

The evolution of computation-driven paradigms in targeted peptide drug design: From predictive modeling to generative AI and clinical translation

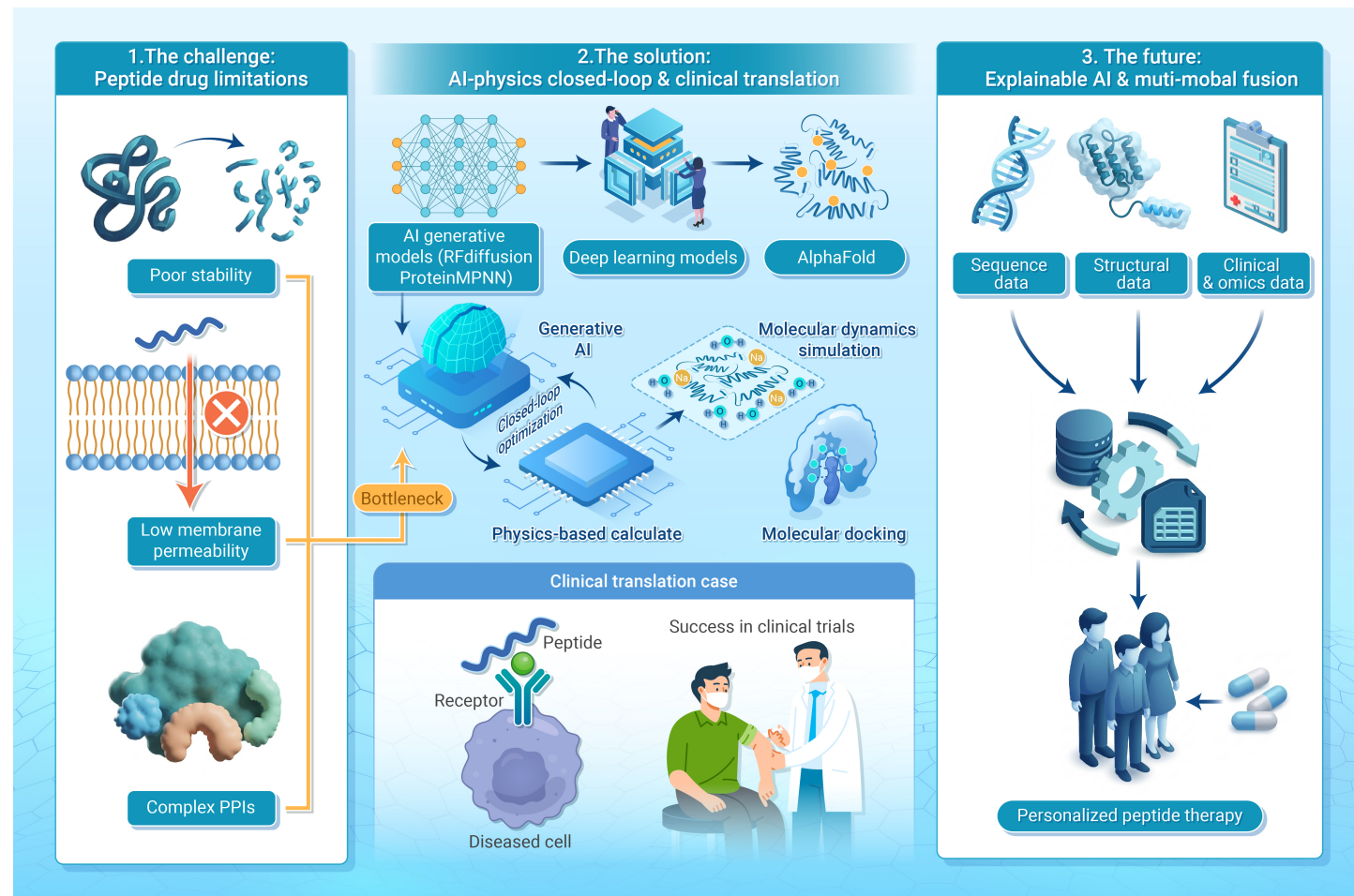
Wenjing Hu,^{1,4} Yuan Sun,^{1,4} Ting Li,¹ Mark Williamson,² Xiaoying Hu,^{3,*} and Maolin Wang^{1,*}

*Correspondence: huxiaoying1127@163.com (X.L.); wangml1240@163.com (M.W.)

Received: January 20, 2026; Accepted: March 13, 2026; Published Online: April 1, 2026; <https://doi.org/10.59717/j.xinn-drugdisc.2026.100009>

© 2026 The Author(s). This is an open access article under the CC BY license (<https://creativecommons.org/licenses/by/4.0/>).

GRAPHICAL ABSTRACT



PUBLIC SUMMARY

- Targeted peptides tackle undruggable targets but face in vivo instability.
- Peptide design shifts from empirical screening to artificial intelligence (AI).
- Generative AI creates novel sequences, bypassing limits of natural databases.
- Integrating AI with physics validation accelerates clinical peptide development.

The evolution of computation-driven paradigms in targeted peptide drug design: From predictive modeling to generative AI and clinical translation

Wenjing Hu,^{1,4} Yuan Sun,^{1,4} Ting Li,¹ Mark Williamson,² Xiaoying Hu,^{3,*} and Maolin Wang^{1,*}

¹Clinical Research Center, The First Affiliated Hospital of Shantou University Medical College, Shantou 545041, China

²Department of Chemistry, University of Cambridge, Cambridge, CB2 1EW, UK

³School of Medicine, Weinan Vocational and Technical College, Weinan 714026, China

⁴These authors contributed equally

*Correspondence: huxiaoying1127@163.com (X.L.); wangml1240@163.com (M.W.)

Received: January 20, 2026; Accepted: March 13, 2026; Published Online: April 1, 2026; <https://doi.org/10.59717/j.xinn-drugdisc.2026.100009>

© 2026 The Author(s). This is an open access article under the CC BY license (<https://creativecommons.org/licenses/by/4.0/>).

Citation: Hu W., Sun Y., Li T., et al. (2026). The evolution of computation-driven paradigms in targeted peptide drug design: From predictive modeling to generative AI and clinical translation. *The Innovation Drug Discovery* 1:100009.

Targeted peptide therapeutics offer a potent solution for undruggable intracellular targets yet their clinical translation remains hampered by poor membrane permeability and metabolic instability. The integration of high-performance computing and artificial intelligence is currently driving a fundamental transition from empirical screening to rational *de novo* design. This review moves beyond a conventional enumeration of tools to construct a strategic framework that integrates physics-based validation with generative deep learning. We critically analyze the synergistic application of molecular dynamics and docking for thermodynamic verification while simultaneously evaluating how diffusion models and protein language models accelerate the exploration of vast chemical spaces. By delineating a closed-loop workflow that incorporates pharmacokinetic constraints into generative algorithms this review not only synthesizes current advancements but also provides a strategic roadmap for seamlessly integrating generative AI with physics-based validation, thereby accelerating the transition of computationally designed peptides from in silico blueprints to viable clinical candidates.

INTRODUCTION

Peptides occupy a unique therapeutic niche, effectively bridging the gap between traditional small-molecule drugs and complex biologics. Chemically defined as chains of 2 to 50 amino acid residues, peptides possess distinct advantages over small molecules due to their physical interaction capabilities. While small molecules are typically rigid and require deep cavities for effective binding, peptides exhibit high conformational flexibility and a larger interfacial surface area (400–1000 Å²). This characteristic enables peptides to selectively bind to flat, broad biological interfaces that lack the deep pockets required by small molecules, thereby addressing targets previously deemed “undruggable”.^{1–4} For example, peptides have shown superior performance over conventional small molecules in targeting intracellular protein-protein interaction (PPI) networks.⁵ Furthermore, peptide therapeutics often offer benefits such as enhanced tissue penetration, potentially lower immunogenicity, and reduced manufacturing costs compared to monoclonal antibodies,^{2,3,6,7} although realizing these advantages depends on specific sequence modifications and formulation considerations. The clinical utility of peptides is exemplified by the success of GLP-1 receptor agonists, such as Liraglutide and Semaglutide. By revolutionizing treatments for metabolic and cardiovascular diseases, these therapies have firmly established peptides as a cornerstone of the modern pharmaceutical industry.^{8–11}

Although peptides offer exciting possibilities for treatment, their clinical application is constrained by distinct pharmacokinetic barriers (Figure 1).¹² The core issue lies in their flexible and changing molecular structure, which makes them highly vulnerable to destruction inside the body. Once introduced into the bloodstream, these molecules are quickly targeted and broken down by enzymes such as DPP-IV. This rapid neutralization significantly shortens their active lifespan, making it difficult to maintain effective therapeutic levels in patients.^{13,14} Furthermore, peptides face a rapid exit from the body through the kidneys. Their relatively small size allows them to be filtered from the bloodstream with high efficiency, leading to extremely brief func-

tional lifespans—often measuring just a few minutes. This quick elimination creates a major therapeutic hurdle, as it becomes difficult to sustain the steady drug levels required to effectively treat the patient.^{15,16} Peptides’ inherent hydrophilicity and large molecular size further restrict their ability to cross biological membranes, creating substantial bioavailability challenges for oral formulations.^{3,17,18} To address these drugability obstacles, traditional medicinal chemistry strategies have predominantly relied on empirical modifications, such as introducing conformational constraints through backbone cyclization, preventing enzymatic cleavage by incorporating non-natural amino acids, or extending circulation time via PEGylation.^{2,19–22} However, these trial-and-error approaches are not only time-consuming and labor-intensive but also struggle to efficiently optimize the delicate balance between peptide activity, stability, and membrane permeability within the vast sequence space.²³

With the rapid advancement of high-performance computing and the exponential accumulation of structural biology data, the drug discovery paradigm is undergoing a profound transformation driven by computational science.²⁴ Molecular modeling techniques have evolved from serving as mere experimental aids to becoming the core engine of rational drug design. We are witnessing a shift from random screening to rational design, grounded in physicochemical principles and powered by artificial intelligence (AI) technologies.^{25–27} Homology modeling and molecular docking enable the prediction of peptide-protein binding interactions at the atomic level, while molecular dynamics (MD) simulations introduce the temporal dimension, revealing the dynamic folding pathways and conformational ensembles of peptides in physiological environments. AI-driven technologies, such as AlphaFold and Protein Language Models (PLMs), have overcome the limitations of traditional methods in predicting flexible peptide structures.¹⁴ The deep integration of these cutting-edge technologies has not only significantly accelerated the discovery of lead compounds but also provided novel approaches to addressing challenges such as peptide conformational instability and transmembrane delivery. This paper systematically traces the evolutionary trajectory of these techniques, with a focus on how computational strategies are empowering the end-to-end design of targeted peptide therapeutics (Figure 2).

MAINSTREAM MOLECULAR MODELING METHODS IN TARGETED PEPTIDE DISCOVERY

The field of peptide drug discovery has evolved from a reliance on traditional trial-and-error screening to a systematic, rational design process (Figure 3).^{25,28–30} This transformation is powered by three foundational modeling techniques: homology modeling, molecular docking, and molecular dynamics (MD) simulations. As highlighted by Mondal et al. and Geng et al., leveraging these computational tools allows the field to move beyond simple guesswork and visualize peptide behavior at the atomic level. These methods allow scientists to accurately predict how a peptide will fold, assess its stability in real-time, and calculate exactly how tightly it will bind to a target protein.^{30,31} Ultimately, this structural insight enables the precise engineering of lead sequences, ensuring high-affinity targeting before physical testing even begins.

