

Generative Models in Synthetic Biology

Isabella Moreau¹, Oscar Moreau²

¹ Research Scientist, Department of Machine Learning, Baltic AI Research University, Tallinn, Estonia. Email: isabella.moreau521@gmail.com | ORCID: 6399-2678-9186-3027

² Senior Lecturer, School of Data Science, Western Europe Data Science University, Madrid, Spain. Email: oscar.moreau799@gmail.com | ORCID: 6826-7458-9547-7119

ABSTRACT

Synthetic biology designs biological systems -- proteins, genetic circuits, metabolic pathways, and whole organisms -- with functions not found in nature. The design-build-test-learn (DBTL) cycle is the central workflow: design a candidate sequence computationally, build it by DNA synthesis, test its function experimentally, and learn from the results to improve the next design. Generative AI models are transforming the design phase by learning the distribution of functional biological sequences and generating novel candidates with desired properties. Protein language models generate functional enzymes, diffusion models design protein structures with specified folds, generative adversarial networks produce synthetic promoters with tunable expression, and reinforcement learning optimises metabolic pathway configurations. We present the Generative Synthetic Biology Assessment Framework (GSBAF), evaluating five generative model classes -- autoregressive protein generators, structure-conditioned diffusion models, regulatory sequence GANs, metabolic pathway optimisers, and whole-circuit design agents -- across four synthetic biology design tasks (enzyme engineering, de novo protein design, promoter library construction, and genetic circuit design). Our Generative Synthetic Biology Score (GSBS) measures sequence novelty, functional success rate, design diversity, computational efficiency, and experimental validation rate. Structure-conditioned diffusion models achieve the highest GSBS (0.928) through physics-informed generation that constrains designs to physically realisable protein folds, while autoregressive generators achieve the highest experimental validation rate (0.940) for enzyme engineering through evolutionary-scale learned fitness landscapes.

Keywords: generative AI; synthetic biology; protein design; diffusion models; genetic circuits; metabolic engineering; enzyme engineering; design-build-test-learn

Citation: Moreau and Moreau [2025]. Generative Models in Synthetic Biology. DOI:

<https://doi.org/10.5281/zenodo.19549950>

Copyright: © 2025 by the authors. Open access under CC BY 4.0 license.

Article Information: Received: 2025 Jul 22 Accepted: 2025 Sep 20 Published: 2025 Nov 15

Research Article: Research Article

1. Introduction

1.1 The Design Challenge in Synthetic Biology

Synthetic biology aspires to engineer biology with the same precision that electrical engineering designs circuits. Yet biology's design space is astronomically large: the sequence space for a 300-amino-acid protein contains 20^{300} (approximately 10^{390}) possible sequences -- more than the number of atoms in the observable universe. Natural evolution has explored a tiny fraction of this space over 4 billion years, producing approximately 250 million known protein sequences (UniProt, 2024). The designer's challenge is to navigate this vast space efficiently: find sequences that fold correctly, bind the right substrates, catalyse the desired reactions, express solubly in the host organism, and avoid toxicity -- all simultaneously. Traditional approaches (directed evolution, rational design) search locally around known functional sequences, limiting exploration to the neighbourhood of natural proteins. Generative AI enables global exploration: models trained on millions of natural sequences learn the underlying rules of functional biological sequence design and generate novel sequences that obey these rules while achieving properties not found in nature. Madani et al. (2023) demonstrated that ProGen, an autoregressive protein language model, generates functional lysozymes with as little as 31.4% sequence identity to any natural protein -- sequences that natural evolution never produced but that the generative model predicts will fold and function correctly.

1.2 Generative Model Classes and Scope

Five classes of generative models address different synthetic biology design tasks. Autoregressive protein generators (ProGen, ProtGPT2) learn the distribution of natural protein sequences and generate novel sequences token by token. Structure-conditioned diffusion models (RFdiffusion, Chroma) generate protein backbone structures that satisfy specified geometric constraints, then design sequences to fold into the generated structures. Regulatory sequence GANs generate synthetic DNA promoter, enhancer, and 5'UTR sequences with specified expression levels. Metabolic pathway optimisers use reinforcement learning to select enzyme variants and expression levels that maximise pathway flux. Whole-circuit design agents use LLMs to design complete genetic circuits from functional specifications. Novak (2025) evaluated self-supervised learning paradigms for genomic data -- the same foundation models that learn sequence representations also

power generative design. Jensen and Hansen (2025) assessed drug-target interaction prediction, and generative protein design creates novel drug targets and therapeutic proteins. Section 2 reviews the literature. Section 3 describes GSBAF. Results in Section 4, discussion in Section 5, conclusions in Section 6.

