Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Structural Biology

Innovative strategies for modeling peptide–protein interactions and rational peptide drug design

Lara Scharbert[1,2] and Birgit Strodel[1,2]

This review highlights cutting-edge techniques for modeling peptide–protein interactions and advancing computer-aided peptide–drug design. We examine significant progress in generating peptide poses through docking and artificial intelligence (AI), assessing peptide flexibility via enhanced molecular dynamics simulations, and analyzing binding interactions through free energy calculations. Additionally, we discuss how these insights can inform the rational design of therapeutic peptides by utilizing free energy metrics and strategic modifications to enhance their binding affinity and therapeutic potential. Looking forward, further integrating AI will be crucial for optimizing peptide design and enhancing drug development efforts.

Addresses
[1] Institute of Theoretical and Computational Chemistry, Heinrich-Heine-University Düsseldorf, Universitaetsstr. 1, Düsseldorf, 40225, Germany
[2] Institute of Biological Information Processing, Structural Biochemistry (IBI-7), Forschungszentrum Jülich, Wilhelm-Johnen-Straße, Jülich, 52428, Germany

Corresponding author: Strodel, Birgit (birgit.strodel@uni-duesseldorf.de)

Current Opinion in Structural Biology 2025, **93**:103083

This review comes from a themed issue on **Biophysical Methods (2025)**

Edited by **Hoi Sung Chung** and **Francesca M. Marassi**

For a complete overview see the Issue and the Editorial

Available online xxx

https://doi.org/10.1016/j.sbi.2025.103083

Introduction

Peptides are estimated to be involved in 15–40 % of all protein-related interactions [1]. These peptide–protein interactions (pepPIs) play crucial roles in processes such as signal transduction, immune regulation, and proteolytic processing, making them attractive targets for therapeutic intervention. Currently, more than 80 approved peptide drugs are available on the global market. Although therapeutic peptides account for only 5 % of the pharmaceutical industry, the market is expected to grow with a projected compound annual growth rate of 10.8 % from 2025 to 2030 [2,3]. Small molecules still dominate the global pharmaceutical market with 75 % of the shares as they offer several advantages, including low production costs, oral bioavailability, and good membrane permeability. However, their small size limits their ability to inhibit interactions involving large surface areas, such as protein–protein interactions (PPIs) [4]. Due to their larger size and inherent flexibility, peptide drugs, in contrast, have a greater potential to target and inhibit PPIs that were previously considered 'undruggable' [5].

In the early days of peptide–drug discovery, scientists were restricted to natural sources. For instance, the first peptide–drug, insulin, was derived from animals, often leading to allergic reactions. With the advent of recombinant technology in the 1980s, clean and selective production of peptides and proteins in cell cultures became feasible. Another major milestone was the invention of a target-based search for novel peptide–drug candidates; the phage display technique enabled the screening of large peptide libraries (up to 10^{10}) against a target protein using bacteriophages [4,6].

With advances in computational power and knowledge, computer-aided drug design (CADD) has become an essential tool in drug discovery. CADD accelerates drug development while reducing costs and resources by enabling virtual screening of vast compound databases and using a knowledge-driven approach to understanding protein–ligand interactions [7]. Building on the principles established in CADD, computer-aided peptide–drug design focuses specifically on optimizing peptide interactions with their targets. A rational peptide design approach necessitates a comprehensive understanding of the underlying pepPIs; therefore, reliable structural information about the involved binding partners is essential. For targets, existing structures can often be utilized as experimental high-resolution methods such as X-ray crystallography and nuclear magnetic resonance spectroscopy have determined over 200,000 protein structures available in the Protein Data Bank (PDB) to date. However, these techniques are complex, time-consuming, costly, and may result in non-native conformations [8,9]. AlphaFold developed by Google DeepMind revolutionized structural biology by leveraging artificial intelligence (AI) to predict protein structures solely from the amino acid sequence, now achieving results in minutes using the AlphaFold server [10,11].

While protein structures for most target proteins can typically be found in databases, establishing the structures of peptides and their interactions with proteins poses greater challenges. This difficulty applies to both experimental and simulation approaches, largely due to the inherent flexibility of peptides, which results in a vast number of possible conformations. Moreover, while small-molecule binding pockets are usually deeply buried and well-defined, peptides rather bind to larger, disordered binding sites on the surface of the protein (Figure 1a vs. b). These issues can lead to uncertainty in predicting binding poses using docking techniques. To address this, molecular dynamics (MD) simulations enable the sampling of the dynamics of pepPIs, facilitating a more holistic analysis of the conformational landscape [7,12].

This review aims to highlight some of the latest and most innovative methods for modeling pepPIs and computer-aided peptide drug design. While comprehensive reviews have previously covered individual aspects of these topics [12,13], we focus on the most recent advancements in the field, including peptide pose generation, evaluation of peptide flexibility and interactions at the binding site, and the rational design of new peptide drugs. These various approaches can be combined into a workflow as shown in Figure 2.

Generation of peptide–protein binding poses

A classic structure-based drug design journey typically starts with the identification of a therapeutic target. If a high-resolution structure of the target protein is not yet available in the PDB, computational methods such as homology modeling—though this approach is becoming increasingly less relevant—and AI approaches like AlphaFold can be employed, to predict its three-dimensional conformation [9]. With the protein structure by hand, the next step is to determine the area where the ligand can bind.

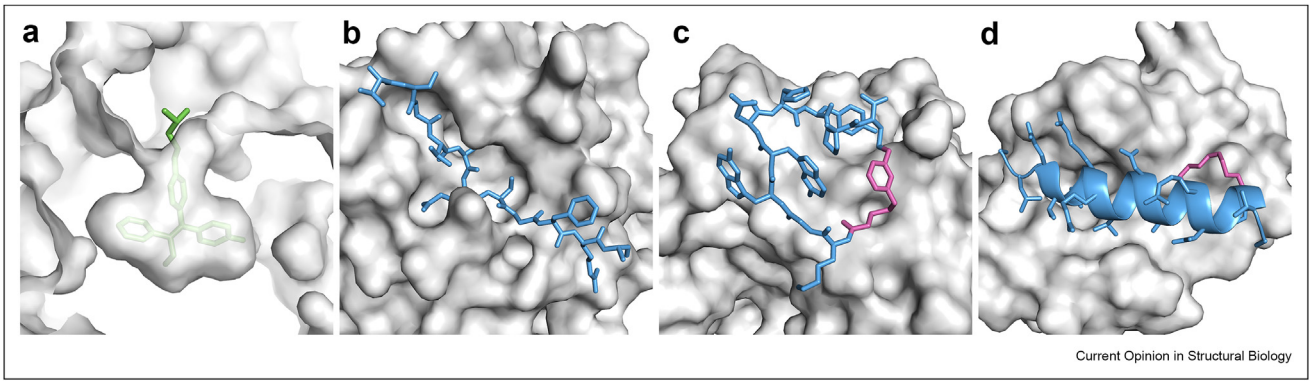
Pocket detection

If no potential binding site is known, there are several open source algorithms for locating and evaluating surface pockets and cavities. Two prominent tools are the web server CASTp 3.0 [14] and Fpocket [15], which utilize a geometry-based method that relies on spheres to detect grooves on the protein surface. In peptide drug discovery, the large, relatively flat interface of pepPIs contrasts sharply with the well-defined and often deeply buried binding pockets commonly found for small molecules. Standard algorithms could therefore be misleading in predicting binding sites for peptides. To address this problem, methods for the identification of pepPI sites were developed. The well-established web server PepSite utilizes a scoring method based on residue binding preferences derived from known peptide–protein structures to predict potential peptide binding sites on protein surfaces [16]. The recently introduced SiteFerret is a novel method for automatic pocket detection that effectively identifies both small-molecule binding pockets and peptide binding sites, utilizing an algorithm that combines hierarchical clustering of virtual probe spheres with the Isolation Forest method to rank putative pockets [17].

