

Medicinal Chemistry and Chemical Biology Highlights

Division of Medicinal Chemistry and Chemical Biology

A Division of the Swiss Chemical Society

Deep Learning-based Design of Peptide Binders

Martin Pacesa*

*Correspondence: Prof. M. Pacesa, E-mail: pacesa@pharma.uzh.ch
Institute of Pharmacology & Toxicology, University of Zurich, Winterthurerstrasse 190, CH-8057 Zurich.

Abstract: Peptides are short chains of amino acids that naturally mediate diverse biological functions, such as intercellular signalling, immune modulation, antimicrobial defence, and others. In recent years, peptides have also emerged as an attractive therapeutic modality. They combine best features of small molecules and biologics, enabling binding to previously undruggable molecular surfaces while remaining compact and amenable to scalable manufacturing. Yet designing peptide binders remains challenging because their interfaces are small, solvent exposed, and often highly dynamic. Recent advances in deep learning have begun to make this problem more tractable by enabling target-conditioned generation and prioritisation of peptide candidates at scale.

Keywords: Binders · Deep learning · Peptide design · Therapeutics

1. What are Peptides and Why are They Interesting

Firstly, by ‘peptide’ I refer to polypeptides shorter than 30 amino acids, to distinguish them from somewhat larger ‘mini-proteins’ (40–100 amino acids), for which current computational protein design pipelines are generally developed.^[1] This cut-off is arbitrary, but I nevertheless find it useful because around this length we start to observe properties that make peptides unique and inherently more challenging to design as binders. These molecules often lack a stable folded state in isolation, present only a limited interaction surface, and can no longer rely on the same degree of shape complementarity and distributed binding energy that larger proteins or antibodies routinely exploit.^[2]

Peptides naturally function as signalling ligands, hormones, neurotransmitters, antimicrobial effectors, trafficking motifs, toxins, and recognition elements.^[3] Their utilisation across the tree of life highlights their evolutionary value as ‘cheap innovators’ because short, flexible sequences can arise readily, be tuned by a small number of mutations, diffuse and turn over rapidly, and engage diverse targets through conformational plasticity.^[4] From a therapeutic perspective, peptides occupy a particularly attractive middle ground between small molecules and larger biologics. Small molecules primarily target hydrophobic pockets but struggle on broad or featureless protein surfaces.^[5] Antibodies, by contrast, can bind most targets with excellent affinity and specificity, but are costly to produce and limited to mainly extracellular applications.^[6]

Peptides can bridge this gap. They give us more atoms to work with and are large enough to engage extended protein surfaces, yet small enough to be chemically synthesised in a scalable way. They can also be tuned through cyclization, stapling, back-

bone modification, terminal capping, PEGylation, lipidation, and incorporation of noncanonical amino acids.^[3] Such modifications can drastically improve their protease resistance, conformational preorganization, pharmacokinetics, permeability, and selectivity.^[3] This shows how the peptide space can be expanded both chemically and structurally far beyond simple linear amino acid chains.

2. Properties and Challenges of Peptide Binders

Conceptually, peptide binder design is fundamentally harder than design of larger protein binders for several reasons (Fig. 1A). The first is simply the modality and interface size. Peptides have less surface area to form favourable molecular interactions, so each contact matters more and small modelling errors carry a larger relative energetic penalty.^[2] The second aspect is entropy. Many peptides are not strongly preorganized in the unbound state, meaning that productive binding must pay a substantial conformational penalty.^[7] The third is solvent exposure. Because peptide interfaces bury less surface area, solvent interactions compete more strongly with intermolecular contacts, meaning that each individual interaction contributes more critically to binding stability. Finally, peptide binding is frequently dynamic. What matters may not be one rigidly bound structure, but an ensemble of partially preorganized conformations, transient complexes, and induced-fit rearrangements.^[8] Current design pipelines do not currently account for this and tend to reduce this problem to a single predicted complex and a handful of confidence metrics.