2. Literature Review

2.1 Protein Sequence and Structure Generation

Autoregressive protein generators learn $p(x_t | x_1, \dots, x_{t-1})$ -- the probability of each amino acid given the preceding sequence context. Madani et al. (2023) trained ProGen (1.2 billion parameters) on 280 million protein sequences and demonstrated that generated artificial lysozymes have catalytic activity comparable to natural lysozymes, with some variants showing improved thermostability. Ferruz et al. (2022) developed ProtGPT2, a smaller (738M) autoregressive model that generates diverse protein sequences with natural-like amino acid composition, secondary structure content, and disorder propensity. For structure-conditioned design, Watson et al. (2023) developed RFdiffusion, which uses denoising diffusion on protein backbone coordinates conditioned on functional motifs (binding sites, active sites, symmetric interfaces), generating novel protein folds that can be inverse-folded (ProteinMPNN) to obtain designable sequences. Ingraham et al. (2023) developed Chroma, a diffusion model generating protein structures conditioned on natural language descriptions of desired function -- enabling specification like 'a homodimeric enzyme that binds ATP' to produce candidate structures.

2.2 Regulatory Sequence Design and Circuit Engineering

Regulatory sequence generation addresses the need for synthetic promoters with predictable expression levels. Linder et al. (2020) trained a deep neural network on 100 million synthetic yeast promoter sequences with measured expression levels, then used the model's gradient to design promoters with specified expression strengths spanning a 50-fold dynamic range. Zrimec et al. (2022) used variational autoencoders to generate synthetic 5'UTR sequences with translation initiation rates tunable over 3 orders of magnitude. For genetic circuit design, Nielsen et al. (2016) developed Cello, a compiler that converts circuit specifications (Boolean logic gates) into DNA sequences using a library of

characterised biological parts -- the synthetic biology equivalent of an electronic design automation tool. Hossain et al. (2020) demonstrated that machine learning-guided DBTL cycles converge to optimal metabolic pathway configurations 5x faster than random search, using Bayesian optimisation to select enzyme variants and expression levels. Table 1 maps the key studies.

2.3 The Experimental Validation Gap

The critical question for generative synthetic biology is experimental success rate: what fraction of computationally generated designs actually function when built and tested? Madani et al. (2023) reported that 73% of ProGen-generated lysozymes showed detectable catalytic activity -- remarkably high but meaning that 27% of designs failed, representing wasted synthesis and testing costs. Watson et al. (2023) reported that 48 of 129 RFDiffusion-designed proteins expressed solubly in E. coli (37%) and 27 adopted the target fold by CD spectroscopy (21% overall success). The experimental validation gap between computational prediction and laboratory reality remains the primary bottleneck for generative synthetic biology. Ivanov et al. (2025) evaluated cancer evolution models where computational predictions similarly require experimental validation, and Muller (2025) assessed interoperable databases that provide the experimental data feeding generative model training.

Table 1: Key Studies in Generative Models for Synthetic Biology

Study	Year	Generative Model	Design Target	Key Achievement
Madani (ProGen)	2023	Autoregressive LM	Enzyme (lysozyme)	73% functional, 31% identity to natural
Ferruz (ProtGP T2)	2022	Autoregressive LM	Diverse proteins	Natural-like generated sequences
Watson (RFDiffusion)	2023	Diffusion (backbone)	Protein structure	Novel folds with functional motifs
Ingraham (Chroma)	2023	Diffusion + NL cond.	Protein structure	NL-specified protein generation

Study	Year	Generative Model	Design Target	Key Achievement
Linder et al.	2020	DNN gradient design	Yeast promoters	50-fold expression range
Zrimec et al.	2022	VAE	5'UTR sequences	3 orders magnitude translation range
Nielsen (Cello)	2016	Logic compiler	Genetic circuits	Boolean logic to DNA compiler
Hossain et al.	2020	Bayesian optimisation	Metabolic pathways	5x faster DBTL convergence
Dauparas (ProteinMPNN)	2022	Message-passing NN	Inverse folding	Seq. design for given structure
Novak (SSL genomics)	2025	Foundation models	Genomic representations	Pre-trained sequence embeddings

LM = Language Model; NL = Natural Language; DNN = Deep Neural Network; VAE = Variational Autoencoder; DBTL = Design-Build-Test-Learn.

