

Advancing Protein Design via Multi-Agent Reinforcement Learning with Pareto-Based Collaborative Optimization

Mingming Zhu^{1*}, Jiahua Rao^{1*†}, Xiaoyu Chen¹, Qianmu Yuan^{1,2}, Yuedong Yang^{1,3†}

¹School of Computer Science and Engineering, Sun Yat-sen University, China

²High Performance Computing Department, National Supercomputing Center in Shenzhen, China

³Guangdong Provincial Key Laboratory of Computational Science, Sun Yat-sen University, Guangzhou, China
{zhumm25, chenxy759}@mail2.sysu.edu.cn, yuanqm3@mail3.sysu.edu.cn, {raojh7, yangyd25}@mail.sysu.edu.cn

Abstract

Protein design is revolutionizing biotechnology, yet existing approaches struggle to balance structural foldability with functional performance. Structure-based models excel at generating stable protein backbones but often overlook critical functional properties, while protein language models capture evolutionary and functional signals but frequently predict sequences lacking structural stability. Integrating these complementary approaches remains challenging due to their inherently conflicting objectives. We present MAProt, a multi-agent framework that synergistically combines structure-based and protein language model-based methods for protein design. Each agent specializes in a distinct aspect of the design objective: the structure-based agent (e.g., ProteinMPNN) ensures compatibility with the target backbone, while protein language model-based agents (e.g., ESM, SaProt) capture evolutionary plausibility and functional potential. To reconcile conflicts and achieve optimal trade-offs, we introduce a Pareto-based negotiation module that enables effective multi-objective coordination and consensus among agents. Extensive experiments on benchmark datasets demonstrate that MAProt achieves a remarkable improvement over state-of-the-art baselines, and generalizes robustly across a range of tasks, including thermodynamic folding stability design, functional protein design, and high-affinity antibody design. These results highlight the power of collaborative optimization for advancing rational protein engineering.

Code — <https://github.com/biomed-AI/MAProt>

Introduction

Protein design is revolutionizing biotechnology by enabling the creation of proteins with customized structural and functional properties (Yeh et al. 2023; Wang et al. 2022). One of the most important challenges is the optimization of protein functional characteristics, which seeks to improve desirable properties, such as stability, binding affinity (Vázquez Torres et al. 2024), or catalytic efficiency (Jiang et al. 2024), through targeted modifications of amino acid sequences. However, the landscape of protein functionality constitutes a vast combinatorial space: as sequence length increases,

the number of possible variants grows exponentially, while only a tiny fraction possesses meaningful biological functions (Anishchenko et al. 2021). As design objectives become more sophisticated, achieving both efficient and reliable protein design remains a fundamental challenge.

Despite remarkable advances in protein design methodologies, effectively balancing structural foldability with functional performance continues to be a central obstacle in the field (Notin et al. 2024). Structural stability and functional properties often impose conflicting requirements in the high-dimensional sequence space: mutations that enhance functional performance may disrupt native foldability and stability. Structure-based approaches, such as ProteinMPNN (Dauparas et al. 2022), are highly effective at ensuring structural integrity but often overlook critical functional attributes, including enzymatic activity, binding affinity, or thermostability (Xu et al. 2025). Consequently, these methods may fail to identify variants with the desired functional properties. To address this limitation, recent methods such as ProteinDPO (Widatalla, Rafailov, and Hie 2024) leverage high-quality thermostability datasets to guide structure-based models through reinforcement learning, enabling the generation of sequences that not only preserve the input backbone structure but also exhibit improved functional performance.

However, the scarcity of large-scale, high-quality experimental fitness datasets (Notin et al. 2023) constrains the ability of current models to systematically learn and balance both structural and functional design objectives. In contrast, protein language models such as ESM (Lin et al. 2023) and SaProt (Su et al. 2024) excel at predicting functional effects by capturing global sequence patterns and evolutionary constraints from vast protein databases. This capability enables them to infer functional relevance even in the absence of extensive labeled fitness data (Zhu et al. 2024). Nevertheless, these models frequently generate sequences that lack proper structural stability or foldability. This contrast highlights the complementary strengths of structure-based and language model-based approaches (Yuan et al. 2022), underscoring the need for a unified framework that integrates both structural and sequence information to enable balanced optimization of protein structure and function.

To address this, recent studies (Fei et al. 2025; Xiong et al. 2025) have explored directly integrating structural models

*These authors contributed equally.

†Corresponding Author.

Copyright © 2026, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

with sequence-based language models, aiming to leverage their complementary strengths during protein design. However, negotiating the conflicting objectives of different models within a unified framework remains an exceptionally challenging task (Rao et al. 2022). Simultaneously optimizing multiple objectives, such as foldability, evolutionary plausibility, and functional fitness, inevitably introduces complex trade-offs, as the requirements imposed by different models may be mutually incompatible for certain mutations. Achieving optimal synergy demands careful reconciliation of these competing objectives (Rao et al. 2024). Thus, the central challenge lies not only in integrating heterogeneous predictive signals, but also in effectively resolving inter-model conflicts and establishing consensus among them. Accordingly, there is a pressing need for more robust and principled strategies to negotiate and optimize conflicting objectives within unified protein design frameworks.

In this work, we propose MAProt, a multi-agent-based protein design framework that synergistically combines structure-based and protein language model-based methods. Within this framework, each agent is responsible for a distinct aspect of the design objective: the structure-based agent (ProteinMPNN) ensures compatibility with the target backbone, while sequence-based agents (ESM, SaProt) capture global sequence features and mutational effects, thereby collectively addressing both structural and functional requirements. By formulating protein design as a multi-objective optimization problem, we introduce a Pareto-based negotiation and conflict resolution module that coordinates the potentially competing objectives of different agents and seeks consensus solutions. This enables MAProt to generate protein sequences that optimally balance structural stability and functional potential, providing an effective strategy for the synergistic optimization of protein design.

Extensive experiments on benchmark datasets demonstrate that our method consistently outperforms state-of-the-art baselines. MAProt achieves up to 14.1% improvement of success rate in stability protein design on the Megascala benchmark, state-of-the-art performance on the GFP protein fitness optimization benchmark, and a 32% success rate in high-affinity antibody design. Moreover, MAProt is highly extensible to diverse model architectures and parameter scales, demonstrating generalizability across protein domains. Ablation studies and interpretability analyses further underscore the critical role of the negotiation and consensus module in enabling effective multi-objective optimization, establishing MAProt as a powerful and generalizable tool for the rational design and engineering of functional proteins. Our contributions are summarized as follows:

- We propose MAProt, a multi-agent protein design framework that synergistically integrates structure-based models and protein language models.
- We introduce a Pareto-based negotiation mechanism for effective coordination and conflict resolution among agents with competing objectives.
- Extensive experiments on benchmark datasets show that MAProt consistently outperforms state-of-the-art baselines across diverse protein design tasks.

