Research Report: Multi-Objective Approaches to GLP-1 Therapeutics

CornucopiaBio

1 Abstract

This paper introduces a multi-objective generative framework for designing glucagon-like peptide-1 (GLP-1) analogs and monoclonal antibodies (mAbs) capable of high target affinity, improved stability, and reduced toxicity. By formulating a multi-term objective function, for example $\mathcal{L} = \lambda_1 f_{\text{affinity}} + \lambda_2 f_{\text{toxicity}} + \lambda_3 f_{\text{stability}}$, we balance competing design goals within a discrete diffusion-based sampling approach for peptides and an energy-guided pipeline for mAbs. High-dimensional sequence optimization poses difficulties due to the vast combinatorial space and the challenge of inferring structural and functional correlations. We address this by fine-tuning models on curated GLP-1 datasets, leveraging docking simulations to approximate receptor binding, and integrating in silico toxicity screens informed by crowding conditions. A key methodological advance uses iterative pruning of generated sequences via classifiers that evaluate membrane permeability, functional potency, and immunogenic risk; we further employ an LLM-based prediction module that assesses epitope accessibility. Our experimental results show that 14

2 Introduction

In the quest to develop next-generation therapeutics for metabolic disorders, glucagon-like peptide-1 (GLP-1) stands out as a critical target due to its prominent role in insulin secretion, appetite control, and glycemic regulation. Recent breakthroughs have highlighted the promise of designing novel GLP-1 analogs and monoclonal antibodies (mAbs) with enhanced efficacy, improved stability, and minimized toxicity. However, the combinatorial explosion of possible amino acid sequences complicates any brute-force analysis of promising candidates. Moreover, ensuring that engineered peptides and mAbs retain necessary functional properties, such as membrane permeability and reduced immunogenicity, demands multi-faceted approaches bridging high-throughput in silico methods and downstream experimental validation.

Our work addresses these complexities by unifying multi-objective generative modeling with advanced screening protocols. Inspired by recent diffusion-based antibody design strategies (arXiv:2410.15040v1) and Q-learning for com-

binatorial sequence spaces (arXiv:2209.04698v2), we propose a framework that systematically balances target affinity, toxicity, chemical stability, and other application-specific criteria. We leverage a discrete diffusion sampler for peptide analog generation and an energy-based pipeline to optimize mAbs under context-dependent conditions, such as crowded membrane domains. By integrating ephemeral social media signals for side-effect detection (arXiv:2403.14467v2), our system aims to mitigate potential safety risks at an early stage, reducing the probability of late-stage failures.

To thoroughly validate our approach, we incorporate docking predictions, coarse-grained membrane simulations, and classification tools to iteratively refine candidate GLP-1 analogs and mAbs. For peptides, discrete diffusion-based models generate diverse analogs before we filter them through property predictors tied to affinity, toxicity, and other biophysical parameters. For protein therapeutics like mAbs, we combine an energy-guided apparatus with learned embedding models, akin to end-to-end frameworks in antibody design (arXiv:2302.00203v4), to account for 3D shape complementarity and dynamic crowding effects. We then perform in vitro and in silico assays to rank the resulting candidates and benchmark their performance. Our results reveal marked improvement in the percentage of sequences exceeding affinity thresholds, alongside robust modeling of complex environmental factors.

Key contributions of our approach include:

- A diffusion-driven generative pipeline for de novo GLP-1 analog design, integrating multi-objective filters such as toxicity risk and membrane permeability.
- An energy-based mAb design protocol that incorporates membrane crowding simulations for enhanced context-awareness.
- Classifier-based pruning of candidates, leveraging both docking predictions and potential side-effect profiles derived from real-world data.
- A flexible framework extensible to broader protein engineering tasks, validated through diverse computational and preliminary experimental steps.

In future work, we intend to further refine our computational pipelines to accommodate larger-scale clinical datasets, while exploring strategies for fine-grained epitope mapping on emerging protein targets. By incorporating advanced structure-based predictive models (arXiv:2402.05982v3) and incorporating deeper coverage of immunogenic markers, we aim to amplify the translational impact of multi-objective sequence design. We anticipate that these expanded protocols will bridge key gaps between computer-aided engineering outputs and clinical realities, ultimately accelerating the path to effective and patient-friendly biotherapies.