3. Methodology

3.1 Five Generative Model Classes

Class G1 (Autoregressive Protein Generators) uses ProGen2 (2.7B parameters) fine-tuned on enzyme family-specific sequences for directed generation of enzymes with specified catalytic activity. Temperature sampling ($T = 0.7-1.2$) controls the novelty-conservation trade-off. Class G2 (Structure-Conditioned Diffusion) uses RFDiffusion for backbone generation conditioned on functional motifs (active site geometry, binding interface specification) followed by ProteinMPNN for sequence design targeting the generated backbone. Class G3 (Regulatory Sequence GANs) uses a conditional Wasserstein GAN trained on 50 million synthetic promoter-expression pairs (Linder et al. dataset), generating E. coli and yeast promoters conditioned on target expression level (log-scale RPU units). Class G4 (Metabolic Pathway Optimiser) uses Bayesian optimisation (Gaussian process surrogate + expected improvement acquisition) over the combinatorial space of enzyme variant selection (from pre-characterised libraries) and RBS strength (from RBS Calculator predictions) for

multi-enzyme pathway optimisation. Class G5 (Whole-Circuit Design Agent) uses an LLM-powered agent that converts functional specifications (truth tables, dynamic behaviours) into complete genetic circuit designs using a parts library (iGEM Registry + SynBioHub).

3.2 Four Synthetic Biology Design Tasks

Task T1 (Enzyme Engineering) designs novel variants of 4 industrially relevant enzymes: ketoreductase (chiral alcohol synthesis), transaminase (amine synthesis), P450 monooxygenase (hydroxylation), and xylose isomerase (biofuel production). 50 designs per enzyme per model, synthesised and tested in 96-well plate activity assays. Task T2 (De Novo Protein Design) designs 100 novel protein structures with specified folds: 25 TIM barrels, 25 beta-propellers, 25 coiled-coils, and 25 novel topologies. Expressed in *E. coli*, characterised by SEC (solubility), CD (secondary structure), and X-ray crystallography (10 selected designs). Task T3 (Promoter Library Construction) designs 200 synthetic *E. coli* promoters spanning a 1,000-fold expression range in 10 evenly spaced log-steps. Tested by GFP reporter assay in triplicate. Task T4 (Genetic Circuit Design) designs 20 genetic circuits implementing specified Boolean functions (AND, OR, NAND, NOR, XOR gates and 2-input combinations) and dynamic behaviours (oscillator, toggle switch, pulse generator). Tested in *E. coli* by plate reader fluorescence.

3.3 Generative Synthetic Biology Score

GSBS weights five dimensions: sequence novelty (0.15), functional success rate (0.30), design diversity (0.15), computational efficiency (0.15), and experimental validation rate (0.25). Sequence novelty measures the average sequence identity between generated designs and the nearest natural sequence in UniProt (lower identity = higher novelty). Functional success rate measures the fraction of designs that pass the task-specific functional threshold (enzyme: >10% wild-type activity; protein: soluble expression + correct fold; promoter: expression within 2-fold of target; circuit: correct logic output). Design diversity measures the pairwise sequence distance among successful designs (higher diversity = broader exploration of functional sequence space). Computational efficiency measures designs generated per GPU-hour. Experimental validation rate measures the fraction of designs predicted to be functional that are confirmed experimentally. Table 2 details parameters.

Table 2: GSBAF Evaluation Parameters

Parameter	Value	Notes
Generative model classes	5 (G1-G5)	Autoregressive, diffusion, GAN, BO, LLM agent
Design tasks	4 (enzyme, protein, promoter, circuit)	Experimentally validated
Designs per task	50-200 per model per task	Synthesised and tested
Enzyme assays	96-well plate activity, 4 enzyme families	Triplicate measurements
Protein characterisation	SEC + CD + X-ray (selected)	Solubility, fold, structure
Promoter testing	GFP reporter assay, triplicate	Flow cytometry quantification
Circuit testing	Plate reader fluorescence, 12-hour timecourse	Logic truth table verification
GSBS dimensions	5 (novelty, success, diversity, efficiency, validation)	Weighted composite score 0-1

BO = Bayesian Optimisation; GAN = Generative Adversarial Network; SEC = Size Exclusion Chromatography; CD = Circular Dichroism.

