Discovery of novel peptide-GPCR signalling systems across evolutionary domains

Luuk R. H. Dijkhof[1], Teemu K.E. Rönkkö[1], Marin Matic[1], & Alexander S. Hauser[1]

https://doi.org/10.17952/38EPS.2026.O63

Department of Drug Design and Pharmacology, Center for Pharmaceutical Data Science Education, University of Copenhagen, 2100 Copenhagen, Denmark

Oral contribution to the EPS 2026 Conference Proceedings

Abstract. Peptide signalling through G protein-coupled receptors (GPCRs) underlies fundamental physiological processes and has repeatedly provided starting points for therapeutic innovation. Yet the canonical set of human peptide hormones represents only a small fraction of the sequence space capable of engaging GPCRs. Here, I summarize a strategy for expanding peptide–receptor discovery beyond annotated human hormones by combining evolutionary and metagenomic sequence mining, structure-aware prediction, benchmarking, and experimental pharmacology. A receptor-aware benchmark shows that modern structure-prediction approaches can discriminate cognate peptide–GPCR pairs from challenging decoys, and that rescoring local receptor–peptide interfaces can further improve classification. Applying related concepts to the human gut microbiome identifies candidate molecular mimics of endogenous peptide hormones, including a microbial peptide with agonist activity at melanocortin receptors. Together, these studies illustrate how evolutionary diversity, computational prioritization and scalable pharmacology can be integrated to uncover previously hidden peptide signalling systems.

Introduction

G protein-coupled receptors (GPCRs) form one of the largest families of cell-surface receptors and translate a remarkable diversity of extracellular signals into intracellular responses. Approximately 800 GPCRs are encoded by the human genome, with hundreds of approved medicines acting on this receptor family. Peptides are particularly important GPCR ligands: they combine high potency and selectivity with a vast, evolutionarily diversified chemical and sequence space. Classic examples also show that useful peptide pharmacology need not originate from humans. Exendin-4 from Gila monster saliva helped establish the therapeutic potential of GLP-1 receptor agonism, salmon calcitonin has been used clinically, and microbial or venom-derived peptides can engage human receptors with striking specificity.[1–3]

At the same time, peptide discovery is undergoing a change in scale. Genome and transcriptome sequencing, metagenomics, peptidomics and improved annotation of small open reading frames are revealing large numbers of secreted or potentially bioactive peptides that are absent from canonical hormone catalogues. Recent work has demonstrated that precursor-first and genome-first discovery can uncover peptides with strong physiological effects even before their cognate receptor is known. The emerging bottleneck is therefore increasingly not whether candidate peptides exist, but how to identify their receptor targets and signalling profiles efficiently.[4–6]

Our approach treats this as a peptide–receptor pairing problem, where computational models can prioritize receptor pairs for experimental testing. A central requirement is rigorous benchmarking: models must distinguish genuine interactions not only from obviously unrelated peptides, but also from biologically plausible alternatives that resemble true ligands or target related binding pockets.

Results and Discussion

Benchmarking peptide–GPCR interaction prediction

A useful benchmark must reflect the many-to-many architecture of peptide signalling: individual peptides may activate multiple receptors, while individual receptors may respond to more than one endogenous ligand. We assembled a benchmark centred on 124 known positive peptide–GPCR interactions. For each receptor, negative examples were selected in two deliberately different regimes. Five dissimilar decoys were sampled from principal agonists of GPCRs with less than 35% sequence similarity, whereas five challenging decoys were drawn from principal agonists of receptors with the most similar binding pockets. The resulting dissimilar and similar decoys had mean binding-pocket similarities of approximately 20% and 62%, respectively. This design asks models to solve both an easier interaction-recognition problem and a more pharmacologically meaningful discrimination problem among related receptor environments.

We compared sequence- and structure-based approaches using model-appropriate confidence or interaction scores, including AlphaFold-derived complex confidence, ESMFold/NeuralPLexer confidence, D-SCRIPT interaction probability and structure-based error metrics. The benchmark highlighted strong discriminatory power from approaches that exploit evolutionary information and showed that performance depends substantially on the difficulty of the decoy set. This is important for prospective discovery: a model that separates a cognate ligand from random peptides may still fail when confronted with related peptide systems.[7]

We therefore asked whether the predicted complex could be evaluated at the local receptor–peptide interface rather than by global confidence alone. Rescoring strategies based on local interaction scores, energetic terms and graph-neural-network representations improved the ability to surface true binders from decoys. In the presented benchmark, interface-focused rescoring improved ROC performance by about 0.11. These results support a practical two-stage strategy in which co-folding proposes plausible complexes and local interface models provide an additional filter before experimental testing. The broader implication is that structural prediction is most useful when embedded in a benchmarked ranking framework rather than interpreted as a binary claim of binding.