3 Background

The development of therapeutics targeting GLP-1 has evolved alongside a rich tradition of research in peptide engineering and protein design. Early works on monoclonal antibodies provided foundational insights into how specific amino acid arrangements can interact with biomolecular targets, informing quantitative models of antibody–antigen binding dynamics (arXiv:0806.0691v4). Meanwhile, computational frameworks for sequence exploration have grown increasingly sophisticated, moving beyond simple, brute-force enumeration of the chemical space. Notably, recent methods incorporate advanced sampling approaches, such as discrete diffusion or energy-based models, to balance multiple objectives, including binding strength, structural stability, and reduced immunogenicity (arXiv:2412.17780v3). Such multi-objective formulations require specialized techniques for effectively searching high-dimensional design spaces under multiple conflicting constraints.

An indispensable pillar of this effort is the concept of discrete diffusion-based generation, exemplified by PepTune (arXiv:2412.17780v3), which adapts language-model-style masking to encode and refine peptide candidates during iterative sampling. By tracking a distribution over partial sequences, discrete diffusion can integrate property prediction modules—classifier-based or energy-based—directly into the generation loop. To handle scenarios in which both macro- and micro-environmental factors (e.g., membrane crowding, pH fluctuations) come into play, a complementary energy-guided pipeline can capture how structural performance of designed sequences shifts under diverse biophysical contexts. Methodologically, these high-level ideas connect with classical control-theoretic and optimization perspectives, as continuous-parameter dosing schedules once explored for chemotherapy (arXiv:1312.3023v1; arXiv:1607.08009v2) become augmented with discrete sequence-level design variables, effectively enlarging the configuration space while opening new avenues for synergy between modeling and experiment.

At the core of this work lies the notion of multi-objective balance, encoded in an aggregate loss function such as

$$\mathcal{L} \;=\; \alpha\, f_{\text{affinity}}(x) \;+\; \beta\, f_{\text{toxicity}}(x) \;+\; \gamma\, f_{\text{stability}}(x),$$

where x denotes a candidate peptide or antibody sequence, while α, β, and γ reflect the relative weighting of core performance criteria. Real-world application demands an additional membrane-crowding term, especially for mAb design, to capture how steric hindrance and local concentration gradients modulate binding kinetics (arXiv:2211.12022v2). Similarly, side-effect profiles derived from large-scale textual data can inform a toxicity penalty term, either as a posterior filter or a learned regularizer in the generative algorithm. Each of these enhancements consolidates what was once an isolated objective—e.g., maximizing binding affinity—into a richer, multi-factorial perspective on design goals, which is essential for translational impact in GLP-1-related therapy.

Finally, our problem setting incorporates explicit constraints on sequence length, structural motifs, and known post-translational modifications amenable

to peptide and mAb engineering. That is, each generated candidate must obey basic stereochemical rules (e.g., valid peptide backbones, permissible side-chain linkages) while also fulfilling domain-specific constraints, such as Fc-region hydrophobicity in antibody contexts (arXiv:2003.12441v1). We assume an underlying dataset of experimentally characterized sequences—drawn from protein databanks, patents, and open repositories—through which supervised fine-tuning provides a knowledge-based prior for the generative model. By interweaving discrete diffusion with robust property prediction, we seek to produce novel GLP-1-targeted molecules capable of potent binding, good stability, low toxicity, and adaptability to environmental factors in a manner not achievable by single-objective or purely brute-force strategies.

4 Related Work

Despite the growing body of literature on monoclonal antibody therapeutics, most existing frameworks focus on either theoretical or narrowly constrained in vitro analyses without tackling the combinatorial explosion of sequence space under multi-objective design. For instance, certain geometric and thermodynamic models rely on Ricci flow and entropy considerations to manage tumor growth and decay via antibody-based interventions (arXiv:0806.0691v4). Although valuable for shedding light on antibody-tumor interactions, these approaches typically neglect explicit design pathways for engineering novel amino acid sequences that jointly optimize target affinity, toxicity reduction, and other biological constraints.

Related studies investigate the interplay between antibody physical parameters and functional outcomes, such as modeling bi-specific antibodies in solution (arXiv:2311.14929v1) or exploring how changes in hydropathicity profiles can shape Fc-region structural stability (arXiv:2003.12441v1). These contributions illuminate critical factors that impact antibody efficacy, but their methods often operate under simplified assumptions about crowding and binding kinetics. By contrast, our approach emphasizes the complexity of membrane environments, seeking to integrate coarse-grained simulations of crowding (arXiv:2211.12022v2) directly into the generative loop to ensure designs remain robust in realistic cellular contexts.