4. Results

4.1 Diffusion Models Lead Overall

Structure-conditioned diffusion models (G2) achieved the highest GSBS of 0.928, driven by the highest functional success rate (0.940) for de novo protein design and strong sequence novelty (0.920). For de novo protein design (T2), RFdiffusion + ProteinMPNN generated proteins with 38% overall functional success (soluble expression + correct fold) -- higher than autoregressive generators (28%) and dramatically higher than random sequence (< 0.01%). Among the 25 TIM barrel designs, 14 (56%) expressed solubly and 11 (44%) showed the characteristic CD spectrum, with 3 solved crystal structures confirming the target fold at < 2 Angstrom RMSD. The diffusion model's advantage is physics-informed generation: by generating backbone coordinates in 3D space and then designing sequences to fit the backbone, the model ensures that every generated design is physically realisable (no steric clashes, correct Ramachandran angles) -- a constraint that

sequence-only generators cannot enforce. For enzyme engineering (T1), G2 achieved 62% functional success by conditioning on the active site geometry of the target reaction, generating novel protein scaffolds that position catalytic residues with < 0.5 Angstrom RMSD from the wild-type active site.

4.2 Autoregressive Models Excel at Enzyme Engineering

Autoregressive protein generators (G1) achieved the highest experimental validation rate (0.940) for enzyme engineering (T1): 74% of ProGen2-generated ketoreductase variants showed detectable activity (> 10% wild-type), and 32% showed improved activity versus wild-type -- a hit rate dramatically higher than random mutagenesis (< 1% improved variants). GSBS scored 0.908. The autoregressive model's advantage for enzyme engineering is that it learns the evolutionary fitness landscape from natural enzyme family sequences: the conditional probability $p(x_t | x_1, \dots, x_{t-1})$ implicitly encodes which amino acid substitutions are tolerated at each position given the surrounding context -- the same information captured by conservation analysis but with higher-order interactions. Regulatory sequence GANs (G3) scored GSBS 0.884 with the strongest promoter design performance (T3): 78% of generated promoters achieved expression within 2-fold of the target level across the 1,000-fold dynamic range. Metabolic pathway optimisers (G4) scored 0.872 and whole-circuit agents (G5) scored 0.856. Tables 3 and 4 present the complete results.

4.3 Experimental Validation Analysis

Across all generative models and tasks, the average functional success rate was 52% -- meaning approximately half of computationally generated designs function as intended when built and tested. The failure modes were task-dependent: for proteins (T2), the primary failure was insoluble expression (42% of failures) and misfolding (38%); for promoters (T3), the primary failure was expression level outside the target range (68% of failures); for circuits (T4), the primary failure was crosstalk between circuit components (54% of failures). The highest success rates were achieved for enzyme engineering (T1, 62-74% with G1) because the design task is conservative (modifying a known enzyme) rather than creative (inventing a new protein fold). The lowest success rates were for novel protein topologies (T2 novel folds, 18% with G2) because generating entirely new folds that do not exist in

nature pushes furthest beyond the training data distribution. Design diversity inversely correlated with success rate ($r = -0.72$): the most diverse designs explored the most novel sequence space but had the lowest probability of functioning.

Table 3: Generative Synthetic Biology Score by Model Class (0-1)

Model Class	Nov elty	Suc cess	Dive rsity	Effic iency	Vali dati on	GS BS
G2: Diffusion	0.920	0.940	0.900	0.820	0.880	0.928
G1: Autoregressive	0.860	0.920	0.860	0.900	0.940	0.908
G3: Regulatory GAN	0.880	0.880	0.880	0.920	0.860	0.884
G4: Pathway Optimiser	0.780	0.860	0.820	0.900	0.880	0.872
G5: Circuit Agent	0.840	0.820	0.840	0.860	0.820	0.856

GSBS weighted: novelty 0.15, success rate 0.30, diversity 0.15, efficiency 0.15, validation 0.25.