Global docking

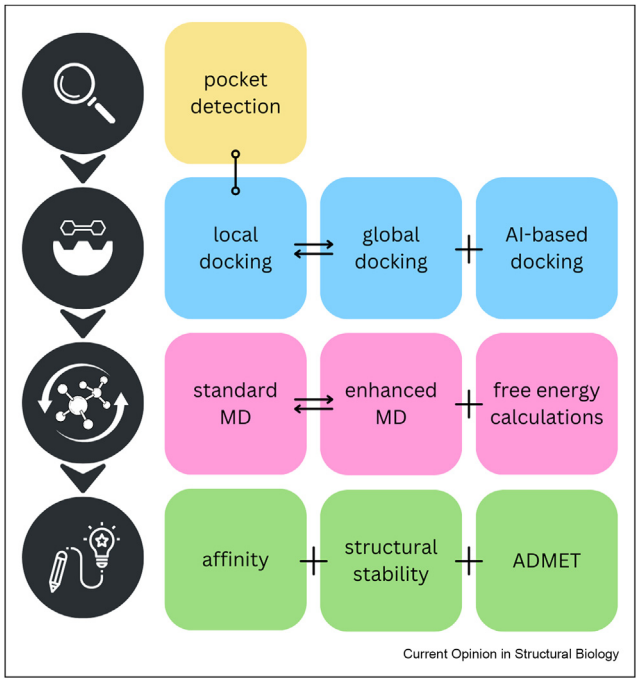
A fast-forward method for determining the structures of peptide–protein complexes—skipping the step of binding pocket detection—is to employ global docking. To this end, the user can choose from a vast range of docking programs that can be classified into protein–protein, peptide–protein, and small-molecule–protein docking. Agrawal et al. compared a diverse set of docking methods for peptide–protein docking using

Figure 1



Representative examples of different protein binding sites occupied by (a) a small-molecule inhibitor (PDB code 3ERT), (b) an extended linear peptide (PDB code 7T70), (c) a cyclic peptide (PDB code 6XIB), and (d) a stapled helical peptide (PDB code 3MK8). The proteins are shown as gray surface and the ligands as sticks, with green for the small molecule, blue for the peptides, purple for peptide modifications, and a blue cartoon for the helix. PDB, Protein Data Bank.

Figure 2



Workflow of a typical peptide–drug design study. The process begins with identifying the peptide binding site on the protein surface if local docking is applied in the next step. Global docking involves the concurrent identification of both the binding site and binding pose. In a consensus docking approach, local or global docking is combined with AI-based pepPI prediction. The peptide pose is then validated and evaluated through standard or enhanced MD simulations, which also quantify peptide flexibility at the binding site and allow for the calculation of binding free energies. Ultimately, the insights gained from pepPI modeling inform peptide design, enhancing the affinity, structural stability, and pharmacological activity of the peptide–drug candidate. AI, artificial intelligence; MD, molecular dynamics; pepPIs, peptide–protein interactions.

133 peptide–protein complexes, focusing on peptides that are 9–15 residues long [18]. They found that FRODOCK [19], a fast global protein–protein docking web server, performed best in blind docking. In the work of Weng et al. a similar, but more comprehensive approach was followed, where they compared 14 docking programs—three for small-molecule–protein docking, three for protein–protein docking, and eight for peptide–protein docking—on a set of 185 peptide–protein complexes, comprising peptides of 5–20 residues [20]. In contrast to Agrawal et al. they found that peptide–protein docking algorithms generally performed better than protein–protein or small-molecule docking algorithms. Notably, in global docking, the web server HPEPDOCK [21], which employs rapid peptide conformation sampling and ensemble docking techniques, produced the best predictions for the entire data set.

Local docking

When the binding site is known, local docking is the preferred and more efficient method. In this category, AutoDock CrankPep (ADCP) [22] has demonstrated the best overall results [20]. The high number of rotatable bonds in peptides contributes to their conformational versatility, which distinguishes them not only from the more rigid small molecules but also from folded proteins. This enormous flexibility is a big challenge for standard docking programs [12]. The peptide docking program ADCP combines the peptide backbone conformation sampling method of CRANKITE [23] with the grid-based receptor representation of AutoDock [24]. The extensive conformational space is efficiently navigated by ADCP using a Metropolis Monte Carlo search strategy, sampling peptide positions and orientations within the receptor's binding site, thereby producing docking poses [22]. Moreover, ADCP is also capable of docking cyclic peptides that are either linked through their backbone or side chain disulfide bonds [25]. Besides peptide flexibility, some docking programs—such as HADDOCK [26], CABS-Dock [27], and FlexPepDock [28]—also account for receptor flexibility by allowing movement of selected side chains and in some cases also refinement of the backbone.

Defining the appropriate size of the binding site can be a challenge in local docking, as, in particular, long peptides can vary dramatically in their full size depending on how expanded or compact they are. Weng et al. studied the impact of the binding site size on the performance of AutoDock or Vina-based Docking programs, like ADCP, by comparing docking boxes generated based on the co-crystallized ligand (ligBS) with those generated based on the peptide length (pepBS) [20]. Compared to the ligBS, the success rates for pepBS were considerably lower for most docking programs, especially in the case of ADCP, and dropped even more with increasing peptide length. Therefore, when using AutoDock or Vina-based docking programs, it is advisable to define the binding site more precisely – preferably by using the co-crystallized ligand as template if available or restricting the docking box to the relevant binding site residues.