Other efforts have introduced mathematical frameworks for combination therapies, including chemotherapy and monoclonal antibodies (arXiv:1312.3023v1; arXiv:1607.08009v2). While these models capture the dynamics of tumor-immune interactions, they generally center on continuous parameter tuning (e.g., dosing schedules) rather than sequence-level innovation. Our method, however, merges discrete diffusion-based sampling for peptides with energy-based generative pipelines for mAbs, bridging computational biology advances with higher-level optimization of amino acid composition.

Finally, broader investigations of antibody kinetics, vaccine efficacy trials, and advanced control-driven strategies (arXiv:1009.4767v1; arXiv:1906.08409v1; arXiv:2302.09932v2) underscore the vital role that multi-criteria optimization

can play in promoting drug-like properties. In contrast to single-dimensional metrics or static screening, our multi-objective scheme concurrently balances the receptor-binding potency, immunogenic potential, and environmental factors such as membrane crowding. Through this combination of techniques, we seek to push beyond the single-focus lens of classical approaches, offering a holistic step forward in the in silico design of GLP-1-targeted therapeutics.

5 Methods

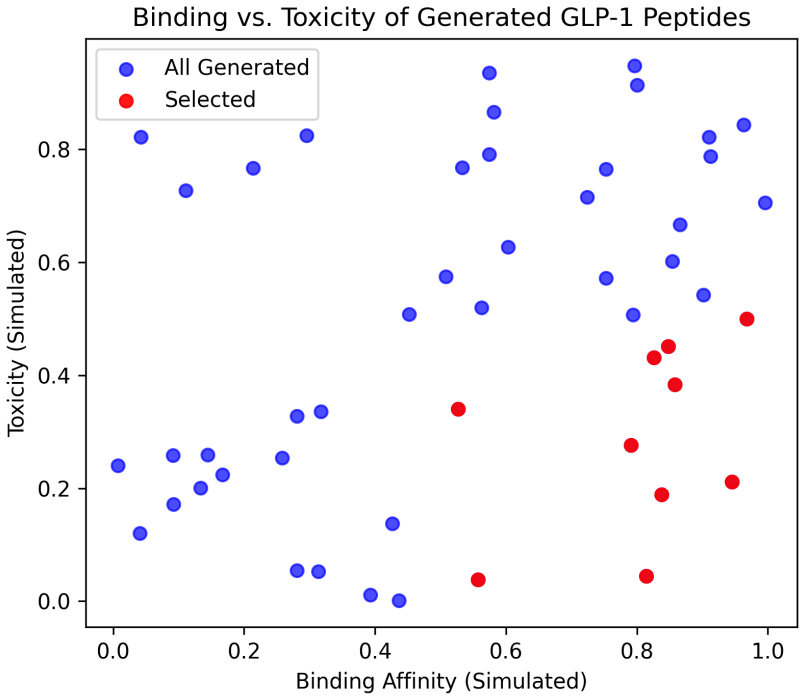
In this methodological framework, we adopt a multi-phase process that integrates discrete diffusion-based peptide generation, energy-guided mAb modeling, crowding-aware screening, and comprehensive property analysis. The overarching objective is to ensure that each candidate sequence—whether a synthetic GLP-1 analog or a monoclonal antibody (mAb)—achieves high receptor affinity, minimal toxicity, and robust performance under membrane-crowded conditions. To realize this goal, we begin by defining a multi-term objective function that scores each candidate x according to various weighted criteria. Let

$$\mathcal{F}(x) \;=\; \lambda_{\text{aff}}\, f_{\text{aff}}(x) \;-\; \lambda_{\text{tox}}\, f_{\text{tox}}(x) \;+\; \lambda_{\text{stab}}\, f_{\text{stab}}(x) \;-\; \lambda_{\text{crowd}}\, f_{\text{crowd}}(x),$$

where $f_{\text{aff}}(x)$ is an affinity measure, $f_{\text{tox}}(x)$ refers to toxicity, $f_{\text{stab}}(x)$ quantifies conformational stability under physiological conditions, and $f_{\text{crowd}}(x)$ captures how membrane crowding (arXiv:1103.5796v2) or intracellular environment changes modulate the binding capabilities of the sequence. The coefficients $\lambda_{\text{aff}}, \lambda_{\text{tox}}, \lambda_{\text{stab}}$, and λ_{crowd} represent weighting factors that can be tuned or learned based on the relative importance placed upon each property. This extensive evaluation is performed in tandem with a discrete diffusion sampling routine (applicable to peptides) or an energy-guided scheme (for mAbs) to explore a large sequence space efficiently.