Microbial peptide mimics as an emerging signalling space

Evolution has repeatedly generated peptides that interact with conserved human receptor architectures, including peptides from animals, plants, microorganisms, venoms, and population variability (**Figure 1**). This motivates systematic searches across biodiversity for sequence or structural motifs compatible with human GPCRs. Candidate-generation strategies can combine profile hidden Markov models, sequence similarity, structural searches, secretion signals and precursor context before moving to receptor co-folding and experimental pharmacology. For disulfide-rich venom peptides such as conotoxins, cysteine connectivity and topology provide an additional layer of information that can be incorporated into prediction. Here we employ structure-aware models because peptide sequence similarity alone can be weak even when receptor engagement is conserved.[2,3]

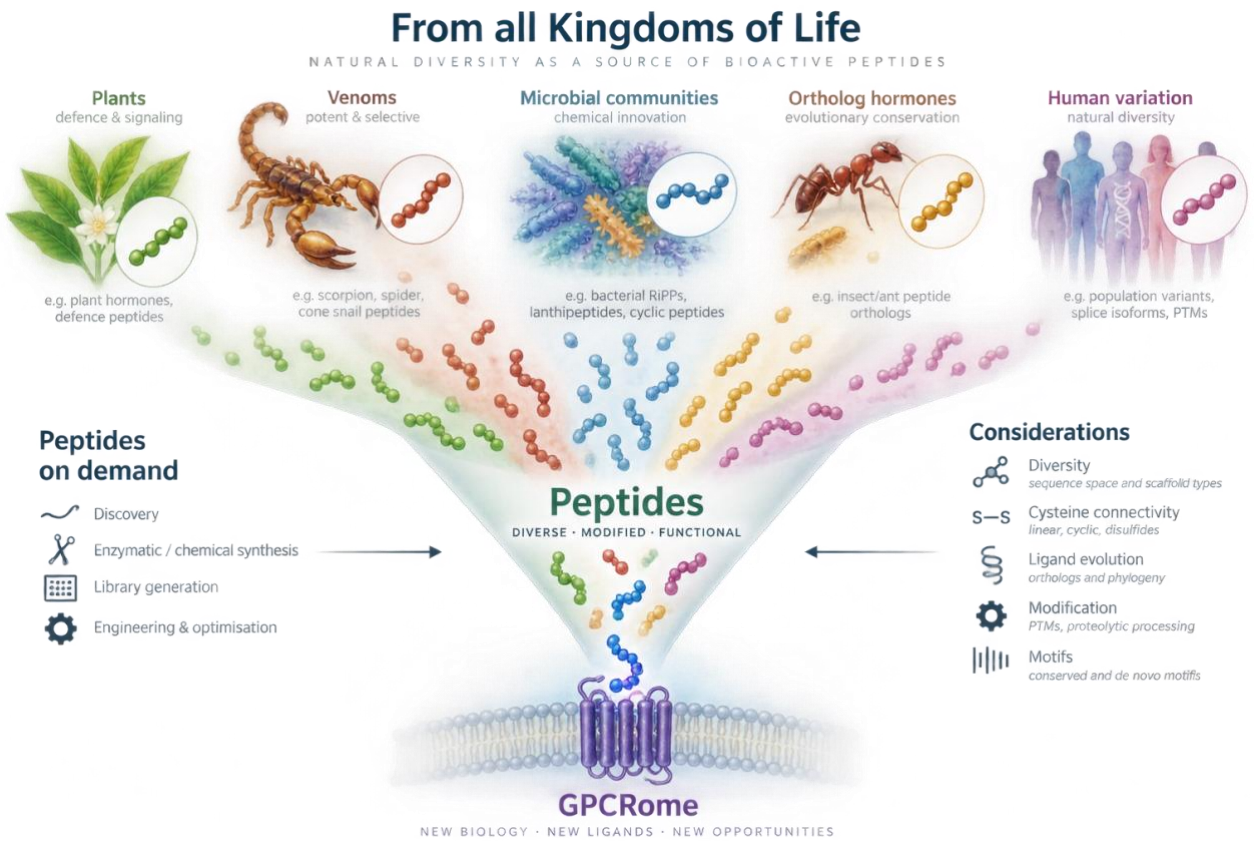


Figure 1: Nature-derived diversity as a source of bioactive peptides targeting GPCRs. Peptide candidates from plants, venoms, microbial communities, evolutionary hormone orthologs and human variation provide diverse sequence and structural space for discovery, engineering and experimental screening against the GPCRome.