AI-based docking

In recent years, AI-based methods for protein structure prediction have emerged, with AlphaFold [10] consistently leading this field and often serving as the benchmark for state-of-the-art comparisons. The release of AlphaFold-Multimer (AF2multi) [29], an update to AlphaFold 2, enabled the prediction of protein–protein complexes, including homomeric and heteromeric protein complexes. While AF2multi was primarily designed for protein–protein complexes, it has also been shown to be effective for peptide–protein docking problems [30,31]. Shanker and Sanner compared AF2multi with

two other deep learning models, OmegaFold and AlphaFold 2 Monomer, as well as their peptide-docking software ADCP [31]. The dataset used in this study contains 99 nonredundant peptide−protein complexes, with the vast majority of peptides having a length of 5−25 amino acids and containing different secondary structures (48 % coil, 34 % helix, and 18 % strand). AF2multi outperformed all other methods with a success rate of 53 %, whereas ADCP achieved only 23 % for the top-ranked solutions. However, when considering all generated solutions beyond the top-ranked poses, ADCP's success rate rises to 62 %, indicating that there is room for improvement in the scoring function. The authors also noted that AF2multi and ADCP performed complementarily, each succeeding on different peptide−protein complexes. AF2multi is suitable for docking long, linear peptides composed of standard amino acids, accounting for both peptide and receptor flexibility. In contrast, ADCP allows peptide flexibility while docking into rigid receptors [22], supporting peptides up to 30 residues, including cyclic structures and noncanonical amino acids. For peptides compatible with both methods, a consensus docking strategy combining ADCP and AF2multi, along with an effective selection mechanism, achieved a 60 % success rate for top-ranking poses and 66 % within the top five [31].

With the release of the AlphaFold Server powered by AlphaFold 3 (AF3), predicting peptide−protein complexes has become easily and rapidly accessible to everyone [11]. Nevertheless, in the case of peptide−protein docking, latest benchmark studies have found that AF2multi performs slightly better than AF3 [32]. The deep learning model ESMFold is a protein structure prediction tool, that, unlike AlphaFold, does not rely on external databases, template searches, or multiple sequence alignments. ESMFold predicts protein structures directly from amino acid sequences, while being up to 60 times faster than AlphaFold [33]. In recent studies, ESMFold's performance in peptide−protein docking was evaluated and compared with the latest AlphaFold versions. Despite generally lower accuracy, ESMFold occasionally outperformed AF3. Given its computational efficiency, ESMFold can be an alternative or a beneficial complement to other methods in pepPI studies [34].

Limitations of docking

A limitation of docking approaches and AI-based methods for predicting peptide−protein complexes is that they provide only static snapshots of the peptides and proteins, excluding their dynamic nature and, in particular, the flexibility of peptides and thus pepPIs, which can be addressed through MD simulations (Figure 2).

Evaluation of poses, interactions, and peptide flexibility

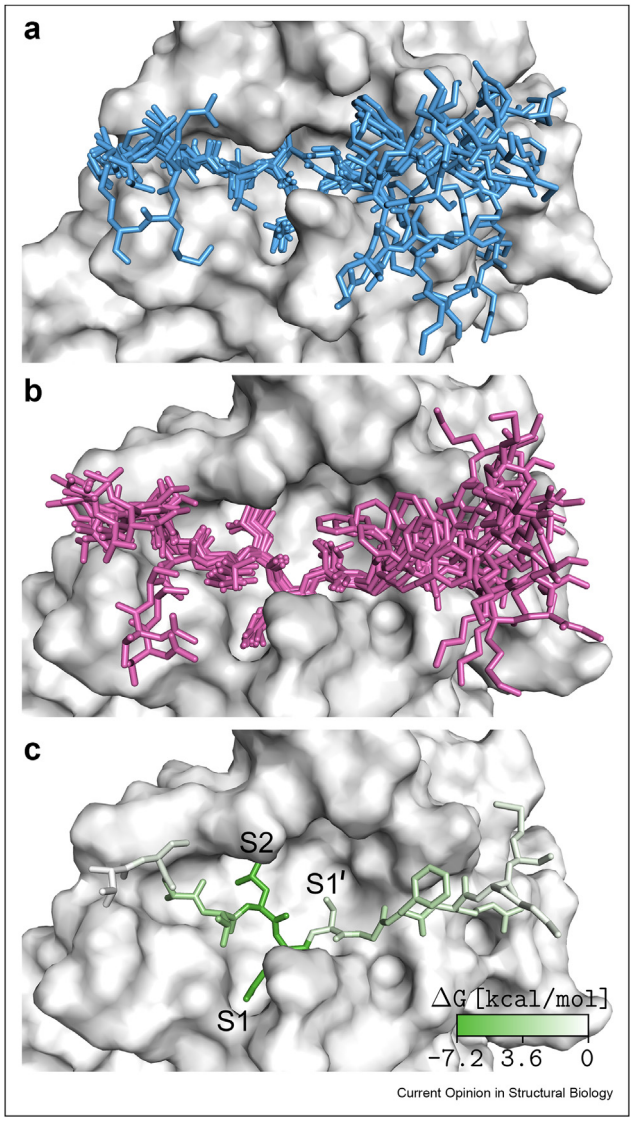
Standard MD

Molecular dynamics simulations have long been a cornerstone in understanding the dynamic nature of biological macromolecules, providing atomistic insights into conformational flexibility and interaction mechanisms that often elude static structural determination methods. Given the limitations of scoring functions in docking methods, evaluating multiple top-ranked poses (e.g. the top 3 or 5; see Figure 3a for a presentation of different docking poses) is advisable rather than relying solely on the highest-ranked pose. Furthermore, when combining physics-based and AI-based methods in a consensus approach, it is crucial to identify which predicted structures are more reliable [31]. MD simulations serve as valuable tools for assessing these docking predictions as they allow for the exploration of the stability and conformational dynamics of peptide−protein complexes over time (see Figure 3b for the peptide flexibility sampled by MD). This dynamic perspective is essential for understanding the interaction mode and affinity of pepPIs, forming the basis for subsequent drug design studies [7]. Recent examples of combining docking and MD simulations for the development of therapeutic peptides include designing peptides to target multiple over-expressed receptors in tumor cells [35], creating peptide inhibitors for the interaction between the Omicron SARS-CoV-2 spike protein and its receptor ACE2 [36], and optimizing a potent macrocyclic peptide that interferes with the protein−protein interaction between the Programmed Cell Death Protein 1 and its Ligand 1 (PD-1/PD-L1) through rational design [37].