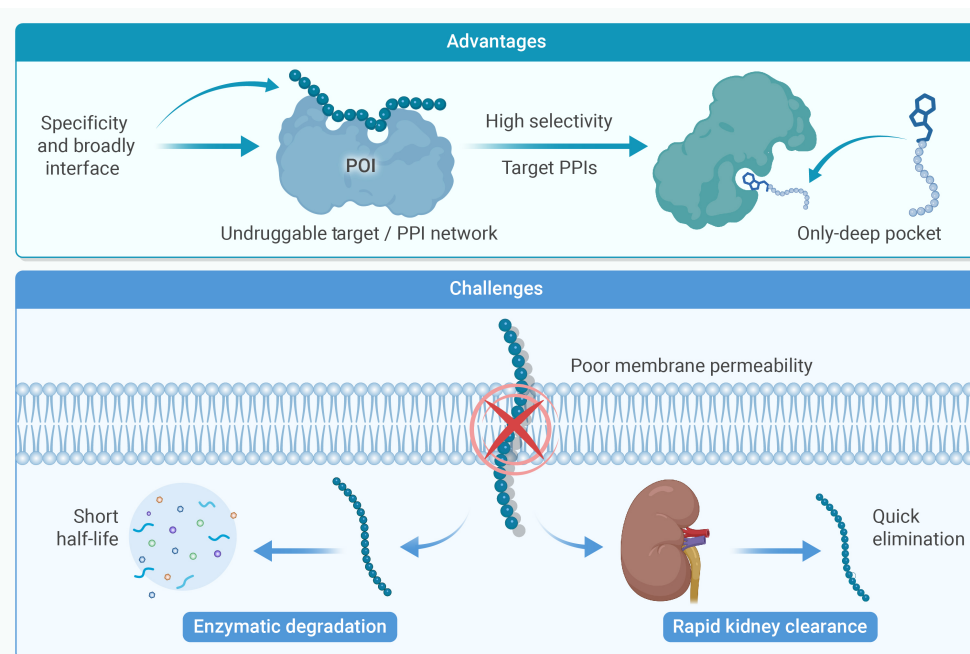


Figure 1. Overview of Peptide Therapeutics This figure summarizes peptides' characteristics as intermediates between small molecules and biologics; their advantages in targeting hard-to-drug interfaces and high tissue penetration; and their pharmacokinetic challenges including rapid enzymatic degradation, renal clearance, and poor membrane permeability.

vastly expanded conformational space due to the inherent flexibility of both their main and side chains, imposing significant demands on the sampling capabilities of docking algorithms.⁴⁸ To accurately reflect interactions in biological systems, modern methodologies have transcended static rigid-body assumptions to incorporate multi-conformation sampling and multi-pose binding concepts. This shift acknowledges that binding is a dynamic equilibrium process governed by multiple energetically favorable conformations rather than a singular structural event.⁴⁹⁻⁵¹

Specific docking strategies encompass template-based, local, and global search protocols

whose applicability depends on the availability of prior structural information.⁵²⁻⁵⁴ Traditional small-molecule docking programs often exhibit limitations when handling longer peptide segments as they fail to effectively capture backbone rearrangements and induced-fit effects.⁴⁹ This deficiency has driven the development of specialized algorithms explicitly optimized to account for peptide chain flexibility and receptor side-chain adjustments.^{50,51} Numerous studies demonstrate that these advanced tools not only explain experimentally observed inhibitory activity and selectivity but also identify key hydrogen-bond networks and hydrophobic anchors, thereby providing explicit structural guidance for subsequent sequence modification and conformation-constrained design (Table S1).⁵⁵

The selection of molecular docking algorithms fundamentally reflects a trade-off between sampling exhaustiveness and atomic resolution (Figure 4). Global docking tools generally prioritize exploring the entire protein surface to locate potential binding sites but often resort to coarse-grained models that may overlook subtle induced-fit changes.^{50,56,65} Conversely, local refinement tools provide high-resolution atomic details but are prone to entrapment in local energy minima if the initial pose deviates significantly from the native state.⁶⁰ Addressing this dichotomy, Ciemny et al. advocate for a hierarchical docking strategy wherein coarse-grained global algorithms are initially employed to capture binding hotspots, followed by high-precision local refinement to optimize interfacial atomic contacts.⁵⁴ Furthermore, for unique topologies such as cyclic peptides, standard algorithms often fail to handle backbone constraints effectively, necessitating specialized tools like AutoDock CrankPep.⁵⁸ Tucker and colleagues exemplified the translational impact of such advanced sampling methodologies through the discovery of the oral PCSK9 inhibitor MK-0616, where computational analysis guided the rigidification of a macrocyclic scaffold to distinctively combine high potency with intestinal permeability.⁷⁵

The core challenges in peptide-protein docking primarily involve handling flexibility, adapting scoring functions, and dealing with the uncertainty of binding sites.^{50,54,74,76-78} Peptides, which typically lack stable conformations in their free state, often undergo substantial rearrangements of their main chains or side chains during binding. Similarly, protein receptors may experience induced conformational changes. This bidirectional flexibility significantly complicates conformation sampling, rendering traditional rigid-body docking methods inadequate for accurately capturing key interaction patterns.⁷⁹⁻⁸¹ To address this, some approaches incorporate induced-fit or conformation-selection models, which, when combined with receptor conformation ensembles, enhance the biological plausibility of the predictions.⁸² Representing a paradigm shift in handling this conformational complexity, emerging AI-native diffusion frameworks such as RAPiDock⁸³ have demon-

Homology modeling

Homology modeling exploits the fundamental evolutionary tenet that proteins or peptides sharing sequence similarity generally fold into comparable three-dimensional structures. Muhammed and Aki-Yalcin explicitly note that this correlation allows for the prediction of unknown target architectures using established templates, rendering the technique indispensable for visualizing targets where experimental structural data remain elusive.³² In the specific context of targeted peptide discovery, this approach serves not only to complement target protein structures but also to reconstruct bioactive peptide conformations derived from natural ligands or functional fragments.^{33,34} The conventional workflow encompasses template selection, sequence alignment, and model refinement,^{32,35-37} where predictive accuracy is intrinsically linked to template quality and sequence identity.³⁸ Specifically, when the sequence identity between the target and template exceeds 30%, the reliability of backbone conformations and structural motifs improves significantly.^{39,40} Beyond static structural prediction, current methodologies have evolved to elucidate how structural features drive biological function. Anighoro posits that integrating these computational models with experimental data facilitates the identification of specific structural attributes governing activity, thereby guiding the rational optimization of peptide sequences.⁴¹ However, challenges persist in modeling highly flexible loop regions or disordered domains^{37,42} where the nonlinear relationship between sequence divergence and structural deviation necessitates rigorous model validation and multi-model comparison.⁴³

Furthermore, homology modeling has established itself as a critical strategy for modulating protein-protein interactions (PPIs) by addressing targets previously deemed undruggable. Grounded in the competitive peptide design logic,⁴⁴ this approach models the complex of a target protein with its natural binding partner to extract interfacial peptide segments as structural templates. By systematically characterizing the binding interface to identify hot spot residues contributing disproportionately to binding free energy, this integrated workflow enables the precise mapping of flat biological interfaces. A prominent application of this utility is observed in the design of inhibitors for the MDM2/MDMX and BCL-2 families where computational modeling successfully guided the transformation of natural fragments into high-affinity inhibitors through the optimization of key interfacial interactions.⁴⁵

Molecular docking

Molecular docking functions as a pivotal mechanism for deciphering peptide-protein recognition patterns by simulating binding conformations and quantifying interaction strengths to provide structural insight for screening and optimization.^{30,46,47} Unlike small-molecule ligands, peptides possess a

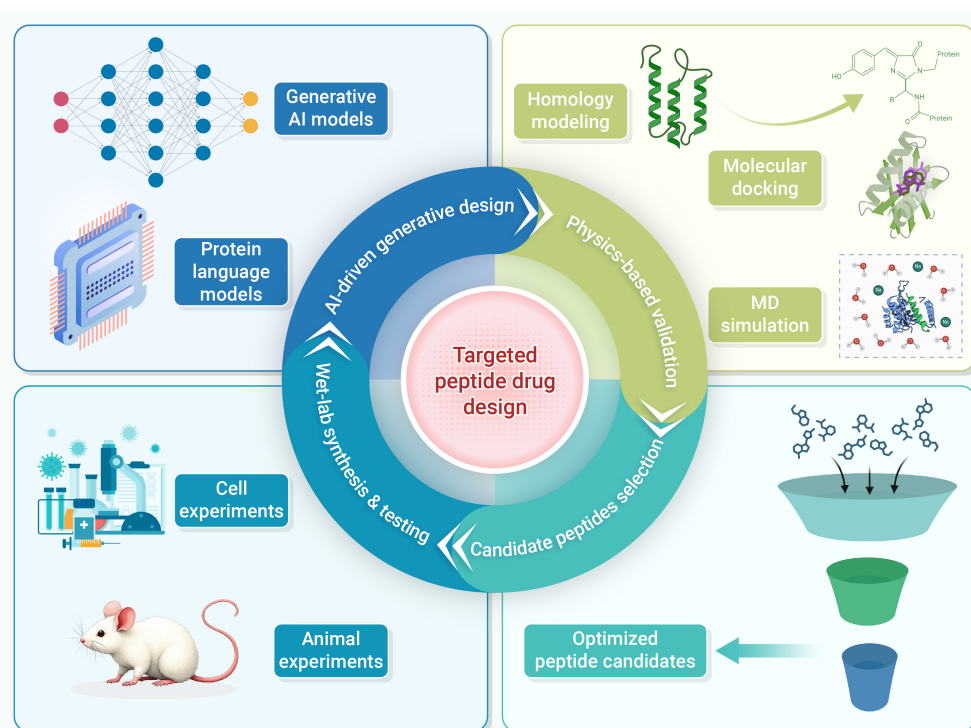


Figure 2. Integrated AI and physics-based workflow for targeted peptide drug design This closed-loop synergizes generative AI models, physics-based validation, and wet-lab experimental feedback to iteratively refine candidates and accelerate therapeutic discovery.

Molecular dynamics simulations

Building on docking predictions, molecular dynamics (MD) simulations offer a temporal perspective to further assess the stability and dynamic behavior of peptide-protein systems.⁸⁵ By solving the atomic equations of motion under classical force fields, MD simulations capture the evolution of molecular systems in solvent environments, a capability that, as demonstrated by Jephthah et al., allows for the direct observation of peptide folding, conformational changes, and interactions over time.⁸⁶ Unlike static docking models, MD accounts for atomic-scale flexibility, solvent effects, and thermal fluctuations, providing significant advantages in revealing binding stability and interaction persistence.⁸⁷ In peptide drug design, MD is extensively applied to evaluate the conformational stability

of candidate sequences, analyze the preservation of key interactions, and compare the impact of various modifications on system dynamics. Lau et al. utilized molecular dynamics simulations to facilitate the engineering of Semaglutide, explicitly verifying that the hydrophilic linker possessed sufficient flexibility to accommodate albumin binding without disrupting the pivotal interactions between the GLP-1 moiety and its receptor. This dynamic validation ensured that the pharmacokinetic extension strategy did not compromise therapeutic efficacy.¹¹ With advancements in computational resources and algorithmic efficiency, MD simulations now encompass larger

ity of candidate sequences, analyze the preservation of key interactions, and compare the impact of various modifications on system dynamics. Lau et al. utilized molecular dynamics simulations to facilitate the engineering of Semaglutide, explicitly verifying that the hydrophilic linker possessed sufficient flexibility to accommodate albumin binding without disrupting the pivotal interactions between the GLP-1 moiety and its receptor. This dynamic validation ensured that the pharmacokinetic extension strategy did not compromise therapeutic efficacy.¹¹ With advancements in computational resources and algorithmic efficiency, MD simulations now encompass larger