Related Work

Protein Design and Optimization

Computational protein design aims to generate amino acid sequences that fold into desired structures and functions. Existing methods are generally divided into two categories. Structure-based models, such as ProteinMPNN and ESM-IF (Hsu et al. 2022), utilize 3D structural information to guide sequence generation, achieving high accuracy in replicating natural protein folds. In contrast, sequence-based methods, including variational autoencoders (Hawkins-Hooker et al. 2021), generative adversarial networks (Repecka et al. 2021), and large language models like ProGen (Madani et al. 2020), learn from evolutionary sequence data to sample diverse and plausible sequences.

Despite these advances, most approaches focus on optimizing general structural or evolutionary fitness, which may not meet specific, application-driven functional requirements. Preference alignment strategies have emerged to guide design toward desired properties; for example, ProteinDPO uses direct preference optimization, while triplet-based distillation (Karimi et al. 2025) and DRAKES (Wang et al. 2024) employ reward-driven fine-tuning. However, these methods depend on high-quality datasets for particular properties and still lack effective mechanisms to integrate and balance these conflicting objectives in protein design.

Multi-objective Collaborative Optimization

Multi-objective optimization in protein design seeks to simultaneously satisfy diverse criteria, such as foldability, evolutionary plausibility, and functional fitness. Recent methods attempt to leverage both complementary strengths of structure- and sequence-based models (Misaki et al. 2025). For example, del Alamo et al. (2025) improves ProteinMPNN for antibody CDR design by integrating the antibody-specific language model AbLang (Olsen, Moal, and Deane 2022), enhancing native-likeness and binding. AICE (Fei et al. 2025) combines structural and evolutionary constraints with inverse folding to predict high-fitness mutations, while ProteinGuide (Xiong et al. 2025) enables guided generation by conditioning ProteinMPNN and ESM3 (Hayes et al. 2025) on auxiliary signals.

Existing frameworks typically lack principled strategies for conflict resolution, often relying on static or heuristic integration. However, directly optimizing multiple objectives simultaneously inevitably introduces complex trade-offs, as competing requirements may be mutually incompatible. To address this challenge, we introduce a Pareto-based multi-agent framework that enables systematic negotiation and consensus-building among heterogeneous models, thereby achieving balanced and synergistic optimization across multiple protein design objectives.

Methodology

Problem Formulation

We formulate protein sequence design as an optimization problem, which seeks to find a sequence $S \in V^L$, where V is the set of standard amino acids and L is the sequence length, that maximizes a specific fitness metric $F =$

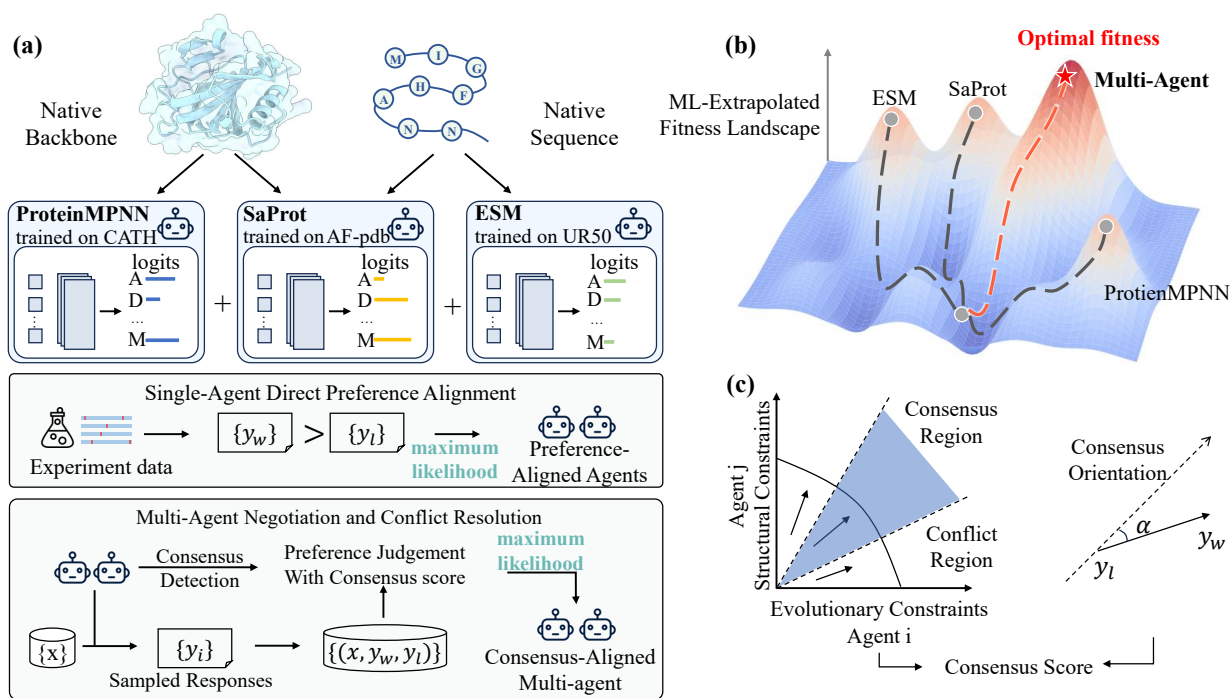


Figure 1: An overview of the MAProt framework. (a) MAProt integrates multiple sequence design models—ProteinMPNN, SaProt, and ESM. Experimental data is first used to align the preferences of individual agents. Subsequently, multi-agent negotiation, incorporating consensus detection and conflict resolution, yields a consensus-aligned solution. (b) Agents collaboratively explore an extrapolated fitness landscape to identify optimal sequences. (c) The consensus and conflict regions among agents are visualized in terms of structural or evolutionary constraints, with a consensus score quantifying agreement.

$f(S, B)$. Here, B denotes a protein backbone structure, and the fitness function f quantifies desired protein properties.