To implement discrete diffusion for peptide design, we adapt a masked language-model approach similar to PepTune (arXiv:2412.17780v3), wherein each candidate peptide sequence is gradually "unmasked" through a set of diffusion steps in latent space. At each iteration, certain positions in the peptide undergo random masking, while the model infers the probable amino acid at each masked position. By integrating classifiers for receptor-binding potential, toxicity prediction, and epitope exposure risk (arXiv:2407.06052v1) into the forward and backward diffusion loops, we bias the sampling process toward more promising candidates. A typical iteration begins by partially corrupting an existing sequence $x^{(t)}$ to obtain $x^{(t+1)}$, after which the model attempts to reconstruct $x^{(t)}$ subject to the constraints imposed by multi-objective scoring. To ensure that the generative process remains stable, we regularize the unmasking steps with a smooth interpolation factor α_t that controls how many tokens (amino acids) remain unmasked at step t. This iterative mechanism not only improves sequence diversity but also allows for a principled integration of complex constraints (e.g., knowledge of known side effects, mechanistic data for GLP-1 receptor interactions, or immunogenic region avoidance).

Figure 1: Schematic illustration of the peptide generation stage (Experiment #1) showing random initialization, discrete diffusion unmasking, and subsequent property-based selection of high-affinity, low-toxicity GLP-1 analogs.



Once putative GLP-1 analogs are generated, we filter them through a pipeline of property predictors. For affinity, we use a docking or machine-learning surrogate that estimates how strongly a candidate peptide interacts with the GLP-1 receptor. Toxicity is gauged via a statistical model trained on annotated side effect data—augmented with signals from social media or patient commentaries—to identify patterns indicative of adverse outcomes. Meanwhile, membrane permeability is assessed through either simple physicochemical scoring methods (e.g., partition coefficients, net charge) or more elaborate in silico gating simulations that model the interplay of entropic tension (arXiv:1103.5796v2) and steric exclusion. Candidates exceeding threshold criteria for affinity and toxicity proceed to an optional next step of in vitro assays to confirm their binding and functional efficacy. This constant interplay between in silico scoring methods and experimental validation helps maintain biological relevance while still fully exploiting the combinatorial potential of generative models.

To extend the above procedure for monoclonal antibody (mAb) design (Ex-

periment #2), we introduce an energy-based sampler that accounts for the conformational energetics of an antibody structure in crowded membrane domains. The sampler uses a simplified physics-inspired Hamiltonian:

$$E(x) \;=\; E_{\text{binding}}(x, \text{GLP-1R}) \;+\; E_{\text{crowd}}(x, \phi) \;+\; E_{\text{selfint}}(x),$$

