

Bayesian Flow-Based Peptide Design for Antimicrobial Resistance Mitigation and Therapeutic Discovery

Aishani Sule

Department of Computer Science, University of Houston, Houston, TX, USA.

aishaniwork@uh.edu

Elsbeth Wulf

Department of Computer Science, University of Central Florida, Orlando, FL, USA.

wulfelsbeth@ucf.edu

Serafin Padron

Department of Computer Science and Engineering, University at Buffalo, Buffalo, NY, USA.

serafinpadron652@buffalo.edu

Abstract

Antimicrobial resistance has become one of the most consequential socio-technical failures of contemporary medicine, linking clinical overuse, agricultural practices, weak diagnostics, uneven regulatory capacity, and global inequities in access to care. Antimicrobial peptides offer a promising therapeutic class because they can disrupt microbial membranes, modulate immunity, and evade some conventional resistance pathways, yet their discovery remains limited by sequence space, toxicity, instability, and costly wet-lab validation. This paper examines Bayesian flow-based peptide design as a system-level strategy for accelerating antimicrobial resistance mitigation and therapeutic discovery. Rather than treating generative modeling as a purely computational exercise, the analysis positions Bayesian flow within a larger architecture that includes multi-modal molecular representation, all-atom rationalization, closed-loop experimental validation, manufacturing infrastructure, clinical translation, biosecurity governance, and equitable access. The discussion emphasizes structural trade-offs among generative fidelity, interpretability, computational cost, data provenance, and regulatory accountability. It also considers robustness under distribution shift, fairness in global health deployment, dual-use risk, environmental sustainability, and policy mechanisms for responsible innovation. The central argument is that Bayesian flow-based peptide design can meaningfully contribute to antimicrobial resistance mitigation only if it is embedded in durable socio-technical systems that connect model development to laboratory verification, production capacity, surveillance, and public governance. The paper concludes with forward-looking priorities for interdisciplinary research and policy.

Keywords

Bayesian flow; antimicrobial peptides; antimicrobial resistance; generative protein design; socio-technical systems; biosecurity governance; clinical translation; sustainable manufacturing.

1. Introduction

Antimicrobial resistance is not merely a biological phenomenon but a systems problem that emerges from clinical prescribing, agricultural antibiotic use, inadequate sanitation, weak infection control, global travel, and uneven access to effective medicines [1], [2]. The World Health Organization has framed resistance as a threat requiring coordinated action across human health, animal health, food systems, and environmental management [3]. In this context, the search for new therapeutic classes is urgent, but it cannot be reduced to the discovery of individual molecules. It requires infrastructures for surveillance, rapid diagnostics, stewardship, manufacturing, distribution, and regulatory oversight. Antimicrobial peptides have attracted sustained attention because they often act through membrane disruption, immunomodulation, and other mechanisms that differ from conventional small-molecule antibiotics [4], [5]. Their potential to address resistant pathogens is considerable, yet the design of peptides with high activity, low toxicity, acceptable stability, and scalable production remains difficult [6].

The difficulty is partly a search problem. Peptide sequence space is vast, and functional constraints are distributed across sequence, structure, charge, hydrophobicity, conformational flexibility, and environmental context. Computational tools have therefore become central to antimicrobial peptide discovery, ranging from database mining and quantitative structure-activity analysis to deep generative models [7]. Protein language models and sequence-based representation learning have demonstrated that large unlabeled datasets can yield useful embeddings for design and prediction [8], [9]. More recently, unified multi-modal Bayesian flow approaches have shown promise for rationalized all-atom protein design by integrating sequence, structure, and chemical information within a probabilistic generative framework [10]. Such methods suggest a pathway toward peptide candidates that are not merely plausible sequences but structurally and functionally rationalized molecules.

