



Multi-objective optimization for designing structurally similar proteins with dissimilar sequences

Ryo Akiba, Yoshitaka Moriwaki, Ryuichiro Ishitani & Naruki Yoshikawa

To cite this article: Ryo Akiba, Yoshitaka Moriwaki, Ryuichiro Ishitani & Naruki Yoshikawa (09 Jan 2026): Multi-objective optimization for designing structurally similar proteins with dissimilar sequences, Science and Technology of Advanced Materials: Methods, DOI: [10.1080/27660400.2025.2611575](https://doi.org/10.1080/27660400.2025.2611575)

To link to this article: <https://doi.org/10.1080/27660400.2025.2611575>



© 2026 The Author(s). Published by National Institute for Materials Science in partnership with Taylor & Francis Group



Accepted author version posted online: 09 Jan 2026.



Submit your article to this journal [↗](#)



Article views: 26



View related articles [↗](#)



View Crossmark data [↗](#)

We propose a multi-objective optimization framework for protein design that simultaneously optimizes structural and sequence similarity. The proposed method can generate proteins whose sequences differ from the reference while maintaining structural similarity.

RESEARCH ARTICLE

Multi-objective optimization for designing structurally similar proteins with dissimilar sequences

Ryo Akiba^a, Yoshitaka Moriwaki^b, Ryuichiro Ishitani^b, and Naruki Yoshikawa^b

^aSchool of Life Science and Technology, Institute of Science Tokyo, Tokyo, Japan

^bMedical Research Laboratory, Institute of Integrated Research, Institute of Science Tokyo, Tokyo, Japan

ARTICLE HISTORY

Compiled December 29, 2025

ABSTRACT

Recent advances in artificial intelligence technologies have accelerated the development of computational protein design techniques. Although the structure of the designed proteins has been the primary focus, diversity of designed amino acid sequences is another important aspect of protein design. To address the trade-off between reducing sequence similarity and improving structural similarity, simultaneous optimization of these two objectives can be effective. We present a method that integrates ProteinMPNN with the multi-objective optimization algorithm NSGA-II to design proteins that retain high structural similarity to a target while exhibiting low sequence similarity. Using Top7 as a reference protein, we demonstrate that our approach can design proteins with lower sequence similarity to the reference compared to the original ProteinMPNN, while maintaining comparable structural similarity.

KEYWORDS

protein design; multi-objective optimization; generative AI

1. Introduction

Protein design is an important research area that can be applied to diverse applications, such as drug discovery and enzyme development. With recent advances in artificial intelligence (AI), various generative AI methods for protein design have been proposed [1]. Among these, two-staged protein design methods have demonstrated superior performance and attracted considerable attention. They divide the protein design process into the generation of three-dimensional structure and the generation of amino acid sequences that fold into the designed structure.

CONTACT Ryuichiro Ishitani and Naruki Yoshikawa (naruki.yoshikawa@alumni.utoronto.ca)

In the first stage of structural design, diffusion models, a type of deep generative model, are widely used [2]. Diffusion-based methods such as RFDiffusion [3] and Chroma [4] generate diverse protein backbones by gradually transforming random noise into meaningful structures. Another approach, known as the hallucination method [5, 6], searches for structures that maximize an objective function using structure prediction models.

In the second stage of sequence design, conventional approaches have relied on physics-based tools that search for sequences based on energy calculations [1]. More recently, machine learning methods using graph neural networks have emerged. For example, ProteinMPNN (pMPNN) [7] employs a message passing neural network that represents protein backbones as graphs, where interatomic distances and orientations are encoded as features. This enables the model to predict amino acid sequences that fold into structures similar to a given backbone. Compared with traditional physics-based methods, pMPNN achieves higher sequence recovery for protein backbones. Moreover, its designs have been experimentally validated, confirming that the generated sequences robustly produce the target structures [6, 3].

However, pMPNN tends to generate similar sequences. This lack of diversity limits exploration of the broader landscape and hinders the discovery of proteins optimized not only for structure but also for additional properties, such as stability and solubility, which are critical in practical applications. Although increasing the temperature parameters is a common strategy to improve the sequence diversity, it reduces structural similarity to the target structure. To improve the diversity of designed proteins, several machine learning methods have been proposed [8, 9, 10].

Since proteins with similar amino acid sequences tend to have similar structures, there is a trade-off between reducing sequence similarity and improving structural similarity. Multi-objective optimization [11] is often employed to simultaneously optimize multiple objective functions in trade-offs. Recently, several studies have explored the introduction of multi-objective optimization algorithms into protein sequence design. For example, Luo et al. proposed a deep learning-guided algorithm that designs amino acid sequences that jointly optimize multiple experimental properties, such as fluorescence intensity and stability, or binding affinities to multiple targets [12]. Hong et al. integrated AI models into the NSGA-II framework and successfully optimized the stability of switch proteins in different conformational states, as well as the stability of proteins in various binding states [13]. However, these studies primarily focus on improving protein performance, and sequence diversity was not explicitly considered.

