peptide interactions that encode information regarding preferred AA-binding preferences. Therefore, this approach allows for the rapid identification of potential peptide binding sites by scanning the protein surfaces with each AA. Finally, in a fully structure-based approach, Lavi and coworkers adapted a method originally used for small-molecule binding site identification [7]. The premise to this approach is that the entire protein surface is scanned with 16 molecular fragments. The critical fragment binding sites are then identified by clustering and contact scoring, and after removing inaccessible sites, the candidate peptide binding sites are correlated with those for which multiple fragment sites exist.

Novel peptide lead identification: peptide properties & *in vitro* phage display screening

Having established the structural basis for how peptides can expand the viable drug space, we now turn to the discovery of novel peptide leads through the use of *in vitro* screening approaches. *In vitro* display and library screening technologies can be used to rapidly generate peptide diversity. Traditionally, the technique of phage display has been used to search through vast numbers of peptides to discover bioactive peptides. Indeed, approved drugs have been discovered using phage display e.g., Kalbitor (Dyax, Burlington MA, USA).

Because it exploits the bacteriophage for peptide synthesis, phage display is, however, limited in its use of nonstandard AAs. This is significant, as incorporation of non-natural functionality and various cyclization chemistries has proven to be a powerful way of converting a peptide lead into a drug with improved potency and PK properties. For example, cyclosporin – a naturally occurring N-to-C-cyclized macrocyclic peptide – is composed of numerous nonstandard AAs and is the only highly orally bioavailable approved peptide drug. Nature, too, exploits cyclization in bioactive peptides (e.g., marine and microbial macrocyclic peptides, depsipeptides and disulfide-rich peptides, such as conotoxins and cyclotides). Importantly, successful examples from these peptide families have led to promising drug development candidates. Indeed, a conotoxin peptide has been approved to treat neuropathic pain (Ziconotide) and the Craik group is using novel cyclization approaches to

develop an orally bioavailable version of the drug [8]. In the near future, optimizing natural peptides into viable drug leads will be accomplished through the use of enhanced high-throughput biological screening, including synthetic peptide libraries as well as *in silico* design and computational screening methods.

Novel peptide lead identification: going beyond phage display screening

Following lead identification, peptide chemists traditionally seek to explore non-natural peptide space through strategies such as non-natural side-chain substitution, *N*-alkylation and the engineering of backbone constraints to address PK liabilities and improve biological potency. While powerful, such systematic approaches have proven to be costly and time consuming. Perhaps the biggest improvements in peptide drug discovery will come through the use of improved *in vitro* display technologies that can efficiently incorporate non-natural AA side chains, nonstandard stereochemistries, methylation and various cyclizations. In what follows, we elaborate such peptide drug discovery that has benefited from the use of enhanced peptide display and screening technologies.

Peptide lead identification & optimization using advanced display & screening technologies: some recent examples

Although not a recent advancement in peptide-based drug discovery, mirror image phage display is a display methodology that can be used to go beyond the natural 20 AAs. As an example, the strategy of mirror image phage display used by Kay and coworkers led to the discovery of potent D-peptide inhibitors of gp41-mediated HIV entry (**Figure 1.3A**) [9].

Yamagishi and coworkers describe the application of a ribosome-expressed library of non-natural cyclic peptides to the discovery of blockers of the E2 ubiquitin ligase E6AP (**Figure 1.3B**) [10]. This work is groundbreaking because it uses a biologically generated *in vitro*-displayed peptide library to target a PPI otherwise deemed to be undruggable. The thioether cyclized 14-mer library (>1012 peptides) included five non-natural AAs and identified an initial lead that included one d-AA and four N-methyl AAs with a Kd of 0.6 nM. Interestingly, when not cyclized the Kd dropped nearly 300-fold. Furthermore, when the N-methyl AAs were replaced by their natural analogs no binding was detected. As a second application of this technology, Kawakami and coworkers screened a macrocyclic library against the extracellular domain of the VEGF receptor (VEGFR) [11]. This resulted in the discovery of several nanomolar inhibitors of VEGFR autophosphorylation

Figure 1.3 Examples of peptides discovered by various screening and design methodologies.