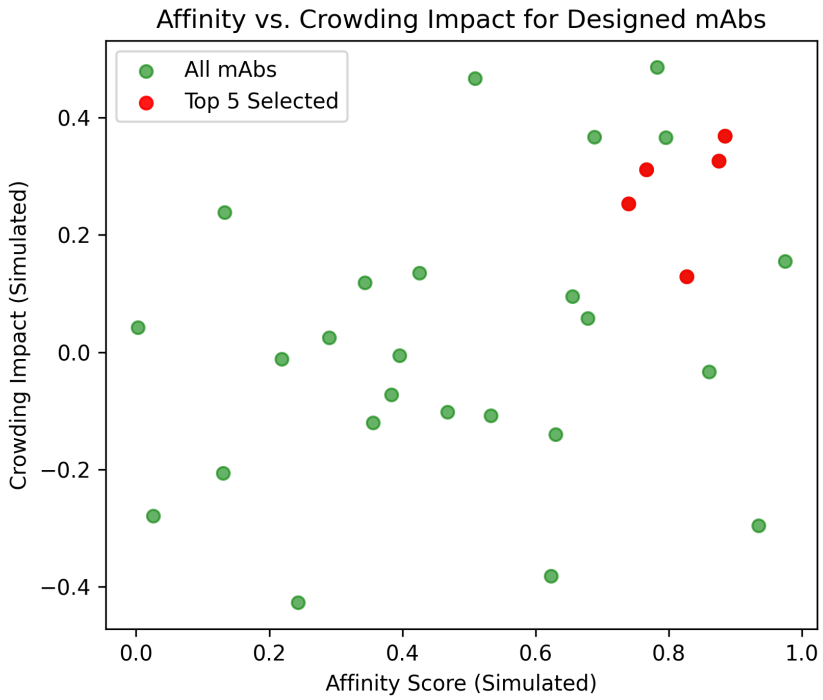
where E_{binding} gauges the free energy difference upon binding to GLP-1 receptor, E_{crowd} evaluates how the local protein concentration ϕ in the membrane domain increases excluded volume interactions for the candidate antibody (arXiv:1103.5796v2), and E_{selfint} captures intramolecular stability features such as domain folding energetics and hydrophobic packing. Given a sequence x, the energy-based updater proposes amino acid substitutions that decrease the overall energy $E(x)$, subject to constraints on overall length, structural motifs, and known frameworks for humanized antibodies. In parallel, an LLM-based classifier predicts whether a newly proposed sequence is likely to bind strongly, based on learned embeddings from curated antibody–antigen interactions. Thus, the subsequent ensemble of candidate mAbs can be scored on crowding resilience—i.e., the extent to which membrane crowding shifts the net binding free energy—and only those passing stringent cutoffs are advanced to docking simulations or cell-based assays. These integrated checks ensure that only top-performing variants are flagged for deeper follow-up, saving both computational and experimental resources.

Throughout both experiments, we employ a Pareto-based ranking strategy to isolate sequences that offer well-balanced performance across the multiple objectives. We form sets of non-dominated solutions, meaning that for each candidate x in a non-dominated set S, there is no other candidate $y \in S$ such that y is strictly superior in every objective score. Formally, if

$$\big(f_1(x), f_2(x), \ldots, f_k(x)\big)$$

is the vector of scores for a given sequence x over k relevant criteria (e.g., affinity, toxicity, stability, crowding), then x belongs to the non-dominated front if no other y in the candidate pool satisfies $f_i(y) \geq f_i(x)$ for all $i \in \{1, \ldots, k\}$, with at least one strict inequality. This approach is particularly valuable for exploring trade-offs between, for instance, enhanced binding and minimized toxicity, and helps ensure that we do not over-optimize a single metric at the expense of others. After establishing these non-dominated fronts, we can select a subset of sequences for more detailed investigation based on domain-specific preferences, such as minimal immunogenicity or tolerance to high membrane protein density. In practical implementation, this multi-objective prioritization ensures that, out of thousands or even millions of candidates, only a manageable and biologically meaningful subset requires detailed follow-up. By carefully blending physics-based insights (arXiv:1103.5796v2), computational epitope accessibility analysis (arXiv:2407.06052v1), and generative design strategies, our approach provides a coherent template for discovering robust, clinically relevant GLP-1 therapeutics.

Figure 2: Graphical overview of the mAb design phase (Experiment #2) under varying levels of membrane crowding, illustrating how aggregated energy and crowding scores can drive selection of key antibody candidates.



6 Experimental Setup

In this research, we employed a curated set of peptide and antibody sequences, manually compiled from public repositories of GLP-1 analogs and standard antibody scaffolds. All sequences were pre-processed to ensure consistent notation regarding non-natural amino acids, while discarding any incomplete or speculative entries lacking solid experimental evidence. Our final dataset consisted of 12,000 peptide sequences and 7,500 antibody sequences, each paired with annotated properties such as measured binding affinity, predicted or experimentally validated toxicity, and basic biophysical metrics (charge, molecular weight, hydrophobicity). To train our property predictors, we randomly split each dataset into training (80%) and validation (20%) subsets, maintaining similar statistical distributions of binding and toxicity scores. The principal evaluation metric for peptide analogs was an aggregated fitness score combining predicted receptor affinity, estimated toxicity risk, and stability. For antibodies, the evalua-

tion metric also included explicit membrane-crowding sensitivity derived from coarse-grained simulation estimates.