In this study, we propose a protein structure design method that combines pMPNN with the multi-objective optimization algorithm NSGA-II [14] to simultaneously optimize two objectives in trade-off: sequence dissimilarity and structural similarity. Using the TOP7 protein as a test case, our method successfully designed novel proteins that are structurally similar to the original protein while exhibiting low sequence similarity, with less than 10% of amino acids identical to the original sequence. These results demonstrate that the performance of generative AI models for protein design can be improved by combining them with multi-objective optimization methods.

2. Method

In this study, we propose an algorithm that combines the protein design method pMPNN, which generates sequences that fold into a given structure, with the multi-objective optimization algorithm NSGA-II to enable the design of proteins that simultaneously improve multiple

evaluation criteria.

2.1. Multi-objective optimization

The problem of simultaneously optimizing multiple objective functions is called a multi-objective optimization problem [11], and it can be formulated as Equation (1).

$$\min_{x \in \Omega} F(x) = (f_1(x), \dots, f_k(x)), \quad (1)$$

where Ω is the domain of the variable x , $k \geq 2$ is the number of objective functions, and $f_i(x)$ ($i = 1, \dots, k$) are objective functions.

Since there is no single solution that simultaneously optimizes all conflicting objective functions, the goal of multi-objective optimization is to find a set of solutions that are not dominated by any other feasible solutions. The variable x is said to dominate another variable x' when Equation (2) holds.

$$(\forall i \in \{1, \dots, k\}, f_i(x) \leq f_i(x')) \wedge (\exists j \in \{1, \dots, k\}, f_j(x) < f_j(x')) \quad (2)$$

A solution x that is not dominated by any other solution x' is called a Pareto optimal solution, and the set of all Pareto optimal solutions is called the Pareto front.

In this study, we aim to design novel proteins whose structures are similar to a known reference protein but exhibit low sequence similarity. This problem can be formulated as a multi-objective optimization problem with two objective functions: the sequence recovery rate $f_{\text{recovery}}(x)$, which measures the sequence similarity between a designed protein and the reference protein, and the structural similarity $f_{\text{structure}}(x)$, which measures similarity between a predicted structure of a designed protein and the known structure of the reference protein. This multi-objective optimization problem with $k = 2$ is expressed as Equation (3).

$$\min_{x \in \Omega} F(x) = (f_{\text{recovery}}(x), f_{\text{structure}}(x)), \quad (3)$$

where Ω is the set of all amino acid sequences with the same residue length n as the reference protein. $f_{\text{recovery}}(x)$ is a function to measure the sequence similarity between the designed protein x and the reference protein r . It is defined as Equation (4).

$$f_{\text{recovery}}(x) = \frac{1}{n} \sum_{i=1}^n \delta(x_i, r_i), \quad (4)$$

where $\delta(x_i, r_i)$ is an indicator function that returns 1 when the amino acids at position i match and 0 otherwise.

$f_{\text{structure}}(x)$ is a function representing the structural similarity between the predicted structure $AF(x)$, the structure predicted by AlphaFold2 [15] from the amino acid sequence x , and the structure of the reference protein R . The AlphaFold2 predictions were performed via LocalColabFold [16]. As a metric for structural similarity, we use the Template Modeling score (TM-score) [17], which assigns higher values to more similar structures. $f_{\text{structure}}(x)$ is defined by

inverting the sign of TM-score as in Equation (5) to treat this as a minimization problem.

$$f_{\text{structure}}(x) = -\text{TM-score}(AF(x), R) \quad (5)$$

TM-scores were computed with USalign [18] using option 1, which superposes two structures by assuming that residues with the same index are equivalent between them, without performing structure-based alignment. Sequence-independent structural alignment can inflate similarity by matching discontinuous fragments, whereas our objective is global positional correspondence along the chain. Residue-by-residue alignment thus provides a more stable objective for optimization.

2.2. NSGA-II and mutation with ProteinMPNN

We used NSGA-II [14] to solve the multi-objective optimization problem. It is a widely used evolutionary algorithm that generates offspring from the parent population by genetic operations, such as crossover and mutation, and selects the next generation using non-dominated sorting and crowding distance to approximate the Pareto front. Instead of genetic operations in the original NSGA-II, we introduced a mutation operation that partially applies pMPNN to randomly selected three consecutive amino acids of a protein (Figure 1). By combining pMPNN and NSGA-II, we could incorporate the efficiency of generative AI algorithms into a conventional multi-objective optimization algorithm. The mutation operation using pMPNN is shown in Algorithm 1.