Enhanced MD

However, traditional MD simulations can be computationally intensive and struggle to efficiently sample rare events, such as transitions between the conformational states of peptides in peptide−protein complexes, where pepPIs may raise energy barriers. To address these limitations, enhanced MD sampling techniques, such as replica exchange MD, accelerated MD, or Gaussian accelerated MD (GaMD), have been refined or developed to improve the sampling efficiency of MD simulations of PPIs or pepPIs [38,39]. With peptide Gaussian accelerated molecular dynamics (Pep-GaMD), Wang et al. introduced a GaMD technique that is specifically tailored for modeling the high flexibility of peptides [40]. Its dual-boost algorithm applies two distinct boost potentials: one specifically targeting the peptide's essential potential energy and another for the remaining potential energy of the entire system. The same group has published a PepBinding

Figure 3



The flexibility of pepPIs exemplified by docking and MD simulations of the SARS-CoV-2 main protease (PDB code 7T70) with its substrate. (a) The top 10 ranked structures after peptide docking with ADCP. (b) The 10 most frequently sampled peptide conformations during a 100 ns MD simulation. (c) Per-residue ΔG_{bind} decomposition of the peptide, from the highest (green) to the lowest (white) contribution to the affinity (measured as negative energies, see color scale in the bottom right corner). Subsites S1 and S2 are occupied by key residues, while S1′ has potential for peptide modifications. *ΔG_{bind}, binding free energy; ADCP, AutoDock CrankPep; AI, artificial intelligence; MD, molecular dynamics; PDB, Protein Data Bank; pepPIs, peptide–protein interactions.*

workflow for predicting peptide–protein structures combining peptide docking, enhanced sampling using Pep-GaMD, and structural clustering [41]. They demonstrated their workflow on seven peptide–protein models generated via docking with HPEPDOCK, initially classified as inaccurate to medium-quality based on the Critical Assessment of PRediction of Interactions criteria. Each model underwent a 200 ns Pep-GaMD simulation, successfully refining them into five medium-quality and two acceptable-quality structures.

Free energy calculations

While docking methods use fast empirical or force field-based scoring functions to estimate the affinity between protein and peptide, free energy calculations combined with MD simulations can achieve a more accurate estimation of binding affinity [42]. To this end, the most commonly used free energy calculation methods are molecular mechanics/Poisson–Boltzmann surface area (MM/PBSA) or molecular mechanics/generalized Born surface area (MM/GBSA), free energy perturbation (FEP), and thermodynamic integration (TI). While end-point methods, like MM/PBSA and MM/GBSA, are more computationally efficient, requiring only a single trajectory, FEP and TI are both pathway-based methods that are considered more accurate but involve much higher computational costs. Therefore, MM/PB(GB)SA has become a popular technique to calculate the binding free energy (ΔG_{bind}) between a protein and ligand [43]. Weng et al. suggested that the choice between MM/PBSA and MM/GBSA may depend on peptide length and other system-specific factors. Their study showed that MM/PBSA yielded higher accuracy for short peptides (5–12 residues), while MM/GBSA performed better for medium-length peptides (20–25 residues) [44]. Furthermore, both methods are sensitive to parameters such as the solute dielectric constant and the choice of force field, and are challenged by the high computational cost and limited accuracy of entropy calculations, typically failing to (correctly) capture conformational entropy [45].

Among the available tools for MM/PB(GB)SA calculations, Amber includes the widely used MMPBSA.py script within the AmberTools package, which enables free energy calculations based on MD snapshots [46]. To enhance accessibility for GROMACS users, Valdés-Tresanco developed gmx_mmpbsa, an adaptation of MMPBSA.py for processing trajectories and topology files generated with GROMACS [47]. Compared with other MM/PB(GB)SA implementations for GROMACS, such as g_mmpbsa and GMXPBSA2.1, gmx_mmpbsa offers several advantages [48]. It is compatible with all GROMACS versions and provides enhanced calculation and analysis features, including ΔG_{bind} calculations with different solvation models, stability assessments, computational alanine scanning, ΔG_{bind} decomposition, and entropy corrections. Furthermore, it supports the Amber, OPLS, and CHARMM force fields and is easy to install using Conda or AmberTools. These advantages have contributed to the increasing use of gmx_mmpbsa for analyzing pepPI sampled by MD simulation in peptide–drug discovery studies [49–52].

Peptide–drug design

Following the identification of promising binding peptides, further refinements are needed to optimize their affinity, stability, and absorption, distribution, metabolism, excretion, and toxicity (ADMET) properties for effective and safe therapeutic development (Figure 2) [53].

Optimizing affinity

Rational design of a peptide drug requires reliable and efficient computational tools for analyzing the affinity of pepPIs. Computational alanine scanning (CAS) is a ΔG_{bind}-based method that systematically replaces amino acids with alanine to identify key residues critical for the affinity and stability of pepPIs. CAS is not only a great method to identify these 'hot spot' residues in pepPIs, it has also helped to elucidate the structure–activity relationship of different peptide drug candidates, by giving insights into which residues are essential for activity and which can be modified to enhance affinity, specificity, or stability [54–56]. Another frequently used method in computational peptide design is the per-residue ΔG_{bind} decomposition, which analyses the energetic contributions of individual amino acids to the overall ΔG_{bind} of a peptide–protein complex (Figure 3c). Weakly contributing residues can be modified, with the effects evaluated through subsequent ΔG_{bind} calculations after mutation. Using this method, researchers were able to propose peptide residue substitutions that led to improved binding affinity [57,58].

There are several tools for ΔG_{bind}-based analyses of pepPIs at the residues level. The BAlaS web application offers a fast and user-friendly platform for performing CAS, requiring only a structure file in PDB format as input [59]. For the analysis of MD trajectories using MM/PB(GB)SA calculations, tools such as CompASM [60], an Amber-based plug-in for Visual Molecular Dynamics (VMD) [61] for conducting CAS, and gmx_MMPBSA [48], which enables both per-residue ΔG_{bind} decomposition and CAS, have been developed. While these methods can identify residues with modification potential, they do not directly suggest specific mutations. In a peptide design approach based on modeling the SARS-CoV-2 main protease with its substrates, Chan et al. [62] utilized the computational saturation mutagenesis algorithm BUDE_SM [63] from the Bristol University Docking Engine (BUDE) [64] to substitute each peptide residue with the corresponding 19 alternative proteinogenic amino acids. *In vitro* inhibition assays confirmed that they successfully translated the BUDE_SM predictions into potent peptide inhibitors.

Structural stability

To enhance a peptide's binding affinity and selectivity, one approach is to maintain its target-bound conformation, which can be achieved through strategies like cyclization. It also improves its stability and pharmacokinetic properties, as cyclization protects peptides from enzymatic degradation and hydrolysis, and reduces its immunogenicity. Peptides can be linked to themselves in different ways, such as head to tail (Figure 1c), side chain to side chain, or head to side chain, using nonpeptidic linkers or groups to facilitate these connections [65,66]. Cyclization is often applied to peptides derived from so-called 'hot segments', an extension of the 'hot spots' concept that defines a continuous binding region essential for a PPI [5]. This was demonstrated by Lopez et al. who utilized MD simulations with enhanced sampling methods to assess the degree of preorganization—the structural conformation that most accurately mimics the 'hot loop'—in a series of cyclic peptides differing in a few point mutations and linker types [67]. They ultimately obtained an inhibitor with the highest degree of preorganization and a 280-fold improvement in binding affinity. In a recent *de novo* peptide design study, the computational generation of novel macrocyclic peptides without relying on preexisting sequences was realized using AI [68]. To this end, Rettie et al. developed RFpeptides, a deep learning framework using denoising diffusion models to design macrocyclic peptides with high affinity for selected protein targets.