Additionally, peptide binders can come in different forms (Fig. 1B). Some bind as amphipathic helices, others as β -hairpins, extended strands, or structured loops.^[2] Many can be largely disordered on their own and only become ordered upon binding.^[2] Some peptides can depend on rigidified cyclic or disulfide-constrained topologies to reduce entropic costs,^[3] while others can remain at least partly fuzzy even in the bound state.^[8] This diversity is often underappreciated in discussions of peptide design. ‘Peptide binder’ is not a single structural class, and a design framework that performs well on helical binders may fail on fuzzy interactors or on compact macrocycles. That heterogeneity is one reason why the field has advanced slower than one might have expected from recent successes in protein design more broadly.

Traditionally, peptide binders have been discovered by display or library-based screening approaches or mined from natural sources. These approaches can be effective, but they offer limited control over binding mode and are not always straightforward to optimize beyond the initial hit. More fundamentally, purely empirical screening is an inefficient way to traverse an enormous and chemically diverse space when the properties we care about, such as affinity and specificity, are highly target-, geometry-, and system-dependent. Computational design offers a more attractive route because it can bias the search toward peptides that are already shaped for a given surface, binding mode, or biochemical

Can you show us your Medicinal Chemistry and Chemical Biology Highlight?

Please contact: Prof. Dr. Nathalie Grob, Dept. of Chemistry and Applied Biosciences, ETH Zurich, Vladimir-Prelog-Weg 1–5/10, CH-8093 Zurich, E-mail: nathalie.grob@pharma.ethz.ch

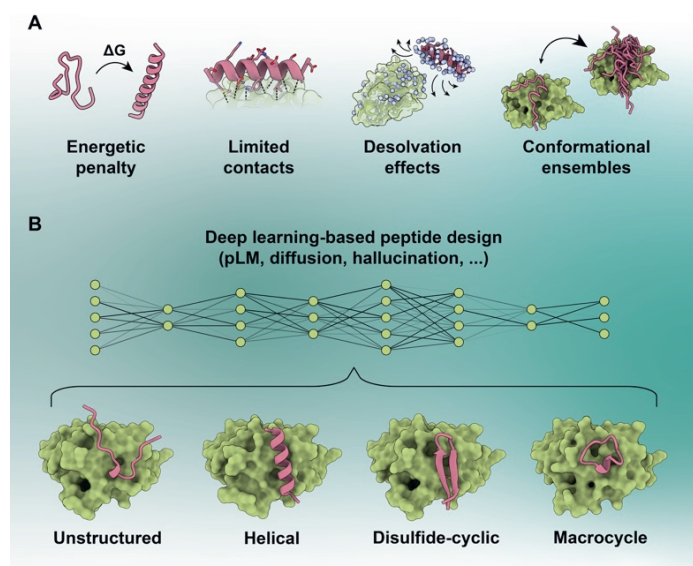


Fig. 1. Deep learning can help overcome some traditional challenges in the design of peptide binders. (A) Peptide binders face several intrinsic challenges compared with larger protein binders, including smaller interaction surfaces, entropic penalties associated with folding upon binding, stronger competition with solvent due to limited burial, and the prevalence of dynamic binding ensembles. Most of these biophysical constraints are not explicitly accounted for in present pipelines. (B) Deep learning pipelines can generate peptide binders to arbitrary targets from sequence or structural representations, producing diverse binding geometries.

and pharmaceutical properties. The real promise is not to replace experiments, but to make experimental screening smaller, more informed, and more rational from the outset.

This is where deep learning has changed the landscape. Previously, peptide design was constrained because candidate peptides had to be placed in a realistic binding geometry and then scored with enough accuracy to separate true binders from the overwhelming number of decoys.^[9] Deep learning changed this by learning transferable constraints on peptide-protein recognition from the data distribution of protein structures and sequence co-evolution, which makes it easier to generate plausible designs. Even so, reliably identifying true binders remains a central challenge, and many methods still rely on imperfect prediction and filtering oracles.^[10]

3. Validated Pipelines for Peptide Binder Design

One emerging direction relies primarily on protein language models (pLMs), which learn latent representations of protein sequences from large sequence databases. These models capture patterns related to residue compatibility, motif usage, and aspects of structure and interaction propensity directly from sequence statistics.^[11] As a result, they can be used to generate peptide candidates conditioned on a target sequence without explicitly requiring structural information. Methods such as PepMLM^[12] and PepPrCLIP^[13] exemplify this approach by generating peptide sequences conditioned on a target and learned co-evolutionary signatures. This is appealing because it does not require an explicit target structure and can therefore be applied to targets that are unstructured or conformationally heterogeneous. However, sequence-only design still relies on interaction features that are implicitly encoded in the model and lacks an explicit representation of the peptide-target interaction context, which can make it difficult to reliably identify true binders.

