AnnoDPO: Protein Functional Annotation Learning with Direct Preference Optimization

Zixuan Jiang [1] **Renjing Xu** [1]

Abstract

Deciphering protein function remains a fundamental challenge in protein representation learning. The task presents significant difficulties for protein language models (PLMs) due to the sheer volume of functional annotation categories and the highly imbalanced distribution of annotated instances across biological ontologies. Inspired by the remarkable success of reinforcement learning from human feedback (RLHF) in large language model (LLM) alignment, we propose AnnoDPO, a novel multi-modal framework for protein function prediction that leverages Direct Preference Optimization (DPO) to enhance annotation learning. Our methodology addresses the dual challenges of annotation scarcity and category imbalance through preference-aligned training objectives, establishing a new paradigm for biological knowledge integration in protein representation learning. We provide the code for AnnoDPO at `https://github.com/AzusaXuan/AnnoDPO`.

1. Introduction

Proteins serve as the central machinery of life, executing crucial biological activities. While high-throughput sequencing technologies (Reuter et al., 2015) have driven exponential growth in sequenced genomes over two decades (Consortium, 2019; Suzek et al., 2015), functionally characterized proteins (Boeckmann et al., 2003; Gasteiger et al., 2001) lag significantly due to structural complexity and challenges in capturing interaction dynamics. This disparity underscores the persistent challenge of accurate, large-scale automated protein function prediction (Radivojac et al., 2013; Friedberg, 2006).

Traditional approaches for functional annotation—including statistical methods and rule-based systems like UniRule—remain widely adopted in protein databases (Consortium, 2019; Doğan et al., 2016; Śledź & Jinek, 2016). However, their reliance on simplified sequence-function mappings often leads to inaccuracies. Deep learning methods (Kulmanov et al., 2018; You et al., 2021; Kulmanov & Hoehndorf, 2020; Kulmanov et al., 2024; Yu et al., 2023; Jang et al., 2024) have recently emerged as superior alternatives, with PLMs (Elnaggar et al., 2021; Brandes et al., 2022; Rives et al., 2021; Meier et al., 2021) revolutionizing prediction capabilities. However, PLMs face two fundamental challenges: discerning subtle sequence variations that induce dramatic functional divergence and overcoming extreme annotation sparsity where fewer than 5% of Swiss-Prot entries contain more than 10 Gene Ontology annotations. These combined limitations maintain a persistent accuracy gap between computational predictions and expert annotations, underscoring the need to integrate domain knowledge into PLM-guided functional inference.

A crucial breakthrough has emerged in LLM alignment through RLHF (Christiano et al., 2017; Ziegler et al., 2019; Ouyang et al., 2022; Bai et al., 2022; Glaese et al., 2022), which enables AI systems to better align with human preferences. Building on these successes in natural language processing, researchers have begun exploring RLHF's potential for protein-related AI applications. Recent demonstrations span controllable protein generation (Liu et al., 2025; Stocco et al., 2024; Widatalla et al., 2024) and protein knowledge assistants (Zhou et al., 2025), establishing RLHF as a viable paradigm for biological sequence modeling. Notably, prior work has not yet explored DPO (Rafailov et al., 2023), a prominent RLHF variant that eliminates reward modeling through direct policy optimization, for protein function annotation prediction.

This study establishes three key contributions: (1) We develop an end-to-end multimodal framework integrating protein sequences with functional annotations, enhanced by contrastive learning during supervised fine-tuning (SFT) to optimize cross-modal feature alignment. (2) We pioneer the adaptation of DPO to protein language models, creating the

[1]The Hong Kong University of Science and Technology (Guangzhou). Correspondence to: Renjing Xu <renjingxu@hkust-gz.edu.cn>.

Proceedings of the ICML 2025 Workshop on Multi-modal Foundation Models and Large Language Models for Life Sciences, Vancouver, Canada. 2025. Copyright 2025 by the author(s).

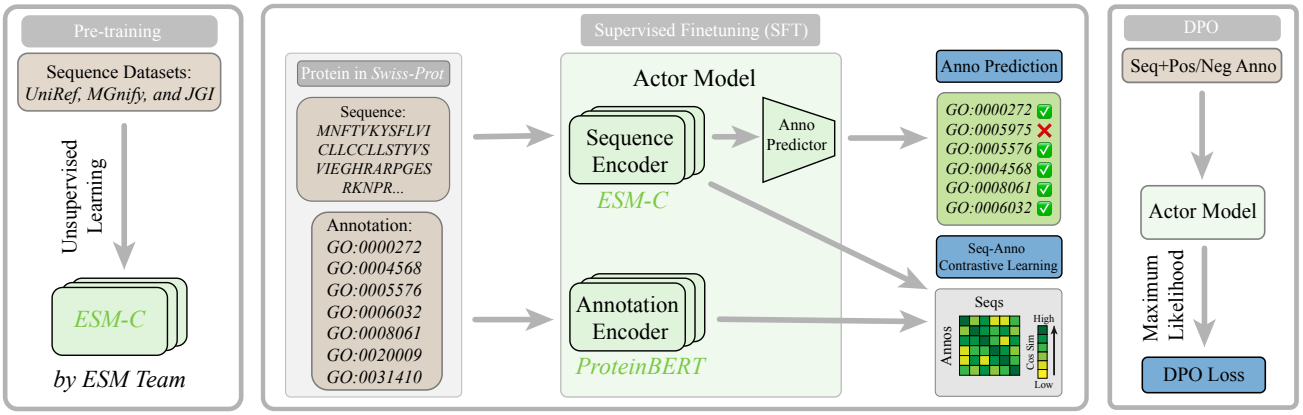

Figure 1. Model architecture and training objectives of AnnoDPO. The training framework is divided into three stages: **Pre-training**: Self-supervised learning of ESM-C on protein sequences from UniRef, MGnify, and JGI (ESM Team, 2024); **SFT**: Dual-objective finetuning with annotation prediction and sequence-annotation contrastive alignment; **DPO**: Preference optimization through positive annotations against negative ones.

first DPO-powered architecture for enhancing functional annotation accuracy. (3) We systematically characterize how DPO reshapes model attention patterns to better capture hierarchical relationships in Gene Ontology annotations.

2. Background

Protein Functional Annotation Prediction Gene Ontology (GO) (Ashburner et al., 2000) provides standardized functional descriptors across three biological domains. Predicting GO terms remains essential for characterizing unannotated proteins. The Enzyme Commission (EC) system (Tipton & Boyce, 2000) classifies enzymes via four-digit catalytic activity codes, while UniProtKB keywords (KW) (Magrane & Consortium, 2011) systematically categorize functional attributes in Swiss-Prot entries. Together, these annotation systems enable comprehensive protein function analysis.