Our framework leverages a set of specialized pretrained models as agents: ProteinMPNN for structure-based sequence generation, ESM for functional evaluation, and SaProt for assessing sequence and structural constraints. Each agent focuses on a distinct aspect of protein design, either generating sequences or evaluating them based on structural, functional, or evolutionary criteria. Through Pareto-based optimization, these agents negotiate to balance multiple objectives and reach consensus on high-quality sequences. This multi-agent strategy enables MAProt to efficiently explore the protein fitness landscape and generate sequences with superior biochemical properties.

Architectures of ProteinMPNN, ESM and SaProt

To leverage the diverse information in protein structures and sequences, MAProt employs three specialized agents, each tailored to a distinct modality of protein data. Specifically, ProteinMPNN leverages structural information by utilizing the 3D backbone coordinates, which are represented as a tensor in $\mathbb{R}^{L \times 3}$, where L denotes the sequence length. This allows ProteinMPNN to generate amino acid sequences $S = (s_1, \dots, s_L)$ that are compatible with the target fold.

ESM is a transformer-based protein language model that processes amino acid sequences $S = (s_1, \dots, s_L)$, where s_i denotes the residue at position i (Yuan et al. 2024). By

capturing long-range dependencies, ESM evaluates the evolutionary plausibility and functional potential of candidate sequences based on patterns learned from large-scale protein sequence datasets. In contrast, SaProt integrates both sequence and structural information by encoding the 3D structure as a structural sequence $F = (f_1, \dots, f_L)$ using Foldseek (Van Kempen et al. 2024), where each f_i denotes the structural token at residue i . These structure tokens are then concatenated with the amino acid sequence $Z = ((s_1, f_1), \dots, (s_L, f_L))$, enabling joint modeling of sequence-structure relationships. Further details are provided in Appendix.

Single-Agent Preference Alignment

We next analyze how each agent, operating independently on its respective modality, develops distinct sequence design preferences, establishing a foundation for subsequent multi-agent integration and consensus.

Building Preference Data To align sequence generation preferences with experimentally measured protein fitness, we construct a set of sequence preference pairs, following prior work ProteinDPO: $\mathcal{D}_{\text{pref}} = (x, y^w, y^l)$, where for each input x , y^w and y^l are candidate sequences with higher and lower experimentally measured fitness, respectively. These pairs are independently generated for each agent using available experimental data, ensuring that preference comparisons are empirically grounded.

Direct Preference Optimization Each agent is initialized from a pretrained model, trained on large-scale sequence or structure data, to model the conditional distribution $y \sim \pi_\theta(\cdot | x)$, where θ denotes model parameters. To align sequence generation with experimental fitness, we fine-tune each agent’s policy to maximize the likelihood of high-fitness sequences. Specifically, we apply Direct Preference Optimization (DPO) (Rafailov et al. 2023) to align the policy π_θ with the preference dataset $\mathcal{D}_{\text{pref}}$, relative to a fixed reference policy π_{ref} .

DPO defines the probability that a candidate sequence y_w is preferred over another candidate y_l in prompt x as:

$$r_\theta(x, y) = \beta \log \frac{\pi_\theta(y | x)}{\pi_{\text{ref}}(y | x)}, \quad (1)$$

$$p_\theta(y_w \succ y_l | x) = \sigma(r_\theta(x, y_w) - r_\theta(x, y_l)),$$

where $\sigma(\cdot)$ denotes the sigmoid function and β is a temperature hyperparameter. The DPO loss for each agent is:

$$\mathcal{L}_{\text{DPO}} = \mathbb{E}_{(x, y_w, y_l) \sim \mathcal{D}_{\text{pref}}} [-\log p_\theta(y_w \succ y_l | x)]. \quad (2)$$

To fine-tune generation preferences toward proteins with desired properties, we jointly minimize the sum of the DPO losses for ProteinMPNN, ESM, and SaProt:

$$\mathcal{L} = \mathcal{L}_{\text{DPO,MPNN}} + \mathcal{L}_{\text{DPO,ESM}} + \mathcal{L}_{\text{DPO,SaProt}}. \quad (3)$$

The preference alignment mechanism fine-tunes each model with DPO to align its generative preferences with its specialized property, such as structural compatibility, evolutionary relevance, or functional potential.

Pareto-based Multi-Agent Preference Agreement

While each agent can independently optimize its sequence generation policy, the protein design often requires integrating multiple, potentially conflicting objectives. To this end, we propose a multi-agent consensus module, inspired by the concept of Pareto optimality in multi-objective optimization, to systematically quantify and leverage inter-agent agreement for robust preference optimization. In this framework, consensus regions correspond to Pareto-optimal solutions that balance multiple desirable properties, thereby reflecting agreement across diverse agent preferences.

For each native protein x in the training set, we generate N sequences $\{y_i\}_{i=1}^N$ using the policy model learned from the previous stage. Each agent independently evaluates every sequence, obtaining three probability scores:

$$P_{\text{Agent}}(y_i) = \pi_{\text{Agent}}(y_i | x). \quad (4)$$

Due to the distinct modeling strategies and domain expertise of each agent, these probabilities can exhibit significant variability. Therefore, a systematic approach is needed to identify candidates with high inter-agent consensus.

Consensus-Guided Preference Pair Construction To quantify consensus and conflict between two agents, we compute the relative preference margin: $M_{A,B}(y_i) = P_A / \sqrt{P_A^2 + P_B^2}$ where P_A and P_B are the probability scores assigned by agents A and B to candidate y_i . We

assign a consensus label $g_{(A,B)}(y_i)$ to each candidate sequence as follows:

$$g_{(A,B)}(y_i) = \begin{cases} 1, & M_{A,B}(y_i) > \tau_1 \\ 0, & \tau_2 < M_{A,B}(y_i) \leq \tau_1 \\ -1, & M_{A,B}(y_i) \leq \tau_2 \end{cases}, \quad (5)$$

where $g_{(A,B)} = 0$ indicates a consensus region where both agents provide similar assessments, while $g_{(A,B)} = 1$ or -1 indicates a conflict region, reflecting a significant preference for one candidate by either agent.

Candidates y_i with the same group label are grouped, reflecting consistent evaluation regimes across agent pairs. Preference pairs $D_{\text{cons}} = (x, y^w, y^l)$ are constructed within each group, ensuring y^w and y^l keep preference consistency.