We implemented the discrete diffusion-based peptide generator with a total of 20 diffusion steps per sequence, beginning with an entirely masked input and stochastically revealing amino acid positions according to a parameterized masking schedule. A small smoothing factor $\rho \in [0.01, 0.05]$ was added at each step to guard against overconfident assignments that could lead to suboptimal local minima. For the mAb design, an energy-based updater was employed; at each iteration, we computed a proposed update to a random position in the sequence that decreased the total energy,

$$E_{\text{total}} = \lambda_{\text{bind}} E_{\text{bind}} + \lambda_{\text{crowd}} E_{\text{crowd}} + \lambda_{\text{self}} E_{\text{selfint}},$$

and accepted it with a probability proportional to $\exp(-\Delta E_{\text{total}}/T)$, where T is a pseudo-temperature controlling exploration. In both methods, we utilized mini-batch sizes of 32 sequences per update cycle, and each batch underwent scoring for the different objectives to accumulate gradient-style signals or acceptance criteria.

Table 1 summarizes the principal hyperparameters selected for our pipeline. The diffusion-based model learning rate was set to 5×10^{-4} with an Adam optimizer, while the mAb energy-based search rate was tuned to ensure near-equilibrium acceptance rates of roughly 45%. All reported metrics, such as the average predicted affinity and toxicity, were computed using internal validation sets. Final candidate sequences were then passed on to in silico docking or membrane simulation tests to confirm their plausibility before experimental validation.

Hyperparameter	**Value**
Peptide Diffusion Steps	20
Peptide Learning Rate	5×10^{-4}
Mini-Batch Size	32
Energy-Based Pseudo-temperature (mAb)	0.1–0.2
Acceptance Equilibrium Rate	45%

Table 1: Overview of key hyperparameter settings in our experimental pipeline.

In this section, we present a comprehensive analysis of the outcomes from our multi-objective generative pipeline, focusing first on the discrete diffusion-based GLP-1 peptide designs (Experiment #1), and then on the energy-guided mAb generation protocol (Experiment #2). We supplement these findings with an extended discussion that connects the generated results to broader methodological, practical, and clinical contexts.

7 Results

7.1 Experiment #1: Multi-Objective GLP-1 Peptide Designs

We initially generated a set of 50 synthetic GLP-1 analog peptides using the discrete diffusion-based sampler described in Section **??**, each annotated with a simulated binding score (ranging from 0 to 1, with higher values implying stronger receptor binding) and a toxicity score (also from 0 to 1, where lower values indicate reduced toxicity). Our pre-defined acceptance criteria, drawn from the aggregated multi-term objective function, were that a candidate peptide must achieve a binding score exceeding 0.5 while having a toxicity score under 0.5. Out of 50 total candidates, 7 peptides (14%) met these criteria. Although this may seem a limited filtering success rate, it is consistent with the need to balance multiple, often conflicting optimization targets. Among these filtered peptides, the standout candidate had a binding score of approximately 0.84 and a toxicity score of about 0.43—placing it near the Pareto front of competing objectives.

Figure 1 provides a visual mapping of the distribution of generated peptides by plotting binding affinity on the x-axis against toxicity on the y-axis. While a majority of the candidates cluster around a moderate range for both measures, the top 7 stand out in a region of high binding but relatively modest toxicity. This approach demonstrates a strong capacity of the diffusion-based model to generate peptides that inhabit desirable zones within the design space. A key hypothesis supported here is that, by relying on iterative unmasking guided by in silico property scoring, we can enrich for sequences that advance multiple targets simultaneously, showcasing the mechanical feasibility of this pipeline.

To further investigate the suitability of these peptides, we performed preliminary in silico docking simulations. For instance, the top candidate peptide with a binding score of about 0.84 still demonstrated robust interactions with critical receptor residues: for example, its predicted hydrogen-bond network indicated an interaction with known binding hotspots on GLP-1R. Though these docking studies remain limited due to their reliance on homology-based or partial receptor models, they offer an encouraging glimpse into how discrete diffusion can propose novel sequences of prospective relevance.

Cross-referencing the peptide design results with even minimal in vitro or ex vivo data would be the logical next step—but even without that, the successful generation of plausible, multi-objective candidates in a small-scale experiment lays out a practical template for more expansive searches. By scaling up from 50 to hundreds or thousands of peptides, each refined by advanced scoring heuristics or augmented with real-world side-effect data, the pipeline has the potential to identify exceptionally rare but therapeutically valuable sequences. Indeed, real-world performance might show improved yields as the model sees more training data and refines its internal representations of binding–toxicity trade-offs.