Protein Multi-modal Learning in Annotation Prediction The integration of PLMs with multi-source data has established multimodal learning as the standard for functional annotation. Key advances include: CLEAN (Yu et al., 2023) aligning enzymes with EC numbers via contrastive learning; ProteinBERT (Brandes et al., 2022) jointly modeling sequences and GO terms; OntoProtein (Zhang et al., 2022) encoding knowledge graphs with textual descriptors. Generation paradigms like ProGen (Madani et al., 2020) utilize function labels for controllable synthesis, while ProtST (Xu et al., 2023) bridges sequences with biomedical texts. Most notably, SaProt (Su et al., 2023) achieves SOTA performance through structure-aware tokenization integrating sequence-structure relationships.

Reinforcement Learning from Human Feedback RLHF methodologies bifurcate into reward-modeling and direct preference optimization paradigms. Reward-based approaches (Stiennon et al., 2020; Ouyang et al., 2022; Christiano et al., 2017; Havrilla et al., 2024; Setlur et al., 2024) employ two-stage training: first learning reward functions from preference data, then optimizing policies via online RL algorithms like PPO (Schulman et al., 2017). Conversely, reward-free methods (Yuan et al., 2023; Song et al., 2024; Dong et al., 2023) bypass explicit reward modeling by directly optimizing language models on preference rankings. Notably, Direct Preference Optimization (DPO) (Rafailov et al., 2023) has emerged as a predominant reward-free approach due to its stable single-stage training and competitive performance. The field continues to debate fundamental trade-offs: reward-based methods' alignment precision versus reward-free approaches' computational efficiency (Li et al., 2023; Xu et al., 2024).

3. Method

Our three-stage training framework (Fig. 1) comprises pre-training, supervised finetuning (SFT) with combined annotation prediction and sequence-annotation contrastive objectives, and Direct Preference Optimization (DPO). The pre-training stage builds upon ESM Cambrian (ESM-C) (ESM Team, 2024), where we employ the 300M parameter variant as our foundational sequence encoder. We elaborate the details of SFT and DPO in the subsequent sections and hyperparameter details in Appendix A.

Dataset Curation and data input We use Swiss-Prot (Boeckmann et al., 2003) as the training set as it is one of the most widely used dataset for protein function. To ensure enough sequences for test, we choose the dataset version

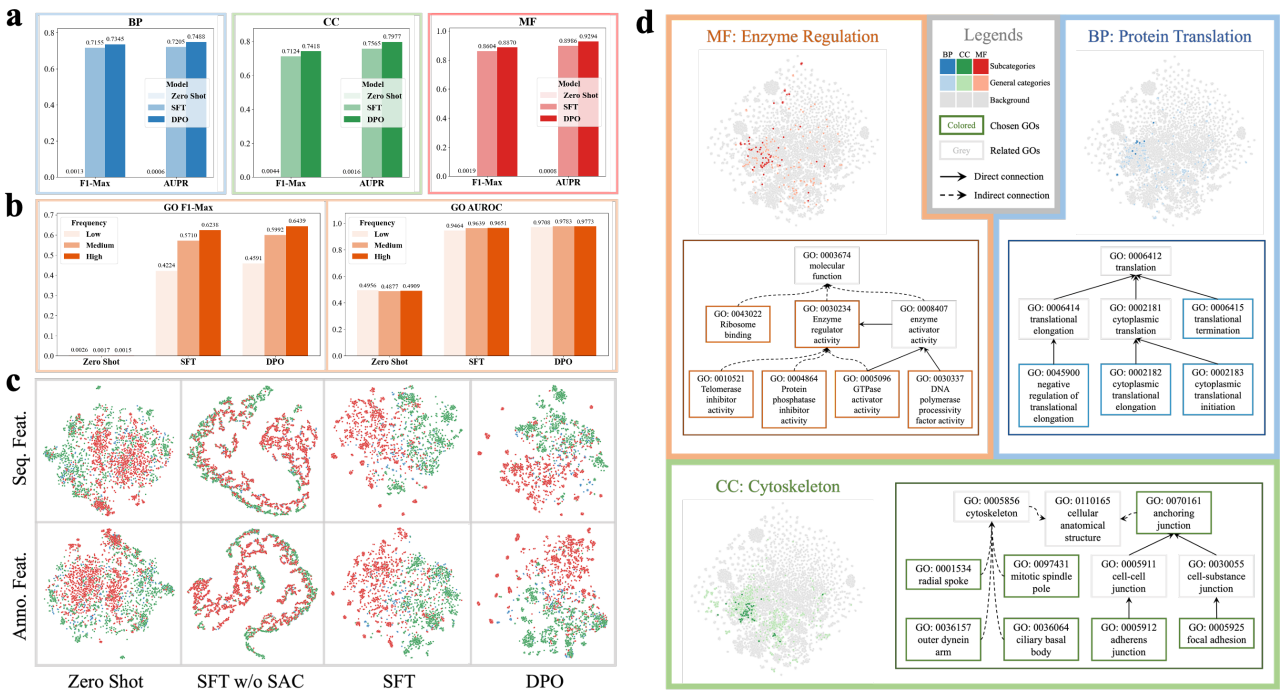

Figure 2. Comprehensive Evaluation of Protein Function Annotation Performance. **(a)** Cross-category performance comparison (numerical results in Tab. 7). **(b)** Robustness analysis across label frequency regimes (numerical results in Tab. 8). **(c)** t-SNE visualization of GO category discriminability in latent space. **(d)** Hierarchical relationship preservation in tightly-related GO term families (additional examples in Appendix D).

updated in Jan. 2010 totaling ∼510,000 sequences and spilit it at the ratio 9:1 for training and testing. Then we select all the sequences updated after that to construct the Swiss-Prot-New dataset totaling ∼60,000 sequences. We demonstrate dataset details in Appendix B.

Supervised Finetuning (SFT) The SFT stage integrates three core components: (1) a pretrained ESM-C sequence encoder (ESM Team, 2024) that converts protein sequences into embeddings, (2) an MLP-based annotation predictor generating GO term probabilities from sequence embeddings, and (3) a de novo trained ProteinBERT annotation encoder (Brandes et al., 2022) that encodes functional annotations. We establish cross-modal alignment through contrastive learning between sequence embeddings and annotation features via the sequence-annotation contrastive loss, while simultaneously optimizing annotation prediction accuracy through standard classification objectives. The mathematical formulations of these dual losses are defined as follows:

Annotation Prediction (AP) Loss This loss is a sum of the categorical cross-entropy over the protein sequences and the binary cross-entropy over the annotations, namely

$$\mathcal{L}_{\text{AP}} = -\sum_{j \in N} \left(y_j^A \log(p_j^A) + (1 - y_j^A) \log(1 - p_j^A) \right), \tag{1}$$

where $N = 7533$ denotes the size of our curated Gene Ontology vocabulary, $y_j^A \in \{0, 1\}$ indicates the presence of the j-th GO term in the ground-truth annotations, and $p_j^A \in [0, 1]$ represents the predicted probability for that term. The GO vocabulary was constructed by retaining terms with over 100 times occurrences in Swiss-Prot, ensuring sufficient statistical support for reliable learning.