Algorithm 1 Mutation Operation Using ProteinMPNN

```

1: procedure MUTATEWITHPROTEINMPNN(Population of sequences  $P$ , Reference
   protein structure  $R$ )
2:    $Q \leftarrow \{\}$  ▷ Initialize the list of offspring
3:   for  $p \in P$  do
4:      $pos \leftarrow$  randomly select one position from  $\{1, \dots, \text{len}(p)\}$ 
5:      $W \leftarrow \{pos - 1, pos, pos + 1\}$ 
6:      $p_{\text{masked}} \leftarrow \text{MASK}(p, W)$  ▷ Mask amino acids at positions  $W$ 
7:      $p_{\text{mutated}} \leftarrow \text{PROTEINMPNN}(R, p_{\text{masked}})$  ▷ Redesign masked region
8:      $Q \leftarrow Q \cup \{p_{\text{mutated}}\}$  ▷ Add generated sequence to list
9:   end for
10:  return  $Q$ 
11: end procedure

```

3. Results and Discussion

To evaluate the proposed algorithm, we conducted an experiment designing novel amino acid sequences using Top7 [19] as the reference protein. We executed the proposed method 24 times, each run starting from an initial population of amino acid sequences randomly generated with different seeds. Each sequence was represented by 92 characters, corresponding to the number of modeled residues in the Protein Data Bank entry of Top7 (PDB ID: 1QYS). The initial population size was set to 20, the number of generations was set to 500, and the temperature parameter of pMPNN was set to 0.3. For comparison, we also generated sequences using the

original pMPNN without any modification or combinations with other techniques (vanilla pMPNN) on the same reference protein under five temperature settings: 0.3, 0.7, 1.0, 2.0, and 3.0. To ensure a fair comparison, each setting was run for 15 hours on a single H100 GPU to match the computational budget used for the proposed method. Note that the reference protein for TM-score calculation was the Top7 structure with selenomethionine residues replaced by methionine residues.

Figure 2 compares the distribution of f_{recovery} values of proteins designed by the proposed method and vanilla pMPNN, and Figure 3 shows the Pareto front of designed proteins obtained from each method. Our method successfully generated proteins in which fewer than 10% of the amino acids are identical to those in the reference protein. Although the vanilla pMPNN could generate proteins with low recovery values by increasing the temperature, their recovery values were higher than the values attained by our method. Furthermore, the vanilla pMPNN could not generate proteins with high structural similarity when the temperature was increased. These results demonstrate that our method outperforms the vanilla pMPNN in generating proteins with low sequence similarity while maintaining high structural similarity.

To quantitatively evaluate the diversity of designed proteins, we calculated the average pairwise sequence distance defined as the normalized Hamming distance, i.e., the proportion of positions at which amino acids of the two sequences differ. Table 1 shows the average pairwise sequence distance of proteins designed by each method. The proposed method exhibited higher sequence diversity than pMPNN at low temperatures (0.3 or 0.7), and comparable to pMPNN at a temperature of 1.0. This indicates that the proposed method could effectively generate diverse proteins as the standard method.

The protein sequence similarity network [20] of the protein designed by the proposed method is shown in Figure 4. A sequence similarity network is a network constructed based on the similarity between protein sequences and is used to visualize relationships and cluster structures within a set of sequences. In the proposed method, proteins derived from different seeds form multiple loose clusters and exhibit distributions that rarely overlap between seeds. This indicates that the diversity of proteins designed by the proposed method arises from the use of different random seeds.

Figure 5 shows the Pareto fronts obtained by the proposed method using different random seeds. It indicates that its performance varies depending on the seed values. To evaluate the quality of these solution sets in comparison with other methods, we calculated the hypervolume indicator [11]. The hypervolume indicator measures the volume of the objective space dominated by the solution set relative to a reference point. In this study, the reference point was defined by combining the worst values for each objective function observed across the union of Pareto fronts from all methods. A larger hypervolume value indicates better convergence and diversity relative to this worst-case boundary. Table 2 shows the comparison of hypervolume values. The proposed method achieved an average hypervolume of 0.261 ± 0.010 , with a maximum of 0.276, which exceeded the values obtained by the other methods. Among the pMPNN settings, the temperature parameter of 1.0 yielded the highest hypervolume value (0.173). Even the minimum value of 0.243 recorded across the 24 runs of the proposed method was higher than this best vanilla pMPNN result. This confirms that the proposed method was efficient in designing proteins with dissimilar sequence while maintaining structural similarity.

To validate the contribution of generative AI within the evolutionary framework, we compared the proposed method with a baseline NSGA-II implementation that employs a standard random mutation operator. In this baseline approach, selected amino acids were

replaced with random residues rather than sequences predicted by pMPNN. Figure 6 illustrates the Pareto fronts obtained by both methods. It indicates that the proteins generated by the random mutation method exhibited lower structural similarity compared to those generated by our method. Furthermore, the random mutation method achieved a hypervolume of 0.099 (Table 2), which is lower than the average of 0.261 achieved by the proposed method. This result demonstrates that the integration of pMPNN enables the efficient exploration of the protein sequence landscape and highlights the crucial role of integrating generative AI methods with conventional multi-objective optimization frameworks.

