



DIFFUSION MODELS FOR MACROCYCLIC PEPTIDE DESIGN USING STABILITY, PERMEABILITY, AND BINDING CONSTRAINTS

Patrick O'Connor^{1*}, Sean Murphy¹

1. *Department of Pharmaceutical Informatics and Analytics, Faculty of Pharmacy, Trinity College Dublin, Dublin, Ireland.*

ARTICLE INFO

Received:

10 January 2026

Received in revised form:

17 March 2026

Accepted:

23 March 2026

Available online:

28 April 2026

Keywords: Diffusion models, Macrocyclic peptides, Peptide therapeutics, Molecular generation, Permeability, Conformational stability

ABSTRACT

Macrocyclic peptides can address challenging therapeutic targets that are often poorly modulated by conventional small molecules or biologics, but their design is difficult because conformational preorganization, membrane permeability, and target binding are interdependent and sometimes competing requirements. Existing generative approaches for peptides often focus on sequence novelty or target affinity without fully accounting for the geometric constraints imposed by cyclization, highlighting the need for models that reason jointly over topology, conformation, and medicinal chemistry properties. This article proposes a diffusion-based generative framework for designing macrocyclic peptides conditioned on predicted conformational stability, membrane permeability, and binding affinity, enabling the generation of candidate macrocycles that satisfy multiple design criteria within a single generative process. The model operates on cyclization-aware three-dimensional coordinates or torsion-angle representations, with property predictors guiding denoising toward molecular structures that exhibit favorable therapeutic profiles. Conceptually, this approach can produce synthetically plausible macrocyclic peptides with desirable permeability, target engagement, and conformational preferences, which should then be evaluated through computational filters, molecular dynamics simulations, and prospective experimental testing before being considered as drug leads. By integrating structural generation with pharmacokinetic and pharmacodynamic constraints, multi-constraint diffusion generation offers a model-oriented strategy for efficiently exploring drug-like macrocyclic peptide chemical space and accelerating the rational design of constrained peptide therapeutics.

This is an open-access article distributed under the terms of the [Creative Commons Attribution-Non Commercial-Share Alike 4.0 License](https://creativecommons.org/licenses/by-nc-sa/4.0/), which allows others to remix, and build upon the work non-commercially.

To Cite This Article: O'Connor P, Murphy S. Diffusion Models for Macrocyclic Peptide Design Using Stability, Permeability, and Binding Constraints. *Pharmacophore*. 2026;17(2):54-62. <https://doi.org/10.51847/x5pK88IN9y>

Introduction

Macrocyclic peptides occupy a valuable therapeutic space because they can present large, structured recognition surfaces while retaining opportunities for chemical optimization beyond conventional linear peptides. Their promise is especially relevant for protein-protein interactions and other difficult targets, but the same structural features that support potency can also impair passive permeability and oral exposure [1]. Design therefore requires balancing target binding, conformational stability, and transport-relevant properties rather than optimizing a single endpoint. Reviews of cyclic peptide drug discovery emphasize that N-methylation, intramolecular hydrogen bonding, macrocycle size, and conformational behavior all contribute to whether a candidate can become drug-like [2, 3].

Computational macrocycle design has traditionally relied on conformational sampling, docking, molecular dynamics, and expert-guided medicinal chemistry to evaluate proposed structures after they have been enumerated. These approaches are essential because macrocycles can adopt environment-dependent conformations that influence permeability and binding, making a single static structure an incomplete design object [4]. Molecular simulations have clarified how solvent exposure, hydrogen-bond rearrangement, and chameleonic conformational behavior shape macrocycle properties [5, 6]. However, post-hoc evaluation alone does not directly generate new macrocyclic chemical matter that is already biased toward a combined stability, permeability, and affinity profile.

Diffusion models provide a flexible generative framework in which molecular structures can be produced through iterative denoising from random noise. Denoising diffusion probabilistic models established a general route for learning complex data distributions [7], and molecular adaptations have extended this principle to equivariant three-dimensional generation [8], molecular conformer generation [9], torsional diffusion [10], and discrete graph diffusion [11]. These developments are

Corresponding Author: Patrick O'Connor; Department of Pharmaceutical Informatics and Analytics, Faculty of Pharmacy, Trinity College Dublin, Dublin, Ireland. E-mail: patrick.oconnor@gmail.com

relevant to macrocyclic peptides because the design problem is inherently geometric and topological. Recent peptide and structure-based diffusion methods suggest that generative models can be conditioned by molecular context and structural constraints rather than limited to unconstrained sequence generation [12-14].