A second class of approaches addresses this limitation more directly by generating peptide binders compatible with a target

interface using structural information. One direction uses diffusion-based generative models, which sample peptide backbones directly in three-dimensional space conditioned on a target structure. An early example is RFpeptides,^[14] which adapts RFdiffusion^[15] to generate macrocyclic peptide scaffolds geometrically compatible with a target interface followed by sequence design. Using this approach, peptide binders were obtained against several protein targets including MCL1, MDM2, GABARAP and RbtA. The macrocycles exhibited affinities ranging from 6 nM to 2 μM and exceptional structural agreement between designs and crystal structures.^[14] More recent frameworks extend this concept to broader peptide binder modalities. For example, BoltzGen integrates structure prediction and generative modelling in an all-atom framework capable of designing multiple binder types, including linear peptides and disulfide-bonded cyclic peptides.^[16] In experimental campaigns, peptide binders were obtained against RagC and the RagA:RagC complex, yielding multiple hits after testing tens of designs, with the best linear peptide exhibiting a 3.5 μM affinity and the best macrocycle approximately 80 μM.

A more widely applied peptide design strategy relies on ‘hallucination’ or predictor-guided optimization. Here, structure prediction networks are used iteratively as design oracles to search the sequence-structure space for peptide-target complexes that satisfy predefined loss function constraints (such as prediction confidence, interface contacts, *etc.*).^[17] AfCycDesign represents an early adaptation of AlphaFold2^[18] for cyclic peptide design.^[19] Using a cyclic residue offset, the pipeline can effectively model macrocycle structures that can be used to engineer binders to targets such as MDM2 and Keap1 with reported inhibitory activity in the nanomolar to micromolar range. Related approaches such as EvoBind^[20–22] and EvoBind2^[23] perform joint sequence-structure searches guided by AlphaFold2 with co-evolutionary information extracted from multiple-sequence alignments and can design both linear and cyclic peptides directly using only the target sequence. EvoBind generated cyclic peptide agonists achieving an EC₅₀ of 32 nM against GLP1R,^[21] and micromolar affinity towards the HIV envelope protein^[20] and RNase A.^[23] Also purely structure-based hallucination pipelines with no provided co-evolutionary priors have been shown to be able to efficiently generate peptide binders. For example, BindCraft^[24] was shown to be able to generate helical peptide binders for the oncoprotein MDM2 and the therapeutic target WDR5 with nanomolar affinity.^[25] Interestingly, the helical peptides could be chemically stapled to lock their conformation and improve their binding affinity.

Finally, an emerging direction seeks to expand peptide design beyond the canonical amino acid alphabet. Most current pipelines operate within the standard twenty residues used in protein structure prediction models. Novel methods such as RareFold^[26] extend this framework to incorporate non-canonical amino acids, enabling design of both linear and cyclic peptides in the micromolar range containing expanded chemical building blocks and allowing exploration of interaction chemistries inaccessible to conventional protein design pipelines. Several recently developed generative frameworks (BoltzGen,^[16] RFdiffusion3,^[27] HalluDesign,^[28] BoltzDesign,^[29] and others) operate at an all-atom level or allow arbitrary chemical constraints and are therefore, in principle, capable of incorporating non-canonical residues, although this capability has not yet been experimentally demonstrated.

4. Filling the Gaps

Overall, I think the field is at an exciting but still developing stage. Deep learning has unquestionably made peptide binder design more accessible and productive than it was even a few years ago. We now have credible examples of computational methods being capable of generating experimentally validated linear and cyclic binders. Yet peptide affinities and hit rates are still lower than what we have come to expect for larger and more structured

binders. This is unsurprising, as current prediction models and downstream filtering oracles still underperform on smaller interfaces, highly flexible binders, or peptides with non-canonical components.^[26,30] In my view, this is now the central bottleneck.