At the same time, the translation of generative peptide design into antimicrobial resistance mitigation depends on system-level choices. A model that produces candidate sequences without validated synthesis routes, toxicity assays, stability testing, and clinical development pathways remains a computational artifact. A deployment that ignores data provenance, biosecurity, or equitable access may create new risks even as it addresses old ones. This paper argues that Bayesian flow-based peptide design should be understood as one component of a broader socio-technical infrastructure. The sections that follow examine antimicrobial resistance as a multi-sector challenge, peptide therapeutics and their design bottleneck, Bayesian flow methods, deployment architecture, robustness and fairness, clinical translation and manufacturing, policy and biosecurity, and future research directions.

2. Antimicrobial Resistance as a Socio-Technical Challenge

The persistence of antimicrobial resistance reflects a mismatch between biological evolution and institutional response. Bacteria can acquire resistance through mutation, horizontal gene transfer, and selection pressures created by antibiotic exposure. Hospitals, farms, wastewater systems, and communities are connected in ways that make resistance a shared infrastructural problem rather than a series of isolated clinical events [1], [2]. When surveillance is fragmented, outbreaks of resistant organisms may be detected late. When diagnostics are unavailable, clinicians may prescribe broad-spectrum antibiotics empirically. When infection prevention is underfunded, resistant organisms circulate more easily. These are not only scientific failures but governance and infrastructure failures [3].

Addressing resistance therefore requires simultaneous action on multiple fronts. Stewardship programs seek to reduce inappropriate antibiotic use, but they depend on laboratory capacity,

clinical decision support, and public trust. Agricultural policy must address antibiotic use in livestock and aquaculture while protecting food security and rural livelihoods. Environmental monitoring must account for pharmaceutical waste, hospital effluent, and agricultural runoff. Global equity is central because low- and middle-income countries often bear the highest burden of resistant infections while having the least access to new diagnostics and therapies [2]. A new peptide therapeutic that is clinically effective but unavailable or unaffordable in high-burden settings would not solve the systemic problem.

Digital and artificial intelligence systems enter this landscape as both enablers and new dependencies. Machine learning can improve outbreak prediction, antibiotic prescribing, diagnostic interpretation, and drug discovery [17], [18]. However, algorithmic systems also require data, compute, maintenance, and governance. If they are deployed without attention to bias, transparency, and accountability, they can reproduce existing inequities or create new forms of opacity in clinical decision-making [19], [20]. For antimicrobial resistance, the risk is that advanced design platforms concentrate capability in high-income institutions while the burden of resistance remains global. An effective strategy must therefore link molecular innovation to capacity building, technology transfer, and public health infrastructure.

3. Peptide Therapeutics and the Design Bottleneck

Antimicrobial peptides are short chains of amino acids that can kill or inhibit microbes through diverse mechanisms. Many are cationic and amphipathic, allowing them to interact with negatively charged microbial membranes, but others act intracellularly or through immunomodulatory pathways [4]. This mechanistic diversity is attractive because it may reduce the ease with which pathogens develop resistance. However, the same properties that enable membrane activity can also produce cytotoxicity against host cells, limiting therapeutic windows. Peptides may be degraded by proteases, have poor oral bioavailability, require parenteral administration, and face manufacturing costs that are higher than those of many small molecules [5], [6].

The design bottleneck is multidimensional. A candidate peptide must balance antimicrobial activity against target pathogens, selectivity for microbial versus host membranes, stability in physiological conditions, solubility, immunogenicity, and manufacturability. These objectives often conflict. Increasing positive charge can improve membrane interaction but may increase toxicity. Increasing hydrophobicity can enhance insertion into membranes but may reduce solubility. Rigidifying structure can improve stability but may reduce broad-spectrum activity. Computational design must therefore navigate a Pareto frontier rather than optimize a single score [6], [7]. Traditional screening and rational design methods have produced important peptides, but their throughput is limited by the cost and time of synthesis and biological testing.

Generative models offer a way to explore sequence and structure space more efficiently. Protein language models can learn statistical regularities from large sequence databases and propose novel sequences with plausible folding properties [8]. Representation learning methods can transfer information from unlabeled protein data to predict function and stability [9]. Structure prediction systems such as AlphaFold have transformed the ability to estimate three-dimensional conformations, while diffusion-based protein design methods have demonstrated de novo generation of protein structures and functional motifs [11], [12]. These advances create a foundation for peptide design, but peptides present specific challenges because their short length, conformational flexibility, and membrane interactions are not always well captured by models trained primarily on globular proteins.