Consensus Detection To robustly evaluate the reliability of each constructed preference pair (y^w, y^l) , we combine structural and continuous consensus measures. Specifically, let $p = y^w - y^l$ denote the vector of probability score differences across agents, with the ideal consensus direction given by $u = (1, 1, 1)$. The continuous consensus is measured by the angular deviation α between p and u :

$$\alpha = \arccos \left(\frac{p \cdot u}{\|p\| \cdot \|u\|} \right). \quad (6)$$

For the structural consensus, we calculate the proportion of consensus labels $g = 0$ within each group, denoted as a , which reflects the level of agreement among agents. These two components are integrated to form the consensus score:

$$c = \left(1 - \frac{2\alpha}{\pi} \right) \cdot (a + \gamma), \quad (7)$$

where γ controls the baseline confidence level, π is circle constant pi.

Consensus-Weighted Loss Formulation Finally, we apply this consensus score as a weighting factor, encouraging the model to prioritize learning from high-confidence pairs and to be less influenced by uncertain or disputed ones. The confidence-weighted loss is:

$$\mathcal{L}_{\text{Weighted}} = \mathbb{E}_{(x, y^w, y^l) \sim \mathcal{D}_{\text{cons}}} [-c \cdot \log p_\theta(y^w \succ y^l | x)]. \quad (8)$$

Multi-Agent Sampling

To leverage the strengths of multiple agents, we employ a logit fusion strategy during sequence generation. At each residue position i , we aggregate information from all M agents by summing their logits to obtain the fused logits:

$$z = \sum_{k=1}^M z^{(k)}. \quad (9)$$

The probability distribution over amino acid types is:

$$P_T(y_i | y_{<i}, x) = \frac{\exp(z_i/T)}{\sum_{j=1}^{20} \exp(z_j/T)}, \quad (10)$$

where T is a temperature hyperparameter. At each position, the next residue is sampled from the distribution with multinomial sampling, enabling stochastic sequence generation.

Method	Pred-ddG(Med) \uparrow	%(ddG > 0)(%) \uparrow	scRMSD(Med) \downarrow	%(scRMSD < 2)(%) \uparrow	Success Rate(%) \uparrow
ESM-IF	-1.898(0.011)	2.9(0.1)	1.105(0.021)	73.5(0.5)	2.0(0.1)
ProteinMPNN	-0.510(0.001)	35.1(0.0)	0.862(0.027)	93.0(0.4)	31.5(0.1)
CG	-0.561(0.045)	36.9(1.1)	0.839(0.012)	90.9(0.6)	34.7(0.9)
SMC	0.659(0.044)	68.5(3.1)	0.841(0.006)	93.8(0.4)	63.6(4.0)
TDS	0.674(0.086)	68.2(2.4)	0.834(0.001)	94.4(1.2)	62.9(2.8)
CFG	-1.186(0.035)	11.0(0.4)	3.146(0.062)	29.4(1.0)	1.3(0.4)
DRAKES	1.095(0.026)	86.4(0.2)	0.918(0.006)	91.8(0.5)	78.6(0.7)
ProteinDPO	1.120(0.078)	89.8(4.9)	1.036(0.034)	82.9(0.6)	73.7(4.7)
MAProt (DPO)	2.143(0.030)	93.8(0.5)	1.012(0.038)	89.0(2.3)	82.9(2.7)
MAProt	1.603(0.073)	94.7(1.2)	0.890(0.032)	94.8(1.7)	89.7(0.2)

Table 1: Comparison of the performance of baselines and MAProt on the Megascale benchmark for protein stability optimization. Results are reported as the mean across 3 random seeds, with standard deviations in parentheses.

Experiments

Experimental Setup

Datasets We evaluate MAProt on three benchmarks: Megascale (Tsuboyama et al. 2023) (~1.8M variants) for stability optimization, GFP (Sarkisyan et al. 2016) (56,806 variants) for functional phenotype design, and AffinityDesign (47 complexes) for high-affinity antibody-antigen engineering. Each dataset targets a distinct protein property. Further details and data splits are provided in the Appendix.

Baselines We compare our method against a comprehensive set of baselines covering diverse protein sequence optimization strategies. On the Megascale and AffinityDesign datasets, we include guidance-based methods such as classifier guidance (CG) (Nisonoff et al. 2025), SMC-based guidance and TDS (Wu et al. 2023); classifier-free guidance (CFG) (Ho and Salimans 2022); pretrained models without task-specific fine-tuning, ProteinMPNN and ESM-IF; as well as reinforcement learning fine-tuned methods, DRAKES, and ProteinDPO. On the GFP dataset, we benchmark against a range of optimization and sampling baselines, including GFlowNets (GFN-AL) (Jain et al. 2022), model-based adaptive sampling (CbAS) (Brookes, Park, and Listgarten 2019), greedy search (AdaLead) (Sinai et al. 2020), Bayesian optimization (BO-qEI) (Wilson et al. 2017), conservative model-based optimization (CoMS) (Trabucco et al. 2021), proximal exploration (PEX) (Ren et al. 2022), and predictor-guided approaches such as VLGP0 (Bogensperger et al. 2025), GWG, and GGS (Kirjner et al. 2024).

Experimental Settings. MAProt is implemented using Pytorch v1.10.2 and trained on a single NVIDIA A800 80GB GPU. All trainable parameters are optimized by Adam (Kingma and Ba 2014) algorithm with a learning rate of 5×10^{-6} . We determined the DPO loss hyperparameter β and temperature T with grid search. For all baselines, they are retrained on the same machine with the same hyperparameter settings reported. Details on hyperparameters and evaluation metrics are in the supplementary materials.

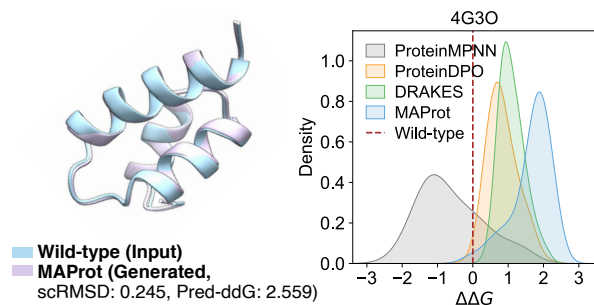


Figure 2: Example of generated proteins. The histograms display the $\Delta\Delta G$ for each generated sequence.

Main Results

Thermodynamic Folding Stability Design. We evaluate our method on the challenging task of protein thermodynamic folding stability design using the Megascale benchmark. Following previous work, we evaluate all methods using the $\Delta\Delta G$ predictor provided by DRAKES. Table 1 summarizes the performance of MAProt and baseline methods. Our method achieves state-of-the-art performance improving success rate by +14.1% and median predicted $\Delta\Delta G$ by +43.1% over DRAKES, demonstrating superior generation of stable protein sequences.