A key limitation is that most current pipelines do not explicitly model the physical determinants of peptide-protein interactions. Instead, they rely on proxy metrics derived from structure prediction models, such as prediction confidence, interface contacts, or geometric compatibility, to rank candidate complexes. While these signals can enrich for plausible binders,^[31] they do not explicitly account for the thermodynamics or kinetics of binding, nor for the energetic balance between bound and unbound states. Important contributions such as solvent competition, entropic penalties associated with peptide folding, and conformational heterogeneity are therefore largely ignored.

I therefore suspect that the next phase of progress will not come simply from scaling or further optimizing existing pipelines, but from fundamentally reconsidering what these models are trained to represent. Improved representations of peptide chemistry and conformational flexibility will likely be necessary to capture the diversity of interactions that short polypeptides can form. This may include models that better represent ensembles of peptide conformations rather than single static structures, incorporate solvent and environmental effects more explicitly, and expand beyond the canonical amino acid alphabet. Advances along these directions could allow future design frameworks to more faithfully learn the physical principles governing peptide-protein recognition and ultimately improve both hit rates and achievable affinities.

Acknowledgement

M. Pacesa is supported by the European Research Council (ERC) Starting Grant 101220545 (NAMPify) and by the University Research Priority Program of the University of Zurich (URPP) Innovative Therapies in Rare Diseases (ITINERARE).