Another popular method to stabilize secondary structure elements like helices is by using staples, which involve cross-linking two or more side chains to provide additional structural support (Figure 1d). The most common form of staples in peptides is the introduction of disulfide bonds. However, synthetic staples can also include various linkers or chemical groups that can react with amino acid side chains to stabilize the desired conformation. Like cyclic peptides, stapled peptides often possess further advantages like enhanced cell penetration and proteolytic stability [69,70].

Absorption, distribution, metabolism, excretion, and toxicity

In addition to high affinity and selectivity, the development of effective and safe drugs requires careful consideration of their ADMET properties [53]. Especially in the context of peptide-based therapeutics, proteolytic cleavage, rapid clearance, and low membrane permeability represent a major challenge [6,71]. To overcome these limitations, strategies like chemical modifications (e.g. acetylation/amidation and nonnatural or d-amino acids), conjugation strategies (e.g.

lipidation, PEGylation, and fusion proteins), and structural stabilization (e.g. stapling or cross-linking) have been developed [72]. Two web servers for ADMET prediction have recently been introduced enabling the assessment of key drug-like properties, such as solubility, bioavailability, blood–brain barrier penetration, and toxicity. ADMET-AI allows rapid screening of up to 1000 molecules (and one million molecules with a local version) at once, making it ideal for high-throughput evaluation [73]. For a more comprehensive approach, ADMETlab 3.0 offers a wider range of ADMET properties, detailed toxicity risk assessments, and an uncertainty estimate module to aid in the candidate selection process [74]. However, it should be noted that most ADMET prediction tools are trained on small molecules and therefore face challenges to accurately predict the pharmacokinetics of therapeutic peptides. Developing a dedicated platform to address these specific ADMET challenges for peptide drugs would be highly beneficial.

Conclusion and outlook

This review emphasizes some of the latest and most innovative methods for modeling peptide–protein interactions (pepPIs) and computer-aided peptide drug design. We addressed recent advancements for the peptide pose generation using docking and AI techniques, the evaluation of peptide flexibility using MD simulations, interactions at the binding site using binding free energy decompositions, and the rational design of new peptide drugs. We recommend employing a consensus approach that integrates physics-based and AI-based docking methods to increase the likelihood of obtaining accurate structural predictions. Enhanced sampling MD simulations can further improve structural reliability, while techniques like MM/PBSA for binding free energy decomposition and computational alanine scanning facilitate the peptide design process. Despite these advancements, a significant challenge remains: the development of a comprehensive workflow as illustrated in Figure 2 that consolidates the best available tools in one platform would greatly benefit the field. Looking ahead, the integration of AI for producing structures of peptide–protein complexes as well as for assessing the ADMET properties of peptides will play an increasingly vital role in peptide design, further enhancing our capabilities in drug development. The ultimate goal is to accurately predict pepPIs so that *in silico* results match the reliability of protein structure predictions with AlphaFold and yield binding energies that correlate with experimentally determined K_D values from binding assays. Achieving this would enable peptide–protein docking to replace experiments, which are challenged by the often transient pepPIs that produce noisy data.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We thank the VolkswagenStiftung for their financial support of this work (grant no. 99761).

Data availability

No data were used for the research described in the article.

References

Papers of particular interest, published within the period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Petsalaki E, Russell RB: **Peptide-mediated interactions in biological systems: new discoveries and applications**. *Curr Opin Biotechnol* 2008, **19**:344–350, https://doi.org/10.1016/j.copbio.2008.06.004.

2. Rossino G, Marchese E, Galli G, Verde F, Finizio M, Serra M, Linciano P, Collina S: **Peptides as therapeutic agents: challenges and opportunities in the green transition era**. *Molecules (Basel)* 2023, **28**:7165, https://doi.org/10.3390/molecules28207165.

3. *Global peptide therapeutics market size & outlook*. 2025 [cited Monday, 10. March 2025]. URL, https://www.grandviewresearch.com/horizon/outlook/peptide-therapeutics-market-size/global.

4. Wang L, Wang N, Zhang W, Cheng X, Yan Z, Shao G, Wang X, Wang R, Fu C: **Therapeutic peptides: current applications and future directions**. *Signal Transduct Targeted Ther* Feb. 2022, **7**, https://doi.org/10.1038/s41392-022-00904-4.

5. London N, Raveh B, Schueler-Furman O: **Druggable protein–protein interactions – from hot spots to hot segments**. *Curr Opin Chem Biol* 2013, **17**:952–959, https://doi.org/10.1016/j.cbpa.2013.10.011.

6. Muttenthaler M, King GF, Adams DJ, Alewood PF: **Trends in peptide drug discovery**. *Nat Rev Drug Discov* 2021, **20**:309–325, https://doi.org/10.1038/s41573-020-00135-8.

7. Sabe VT, Ntombela T, Jhamba LA, Maguire GE, Govender T, Naicker T, Kruger HG: **Current trends in computer aided drug design and a highlight of drugs discovered via computational techniques: a review**. *Eur J Med Chem* 2021, **224**, 113705, https://doi.org/10.1016/j.ejmech.2021.113705.

8. RCSB: ***RCSB protein Data Bank*** [cited Tuesday, 11. March 2025]. URL, https://www.rcsb.org.

9. Bertoline LMF, Lima AN, Krieger JE, Teixeira SK: **Before and after alphafold2: an overview of protein structure prediction**. *Front Bioinform* 2023, **3**, 1120370, https://doi.org/10.3389/fbinf.2023.1120370.

10. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K, Bates R, Žídek A, Potapenko A, *et al.*: **Highly accurate protein structure prediction with alphafold**. *Nature* 2021, **596**:583–589.

11. Abramson J, Adler J, Dunger J, Evans R, Green T, Pritzel A, Ronneberger O, Willmore L, Ballard AJ, Bambrick J, Bodenstein SW, Evans DA, Hung C-C, O'Neill M, Reiman D, Tunyasuvunakool K, Wu Z, Žemgulytė A, Arvaniti E, Beattie C, Bertolli O, Bridgland A, Cherepanov A, Congreve M, Cowen-Rivers AI, Cowie A, Figurnov M, Fuchs FB, Gladman H, Jain R, Khan YA, Low CMR, Perlin K, Potapenko A, Savy P, Singh S, Stecula A, Thillaisundaram A, Tong C, Yakneen S, Zhong ED, Zielinski M, Žídek A, Bapst V, Kohli P, Jaderberg M, Hassabis D, Jumper JM: **Accurate structure prediction of biomolecular interactions with alphafold 3**. *Nature* 2024, **630**:493–500, https://doi.org/10.1038/s41586-024-07487-w.

12. Mondal A, Chang L, Perez A: **Modelling peptide–protein complexes: docking, simulations and machine learning**. *QRB Discov* 2022, **3**:e17, https://doi.org/10.1017/qrd.2022.14.

13. Wang X, Ni D, Liu Y, Lu S: **Rational design of peptide-based inhibitors disrupting protein-protein interactions**. *Front Chem* May 2021, **9**, https://doi.org/10.3389/fchem.2021.682675.