Sequence-Annotation Contrastive (SAC) Loss $\mathcal{L}_{\text{SAC}}$ implements bidirectional alignment between sequence features h^S and annotation features h^A through normalized feature matching. Given a positive pair $(\mathrm{h}^S, \mathrm{h}^A)$ where i indexes protein sequences and j indexes functional annotations, the loss computes symmetrized similarity distributions over negative samples:

$$\mathcal{L}_{SAC} = -\frac{1}{2} \sum_{(i,j)} \left(\log \frac{\exp\left(\frac{\mathrm{h}_i^S \cdot \mathrm{h}_j^A}{\tau}\right)}{\sum_k \exp\left(\frac{\mathrm{h}_i^S \cdot \mathrm{h}_k^A}{\tau}\right)} + \log \frac{\exp\left(\frac{\mathrm{h}_i^S \cdot \mathrm{h}_j^A}{\tau}\right)}{\sum_k \exp\left(\frac{\mathrm{h}_k^S \cdot \mathrm{h}_j^A}{\tau}\right)} \right). \tag{2}$$

Here τ is the temperature hyperparameter scaling similarity magnitudes, and summation indices k traverse randomly sampled negative annotations or sequences. The dual logarithmic terms enforce mutual retrievability constraints: protein sequences should distinguish their true annotations from decoys, while annotations should identify their corresponding sequences.

Direct Preference Optimization (DPO) Loss By parameterizing human preference probabilities through the optimal policy π_θ rather than explicit reward modeling, we derive the Direct Preference Optimization (DPO) objective:

$$\mathcal{L}_{\text{DPO}}(\pi_\theta; \pi_{\text{ref}}) = -\mathbb{E}_{(x,y_w,y_l)\sim\mathcal{D}}\left[\log\sigma\left(\beta\log\frac{\pi_\theta(y_w|x)}{\pi_{\text{ref}}(y_w|x)} - \beta\log\frac{\pi_\theta(y_l|x)}{\pi_{\text{ref}}(y_l|x)}\right)\right], \tag{3}$$

where x denotes input protein sequences, y_w represents ground-truth functional annotations from Swiss-Prot, and y_l corresponds to synthetic negatives. The reference policy π_{ref} preserves knowledge from the supervised fine-tuned model, while the temperature parameter $\beta > 0$ controls deviation from this baseline. The sigmoid function $\sigma(\cdot)$ converts log-probability differences into preference likelihoods.

4. Experiments

Performance Evaluation in Gene Ontology Subcategories We conducted a comprehensive evaluation of model performance across GO subcategories. The testset sequences were stratified by these three ontological categories and evaluated using zero-shot, SFT, and DPO models. Quantitative analysis employing F1-Max and AUPR metrics revealed substantial performance disparities (Fig. 2a). The zero-shot approach demonstrated minimal predictive capability (F1-Max less than 0.1 across all categories), while DPO consistently outperformed SFT, achieving relative F1-Max improvements of 2.7%, 4.1%, and 3.1% in Biological Process (BP), Cellular Component (CC), and Molecular Function (MF) categories respectively.

Long-Tail Distribution Adaptation Analysis To investigate model robustness against label frequency imbalance, we categorized GO terms into three frequency groups: low-frequency ($< 1\%$ occurrence), medium-frequency (1-10%), and high-frequency ($> 10\%$). All models were evaluated on testset (Fig. 2b). DPO exhibited superior performance across all frequency regimes, particularly demonstrating 8.7%, 4.9% and 3.2% F1-Max improvements over SFT in the low, medium and high-frequency categories. This underscores DPO model's enhanced capability in managing rare annotations through its preference optimization framework.

General GO Category Discriminability We visualize single-category GO annotations (BP/CC/MF) from Swiss-Prot-New via t-SNE. Both sequence and annotation features form distinct clusters aligned with biological categories (Fig. 2c). DPO demonstrates clearer separation than other baselines, particularly between molecular functions and cellular components, indicating enhanced ability to distinguish functional categories.

Fine-Grained Ontological Relationship Learning To examine hierarchical relationship capture within GO categories, we selected tightly-related GO term families (e.g., enzyme regulation in MF, protein translation in BP, cytoskeleton in CC) and visualized their sequence embeddings. Fig. 2d demonstrates that DPO-learned features preserve ontological proximity, with related terms forming distinct subclusters. This hierarchical structure awareness enables more biologically meaningful annotation predictions.

Ablation Study Our systematic ablation analysis (Tab. 1) reveals critical architectural contributions to model performance. The zero-shot model shows minimal functionality, while SFT model achieves substantial improvement. The integration of LoRA adapters provides additional gains, demonstrating the effectiveness of parameter-efficient fine-tuning. Our DPO models significantly outperform previous baselines, where DPO model with model-predicted annotations as negatives achieves state-of-the-art performance. Notably, the contrastive learning component proves essential for its removal degrades GO F1-Max by 67.9% compared to full SFT.

Table 1. Ablation study on the model structure.

Model version	F1-Max	Recall	AUROC
Zero Shot	0.0016	0.4687	0.4941
SFT w/o SAC	0.2419	0.0686	0.9358
SFT	0.7533	0.6031	0.9891
SFT LoRA	0.7683	0.6332	0.9915
DPO w/ msk noise	0.7796	0.6192	0.9961
DPO w/ pred	**0.7947**	**0.7027**	**0.9979**

5. Conclusion

In this study, we present a novel framework for protein functional annotation prediction by integrating Direct Preference Optimization into a multimodal learning pipeline. Our method addresses annotation sparsity through two synergistic mechanisms: contrastive alignment between sequence embeddings and GO term features during supervised fine-tuning and direct optimization of human-curated annotation preferences via DPO, circumventing reward modeling complexities. Experimental results demonstrate enhanced discriminability across GO categories compared to conventional approaches, with latent space visualizations revealing clear separation of biological processes, molecular func-

tions, and cellular components. While current performance is constrained by existing annotation biases in Swiss-Prot, this work establishes a paradigm for incorporating evolving functional knowledge through preference-aware learning, enabling adaptive integration of new annotation evidence without architectural modification.