MAProt robustly optimizes folding stability while maintaining structural integrity. DRAKES preserves structure but has limited exploration, yielding smaller $\Delta\Delta G$ gains. Multi-agent frameworks without consensus (e.g., MAProt (DPO)) can improve $\Delta\Delta G$, but often at the cost of structural quality due to conflicting objectives. In contrast, MAProt’s consensus-driven negotiation effectively balances evolutionary and structural constraints, delivering superior stability without sacrificing integrity, demonstrating the advantage of our multi-agent consensus optimization approach.

To further demonstrate the effectiveness of MAProt, we analyze the distribution of predicted $\Delta\Delta G$ values for generated sequences (Figure 2). MAProt’s distribution is shifted toward higher stability, with a substantially larger proportion of successful designs than other methods. For example, the

Method	Medium difficulty			Hard difficulty		
	Fitness \uparrow	Diversity	Novelty	Fitness \uparrow	Diversity	Novelty
GFN-AL	0.09 \pm 0.1	25.1 \pm 0.5	213 \pm 2.2	0.1 \pm 0.2	23.6 \pm 1.0	214 \pm 4.2
CbAS	0.14 \pm 0.1	9.7 \pm 1.1	7.2 \pm 0.4	0.18 \pm 0.0	9.6 \pm 1.3	7.8 \pm 0.4
AdaLead	0.56 \pm 0.0	3.5 \pm 0.2	2.0 \pm 0.0	0.18 \pm 0.0	5.6 \pm 0.5	2.8 \pm 0.4
BOqei	0.20 \pm 0.0	19.3 \pm 0.0	0.0 \pm 0.0	0.0 \pm 0.5	94.6 \pm 71	54.1 \pm 81
CoMS	0.00 \pm 0.1	133 \pm 25	192 \pm 12	0.0 \pm 1.5	144 \pm 7.5	201 \pm 3.0
PEX	0.47 \pm 0.0	3.0 \pm 0.1	3.0 \pm 0.0	0.34 \pm 0.3	1.3 \pm 0.3	1.3 \pm 0.3
gg-dWJS	0.55 \pm 0.1	52.3 \pm 3.4	16.3 \pm 5.7	0.61 \pm 0.1	68.0 \pm 5.6	44.8 \pm 47
GWG	0.10 \pm 0.0	33.0 \pm 0.0	12.8 \pm 0.4	0.0 \pm 0.0	4.2 \pm 7.0	7.6 \pm 1.1
GGS	0.76 \pm 0.0	3.7 \pm 0.2	5.0 \pm 0.0	0.74 \pm 0.0	3.6 \pm 0.1	8.0 \pm 0.0
VLGPO, predictor g_ϕ	<u>0.87</u> \pm 0.0	4.31 \pm 0.1	6.0 \pm 0.0	0.75 \pm 0.0	3.1 \pm 0.2	6.0 \pm 0.0
VLGPO, smoothed $g_{\bar{\phi}}$	0.84 \pm 0.0	2.06 \pm 0.1	5.0 \pm 0.0	0.78 \pm 0.0	2.5 \pm 0.2	6.0 \pm 0.0
MAProt	0.88 \pm 0.0	3.8 \pm 0.4	7.0 \pm 0.0	0.87 \pm 0.0	3.4 \pm 0.5	8.0 \pm 0.0

Table 2: Comparison of the performance of baselines and MAProt on the GFP benchmark for fluorescence fitness optimization. Results are reported as the mean across 5 random seeds, with standard deviations in parentheses.

Method	Δ Affinity(Med) \downarrow	$\%(\Delta$ Affinity < 0) \uparrow (%)	Recovery(Med) \uparrow	$\%(\text{Recovery} > 0.7)$ (%) \uparrow	Success(%) \uparrow
ProteinMPNN	0.889(0.006)	3.3(0.1)	0.543(0.000)	0.0(0.0)	0.0(0.0)
ProteinMPNN(SFT)	0.342(0.102)	27.4(2.8)	0.478(0.005)	0.0(0.0)	0.0(0.0)
ProteinDPO	0.345(0.027)	34.3(0.9)	0.542(0.001)	0.0(0.0)	0.0(0.0)
MAProt (Zero-shot)	0.301(0.002)	23.7(0.1)	<u>0.708(0.002)</u>	<u>65.0(0.2)</u>	14.2(0.1)
MAProt (DPO)	<u>0.222(0.010)</u>	31.8(0.6)	0.706(0.000)	58.8(0.5)	<u>22.3(0.9)</u>
MAProt	0.153(0.020)	<u>32.4(2.6)</u>	0.763(0.001)	98.8(0.1)	32.1(2.5)

Table 3: Comparison of the performance of baselines and MAProt on the AffinityDesign benchmark for antibody affinity optimization. Results are reported as the mean across 3 random seeds, with standard deviations in parentheses.

designed sequence for 4G3O closely aligns with the original backbone (scRMSD = 0.245) while achieving significantly improved stability, also indicating that MAProt enhances stability without compromising structural integrity.

Functional Protein Design: Green Fluorescent Protein. We further evaluate MAProt on the challenging task of functional protein design using the GFP benchmark. Unlike stability optimization, this task requires models to generalize to novel, high-fluorescence sequences across a vast and diverse sequence space.

Table 2 presents the results on the GFP benchmark, comparing MAProt with baselines across both medium and hard difficulty tasks. MAProt consistently outperforms all baselines in the primary objective, achieving mean fitness scores of 0.88 (medium) and 0.87 (hard), establishing a state-of-the-art performance. The improvement is particularly pronounced on the hard task, where MAProt achieves up to 12% higher fitness than the second-best method (VLGPO with 0.78), demonstrating robust generalization and optimization.

In addition, MAProt maintains competitive diversity and novelty scores, indicating that generated sequences are not only highly functional but also sufficiently distinct. This superior performance can be attributed to MAProt’s ability to effectively coordinate multi-agent preferences, balancing functional optimization with sequence diversity. By contrast,

single-agent or less coordinated approaches tend to converge to local optima or produce redundant variants, as reflected in their lower novelty and diversity metrics.

High-Affinity Antibody Design. To further demonstrate the practicality and robustness of MAProt, we apply it to high-affinity antibody sequence design, a task requiring simultaneous optimization of multi-chain complexes for structural compatibility and binding affinity, closely reflecting real-world therapeutic antibody engineering challenges.