This article frames macrocyclic peptide design as a conditional diffusion problem in which the model proposes cyclized peptide structures while explicitly accounting for conformational stability, permeability, and binding. The central thesis is that property-guided denoising could unify generative chemistry with pharmacokinetic and pharmacodynamic design by steering samples toward macrocycles that are both structurally plausible and therapeutically relevant. Such a framework builds on geometric latent diffusion for three-dimensional molecules [15], full-atom peptide diffusion [13], and emerging de novo macrocycle design approaches that connect deep learning with protein-binding macrocycles [16]. The goal is not to claim experimental performance, but to define a conceptual model that should be evaluated rigorously before practical deployment.

Background

Macrocyclic Peptides as Therapeutics

Macrocyclic peptides are distinguished by ring-constrained backbones that can reduce entropic penalties upon binding and stabilize bioactive conformations. Their oral bioavailability and permeability are strongly influenced by conformational shielding, intramolecular hydrogen bonds, N-methylation patterns, and the ability to alter polarity across environments [1, 2]. Chameleonic macrocycles can expose polar groups in aqueous settings while hiding them in membrane-like environments, a behavior that has been studied through solution conformational analysis and simulation [5, 6]. Because these properties emerge from three-dimensional structure rather than sequence alone, macrocyclic peptide design requires representations that preserve geometry, topology, and environment-dependent conformational preferences.

Generative Models for Peptide Design

Earlier molecular generative strategies, including reinforcement learning and sequence-based design, showed that machine learning can explore chemical space under optimization pressure. Deep reinforcement learning for molecular de novo design demonstrated how property-oriented objectives could guide molecular generation [17], while benchmarking efforts such as MOSES highlighted the need to evaluate validity, novelty, uniqueness, and distributional realism in generated molecules [18]. For macrocyclic peptides, however, simple string or graph generation can struggle to enforce ring closure, stereochemistry, backbone geometry, and realistic conformational ensembles. Transformer-based macrocyclic peptide generation illustrates the value of learned peptide representations [19], but diffusion models offer a complementary route because they can generate structures through gradual geometric refinement.

Diffusion Models for Molecular Generation

Diffusion models learn to reverse a noise process, enabling samples to be generated by progressive denoising rather than by direct one-step prediction. In molecular design, this framework has been adapted to equivariant coordinate models that respect three-dimensional symmetries [8], conformer generation methods that model molecular geometry [9], and torsional diffusion models that focus on internal degrees of freedom [10]. Discrete denoising diffusion extends the same logic to graph structures, which is important when atoms and bonds are generated or modified [11]. For macrocyclic peptides, these paradigms suggest that cyclic backbones could be represented through coordinated graph, torsion, and three-dimensional geometric diffusion processes.

Property Prediction for Macrocyclic Peptides

Property prediction for macrocyclic peptides requires models that connect structure with permeability, conformational stability, and binding behavior. CycPeptMPDB provides a curated resource for cyclic peptide membrane permeability, making it relevant for training or calibrating permeability predictors [20], while CycPeptMP integrates multi-level molecular features to support permeability modeling [21]. Conformational stability can be estimated conceptually from molecular dynamics-derived features such as structural persistence, solvent exposure, and hydrogen-bond satisfaction, which have been used to interpret macrocycle behavior in different environments [4, 22]. Binding can be assessed through docking, structure-based design, or affinity-oriented predictors, but these estimates should be treated as surrogate signals rather than experimental truth [23, 24].

Conditional Generation and Multi-Objective Diffusion

Conditional generation aims to steer a generative model toward desired property profiles instead of sampling only from the average training distribution. In molecular diffusion, conditioning may be introduced through property embeddings, classifier guidance, classifier-free guidance, or target-structure context, depending on whether the model is generating graphs, conformers, or protein-bound ligands [11, 12]. Multi-objective macrocyclic peptide design is especially demanding because improving target binding may reduce permeability, while increasing conformational rigidity may either help or hinder transport depending on the exposed polar surface. Recent work on preference-based optimization for permeability-optimized target-binding macrocycles illustrates the conceptual importance of balancing multiple macrocycle objectives rather than treating each property independently [25].

Model Development Overview

High-Level Generation Pipeline

The proposed pipeline begins with a diffusion model trained on known macrocyclic peptide structures so that it learns a distribution over cyclized peptide geometries rather than unconstrained peptide sequences. During generation, a desired property profile would be encoded as a conditioning signal containing stability, permeability, and binding objectives. The reverse denoising process would then transform a noisy initial representation into a candidate macrocycle whose topology and conformation are consistent with the learned macrocyclic peptide manifold. This design follows the general denoising principle introduced for diffusion models [7] while drawing on molecular coordinate diffusion methods for three-dimensional generation [8, 15].