Impact Statement

This paper pioneers the integration of contrastive learning with Direct Preference Optimization to address critical challenges in protein functional annotation: annotation sparsity and cross-modal misalignment. By eliminating reward modeling dependencies and enabling direct optimization of biological preferences, this work accelerates the discovery of uncharacterized protein functions while providing a blueprint for dynamic integration of evolving functional evidence in computational biology. The methodology extends beyond annotation prediction, offering a generalizable paradigm for human-preference-aligned learning in biological sequence analysis.

References

Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., Davis, A. P., Dolinski, K., Dwight, S. S., Eppig, J. T., et al. Gene ontology: tool for the unification of biology. *Nature genetics*, 25(1):25–29, 2000.

Bai, Y., Jones, A., Ndousse, K., Askell, A., Chen, A., DasSarma, N., Drain, D., Fort, S., Ganguli, D., Henighan, T., et al. Training a helpful and harmless assistant with reinforcement learning from human feedback. *arXiv preprint arXiv:2204.05862*, 2022.

Boeckmann, B., Bairoch, A., Apweiler, R., Blatter, M.-C., Estreicher, A., Gasteiger, E., Martin, M. J., Michoud, K., O'Donovan, C., Phan, I., et al. The swiss-prot protein knowledgebase and its supplement trembl in 2003. *Nucleic acids research*, 31(1):365–370, 2003.

Brandes, N., Ofer, D., Peleg, Y., Rappoport, N., and Linial, M. Proteinbert: a universal deep-learning model of protein sequence and function. *Bioinformatics*, 38(8):2102–2110, 2022.

Christiano, P. F., Leike, J., Brown, T., Martic, M., Legg, S., and Amodei, D. Deep reinforcement learning from human preferences. *Advances in neural information processing systems*, 30, 2017.

Consortium, U. Uniprot: a worldwide hub of protein knowledge. *Nucleic acids research*, 47(D1):D506–D515, 2019.

Doğan, T., MacDougall, A., Saidi, R., Poggioli, D., Bateman, A., O'Donovan, C., and Martin, M. J. Uniprot-daac: domain architecture alignment and classification, a new method for automatic functional annotation in uniprotkb. *Bioinformatics*, 32(15):2264–2271, 2016.

Dong, H., Xiong, W., Goyal, D., Zhang, Y., Chow, W., Pan, R., Diao, S., Zhang, J., Shum, K., and Zhang, T. Raft: Reward ranked finetuning for generative foundation model alignment. *arXiv preprint arXiv:2304.06767*, 2023.

Elnaggar, A., Heinzinger, M., Dallago, C., Rehawi, G., Wang, Y., Jones, L., Gibbs, T., Feher, T., Angerer, C., Steinegger, M., et al. Prottrans: Toward understanding the language of life through self-supervised learning. *IEEE transactions on pattern analysis and machine intelligence*, 44(10):7112–7127, 2021.

ESM Team. Esm cambrian: Revealing the mysteries of proteins with unsupervised learning. *EvolutionaryScale Website https://evolutionaryscale.ai/blog/esm-cambrian*, 2024.

Friedberg, I. Automated protein function prediction—the genomic challenge. *Briefings in bioinformatics*, 7(3): 225–242, 2006.

Gasteiger, E., Jung, E., and Bairoch, A. Swiss-prot: connecting biomolecular knowledge via a protein database. *Current issues in molecular biology*, 3(3):47–55, 2001.

Glaese, A., McAleese, N., Trębacz, M., Aslanides, J., Firoiu, V., Ewalds, T., Rauh, M., Weidinger, L., Chadwick, M., Thacker, P., et al. Improving alignment of dialogue agents via targeted human judgements. *arXiv preprint arXiv:2209.14375*, 2022.

Havrilla, A., Raparthy, S., Nalmpantis, C., Dwivedi-Yu, J., Zhuravinskyi, M., Hambro, E., and Raileanu, R. Glore: When, where, and how to improve llm reasoning via global and local refinements. *arXiv preprint arXiv:2402.10963*, 2024.

Jang, Y. J., Qin, Q.-Q., Huang, S.-Y., Peter, A. T. J., Ding, X.-M., and Kornmann, B. Accurate prediction of protein function using statistics-informed graph networks. *Nature Communications*, 15(1):6601, 2024.

Kulmanov, M. and Hoehndorf, R. Deepgoplus: improved protein function prediction from sequence. *Bioinformatics*, 36(2):422–429, 2020.

Kulmanov, M., Khan, M. A., and Hoehndorf, R. Deepgo: predicting protein functions from sequence and interactions using a deep ontology-aware classifier. *Bioinformatics*, 34(4):660–668, 2018.

Kulmanov, M., Guzmán-Vega, F. J., Duek Roggli, P., Lane, L., Arold, S. T., and Hoehndorf, R. Protein function prediction as approximate semantic entailment. *Nature Machine Intelligence*, 6(2):220–228, 2024.

Li, Z., Xu, T., and Yu, Y. Policy optimization in rlhf: The impact of out-of-preference data. *arXiv preprint arXiv:2312.10584*, 2023.

Liu, X., Liu, Y., Chen, S., and Hu, W. Controllable protein sequence generation with llm preference optimization. *arXiv preprint arXiv:2501.15007*, 2025.

Madani, A., McCann, B., Naik, N., Keskar, N. S., Anand, N., Eguchi, R. R., Huang, P.-S., and Socher, R. Progen: Language modeling for protein generation. *arXiv preprint arXiv:2004.03497*, 2020.

Magrane, M. and Consortium, U. Uniprot knowledgebase: a hub of integrated protein data. *Database*, 2011:bar009, 2011.

Meier, J., Rao, R., Verkuil, R., Liu, J., Sercu, T., and Rives, A. Language models enable zero-shot prediction of the effects of mutations on protein function. *Advances in neural information processing systems*, 34:29287–29303, 2021.

Ouyang, L., Wu, J., Jiang, X., Almeida, D., Wainwright, C., Mishkin, P., Zhang, C., Agarwal, S., Slama, K., Ray, A., et al. Training language models to follow instructions with human feedback. *Advances in neural information processing systems*, 35:27730–27744, 2022.

Radivojac, P., Clark, W. T., Oron, T. R., Schnoes, A. M., Wittkop, T., Sokolov, A., Graim, K., Funk, C., Verspoor, K., Ben-Hur, A., et al. A large-scale evaluation of computational protein function prediction. *Nature methods*, 10(3):221–227, 2013.

Rafailov, R., Sharma, A., Mitchell, E., Manning, C. D., Ermon, S., and Finn, C. Direct preference optimization: Your language model is secretly a reward model. *Advances in Neural Information Processing Systems*, 36: 53728–53741, 2023.