Figure 7 shows the distribution of $f_{\text{structure}}$ and f_{recovery} for proteins obtained from a single run of the proposed method starting from one random seed. In the initial generations, only proteins with TM-scores around 0.2 to 0.4 were obtained. However, as the generations progressed, proteins with higher TM-scores were obtained while maintaining low recovery values.

As a demonstrative example, we selected one protein designed by our method and superimposed its structures predicted by AlphaFold2, which was referenced during the design, and AlphaFold3 [21], which was not involved in the design process, onto the target structure using USalign (Figure 8). The structure predicted by AlphaFold2 showed a high pLDDT value of 93.12. Superimposition with the target structure (Figure 8a) yielded a TM-score of 0.96, indicating a very close match and supporting the effectiveness of the design. In addition, the AlphaFold3 prediction, despite being unused during design, had high confidence with a pTM score of 0.84 and a pLDDT value of 90.17. This structure also aligned closely with the reference, with a TM-score of 0.94 and RMSD of 0.93 Å... upon superimposition (Figure 8). This high structural agreement obtained from an independent prediction method suggests that the design is not simply overfitting to the AlphaFold2, but is likely capable of forming the intended target structure experimentally.

4. Conclusion

In this paper, we proposed a method that combines pMPNN with a multi-objective optimization algorithm to design proteins with high structural similarity but low sequence similarity. Evaluation experiments showed that the proposed method designs proteins with lower sequence similarity to the reference compared to the vanilla pMPNN, while maintaining comparable structural similarity. While this study used a similarity of predicted three-dimensional structure as the objective function, the framework of the multi-objective optimization method presented here can be applied to other objective functions as well.

Although the proposed method was more effective than the vanilla pMPNN, several limitations remain. The primary limitation is the large computational cost of the proposed method. The current algorithm requires many evaluation function calls, because of the nature of the evolutionary algorithm that modifies only a small portion of sequence at a time. While we utilized relatively fast computational evaluation functions in this study, our method can be extended to other functions, including experimental evaluations, which may require a longer time to evaluate. Therefore, further improvement in design efficiency is necessary, potentially by incorporating more efficient generative AI approaches, such as protein language models, to perform protein design with fewer evaluation calls. Another limitation is the generalizability of our approach. This study only validated our method using the Top7 protein, which is known to be highly designable. We will evaluate the adaptability to other proteins in the future. Finally, the applicability of the proposed method depends on the availability of reliable evaluation functions.

In this study, we employ AlphaFold2 for protein folding evaluation, which offers a reasonable degree of reliability. However, as current computational methods still lack sufficient accuracy in predicting the folding of de novo designed proteins, experimental validation remains necessary. Experimental validation of the structural similarity of the designed proteins is part of our future work.

Additional information

Data availability

The code and data used in this work are available on GitHub (<https://github.com/Aketami23/Multipurpose-Optimization-of-Protein-Sequences>).

Funding

This research was partially supported by the Research Support Project for Life Science and Drug Discovery (Basis for Supporting Innovative Drug Discovery and Life Science Research (BINDS)) from AMED under Grant Number JP25ama121012 for R.I. This work was also supported by JSPS KAKENHI Grant-in-Aid for Transformative Research Areas (A) under Grant Numbers 25H01570 and 25H02250 for R.I and JSPS KAKENHI Grant Number JP25K21333 for N.Y. Additional support was provided by the Medical Research Center Initiative for High Depth Omics at the Institute of Science Tokyo, Nanken-Kyoten at the Institute of Science Tokyo 2025, and the Multilayered Stress Diseases project (JPMXP1323015483) at the Institute of Science Tokyo. This work utilized computational resources from the TSUBAME4.0 supercomputer at the Institute of Science Tokyo.

Conflict of interest

The authors report there are no competing interests to declare.

References

- [1] Kortemme T. De novo protein design—from new structures to programmable functions. *Cell*. 2024;187(3):526–544.
- [2] Yim J, Stärk H, Corso G, et al. Diffusion models in protein structure and docking. *Wiley Interdiscip Rev Comput Mol Sci*. 2024;14(2):e1711.
- [3] Watson JL, Juergens D, Bennett NR, et al. De novo design of protein structure and function with RFdiffusion. *Nature*. 2023;620(7976):1089–1100.
- [4] Ingraham JB, Baranov M, Costello Z, et al. Illuminating protein space with a programmable generative model. *Nature*. 2023;623(7989):1070–1078.
- [5] Anishchenko I, Pellock SJ, Chidyausiku TM, et al. De novo protein design by deep network hallucination. *Nature*. 2021;600(7889):547–552.
- [6] Wicky BI, Milles LF, Courbet A, et al. Hallucinating symmetric protein assemblies.

Science. 2022;378(6615):56–61.