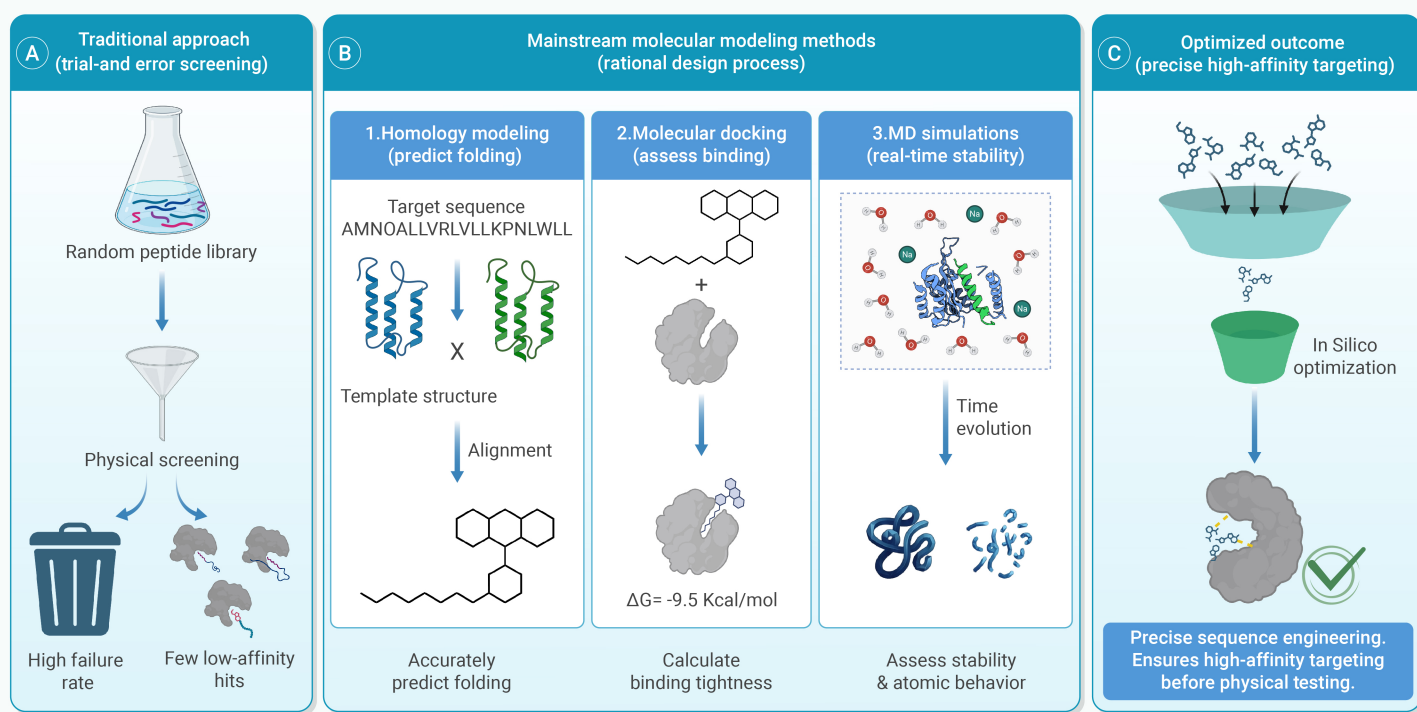


Figure 3. Workflow illustrating the transition from traditional trial-and-error to rational peptide drug design The traditional approach (A) relies on inefficient physical screening of random peptide libraries, often yielding high failure rates. The rational design process (B) leverages computational modeling tools: Homology Modeling to predict folding, Molecular Docking to calculate binding tightness (affinity), and Molecular Dynamics (MD) simulations to assess real-time atomic stability. This in silico optimization enables precise engineering of lead sequences, ensuring high-affinity targeting before physical testing begins (C).

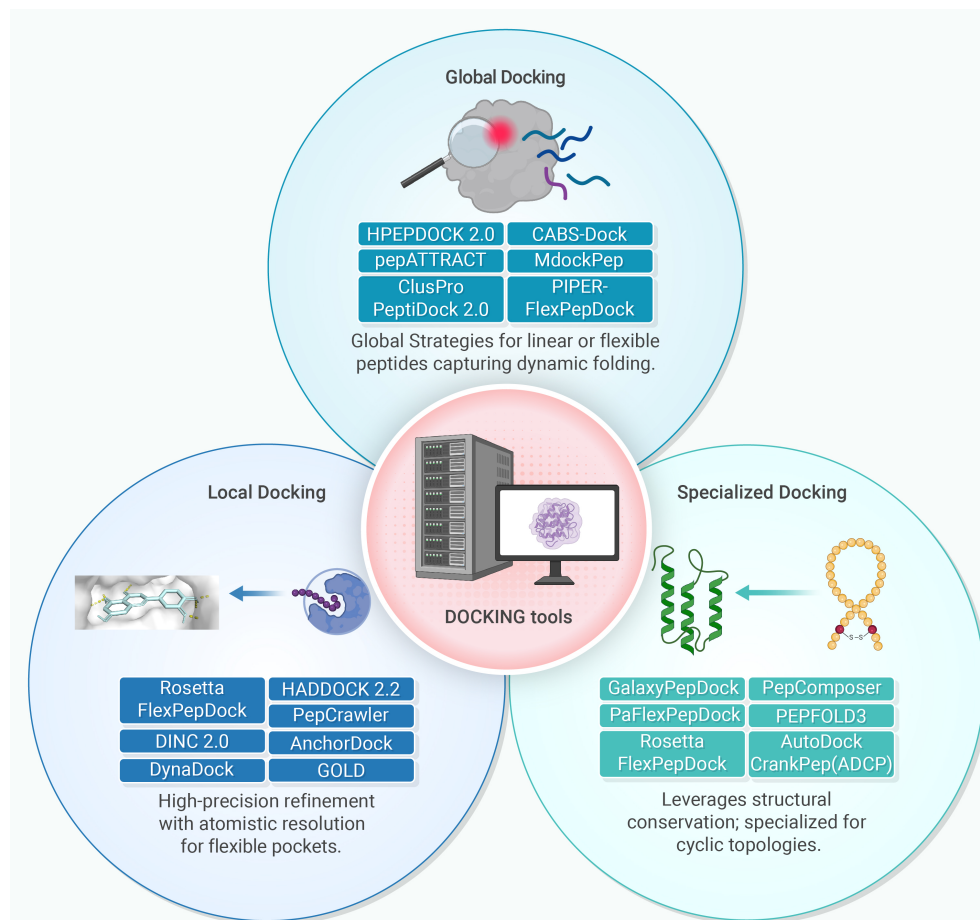


Figure 4. Classification of strategies and commonly used tools for peptide-protein molecular docking

peptide and biological activity data. This integration enhances predictive efficiency and design scalability while marking a transformative shift in the peptide discovery process. AI methodologies in peptide discovery can be broadly categorized into two distinct functional streams: predictive models for determining structure and function, and generative models for creating novel sequences to fulfill specific therapeutic criteria (Figure 5).^{96,97}

Predictive AI: from sequence to structure and function

Predictive models primarily focus on the accurate mapping of peptide sequences to their three-dimensional conformations and biological activities. While traditional methods trained on functional annotations established the groundwork, the advent of AlphaFold has revolutionized structural insight by offering high-precision predictions. AlphaFold employs evolutionary information via multiple sequence alignments and attention mechanisms to capture long-range residue correlations.³³ This capability is particularly powerful for modeling peptide-protein complexes and assessing potential binding modes, effectively addressing the challenge of induced-fit conformational

systems and longer timescales, facilitating a deeper exploration of the complex conformational space of peptides.^{88,89}

The choice of solvent models and force field parameters is critical in peptide molecular dynamics simulations. Explicit solvent models provide a more accurate representation of hydrogen bond networks and hydration effects but at the cost of higher computational demands.⁹⁰ In contrast, implicit solvent models offer greater computational efficiency, though they may sacrifice precision in capturing fine-scale interactions.⁹¹

Beyond the trade-offs between solvent accuracy and computational cost, a persistent challenge in peptide simulations involves effectively traversing rough energy landscapes populated by numerous local minima. To mitigate these sampling limitations, enhanced sampling techniques such as Replica Exchange Molecular Dynamics (REMD) and hybrid hierarchical strategies combining Accelerated Molecular Dynamics (aMD) with metadynamics have gained prominence. REMD facilitates the traversal of energy barriers through the utilization of multiple temperature replicas, a capability that has been effectively demonstrated in the structural characterization of amyloid peptides.⁹² Similarly, hybrid approaches enhance predictive accuracy by employing aMD for rapid conformational scanning alongside metadynamics for precise free energy estimation.⁹³ Furthermore, Weber et al. demonstrated that the integration of Machine Learning (ML) algorithms effectively guides molecular dynamics simulations to efficiently explore key conformational regions and accelerate convergence, thereby addressing the complexities of high-dimensional conformational sampling and rare transition events.⁹⁴

STRUCTURE MODELING IN MACHINE LEARNING AND TARGETED PEPTIDE DISCOVERY

As peptide drug design evolves from experience-based methods to data-driven approaches, artificial intelligence (AI) technologies are emerging as key enablers that bridge sequence, structure, and function spaces.⁹⁵ Unlike traditional molecular modeling techniques relying on explicit physical assumptions and computationally expensive energy functions, machine learning (ML) and deep learning (DL) methods identify underlying patterns from existing

changes that often confound traditional docking algorithms.^{98,99} By providing near-experimental accuracy, AlphaFold serves as a rigorous structural validator for candidate sequences derived from biological screening.

Complementing the high-precision capabilities of alignment-based methods, emerging protein language models such as ESMFold represent a significant leap in computational efficiency. Unlike AlphaFold which relies on computationally intensive multiple sequence alignments, ESMFold leverages vast sequence databases to predict structures directly from single sequences.¹⁰⁰ This architecture offers a speed advantage of up to an order of magnitude, making it uniquely reliable for high-throughput screening campaigns where millions of candidate peptides must be evaluated rapidly. Consequently, a tiered predictive strategy has emerged in which ESMFold is utilized for rapid large-scale filtration while AlphaFold is reserved for the detailed structural refinement of top-ranking candidates.¹⁰¹

To navigate this diverse methodological landscape, a critical comparison of these foundational technologies reveals distinct trade-offs between predictive accuracy, computational scalability, and the ability to capture conformational heterogeneity (Table 1). Although evolutionary algorithms like AlphaFold have achieved breakthrough accuracy in structure prediction, their reliance on multiple sequence alignment data constitutes an intrinsic bottleneck that leads to significantly reduced confidence when modeling *de novo* designed artificial sequences or orphan peptides lacking homologous data.^{33,98} Protein language models circumvent this limitation by learning deep grammatical patterns of amino acid sequences to achieve millisecond-level inference speeds, making ESMFold an ideal choice for screening ultra-large sequence libraries despite a slight compromise in accuracy for low-complexity sequences.¹⁰⁰ Cordoves-Delgado and García-Jacas emphasize that one must weigh screening throughput against structural detail, suggesting the utilization of language models for preliminary screening tasks involving millions of molecules while reserving AlphaFold for the lead optimization phase where its ability to precisely characterize interfacial steric hindrance remains irreplaceable.¹⁰¹ A rational computational pipeline should flexibly switch between these approaches according to the specific requirements of the project stage.

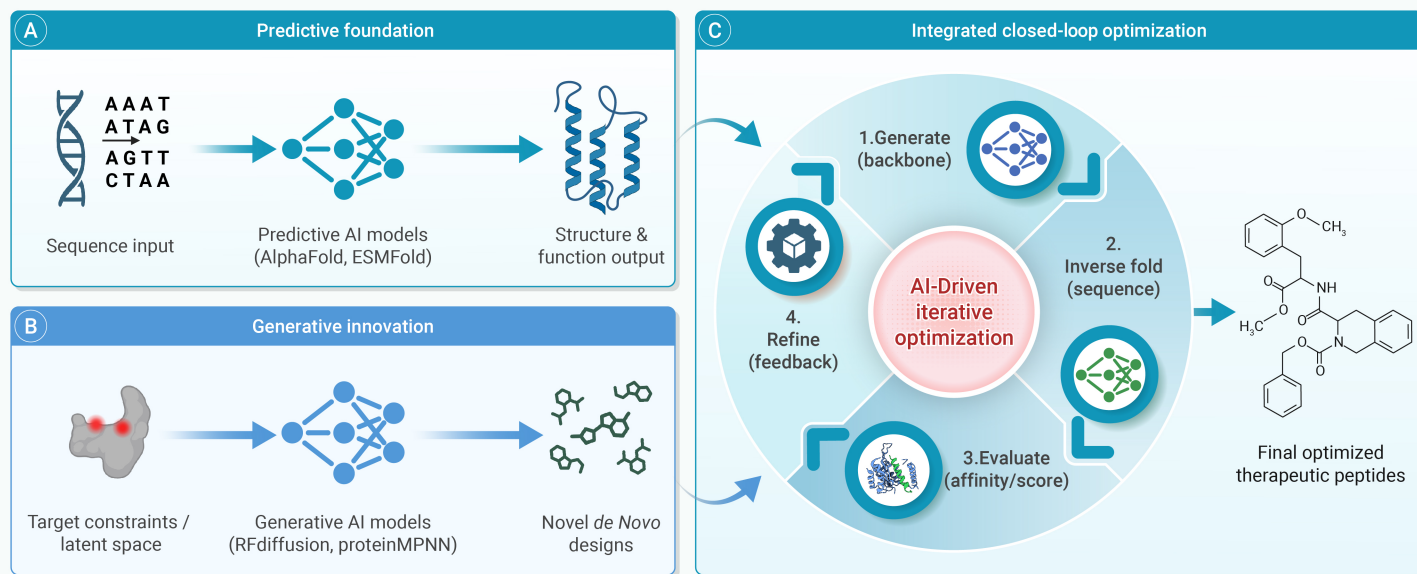


Figure 5. AI-driven workflow for peptide discovery (A) Predictive models map peptide sequences to structures. (B) Generative models enable *de novo* design based on target constraints. (C) An integrated closed-loop synergizes these streams through iterative generation, evaluation, and refinement to produce optimized therapeutics.