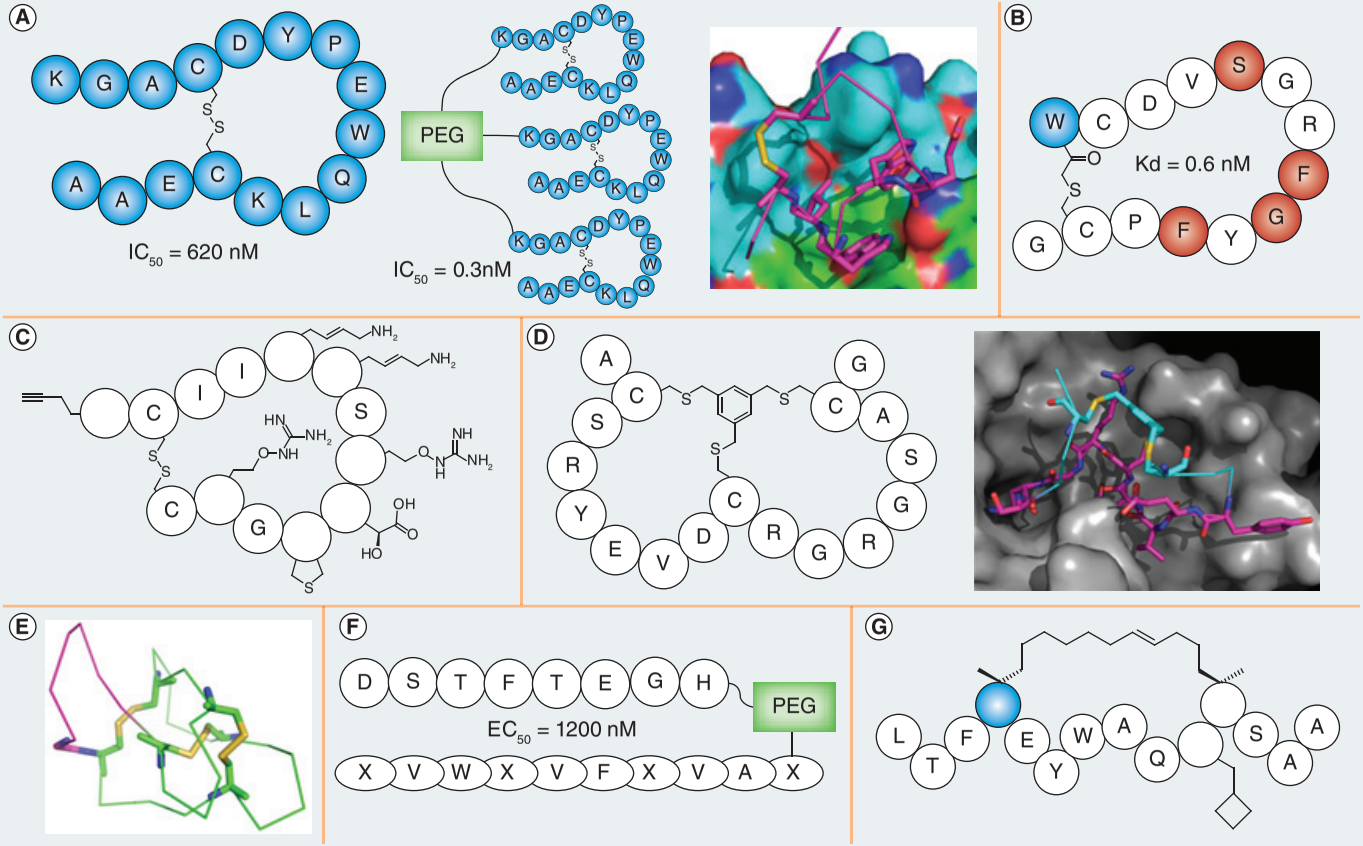


Figure 1.3 Examples of peptides discovered by various screening and design methodologies. (cont.).

In all depictions of a sequence, the N-terminal starts in the upper left corner. Spheres indicated α-amino acids (AAs). Ellipses indicate β-AAs. White indicates l-AAs. Cyan indicates d-AAs. Red indicates N-methyl AAs. Circles without letters indicate non-natural side chains, in which case the side chain is explicitly drawn. In all cases cyclization is explicitly drawn. **(A)** An example of inhibitors of gp41-mediated HIV entry [9]. The left frame shows a single peptide that consists of all d-AAs. The second frame depicts the trimeric peptide. The final frame shows a co-crystal structure with the monomeric peptide (magenta) and the gp41 trimer (cyan and green). **(B)** An example of a ubiquitin ligase E6AP blocker [10]. **(C)** A thrombin inhibitor [12]. **(D)** An inhibitor of urokinase-type plasminogen activator [13]. The first frame shows the sequence. The second frame shows the bicyclic peptide (magenta) as co-crystallized with uPA (gray). **(E)** A bradykinin agonist engineered into a cyclotide scaffold [15]. The magenta loop is the loop that was replaced with the known bradykinin antagonist. **(F)** A β-peptide antagonist of the class B GPCR GLP-1R [20]. The Xs are solvent-exposed charged residues intended to form salt bridges to stabilize the β-helix. **(G)** A stapled helical peptide that acts as a dual inhibitor of MDM2 and MDMX [21].