[7] Dauparas J, Anishchenko I, Bennett N, et al. Robust deep learning–based protein sequence design using ProteinMPNN. *Science*. 2022;378(6615):49–56.

[8] Bryant DH, Bashir A, Sinai S, et al. Deep diversification of an AAV capsid protein by machine learning. *Nat Biotechnol*. 2021;39(6):691–696.

[9] Sturmfels P, Rao R, Verkuil R, et al. Seq2MSA: a language model for protein sequence diversification. In: *Machine Learning in Structural Biology Workshop, NeurIPS 2022*; 2022.

[10] Silva LA, Meynard-Piganeau B, Lucibello C, et al. Fast uncovering of protein sequence diversity from structure. In: *The Thirteenth International Conference on Learning Representations*; 2025. Available from: <https://openreview.net/forum?id=1iuaxjssVp> .

[11] Coello CAC, Lamont GB, Veldhuizen DAV. *Evolutionary algorithms for solving multi-objective problems*. Springer, New York (NY); 2007.

[12] Luo J, Ding K, Luo Y. Pareto-optimal sampling for multi-objective protein sequence design. *iScience*. 2025;28(3).

[13] Hong L, Kortemme T. An integrative approach to protein sequence design through multiobjective optimization. *PLoS Comput Biol*. 2024;20(7):e1011953.

[14] Deb K, Pratap A, Agarwal S, et al. A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Trans Evol Comput*. 2002;6(2):182–197.

[15] Jumper J, Evans R, Pritzel A, et al. Highly accurate protein structure prediction with AlphaFold. *Nature*. 2021;596(7873):583–589.

[16] Mirdita M, Schütze K, Moriwaki Y, et al. ColabFold: making protein folding accessible to all. *Nat Methods*. 2022;19:679–682.

[17] Zhang Y, Skolnick J. Scoring function for automated assessment of protein structure template quality. *Proteins: Struct Funct Bioinf*. 2004;57(4):702–710.

[18] Zhang C, Shine M, Pyle AM, et al. US-align: universal structure alignments of proteins, nucleic acids, and macromolecular complexes. *Nat Methods*. 2022;19(9):1109–1115.

[19] Kuhlman B, Dantas G, Ireton GC, et al. Design of a novel globular protein fold with atomic-level accuracy. *Science*. 2003;302(5649):1364–1368.

[20] MiguelMSandin. SSNetworks ; 2022. Available from: <https://doi.org/10.5281/zenodo.7430471> .

[21] Abramson J, Adler J, Dunger J, et al. Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature*. 2024;630(8016):493–500.

[22] Schrödinger, LLC. The PyMOL molecular graphics system, version 2.5; 2021. Available from the PyMOL website: <https://pymol.org/>.

Table 1. Average pairwise normalized Hamming distance of proteins designed by each method. Proteins were filtered with pLDDT ≥ 90 and TM-score ≥ 0.9 , and 384 proteins were sampled to reduce calculation time. The samples for the proposed method were obtained by aggregating 16 samples from each of 24 independent seeds.

Method	Avg. distance
Proposed method	0.716
ProteinMPNN (temp = 1.0)	0.739
ProteinMPNN (temp = 0.7)	0.627
ProteinMPNN (temp = 0.3)	0.385

Table 2. Comparison of hypervolume indicator values. The proposed method shows the mean and standard deviation across 24 independent runs. The other methods show the value for a single run with the same computational budget.

Method	Hypervolume
Proposed Method	0.261 ± 0.010
ProteinMPNN (temp = 0.3)	0.071
ProteinMPNN (temp = 0.7)	0.127
ProteinMPNN (temp = 1.0)	0.173
ProteinMPNN (temp = 2.0)	0.137
ProteinMPNN (temp = 3.0)	0.136
NSGA-II (Random Mutation)	0.099

Figure 1. Overview of the proposed protein design workflow integrating NSGA-II and ProteinMPNN. At each generation, local sequence mutations are introduced by ProteinMPNN, followed by structure prediction and evaluation of sequence recovery and structural similarity to the target protein. By repeating selection and mutation, proteins that optimize both sequence recovery and structural similarity are progressively obtained.

Figure 2. Comparison of sequence similarity for proteins obtained by the proposed method and the vanilla pMPNN with different temperatures. For the proposed method, the proteins with the lowest sequence similarity that satisfy TM-score ≥ 0.90 and pLDDT ≥ 90 are plotted for each run. For pMPNN, the generated proteins were plotted after filtering by TM-score ≥ 0.90 and

pLDDT \geq 90. No generated protein passed this filter when the temperature was set to 2.0 or 3.0.

Figure 3. Pareto fronts obtained by the proposed method and the vanilla pMPNN with different temperatures. The proposed method shows the Pareto front from a single representative run. The Pareto front for pMPNN was constructed from the proteins designed by using the same compute budget (identical GPU and runtime) as the proposed method.