The human gut microbiome provides a particularly compelling test case because microbial molecules are continuously produced in close proximity to the host. Gut bacteria are already known to generate small molecules and peptides that resemble or modulate human signalling systems. We therefore developed an in silico workflow that starts from shotgun metagenomic sequence data, translates the microbial sequence space into candidate proteins and searches for peptide segments that mimic endogenous human ligands. In the presented analysis, approximately 150 million gut microbial proteins were searched using endogenous human peptides as reference signals, followed by AlphaFold 3-based co-folding, local rescoring, and prioritization against candidate GPCR targets.[8,9]

This procedure yielded 89 tentative microbial peptide candidates spanning 11 GPCR families. These candidates are hypotheses rather than established endogenous microbial ligands, but they demonstrate the scale of the accessible search space. Initial experimental testing provided a first proof of principle: one gut microbial peptide showed agonist activity at melanocortin receptors in cell-based dose–response experiments. The observation is notable because melanocortin signalling participates in processes ranging from inflammatory regulation to energy homeostasis. Whether such microbial peptides reach relevant concentrations, compartments, and receptors *in vivo* remains unresolved and will require additional physiological experiments.

This example is therefore a demonstration of a discovery route that may lead to the discovery of novel microbiome-to-host physiology. Metagenomic sequence space can be reduced from hundreds of millions of proteins to a tractable set of structurally plausible receptor–peptide pairs, which can then be tested experimentally. Similar workflows can be extended to other microbial communities and to biodiversity-derived peptide libraries from different domains of life.

Perspective

Taken together, these studies suggest a general framework for peptide discovery. Evolution supplies candidate ligands far beyond the canonical human hormone set. Benchmarked computational models narrow the receptor pairs worth testing and local interface scoring can add resolution when related receptors or ligands create difficult decoys. Experimental validation remains essential, particularly because structural confidence does not establish receptor activation, potency, efficacy or physiological exposure. The most productive role of current AI and structure-prediction methods is therefore as prioritization engines coupled tightly to pharmacology. As sequence databases, metagenomic cohorts and structural models continue to expand, this integration should make it increasingly feasible to map previously inaccessible regions of the peptide–GPCR signalling landscape.

Acknowledgements

This research was funded by Biodiversa+, the European Biodiversity Partnership, in the context of the ‘Biodiversity2Drugs’ project under the 2023-2024 BiodivNBS joint call and the Innovation Fund Denmark. We also acknowledge support by the Carlsberg Foundation (grant CF20-0248). We also acknowledge collaborators and partners across the peptide-discovery network, including Christian Gruber, Markus Muttenthaler, Dominik J. Windisch. We thank the EPS 2026 organizers for the invitation and discussion.

References

1. Sánchez Lorente, J. et al. GPCR drug discovery: new agents, targets and indications. Nat. Rev. Drug Discov. 24, 458–479 (2025).
2. Hermes, J. et al. Peptide ligand discovery of G protein-coupled receptors. Nat. Rev. Methods Primers 6, 52 (2026).
3. Gruber, C. W. et al. Biodiversity2Drugs—Renaissance of exploring nature-derived peptides for GPCR ligand discovery. Br. J. Pharmacol. 182, 3691–3694 (2025).
4. Foster, S. R. et al. Discovery of human signaling systems: pairing peptides to G protein-coupled receptors. Cell 179, 895–908.e21 (2019).
5. Coassolo, L. et al. Prohormone cleavage prediction uncovers a non-incretin anti-obesity peptide. Nature 641, 192–201 (2025).
6. Deutsch, E. W. et al. Expanding the human proteome with microproteins and peptideins. Nature 654, 813–825 (2026).
7. Hoegen Dijkhof, L. R., Rönkkö, T. K. E., von Vegesack, H. C., Lenzing, J. & Hauser, A. S. Deep learning in GPCR drug discovery: benchmarking the path to accurate peptide binding. Brief. Bioinform. 26, bbaf186 (2025).
8. Cohen, L. J. et al. Commensal bacteria make GPCR ligands that mimic human signalling molecules. Nature 549, 48–53 (2017).
9. Fan, Y. et al. Polypeptides synthesized by common bacteria in the human gut improve rodent metabolism. Nat. Microbiol. 10, 1918–1939 (2025).