Generative AI: *de novo* design and space exploration

To navigate the rapidly evolving landscape of generative peptide design, this review proposes a classification framework based on the primary data modality driving the generative process. Specifically, these categories are differentiated by the nature of the input representation ranging from one-dimensional sequences to three-dimensional coordinates, the underlying optimization objective such as maximizing chemical diversity versus ensuring geometric complementarity, and whether the model explicitly generates spatial structures or latent sequence embeddings.^{106,107} This framework categorizes methodologies into sequence generation prioritization, structure generation prioritization, and sequence-structure co-optimization. To provide an actionable roadmap for computational design, a systematic evaluation of these generative frameworks reveals critical distinctions in operational requirements, topological constraints, and open-source accessibility (Table 2). The paradigm emphasizing sequence generation prioritizes the exploration of sequence space to identify novel motifs with desired physicochemical properties. Representative models such as Variational Autoencoders and Generative Adversarial Networks operate by mapping discrete amino acid sequences into a continuous latent space.^{106,108} By sampling from this manifold, these models generate diverse candidates ranging from antimicrobial peptides to cell-penetrating peptides that expand beyond the diversity found in natural databases. This approach proves particularly effective for functional peptides where the three-dimensional structure is dynamic or less strictly defined.

Addressing the geometric constraints of high-affinity binding, frontier frameworks such as RFdiffusion and the recently developed RFdiffusion3 have revolutionized the field by treating the binding pocket of the target protein as a conditional input to generate peptide backbones *de novo*.^{107,117} By treating the binding pocket of the target protein as a conditional input, these models prioritize geometric complementarity. They function by progressively denoising random coordinates to create backbones that precisely match the target interface. This approach marks a decisive shift for targeting undruggable sites such as flat protein-protein interaction surfaces where precise structural matching is prerequisite to binding. Cao et al. demonstrated the translational efficacy of this paradigm by rapidly developing *de novo* SARS-CoV-2 miniprotein inhibitors which achieved picomolar affinity in months. This computational strategy successfully bypassed the lengthy affinity maturation process inherent to traditional antibody discovery and validated the capability of generative models to accelerate therapeutic response.¹¹⁸ Expanding this landscape, emerging all-to-all interaction design paradigms exemplified by ODesign have begun to address the complexities of dynamic coupling.¹⁰⁹ By unifying multimodal biomolecular interactions within a single

generative framework rather than treating the target as a static constraint, these approaches signify a rapid evolution toward multi-modal and interaction-aware generative models that capture the reciprocal structural adaptations between peptides and their targets.

Despite the transformative potential of these tools, a significant gap remains between *in silico* generation and wet-lab synthesizability. Translating theoretically perfect backbones into physically viable molecules represents the most critical bottleneck in generative design. Current diffusion models, while adept at satisfying geometric constraints, are prone to "structural hallucination".¹⁰⁷ This phenomenon occurs when the model generates backbones that are geometrically complementary to the target but violate fundamental physicochemical principles required for stability.

Specifically, these hallucinated scaffolds often exhibit orphaned secondary structures, lack valid hydrophobic cores, or introduce excessive solvent-exposed hydrophobic surface areas. Consequently, although these computationally predicted designs may achieve high confidence scores or theoretical affinity in evaluations, they frequently fail in experimental validation due to aggregation, poor solubility, or the inability to adopt a stable folded state in solution. Therefore, as demonstrated by Dauparas et al.,¹¹⁴ sequence recovery tools like ProteinMPNN serve not merely as a method to solve the inverse folding problem but as essential filters connecting geometric design with chemical reality.

ADMET and developability assessment

While generative artificial intelligence models and structural predictors have revolutionized the optimization of target affinity, high binding affinity *in silico* does not inherently translate to *in vivo* therapeutic efficacy. The clinical attrition of peptide drug candidates is predominantly driven by poor absorption, distribution, metabolism, excretion, and toxicity profiles, including rapid proteolytic degradation and membrane impermeability. This translational gap is particularly pronounced for linear peptides. Lacking the conformational constraints of cyclic counterparts, novel linear scaffolds generated *de novo* are highly susceptible to enzymatic cleavage and often struggle to cross lipid bilayers. Integrating property prediction directly into the computational design pipeline is imperative.

Recent advancements in machine learning and deep learning have established robust sequence based and graph based models capable of accurately predicting key peptide developability metrics prior to physical synthesis. To address the challenge of undruggable intracellular targets, predictive tools identify and design sequence motifs that facilitate membrane translocation. Sequence based models and machine learning frameworks, such as CellPPD and MLCPP, leverage extensive datasets of known translocating sequences

Table 1. Common AI methods for predicting peptide 3D structures using sequence information

Tool	Synopsis	Availability	Dynamic & Multi-Conformation Support	Scope & Constraint Handling
AlphaFold ^{33,98}	Alignment-based deep learning model capable of high-accuracy protein and peptide structure prediction; adapted for cyclic peptides and supports predictions from single sequences.	AlphaFold Server is freely available for non-commercial research; Code available at https://github.com/google-deepmind/alphafold ; AlphaFold 3 code is downloadable for academic use, with training weights available upon request.	Primarily predicts a single static structure. However, advanced protocols like AF-Cluster utilize MSA clustering to unlock the prediction of multiple conformations and alternative states.	Scope: Monomer and Complex. Constraints: Native linear support; cyclic peptides require specialized inputs or AlphaFold 3.
ColabFoldv 1.5.5 ¹⁰²	A user-friendly interface implementing AlphaFold technology that allows batch predictions of structures in a single session.	A web server-like interface implemented as a Google Colab notebook utilizing AlphaFold2 models, with open-source code available at https://github.com/sokrypton/ColabFold , supporting both local installation and deployment on HPC clusters.	Supports MSA subsampling and random seeding to sample structural diversity, allowing for the generation of conformational ensembles beyond a single folded state.	Scope: Monomer and Complex. Constraints: Primarily linear; supports cyclic constraints via custom formatting.
HighFold v1.0 ¹⁰³	AlphaFold implementation designed for predicting cyclic peptides and complexes, incorporating constraints like head-to-tail connections and disulfide bridges.	HighFold utilizes AlphaFold2 models for local deployment, with code available on GitHub at https://github.com/hongliangduan/HighFold , and a cyclic design example provided at https://github.com/sokrypton/ColabDesign/blob/main/af/examples/af_cyc_design.ipynb .	Static/Constraint-based. Focuses on refining a single, high-confidence cyclic conformation rather than exploring dynamic ensembles.	Scope: Monomer and Complex. Constraints: Explicit support for cyclic constraints including head-to-tail cyclization and disulfide bridges.
AfCycDesign ¹⁰⁴	AlphaFold implementation optimized for the rapid, accurate prediction and design of cyclic peptide monomers, addressing limitations of insufficient training data.	Implemented within the ColabDesign framework utilizing AlphaFold, this tool is accessible via Neurosnap or Google Colab without local installation, with source code available at https://github.com/sokrypton/ColabDesign/blob/main/af/examples/af_cyc_design.ipynb .	Optimized for generating a specific, stable scaffold design rather than capturing conformational flexibility.	Scope: Monomer only. Constraints: Specialized for head-to-tail cyclic monomers.
ESMFold-2 ¹⁰¹	A deep learning model built on Meta AI's ESM-2 PLM that predicts 3D structures directly from amino acid sequences without relying on extensive homology or alignments.	Available as an open-source standalone package requiring no commercial license, with code and pre-trained weights on GitHub at https://github.com/facebookresearch/esm and API access via BioLM.ai.	Designed for ultra-fast prediction of a deterministic static structure; lacks the MSA-based diversity mechanisms used for multi-conformation sampling.	Scope: Monomer. Constraints: Strictly linear; no native cyclic support.
PepFlow ¹⁰⁵	Trained on peptides of 15 or fewer residues, this deep-learning model predicts dynamic folding patterns based on energy landscapes, accurately recapitulating structures and experimental ensembles at a fraction of the time required by traditional approaches.	Available as a GitLab repository with code and documentation at https://gitlab.com/oabdin/pepflow .	Explicitly designed to predict dynamic folding pathways and recapitulate experimental conformational ensembles, capturing peptide flexibility.	Scope: Monomer. Constraints: Linear peptides.

to extract physicochemical features that govern cellular uptake.^{119,120} The rapid clearance of peptides necessitates accurate *in silico* predictions of their pharmacokinetic lifespan. Emerging artificial intelligence tools like PlifePred focus on predicting sequence specific half lives in biological matrices and identifying specific protease cleavage sites.¹²¹ Preventing adverse immune responses and off target toxicity is another fundamental developability constraint. Classifiers such as ToxinPred utilize support vector machines and quantitative structure toxicity relationship modeling to predict hemolytic activity and general toxicity based on amino acid composition and specific dipeptide frequencies.¹²²

Relying solely on sequence based property predictors can occasionally yield false positives, particularly for highly novel linear scaffolds that venture

into unexplored chemical spaces. State of the art computational workflows integrate these one dimensional sequence predictors with three dimensional structural and dynamic validation. Once a candidate sequence passes the initial screening, predictive structural models like AlphaFold are employed to visualize the solvent accessible surface area and map the electrostatic potential of predicted protease cleavage sites. Physics based evaluations using molecular dynamics simulation engines like GROMACS provide a rigorous validation tier.¹²³ By simulating the peptide in an explicit aqueous environment, temporal metrics such as the Root Mean Square Deviation and the radius of gyration can be systematically tracked. This dynamic analysis crucially distinguishes between physically stable linear sequences and those prone to rapid structural collapse or aggregation in physiological conditions.

Table 2. Comprehensive summary of mainstream generative AI frameworks for peptide design and inverse folding

Tool	Primary Function / Modality	Input Requirements	Applicable Scenarios & Topological Constraints	Open-Source Availability (GitHub Links)
RFdiffusion ¹⁰⁷	De novo Structure Generation (Diffusion)	Target protein 3D structure; defined binding pocket or motif (optional).	Scenarios: High-affinity target-conditioned binder design. Constraints: Originally optimized for rigid mini-proteins. When migrating the algorithm to linear peptide generation, backbones often require rigorous MD validation to prevent unfolding in solution.	github.com/RosettaCom mons/RFdiffusion
ODesign ¹⁰⁹	Multimodal Biomolecular Interaction Design	Target constraints; multi-state or multimodal biomolecular data.	Scenarios: Dynamic all-to-all biomolecular interaction design, capturing reciprocal structural adaptations during binding. Constraints: Represents an emerging paradigm for multi-state coupling; highly valuable for modeling the induced-fit mechanisms often observed when linear peptides bind to targets.	https://odesign.lglab.ac. cn
BindCraft ¹¹⁰	Automated Binder Design Pipeline	Target protein structure; binding site coordinates	Scenarios: An end-to-end automated pipeline integrating RFdiffusion and ProteinMPNN to streamline in silico design. Constraints: Significantly lowers the barrier to entry for generating linear and structurally constrained binders, providing built-in confidence filtering.	https://github.com/mart inpacesa/BindCraft
PepMLM ¹¹¹	Target-Aware Peptide Generation (Masked Language Model)	Target protein sequence or representation.	Scenarios: Direct sequence-to-sequence generation of target-binding peptides, completely bypassing 3D structure dependency. Constraints: Exceptionally suited for generating highly flexible linear peptides where structure-based models struggle, though the output lacks explicit spatial binding poses prior to docking validation.	https://github.com/prog rammablebio/pepmlm
ProGen2 / ProtGPT2 ¹¹²	1D Sequence Generation (Protein LLMs)	Starting sequence prompt; desired properties (e.g., family, function).	Scenarios: Unconditional or property-guided generation of sequence spaces. Constraints: Operates exclusively in 1D sequence space. Generated linear sequences often lack defined 3D geometry and require subsequent structure prediction.	https://github.com/sale sforce/progen
EvoDiff ¹¹³	Sequence-based Generation (Discrete Diffusion)	Protein/peptide sequence lengths; partial sequence contexts.	Scenarios: Generation directly in sequence space, bypassing the need for 3D structural templates. Constraints: Particularly advantageous for intrinsically disordered regions and flexible linear peptides that lack defined 3D geometries.	https://github.com/micr osoft/evodiff
ProteinMPNN ¹¹⁴	Inverse Folding (Structure to Sequence)	3D backbone coordinates generated by diffusion models.	Scenarios: Crucial sequence recovery step. Translates theoretical geometries into physically viable sequences. Constraints: Natively operates on the 20 canonical amino acids; requires customized retraining to accurately recover sequences with non-natural modifications.	https://github.com/dau paras/ProteinMPNN
ESM-IF1 ¹¹⁵	Inverse Folding (Structure to Sequence)	3D backbone coordinates.	Scenarios: A powerful alternative to ProteinMPNN, leveraging the ESM language model architecture for sequence recovery. Constraints: Highly robust for diverse backbone geometries, though similarly restricted to standard amino acid representations.	github.com/facebookres earch/esm
LigandMPNN ¹¹⁶	Generalized Inverse Folding	3D backbone coordinates; spatial coordinates of ligands or modified residues.	Scenarios: Advanced sequence recovery explicitly conditioned on non-protein atoms. Constraints: Overcomes traditional limitations by allowing sequence design in the presence of non-canonical amino acids, metals, and small molecules.	https://github.com/dau paras/LigandMPNN