Bayesian flow-based approaches add a probabilistic layer that can integrate multiple modalities and uncertainty. In principle, such methods can combine sequence information, structural constraints, and chemical properties in a unified generative process [10]. This is important because antimicrobial peptide function is not determined by sequence alone; it depends on conformation, aggregation, membrane composition, and host environment. A model that can represent these dependencies while quantifying uncertainty can help prioritize candidates for synthesis and testing. Yet the design bottleneck does not disappear with better generation. It shifts toward validation, interpretation, and integration with experimental workflows.

4. Bayesian Flow-Based Generative Modeling for Peptide Design

Bayesian flow models belong to a broader family of generative methods that transform simple probability distributions into complex data distributions through a sequence of transformations. Diffusion models, for example, learn to reverse a noise process and have achieved strong results in image and molecular generation [13], [14]. Variational autoencoders and generative adversarial networks represent earlier approaches to learning latent representations and generating samples [15], [16]. Bayesian flow methods extend these ideas by framing generation as an inference process that can incorporate Bayesian updating, multi-modal inputs, and uncertainty calibration. The appeal for peptide design is that such models can potentially reason over sequence, structure, and chemical context without reducing the problem to a single modality.

A unified multi-modal Bayesian flow approach for all-atom protein design illustrates the direction of travel [10]. In system terms, the value of such a method lies not only in sample quality but in its ability to coordinate heterogeneous representations. Sequence tokens, atomic coordinates, side-chain conformations, and physicochemical descriptors can be treated as related views of the same molecular object. A Bayesian formulation can, in principle, update beliefs about candidate structures as new constraints are introduced, such as target binding sites, membrane interaction motifs, or stability requirements. This matters for antimicrobial peptides because design objectives are conditional and context-dependent. A peptide intended for bloodstream infections faces different constraints from one intended for topical or pulmonary delivery.

However, multi-modal generative design introduces structural trade-offs. Greater model complexity can improve expressive power but reduces interpretability and increases computational cost. All-atom representations may capture fine-grained chemistry but require more data and compute than sequence-only models. Bayesian uncertainty estimates can guide active learning, but they are only as good as the prior and likelihood assumptions embedded in the model. If the training data overrepresents certain peptide families, pathogens, or assay conditions, the model may generate candidates that appear novel but are poorly suited to real-world diversity. There is also a risk of overfitting to *in silico* metrics that do not correlate with wet-lab activity or clinical efficacy.

From a system architecture perspective, Bayesian flow-based peptide design should be connected to validation loops. Generated candidates can be ranked by predicted activity, toxicity, stability, and synthesizability. Selected peptides can then be produced and tested *in vitro* against panels of resistant pathogens. Results can be fed back into the model through active learning or Bayesian updating. This closed-loop approach transforms generative modeling from a one-time computational task into an ongoing experimental program. It also creates data infrastructure requirements: standardized assays, metadata capture, negative

results, and reproducible protocols. Without such infrastructure, the model cannot learn reliably from experience.

5. System Architecture and Deployment Considerations

Deploying Bayesian flow-based peptide design at scale requires an architecture that spans data ingestion, model training, inference, experimental orchestration, and regulatory documentation. Data sources may include public peptide databases, genomic surveillance, structural repositories, assay results, and proprietary laboratory records [7], [8], [9]. These sources differ in format, quality, licensing, and privacy constraints. A robust system must track provenance, version data, and document transformations. It must also handle missing and noisy labels, because biological assay data are often heterogeneous and context-dependent. The architecture is therefore as much a data governance system as a machine learning system.