We benchmark MAProt on the AffinityDesign dataset against state-of-the-art baselines (Table 3). MAProt achieves a 32.1% success rate, a high sequence recovery rate (0.763), and strong binding affinity (mean Δ affinity of 0.153), significantly outperforming all competitors. Compared to the best baseline MAProt (DPO), MAProt achieves a 43.9% higher success rate and further reduces mean Δ affinity (lower is better), demonstrating superior performance on both key metrics. These results highlight MAProt’s strength in integrating multiple agent perspectives—including structure, sequence, and function—when navigating the complex antibody sequence space.

Ablation Study

To assess the contribution of each component in MAProt, we conduct ablation experiments on the Megascale dataset.

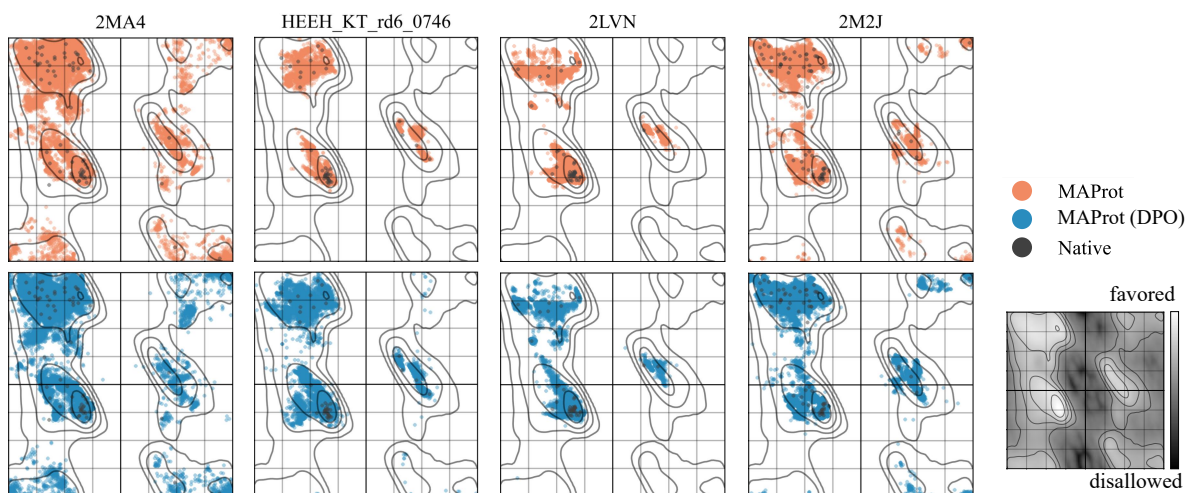


Figure 3: Ramachandran plots comparing dihedral angle ψ and ϕ distributions for generated versus native proteins; inset shows favored (light) and disallowed (dark) regions.

Method	Training	Success Rate(%) \uparrow
ProteinMPNN	Zero-shot	31.5(0.1)
MAProt	Zero-shot	54.4(0.9)
ProteinMPNN	SFT	66.7(3.5)
MAProt	SFT	69.8(1.3)
ProteinMPNN	DPO	79.1(3.5)
ProteinMPNN+ESM	DPO	82.2(3.7)
ProteinMPNN+SaProt	DPO	84.1(2.2)
MAProt	DPO	82.9(2.7)
MAProt	Consensus	89.7(0.2)

Table 4: Ablation of key components and training strategies on Megascale. +ESM/+SaProt denotes inclusion of the respective language model agents.

Multi-Agent Collaboration Optimization As shown in Table 4, the multi-agent framework consistently outperforms single-agent approaches. With DPO, it achieves a success rate of 82.9% versus 79.1% for the single-agent baseline. For SFT, the multi-agent setting reaches 69.8% compared to 66.7%. This shows that collaborative optimization leads to more robust and diverse sequence generation. With full integration, MAProt achieves the highest success rate of 89.7%, surpassing both the DPO-only multi-agent model (82.9%) and dual-agent settings (e.g., ProteinMPNN+SaProt). These results highlight that combining diverse agent knowledge and advanced strategies in MAProt substantially improves protein design performance over standard approaches.

Hyper-Parameter Analysis We further analyze the impact of the preference sharpness parameter β in multi-agent learning. As summarized in Table S3 (see Appendix for details), low β values (e.g., 0.1) enforce strong preference adherence but cause large structural deviations, while high

β (1.0) better preserves structure but reduces fitness and success rate. An intermediate setting ($\beta = 0.5$) achieves the best balance, yielding the highest success rate (89.7%) and robust performance. These results suggest that moderate preference strength is crucial for effective alignment without compromising structural quality.

Visualization Analysis

As shown in Figure 3, Ramachandran plots for MAProt and MAProt (DPO) generated structures are concentrated within favored regions, closely matching native proteins. This confirms that generated structures are geometrically plausible and free from steric clashes. MAProt (top row, orange) shows a more focused distribution in the most favored regions, indicating higher structural stability and fidelity to native conformations. It also covers a broader range of allowed regions than MAProt (DPO), demonstrating greater conformational diversity. Notably, MAProt avoids significant populations in disallowed regions, confirming the absence of geometric conflicts, while MAProt (DPO) shows some residues in these unfavorable areas.

Conclusion

In this work, we present MAProt, a flexible and robust multi-agent protein design framework that unifies diverse structural and functional objectives through negotiation and consensus. Extensive experiments show that MAProt consistently outperforms state-of-the-art methods across multiple benchmarks, achieving higher success rates and superior functional properties, as further supported by ablation and visualization analyses. In the future, we would like to extend MAProt to enable joint structure and sequence co-design, and apply it to more complex protein assemblies and de novo design tasks. Moreover, integrating experimental feedback and active learning will further enhance MAProt’s practical impact, supporting broader applications in protein engineering and therapeutic discovery.

Acknowledgments

This study has been supported by the Guangdong S&T Program [2024B111140001], the China Postdoctoral Science Foundation [2025M771540, GZB20250391], the Youth S&T Talent Support Program of GDSTA [SKXRC2025177], the Young S&T Talent Support Program of Guangdong Precision Medicine Application Association [YSTTGDPMAA2025], and the Lingang Laboratory [LGL-8888].