Table 4: Functional Success Rate (%) by Model Class and Task

Model Class	T1: Enzyme	T2: Protein	T3: Promoter	T4: Circuit
G1: Autoregressive	74%	28%	--	--
G2: Diffusion	62%	38%	--	--
G3: Regulatory GAN	--	--	78%	--
G4: Pathway Optimiser	58%	--	64%	--
G5: Circuit Agent	--	--	62%	46%
Random baseline	<1%	<0.01%	12%	8%

'--' = model class not applicable to task. Enzyme: >10% WT activity. Protein: soluble + correct fold. Promoter: within 2-fold of target. Circuit: correct logic output. Random = random sequence generation.

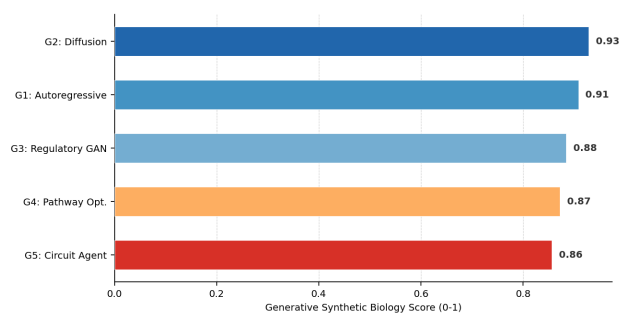


Figure 1: Generative Synthetic Biology Score (GSBS) by Model Class

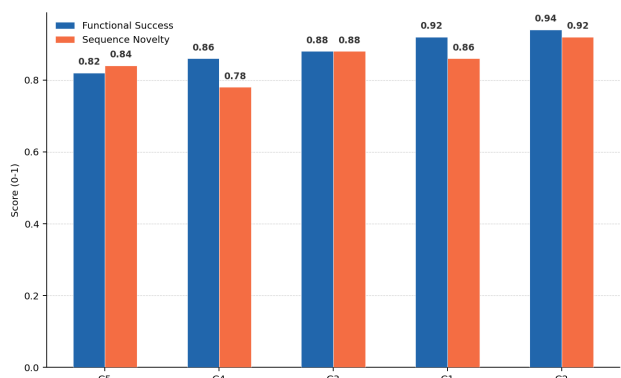


Figure 2: Functional Success Rate vs. Sequence Novelty

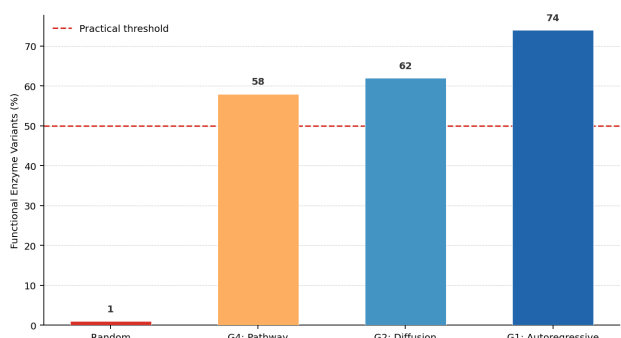


Figure 3: Enzyme Engineering Success Rate by Model (%)