Compute infrastructure is another critical dimension. Training large multi-modal generative models requires significant hardware, energy, and engineering expertise. Inference may be less demanding but still needs to support interactive design sessions, batch screening, and integration with laboratory information management systems. Some institutions may choose cloud-based deployment, while others may require on-premises systems for data sovereignty or security. Hybrid architectures can balance scalability with control, but they introduce complexity in synchronization, access management, and fault tolerance. The choice affects reproducibility, cost, and the ability to collaborate across institutions and countries.

The deployment environment also shapes scientific validity. A model trained on data from one region may not generalize to pathogens prevalent in another. A peptide design pipeline optimized for one synthesis platform may produce sequences that are difficult to manufacture elsewhere. A user interface that hides uncertainty may encourage overconfidence, while one that exposes too much complexity may be unusable in clinical or public health settings. System designers must therefore consider human factors, workflow integration, and decision support. The goal is not to replace microbiologists or clinicians but to augment their judgment with transparent, well-calibrated computational evidence [17], [18].

Maintainability and sustainability are often overlooked. Machine learning systems are not static; they require updates, monitoring, and retraining as new data emerge and as pathogens evolve. A model that performs well at deployment may degrade over time if resistance patterns shift or if assay protocols change. Continuous evaluation, drift detection, and rollback mechanisms are necessary. Documentation for regulatory review must be maintained throughout the lifecycle. These requirements increase cost, but they are essential for trust and long-term impact. In the context of antimicrobial resistance, where time horizons span decades, sustainable infrastructure is not optional.

6. Robustness, Fairness, and Governance

Robustness in peptide design means more than resistance to adversarial inputs. It means reliable performance across diverse pathogens, assay conditions, host environments, and manufacturing constraints. Distribution shift is a major challenge because the biological world is non-stationary. Bacteria evolve, clinical practices change, and new resistance mechanisms emerge. A generative model trained on historical data may propose peptides that are effective against past strains but not against current or future threats. Robust systems must therefore incorporate surveillance data, prospective validation, and mechanisms for rapid updating [2],

[3]. They must also account for uncertainty in a way that informs decision-making rather than obscuring it.

Fairness is equally important. Antimicrobial resistance disproportionately affects populations with limited access to clean water, sanitation, diagnostics, and healthcare. If advanced peptide design platforms are concentrated in well-resourced institutions, the benefits may accrue primarily to wealthy patients and markets. Fairness requires attention to affordability, licensing, technology transfer, and local manufacturing capacity. It also requires inclusive datasets that represent pathogens and patient populations globally. A model trained primarily on data from high-income settings may misestimate efficacy or toxicity in other contexts [19], [20]. Governance frameworks should therefore include equity criteria in funding, evaluation, and deployment decisions.

Governance must also address dual-use risk. Generative protein design can be used for beneficial therapeutics, but it can also lower barriers to creating harmful biological agents. The same methods that design antimicrobial peptides could, in principle, be misused to enhance virulence or evade detection. National and international frameworks for biosecurity have begun to address synthetic biology and artificial intelligence, but the pace of governance often lags behind technical capability [21], [22]. Effective governance requires a combination of sequence screening, access controls, responsible disclosure, and international cooperation. It should not stifle legitimate research, but it must recognize that computational design platforms are part of a broader biosecurity landscape.

Transparency and accountability are central to governance. Regulators, clinicians, and publics need to understand how a candidate peptide was designed, what data were used, what uncertainties remain, and how safety was evaluated. This does not mean disclosing proprietary details that would undermine innovation, but it does require auditable processes and clear lines of responsibility. In clinical settings, algorithmic recommendations should be integrated with human oversight and documented in medical records. In public health settings, model outputs should be traceable to surveillance data and epidemiological assumptions. Governance is thus not a bureaucratic add-on but a design requirement for trustworthy systems [19], [20].

7. Clinical Translation, Manufacturing, and Sustainability

Clinical translation of peptide therapeutics is challenging. Peptides often have poor oral bioavailability, short half-lives, and potential immunogenicity. Formulation strategies such as liposomes, nanoparticles, pegylation, and prodrug approaches can improve delivery, but they add complexity and cost [5], [6]. For antimicrobial peptides, the site of infection matters. A peptide for skin infections may be delivered topically, while one for systemic infections must survive proteases and reach effective concentrations in target tissues. Clinical trials must therefore evaluate not only antimicrobial activity but also pharmacokinetics, safety, and patient outcomes. Generative design can accelerate candidate discovery, but it cannot replace rigorous clinical development.