Core Input Representations

The model could represent each peptide through internal coordinates, such as backbone torsion angles constrained by ring closure, or through full three-dimensional atomic coordinates with explicit cyclization bonds. Torsional representations are attractive because they describe chemically meaningful degrees of freedom and have been effective for molecular conformer generation [10], whereas coordinate-based models can learn equivariant geometric relationships directly [8, 9]. For peptides, full-atom diffusion is especially relevant because side-chain placement, backbone conformation, and cyclization geometry jointly determine binding and permeability [13]. Property predictors would consume the same or related representations to estimate continuous conditioning signals for stability, permeability, and affinity.

Design Principles

A macrocyclic peptide diffusion model must enforce topological consistency so that generated structures remain cyclized and chemically interpretable. It must also support multi-property conditioning, because drug-like macrocycles require simultaneous control over conformation, transport-related polarity, and target recognition rather than sequential optimization of isolated traits [1, 26]. Diversity is equally important, since the model should explore alternative ring sizes, side-chain patterns, and conformational families without collapsing onto known scaffolds. These principles align with molecular generation benchmarks that emphasize validity, novelty, uniqueness, and distributional quality as core evaluation dimensions [18].

Table 1 defines how each macrocyclic peptide design constraint can be translated into representation signals, diffusion-guidance functions, and medicinal chemistry interpretation.

Table 1. Design-constraint logic for conditional diffusion generation of macrocyclic peptides

Design constraint	Macrocyclic peptide design meaning	Candidate representation signal	Role during diffusion denoising	Main trade-off created	Practical interpretation for medicinal chemistry
Cyclization topology	Ensures that the generated molecule remains a true macrocycle rather than a linear or chemically incoherent peptide	Explicit ring-closing bond, cyclic graph connectivity, constrained torsion-angle geometry, stereochemical encoding	Restricts the denoising trajectory to chemically interpretable macrocyclic structures	Stronger topological control may reduce generative diversity if enforced too rigidly	Generated candidates should be checked first for ring closure, stereochemical plausibility, and feasible macrocyclization chemistry
Conformational stability	Captures whether the macrocycle adopts persistent, design-relevant conformations rather than highly unstable ensembles	Backbone torsion persistence, intramolecular hydrogen bonding, solvent exposure, conformer clustering, MD-derived stability descriptors	Biases denoising toward preorganized structures with stable conformational preferences	Excessive rigidity may improve binding preorganization but reduce adaptability or permeability	Stable conformations should be interpreted as design hypotheses requiring MD and experimental structural confirmation
Membrane permeability	Represents the likelihood that the macrocycle can cross biological membranes despite peptide polarity and size	Predicted permeability score, polar surface shielding, N-methylation pattern, intramolecular hydrogen-bonding profile, chameleonic behavior markers	Guides sampling toward conformations and substituent patterns associated with improved transport	Permeability optimization may conflict with exposed polar groups required for target recognition	Permeability-guided candidates should be prioritized for assay testing rather than accepted from prediction alone
Target binding	Represents compatibility between the macrocycle and a	Docking score, target-context embedding, protein-peptide interface	Conditions denoising toward structures	Maximizing predicted binding may produce polar, rigid, or bulky	Binding scores should be used as ranking signals, not

	target pocket, protein interface, or binding conformation	geometry, affinity surrogate, pharmacophore alignment	compatible with target engagement	structures with poor permeability or synthesis feasibility	as proof of biological activity
Synthetic accessibility	Assesses whether the generated macrocycle can realistically be made, purified, and optimized	Ring strain estimate, residue availability, non-canonical residue burden, protecting-group complexity, cyclization feasibility	Can be applied as an auxiliary guidance signal or post-generation filter	Synthesis-aware filtering may discard novel structures that are computationally attractive but difficult to realize	Chemist review is essential before advancing generated structures to synthesis
Diversity and novelty	Prevents the model from reproducing only known scaffolds or collapsing into narrow macrocycle families	Scaffold similarity, ring-size distribution, side-chain diversity, uniqueness, distance from training structures	Encourages exploration of alternative macrocycle families during sampling	High novelty may increase uncertainty in property predictors and synthesis feasibility	Novel candidates should be advanced only when novelty is paired with interpretable constraint satisfaction
Experimental falsifiability	Ensures that outputs can be tested through real assays and not only through computational scoring	Defined assay plan, measurable permeability endpoint, binding assay, structural characterization, MD validation pathway	Does not directly shape denoising unless encoded as a design objective, but governs candidate selection	Computationally attractive candidates may fail when exposed to assay constraints	The final output should be a prioritized experimental hypothesis, not a declared drug lead

Macrocyclic Peptide Representations and Data

Dataset Compilation

A training corpus for the proposed model would combine macrocyclic peptide structures from public structural repositories, curated permeability databases, and literature-derived medicinal chemistry reports. CycPeptMPDB is directly relevant because it links cyclic peptide structures with membrane permeability information [20], while permeability prediction models based on multi-level molecular descriptors can provide surrogate labels when experimental values are unavailable [21]. Structural examples from protein-bound macrocycles and computationally designed peptide macrocycle inhibitors can enrich the model with target-recognition geometries [16, 24]. Molecular dynamics and enhanced conformational sampling could augment each peptide with plausible conformers, reflecting the observation that macrocycles may adopt distinct conformations across solvent environments [4, 22].