EC_{50}: Half maximal effective concentration; GLP-1R: Glucagon-like peptide-1 receptor; GPCR: G-protein coupled receptor; PEG: Polyethylene glycol.

and also showed that the newly discovered macrocycles blocked VEGFR signaling via the protein–protein interactions between VEGF and the extracellular domain of VEGFR. These peptide leads may offer significant selectivity advantages over ATP-competitive small-molecule inhibitors of VEGFR.

A recent discovery of thrombin inhibitors demonstrates the successful application of an mRNA display technique incorporating non-natural AAs into a cyclic peptide library (**Figure 1.3C**) [12]. This pioneering work by Szostak and coworkers demonstrates the advantages of using non-natural AAs. The library incorporated eight natural and 12 non-natural AAs. The equivalent library incorporating the 20 natural AAs was screened in parallel with the non-natural library. Both libraries produced low nanomolar thrombin inhibitors. No similarity, however, was reported between the various inhibitors. Furthermore, when the non-natural AAs were replaced with their natural counterparts, the resulting peptides showed no activity.

A final example demonstrates the generation of greater diversity using a chemical 'post-translational' modification strategy starting with a linear phage display library. The library was designed to incorporate three cysteine residues, resulting in a global bicyclic structure. Ultimately, a library of more than 109 bicyclic peptides was screened for inhibitors of human urokinase-type plasminogen activator (**Figure 1.3D**) [13]. The most potent peptide exhibited a Ki of 4 nM. As a control, they synthesized the equivalent peptide with the cysteines replaced by serines, preventing cyclization. The peptide potency decreased by approximately 4000-fold, further implicating the importance of rigid templates in peptide drug discovery.

Engineering natural peptide scaffolds

Numerous studies have exploited the remarkable characteristics of the cyclotide family of peptides [14]. Briefly, the cyclotides are a family of naturally-occurring disulfide rich peptides. They are unusually stable relative to peptides of comparable size. This enhanced stability is often attributed to a cyclic cysteine knot, which is composed of three disulfide bonds and an N-to-C-cyclized backbone. The cyclotides have six cysteine loops, four of which are amenable to sequence changes. Several groups have altered the sequence of one of these loops to that of a known bioactive peptide, thereby creating a cyclotide with new biological activity. As an example, Wong and coworkers replaced loop 6 of the cyclotide kalata B with a nine-residue bradykinin antagonist (**Figure 1.3E**) [15]. This resulted in a hybrid peptide with bradykinin antagonist properties and the enhanced PK properties of a cyclotide. Ultimately, the peptide was shown to be orally bioavailable.

Floudas and coworkers have developed a computational methodology to improve biological activity by stabilizing the active conformation of a peptide

through residue replacement [16]. Briefly, given desired Cα–Cα distances, they used a coarse grain approach to select sequences most likely to stabilize the conformation. The sequences were subjected to further calculations to calculate the extent to which a sequence would stabilize a target conformation. In their most recent application, they applied the method to the design of agonists, partial agonists and antagonists of the complement component 3a receptor (C3aR) [17]. Since the structure of C3aR is unknown, they applied the design method starting from the structure of the C-terminal 15 AAs of C3a. Importantly, they were able to improve a 170 nM agonist with 38% maximal stimulation into a 25 nM agonist with a 72% maximal stimulation.

In another example of data-driven peptide design, the approach of analysis of correlated mutations was applied to the discovery of biologically active peptides [18]. Using this method, several bioactive helical peptides were identified. For example, a shortened version of gp96 that significantly blocked the production of several chemokines in lipopolysaccharide-stimulated peripheral blood mononuclear cells was identified. The peptide was also shown to be active *in vivo* in a mouse model. A second helical peptide, designed from clusterin, was shown to inhibit the growth of the human lung carcinoma cell line A49 with an IC_{50} of 0.95 nM. Finally, a series of peptides were designed from the angiopoietin family to disrupt the helix–helix interactions critical for the activity of the angiopoietin proteins [19]. Out of 20 peptides tested, 11 showed significant antiangiogenic activity; the peptides also showed *in vivo* antiangiogenic activity.

