Protein Sequence Design in a Latent Space via Model-based Reinforcement Learning

Minji Lee[O,1] Luiz Felipe Vecchietti[O,3] Hyunkyu Jung[1,3,4] Hyunjoo Ro[4] Meeyoung Cha[3,1] Ho Min Kim[2,4]

[1]School of Computing, KAIST [2]Graduate School of Medical Science and Engineering, KAIST
[3]Data Science Group, Institute for Basic Science [4]Center for Biomolecular and Cellular Structure, Institute for Basic Science

1. Introduction

Protein sequence design

- Enhancing the functionality of a protein
- Enhancing the cellular fitness of an organism
- Directed evolution, data driven methods

Challenges of protein sequence design

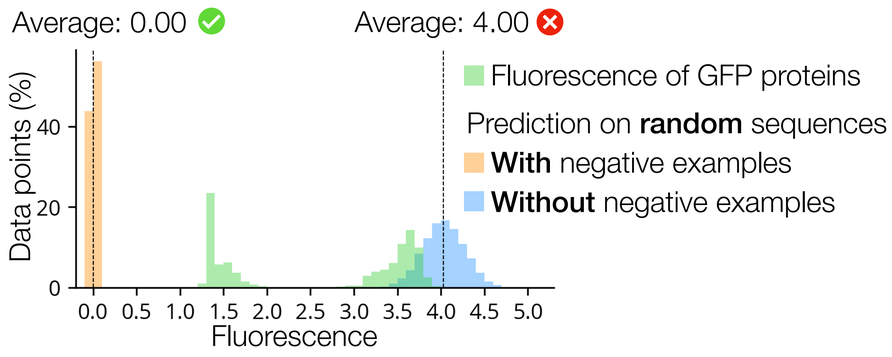
- Vast search space
- Non-functional >> Functional sequences

Data-driven methods

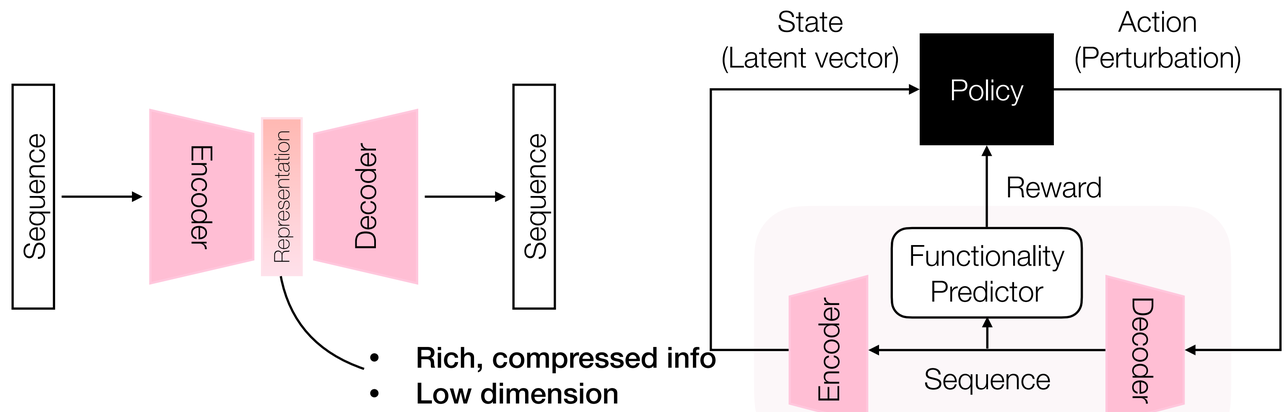
- Reinforcement learning [1]
- Bayesian optimization [2,3,4]
- Generative models [5]