Filtering out candidates with unfavorable pharmacokinetic traits early in the pipeline serves as an essential checkpoint that ensures downstream optimization efforts are focused exclusively on viable, drug-like candidates.

The closed-loop integration

The synergy between predictive and generative models establishes a robust end-to-end design framework that optimizes therapeutic potential. This workflow typically begins with the generation of high-quality peptide backbones using diffusion models, followed by the application of graph neural network tools like ProteinMPNN to perform the inverse folding step.¹¹⁴ ProteinMPNN accurately translates the generated three-dimensional backbones into specific amino acid sequences that are physically capable of folding into the designed structures. This structural-to-sequence translation ensures that the geometric intent of the generative model is preserved in the final chemical composition of the peptide.

Following sequence recovery, predictive tools like AlphaFold or structure-aware machine learning models act as critical filters to evaluate structural confidence and interface plausibility within a generation-evaluation-feedback optimization loop.^{124,125} Rather than directly calculating thermodynamic binding affinity, these models utilize confidence metrics such as the Predicted Aligned Error (PAE) and interface predicted Template Modeling score (ipTM) as rapid proxies to rank potential binders prior to rigorous physics-based validation. In this iterative process, the predicted structural confidence and interface quality scores are fed back to the generative model to refine the search space and adjust distribution parameters. This active learning approach gradually focuses the model on regions of the chemical space with optimal binding interfaces, thereby reducing the rate of false positives. Crucially, to ensure clinical translatability, this loop must transcend simple affinity optimization by embedding sequence- and structure-based property predictors to enforce pharmacokinetic and developability constraints via a multi-objective optimization strategy.⁹⁷ This framework employs hard rule-based filters for the immediate binary rejection of generated sequences violating essential physicochemical thresholds, such as excessive molecular weight or high aggregation propensity scores, alongside multi-objective reward functions that expand the scoring metric in active learning iterations to include terms for solubility, proteolytic stability, and membrane permeability proxies.¹²⁴ Moreover, negative design signals derived from AI-driven developability profiling are incorporated to explicitly penalize features associated with immunogenicity or metabolic toxicity, effectively guiding the generative distribution away from valid but undruggable chemical spaces.¹¹⁴ This integrated workflow explicitly incorporates three-dimensional conformational information to significantly improve the specificity and success rate of targeted peptide therapeutics compared to traditional random screening methods.⁹⁷

AI modeling of peptides with non-Canonical amino acids

Overcoming the limitations imposed by the twenty canonical amino acids represents a critical frontier in peptide drug design. The incorporation of non canonical amino acids provides an essential strategy for enhancing proteolytic stability, modulating binding affinity, and engineering specific physicochemical properties. The reliance of current artificial intelligence frameworks on databases predominantly composed of natural proteins creates a significant bottleneck for the computational modeling of these chemically modified sequences.

A systematic evaluation of state of the art predictive tools reveals varying degrees of support for non canonical residues. Traditional deep learning models, including AlphaFold2 and ESMFold, are constrained by their training on the standard twenty amino acid alphabet, rendering them incapable of directly predicting conformations containing unnatural side chains or D-amino acids. While physics based suites like Rosetta accommodate these modifications, they require computationally expensive manual parameterization for each novel residue. The recent introduction of AlphaFold 3 signifies a major computational advancement, natively supporting a broad spectrum of modified residues and non standard ligands through generalized all atom representations rather than strict residue level tokenization.¹²⁶

In the realm of generative artificial intelligence, protein language models encounter fundamental tokenization limitations. Because these models utilize

a predefined discrete vocabulary corresponding to canonical amino acids, incorporating non standard residues disrupts the learned grammatical patterns. Addressing this limitation necessitates innovative representational strategies. One immediate approach involves expanding the model vocabulary and fine tuning pre trained networks on customized datasets containing non standard tokens. Alternatively, shifting from residue based representations to generalized chemical string formats, such as the Simplified Molecular Input Line Entry System or Self Referencing Embedded Strings, allows language models to process peptides as arbitrary chemical graphs rather than restricted amino acid chains.¹²⁷ Furthermore, emerging continuous space generative frameworks bypass one dimensional tokenization entirely. By generating three dimensional all atom coordinates directly via flow matching or atomistic diffusion paradigms, these structure based approaches inherently accommodate the geometric and chemical diversity of non canonical amino acids, seamlessly integrating them into the de novo design pipeline.¹⁰⁵

Strategic implementation and methodological constraints of AI paradigms

The effective application of artificial intelligence in peptide therapeutics necessitates a nuanced understanding of the intrinsic trade-off between computational throughput and structural resolution characterizing distinct algorithmic architectures. Protein language models including ESMFold prioritize inference speed to facilitate the extensive filtration of massive sequence libraries while evolutionary-based platforms such as AlphaFold provide the high-fidelity atomic coordinates essential for lead optimization despite their reliance on multiple sequence alignments.^{100,101} This dependence on evolutionary history restricts the utility of alignment-based methods when modeling orphan peptides or *de novo* designed sequences that lack homologous templates in existing databases.^{33,98} Generative diffusion frameworks excel in constructing geometric complementarity against undruggable protein interfaces yet frequently yield thermodynamically unstable backbones that require rigorous physical validation to eliminate these structural hallucinations.^{107,114} Practically, identifying these artifacts requires a hierarchical "validation ladder" to ensure resource efficiency. The initial tier involves quick energy minimization or relaxation, where candidates exhibiting severe steric clashes or highly unfavorable energy scores quantified by standard metrics like Rosetta energy units are immediately discarded.¹¹⁸ Surviving designs subsequently undergo short-timescale molecular dynamics simulations ranging from 10 to 50 nanoseconds, where concrete failure criteria including rapid structural collapse, high Root Mean Square Deviation (RMSD) fluctuations, or the disintegration of designed hydrogen-bond networks serve as reliable proxies for thermodynamic instability.⁸⁶ Only sequences demonstrating robust conformational persistence across these stages warrant advancement to computationally intensive free energy calculations or wet-lab synthesis. A robust discovery pipeline consequently demands the hierarchical integration of these diverse methodologies where rapid generative exploration precedes high-precision structural refinement to successfully reconcile geometric novelty with biological feasibility.^{97,124}

CLINICAL TRANSLATION CASES OF COMPUTATIONALLY DESIGNED PEPTIDES

The translatability of computational design strategies from theoretical concepts to clinical utility is powerfully illustrated by the rational engineering of Semaglutide. The natural GLP-1 backbone suffers from rapid enzymatic degradation and renal clearance which necessitated a shift beyond empirical chemical modification. Computational structural analysis guided the precise substitution of amino acids and the incorporation of a C18 fatty acid diacid via a hydrophilic spacer. This design strategy induces a depot effect through reversible albumin binding while maintaining a specific linker geometry that avoids steric interference with the receptor active site. The success of Semaglutide demonstrates that pharmacokinetic properties can be engineered as intrinsic structural features rather than relying on retrospective chemical adaptations.^{11,128,129}

The development of the macrocyclic PCSK9 inhibitor MK-0616 extends rational design into the challenging domain of oral bioavailability for large biologic-like peptides. Inhibiting flat protein-protein interaction surfaces typi-

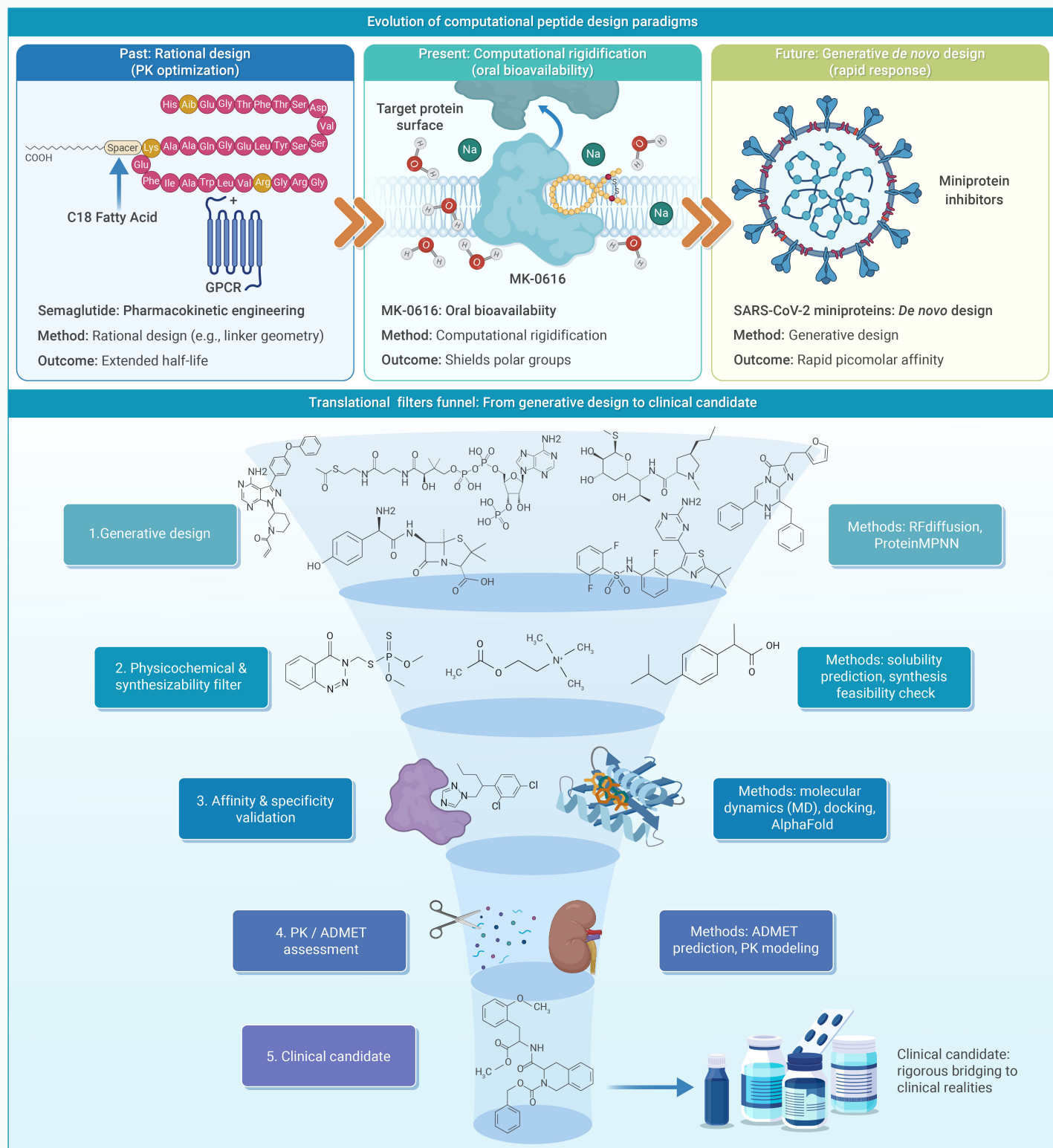


Figure 6. Evolution of computational peptide drug design paradigms and the multi-level translational validation framework This figure reveals the evolution of design strategies, from rational engineering to generative AI, and illustrates the necessary multi-stage filtering process required to bridge computational predictions with successful clinical translation.