References

- Anishchenko, I.; Pellock, S. J.; Chidyausiku, T. M.; Ramelot, T. A.; Ovchinnikov, S.; Hao, J.; Bafna, K.; Norn, C.; Kang, A.; Bera, A. K.; et al. 2021. De novo protein design by deep network hallucination. *Nature*, 600(7889): 547–552.
- Bogensperger, L.; Narnhofer, D.; Allam, A.; Schindler, K.; and Krauthammer, M. 2025. A Variational Perspective on Generative Protein Fitness Optimization. *arXiv:2501.19200*.
- Brookes, D.; Park, H.; and Listgarten, J. 2019. Conditioning by adaptive sampling for robust design. In Chaudhuri, K.; and Salakhutdinov, R., eds., *Proceedings of the 36th International Conference on Machine Learning*, volume 97 of *Proceedings of Machine Learning Research*, 773–782. PMLR.
- Dauparas, J.; Anishchenko, I.; Bennett, N.; Bai, H.; Ragotte, R. J.; Milles, L. F.; Wicky, B. I.; Courbet, A.; de Haas, R. J.; Bethel, N.; et al. 2022. Robust deep learning-based protein sequence design using ProteinMPNN. *Science*, 378(6615): 49–56.
- del Alamo, D.; Frick, R.; Truan, D.; and Karpiak, J. 2025. Adapting ProteinMPNN for antibody design without retraining. *bioRxiv*, 2025–05.
- Fei, H.; Li, Y.; Liu, Y.; Wei, J.; Chen, A.; and Gao, C. 2025. Advancing protein evolution with inverse folding models integrating structural and evolutionary constraints. *Cell*.
- Hawkins-Hooker, A.; Depardieu, F.; Baur, S.; Couairon, G.; Chen, A.; and Bikard, D. 2021. Generating functional protein variants with variational autoencoders. *PLoS computational biology*, 17(2): e1008736.
- Hayes, T.; Rao, R.; Akin, H.; Sofroniew, N. J.; Oktay, D.; Lin, Z.; Verkuil, R.; Tran, V. Q.; Deaton, J.; Wiggert, M.; et al. 2025. Simulating 500 million years of evolution with a language model. *Science*, 387(6736): 850–858.
- Ho, J.; and Salimans, T. 2022. Classifier-Free Diffusion Guidance. Preprint, *arXiv:2207.12598*.
- Hsu, C.; Verkuil, R.; Liu, J.; Lin, Z.; Hie, B.; Sercu, T.; Lerer, A.; and Rives, A. 2022. Learning inverse folding from millions of predicted structures. In *International conference on machine learning*, 8946–8970. PMLR.
- Jain, M.; Bengio, E.; Hernandez-Garcia, A.; Rector-Brooks, J.; Dossou, B. F. P.; Ekbote, C. A.; Fu, J.; Zhang, T.; Kilgour, M.; Zhang, D.; Simine, L.; Das, P.; and Bengio, Y. 2022. Biological Sequence Design with GFlowNets. In Chaudhuri, K.; Jegelka, S.; Song, L.; Szepesvari, C.; Niu, G.; and Sabato, S., eds., *Proceedings of the 39th International Conference on Machine Learning*, volume 162 of *Proceedings of Machine Learning Research*, 9786–9801. PMLR.
- Jiang, K.; Yan, Z.; Di Bernardo, M.; Sgrizzi, S. R.; Villiger, L.; Kayabolen, A.; Kim, B.; Carscadden, J. K.; Hiraizumi, M.; Nishimasu, H.; et al. 2024. Rapid in silico directed evolution by a protein language model with EVOLVEpro. *Science*, 387(6732): eadr6006.
- Karimi, M.; Banerjee, S.; Jaakkola, T.; Dubrov, B.; Shang, S.; and Benson, R. 2025. Data Distillation for extrapolative protein design through exact preference optimization. In *The Thirteenth International Conference on Learning Representations*.
- Kingma, D. P.; and Ba, J. 2014. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.
- Kirjner, A.; Yim, J.; Samusevich, R.; Bracha, S.; Jaakkola, T. S.; Barzilay, R.; and Fiete, I. R. 2024. Improving Protein Optimization with Smoothed Fitness Landscapes. In *Proceedings of the 12th International Conference on Learning Representations*. ICLR 2024.
- Lin, Z.; Akin, H.; Rao, R.; Hie, B.; Zhu, Z.; Lu, W.; Smetanin, N.; Verkuil, R.; Kabeli, O.; Shmueli, Y.; et al. 2023. Evolutionary-scale prediction of atomic-level protein structure with a language model. *Science*, 379(6637): 1123–1130.
- Madani, A.; McCann, B.; Naik, N.; Keskar, N. S.; Anand, N.; Eguchi, R. R.; Huang, P.-S.; and Socher, R. 2020. Progen: Language modeling for protein generation. *arXiv preprint arXiv:2004.03497*.
- Misaki, K.; Inoue, Y.; Imajuku, Y.; Kuroki, S.; Nakamura, T.; and Akiba, T. 2025. Wider or deeper? scaling llm inference-time compute with adaptive branching tree search. *arXiv preprint arXiv:2503.04412*.
- Nisonoff, H.; Xiong, J.; Allenspach, S.; and Listgarten, J. 2025. Unlocking Guidance for Discrete State-Space Diffusion and Flow Models. Preprint, *arXiv:2406.01572*.
- Notin, P.; Kollasch, A.; Ritter, D.; Van Niekerk, L.; Paul, S.; Spinner, H.; Rollins, N.; Shaw, A.; Orenbuch, R.; Weitzman, R.; et al. 2023. Proteingym: Large-scale benchmarks for protein fitness prediction and design. *Advances in Neural Information Processing Systems*, 36: 64331–64379.
- Notin, P.; Rollins, N.; Gal, Y.; Sander, C.; and Marks, D. 2024. Machine learning for functional protein design. *Nature biotechnology*, 42(2): 216–228.
- Olsen, T. H.; Moal, I. H.; and Deane, C. M. 2022. AbLang: an antibody language model for completing antibody sequences. *Bioinformatics Advances*, 2(1): vbac046.
- Rafailov, R.; Sharma, A.; Mitchell, E.; Manning, C. D.; Ermon, S.; and Finn, C. 2023. Direct preference optimization: Your language model is secretly a reward model. *Advances in neural information processing systems*, 36: 53728–53741.
- Rao, J.; Xie, J.; Yuan, Q.; Liu, D.; Wang, Z.; Lu, Y.; Zheng, S.; and Yang, Y. 2024. A variational expectation-maximization framework for balanced multi-scale learning of protein and drug interactions. *Nature Communications*, 15(1): 4476.
- Rao, J.; Zheng, S.; Lu, Y.; and Yang, Y. 2022. Quantitative evaluation of explainable graph neural networks for molecular property prediction. *Patterns*, 3(12).