Cyclization-Aware Representation

Cyclization-aware representation requires the ring-closing bond to be encoded explicitly rather than treated as a post-processing correction. Internal-coordinate schemes can incorporate ring-closure constraints through torsion angles and bond geometry, while three-dimensional coordinate models can represent the macrocycle as an equivariant structure whose atoms and bonds must remain geometrically compatible [8, 10]. This is important because macrocycle conformational sampling is challenging, and failure to capture the relevant conformations can distort predictions of both permeability and binding [4, 22]. Accurate structure prediction for cyclic peptides containing non-canonical residues further underscores the need to represent stereochemistry, ring topology, and all-atom geometry together [27].

Feature Engineering for Property Conditioning

Each training peptide would be associated with a property vector summarizing predicted or measured stability, permeability, and target binding. Permeability labels could be derived from curated cyclic peptide measurements and learned feature models [20, 21], while conformational descriptors could be extracted from simulation analyses of hydrogen bonding, solvent exposure, and structural persistence [5, 6]. Binding labels could come from docking, structure-based macrocycle design, or target-specific protein-peptide complex modeling, with the understanding that such labels are approximations rather than definitive biological outcomes [23, 24]. These property vectors would be embedded into the diffusion model so that denoising is guided by medicinal chemistry objectives rather than by structural realism alone.

Diffusion Model Architecture

Forward and Reverse Process for Macrocycle Coordinates

In the forward process, noise would be added gradually to the macrocyclic peptide representation, either by perturbing three-dimensional coordinates or by corrupting torsion-angle states. The reverse process would learn to denoise these perturbed structures using an equivariant neural network, graph model, or transformer that respects molecular geometry and peptide connectivity [8, 11]. Geometric diffusion and conformer diffusion methods show how molecular structures can be reconstructed through learned denoising of coordinates or conformational degrees of freedom [9, 15]. For macrocyclic

peptides, the reverse model would also need to preserve ring closure so that the denoised sample remains a chemically valid macrocycle.

Conditioning Mechanism

The conditioning mechanism would inject stability, permeability, and binding objectives into the denoising network at each reverse step. This could be implemented through feature-wise modulation, cross-attention to property embeddings, or guidance from auxiliary predictors trained on macrocycle-relevant labels. Structure-based diffusion models demonstrate how generative sampling can be shaped by binding-site context [12], while peptide binder diffusion indicates how geometric generation can be adapted to protein-interaction design [14]. For macrocyclic peptides, the same principle would be extended to a multi-constraint setting in which the model is guided toward conformationally stable, permeable, and target-compatible structures. **Figure 1** illustrates the proposed conditional diffusion architecture for generating macrocyclic peptide candidates under simultaneous conformational stability, permeability, and target-binding constraints.

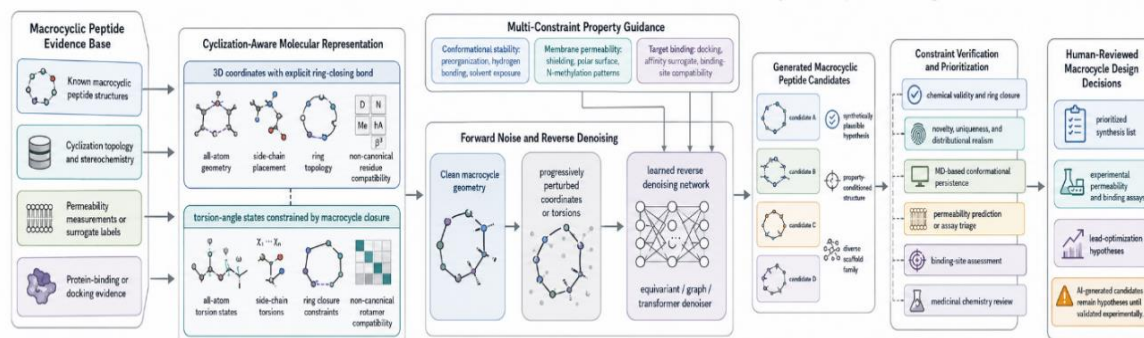


Figure 1. Conditional diffusion architecture for multi-constraint macrocyclic peptide design.