Received: March 8, 2026

- [1] D. R. Fox, C. Taveneau, J. Clement, R. Grinter, G. J. Knott, *Structure* **2025**, *33*, 1631, <https://doi.org/10.1016/j.str.2025.08.007>.
- [2] N. London, D. Movshovitz-Attias, O. Schueler-Furman, *Structure* **2010**, *18*, 188, <https://doi.org/10.1016/j.str.2009.11.012>.
- [3] M. Muttenthaler, G. F. King, D. J. Adams, P. F. Alewood, *Nat. Rev. Drug Discov.* **2021**, *20*, 309, <https://doi.org/10.1038/s41573-020-00135-8>.
- [4] N. E. Davey, M. S. Cyert, A. M. Moses, *Cell Commun. Signal.* **2015**, *13*, 43, <https://doi.org/10.1186/s12964-015-0120-z>.
- [5] M. R. Arkin, J. A. Wells, *Nat. Rev. Drug Discov.* **2004**, *3*, 301, <https://doi.org/10.1038/nrd1343>.
- [6] P. J. Carter, G. A. Lazar, *Nat. Rev. Drug Discov.* **2018**, *17*, 197, <https://doi.org/10.1038/nrd.2017.227>.
- [7] S. Hadži, J. Lah, *Protein Sci.* **2022**, *31*, e4370, <https://doi.org/10.1002/pro.4370>.
- [8] S. L. Shammass, *Curr. Opin. Struct. Biol.* **2017**, *42*, 155, <https://doi.org/10.1016/j.sbi.2017.02.003>.
- [9] A. Mondal, L. Chang, A. Perez, *QRB Discov.* **2022**, *3*, e17, <https://doi.org/10.1017/qr.2022.14>.
- [10] T.-S. Cotet, I. Krawczuk, M. Pacesa, L. Nickel, B. E. Correia, N. Haas, A. Qamar, C. A. Challacombe, P. Kidger, C. Ferragu, A. Naka, L. V. Castorina, K. Subr, T. Kluonis, M. J. Stam, S. M. Unal, C. W. Wood, F. Stocco, N. Ferruz, Y. Kurumida, C. N. Calia, F. Paesani, L. de A. Machado, E. Belot, A. Gitter, M. J. Campbell, L. Hallee, A. C. Organizers, *bioRxiv* **2025**, 2025.04.17.648362, <https://doi.org/10.1101/2025.04.17.648362>.
- [11] Z. Zhang, H. K. Wayment-Steele, G. Bixi, H. Wang, D. Kern, S. Ovchinnikov, *Proc. Natl. Acad. Sci. U. S. A.* **2024**, *121*, e2406285121, <https://doi.org/10.1073/pnas.2406285121>.
- [12] L. T. Chen, Z. Quinn, M. Dumas, C. Peng, L. Hong, M. Lopez-Gonzalez, A. Mestre, R. Watson, S. Vincoff, L. Zhao, J. Wu, A. Stavrand, M. Schaeppers-Cheu, T. Z. Wang, D. Srijay, C. Monticello, P. Vure, R. Pulugurta, S. Pertsemliadis, K. Kholina, S. Goel, M. P. DeLisa, J.-T. A. Chi, R. Truant, H. C. Aguilar, P. Chatterjee, *Nat. Biotechnol.* **2025**, *1*, <https://doi.org/10.1038/s41587-025-02761-2>.
- [13] S. Bhat, K. Palepu, L. Hong, J. Mao, T. Ye, R. Iyer, L. Zhao, T. Chen, S. Vincoff, R. Watson, T. Z. Wang, D. Srijay, V. S. Kavirayuni, K. Kholina, S. Goel, P. Vure, A. J. Deshpande, S. H. Soderling, M. P. DeLisa, P. Chatterjee, *Sci. Adv.* **2025**, *11*, eadr8638, <https://doi.org/10.1126/sciadv.adr8638>.
- [14] S. A. Rettie, D. Juergens, V. Adebomi, Y. F. Bueso, Q. Zhao, A. N. Leveille, A. Liu, A. K. Bera, J. A. Wilms, A. Üffing, A. Kang, E. Brackenbrough, M. Lamb, S. R. Gerben, A. Murray, P. M. Levine, M. Schneider, V. Vasireddy, S. Ovchinnikov, O. H. Weiergräber, D. Willbold, J. A. Kritzer, J. D. Mougous, D. Baker, F. DiMaio, G. Bhardwaj, *Nat. Chem. Biol.* **2025**, *21*, 1, <https://doi.org/10.1038/s41589-025-01929-w>.
- [15] J. L. Watson, D. Juergens, N. R. Bennett, B. L. Trippe, J. Yim, H. E. Eisenach, W. Ahern, A. J. Borst, R. J. Ragotte, L. F. Milles, B. I. M. Wicky, N. Hanikel, S. J. Pellock, A. Courbet, W. Sheffler, J. Wang, P. Venkatesh, I. Sappington, S. V. Torres, A. Lauko, V. D. Bortoli, E. Mathieu, S. Ovchinnikov, R. Barzilay, T. S. Jaakkola, F. DiMaio, M. Baek, D. Baker, *Nature* **2023**, *620*, 1089, <https://doi.org/10.1038/s41586-023-06415-8>.
- [16] H. Stark, A. Faltings, M. Choi, Y. Xie, E. Hur, T. O'Donnell, A. Bushuiev, T. U. car, S. Passaro, W. Mao, M. Reveiz, R. Bushuiev, T. s Pluskal, J. Sivic, K. Kreis, A. Vahdat, S. Ray, J. T. Goldstein, A. Savinov, J. A. Hambalek, A. Gupta, D. A. Taquiri-Diaz, Y. Zhang, A. K. Hatstat, A. Arada, N. H. Kim, E. Tackie-Yarboi, D. Boselli, L. Schneider, C. C. Liu, G.-W. Li, D. Hnisz, D. M. Sabatini, W. F. DeGrado, J. Wohlwend, G. Corso, R. Barzilay, T. Jaakkola, *bioRxiv* **2025**, 2025.11.20.689494, <https://doi.org/10.1101/2025.11.20.689494>.