- Ren, Z.; Li, J.; Ding, F.; Zhou, Y.; Ma, J.; and Peng, J. 2022. Proximal Exploration for Model-guided Protein Sequence Design. In Chaudhuri, K.; Jegelka, S.; Song, L.; Szepesvari, C.; Niu, G.; and Sabato, S., eds., *Proceedings of the 39th International Conference on Machine Learning*, volume 162 of *Proceedings of Machine Learning Research*, 18520–18536. PMLR.
- Repecka, D.; Jauniskis, V.; Karpus, L.; Rembeza, E.; Rokaitis, I.; Zrimec, J.; Poviloniene, S.; Laurynenas, A.; Viknander, S.; Abuajwa, W.; et al. 2021. Expanding functional protein sequence spaces using generative adversarial networks. *Nature Machine Intelligence*, 3(4): 324–333.
- Sarkisyan, K. S.; Bolotin, D. A.; Meer, M. V.; Usmanova, D. R.; Mishin, A. S.; Sharonov, G. V.; Ivankov, D. N.; Bozhanova, N. G.; Baranov, M. S.; Soylemez, O.; Bogatyreva, N. S.; Vlasov, P. K.; Egorov, E. S.; Logacheva, M. D.; Kondrashov, A. S.; Chudakov, D. M.; Putintseva, E. V.; Mamedov, I. Z.; Tawfik, D. S.; Lukyanov, K. A.; and Kondrashov, F. A. 2016. Local fitness landscape of the green fluorescent protein. *Nature*, 533(7603): 397–401.
- Sinai, S.; Wang, R.; Whatley, A.; Slocum, S.; Locane, E.; and Kelsic, E. D. 2020. AdaLead: A simple and robust adaptive greedy search algorithm for sequence design. arXiv:2010.02141.
- Su, J.; Han, C.; Zhou, Y.; Shan, J.; Zhou, X.; and Yuan, F. 2024. SaProt: Protein Language Modeling with Structure-aware Vocabulary. In Kim, B.; Yue, Y.; Chaudhuri, S.; Fragkiadaki, K.; Khan, M.; and Sun, Y., eds., *International Conference on Representation Learning*, volume 2024, 6987–7009.
- Trabucco, B.; Kumar, A.; Geng, X.; and Levine, S. 2021. Conservative Objective Models for Effective Offline Model-Based Optimization. In Meila, M.; and Zhang, T., eds., *Proceedings of the 38th International Conference on Machine Learning*, volume 139 of *Proceedings of Machine Learning Research*, 10358–10368. PMLR.
- Tsuboyama, K.; Dauparas, J.; Chen, J.; Laine, E.; Mohseni Behbahani, Y.; Weinstein, J. J.; Mangan, N. M.; Ovchinnikov, S.; and Rocklin, G. J. 2023. Mega-scale experimental analysis of protein folding stability in biology and design. *Nature*, 620(7973): 434–444.
- Van Kempen, M.; Kim, S. S.; Tumescheit, C.; Mirdita, M.; Lee, J.; Gilchrist, C. L.; Söding, J.; and Steinegger, M. 2024. Fast and accurate protein structure search with Foldseek. *Nature biotechnology*, 42(2): 243–246.
- Vázquez Torres, S.; Leung, P. J.; Venkatesh, P.; Lutz, I. D.; Hink, F.; Huynh, H.-H.; Becker, J.; Yeh, A. H.-W.; Juergens, D.; Bennett, N. R.; et al. 2024. De novo design of high-affinity binders of bioactive helical peptides. *Nature*, 626(7998): 435–442.
- Wang, C.; Uehara, M.; He, Y.; Wang, A.; Biancalani, T.; Lal, A.; Jaakkola, T.; Levine, S.; Wang, H.; and Regev, A. 2024. Fine-tuning discrete diffusion models via reward optimization with applications to dna and protein design. arXiv preprint arXiv:2410.13643.
- Wang, J.; Lianza, S.; Juergens, D.; Tischer, D.; Watson, J. L.; Castro, K. M.; Ragotte, R.; Saragovi, A.; Milles, L. F.; Baek, M.; et al. 2022. Scaffolding protein functional sites using deep learning. *Science*, 377(6604): 387–394.
- Widatalla, T.; Rafailov, R.; and Hie, B. 2024. Aligning protein generative models with experimental fitness via direct preference optimization. *bioRxiv*, 2024–05.
- Wilson, J. T.; Moriconi, R.; Hutter, F.; and Deisenroth, M. P. 2017. The reparameterization trick for acquisition functions. arXiv:1712.00424.
- Wu, L.; Trippe, B.; Naesseth, C.; Blei, D.; and Cunningham, J. P. 2023. Practical and Asymptotically Exact Conditional Sampling in Diffusion Models. In Oh, A.; Naumann, T.; Globerson, A.; Saenko, K.; Hardt, M.; and Levine, S., eds., *Advances in Neural Information Processing Systems*, volume 36, 31372–31403. Curran Associates, Inc. NeurIPS 2023.
- Xiong, J.; Nisonoff, H.; Lukarska, M.; Gaur, I.; Oltrogge, L. M.; Savage, D. F.; and Listgarten, J. 2025. Guide your favorite protein sequence generative model. arXiv preprint arXiv:2505.04823.
- Xu, D.; Rao, J.; Zhu, M.; Zhang, J.; Lu, W.; Zheng, S.; and Yang, Y. 2025. Accurately Predicting Protein Mutational Effects via a Hierarchical Many-Body Attention Network. In *Thirty-ninth Annual Conference on Neural Information Processing Systems*.
- Yeh, A. H.-W.; Norn, C.; Kipnis, Y.; Tischer, D.; Pellock, S. J.; Evans, D.; Ma, P.; Lee, G. R.; Zhang, J. Z.; Anishchenko, I.; et al. 2023. De novo design of luciferases using deep learning. *Nature*, 614(7949): 774–780.
- Yuan, Q.; Chen, J.; Zhao, H.; Zhou, Y.; and Yang, Y. 2022. Structure-aware protein–protein interaction site prediction using deep graph convolutional network. *Bioinformatics*, 38(1): 125–132.
- Yuan, Q.; Tian, C.; Song, Y.; Ou, P.; Zhu, M.; Zhao, H.; and Yang, Y. 2024. GPSFun: geometry-aware protein sequence function predictions with language models. *Nucleic Acids Research*, 52(W1): W248–W255.
- Zhu, M.; Song, Y.; Yuan, Q.; and Yang, Y. 2024. Accurately predicting optimal conditions for microorganism proteins through geometric graph learning and language model. *Communications Biology*, 7(1): 1709.