cally demands molecular sizes that exceed the limits of intestinal permeability. Computational rigidification of the macrocyclic scaffold reconciled this conflict by locking the peptide into a bioactive conformation. This structural constraint minimizes the entropic penalty of binding and optimizes intramolecular hydrogen bonding to shield polar groups during membrane transit. This case validates a design paradigm where structural rigidification

decouples molecular weight from bioavailability to enable the oral delivery of intracellular therapeutics.^{75,130,131}

Advancements in computational power have facilitated a transition toward *de novo* generation as demonstrated by the rapid development of SARS-CoV-2 miniprotein inhibitors. Generative algorithms addressed the urgent need for therapeutics by designing synthetic scaffolds based solely on geometric

complementarity to the viral Spike protein. These computational designs achieved picomolar affinity and potent neutralization in animal models within months by creating rigid interfaces that precisely match the target topography. This achievement highlights the emergence of a compute-first discovery model where high-affinity therapeutics are generated on demand to address emerging biological threats with unprecedented speed and precision.^{107,118}

Despite these transformative successes, the translation of computational blueprints into clinical realities frequently confronts the divergence between theoretical predictions and experimental behavior. A primary obstacle lies in the synthetic feasibility of *de novo* designed sequences where generative models focusing exclusively on geometric complementarity often introduce excessive hydrophobic surface areas or aggregation-prone motifs that preclude successful chemical synthesis and aqueous solubility.^{107,114} Furthermore, high binding affinity validated in silico does not inherently translate to in vivo therapeutic efficacy as computational optimization often neglects crucial pharmacokinetic parameters such as proteolytic susceptibility and renal clearance rates which dictate the biological half-life of the final candidate.^{12,13} The rigorous bridging of this translational gap necessitates the incorporation of explicit solubility constraints and negative design filters during the generative phase to ensure that the resultant molecules possess both structural integrity and drug-like physicochemical properties appropriate for physiological environments (Figure 6).¹¹⁴

DISCUSSION AND CONCLUSIONS

The landscape of peptide drug discovery is undergoing a fundamental transformation driven by artificial intelligence technologies including AlphaFold and protein language models. AlphaFold has effectively resolved the sequence-to-structure prediction challenge, while generative AI tools such as RFDiffusion revolutionize the rational design pipeline from target identification to drug candidate generation. These advancements substantially expand the exploration of peptide conformational space and enable the precise targeting of sites previously deemed intractable, which marks a definitive paradigm shift in peptide research from traditional discovery to innovative creation.

Traditional strategies grounded in homology modeling and molecular docking have historically underpinned rational design; yet, their dependence on homologous templates and limited sampling efficiency restricts their utility when targeting dynamic undruggable interfaces. Generative artificial intelligence transcends these boundaries by exploring vast chemical spaces to identify novel scaffolds, while molecular dynamics simulations remain essential for providing the thermodynamic validation and atomic-level resolution that statistical models alone cannot guarantee. An optimal discovery architecture therefore integrates these paradigms, where AI-driven generation precedes physics-based refinement to reconcile computational speed with biological accuracy.

Translating high-affinity peptide binders into clinical drugs continues to face substantial impediments despite significant advances in computational design. Current peptide therapeutics often rely on complex chemical modifications like non-natural amino acid substitutions and backbone cyclization to circumvent inherent issues involving metabolic instability and poor membrane permeability. A critical limitation remains that most existing deep generative models are trained primarily on natural amino acid systems, which excel at designing binding scaffolds with high geometric compatibility but struggle to incorporate non-standard chemical modifications or predict in vivo pharmacokinetic properties. Emerging computational strategies are increasingly integrating specific modification constraints directly into the design process to address this disparity. Recent advancements in structure prediction tools such as HighFold now support the explicit modeling of cyclic constraints including disulfide bridges and head-to-tail cyclization, which enables the accurate prediction of constrained peptide geometries.

Furthermore, a critical frontier for the next generation of peptide design is the expansion of the chemical vocabulary beyond the standard twenty natural amino acids. Current generative models remain largely confined to natural backbones which limits their ability to address key developability challenges such as rapid proteolytic degradation and poor membrane permeability. Future frameworks must explicitly incorporate non-canonical amino acids

and diverse chemical modifications into the generative process. By extending the design space to include these chemically modified building blocks, AI models will be better equipped to directly optimize intrinsic stability and bioavailability thereby bridging the gap between computational hits and clinical drug candidates.

Finally, the next generation of computational design necessitates the integration of Explainable Artificial Intelligence to elucidate the specific structural features driving predictive outcomes, thereby providing mechanistic insights into the underlying structure-activity relationships. Furthermore, the evolution of these predictive frameworks relies on the advancement of Multimodal Data Fusion strategies that synergize one-dimensional sequence information with three-dimensional geometric data and wet-lab functional readouts to construct robust models capable of accurately navigating the complex fitness landscape of peptide therapeutics. This convergence of interpretable algorithmic logic with comprehensive biological data integration holds the potential to overcome the critical barrier between computational screening and clinical application, thereby accelerating the development of precise and translatable targeted peptide drugs.

REFERENCES

- Sharma K., Sharma K. K., Sharma A., et al. (2023). Peptide-based drug discovery: Current status and recent advances. *Drug Discov Today* **28**:103464. DOI:10.1016/j.drudis.2022.103464
- Muttenthaler M., King G. F., Adams D. J., et al. (2021). Trends in peptide drug discovery. *Nat Rev Drug Discov* **20**:309–325. DOI:10.1038/s41573-020-00135-8
- Wang L., Wang N., Zhang W., et al. (2022). Therapeutic peptides: current applications and future directions. *Signal Transduct Target Ther* **7**:48. DOI:10.1038/s41392-022-00904-4
- Nhàn N. T. T., Yamada T. and Yamada K. H. (2023). Peptide-Based Agents for Cancer Treatment: Current Applications and Future Directions. *Int J Mol Sci* **24**. DOI:10.3390/ijms241612931.
- Afonso A. L., Cavaleiro C. T., Castanho M., et al. (2025). The Potential of Peptide-Based Inhibitors in Disrupting Protein-Protein Interactions for Targeted Cancer Therapy. *Int J Mol Sci* **26**. DOI:10.3390/ijms26073117.
- Rivera-Jiménez J., Berraquero-García C., Pérez-Gálvez R., et al. (2022). Peptides and protein hydrolysates exhibiting anti-inflammatory activity: sources, structural features and modulation mechanisms. *Food Funct* **13**:12510–12540. DOI:10.1039/d2fo02223k
- Yang N. J. and Hinner M. J. (2015). Getting across the cell membrane: an overview for small molecules, peptides, and proteins. *Methods Mol Biol* **1266**:29–53. DOI:10.1007/978-1-4939-2272-7_3
- Müller T. D., Blüher M., Tschöp M. H., et al. (2022). Anti-obesity drug discovery: advances and challenges. *Nat Rev Drug Discov* **21**:201–223. DOI:10.1038/s41573-021-00337-8
- Drucker D. J. (2025). GLP-1-based therapies for diabetes, obesity and beyond. *Nat Rev Drug Discov* **24**:631–650. DOI:10.1038/s41573-025-01183-8
- Drucker D. J. (2024). Prevention of cardiorenal complications in people with type 2 diabetes and obesity. *Cell Metab* **36**:338–353. DOI:10.1016/j.cmet.2023.12.018
- Lau J., Bloch P., Schäffer L., et al. (2015). Discovery of the Once-Weekly Glucagon-Like Peptide-1 (GLP-1) Analogue Semaglutide. *J Med Chem* **58**:7370–7380. DOI:10.1021/acs.jmedchem.5b00726
- Pei J., Gao X., Pan D., et al. (2022). Advances in the stability challenges of bioactive peptides and improvement strategies. *Curr Res Food Sci* **5**:2162–2170. DOI:10.1016/j.crf.2022.10.031
- Drucker D. J. (2020). Advances in oral peptide therapeutics. *Nat Rev Drug Discov* **19**:277–289. DOI:10.1038/s41573-019-0053-0
- Agoni C., Fernández-Díaz R., Timmons P. B., et al. (2025). Molecular Modelling in Bioactive Peptide Discovery and Characterisation. *Biomolecules* **15**. DOI:10.3390/biom15040524.
- Yao J. F., Yang H., Zhao Y. Z., et al. (2018). Metabolism of Peptide Drugs and Strategies to Improve their Metabolic Stability. *Curr Drug Metab* **19**:892–901. DOI:10.2174/1389200219666180628171531
- Tan H., Su W., Zhang W., et al. (2018). Recent Advances in Half-life Extension Strategies for Therapeutic Peptides and Proteins. *Curr Pharm Des* **24**:4932–4946. DOI:10.2174/1381612825666190206105232
- Parida P., Prusty A. K., Patro S. K., et al. (2024). Current Advancements on Oral Protein and Peptide Drug Delivery Approaches to Bioavailability: Extensive Review on Patents. *Recent Adv Drug Deliv Formul* **18**:227–246. DOI:10.2174/0126673878299775240719061653
- Haddadzadeh S., Dorkoosh F. and Bernkop-Schnürch A. (2022). Oral delivery of therapeutic peptides and proteins: Technology landscape of lipid-based nanocarriers. *Adv Drug Deliv Rev* **182**:114097. DOI:10.1016/j.addr.2021.114097
- Chan L. Y., Zhang V. M., Huang Y. H., et al. (2013). Cyclization of the antimicrobial peptide gomesin with native chemical ligation: influences on stability and bioactivity. *Chembiochem* **14**:617–624. DOI:10.1002/cbic.201300034
- Costa L., Sousa E. and Fernandes C. (2023). Cyclic Peptides in Pipeline: What Future for These Great Molecules? *Pharmaceuticals (Basel)* **16**. DOI:10.3390/ph16070996.
- Adessi C. and Soto C. (2002). Converting a peptide into a drug: strategies to improve stability and bioavailability. *Curr Med Chem* **9**:963–978. DOI:10.2174/0929867024606731
- Gentilucci L., De Marco R. and Cerisoli L. (2010). Chemical modifications designed