- [17] C. Frank, A. Khoshouei, L. Fuß, D. Schiwietz, D. Putz, L. Weber, Z. Zhao, M. Hattori, S. Feng, Y. de Stigter, S. Ovchinnikov, H. Dietz, *Science* **2024**, *386*, 439, <https://doi.org/10.1126/science.adq1741>.
- [18] J. Junyer, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Thyasuvunakool, R. Bates, A. Židek, A. Potapenko, A. Bridgland, C. Meyer, S. A. A. Kohl, A. J. Ballard, A. Cowie, B. Romera-Paredes, S. Nikolov, R. Jain, J. Adler, T. Back, S. Petersen, D. Reiman, E. Clancy, M. Zielinski, M. Steinegger, M. Pacholska, T. Berghammer, S. Bodenstein, D. Silver, O. Vinyals, A. W. Senior, K. Kavukcuoglu, P. Kohli, D. Hassabis, *Nature* **2021**, *596*, 583, <https://doi.org/10.1038/s41586-021-03819-2>.
- [19] S. A. Rettie, K. V. Campbell, A. K. Bera, A. Kang, S. Kozlov, Y. F. Bueso, J. D. L. Cruz, M. Ahlrichs, S. Cheng, S. R. Gerben, M. Lamb, A. Murray, V. Adebomi, G. Zhou, F. DiMaio, S. Ovchinnikov, G. Bhardwaj, *Nat. Commun.* **2025**, *16*, 4730, <https://doi.org/10.1038/s41467-025-59940-7>.
- [20] D. Daumiller, F. Giammarino, Q. Li, A. Sönnnerborg, R. C. Diez, P. Bryant, *Antivir. Res.* **2026**, *246*, 106336, <https://doi.org/10.1016/j.antiviral.2025.106336>.
- [21] Q. Li, E. Wiita, T. Helleday, P. Bryant, *bioRxiv* **2025**, 2025.06.06.658268, <https://doi.org/10.1101/2025.06.06.658268>.
- [22] P. Bryant, A. Elofsson, *Biorxiv* **2022**, 2022.07.23.501214, <https://doi.org/10.1101/2022.07.23.501214>.
- [23] Q. Li, E. N. Vlachos, P. Bryant, *Commun. Chem.* **2025**, *8*, 211, <https://doi.org/10.1038/s42004-025-01601-3>.
- [24] M. Pacesa, L. Nickel, C. Schellhaas, J. Schmidt, E. Pyatova, L. Kissling, P. Barendse, J. Choudhury, S. Kapoor, A. Alcaraz-Serna, Y. Cho, K. H. Ghamary, L. Vinué, B. J. Yachnin, A. M. Wollacott, S. Buckley, A. H. Westphal, S. Lindhoud, S. Georgeon, C. A. Goverde, G. N. Hatzopoulos, P. Gönczy, Y. D. Muller, G. Schwank, D. C. Swarts, A. J. Vecchio, B. L. Schneider, S. Ovchinnikov, B. E. Correia, *Nature* **2025**, *1*, <https://doi.org/10.1038/s41586-025-09429-6>.
- [25] M. Filius, T. Patsos, H. Minnee, G. Turco, H. E. Chong, J. Liu, M. Gnatzy, R. S. M. Rooth, A. C. H. Liu, R. D. T. Ta, I. H. A. Rijk, S. Ziani, F. J. Boxman, S. J. Pomplun, *ACS Chem. Biol.* **2025**, *20*, 2991, <https://doi.org/10.1021/acscchembio.5c00774>.
- [26] Q. Li, D. Daumiller, F. Zuo, H. Marcotte, Q. Pan-Hammarström, P. Bryant, *bioRxiv* **2025**, 2025.05.19.654846, <https://doi.org/10.1101/2025.05.19.654846>.
- [27] J. K. V. Butcher, R. Krishna, R. Mitra, R. I. Brent, Y. Li, N. Corley, P. Kim, J. Funk, S. V. Mathis, S. Salike, A. Muraishi, H. Eisenach, T. R. Thompson, J. Chen, Y. Politanska, E. Sehgal, B. Coventry, O. Zhang, B. Qiang, K. Didi, M. Kazman, F. DiMaio, D. Baker, *bioRxiv* **2025**, 2025.09.18.676967, <https://doi.org/10.1101/2025.09.18.676967>.
- [28] M. Fang, C. Wang, J. Shi, F. Lian, Q. Jin, Z. Wang, Y. Zhang, Z. Cui, Y. Wang, Y. Ke, Q. Han, L. Cao, *bioRxiv* **2025**, 2025.11.08.686881, <https://doi.org/10.1101/2025.11.08.686881>.
- [29] Y. Cho, M. Pacesa, Z. Zhang, B. E. Correia, S. Ovchinnikov, *bioRxiv* **2025**, 2025.04.06.647261, <https://doi.org/10.1101/2025.04.06.647261>.
- [30] T. Tsaban, J. K. Varga, O. Avraham, Z. Ben-Aharon, A. Khrumushin, O. Schueler-Furman, *Nat. Commun.* **2022**, *13*, 176, <https://doi.org/10.1038/s41467-021-27838-9>.
- [31] M. D. Overath, A. Rygaard, C. P. Jacobsen, V. Brasas, O. Morell, P. Sormanni, T. P. Jenkins, *bioRxiv* **2025**, <https://doi.org/10.1101/2025.08.14.670059>.