Training and Sampling

Training would optimize the model to recover clean macrocyclic peptide structures from noisy versions while learning how property conditions relate to the denoising trajectory. The score-matching logic underlying diffusion models [7] can be paired with torsional or coordinate representations that are better suited to molecular geometry [9, 10]. At inference, the model would begin from noise and progressively refine a candidate structure under a specified property profile, conceptually producing macrocycles that satisfy design intentions before downstream filtering. Because learned generation can still propose unrealistic molecules, evaluation should include molecular-generation quality checks such as validity, novelty, and uniqueness alongside macrocycle-specific constraints [18].

Incorporating Stability, Permeability, And Binding Constraints

Stability Constraint

A stability constraint would encourage the model to generate macrocycles whose conformations are preorganized for the intended binding mode while remaining physically plausible in solution. This signal could be derived from molecular dynamics or enhanced sampling descriptors that capture persistence of backbone geometry, intramolecular hydrogen bonding, and solvent-exposure patterns, all of which are central to macrocycle conformational behavior [4, 5]. Strategies for tuning cyclic peptide conformations show that small chemical changes can strongly alter the accessible conformational ensemble [28]. In a diffusion framework, such descriptors would act as conditioning variables or guidance signals that bias denoising toward structures with stable, design-relevant conformational preferences.

Permeability Constraint

A permeability constraint would guide generation toward macrocycles predicted to cross membranes more effectively while retaining the structural features required for binding. Cyclic peptide permeability is closely linked to conformational shielding, polar surface presentation, and substituent patterns such as N-methylation, and curated permeability resources provide a basis for learning these relationships [1, 20]. Predictive models that integrate multiple molecular feature levels can support permeability-aware conditioning, although their predictions should be interpreted as surrogate design signals rather than definitive experimental outcomes [21]. Within the diffusion model, the permeability score would be embedded into the denoising process so that candidate structures are sampled from regions of macrocycle space expected to have favorable transport behavior.

Binding Affinity Constraint

A binding affinity constraint would condition the model on a target-specific signal derived from docking, structure-based scoring, or free-energy-inspired surrogate prediction. Structure-based macrocycle design has shown how macrocyclization can

improve ligand organization and target engagement when applied to appropriate binding geometries [23]. Computationally designed macrocyclic inhibitors and deep learning approaches for protein-binding macrocycles further demonstrate that target recognition can be incorporated into macrocycle design workflows [16, 24]. In the proposed diffusion model, the affinity condition would be balanced with stability and permeability so that generation does not simply maximize predicted binding at the expense of drug-like macrocycle behavior.

Evaluating Generated Macrocycles

In-Silico Quality Metrics

Generated macrocycles should first be evaluated for chemical and topological validity, including whether the ring closure is geometrically consistent and whether stereochemistry is interpretable. General molecular generation benchmarks emphasize validity, novelty, uniqueness, and distributional similarity as core quality dimensions [18], but macrocyclic peptides also require cyclization-aware checks that assess backbone closure, conformational plausibility, and compatibility with non-canonical residues. Discrete graph diffusion and geometric diffusion models provide useful precedents for evaluating whether generated structures obey molecular constraints in graph and coordinate spaces [8, 11]. Constraint satisfaction should therefore be described qualitatively as the degree to which generated candidates align with target profiles for stability, permeability, and binding, rather than as unsupported numerical success rates.

Prospective Validation with MD and Synthesis

After computational filtering, selected candidates should undergo molecular dynamics simulations to assess whether their predicted conformations remain stable under relevant solvent or binding conditions. MD-assisted cyclic peptide structure prediction illustrates how simulation and machine learning can be combined to understand conformational preferences [29], while macrocycle simulation studies show that environment-dependent conformational switching can be critical for interpreting permeability and binding [6, 22]. Synthesis and experimental testing would then be needed to determine whether the proposed designs actually show the intended permeability and target engagement. This prospective validation step is essential because diffusion-generated structures remain hypotheses until confirmed by physical assays.

Integration Into Macrocyclic Drug Discovery Workflow

Hit Expansion and Lead Optimization

In hit expansion, the model could start from a known macrocyclic peptide scaffold and generate chemically related analogues that preserve key recognition elements while exploring alternative side chains, stereochemical patterns, and ring conformations. This use case aligns with the broader medicinal chemistry role of macrocycles as tunable scaffolds for difficult targets [3, 26]. Deep reinforcement learning has previously shown how molecular generation can be guided by design objectives [17], and diffusion provides a structure-oriented alternative in which analogues are refined through denoising rather than sequential action selection. The resulting candidates would enter a lead-optimization loop where medicinal chemists assess whether predicted improvements are chemically meaningful and experimentally testable.

Target-Specific and Pan-Target Design

For target-specific design, the diffusion model could condition on a binding site, peptide-target complex geometry, or target-derived affinity predictor. Structure-based equivariant diffusion models show that target context can guide molecular generation [12], and peptide binder diffusion extends this idea toward protein-interaction settings [14]. In a target-agnostic mode, the model could instead generate macrocycles biased toward general drug-like properties such as permeability, conformational shielding, and structural diversity, which are valuable for screening collections [1, 25]. These two modes would support complementary workflows: focused optimization around a biological target and broader exploration of macrocycle chemical space.