Reuter, J. A., Spacek, D. V., and Snyder, M. P. High-throughput sequencing technologies. *Molecular cell*, 58 (4):586–597, 2015.

Rives, A., Meier, J., Sercu, T., Goyal, S., Lin, Z., Liu, J., Guo, D., Ott, M., Zitnick, C. L., Ma, J., et al. Biological structure and function emerge from scaling unsupervised learning to 250 million protein sequences. *Proceedings of the National Academy of Sciences*, 118(15):e2016239118, 2021.

Schulman, J., Wolski, F., Dhariwal, P., Radford, A., and Klimov, O. Proximal policy optimization algorithms. *arXiv preprint arXiv:1707.06347*, 2017.

Setlur, A., Nagpal, C., Fisch, A., Geng, X., Eisenstein, J., Agarwal, R., Agarwal, A., Berant, J., and Kumar, A. Rewarding progress: Scaling automated process verifiers for llm reasoning. *arXiv preprint arXiv:2410.08146*, 2024.

Śledź, P. and Jinek, M. Structural insights into the molecular mechanism of the m6a writer complex. *elife*, 5:e18434, 2016.

Song, F., Yu, B., Li, M., Yu, H., Huang, F., Li, Y., and Wang, H. Preference ranking optimization for human alignment. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 38, pp. 18990–18998, 2024.

Stiennon, N., Ouyang, L., Wu, J., Ziegler, D., Lowe, R., Voss, C., Radford, A., Amodei, D., and Christiano, P. F. Learning to summarize with human feedback. *Advances in neural information processing systems*, 33:3008–3021, 2020.

Stocco, F., Artigues-Lleixa, M., Hunklinger, A., Widatalla, T., Guell, M., and Ferruz, N. Guiding generative protein language models with reinforcement learning. *arXiv preprint arXiv:2412.12979*, 2024.

Su, J., Han, C., Zhou, Y., Shan, J., Zhou, X., and Yuan, F. Saprot: Protein language modeling with structure-aware vocabulary. *bioRxiv*, pp. 2023–10, 2023.

Suzek, B. E., Wang, Y., Huang, H., McGarvey, P. B., Wu, C. H., and Consortium, U. Uniref clusters: a comprehensive and scalable alternative for improving sequence similarity searches. *Bioinformatics*, 31(6):926–932, 2015.

Tipton, K. and Boyce, S. History of the enzyme nomenclature system. *Bioinformatics*, 16(1):34–40, 2000.

Widatalla, T., Rafailov, R., and Hie, B. Aligning protein generative models with experimental fitness via direct preference optimization. *bioRxiv*, pp. 2024–05, 2024.

Xu, M., Yuan, X., Miret, S., and Tang, J. Protst: Multimodality learning of protein sequences and biomedical texts. In *International Conference on Machine Learning*, pp. 38749–38767. PMLR, 2023.

Xu, S., Fu, W., Gao, J., Ye, W., Liu, W., Mei, Z., Wang, G., Yu, C., and Wu, Y. Is dpo superior to ppo for llm alignment? a comprehensive study. *arXiv preprint arXiv:2404.10719*, 2024.

You, R., Yao, S., Mamitsuka, H., and Zhu, S. Deepgraphgo: graph neural network for large-scale, multispecies protein function prediction. *Bioinformatics*, 37(Supplement_1): i262–i271, 2021.

Yu, T., Cui, H., Li, J. C., Luo, Y., Jiang, G., and Zhao, H. Enzyme function prediction using contrastive learning. *Science*, 379(6639):1358–1363, 2023.

Yuan, Z., Yuan, H., Tan, C., Wang, W., Huang, S., and Huang, F. Rrhf: Rank responses to align language models with human feedback without tears. *arXiv preprint arXiv:2304.05302*, 2023.

Zhang, N., Bi, Z., Liang, X., Cheng, S., Hong, H., Deng, S., Lian, J., Zhang, Q., and Chen, H. Ontoprotein: Protein pretraining with gene ontology embedding. *arXiv preprint arXiv:2201.11147*, 2022.

Zhou, X., Han, C., Zhang, Y., Su, J., Zhuang, K., Jiang, S., Yuan, Z., Zheng, W., Dai, F., Zhou, Y., et al. Decoding the molecular language of proteins with evolla. *bioRxiv*, pp. 2025–01, 2025.

Ziegler, D. M., Stiennon, N., Wu, J., Brown, T. B., Radford, A., Amodei, D., Christiano, P., and Irving, G. Fine-tuning language models from human preferences. *arXiv preprint arXiv:1909.08593*, 2019.

A. Model and Training Details

Table 2. Model architecture hyperparameters

Parameter	Value
Sequence Length	512
Annotation Classes	7533
Annotation Encoder Attention Head Dimension	64
Annotation Encoder Attention Heads	8
Annotation Encoder Depth	12
Annotation Encoder Hidden Dimension	960
Annotation Encoder Global Dimension	512
Annotation Predictor Dropout Rate	0.1
Annotation Predictor Residual Blocks	2

Table 3. SFT hyperparameters

Parameter	Value
Batch Size per GPU	128
Base Learning Rate	5e-5
Minimum Learning Rate	5e-7
Warmup Initial Learning Rate	5e-7
Warmup Epochs	3
Finetuning Epochs	80
Learning Rate Decay Rate	0.95

Table 4. DPO hyperparameters

Parameter	Value
Batch Size per GPU	48
DPO Beta	0.1
Number of Augmentations	3-10
Training Weight	0.01-1.0
DPO Loss Weight	0.01-1.0
KL Divergence Weight	0.1-1.0
NLL Loss Weight	0.01-100
Diversity Loss Weight	1.0
SAC Loss Weight	1.0
Alpha Balance Factor	1.0
Warmup Steps	1% of total steps
DPO Total Epochs	20
Base Learning Rate	5e-5
Minimum Learning Rate	5e-7
Warmup Learning Rate	5e-7

B. Dataset Details

Table 5. Classification of GO terms by functional category and annotation frequency.

Classification	Amount
Function	
CC (Cellular Component)	962
BP (Biological Process)	3346
MF (Molecular Function)	3225
Total	7533
Frequency	
Low	4120
Medium	2680
High	733
Total	7533

Table 6. Dataset sequence counts with annotation inclusion criteria: training set totals, test set for any GO category occurrence, and Swiss-Prot-New for exclusive single-category GO terms and frequency-based counts.