The work of Schepartz and coworkers is noteworthy with respect to the design of a novel peptide agonist of the GLP1R, a class B GPCR (**Figure 1.3F**) [20]. As a starting point for their investigation, they exploited the observation that β-peptides can mimic α-helices. In addition, they exploited the knowledge of the binding and activation mechanism of exendin-4, a GLP-1 peptide agonist. Exendin-4 is 30 AA long and consists of a C-terminal helix and a disordered N-terminus. The helix binds the extracellular domain of GLP1R; the N-terminus is critical for activation. Initially, a β-helix was designed to mimic the C-terminal helix of exendin, resulting in 2 β-helicies that competed effectively for exendin-4 binding. The β-peptides were then linked to the N-terminus of GLP-1 via a PEG linker, and a lead peptide having a half maximal effective concentration (EC_{50}) of approximately 1 μM was identified.

As a final example of peptide engineering, the work of Sawyer and coworkers on stapled α-helical peptides shows the successful lead optimization of a dual inhibitor of MDM2 and MDMX with potent activities *in vitro* and *in vivo* [21]. The project started with a phage peptide that had a high affinity for binding to MDM2 and MDMX. Several modifications of the stapled analog led to an optimized stapled peptide analog, ATSP-7041, which was shown

to exhibit potent cellular activity and robust cellular uptake which were correlated with its amphipathicity, topochemical and solubility properties.

The importance of conformation in determining peptide properties & function

A challenging aspect of understanding the behavior of peptides is conformation. Small molecules typically have approximately 10–15 rotatable bonds. A peptide of only 15 AAs can easily eclipse 50 rotatable bonds. This greater flexibility allows peptides to display multiple faces of which some may be critical for biological function, while others may deleterious to such activity. As mentioned above, peptide conformational flexibility can also impede high-affinity binding. Hence, cyclization is often used to promote a desired conformation and minimize undesirable conformations. Several examples of how cyclization can affect biological activity were given in the preceding sections. Here, we focus on the critical role conformation plays in determining other peptide properties.

Conformational flexibility is increasingly recognized as important for cellular permeability [22]. Nevertheless, our current understanding of peptide cellular penetration is in its infancy. Beck and coworkers reported two conformational families of N-methylated cyclo(d-Ala-Ala5) analogs as putative templates for cell permeability and with the potential for bioavailability [23]. Eight of 54 analogs showed significant Caco-2 permeability. Interestingly, the data suggest that the permeability is not due to passive diffusion because the peptides are not permeable in the artificial membrane permeability (PAMPA) assay and the Caco-2 permeability does not correlate with polarity. Thus, the differences in permeability were attributed to an active transport mechanism. Lokey and coworkers provide evidence that conformational flexibility plays a significant role in passive diffusion as well. They synthesized nine diastereomers of cyclo[Leu-Leu-Leu-Pro-Tyr] and found nearly 2 log units difference in logP as measured in the PAMPA model [24]. The same group showed excellent agreement between a computational comparison of the conformational behavior of eight cyclic hexa- and hepta-peptides in low versus high dielectric media compared with experimentally measured PAMPA logP values [25]. While these results are with small peptides, we expect that the impact of conformation on PK properties will be amplified for larger peptides.

Using computation to proficiently enable novel peptide discovery

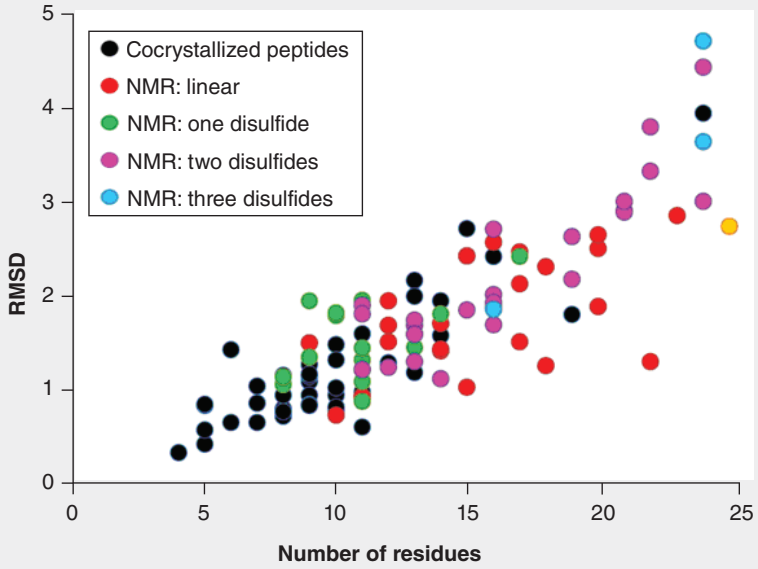
In a 2010 review article, we called for the synergistic use of computation, chemistry and biology to enable novel peptide drug discovery; other groups have expanded on that call [26,27]. Advantages of computational methods include quick turnaround, relatively low operational costs, the calculation

of static and dynamic structural models and the calculation of energetic quantities. Moreover, *in silico* methods can be expanded to include exotic peptide chemistries, including nonstandard side chains, N-methylation, termini modifications, β-peptides and various cyclizations. Some disadvantages include force field and scoring function inaccuracies and imperfect conformational search algorithms. In a previous section, we discussed the use of computational methods to identify potential peptide binding sites. Here we explore other uses of computation in peptide drug discovery.