Evaluation Strategy

Generation Quality and Constraint Satisfaction

The evaluation strategy should examine whether generated molecules are valid macrocycles, whether they are novel relative to the training set, and whether their predicted properties align with the requested conditioning profile. Molecular generation benchmarks such as MOSES provide a conceptual foundation for evaluating generated chemical sets through validity, uniqueness, novelty, and distributional realism [18]. For macrocyclic peptides, these criteria should be extended to include ring topology, stereochemical consistency, conformational plausibility, and agreement between conditioned and predicted property profiles. Results should be expressed conceptually and comparatively, avoiding unsupported numerical claims about generation counts or success rates.

Retrospective Validation

Retrospective validation would test whether the model can reproduce or rediscover known drug-like macrocycles when conditioned on their approximate property profiles. This assessment is relevant because known macrocyclic drugs and cyclic peptide natural products define useful examples of the permeability and bioavailability frontier [1]. Held-out macrocycles from curated permeability resources and structural datasets could be used to ask whether the model samples structures consistent

with observed macrocycle behavior [20, 21]. Such validation would not prove prospective utility, but it would reveal whether the learned diffusion process captures recognizable relationships among cyclization, conformation, and medicinal chemistry properties.

Prospective Experimental Hit Rate

Prospective evaluation would require synthesis and testing of selected AI-designed macrocycles, but the present article treats this step as a proposed validation strategy rather than a reported experiment. Candidate selection should combine predicted permeability, target binding, conformational stability, and medicinal chemistry review, reflecting the multi-constraint design logic used throughout the framework [16, 25]. Experimental assays could then examine whether designed structures show the expected permeability and binding behavior, while MD simulations could help interpret discrepancies between predicted and observed properties [6, 29]. The most important outcome would be whether the workflow generates experimentally informative hypotheses for macrocycle optimization, not whether a single computational score appears favorable.

Table 2 provides a staged validation framework for interpreting diffusion-generated macrocycles as experimentally testable hypotheses rather than confirmed therapeutic leads.

Table 2. Validation framework for moving diffusion-generated macrocycles from computational candidates to experimental hypotheses

Validation layer	Primary question addressed	Recommended assessment	Failure mode detected	Decision consequence
Chemical and topological validity	Is the generated structure a chemically interpretable macrocyclic peptide?	Check ring closure, valence consistency, stereochemistry, residue identity, bond geometry, and macrocycle topology	Linearized peptide, broken ring, impossible valence, ambiguous stereochemistry, unrealistic bond geometry	Remove invalid structures before property scoring or medicinal chemistry review
Representation consistency	Does the generated molecule remain coherent across graph, coordinate, and torsion representations?	Compare cyclic graph structure with 3D coordinates and torsion-angle reconstruction	A structure appears valid in one representation but violates closure or geometry in another	Retain only candidates that remain valid across the model's representational layers
Constraint satisfaction	Does the candidate match the requested stability, permeability, and binding profile?	Re-score generated candidates with independent property predictors and compare predicted outputs with conditioning targets	Candidate satisfies one objective while failing the others; denoising overfits to a single property	Prioritize balanced candidates rather than those with one extreme predicted score
Conformational robustness	Does the candidate maintain plausible conformations under simulated conditions?	Molecular dynamics, enhanced sampling, conformer ensemble analysis, hydrogen-bond persistence, solvent-exposure profiling	Predicted favorable conformation collapses, switches unpredictably, or exposes unfavorable polarity	Advance only candidates with interpretable conformational behavior or clear testable uncertainty
Permeability plausibility	Is predicted transport behavior consistent with macrocycle physicochemical logic?	Evaluate predicted permeability, polar shielding, N-methylation pattern, intramolecular hydrogen bonding, and chameleonic features	Predicted permeability is driven by model artifact rather than chemically interpretable features	Select candidates for permeability assay only when predictions align with structural rationale
Target-engagement plausibility	Is the macrocycle compatible with the intended binding site or protein interface?	Docking, protein-peptide complex modeling, pharmacophore review, interface contact analysis, binding-site steric assessment	Predicted affinity reflects unrealistic pose, steric clash, or unsupported scoring artifact	Use binding evidence as a triage signal requiring biochemical confirmation
Synthetic tractability	Can the candidate reasonably be synthesized and optimized?	Medicinal chemistry review, residue availability, ring strain inspection, cyclization route assessment, non-canonical residue burden	Attractive generated structure is impractical, unstable, or too complex to synthesize	Remove or redesign candidates before experimental allocation
Prospective experimental validation	Does the generated macrocycle show the expected behavior in real assays?	Synthesis, permeability testing, binding assay, structural characterization, and iterative design feedback	Computational predictions fail under experimental conditions	Treat outcomes as feedback for model recalibration, not as simple success or failure