Dataset	Classification	Amount
Training Set	Total	483,285
Test Set	*Function*	
	BP	43,740
	CC	41,225
	MF	46,770
	Total	53,563
Swiss-Prot-New	*Frequency*	
	Low (<1%)	2,501
	Medium (1%∼10%)	15,625
	High (>10%)	37,148
	Function	
	BP	385
	CC	1,829
	MF	1,970
	Total	37,972

C. Experiment Details

Table 7. Quantitative performance metrics across GO subcategories.

Model version	BP					CC					MF				
	Recall	Precision	F1-Max	AUROC	AUPR	Recall	Precision	F1-Max	AUROC	AUPR	Recall	Precision	F1-Max	AUROC	AUPR
Zero Shot	0.4599	0.0006	0.0013	0.4752	0.0006	0.3597	0.0014	0.0044	0.4430	0.0016	0.5571	0.0008	0.0019	0.5450	0.0008
SFT	0.5381	0.9575	0.7155	0.9875	0.7205	0.5579	0.8990	0.7124	0.9933	0.7565	0.7663	0.9591	0.8604	0.9938	0.8986
DPO	0.6075	0.9171	0.7345	0.9958	0.7488	0.6344	0.8742	0.7418	0.9975	0.7977	0.8329	0.9481	0.8870	0.9992	0.9294

Table 8. Quantitative performance of robustness evaluation across annotation frequency groups.

Model version	GO F1-Max			GO Recall			GO AUROC		
	Low	**Medium**	**High**	**Low**	**Medium**	**High**	**Low**	**Medium**	**High**
Zero shot	0.0026	0.0017	0.0015	0.4843	0.4609	0.4549	0.4956	0.4877	0.4909
SFT	0.4224	0.5710	0.6238	0.2197	0.4275	0.5110	0.9464	0.9639	0.9651
DPO	0.4591	0.5992	0.6439	0.3114	0.4985	0.5620	0.9708	0.9783	0.9773

D. Additional Experiment Results

Table 9. Additional GOs of biological process.

Category	Subcategory	Sequence amount	GO ID	Term
Biological Process	Apoptosis	56	GO:0006915	Apoptotic process
			GO:2001235	Positive regulation of apoptotic signaling pathway
			GO:0043027	Cysteine-type endopeptidase inhibitor activity involved in apoptotic process
	Cell Cycle Regulation	11	GO:0007050	Regulation of cell cycle
			GO:2000045	Regulation of G1/S transition of mitotic cell cycle
			GO:0007049	Cell cycle
			GO:0070192	Chromosome organization involved in meiosis
			GO:0007142	Male meiosis II
	Cell Differentiation	30	GO:0048741	Skeletal muscle fiber development
			GO:0021954	Central nervous system neuron development
			GO:0048513	Animal organ development
			GO:0048666	Neuron development
			GO:0045595	Regulation of cell differentiation
			GO:0060173	Limb development
	DNA Replication Repair	32	GO:0006271	DNA strand elongation involved in DNA replication
			GO:0032297	Negative regulation of DNA-templated DNA replication initiation
			GO:0071897	DNA biosynthetic process
			GO:0006290	Pyrimidine dimer repair
			GO:0006267	Pre-replicative complex assembly involved in nuclear cell cycle DNA replication
	Metabolism	300	GO:0006739	NADP metabolic process
			GO:0006644	Phospholipid metabolic process
			GO:0016042	Lipid catabolic process
			GO:0019563	Glycerol catabolic process
			GO:0006083	Acetate metabolic process
	Protein Modification	24	GO:0031398	Positive regulation of protein ubiquitination
			GO:0035871	Protein K11-linked deubiquitination
			GO:0071569	Protein ufmylation
			GO:0001934	Positive regulation of protein phosphorylation
			GO:0035307	Positive regulation of protein dephosphorylation
			GO:0031146	SCF-dependent proteasomal ubiquitin-dependent protein catabolic process
	Protein Translation	35	GO:0002183	Cytoplasmic translational initiation
			GO:0006415	Translational termination
			GO:0045900	Negative regulation of translational elongation
			GO:0002182	Cytoplasmic translational elongation
	RNA Processing	79	GO:0031167	rRNA methylation
			GO:0000288	Nuclear-transcribed mRNA catabolic process, deadenylation-dependent decay
			GO:0000967	rRNA 5'-end processing
			GO:0006406	mRNA export from nucleus
			GO:0000956	Nuclear-transcribed mRNA catabolic process
	Signaling	8	GO:0038166	Angiotensin-activated signaling pathway
			GO:0007259	Cell surface receptor signaling pathway via JAK-STAT
			GO:0033209	Tumor necrosis factor-mediated signaling pathway
			GO:0030520	Estrogen receptor signaling pathway
			GO:0010469	Regulation of signaling receptor activity

Table 10. Additional GOs of cellular component.

Category	Subcategory	Sequence amount	GO ID	Term
Cellular Component	Chromatin Nucleosome	145	GO:0005721	Pericentric heterochromatin
			GO:0000779	Condensed chromosome, centromeric region
			GO:0000792	Heterochromatin
			GO:0031519	PcG protein complex
			GO:0005694	Chromosome
	Chromosome-related	35	GO:0000922	Spindle pole
			GO:0000940	Outer kinetochore
			GO:1990879	CST complex
			GO:0000930	Gamma-tubulin complex
			GO:0035371	Microtubule plus-end
	Cytoskeleton	136	GO:0005925	Focal adhesion
			GO:0005912	Adherens junction
			GO:0070161	Anchoring junction
			GO:0097431	Mitotic spindle pole
			GO:0036064	Ciliary basal body
			GO:0036157	Outer dynein arm
			GO:0001534	Radial spoke
	ER-Golgi	592	GO:0005789	Endoplasmic reticulum membrane
			GO:0090158	Endoplasmic reticulum membrane organization
			GO:0005784	Sec61 translocon complex
			GO:0005802	Trans-Golgi network
	Membrane Complexes	66	GO:0009897	External side of plasma membrane
			GO:0031241	Periplasmic side of cell outer membrane
			GO:0098982	GABA-ergic synapse
			GO:0045211	Postsynaptic membrane
			GO:0005921	Gap junction
			GO:0005922	Connexin complex
			GO:0034707	Chloride channel complex
			GO:0030867	Rough endoplasmic reticulum membrane
	Mitochondrial	146	GO:0005759	Mitochondrial matrix
			GO:0005744	Mitochondrial inner membrane presequence translocase complex
			GO:0030964	NADH dehydrogenase complex
			GO:0070469	Respiratory chain
			GO:0042645	Mitochondrial nucleoid
			GO:0005761	Mitochondrial ribosome
	Nuclear Membrane Pore	52	GO:0031965	Nuclear membrane
			GO:0071765	Nuclear inner membrane organization
			GO:0031080	Nuclear pore complex
	Protein Degradation	16	GO:0000151	Ubiquitin ligase complex
			GO:0019005	SCF ubiquitin ligase complex
			GO:0031464	Cul4-RING E3 ubiquitin ligase complex
	RNA Processing Complexes	48	GO:0005681	Spliceosomal complex
			GO:0071006	U2-type catalytic step 1 spliceosome
			GO:0071007	U2-type catalytic step 2 spliceosome
			GO:0089701	U2 snRNP
			GO:0005685	U1 snRNP
			GO:0005849	mRNA cleavage factor complex
	Transcription Complexes	731	GO:0005666	RNA polymerase III complex
			GO:0000428	DNA-directed RNA polymerase complex
			GO:0016580	Sin3 complex
			GO:0016592	Mediator complex
			GO:0030880	RNA polymerase complex
			GO:0005673	Transcription factor TFIIE complex
			GO:0016586	RSC-type complex
			GO:0032783	Super elongation complex
			GO:0090575	RNA polymerase II transcription factor complex
			GO:0000118	Histone deacetylase complex