The 2010 review article seems to have stimulated interest and several prominent groups and companies are now reporting results for virtual peptide docking. For example, researchers associated with the Rosetta project have reported results for the peptide docking program, FlexPepDock [28]. Similarly, researchers at Schrodinger recently reported the results of a peptide docking study using a modified version of the docking program Glide [29], and researchers at the University of Utrecht (The Netherlands) have described their peptide docking program Haddock [30]. Given the correct binding site, these programs can reproduce co-crystallized binding modes in the top 10 poses approximately 30–50% of the time starting with an apoprotein structure. The biggest difference between the results with peptides and small molecules is the amount of CPU time needed. For small molecules the times involved are often seconds to minutes, whereas the CPU times for peptides can be in the order of days to weeks depending on the size of the peptides. For problems such as identifying the binding mode for a known binding peptide this amount of CPU time is trivial. For peptide discovery work, where scientists are often interested in evaluating massive numbers of peptides, the CPU requirements become the limiting factor.

Novel computational methods have also been developed for predicting the 3D structures of linear and cyclic peptides with up to approximately 30 AAs [31,32]. These methods are, however, limited by their use of knowledge extracted from structures in the PDB. This limits their ability to function with important families of non-natural AAs. We have developed an alternative to such knowledge-based approaches. Briefly, our approach (CMDpeptide^{SM}, CT, USA) calculates conformational ensembles using an optimized Monte Carlo algorithm and rank-orders the conformations using an optimized, physics-based scoring function. In most cases, CMDpeptide^{SM} can locate distinctly near-native structures. We illustrate some results of our method in **Figures 1.4–1.6**; where possible, we have included cyclic peptides and peptides incorporating non-natural AAs. For example, **Figure 1.5** shows the results with cyclosporin and **Figure 1.6** shows results with two naturally occurring disulfide-rich peptides, uroguanylin and an α-conotoxin, and a stapled α-helix.

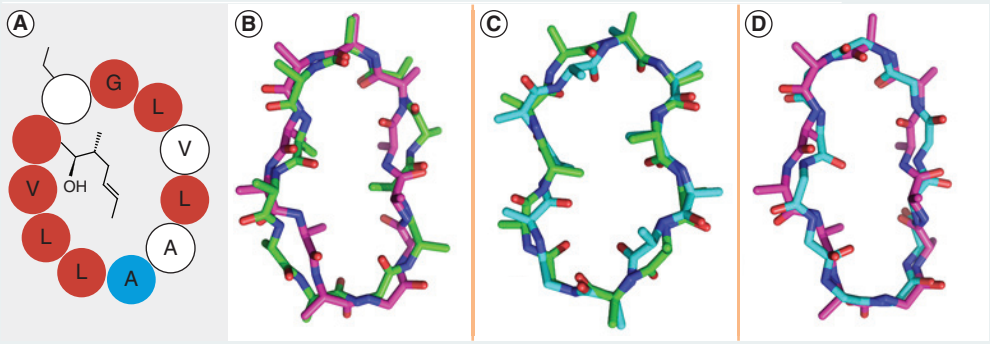
Figure 1.4 The overall results of CMDpeptide^{SM} to identify native conformations on a set of peptides with experimental structures.



The different colors indicate whether the peptide structure is from an x-ray co-crystal structure (black), a linear NMR structure (red) or cyclic NMR structure with one (green), two (magenta) or three (blue) disulfide bonds. The RMSD is the lowest RMSD found within the full ensemble generated and is given in units of Å.
RMSD: Root-mean-square deviation.

Indeed, we have developed a versatile, multimodule, fully integrated, computational peptide drug discovery platform called CMDinventus^{SM}, of which CMDpeptide^{SM} is a single module. Unlike other approaches, CMDinventus^{SM} is rigorously physics-based, has been developed from the ground-up to address the peptide discovery problem and has been parameterized to work with proteogenic and nonproteogenic peptide chemistries. CMDinventus^{SM} has been and is being used to design multiple lead series for disparate protein-protein interaction targets. CMDinventus^{SM} can be readily configured to perform massive *in silico* natural and non-natural peptide screening experiments and combined with *in vitro* technologies to make our vision of synergistic peptide drug discovery a fruitful reality.