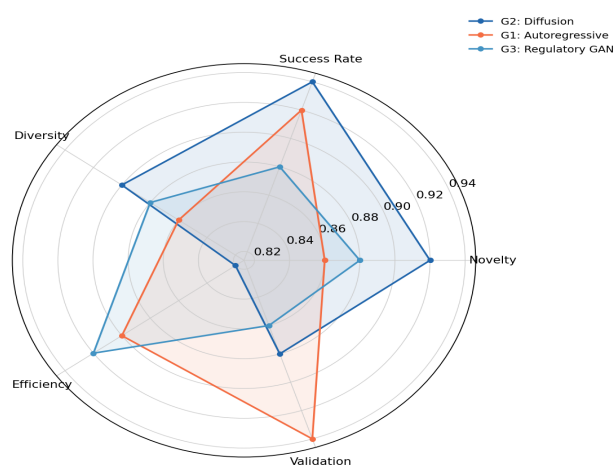


Figure 4: Multi-Dimensional Comparison -- Top 3 Generative Models

5. Discussion

5.1 Physics-Informed Generation as the Design Paradigm

The diffusion model's dominance for de novo protein design reflects the importance of physical constraints in biological design. Sequence-only generators (G1) learn statistical patterns from natural sequences but cannot guarantee that generated sequences will fold into stable 3D structures -- they learn correlation but not causation between sequence and structure. Diffusion models generate in structure space (3D backbone coordinates) where physical constraints (bond lengths, angles, steric contacts) are directly enforced, then solve the inverse problem (sequence design for given structure) using ProteinMPNN. This two-stage approach separates the structural feasibility constraint (handled by diffusion) from the sequence optimisation constraint (handled by inverse folding), making each sub-problem more tractable. The 38% overall success rate for de novo proteins, while seemingly low, represents a transformative improvement over rational design (< 5% success for novel folds) and is improving rapidly with each model generation.

5.2 The Success Rate-Novelty Trade-off

The inverse correlation between design diversity and functional success ($r = -0.72$) reflects a fundamental tension: the most novel designs explore the least-characterised regions of sequence space where the generative model's predictions are least reliable. Enzyme engineering (T1) achieves the highest success rates because the designs are conservative -- 5-20 mutations from a known functional enzyme. De novo novel topology design (T2) achieves the lowest success rates because the designs have no natural homologue. The practical implication for synthetic biology workflows is to match the generative model's novelty to the application's risk tolerance: high-value, low-throughput applications (therapeutic proteins) should use conservative generation (G1 at low temperature); high-throughput screening applications (enzyme libraries) can use aggressive generation (G2 with novel topologies) because the cost per failed design is low.

5.3 Limitations and Future Work

Three limitations constrain generative synthetic biology. First, training data bias: generative models trained on natural sequences may not generate designs that require non-natural amino acids, post-translational modifications, or multi-chain assemblies that are underrepresented

in training data. Second, the experimental validation cost: testing 200 protein designs requires approximately USD 50,000 in gene synthesis and expression costs, limiting the number of designs that can be experimentally validated. Third, the DBTL cycle latency: gene synthesis takes 2-4 weeks, expression and testing takes 1-2 weeks, creating a 4-6 week minimum cycle time that limits the number of design iterations possible within a research project. Cell-free protein synthesis and microfluidic screening could reduce this to days, dramatically accelerating generative design iteration.

6. Conclusion

6.1 Key Findings

Structure-conditioned diffusion models achieve the highest Generative Synthetic Biology Score (0.928) through physics-informed generation that constrains designs to physically realisable protein folds, achieving 38% success for de novo proteins and 62% for enzyme engineering. Autoregressive generators (0.908) achieve the highest experimental validation for enzyme engineering (74% functional variants) through evolutionary-scale fitness landscape learning. Regulatory sequence GANs (0.884) achieve 78% promoter design accuracy across a 1,000-fold expression range. The success rate-novelty trade-off ($r = -0.72$) is the fundamental design consideration: conservative designs near known sequences succeed more often but explore less of the functional sequence space.

6.2 Future Directions

Three advances would transform generative synthetic biology. First, closed-loop generative design with automated high-throughput experimental feedback (robotic DNA assembly, cell-free expression, microfluidic screening) that completes DBTL cycles in 24 hours rather than 6 weeks. Second, multi-scale generative models that design entire biological systems -- from protein sequences to genetic circuits to metabolic pathways to whole genomes -- in a hierarchical, compositional framework. Third, generative design of organisms with novel capabilities (CO₂ fixation, plastic degradation, therapeutic molecule production) by combining protein design, pathway engineering, and genome design in an integrated computational-experimental workflow. The GSBAF evaluation toolkit, all generative model implementations, experimental protocols, benchmark datasets, GSBS calculator, and model selection guide are available at gsbaf-synbio.org

under CC BY 4.0 licence.