Manufacturing is another bottleneck. Solid-phase peptide synthesis is well established, but scale-up can be expensive, especially for longer or modified peptides. Environmental impact is significant because peptide synthesis uses solvents, reagents, and energy. Sustainable manufacturing requires greener chemistries, solvent recovery, waste reduction, and efficient purification. These considerations should inform design choices early. A peptide that is highly active but impossible to manufacture at scale or with acceptable environmental cost is not a

viable public health solution. System-level design should therefore include manufacturability and life-cycle assessment as objectives alongside biological activity [5], [6].

Regulatory pathways for AI-designed peptides are still evolving. Regulators must evaluate the safety and efficacy of the final product, but they may also need to understand the computational methods used in discovery. This creates questions about validation, reproducibility, and post-market surveillance. If a model is updated after approval, how does that affect the approved product? If a peptide is designed for a specific pathogen, how should resistance emergence be monitored? These questions require collaboration among developers, clinicians, regulators, and public health agencies. They also highlight the need for standards in data reporting, model documentation, and clinical evidence generation [17], [18].

Sustainability also includes the broader health system. Introducing a new peptide therapeutic requires supply chains, cold storage, prescribing guidelines, and monitoring. In low-resource settings, these requirements may be difficult to meet. A sustainable strategy should therefore consider simplified formulations, heat stability, and affordable production. It should also integrate with existing stewardship programs to avoid inappropriate use that could accelerate resistance. The goal is not simply to add a new drug but to strengthen the system that delivers it. Bayesian flow-based design can contribute to this goal if it is paired with manufacturing innovation, health economics, and implementation science.

8. Policy Implications and Biosecurity

Policy for AI-enabled peptide design must balance innovation, access, and security. Public funding agencies can support open datasets, benchmark tasks, and shared infrastructure that reduce duplication and increase reproducibility. They can also require biosecurity risk assessments for generative protein design projects. International agreements can promote common standards for sequence screening and responsible research conduct. However, policy must be adaptive because the technology is changing rapidly. Static rules may quickly become obsolete, while vague principles may fail to guide practice. A combination of risk-based regulation, professional norms, and technical safeguards is likely to be most effective [21], [22].

Biosecurity risks are not hypothetical. Generative models can produce sequences that are not found in nature, and the tools for synthesis are becoming more accessible. This creates a tension between openness and control. Too much restriction could slow beneficial research and widen inequities; too little could enable misuse. One approach is to focus on choke points, such as DNA synthesis screening, equipment controls, and know-your-customer requirements. Another is to invest in detection and response capabilities, including genomic surveillance and rapid diagnostics. Governance should also be international, because pathogens and data do not respect borders. The World Health Organization's action plan on antimicrobial resistance provides a foundation, but it must be updated for the age of generative biology [3].

Policy should also address intellectual property and access. If AI-designed peptides are patented broadly, they may be unaffordable in high-burden settings. Alternative models such as patent pools, open licensing, and tiered pricing could improve access. Public research funders can attach access conditions to grants. International collaborations can build local capacity for design, testing, and manufacturing. These measures are not only ethical but also strategically important, because resistance anywhere can spread everywhere. A fragmented response that leaves some regions without new therapies will ultimately undermine global health security [2], [3].

Finally, policy must support evaluation. It is not enough to claim that Bayesian flow-based design accelerates discovery; the claim must be tested against real-world outcomes. Metrics should include time to candidate, success rate in preclinical and clinical trials, cost per validated peptide, equity of access, and environmental impact. Independent evaluation can identify what works, what fails, and where resources should be directed. This requires funding for interdisciplinary research that combines computational science, microbiology, health economics, and policy analysis. Without such evaluation, the field risks repeating cycles of hype and disappointment.