14. Tian W, Chen C, Lei X, Zhao J, Liang J: **Castp 3.0: computed atlas of surface topography of proteins**. *Nucleic Acids Res* 2018, **46**:W363–W367, https://doi.org/10.1093/nar/gky473.

15. Le Guilloux V, Schmidtke P, Tuffery P: **Fpocket: an open source platform for ligand pocket detection**. *BMC Bioinf* 2009, **10**:168, https://doi.org/10.1186/1471-2105-10-168.

16. Trabuco LG, Lise S, Petsalaki E, Russell RB: **Pepsite: prediction of peptide-binding sites from protein surfaces**. *Nucleic Acids Res* 2012, **40**:W423–W427, https://doi.org/10.1093/nar/gks398.

17. •• Gagliardi L, Rocchia W: **Siteferret: beyond simple pocket identification in proteins**. *J Chem Theor Comput* 2023, **19**: 5242–5259, https://doi.org/10.1021/acs.jctc.2c01306.
SiteFerret is a powerful and versatile tool for identifying all types of binding pockets on protein surfaces, including those for peptides. Using AI techniques, it accurately predicts and ranks binding sites, offering a reliable method for subsequent local docking.

18. Agrawal P, Singh H, Srivastava HK, Singh S, Kishore G, Raghava GPS: **Benchmarking of different molecular docking methods for protein-peptide docking**. *BMC Bioinf* Feb. 2019, **19**, https://doi.org/10.1186/s12859-018-2449-y.

19. Ramírez-Aportela E, López-Blanco JR, Chacón P: **Frodock 2.0: fast protein–protein docking server**. *Bioinformatics* 2016, **32**: 2386–2388, https://doi.org/10.1093/bioinformatics/btw141.

20. Weng G, Gao J, Wang Z, Wang E, Hu X, Yao X, Cao D, Hou T: **Comprehensive evaluation of fourteen docking programs on protein–peptide complexes**. *J Chem Theor Comput* 2020, **16**: 3959–3969, https://doi.org/10.1021/acs.jctc.9b01208.

21. Zhou P, Jin B, Li H, Huang S-Y: **HPEPDOCK: a web server for blind peptide–protein docking based on a hierarchical algorithm**. *Nucleic Acids Res* 2018, **46**:W443–W450, https://doi.org/10.1093/nar/gky357.

22. Zhang Y, Sanner MF: **AutoDock CrankPep: combining folding and docking to predict protein-peptide complexes**. *Bioinformatics* 2019, **35**:5121–5127, https://doi.org/10.1093/bioinformatics/btz459.

23. Podtelezhnikov AA, Wild DL: **CRANKITE: a fast polypeptide backbone conformation sampler**. *Source Code Biol Med* 2008, **3**:12, https://doi.org/10.1186/1751-0473-3-12.

24. Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, Goodsell DS, Olson AJ: **AutoDock4 and AutoDockTools4: automated docking with selective receptor flexibility**. *J Comput Chem* 2009, **30**:2785–2791, https://doi.org/10.1002/jcc.21256.

25. Zhang Y, Sanner MF: **Docking flexible cyclic peptides with AutoDock CrankPep**. *J Chem Theor Comput* 2019, **15**: 5161–5168, https://doi.org/10.1021/acs.jctc.9b00557.

26. Trellet M, Melquiond AS, Bonvin AM: **A unified conformational selection and induced fit approach to protein-peptide docking**. *PLoS One* 2013, **8**, e58769.

27. Kurcinski M, Jamroz M, Blaszczyk M, Kolinski A, Kmiecik S: **Cabs-dock web server for the flexible docking of peptides to proteins without prior knowledge of the binding site**. *Nucleic Acids Res* 2015, **43**:W419–W424.

28. London N, Raveh B, Cohen E, Fathi G, Schueler-Furman O: **Rosetta FlexPepDock web server—high resolution modeling of peptide–protein interactions**. *Nucleic Acids Res* 2011, **39**: W249–W253, https://doi.org/10.1093/nar/gkr431.

29. Evans R, O'Neill M, Pritzel A, Antropova N, Senior A, Green T, Žídek A, Bates R, Blackwell S, Yim J, Ronneberger O, Bodenstein S, Zielinski M, Bridgland A, Potapenko A, Cowie A, Tunyasuvunakool K, Jain R, Clancy E, Kohli P, Jumper J, Hassabis D: **Protein complex prediction with alphafold-multimer**. *bioRxiv* 2022, https://doi.org/10.1101/2021.10.04.463034.

30. Johansson-Åkhe I, Wallner B: **Improving peptide-protein docking with AlphaFold-Multimer using forced sampling**. *Front Bioinform* 2022, **2**, 959160, https://doi.org/10.3389/fbinf.2022.959160.

31. •• Shanker S, Sanner MF: **Predicting protein–peptide interactions: benchmarking deep learning techniques and a comparison with focused docking**. *J Chem Inf Model* 2023, **63**: 3158–3170, https://doi.org/10.1021/acs.jcim.3c00602.
This study evaluates the efficacy of deep learning models, including AlphaFold2, in predicting pepPIs and compares their performance to the docking method AutoDock CrankPep. AlphaFold2 achieves a docking success rate of 53%, while AutoDock CrankPep samples correct solutions for about 62% of complexes when considering all generated solutions. Notably, a consensus approach combining both methods yields a 60% success rate for the top-ranking results, providing valuable insights for advancing computational strategies in modeling pepPIs.

32. Manshour N, Ren J, Esmaili F, Bergstrom E, Xu D: **Comprehensive evaluation of alphafold-multimer, alphafold3 and colabfold, and scoring functions in predicting protein-peptide complex structures**. *bioRxiv* 2024, https://doi.org/10.1101/2024.11.11.622992.

33. Lin Z, Akin H, Rao R, Hie B, Zhu Z, Lu W, Smetanin N, Verkuil R, Kabeli O, Shmueli Y, dos Santos Costa A, Fazel-Zarandi M, Sercu T, Candido S, Rives A: **Evolutionary-scale prediction of atomic-level protein structure with a language model**. *Sci Technol Humanit* 2023, **379**:1123–1130, https://doi.org/10.1126/science.ade2574.

34. • Zalewski M, Wallner B, Kmiecik S: **Protein–peptide docking with ESMFold language model**. *J Chem Theor Comput* Mar. 2025, https://doi.org/10.1021/acs.jctc.4c01585.
In this study, the ESMFold language model for peptide–protein docking is compared to state of the art docking programs like AlphaFold 3. While ESMFold's accuracy is generally lower than that of AlphaFold, it offers excellent computational efficiency. This makes ESMFold a valuable tool for high-throughput pose prediction studies and a useful complement to other methods in a consensus approach.

35. Rico MI, Lebedenko CG, Mitchell SM, Banerjee IA: **Molecular dynamics simulations, docking and mmgbsa studies of newly designed peptide-conjugated glucosyloxy stilbene derivatives with tumor cell receptors**. *Mol Divers* 2022, **26**: 2717–2743, https://doi.org/10.1007/s11030-021-10354-9.