Figure 1.5 Cyclosporin.

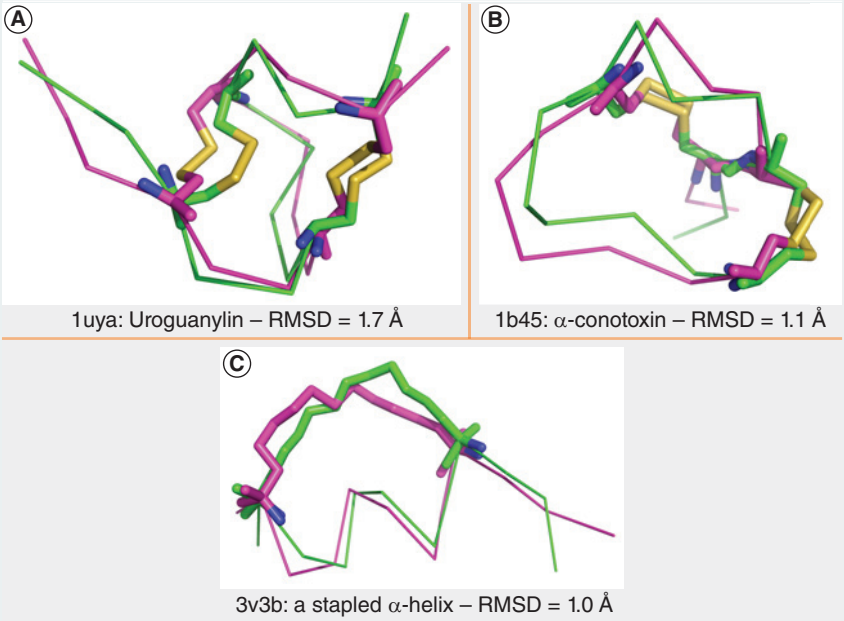


(A) The sequence of cyclosporin. The sequence depiction is as described in **Figure 1.3**. **(B)** A comparison of the two conformations from co-crystal structures. The conformation in green is cyclosporin as bound to cyclophillin A (1cwa). The conformation in magenta is cyclosporin as bound to an antibody. **(C)** The closest conformation to the conformation bound to cyclophillin A. **(D)** The closest conformation to that bound to the antibody. The two conformations in cyan are the closest from the CMDpeptide^{SM} conformational ensemble to the given bound conformation. The RMSD between the two bound conformations **(A)** is 1.3 Å while the RMSDs between the nearest computationally generated conformation **(B)** and **(C)** are 0.91 and 0.95 Å, respectively.

Summary & conclusion

The renewed excitement around peptide drug discovery is both genuine and well justified. Peptides offer an opportunity to target novel regions of biological space, for example, intracellular PPIs, not suitable for either small molecules or antibodies. Even with the advent of display technologies capable of incorporating more non-natural functionality, the challenge of using peptides to transform drug discovery should not be underestimated. Indeed, as the numbers of possible cyclizations and non-natural AAs increases, display technology will only be able to scratch the surface in terms of the potential exploitable peptide chemical space. To fully exploit peptides for drug discovery, it will be necessary to combine the best of existing *in silico* and *in vitro* approaches. For example, through the use of computation, display technologies could be focused on the scaffolds and side chains that achieved an optimized pharmacophore to a particular therapeutic target, including the possibility of designing multitarget specificity. Unquestionably, the importance of PK, including protease stability, distribution, cell penetration, toxicity and immunogenicity, must be considered early in the peptide drug discovery process. As our understanding of these phenomena improves, the opportunity to harness predictive knowledge from leveraging

Figure 1.6 Three macrocyclic peptides.



In all cases, the green conformation is the experimentally observed conformation either from an NMR structure or a co-crystal structure and the magenta conformation is that from the CMDpeptide^{SM} ensemble closest to the native conformation. **(A)** Uroguanylin (PDB code 1uya; RMSD: 1.7 Å). **(B)** An α-conotoxin (pdb code 1b45; RMSD: 1.1 Å). **(C)** A stapled α-helix (pdb code 3v3b; RMSD: 1.0 Å).
RMSD: Root-mean-square deviation.

the structure–property relationships of peptides will fulfill the promise of innovative technologies and development of breakthrough medicines.

Financial and competing interests disclosure

J Audie, DJ Diller and M Jarosinski are employed by CMDBioscience, a start-up company that specializes in computer-assisted peptide drug design. TK Sawyer is Emeritus Founding Chair of the CMDBioscience SAB. The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

No writing assistance was utilized in the production of this manuscript.