- to improve peptide stability: incorporation of non-natural amino acids, pseudo-peptide bonds, and cyclization. *Curr Pharm Des* **16**:3185–3203. DOI:10.2174/138161210793292555
23. Fosgerau K. and Hoffmann T. (2015). Peptide therapeutics: current status and future directions. *Drug Discov Today* **20**:122–128. DOI:10.1016/j.drudis.2014.10.003
 24. Schneider P., Walters W. P., Plowright A. T., et al. (2020). Rethinking drug design in the artificial intelligence era. *Nat Rev Drug Discov* **19**:353–364. DOI:10.1038/s41573-019-0050-3
 25. Hashemi Z. S., Zarei M., Fath M. K., et al. (2021). In silico Approaches for the Design and Optimization of Interfering Peptides Against Protein-Protein Interactions. *Front Mol Biosci* **8**:669431. DOI:10.3389/fmolb.2021.669431
 26. Vidal-Limon A., Aguilar-Toalá J. E. and Liceaga A. M. (2022). Integration of Molecular Docking Analysis and Molecular Dynamics Simulations for Studying Food Proteins and Bioactive Peptides. *J Agric Food Chem* **70**:934–943. DOI:10.1021/acs.jafc.1c06110
 27. Wang Y., Stebe K. J., de la Fuente-Nunez C., et al. (2024). Computational Design of Peptides for Biomaterials Applications. *ACS Appl Bio Mater* **7**:617–625. DOI:10.1021/acsabm.2c01023
 28. Kazmirschuk T. D. D., Bradbury-Jost C., Withey T. A., et al. (2023). Peptides of a Feather: How Computation Is Taking Peptide Therapeutics under Its Wing. *Genes (Basel)* **14**. DOI:10.3390/genes14061194
 29. Allison J. R. (2020). Computational methods for exploring protein conformations. *Biochem Soc Trans* **48**:1707–1724. DOI:10.1042/bst20200193
 30. Mondal A., Chang L. and Perez A. (2022). Modelling peptide-protein complexes: docking, simulations and machine learning. *QRB Discov* **3**:e17. DOI:10.1017/qrd.2022.14
 31. Geng H., Chen F., Ye J., et al. (2019). Applications of Molecular Dynamics Simulation in Structure Prediction of Peptides and Proteins. *Comput Struct Biotechnol J* **17**:1162–1170. DOI:10.1016/j.csbj.2019.07.010
 32. Muhammed M. T. and Aki-Yalcin E. (2019). Homology modeling in drug discovery: Overview, current applications, and future perspectives. *Chem Biol Drug Des* **93**:12–20. DOI:10.1111/cbdd.13388
 33. Jumper J., Evans R., Pritzel A., et al. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature* **596**:583–589. DOI:10.1038/s41586-021-03819-2
 34. Baek M., DiMaio F., Anishchenko I., et al. (2021). Accurate prediction of protein structures and interactions using a three-track neural network. *Science* **373**:871–876. DOI:10.1126/science.abb8754
 35. John B. and Sali A. (2003). Comparative protein structure modeling by iterative alignment, model building and model assessment. *Nucleic Acids Res* **31**:3982–3992. DOI:10.1093/nar/gkg460
 36. MacCallum J. L., Pérez A., Schindiers M. J., et al. (2011). Assessment of protein structure refinement in CASP9. *Proteins* **79** Suppl 10:74–90. DOI:10.1002/prot.23131
 37. Schmidt T., Bergner A. and Schwede T. (2014). Modelling three-dimensional protein structures for applications in drug design. *Drug Discov Today* **19**:890–897. DOI:10.1016/j.drudis.2013.10.027
 38. Floudas C. A. (2007). Computational methods in protein structure prediction. *Biotechnol Bioeng* **97**:207–213. DOI:10.1002/bit.21411
 39. Fiser A. (2010). Template-based protein structure modeling. *Methods Mol Biol* **673**:73–94. DOI:10.1007/978-1-60761-842-3_6
 40. Xiang Z. (2006). Advances in homology protein structure modeling. *Curr Protein Pept Sci* **7**:217–227. DOI:10.2174/138920306777452312
 41. Anighoro A. (2022). Deep Learning in Structure-Based Drug Design. *Methods Mol Biol* **2390**:261–271. DOI:10.1007/978-1-0716-1787-8_11
 42. Rodrigues J. P., Melquiond A. S., Karaca E., et al. (2013). Defining the limits of homology modeling in information-driven protein docking. *Proteins* **81**:2119–2128. DOI:10.1002/prot.24382
 43. Haddad Y., Adam V. and Heger Z. (2020). Ten quick tips for homology modeling of high-resolution protein 3D structures. *PLoS Comput Biol* **16**:e1007449. DOI:10.1371/journal.pcbi.1007449
 44. Baeza J., Bedoya M., Cruz P., et al. (2025). Main methods and tools for peptide development based on protein-protein interactions (PPIs). *Biochem Biophys Res Commun* **758**:151623. DOI:10.1016/j.bbrc.2025.151623
 45. Gupta V. A., Ackley J., Kaufman J. L., et al. (2021). BCL2 Family Inhibitors in the Biology and Treatment of Multiple Myeloma. *Blood Lymphat Cancer* **11**:11–24. DOI:10.2147/blott.S245191
 46. Shuli Z., Linlin L., Li G., et al. (2022). Bioinformatics and Computer Simulation Approaches to the Discovery and Analysis of Bioactive Peptides. *Curr Pharm Biotechnol* **23**:1541–1555. DOI:10.2174/1389201023666220106161016
 47. Hetényi C. and van der Spoel D. (2002). Efficient docking of peptides to proteins without prior knowledge of the binding site. *Protein Sci* **11**:1729–1737. DOI:10.1110/ps.0202302
 48. Audie J. and Swanson J. (2012). Recent work in the development and application of protein-peptide docking. *Future Med Chem* **4**:1619–1644. DOI:10.4155/fmc.12.99
 49. Kellenberger E., Rodrigo J., Muller P., et al. (2004). Comparative evaluation of eight docking tools for docking and virtual screening accuracy. *Proteins* **57**:225–242. DOI:10.1002/prot.20149
 50. Weng G., Gao J., Wang Z., et al. (2020). Comprehensive Evaluation of Fourteen Docking Programs on Protein-Peptide Complexes. *J Chem Theory Comput* **16**:3959–3969. DOI:10.1021/acs.jctc.9b01208
 51. Zhao H., Jiang D., Shen C., et al. (2024). Comprehensive Evaluation of 10 Docking Programs on a Diverse Set of Protein-Cyclic Peptide Complexes. *J Chem Inf Model* **64**:2112–2124. DOI:10.1021/acs.jcim.3c01921
 52. Agrawal P., Singh H., Srivastava H. K., et al. (2019). Benchmarking of different molecular docking methods for protein-peptide docking. *BMC Bioinformatics* **19**:426. DOI:10.1186/s12859-018-2449-y
 53. Kundrotas P. J., Zhu Z., Janin J., et al. (2012). Templates are available to model nearly all complexes of structurally characterized proteins. *Proc Natl Acad Sci U S A* **109**:9438–9441. DOI:10.1073/pnas.1200678109
 54. Cierny M., Kurcinski M., Kamel K., et al. (2018). Protein-peptide docking: opportunities and challenges. *Drug Discov Today* **23**:1530–1537. DOI:10.1016/j.drudis.2018.05.006
 55. Sadybekov A. V. and Katritch V. (2023). Computational approaches streamlining drug discovery. *Nature* **616**:673–685. DOI:10.1038/s41586-023-05905-z
 56. Zhou P., Jin B., Li H., et al. (2018). HPEPDOCK: a web server for blind peptide-protein docking based on a hierarchical algorithm. *Nucleic Acids Res* **46**:W443–W450. DOI:10.1093/nar/gky357
 57. Kurcinski M., Badaczewska-Dawid A., Kolinski M., et al. (2020). Flexible docking of peptides to proteins using CABS-dock. *Protein Sci* **29**:211–222. DOI:10.1002/pro.3771
 58. Zhang Y. and Sanner M. F. (2019). AutoDock CrankPep: combining folding and docking to predict protein-peptide complexes. *Bioinformatics* **35**:5121–5127. DOI:10.1093/bioinformatics/btz459
 59. Lee H., Heo L., Lee M. S., et al. (2015). GalaxyPepDock: a protein-peptide docking tool based on interaction similarity and energy optimization. *Nucleic Acids Res* **43**:W431–435. DOI:10.1093/nar/gkv495
 60. London N., Raveh B., Cohen E., et al. (2011). Rosetta FlexPepDock web server—high resolution modeling of peptide-protein interactions. *Nucleic Acids Res* **39**:W249–253. DOI:10.1093/nar/gkr431
 61. Donsky E. and Wolfson H. J. (2011). PepCrawler: a fast RRT-based algorithm for high-resolution refinement and binding affinity estimation of peptide inhibitors. *Bioinformatics* **27**:2836–2842. DOI:10.1093/bioinformatics/btr498
 62. Dominguez C., Boelens R. and Bonvin A. M. (2003). HADDOCK: a protein-protein docking approach based on biochemical or biophysical information. *J Am Chem Soc* **125**:1731–1737. DOI:10.1021/ja026939x
 63. de Vries S. J., Rey J., Schindler C. E. M., et al. (2017). The pepATTRACT web server for blind, large-scale peptide-protein docking. *Nucleic Acids Res* **45**:W361–W364. DOI:10.1093/nar/gkx335
 64. Alam N., Goldstein O., Xia B., et al. (2017). High-resolution global peptide-protein docking using fragments-based PIPER-FlexPepDock. *PLoS Comput Biol* **13**:e1005905. DOI:10.1371/journal.pcbi.1005905
 65. Porter K. A., Xia B., Beglov D., et al. (2017). ClusPro PeptiDock: efficient global docking of peptide recognition motifs using FFT. *Bioinformatics* **33**:3299–3301. DOI:10.1093/bioinformatics/btx216
 66. Antes I. (2010). DynaDock: A new molecular dynamics-based algorithm for protein-peptide docking including receptor flexibility. *Proteins* **78**:1084–1104. DOI:10.1002/prot.22629
 67. Ben-Shimon A. and Niv M. Y. (2015). AnchorDock: Blind and Flexible Anchor-Driven Peptide Docking. *Structure* **23**:929–940. DOI:10.1016/j.str.2015.03.010
 68. Xu X., Yan C. and Zou X. (2018). MDockPeP: An ab-initio protein-peptide docking server. *J Comput Chem* **39**:2409–2413. DOI:10.1002/jcc.25555
 69. Lamiae A., Thévenet P., Rey J., et al. (2016). PEP-FOLD3: faster de novo structure prediction for linear peptides in solution and in complex. *Nucleic Acids Res* **44**:W449–454. DOI:10.1093/nar/gkw329
 70. Antunes D. A., Moll M., Devaurs D., et al. (2017). DINC 2.0: A New Protein-Peptide Docking Webserver Using an Incremental Approach. *Cancer Res* **77**:e55–e57. DOI:10.1158/0008-5472.Can-17-0511
 71. Jones G., Willett P., Glen R. C., et al. (1997). Development and validation of a genetic algorithm for flexible docking. *J Mol Biol* **267**:727–748. DOI:10.1006/jmbi.1996.0897
 72. Raveh B., London N., Zimmerman L., et al. (2011). Rosetta FlexPepDock ab-initio: simultaneous folding, docking and refinement of peptides onto their receptors. *PLoS One* **6**:e18934. DOI:10.1371/journal.pone.0018934
 73. Li H., Lu L., Chen R., et al. (2014). PaFlexPepDock: parallel ab-initio docking of peptides onto their receptors with full flexibility based on Rosetta. *PLoS One* **9**:e94769. DOI:10.1371/journal.pone.0094769
 74. London N., Raveh B. and Schueler-Furman O. (2012). Modeling peptide-protein interactions. *Methods Mol Biol* **857**:375–398. DOI:10.1007/978-1-61779-588-6_17
 75. Tucker T. J., Embrey M. W., Alleyne C., et al. (2021). A Series of Novel, Highly Potent, and Orally Bioavailable Next-Generation Tricyclic Peptide PCSK9 Inhibitors. *J Med Chem* **64**:16770–16800. DOI:10.1021/acs.jmedchem.1c01599
 76. Rentzsch R. and Renard B. Y. (2015). Docking small peptides remains a great challenge: an assessment using AutoDock Vina. *Brief Bioinform* **16**:1045–1056. DOI:10.1093/bib/bbv008
 77. Huang S. Y. (2015). Exploring the potential of global protein-protein docking: an overview and critical assessment of current programs for automatic ab initio docking. *Drug Discov Today* **20**:969–977. DOI:10.1016/j.drudis.2015.03.007
 78. Yan C., Xu X. and Zou X. (2016). Fully Blind Docking at the Atomic Level for Protein-Peptide Complex Structure Prediction. *Structure* **24**:1842–1853. DOI:10.1016/j.str.2016.07.021
 79. Koshland D. E., Jr., Némethy G. and Filmer D. (1966). Comparison of experimental binding data and theoretical models in proteins containing subunits. *Biochemistry* **5**:365–385. DOI:10.1021/bi00865a047
 80. Antunes D. A., Devaurs D. and Kavrakli L. E. (2015). Understanding the challenges of protein flexibility in drug design. *Expert Opin Drug Discov* **10**:1301–1313. DOI:10.1517/17460441.2015.1094458
 81. Buonfiglio R., Recanatini M. and Masetti M. (2015). Protein Flexibility in Drug Discovery: From Theory to Computation. *ChemMedChem* **10**:1141–1148. DOI:10.1002/cmdc.201500086
 82. Miller E. B., Murphy R. B., Sindhikara D., et al. (2021). Reliable and Accurate Solution to the Induced Fit Docking Problem for Protein-Ligand Binding. *J Chem Theory Comput* **17**:2630–2639. DOI:10.1021/acs.jctc.1c00136
 83. Zhao H., Zhang O., Jiang D., et al. (2025). Protein-peptide docking with a rational and accurate diffusion generative model. *Nature Machine Intelligence* **7**:1308–1321. DOI:10.1038/s42256-025-01077-9