36. Jin J, Wang N, Wang S, Alsaikhan F, Ahmad I: **Targeting the main sars-cov-2 pathways with peptide inhibitors by molecular docking and molecular simulation approaches**. *J Mol Liq* 2024, **398**, 124193, https://doi.org/10.1016/j.molliq.2024.124193.

37. Miao Q, Zhang W, Zhang K, Li H, Zhu J, Jiang S: **Rational design of a potent macrocyclic peptide inhibitor targeting the pd-1/pd-l1 protein–protein interaction**. *RSC Adv* 2021, **11**: 23270–23279, https://doi.org/10.1039/d1ra03118j.

38. Siebenmorgen T, Zacharias M: **Efficient refinement and free energy scoring of predicted protein–protein complexes using replica exchange with repulsive scaling**. *J Chem Inf Model* 2020, **60**:5552–5562, https://doi.org/10.1021/acs.jcim.0c00853.

39. Wang J, Alekseenko A, Kozakov D, Miao Y: **Improved modeling of peptide-protein binding through global docking and accelerated molecular dynamics simulations**. *Front Mol Biosci* 2019, **6**, https://doi.org/10.3389/fmolb.2019.00112.

40. Wang J, Miao Y: **Peptide Gaussian accelerated molecular dynamics (Pep-GaMD): enhanced sampling and free energy and kinetics calculations of peptide binding**. *J Chem Phys* 2020, **153**, 154109, https://doi.org/10.1063/5.0021399.

41. •• Wang J, Koirala K, Do HN, Miao Y: **PepBinding: a workflow for predicting peptide binding structures by combining peptide docking and peptide Gaussian accelerated molecular dynamics simulations**. *J Phys Chem B* 2024, **128**:7332–7340, https://doi.org/10.1021/acs.jpcb.4c02047.
PepBinding is a workflow that combines peptide docking and Gaussian accelerated molecular dynamics (Pep-GaMD) simulations to predict peptide binding structures. The authors showed that while peptide docking alone provided medium-to-poor quality models, the addition of Pep-GaMD significantly improved the results, offering a more accurate prediction of pepPIs. This approach serves as a valuable component of an efficient peptide drug discovery workflow.

42. Malmstrom RD, Watowich SJ: **Using free energy of binding calculations to improve the accuracy of virtual screening predictions**. *J Chem Inf Model* 2011, **51**:1648–1655, https://doi.org/10.1021/ci200126v.

43. Wang S, Sun X, Cui W, Yuan S: **MM/PB(GB)SA benchmarks on soluble proteins and membrane proteins**. *Front Pharmacol* 2022, **13**, 1018351, https://doi.org/10.3389/fphar.2022.1018351.

44. Weng G, Wang E, Chen F, Sun H, Wang Z, Hou T: **Assessing the performance of mm/pbsa and mm/gbsa methods. 9. prediction reliability of binding affinities and binding poses for protein–peptide complexes**. *Phys Chem Chem Phys* 2019, **21**: 10135–10145, https://doi.org/10.1039/c9cp01674k.

45. Wang S, Sun X, Cui W, Yuan S: **MM/PB(GB)SA benchmarks on soluble proteins and membrane proteins**. *Front Pharmacol* 2022, **13**, https://doi.org/10.3389/fphar.2022.1018351.

46. Miller BR, McGee TD, Swails JM, Homeyer N, Gohlke H, Roitberg AE: **MMPBSA.py: an efficient program for end-state free energy calculations**. *J Chem Theor Comput* 2012, **8**: 3314–3321, https://doi.org/10.1021/ct300418h.

47. Abraham MJ, Murtola T, Schulz R, Páll S, Smith JC, Hess B, Lindahl E: **Gromacs: high performance molecular simulations through multi-level parallelism from laptops to supercomputers**. *SoftwareX* 2015, **1**:19–25, https://doi.org/10.1016/j.softx.2015.06.001.

48. Valdés-Tresanco MS, Valdés-Tresanco ME, Valiente PA, Moreno E: **gmx_MMPBSA: a new tool to perform end-state free energy calculations with GROMACS**. *J Chem Theor Comput* 2021, **17**:6281–6291, https://doi.org/10.1021/acs.jctc.1c00645.

49. Chen S, Lin T, Basu R, Ritchey J, Wang S, Luo Y, Li X, Pei D, Kara LB, Cheng X: **Design of target specific peptide inhibitors using generative deep learning and molecular dynamics simulations**. *Nat Commun* 2024, **15**:1611, https://doi.org/10.1038/s41467-024-45766-2.

50. • Wu Y, Deng S, Wang X, Thunders M, Qiu J, *et al.*: **Discovery and mechanism of action of a novel antimicrobial peptide from an earthworm**. *Microbiol Spectr* 2023, **11**:e03206–e03222, https://doi.org/10.1128/spectrum.03206-22.

In this comprehensive approach, extensive wet-laboratory experimental techniques and computational modeling methods are combined to identify and characterize a novel antimicrobial peptide (AMP) from earthworms. This combination enhances the understanding of AMP mechanisms and supports the design of new peptide-based antibiotics.

51. Niu P, Tao Y, Meng Q, Huang Y, Li S, Ding K, Ma D, Ye Z, Fan M: **Discovery of novel macrocyclic derivatives as potent and selective cyclin-dependent kinase 2 inhibitors**. *Bioorg Med Chem* 2024, **104**, 117711, https://doi.org/10.1016/j.bmc.2024.117711.

52. Eberle RJ, Sevenich M, Gering I, Scharbert L, Strodel B, Lakomek NA, Santur K, Mohrlüder J, Coronado MA, Willbold D: **Discovery of all-D-peptide inhibitors of SARS-CoV-2 3C-like protease**. *ACS Chem Biol* 2023, **18**:315–330, https://doi.org/10.1021/acschembio.2c00735.

53. Ferreira LL, Andricopulo AD: **ADMET modeling approaches in drug discovery**. *Drug Discov Today* 2019, **24**:1157–1165, https://doi.org/10.1016/j.drudis.2019.03.015.

54. Valiente PA, Wen H, Nim S, Lee J, Kim HJ, Kim J, Perez-Riba A, Paudel YP, Hwang I, Kim K-D, Kim S, Kim PM: **Computational design of potent D-peptide inhibitors of SARS-CoV-2**. *J Med Chem* 2021, **64**:14955–14967, https://doi.org/10.1021/acs.jmedchem.1c00655.

55. Amarasinghe KN, De Maria L, Tyrchan C, Eriksson LA, Sadowski J, Petrović D: **Virtual screening expands the non-natural amino acid palette for peptide optimization**. *J Chem Inf Model* 2022, **62**:2999–3007, https://doi.org/10.1021/acs.jcim.2c00193.