Summary

- There is a large physiochemical and biological gap between small molecules and antibodies, a gap that peptides have the potential to fill.
- The key challenges for peptides to become drugs are largely pharmacokinetic, including oral bioavailability, cellular penetration, proteolytic stability and renal clearance.
- The key to addressing the pharmacokinetic problems associated with peptides is the incorporation of more non-natural functionality such as D-amino acids, N-methylation and cyclization.
- Understanding the complex conformational behavior of peptides is critical for understanding many of their properties and ultimately for engineering desirable properties into the next generation of peptide-based drugs.
- To fully exploit peptides for transformational drug discovery, it will be necessary to integrate the most proficient of both *in silico* and *in vitro* methodologies.

Key terms

Peptide pharmacokinetics:	the key pharmacokinetic (PK) challenges with peptides are proteolytic stability, cellular penetration and renal clearance. This differs from small molecules, which encounter cytochrome P450 metabolism and idiosyncratic toxicity.
Peptide conformation:	the complex conformational behavior of peptides is another critical difference from small molecules. Conformation is crucial for a peptide's pharmacokinetic and pharmacologic behavior. Peptides can adopt several different faces, each of which could be responsible for a different biological activity or characteristic.
Non-natural functionality:	the key to addressing many of the shortcomings of peptides is to incorporate non-natural functionality. This includes cyclization, D-amino acids, N-methyl amino acids and non-natural side chains.
Display technology:	includes any technology that relies on a biological mechanism to synthesize and display a large library of peptides for screening against pharmaceutical targets. This includes classic phage display and the more recent development of mRNA display technologies that have allowed the incorporation of significantly greater non-natural functionality in the resulting libraries.

References

1. Reichert J. Development Trends for Peptide Therapeutics. Peptide Therapeutic Foundation, San Diego, California, USA. (2010).
2. Dimasi JA, Feldman L, Seckler A, Wilson A. Trends in risks associated with new drug development: success rates for investigational drugs. *Clin. Pharmacol. Ther.* 87(3), 272–277 (2010).
3. Hoare SR. Mechanisms of peptide and nonpeptide ligand binding to Class B G-protein-coupled receptors. *Drug Discov. Today* 10(6), 417–427 (2005).
4. London N, Movshovitz-Attias D, Schueler-Furman O. The structural basis of

peptide–protein binding strategies. *Structure* 18(2), 188–199 (2010).