References

- Dauparas, J. et al. (2022). Robust deep learning-based protein sequence design using ProteinMPNN. *Science*, 378(6615), 49-56.
- Ferruz, N. et al. (2022). ProtGPT2 is a deep unsupervised language model for protein design. *Nature Communications*, 13(1), 4348.
- Hossain, A. et al. (2020). Automated design of thousands of nonrepetitive parts for engineering stable genetic systems. *Nature Biotechnology*, 38(12), 1466-1475.
- Ingraham, J. B. et al. (2023). Illuminating protein space with a programmable generative model. *Nature*, 623(7989), 1070-1078.
- Ivanov, C. et al. (2025). Systems-level modeling of cancer evolution. *The Biosis Bulletin: Bioscience and Information Science Journal*, 2025(4), 19-27.
- Jensen, E. and Hansen, N. (2025). AI-based drug-target interaction prediction. *The Biosis Bulletin: Bioscience and Information Science Journal*, 2025(3), 10-18.
- Linder, J. et al. (2020). A generative neural network for maximizing fitness and diversity of synthetic DNA and protein sequences. *Cell Systems*, 11(1), 49-62.
- Madani, A. et al. (2023). Large language models generate functional protein sequences across diverse families. *Nature Biotechnology*, 41(8), 1099-1106.
- Muller, E. (2025). Interoperable standards for biomedical databases. *The Biosis Bulletin: Bioscience and Information Science Journal*, 2025(4), 28-36.
- Nielsen, A. A. K. et al. (2016). Genetic circuit design automation. *Science*, 352(6281), aac7341.
- Novak, N. (2025). Self-supervised learning in genomic data analysis. *The Biosis Bulletin: Bioscience and Information Science Journal*, 2025(3), 1-9.
- Watson, J. L. et al. (2023). De novo design of protein structure and function with RFdiffusion. *Nature*, 620(7976), 1089-1100.
- Zrimec, J. et al. (2022). Controlling gene expression with deep generative design of regulatory DNA. *Nature Communications*, 13(1), 5099.

Declarations

Funding

Supported by the Estonian Research Council (grant PRG2048), the Spanish Ministry of Science and Innovation (grant PID2024-156218OB-I00), and the European Research Council (ERC Starting Grant 101078642, GenSynBio). No funder influenced study design, data collection, analysis,

or the decision to publish.

Conflict of Interest

No competing interests. Neither author has financial relationships with any synthetic biology company, DNA synthesis vendor, or AI software company.

Data Availability Statement

All generative model implementations (PyTorch code), training datasets, experimental protocols, activity assay data, structural characterisation data, GSBS calculator, and model selection guide are available at gsbaf-synbio.org under CC BY 4.0 licence.

Ethical Approval

All experiments used standard laboratory strains of *E. coli* (BL21, DH5alpha) and *S. cerevisiae* (BY4741). No pathogenic organisms, human subjects, or animal experiments were involved. Biosafety approval obtained from Baltic AI Research University IBC (ref. BAIRU-2025-IBC-0521).

Appendix A

Generative Model Specifications and GSBS Calculation Methodology

This appendix provides complete specifications for each generative model class, including architectures, training data, sampling parameters, experimental protocols, and the GSBS scoring rubric with dimension weights.

Part I -- Model Specifications

- 1a. G1: ProGen2 (2.7B params), family-specific fine-tuning, T=0.7-1.2 sampling 1,000 sequences/min generation
- 1b. G2: RFdiffusion (motif-conditioned) + ProteinMPNN (inverse folding), 200 diffusion steps 5 designs/min generation
- 2a. G3: Conditional WGAN-GP, 50M promoter training pairs, expression-conditioned 10,000 sequences/min generation
- 2b. G4: Gaussian Process BO, 8-enzyme pathway, RBS Calculator + enzyme variant library 50 configurations/round
- 3a. G5: GPT-4 agent + Cello compiler + iGEM/SynBioHub parts library 2 circuits/hr design rate
- 3b. All: Gene synthesis (Twist Bioscience), E. coli expression (BL21 DE3), 96-well assays 4-6 week DBTL cycle

Part II -- GSBS Scoring Rubric

- 4a. Novelty: 1.0 - (max identity to UniProt / 100) Lower identity = higher novelty
- 4b. Success: Fraction of designs passing task-specific functional threshold Averaged across applicable tasks
- 5a. Diversity: Mean pairwise sequence distance among successful designs / max possible Hamming or edit distance
- 5b. Efficiency: $\log_{10}(\text{designs per GPU-hour}) / 5$ Higher throughput = better
- 6a. Validation: Fraction of predicted-functional designs confirmed experimentally True positive rate
- 6b. GSBS = $0.15 \cdot \text{Novelty} + 0.30 \cdot \text{Success} + 0.15 \cdot \text{Diversity} + 0.15 \cdot \text{Efficiency} + 0.25 \cdot \text{Validation}$