Table 11. Additional GOs of molecular function.

Category	Subcategory	Sequence amount	GO ID	Term
Molecular Function	Hydrolase Activity	141	GO:0016798	Hydrolase activity, acting on glycosyl bonds
			GO:0070004	Cysteine-type exopeptidase activity
			GO:0008234	Cysteine-type peptidase activity
			GO:0004045	Aminoacyl-tRNA hydrolase activity
			GO:0016920	Pyroglutamyl-peptidase activity
			GO:0004843	Thiol-dependent deubiquitinase activity
	Transferase Activity	54	GO:0016765	Transferase activity, transferring alkyl or aryl groups
			GO:0008318	Protein prenyltransferase activity
			GO:0004057	Arginyl-tRNA–protein transferase activity
			GO:0015019	Heparan-alpha-glucosaminide N-acetyltransferase activity
			GO:0008791	Arginine N-succinyltransferase activity
			GO:0047173	Phosphatidylcholine-retinol O-acyltransferase activity
	Oxidoreductase Activity	9	GO:0016714	Oxidoreductase activity, acting on paired donors
			GO:0004174	Electron-transferring-flavoprotein dehydrogenase activity
			GO:0004471	Malate dehydrogenase (decarboxylating) (NAD+) activity
			GO:0047111	Formate dehydrogenase (cytochrome-c-553) activity
			GO:0004665	Prephenate dehydrogenase (NADP+) activity
			GO:0046553	D-malate dehydrogenase (decarboxylating) (NAD+) activity
			GO:0003834	Beta-carotene 15,15'-dioxygenase activity
			GO:0016630	Protochlorophyllide reductase activity
	Kinase Phosphatase Activity	18	GO:0106311	Protein serine/threonine kinase activity
			GO:0004797	Thymidine kinase activity
			GO:0004703	G protein-coupled receptor kinase activity
			GO:0008673	2-dehydro-3-deoxygluconokinase activity
			GO:0004331	Fructose-2,6-bisphosphate 2-phosphatase activity
	Ion Transport	33	GO:0015087	Cobalt ion transmembrane transporter activity
			GO:0008324	Cation transmembrane transporter activity
			GO:0005221	Intracellularly cyclic nucleotide-activated monoatomic cation channel activity
			GO:0005223	Intracellularly cGMP-activated cation channel activity
			GO:0008308	Voltage-gated monoatomic anion channel activity
			GO:0015444	P-type magnesium transporter activity
	Organic Molecule Transport	13	GO:0015187	Glycine transmembrane transporter activity
			GO:0015181	L-arginine transmembrane transporter activity
			GO:0005324	Long-chain fatty acid transmembrane transporter activity
			GO:0015221	Lipopolysaccharide transmembrane transporter activity
			GO:0090482	Vitamin transmembrane transporter activity
			GO:0015655	Alanine:sodium symporter activity
			GO:0015189	L-lysine transmembrane transporter activity
	DNA Binding	30	GO:1990837	Sequence-specific double-stranded DNA binding
			GO:0000986	Cis-regulatory region sequence-specific DNA binding
			GO:0000404	Heteroduplex DNA loop binding
			GO:0043138	3'-5' DNA helicase activity
			GO:1990970	Trans-activation response element binding
	RNA Binding	15	GO:0008143	Poly(A) binding
			GO:0045131	Pre-mRNA branch point binding
			GO:0001070	RNA binding transcription factor activity
			GO:0030619	U1 snRNA binding
			GO:0033897	Ribonuclease T2 activity
	Signal Protein Binding	75	GO:0005132	Type I interferon receptor binding
			GO:0005164	Tumor necrosis factor receptor binding
			GO:0008190	Eukaryotic initiation factor 4E binding
			GO:0031072	Heat shock protein binding
			GO:0008013	Beta-catenin binding
			GO:0044325	Ion channel binding
	Enzyme Regulation	355	GO:0005096	GTPase activator activity
			GO:0004864	Protein phosphatase inhibitor activity
			GO:0030234	Enzyme regulator activity
			GO:0043022	Ribosome binding
			GO:0010521	Telomerase inhibitor activity
			GO:0030337	DNA polymerase processivity factor activity

Figure 3. Additional results of sequences with biological process related GOs in fine-grained ontological relationship learning task.