56. Shahab M, Khan SS, Zulfat M, Bin Jardan YA, Mekonnen AB, Bourhia M, Zheng G: **In silico mutagenesis-based designing of oncogenic SHP2 peptide to inhibit cancer progression**. *Sci Rep* 2023, **13**, https://doi.org/10.1038/s41598-023-37020-4.

57. Khan A, Umbreen S, Hameed A, Fatima R, Zahoor U, Babar Z, Waseem M, Hussain Z, Rizwan M, Zaman N, Ali S, Suleman M, Shah A, Ali L, Ali SS, Wei D-Q: **In silico mutagenesis-based remodelling of SARS-CoV-1 peptide (ATLQAIAS) to inhibit SARS-CoV-2: structural-dynamics and free energy calculations**. *Interdiscipl Sci Comput Life Sci* 2021, **13**:521–534, https://doi.org/10.1007/s12539-021-00447-2.

58. Ghobadi Z, Mahnam K, Shakhsi-Niaei M: **In-silico design of peptides for inhibition of HLA-A*03-KLIETYFSK complex as a new drug design for treatment of multiples sclerosis disease**. *J Mol Graph Model* 2022, **111**, 108079, https://doi.org/10.1016/j.jmgm.2021.108079.

59. Wood CW, Ibarra AA, Bartlett GJ, Wilson AJ, Woolfson DN, Sessions RB: **BAlaS: fast, interactive and accessible computational alanine-scanning using BudeAlaScan**. *Bioinformatics* 2020, **36**:2917–2919, https://doi.org/10.1093/bioinformatics/btaa026.

60. Ribeiro JV, Cerqueira NM, Moreira IS, Fernandes PA, Ramos MJ: **CompASM: an Amber-VMD alanine scanning mutagenesis plug**. *Theor Chem Acc* 2012, **131**:1271, https://doi.org/10.1007/s00214-012-1271-2.

61. Humphrey W, Dalke A, Schulten K: **VMD: visual molecular dynamics**. *J Mol Graph* 1996, **14**:33–38, https://doi.org/10.1016/0263-7855(96)00018-5.

62. Chan HTH, Moesser MA, Walters RK, Malla TR, Twidale RM, John T, Deeks HM, Johnston-Wood T, Mikhailov V, Sessions RB, Dawson W, Salah E, Lukacik P, Strain-Damerell C, Owen CD, Nakajima T, Świderek K, Lodola A, Moliner V, Glowacki DR, Spencer J, Walsh MA, Schofield CJ, Genovese L, Shoemark DK, Mulholland AJ, Duarte F, Morris GM: **Discovery of SARS-CoV-2 Mpro peptide inhibitors from modelling substrate and ligand binding**. *Chem Sci* 2021, **12**:13686–13703, https://doi.org/10.1039/d1sc03628a.

63. Hetherington K, Dutt S, Ibarra AA, Cawood EE, Hobor F, Woolfson DN, Edwards TA, Nelson A, Sessions RB, Wilson AJ: **Towards optimizing peptide-based inhibitors of protein–protein interactions: predictive saturation variation scanning (PreSaVS)**. *RSC Chem Biol* 2021, **2**:1474–1478, https://doi.org/10.1039/d1cb00137j.

64. McIntosh-Smith S, Price J, Sessions RB, Ibarra AA: **High performance in silico virtual drug screening on many-core processors**. *Int J High Perform Comput Appl* 2014, **29**:119–134, https://doi.org/10.1177/1094342014528252.

65. Vu QN, Young R, Sudhakar HK, Gao T, Huang T, Tan YS, Lau YH: **Cyclisation strategies for stabilising peptides with irregular conformations**. *RSC Med Chem* 2021, **12**:887–901, https://doi.org/10.1039/d1md00098e.

66. Marsault E, Peterson ML: **Macrocycles are great cycles: applications, opportunities, and challenges of synthetic macrocycles in drug discovery**. *J Med Chem* 2011, **54**: 1961–2004, https://doi.org/10.1021/jm1012374.

67. • Fonseca Lopez F, Miao J, Damjanovic J, Bischof L, Braun MB, *et al.*: **Computational prediction of cyclic peptide structural ensembles and application to the design of keap1 binders**. *J Chem Inf Model* 2023, **63**:6925–6937, https://doi.org/10.1021/acs.jcim.3c01337.

This MD simulation study explored the structural conformations of cyclic peptides. Aiming to mimic the binding interface of a PPI, the authors designed peptides that effectively inhibit Keap1, a protein involved in the cellular stress response, providing valuable insights into the development of potential cyclic peptide-based drugs.

68. Rettie SA, Juergens D, Adebomi V, Bueso YF, Zhao Q, Leveille AN, Liu A, Bera AK, Wilms JA, Üffing A, Kang A, Brackenbrough E, Lamb M, Gerben SR, Murray A, Levine PM, Schneider M, Vasireddy V, Ovchinnikov S, Weiergräber OH, Willbold D, Kritzer JA, Mougous JD, Baker D, DiMaio F, Bhardwaj G: **Accuratede novodesign of high-affinity protein binding macrocycles using deep learning**. *bioRxiv* 2024, https://doi.org/10.1101/2024.11.18.622547.

69. Bluntzer MTJ, O'Connell J, Baker TS, Michel J, Hulme AN: **Designing stapled peptides to inhibit protein-protein interactions: an analysis of successes in a rapidly changing field**. *Pept Sci* 2021, **113**(1), e24191, https://doi.org/10.1002/pep2.24191.

70. Li Y, Wu M, Fu Y, Xue J, Yuan F, Qu T, Rissanou AN, Wang Y, Li X, Hu H: **Therapeutic stapled peptides: efficacy and molecular targets**. *Pharmacol Res* 2024, **203**, 107137, https://doi.org/10.1016/j.phrs.2024.107137.

71. Barman P, Joshi S, Sharma S, Preet S, Sharma S, Saini A: **Strategic approaches to improvise peptide drugs as next generation therapeutics**. *Int J Pept Res Therapeut* 2023, **29**(4): 61, https://doi.org/10.1007/s10989-023-10524-3.

72. Akbarian M, Chen S-H: **Instability challenges and stabilization strategies of pharmaceutical proteins**. *Pharmaceutics* 2022, **14**:2533, https://doi.org/10.3390/pharmaceutics14112533.

73. Swanson K, Walther P, Leitz J, Mukherjee S, Wu JC, Shivnaraine RV, Zou J: **ADMET-AI: a machine learning ADMET platform for evaluation of large-scale chemical libraries**. *Bioinformatics* 2024, **40**, btae416, https://doi.org/10.1093/bioinformatics/btae416.

74. Fu L, Shi S, Yi J, Wang N, He Y, Wu Z, Peng J, Deng Y, Wang W, Wu C, Lyu A, Zeng X, Zhao W, Hou T, Cao D: **ADMETlab 3.0: an updated comprehensive online ADMET prediction platform enhanced with broader coverage, improved performance, API functionality and decision support**. *Nucleic Acids Res* 2024, **52**:W422–W431, https://doi.org/10.1093/nar/gkae236.