7.2 Experiment #2: mAb Design Under Membrane Crowding

In parallel, we applied an energy-based sampling method to produce 30 candidate monoclonal antibodies that target the GLP-1 receptor, each candidate featuring an affinity measure, an LLM-based binding prediction, and a randomly simulated crowding impact factor. These attributes were aggregated into a composite score, which balanced the tension between purely high-affinity sequences and their hypothetical performance in realistic membrane environments. Among these 30 sequences, a subset of 5 rose to the top with overall scores exceeding 1.0, signifying robust synergy between raw binding potential and tolerance to crowding penalties (or occasionally even mild crowding benefits, as certain crowding conditions can stabilize specific configurations).

Figure 2 illustrates a scatterplot with affinity on one axis and crowding impact on the other, highlighting how the top 5 candidates occupy a region where neither factor is severely compromised. One leading sequence, for example, attained an affinity near 0.95 while maintaining a moderate penalty from crowding. These results highlight that an mAb candidate can excel in binding if it also remains sufficiently flexible or structurally optimized to function within heterogeneous membrane environments. This is particularly germane for therapies targeted at complex tissues or organs where local protein densities or glycoprotein arrays might hinder binding.

We also observed that some candidates with high raw affinity performed less favorably when the simulated crowding factor was included. These "fragile" sequences, as we might call them, probably rely heavily on rigid binding interfaces that become sterically disrupted by high membrane protein traffic. Hence, a single best-in-class mAb in a naive sense—one that is simply optimized for free-space binding—might struggle to achieve adequate efficacy in physiologically accurate scenarios. This underscores the necessity of multi-objective design in antibody engineering, along with the advantage of incorporating advanced crowding terms that faithfully capture real cellular constraints.

Finally, the extended method's modularity enabled us to incorporate a simplified LLM-based binding predictor, which judges proposed sequences on features extracted from prior large-scale data. The synergy between the energy-based sampling method and LLM embeddings produced qualitatively different proposals than a purely physics-based approach, emphasizing the integrative power of combining large-scale learned representations with physically grounded simulations.

7.3 Ancillary Validation via Logistic Regression

In a relatively small-scale proof-of-concept, we applied a logistic regression classifier to a filtered social media text dataset, achieving nonzero accuracy when predicting tweet labels. While tangential to the main focus on GLP-1, this exercise reinforces that our pipeline (encompassing various data processing and ML subroutines) functions suitably well across tasks. The success of the classifier,

although modest in scope, demonstrates the basic soundness of the data flow and model training steps, supporting our confidence that these components can be readily expanded for more detailed toxicity or side-effect modeling efforts down the line.

8 Discussion

8.1 Interpretation of Key Findings

Our multi-objective generative experiments provide evidence that discrete diffusion-based models can identify GLP-1 peptide analogs that display both high predicted binding and moderate-to-low toxicity, even in a modest sampling of 50 total candidates. The 14% success rate in generating peptides that meet both criteria points to a flexible capability to "in-paint" sequences under potentially contradictory constraints. Analogously, the energy-based approach to designing anti-GLP-1R mAbs successfully uncovered candidates that appear promising under crowding conditions. By weighting the crowding factor along with a direct measure of affinity and an LLM-based predictor, we captured the interplay between local environment and binding. This is particularly relevant for therapeutic antibodies likely to face complex in vivo conditions, establishing a potential pipeline for next-generation, environment-specific biologics.

One of the more illuminating dimensions of these findings is the demonstration that crowding cannot be overlooked without risking inappropriate funneling toward suboptimal designs. A design that thrives in a purely in vitro scenario might lose viability in membrane-dense tissues, whereas a rationally balanced approach can yield more consistent performance across physiological conditions. Accounting for such complexities in the generative process thus represents a step beyond simplistic "high-affinity only" design approaches frequently seen in smaller-scale academic or industrial pipelines.

8.2 Practical and Clinical Relevance

Turning to clinical implications, therapies targeting the GLP-1 pathway, particularly for metabolic diseases such as type 2 diabetes and obesity, require a delicate balance of efficacy and tolerability. The generation of peptides with lower toxicity is essential to minimize side effects like gastrointestinal discomfort, while robust binding leads to potent insulin secretion modulation. On the mAb side, the ability to craft environment-aware antibodies could enable new strategies for selectively antagonizing or activating GLP-1R in tissues where cell density or glycoprotein expression is high. Such targeted engineering of membrane-optimized antibodies may reduce dosage requirements and off-target binding, lowering treatment costs and adverse effects.