Limitations

Reliance on Surrogate Property Predictors

The proposed framework depends heavily on surrogate predictors for stability, permeability, and binding, and each predictor may be uncertain outside the chemical space used for training. Permeability models built from curated cyclic peptide data are valuable, but they can still be sensitive to representation choices, molecular feature coverage, and the scarcity of comparable measurements [20, 21]. Stability signals from MD or conformational sampling may also depend on force-field assumptions and sampling completeness, particularly for flexible or chameleonic macrocycles [4, 22]. Binding predictors introduce additional uncertainty because docking and structure-based scores may not fully capture solvent effects, induced fit, or the thermodynamics of macrocyclic peptide recognition [23].

Synthetic Accessibility

A diffusion model can be designed to preserve cyclization topology, but that does not guarantee that the generated macrocycle is synthetically accessible or practical to optimize. Macrocyclic peptide synthesis may be complicated by ring strain, protecting-group strategy, stereochemical complexity, non-canonical residues, and poor cyclization efficiency, all of which can make apparently attractive designs difficult to realize. Macrocyclic chemical space is broad and structurally diverse, so generated candidates should be evaluated for medicinal chemistry tractability rather than accepted solely because they satisfy computational property constraints [26]. Future versions of the framework should include synthesis-aware scoring or retrosynthetic feasibility checks alongside stability, permeability, and binding guidance.

Conclusion

A diffusion model for macrocyclic peptide design could provide a unified generative framework for proposing cyclized peptide structures under multiple medicinal chemistry constraints. By operating on torsion-angle or three-dimensional coordinate representations, such a model could explicitly account for the geometry and topology that make macrocyclic peptides distinct from linear peptides and small molecules. The central idea is to guide denoising with stability, permeability, and binding signals so that candidate structures are generated within a therapeutically relevant design space.

The major strength of this approach is simultaneous optimization of properties that are usually evaluated sequentially. Instead of generating peptide structures first and applying permeability, stability, and binding filters afterward, conditional diffusion could incorporate these objectives during molecular construction. This makes the framework well suited to macrocyclic peptides, where conformational preorganization, transport behavior, and target recognition are deeply interconnected.

Important challenges remain before such a model could be considered reliable for drug discovery decisions. Surrogate property predictors may be inaccurate for unusual macrocycles, generated structures may be difficult to synthesize, and computationally favorable conformations may not persist under experimental conditions. Prospective validation is therefore essential, and model outputs should be treated as prioritized hypotheses rather than confirmed leads.

The next stage for the field should be an integrated computational-experimental campaign that benchmarks multi-constraint macrocycle generation against real synthesis, permeability testing, binding assays, and structural characterization. Such studies would clarify where diffusion models genuinely improve macrocyclic peptide discovery and where human medicinal chemistry judgment remains indispensable. If validated carefully, diffusion-based macrocycle design could become a powerful tool for accelerating next-generation peptide therapeutics.

Acknowledgments: None

Conflict of interest: None

Financial support: None

Ethics statement: None

References

1. Naylor MR, Bockus AT, Blanco MJ, Lokey RS. Cyclic peptide natural products chart the frontier of oral bioavailability in the pursuit of undruggable targets. *Curr Opin Chem Biol.* 2017;38:141-7.
2. Walport LJ, Obexer R, Suga H. Strategies for transitioning macrocyclic peptides to cell-permeable drug leads. *Curr Opin Biotechnol.* 2017;48:242-50.
3. Saunders GJ, Yudin AK. A focus on the discovery of potent and selective cyclic peptide scaffolds for drug development. *Chem Sci.* 2022;13(44):12942-4.
4. Poongavanam V, Danelius E, Peintner S, Alcaraz L, Caron G, Cummings MD, et al. Conformational sampling of macrocyclic drugs in different environments: can we find the relevant conformations? *ACS Omega.* 2018;3(9):11742-57.