9. Future Directions

Future research should advance Bayesian flow methods in several directions. First, models need to handle uncertainty more effectively, especially in out-of-distribution settings. This includes calibration, ensembling, and Bayesian updating as new experimental data arrive [10], [13], [14]. Second, multi-modal representations should be extended to include membrane composition, host immune response, and pharmacokinetic properties. Peptide activity is context-dependent, and models that ignore context may produce misleading predictions. Third, active learning and closed-loop experimentation should be integrated into standard workflows. Laboratories can use robotic synthesis and high-throughput assays to generate data that improve models iteratively [8], [9].

Another direction is federated and privacy-preserving learning. Many valuable datasets are held by hospitals, companies, and government agencies that cannot share raw data. Federated learning can allow models to be trained across institutions without centralizing sensitive information. However, it introduces challenges in communication, heterogeneity, and governance. For antimicrobial resistance, federated approaches could combine surveillance data from multiple countries while respecting national sovereignty. They could also enable smaller institutions to participate in model development without building large compute infrastructure. These are system-level innovations as much as algorithmic ones [17], [18].

Interpretability is also critical. Clinicians and regulators are more likely to trust a design if they can understand why a peptide was proposed. Attribution methods, structural visualizations, and uncertainty reports can help. However, interpretability must be meaningful rather than superficial. It should connect model outputs to biological mechanisms and experimental evidence. This requires collaboration between computational scientists and microbiologists. It also requires standards for reporting model behavior, limitations, and failure modes. Without interpretability, adoption will remain limited to research settings [19], [20].

Finally, the field should invest in benchmarks and shared infrastructure. Benchmarks for antimicrobial peptide design should include diverse pathogens, realistic assay conditions, and multiple objectives such as activity, toxicity, and stability. They should be periodically updated to prevent overfitting. Shared data platforms can reduce duplication and improve reproducibility. Regulatory science can develop templates for documenting AI-designed therapeutics. These collective goods are not glamorous, but they are essential for translating Bayesian flow-based design into durable public health impact [7], [10], [21].

10. Conclusion

Bayesian flow-based peptide design offers a promising computational pathway for antimicrobial resistance mitigation and therapeutic discovery. By integrating sequence, structure, and chemical information within a probabilistic generative framework, such

methods can propose candidates that are more rationalized and context-aware than sequence-only approaches. However, the value of these methods depends on the systems in which they are embedded. Without validated experimental loops, robust data infrastructure, manufacturing capacity, fair access policies, and biosecurity governance, generative design will remain an isolated technical achievement rather than a public health solution. The analysis in this paper emphasizes that antimicrobial resistance is a socio-technical challenge requiring coordinated action across disciplines and sectors. Bayesian flow models can contribute to this effort if they are designed for uncertainty, interpretability, sustainability, and equity. Future work should focus on closed-loop validation, multi-modal representation, federated learning, benchmarks, and policy integration. The overarching lesson is that molecular innovation and institutional design must advance together. Only then can AI-enabled peptide discovery meaningfully reduce the burden of antimicrobial resistance.