5. Yip KY, Utz L, Sitwell S *et al.* Identification of specificity determining residues in peptide recognition domains using an information theoretic approach applied to large-scale binding maps. *BMC Biol.* 9, 53 (2011).
6. Trabuco LG, Lise S, Petsalaki E, Russell RB. PepSite: prediction of peptide-binding sites from protein surfaces. *Nucleic Acids Res.* 40, W423–W427 (2012).
7. Lavi A, Ngan CH, Movshovitz-Attias D *et al.* Detection of peptide-binding sites on protein surfaces: the first step toward the modeling and targeting of peptide-mediated interactions. *Proteins* 81(12), 2096–2105 (2013).
8. Clark RJ, Akcan M, Kaas Q, Daly NL, Craik DJ. Cyclization of conotoxins to improve their biopharmaceutical properties. *Toxicon* 59(4), 446–455 (2012).
9. Welch BD, Francis JN, Redman JS *et al.* Design of a potent d-peptide HIV-1 entry inhibitor with a strong barrier to resistance. *J. Virol.* 84(21), 11235–11244 (2010).
10. Yamagishi Y, Shoji I, Miyagawa S *et al.* Natural product-like macrocyclic N-methyl-peptide inhibitors against a ubiquitin ligase uncovered from a ribosome-expressed *de novo* library. *Chem. Biol.* 18(12), 1562–1570 (2011).
11. Kawakami T, Ishizawa T, Fujino T, Reid PC, Suga H, Murakami H. *In vitro* selection of multiple libraries created by genetic code reprogramming to discover macrocyclic peptides that antagonize VEGFR2 activity in living cells. *ACS Chem. Biol.* 8(6), 1205–1214 (2013).
12. Schlippe YV, Hartman MC, Josephson K, Szostak JW. *In vitro* selection of highly modified cyclic peptides that act as tight binding inhibitors. *J. Am. Chem Soc.* 134(25), 10469–10477 (2012).
13. Angelini A, Cendron L, Chen S *et al.* Bicyclic peptide inhibitor reveals large contact interface with a protease target. *ACS Chem. Biol.* 7(5), 817–821 (2012).
14. Poth AG, Chan LY, Craik DJ. Cyclotides as grafting frameworks for protein engineering and drug design applications. *Biopolymers* 100(5), 480–491 (2013).
15. Wong CT, Rowlands DK, Wong CH *et al.* Orally active peptidic bradykinin B1 receptor antagonists engineered from a cyclotide scaffold for inflammatory pain treatment. *Angew Chem. Int. Ed. Engl.* 51(23), 5620–5624 (2012).
16. Fung HK, Floudas CA, Taylor MS, Zhang L, Morikis D. Toward full-sequence *de novo* protein design with flexible templates for human beta-defensin-2. *Biophys. J.* 94(2), 584–599 (2008).
17. Bellows-Peterson ML, Fung HK, Floudas CA *et al. De novo* peptide design with C3a receptor agonist and antagonist activities: theoretical predictions and experimental validation. *J. Med. Chem.* 55(9), 4159–4168 (2012).
18. Kliger Y, Levy O, Oren A *et al.* Peptides modulating conformational changes in secreted chaperones: from *in silico* design to preclinical proof of concept. *Proc. Natl Acad. Sci. USA* 106(33), 13797–13801 (2009).
19. Palmer GM, Tiran Z, Zhou Z *et al.* A novel angiopoietin-derived peptide displays anti-angiogenic activity and inhibits tumour-induced and retinal neovascularization. *Br. J. Pharmacol.* 165(6), 1891–1903 (2012).
20. Denton EV, Craig CJ, Pongratz RL *et al.* A beta-peptide agonist of the GLP-1 receptor, a Class B GPCR. *Org. Lett.* 15(20), 5318–5321 (2013).
21. Chang YS, Graves B, Guerlavais V *et al.* Stapled alpha-helical peptide drug development: a potent dual inhibitor of MDM2 and MDMX for p53-dependent cancer therapy. *Proc. Natl Acad. Sci. USA* 110(36), E3445–E3454 (2013).
22. Bockus AT, Mcewen CM, Lokey RS. Form and function in cyclic peptide natural products: a pharmacokinetic perspective. *Curr. Top. Med. Chem.* 13(7), 821–836 (2013).
23. Beck JG, Chatterjee J, Laufer B *et al.* Intestinal permeability of cyclic peptides: common key backbone motifs identified. *J. Am. Chem. Soc.* 134(29), 12125–12133 (2012).
24. Rezai T, Yu B, Millhauser GL, Jacobson MP, Lokey RS. Testing the conformational hypothesis of passive membrane permeability using synthetic cyclic peptide diastereomers. *J. Am. Chem. Soc.* 128(8), 2510–2511 (2006).
25. Rezai T, Bock JE, Zhou MV, Kalyanaraman C, Lokey RS, Jacobson MP. Conformational flexibility, internal hydrogen bonding, and passive membrane permeability: successful *in silico* prediction of the relative permeabilities of

cyclic peptides. *J. Am. Chem. Soc.* 128(43), 14073–14080 (2006).

26. Audie J, Boyd C. The synergistic use of computation, chemistry and biology to discover novel peptide-based drugs: the time is right. *Curr. Pharm. Des.* 16(5), 567–582 (2010).
27. Honarparvar B, Govender T, Maguire GE, Soliman ME, Kruger HG. Integrated approach to structure-based enzymatic drug design: molecular modeling, spectroscopy, and experimental bioactivity. *Chem. Rev.* 114(1), 493–537 (2014).
28. Raveh B, London N, Zimmerman L, Schueler-Furman O. Rosetta FlexPepDock ab-initio: simultaneous folding, docking and refinement of peptides onto their receptors. *PLoS ONE* 6(4), e18934 (2011).
29. Tubert-Brohman I, Sherman W, Repasky M, Beuming T. Improved docking of polypeptides with glide. *J. Chem. Inf. Model.* 53(7), 1689–1699 (2013).
30. Trellet M, Melquiond AS, Bonvin AM. A unified conformational selection and induced fit approach to protein–peptide docking. *PLoS ONE* 8(3), e58769 (2013).
31. Beaufays J, Lins L, Thomas A, Brasseur R. *In silico* predictions of 3D structures of linear and cyclic peptides with natural and non-proteinogenic residues. *J. Pept. Sci.* 18(1), 17–24 (2012).
32. Thevenet P, Shen Y, Maupetit J, Guyon F, Derreumaux P, Tuffery P. PEP-FOLD: an updated *de novo* structure prediction server for both linear and disulfide bonded cyclic peptides. *Nucleic Acids Res.* 40, W288–W293 (2012).