5. Danelius E, Poongavanam V, Peintner S, Wieske LH, Erdélyi M, Kihlberg J. Solution conformations explain the chameleonic behaviour of macrocyclic drugs. *Chem Eur J*. 2020;26(23):5231-44.
6. Sethio D, Poongavanam V, Xiong R, Tyagi M, Duy Vo D, Lindh R, et al. Simulation reveals the chameleonic behavior of macrocycles. *J Chem Inf Model*. 2022;63(1):138-46.
7. Ho J, Jain A, Abbeel P. Denoising diffusion probabilistic models. *Adv Neural Inf Process Syst*. 2020;33:6840-51.
8. Hoogeboom E, Satorras VG, Vignac C, Welling M. Equivariant diffusion for molecule generation in 3D. In: *Proceedings of the 39th International Conference on Machine Learning*; 2022. PMLR. p. 8867-87.
9. Xu M, Yu L, Song Y, Shi C, Ermon S, Tang J. GeoDiff: a geometric diffusion model for molecular conformation generation. *arXiv:2203.02923*. 2022.
10. Jing B, Corso G, Chang J, Barzilay R, Jaakkola T. Torsional diffusion for molecular conformer generation. *Adv Neural Inf Process Syst*. 2022;35:24240-53.
11. Vignac C, Krawczuk I, Siraudin A, Wang B, Cevher V, Frossard P. DiGress: discrete denoising diffusion for graph generation. *arXiv:2209.14734*. 2022.
12. Schneuing A, Harris C, Du Y, Didi K, Jamasb A, Igashov I, et al. Structure-based drug design with equivariant diffusion models. *Nat Comput Sci*. 2024;4(12):899-909.
13. Kong X, Jia Y, Huang W, Liu Y. Full-atom peptide design with geometric latent diffusion. *Adv Neural Inf Process Syst*. 2024;37:74808-39.
14. Dong B, Li S, Hou C, Xu D. PPI-Diff: de novo generation of peptide binders via resolution-aware geometric diffusion. *Biomolecules*. 2026;16(4):528.
15. Xu M, Powers AS, Dror RO, Ermon S, Leskovec J. Geometric latent diffusion models for 3D molecule generation. In: *Proceedings of the 40th International Conference on Machine Learning*; 2023. PMLR. p. 38592-610.
16. Rettie SA, Juergens D, Adebomi V, Bueso YF, Zhao Q, Leveille AN, et al. Accurate de novo design of high-affinity protein-binding macrocycles using deep learning. *Nat Chem Biol*. 2025:1-9.
17. Olivecrona M, Blaschke T, Engkvist O, Chen H. Molecular de-novo design through deep reinforcement learning. *J Cheminform*. 2017;9(1):48.
18. Polykovskiy D, Zhebrak A, Sanchez-Lengeling B, Golovanov S, Tatanov O, Belyaev S, et al. Molecular sets (MOSES): a benchmarking platform for molecular generation models. *Front Pharmacol*. 2020;11:565644.
19. Xu X, Xu C, He W, Wei L, Li H, Zhou J, et al. HELM-GPT: de novo macrocyclic peptide design using generative pre-trained transformer. *Bioinformatics*. 2024;40(6):btae364.
20. Li J, Yanagisawa K, Sugita M, Fujie T, Ohue M, Akiyama Y. CycPeptMPDB: a comprehensive database of membrane permeability of cyclic peptides. *J Chem Inf Model*. 2023;63(7):2240-50.
21. Li J, Yanagisawa K, Akiyama Y. CycPeptMP: enhancing membrane permeability prediction of cyclic peptides with multi-level molecular features and data augmentation. *Brief Bioinform*. 2024;25(5):bbae417.
22. Tang X, Kokot J, Waibl F, Fernández-Quintero ML, Kamenik AS, Liedl KR. Addressing challenges of macrocyclic conformational sampling in polar and apolar solvents: lessons for chameleonicity. *J Chem Inf Model*. 2023;63(22):7107-23.
23. Cummings MD, Sekharan S. Structure-based macrocycle design in small-molecule drug discovery and simple metrics to identify opportunities for macrocyclization of small-molecule ligands. *J Med Chem*. 2019;62(15):6843-53.
24. Mulligan VK, Workman S, Sun T, Rettie S, Li X, Worrall LJ, et al. Computationally designed peptide macrocycle inhibitors of New Delhi metallo- β -lactamase 1. *Proc Natl Acad Sci USA*. 2021;118(12):e2012800118.
25. Sun H, Tan H, Chu Y, Li J, Wang R, Wei DQ. Design of permeability-optimized target-binding macrocycles via direct preference optimization. *Chem Sci*. 2026. [Epub ahead of print].
26. Viarengo-Baker LA, Brown LE, Rzepiela AA, Whitty A. Defining and navigating macrocycle chemical space. *Chem Sci*. 2021;12(12):4309-28.
27. Wu D, Zou Y. Accurate 3D structure prediction of small cyclic peptides containing non-canonical amino acid residues using an all-atom diffusion model with stereogenic implementation. *J Chem Inf Model*. 2026. [Epub ahead of print].
28. Jwad R, Weissberger D, Hunter L. Strategies for fine-tuning the conformations of cyclic peptides. *Chem Rev*. 2020;120(17):9743-89.
29. Miao J, Descoteaux ML, Lin YS. Structure prediction of cyclic peptides by molecular dynamics + machine learning. *Chem Sci*. 2021;12(44):14927-36.