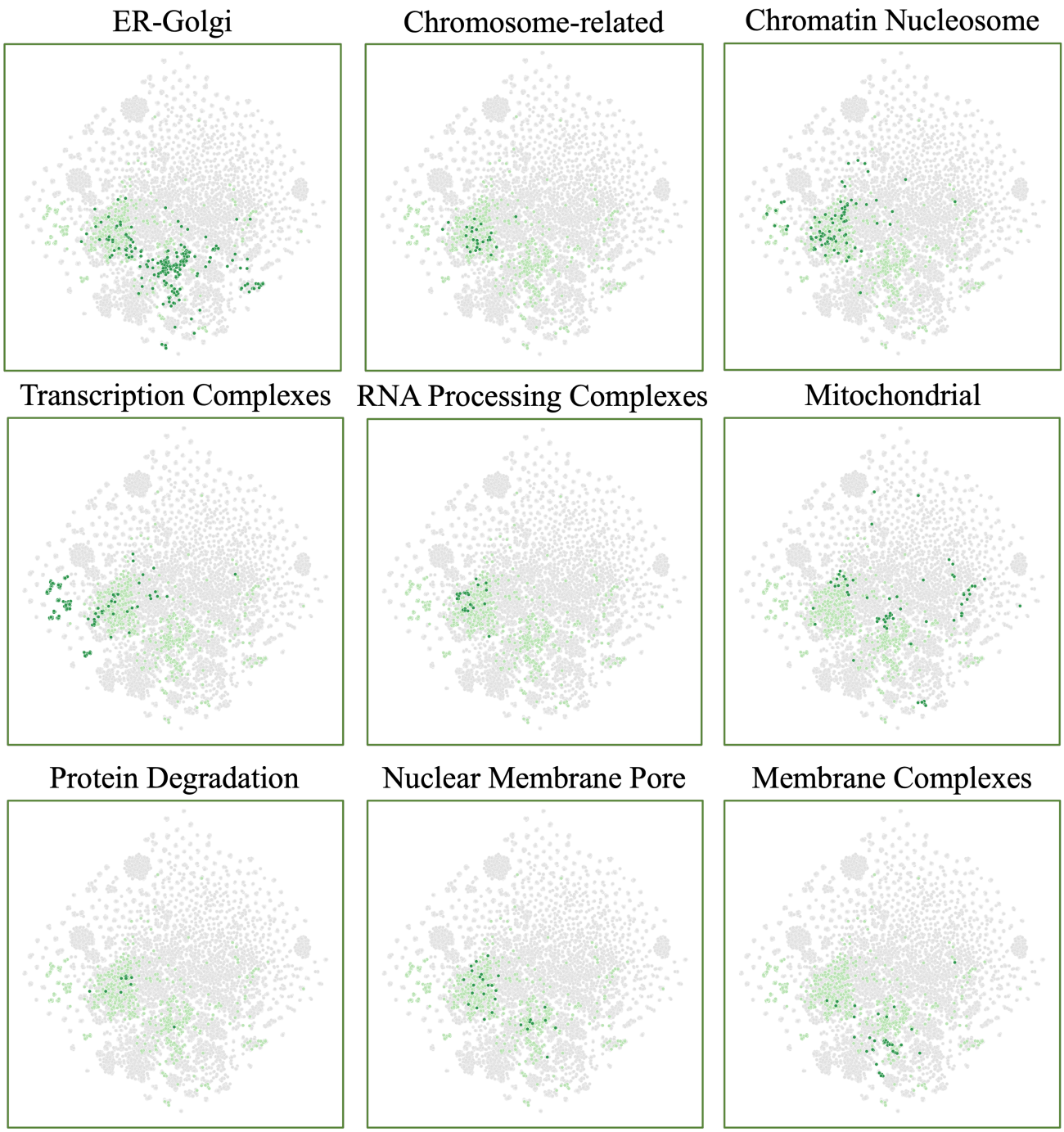

Figure 4. Additional results of sequences with cellular component related GOs in fine-grained ontological relationship learning task.

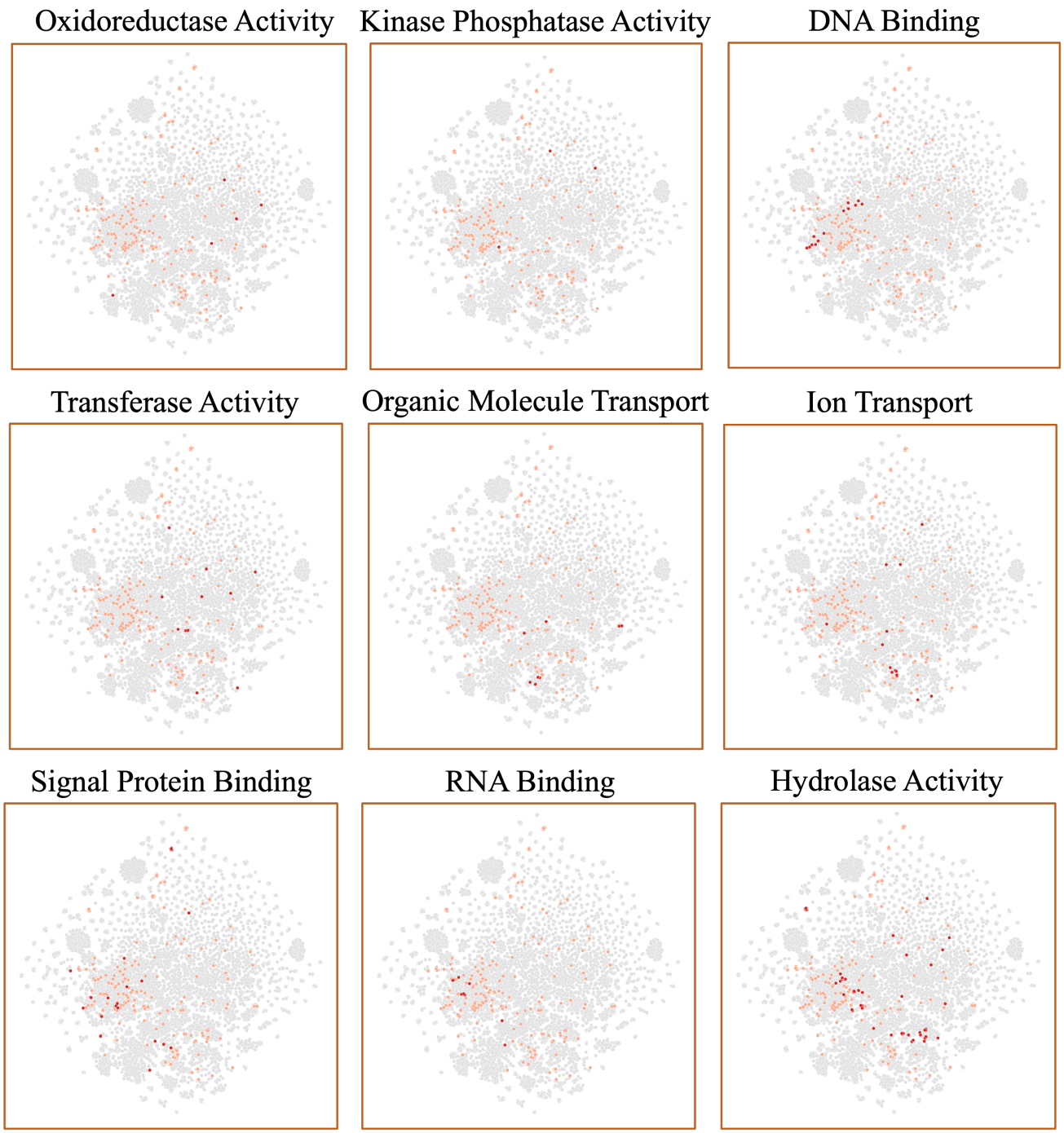

Figure 5. Additional results of sequences with molecular function related GOs in fine-grained ontological relationship learning task.