Figure 4. t-SNE plot of the Sequence Similarity Network (SSN) [20] of proteins generated by the proposed method and pMPNN. The distance matrix was calculated using Hamming distance. Points represent sequences generated by the proposed method from different random seeds that satisfy the criteria TM-score \geq 0.9 and pLDDT \geq 90. Up to 300 sequences are sampled from the Pareto front of each seed and are color-coded by seed.

Figure 5. Pareto front obtained for each random seed.

Figure 6. Comparison of Pareto fronts obtained by NSGA-II with the proposed pMPNN-based mutation strategy and with random mutation.

Figure 7. Distribution of $f_{\text{structure}}$ and f_{recovery} for proteins obtained from a single run of the proposed method. Each protein is displayed in a different color according to the order in which it was obtained.

Figure 8. Predicted three-dimensional structure of the protein designed by the proposed method. PyMol [22] was used for visualization.

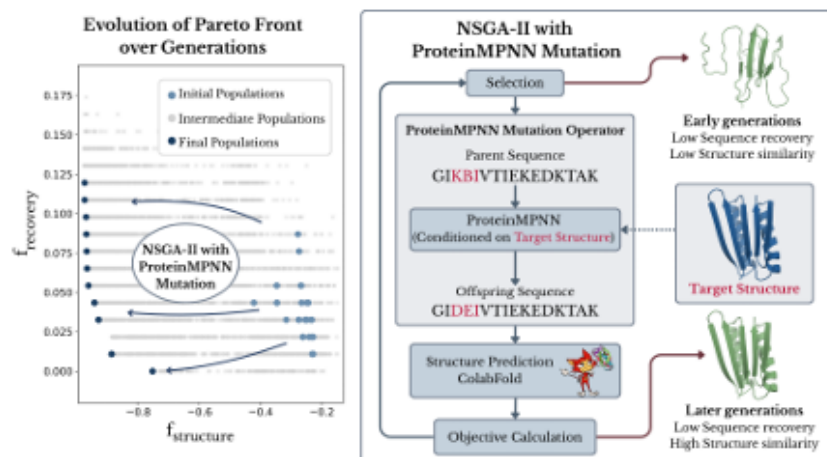


Figure 1. Overview of the proposed protein design workflow integrating NSGA-II and ProteinMPNN. At each generation, local sequence mutations are introduced by ProteinMPNN, followed by structure prediction and evaluation of sequence recovery and structural similarity to the target protein. By repeating selection and mutation, proteins that optimize both sequence recovery and structural similarity are progressively obtained.

ACCEPTED MANUSCRIPT

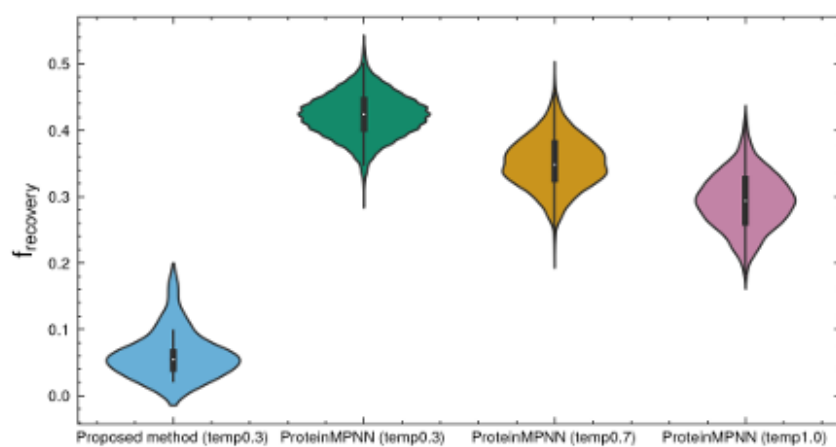


Figure 2. Comparison of sequence similarity for proteins obtained by the proposed method and the vanilla pMPNN with different temperatures. For the proposed method, the proteins with the lowest sequence similarity that satisfy $\text{TM-score} \geq 0.90$ and $\text{pLDDT} \geq 90$ are plotted for each run. For pMPNN, the generated proteins were plotted after filtering by $\text{TM-score} \geq 0.90$ and $\text{pLDDT} \geq 90$. No generated protein passed this filter when the temperature was set to 2.0 or 3.0.