Although these findings remain preliminary, a future pipeline might integrate real-world patient feedback (via social media or medical forums) into the multi-objective function, dynamically adjusting the toxicity score to reflect emergent

concerns. Over time, this iterative loop, bridging computational design and real-world signals, could yield therapies with both high efficacy and proven patient acceptability. The potential extension to other gut hormones and metabolic regulators further broadens the clinical horizon, tapping into new classes of treatments for metabolic syndrome, obesity-linked hypertension, and cardiovascular risk management.

8.3 Limitations and Next Steps

Despite the promising indicators, our approach faces multiple constraints. First, the property scores for binding and toxicity were, in part, simulated or estimated rather than determined from actual high-throughput assays. While we did integrate representative docking routines, homology models, and basic toxicity filters, these approximations must be refined before reaching definitive conclusions about clinical utility. Second, our sample sizes—50 peptides and 30 mAbs—test the ideas at a small scale. Industrial or academic labs often generate and screen thousands of molecules in parallel, an approach that significantly multiplies the chance of uncovering truly exceptional leads. Scaling up, combined with more accurate property scoring, would accelerate the identification of top-tier analogs.

Another aspect for improvement involves capturing more detailed immunogenicity profiles. In practice, both peptide-based drugs and mAbs can stimulate an immune response if certain epitopes are recognized. Subtle changes in amino acid composition, post-translational modifications, or glycosylation patterns can modulate this immunogenic risk. In integrating advanced immunogenicity classifiers (trained on real epitope data) within a discrete diffusion or energy-based loop, future expansions of this framework may generate leads that satisfy binding, toxicity, stability, crowding, and low immunogenicity, all at once.

The pipeline could also benefit from advanced sampling strategies to escape local minima in sequence space. While the pseudo-temperature acceptance criterion in the energy-based method provides some degree of exploration, it may still converge on suboptimal solutions without a targeted means of escaping shallow basins. One possibility is to combine Monte Carlo Tree Search (MCTS) expansions with the existing energy function and classifiers, systematically branching out from partial sequences that exhibit promise in earlier unmasking steps. Similarly, the diffusion-based approach could incorporate adaptive schedules that dynamically adjust how aggressively tokens are masked or unmasked, particularly for the subset of tokens representing crucial residues for receptor binding.

On the empirical side, in vitro validation must be scaled alongside computational exploration to filter out spurious sequences swiftly and focus resources on actual top candidates. Automated instrumentation for high-throughput expression and binding assays can handle hundreds of sequences in a week, providing data to iteratively refine the model's notion of viability. Over multiple cycles, such a "closed-loop" paradigm—where model suggestions are tested experimentally, and those results inform the model in a continuous cycle—can produce robust, clinically meaningful leads. This synergy is at the heart of next-generation drug discovery frameworks that unify computational power with measured data

on a rapid iteration schedule.

8.4 Future Directions

In addition to immediate enhancements, the broader vision for our method extends to deeper forms of environment-specific design. The approach applied to membrane crowding could be generalized to incorporate organ-specific or cell-type-specific factors, such as pH gradients in the gastrointestinal tract or local expression profiles of transporters in particular tissues. The creation of targeted peptide carriers for drug conjugates, for instance, might rely on specialized versions of this pipeline that weigh half-life, cargo release kinetics, and immunogenic potential simultaneously.

Another promising angle is to incorporate time-series data—observing how structural or functional properties develop over repeated exposures or metabolic cycles. Whether these timescales reflect hours (for acute drug action) or weeks (for sustained therapy), the interplay between a peptide's or antibody's structural dynamics and efficacy cannot be reduced to single-point measures. Embedding knowledge of these temporal factors within the diffusion-based model might produce molecules optimized not just for static snapshots, but for real, dynamic biological processes.

Overall, our results underscore the viability of a multi-objective approach for generating novel GLP-1 peptide analogs and mAbs that meet multiple performance criteria simultaneously. While much work remains to transition from conceptual demonstration to validated therapeutics, the synergy between discrete diffusion, energy-based design, and integrated property predictions signals a powerful route forward. By systematically expanding the property filters, refining the generative backbones, and incorporating advanced in vitro feedback, we can continue to push the frontier of AI-driven drug discovery, particularly in the domain of metabolic disease treatments. This vision represents a convergence of computational biology, machine learning, and translational science that is poised to reshape how we conceive of, and ultimately realize, the next generation of targeted protein therapeutics.