84. Ciemny M. P., Debinski A., Paczkowska M., et al. (2016). Protein-peptide molecular docking with large-scale conformational changes: the p53-MDM2 interaction. *Sci Rep* **6**:37532. DOI:10.1038/srep37532
85. De Vivo M., Masetti M., Bottegoni G., et al. (2016). Role of Molecular Dynamics and Related Methods in Drug Discovery. *J Med Chem* **59**:4035–4061. DOI:10.1021/acs.jmedchem.5b01684
86. Jephthah S., Pesce F., Lindorff-Larsen K., et al. (2021). Force Field Effects in Simulations of Flexible Peptides with Varying Polyproline II Propensity. *J Chem Theory Comput* **17**:6634–6646. DOI:10.1021/acs.jctc.1c00408
87. Salmaso V. and Moro S. (2018). Bridging Molecular Docking to Molecular Dynamics in Exploring Ligand-Protein Recognition Process: An Overview. *Front Pharmacol* **9**:923. DOI:10.3389/fphar.2018.00923
88. Chen C. H., Melo M. C., Berglund N., et al. (2020). Understanding and modelling the interactions of peptides with membranes: from partitioning to self-assembly. *Curr Opin Struct Biol* **61**:160–166. DOI:10.1016/j.sbi.2019.12.021
89. Zhang S., Lv S., Fu X., et al. (2021). Peptide Dynamics Simulations Study of the Interactions between Human Dipeptidyl-Peptidase III and Two Substrates. *Molecules* **26**. DOI:10.3390/molecules26216492
90. Florová P., Sklenovský P., Banáš P., et al. (2010). Explicit Water Models Affect the Specific Solvation and Dynamics of Unfolded Peptides While the Conformational Behavior and Flexibility of Folded Peptides Remain Intact. *J Chem Theory Comput* **6**:3569–3579. DOI:10.1021/ct1003687
91. Zhang J., Zhang H., Wu T., et al. (2017). Comparison of Implicit and Explicit Solvent Models for the Calculation of Solvation Free Energy in Organic Solvents. *J Chem Theory Comput* **13**:1034–1043. DOI:10.1021/acs.jctc.7b00169
92. Turner M., Mutter S. T., Kennedy-Britten O. D., et al. (2019). Replica exchange molecular dynamics simulation of the coordination of Pt(II)-Phenanthroline to amyloid- β . *RSC Adv* **9**:35089–35097. DOI:10.1039/c9ra04637b
93. Clayton J., Baweja L. and Wereszczynski J. (2022). Peptide Dynamics and Metadynamics: Leveraging Enhanced Sampling Molecular Dynamics to Robustly Model Long-Timescale Transitions. *Methods Mol Biol* **2405**:151–167. DOI:10.1007/978-1-0716-1855-4_8
94. Weber J. K., Morrone J. A., Kang S. G., et al. (2023). Unsupervised and supervised AI on molecular dynamics simulations reveals complex characteristics of HLA-A2-peptide immunogenicity. *Brief Bioinform* **25**. DOI:10.1093/bib/bbad504.
95. Perpetuo L., Klein J., Ferreira R., et al. (2021). How can artificial intelligence be used for peptidomics. *Expert Rev Proteomics* **18**:527–556. DOI:10.1080/14789450.2021.1962303
96. Chen Z., Wang R., Guo J., et al. (2024). The role and future prospects of artificial intelligence algorithms in peptide drug development. *Biomed Pharmacother* **175**:116709. DOI:10.1016/j.biopha.2024.116709
97. Goles M., Daza A., Cabas-Mora G., et al. (2024). Peptide-based drug discovery through artificial intelligence: towards an autonomous design of therapeutic peptides. *Brief Bioinform* **25**. DOI:10.1093/bib/bbae275.
98. Tsaban T., Varga J. K., Avraham O., et al. (2022). Harnessing protein folding neural networks for peptide-protein docking. *Nat Commun* **13**:176. DOI:10.1038/s41467-021-27838-9
99. McDonald E. F., Jones T., Plate L., et al. (2023). Benchmarking AlphaFold2 on peptide structure prediction. *Structure* **31**:111–119.e112. DOI:10.1016/j.str.2022.11.012
100. Lin Z., Akin H., Rao R., et al. (2023). Evolutionary-scale prediction of atomic-level protein structure with a language model. *Science* **379**:1123–1130. DOI:10.1126/science.ade2574
101. Cordoves-Delgado G. and García-Jacas C. R. (2024). Predicting Antimicrobial Peptides Using ESMFold-Predicted Structures and ESM-2-Based Amino Acid Features with Graph Deep Learning. *J Chem Inf Model* **64**:4310–4321. DOI:10.1021/acs.jcim.3c02061
102. Mirdita M., Schütze K., Moriwaki Y., et al. (2022). ColabFold: making protein folding accessible to all. *Nat Methods* **19**:679–682. DOI:10.1038/s41592-022-01488-1
103. Zhang C., Zhang C., Shang T., et al. (2024). HighFold: accurately predicting structures of cyclic peptides and complexes with head-to-tail and disulfide bridge constraints. *Brief Bioinform* **25**. DOI:10.1093/bib/bbae215.
104. Rettie S. A., Campbell K. V., Bera A. K., et al. (2023). Cyclic peptide structure prediction and design using AlphaFold. *bioRxiv*. DOI:10.1101/2023.02.25.529956.
105. Lee J. S. (2025). Design of peptides with non-canonical amino acids using flow matching. *the preprint server for biology*. DOI:10.1101/2025.07.31.667780.
106. Mao J., Guan S., Chen Y., et al. (2023). Application of a deep generative model produces novel and diverse functional peptides against microbial resistance. *Comput Struct Biotechnol J* **21**:463–471. DOI:10.1016/j.csbj.2022.12.029
107. Watson J. L., Juergens D., Bennett N. R., et al. (2023). De novo design of protein structure and function with RFdiffusion. *Nature* **620**:1089–1100. DOI:10.1038/s41586-023-06415-8
108. Wan F., Kontogiorgos-Heintz D. and de la Fuente-Nunez C. (2022). Deep generative models for peptide design. *Digit Discov* **1**:195–208. DOI:10.1039/d1dd00024a
109. Zhang O., Zhang X., Lin H., et al. (2025). ODesign: A World Model for Biomolecular Interaction Design.
110. Pacesa M., Nickel L., Schellhaas C., et al. (2025). One-shot design of functional protein binders with BindCraft. *Nature* **646**:483–492. DOI:10.1038/s41586-025-09429-6
111. Chen T., Dumas M., Watson R., et al. (2024). PepMLM: Target Sequence-Conditioned Generation of Therapeutic Peptide Binders via Span Masked Language Modeling. *ArXiv*.
112. Nijkamp E., Ruffolo J. A., Weinstein E. N., et al. (2023). ProGen2: Exploring the boundaries of protein language models. *Cell Syst* **14**:968–978.e963. DOI:10.1016/j.cels.2023.10.002
113. Alamdari S., Thakkar N., van den Berg R., et al. (2023). Protein generation with evolutionary diffusion: sequence is all you need. *bioRxiv*:2023.2009.2011.556673. DOI:10.1101/2023.09.11.556673.
114. Dauparas J., Anishchenko I., Bennett N., et al. (2022). Robust deep learning-based protein sequence design using ProteinMPNN. *Science* **378**:49–56. DOI:10.1126/science.ade2187
115. Hsu C., Verkuil R., Liu J., et al. (2022). Learning inverse folding from millions of predicted structures. *bioRxiv*:2022.2004.2010.487779. DOI:10.1101/2022.04.10.487779.
116. Dauparas J., Lee G. R., Pecoraro R., et al. (2025). Atomic context-conditioned protein sequence design using LigandMPNN. *Nat Methods* **22**:717–723. DOI:10.1038/s41592-025-02626-1
117. Butcher J., Krishna R., Mitra R., et al. (2025). De novo Design of All-atom Biomolecular Interactions with RFdiffusion3. *bioRxiv*. DOI:10.1101/2025.09.18.676967.
118. Cao L., Goreshtnik I., Coventry B., et al. (2020). De novo design of picomolar SARS-CoV-2 miniprotein inhibitors. *Science* **370**:426–431. DOI:10.1126/science.abd9909
119. Gautam A., Chaudhary K., Kumar R., et al. (2013). In silico approaches for designing highly effective cell penetrating peptides. *J Transl Med* **11**:74. DOI:10.1186/1479-5876-11-74
120. Manavalan B., Subramaniyam S., Shin T. H., et al. (2018). Machine-Learning-Based Prediction of Cell-Penetrating Peptides and Their Uptake Efficiency with Improved Accuracy. *J Proteome Res* **17**:2715–2726. DOI:10.1021/acs.jproteome.8b00148
121. Mathur D., Singh S., Mehta A., et al. (2018). In silico approaches for predicting the half-life of natural and modified peptides in blood. *PLoS One* **13**:e0196829. DOI:10.1371/journal.pone.0196829
122. Gupta S., Kapoor P., Chaudhary K., et al. (2013). In silico approach for predicting toxicity of peptides and proteins. *PLoS One* **8**:e73957. DOI:10.1371/journal.pone.0073957
123. Abraham M. J., Murtola T., Schulz R., et al. (2015). GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1-2**:19–25. DOI:https://doi.org/10.1016/j.softx.2015.06.001.
124. Ding X., Chen X., Sullivan E. E., et al. (2024). Fast, accurate ranking of engineered proteins by target-binding propensity using structure modeling. *Mol Ther* **32**:1687–1700. DOI:10.1016/j.ymthe.2024.04.003
125. Chang L. and Perez A. (2023). Ranking Peptide Binders by Affinity with AlphaFold. *Angew Chem Int Ed Engl* **62**:e202213362. DOI:10.1002/anie.202213362
126. Abramson J., Adler J., Dunger J., et al. (2024). Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* **630**:493–500. DOI:10.1038/s41586-024-07487-w
127. Feller A. L. and Wilke C. O. (2024). Peptide-specific chemical language model successfully predicts membrane diffusion of cyclic peptides. *bioRxiv*:2024.2008.2009.607221. DOI:10.1101/2024.08.09.607221.
128. Liu Y., Wen W., Xu B., et al. (2025). Molecular dynamics insights into the binding interactions of semaglutide with human serum albumin. *J Biomol Struct Dyn*:1–18. DOI:10.1080/07391102.2025.2532848.
129. Knudsen L. B. and Lau J. (2019). The Discovery and Development of Liraglutide and Semaglutide. *Front Endocrinol (Lausanne)* **10**:155. DOI:10.3389/fendo.2019.00155
130. Johns D. G., Campeau L. C., Banka P., et al. (2023). Orally Bioavailable Macrocyclic Peptide That Inhibits Binding of PCSK9 to the Low Density Lipoprotein Receptor. *Circulation* **148**:144–158. DOI:10.1161/circulationaha.122.063372
131. Fang P., Pang W. K., Xuan S., et al. (2024). Recent advances in peptide macrocyclization strategies. *Chem Soc Rev* **53**:11725–11771. DOI:10.1039/d3cs01066j

FUNDING AND ACKNOWLEDGMENTS

This work was supported by the Guangdong Basic and Applied Basic Research Foundation (2024A1515030013), State Key Laboratory of Southwestern Chinese Medicine Resources Foundation (FSKLTTCM202312). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. The authors have no further acknowledgements to declare.

AUTHOR CONTRIBUTIONS

All authors contributed to the conception and design of this study. W.H. and Y.S. conducted the reference search and write the manuscript. T.L. and M.W. performed the quality assessment of the literature. The initial draft was written by W.H. and Y.S., while X.H. and M.W. edited and polished the final manuscript. This study was conducted under the guidance of X.H. and M.W. All authors contributed to the manuscript and approved the final version.

DECLARATION OF INTERESTS

The authors declare no competing interests.

DATA AND CODE AVAILABILITY

Data are available from the corresponding author upon reasonable request.

SUPPLEMENTAL INFORMATION

It can be found online at <https://doi.org/10.59717/j.xinn-drugdisc.2026.100009>.