ACCEPTED MANUSCRIPT

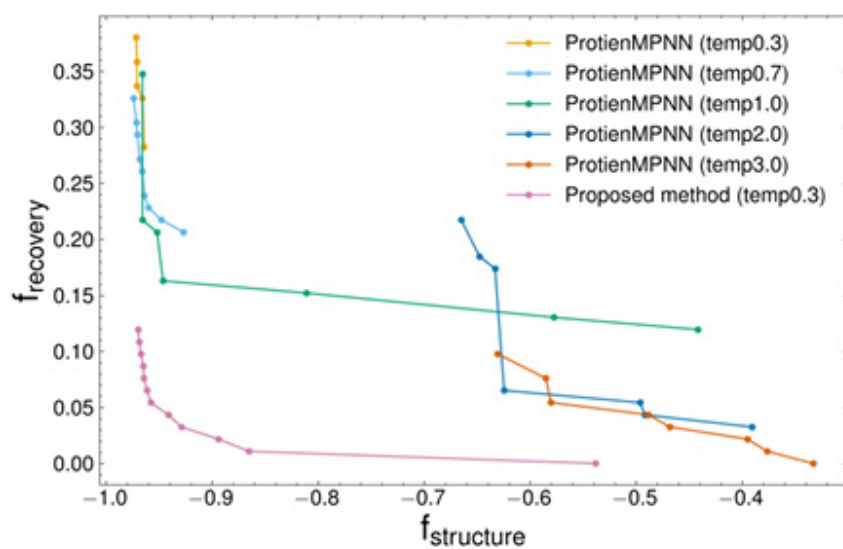


Figure 3. Pareto fronts obtained by the proposed method and the vanilla pMPNN with different temperatures. The proposed method shows the Pareto front from a single representative run. The Pareto front for pMPNN was constructed from the proteins designed by using the same compute budget (identical GPU and runtime) as the proposed method.

ACCEPTED MANUSCRIPT

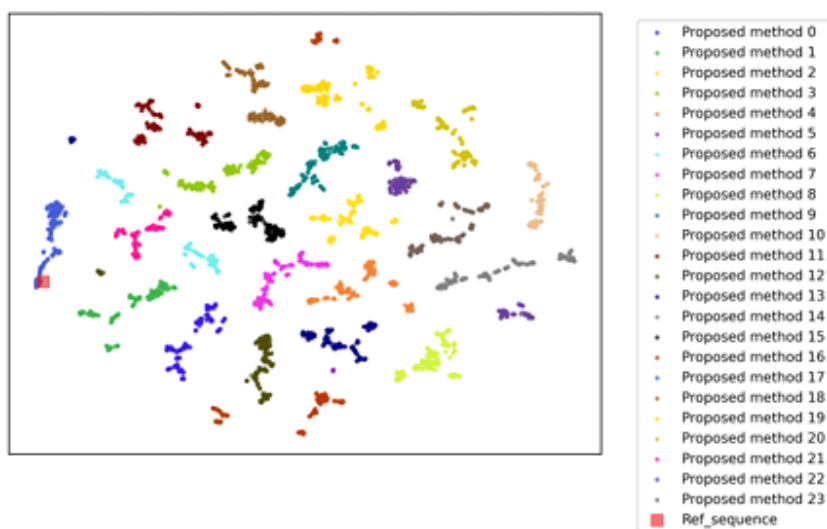


Figure 4. t-SNE plot of the Sequence Similarity Network (SSN) [20] of proteins generated by the proposed method and pMPNN. The distance matrix was calculated using Hamming distance. Points represent sequences generated by the proposed method from different random seeds that satisfy the criteria $TM\text{-score} \geq 0.9$ and $pLDDT \geq 90$. Up to 300 sequences are sampled from the Pareto front of each seed and are color-coded by seed.

ACCEPTED MANUSCRIPT

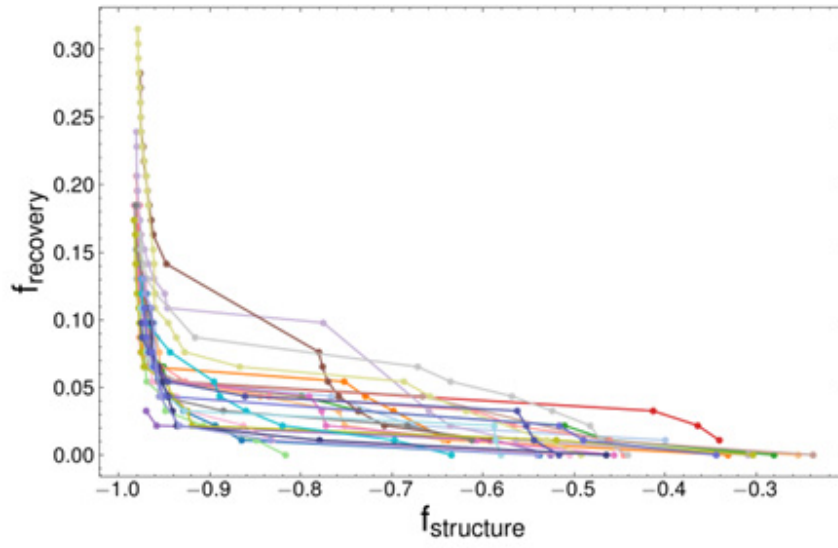


Figure 5. Pareto front obtained for each random seed.