References

1. Ventola, C. L. (2015). The antibiotic resistance crisis: Part 1: Causes and threats. *P & T*, 40(4), 277–283.
2. Laxminarayan, R., Duse, A., Wattal, C., Zaidi, A. K. M., Wertheim, H. F. L., Sumpradit, N., Vlieghe, E., Hara, G. L., Gould, I. M., Goossens, H., Greko, C., So, A. D., Bigdeli, M., Tomson, G., Woodhouse, W., Ombaka, E., Peralta, A. Q., Qamar, F. N., Mir, F., ... Cars, O. (2013). Antibiotic resistance—the need for a global solutions. *The Lancet Infectious Diseases*, 13(12), 1057–1098.
3. World Health Organization. (2015). Global action plan on antimicrobial resistance. World Health Organization.
4. Fjell, C. D., Hiss, J. A., Hancock, R. E. W., & Schneider, G. (2012). Designing antimicrobial peptides: Form follows function. *Nature Reviews Drug Discovery*, 11(1), 37–51.
5. Mahlapuu, M., Håkansson, J., Ringstad, L., & Björn, C. (2016). Antimicrobial peptides: An emerging category of therapeutic agents. *Frontiers in Cellular and Infection Microbiology*, 6, 194.
6. Torres, M. D. T., Sothiselvam, S., Lu, T. K., & de la Fuente-Nunez, C. (2019). Peptide design principles for antimicrobial applications. *Journal of Molecular Biology*, 431(18), 3547–3567.
7. Porto, W. F., Pires, Á. S., & Franco, O. L. (2017). Computational tools for exploring sequence databases as a resource for antimicrobial peptides. *Biotechnology Advances*, 35(3), 337–349.
8. Madani, A., McCann, B., Naik, N., Keskar, N. S., Anand, N., Eguchi, R. R., Huang, P. S., & Socher, R. (2020). ProGen: Language modeling for protein generation. *arXiv preprint arXiv:2004.03497*.
9. Alley, E. C., Khimulya, G., Biswas, S., AlQuraishi, M., & Church, G. M. (2019). Unified rational protein engineering with sequence-based deep representation learning. *Nature Methods*, 16(12), 1315–1322.
10. Wu, H., Song, Y., Zhang, Z., Zhang, Z., Zhou, H., Ma, W. Y., & Liu, J. (2026). Rationalized All-Atom Protein Design with Unified Multi-Modal Bayesian Flow. *Advances in Neural Information Processing Systems*, 38, 3019-3050.

11. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Židek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S. A. A., Ballard, A. J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., ... Hassabis, D. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature*, 596(7873), 583–589.
12. Watson, J. L., Juergens, D., Bennett, N. R., Trippe, B. L., Yim, J., Eisenach, H. E., Ahern, W., Borst, A. J., Ragotte, R. J., Milles, L. F., Wicky, B. I. M., Hanikel, N., Pellock, S. J., Courbet, A., Sheffler, W., Wang, J., Venkatesh, P., Sappington, I., Torres, S. V., ... Baker, D. (2023). De novo design of protein structure and function with RFdiffusion. *Nature*, 620(7976), 1089–1100.
13. Sohl-Dickstein, J., Weiss, E. A., Maheswaranathan, N., & Ganguli, S. (2015). Deep unsupervised learning using nonequilibrium thermodynamics. *Proceedings of the 32nd International Conference on Machine Learning*, 37, 2256–2265.
14. Ho, J., Jain, A., & Abbeel, P. (2020). Denoising diffusion probabilistic models. *Advances in Neural Information Processing Systems*, 33, 6840–6851.
15. Kingma, D. P., & Welling, M. (2014). Auto-encoding variational Bayes. *arXiv preprint arXiv:1312.6114*.
16. Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2014). Generative adversarial nets. *Advances in Neural Information Processing Systems*, 27, 2672–2680.
17. Topol, E. J. (2019). High-performance medicine: The convergence of human and artificial intelligence. *Nature Medicine*, 25(1), 44–56.
18. Rajkomar, A., Dean, J., & Kohane, I. (2019). Machine learning in medicine. *New England Journal of Medicine*, 380(14), 1347–1358.
19. Char, D. S., Shah, N. H., & Magnus, D. (2018). Implementing machine learning in health care—Addressing ethical challenges. *New England Journal of Medicine*, 378(11), 981–983.
20. Vayena, E., Blasimme, A., & Cohen, I. G. (2018). Machine learning in medicine: Addressing ethical challenges. *PLoS Medicine*, 15(11), e1002689.
21. National Academies of Sciences, Engineering, and Medicine. (2018). *Biodefense in the age of synthetic biology*. National Academies Press.
22. Oye, K. A., Esvelt, K., Appleton, E., Catteruccia, F., Church, G., Kuiken, T., Lightfoot, S. B.-Y., McNamara, J., Smidler, A., & Collins, J. J. (2014). Regulating gene drives. *Science*, 345(6197), 626–628.