ACCEPTED MANUSCRIPT

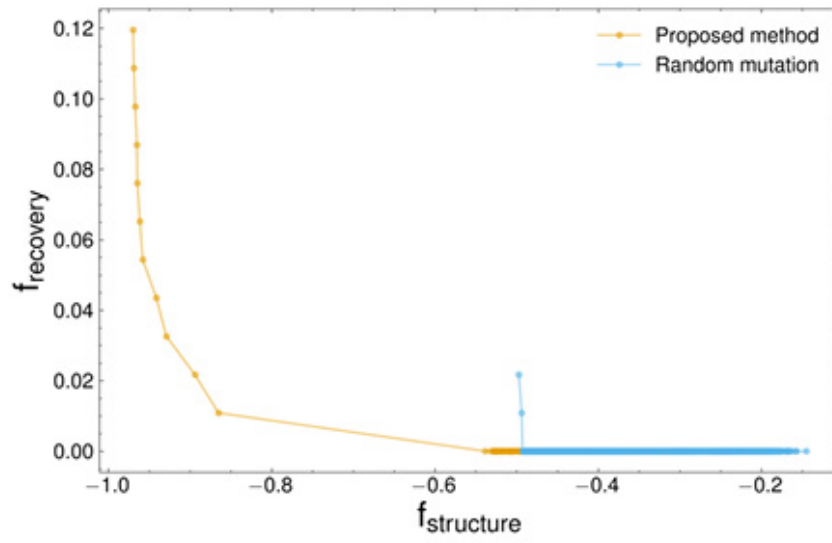


Figure 6. Comparison of Pareto fronts obtained by NSGA-II with the proposed pMPNN-based mutation strategy and with random mutation.

ACCEPTED MANUSCRIPT

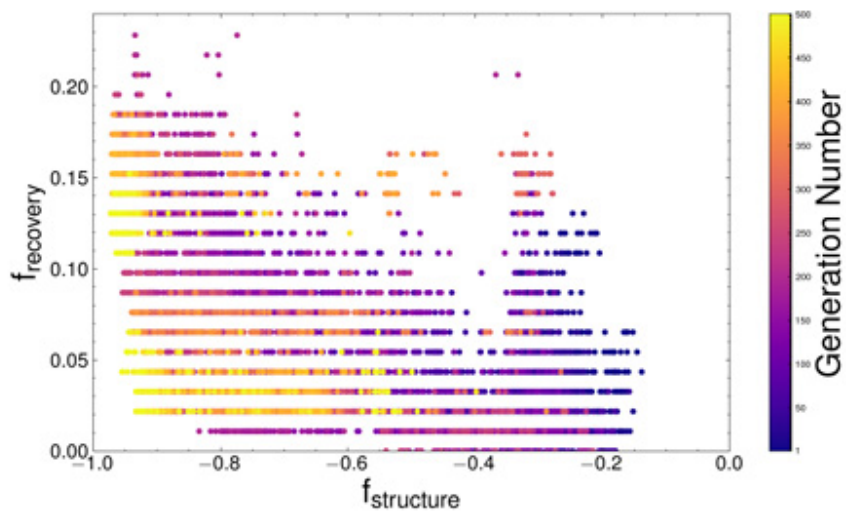
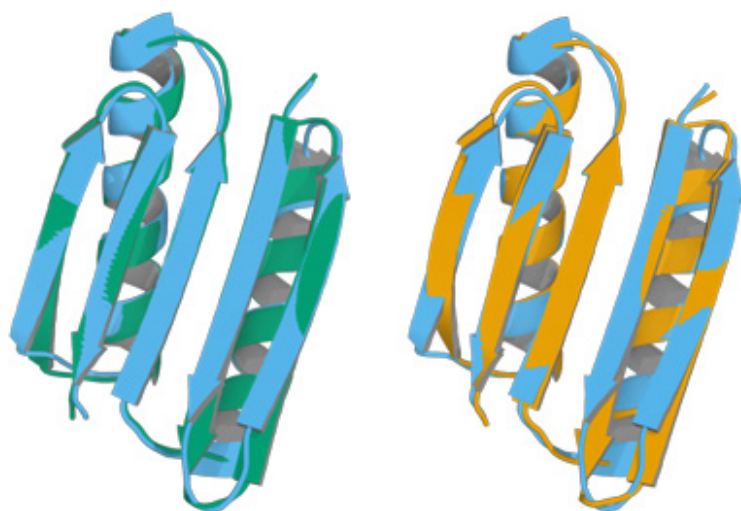


Figure 7. Distribution of $f_{\text{structure}}$ and f_{recovery} for proteins obtained from a single run of the proposed method. Each protein is displayed in a different color according to the order in which it was obtained.

ACCEPTED MANUSCRIPT



(a) Predicted structure by AlphaFold 2 (green) and target structure (light blue). (b) Predicted structure by AlphaFold 3 (orange) and target structure (light blue)

Figure 8. Predicted three-dimensional structure of the protein designed by the proposed method. PyMol [22] was used for visualization.

ACCEPTED IN