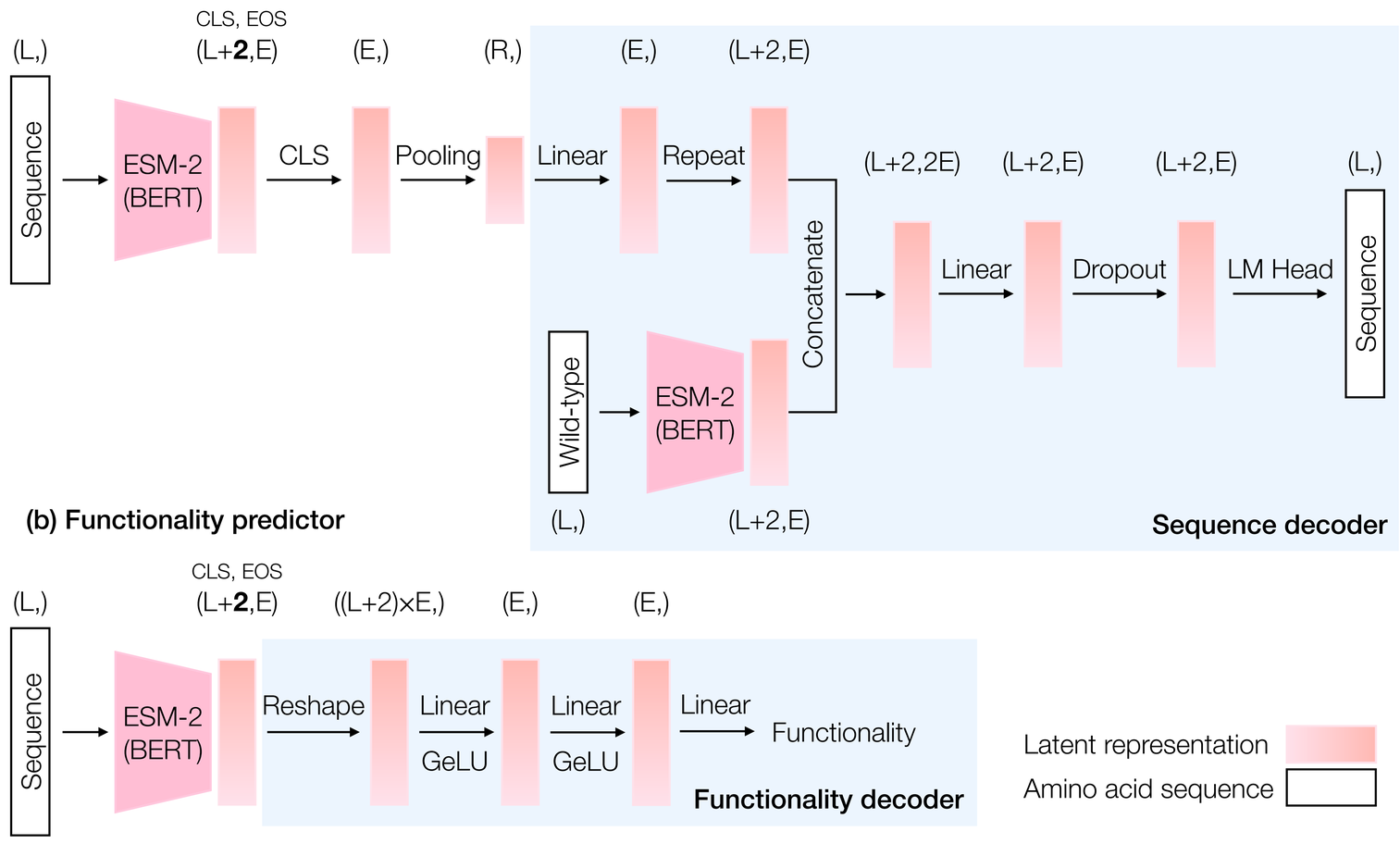
Still challenging to generate optimized sequences that are experimentally validated. Why?

1. **Inefficiency.** Optimization as amino acid addition/mutation.
2. **Oracle** trained without negative examples.



We model protein sequence design as a Markov Decision Process (MDP) to optimize the latent representation by learned perturbation



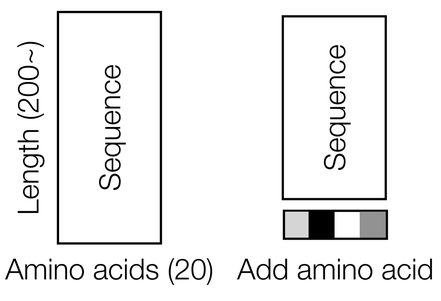


Baseline (learned mutation)

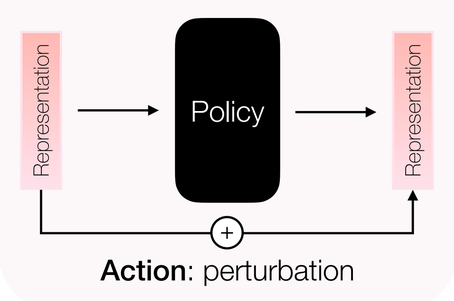
State: sequence **Action**: mutation

Previous approach (learned addition)

State: sequence **Action**: addition



Our approach



2. Methodology

Sequence encoder

1. Pre-trained protein language model encoder is used to obtain latent embeddings
2. Dimensionality reduction → Used as a state of RL agent

Sequence decoder

1. Recover embeddings from reduced representation
2. Pre-trained decoder head to recover sequence

Protein functionality prediction

- Predict functionality from sequence
- Use pre-trained protein language model as a backbone
- Optimization oracle (reward) and evaluation oracle trained separately to prevent information leakage

Model-based reinforcement learning

Trains a policy using an off-policy RL algorithm that models reward function based on the functionality predictor

3. Experimental setup

Datasets: 2 proteins with different length and function

- Green fluorescent protein (GFP) [6]
- Imidazoleglycerol-phosphate dehydratase (His3) [7]

Evaluation

1. Optimize 100 mutants of the protein
2. Evaluate top-10 sequences

4. Results

Model	Performance	Novelty	Original	dist(WT)	Diversity	Chromophore
Ours	**3.491** ± 0.352	8.451	100%	7.700	6.311	100%
Directed evolution	3.287 ± 0.237	7.704	-	6.849	4.858	100%
CbAS	3.155 ± 0.153	7.712	80%	6.900	1.956	100%
Random-1	2.824 ± 0.100	6.611	80%	7.186	7.716	100%
Random-5	2.280 ± 0.275	13.91	100%	9.950	12.37	90%
Random-P	1.511 ± 0.797	14.71	100%	14.15	14.62	100%
BO	0.581 ± 0.095	36.96	100%	36.70	6.867	100%
DynaPPO	0.004 ± 0.003	218.9	100%	219.3	224.1	0%
GFlowNet	0.000 ± 0.002	199.4	100%	200.1	12.53	0%

Model	Performance	Novelty	Original	dist(WT)	Diversity
Ours	**0.945** ± 0.091	8.361	60%	10.95	3.521
CbAS	0.749 ± 0.157	7.287	90%	4.700	2.356
Random-1	0.858 ± 0.058	7.372	80%	7.350	7.716
Random-5	0.678 ± 0.096	9.777	100%	8.950	12.37
Directed evolution	0.616 ± 0.110	6.889	-	6.710	6.942
DynaPPO	-0.201 ± 0.142	27.41	100%	26.70	27.47
BO	-0.313 ± 0.065	26.17	100%	27.50	4.756

5. Ablation studies

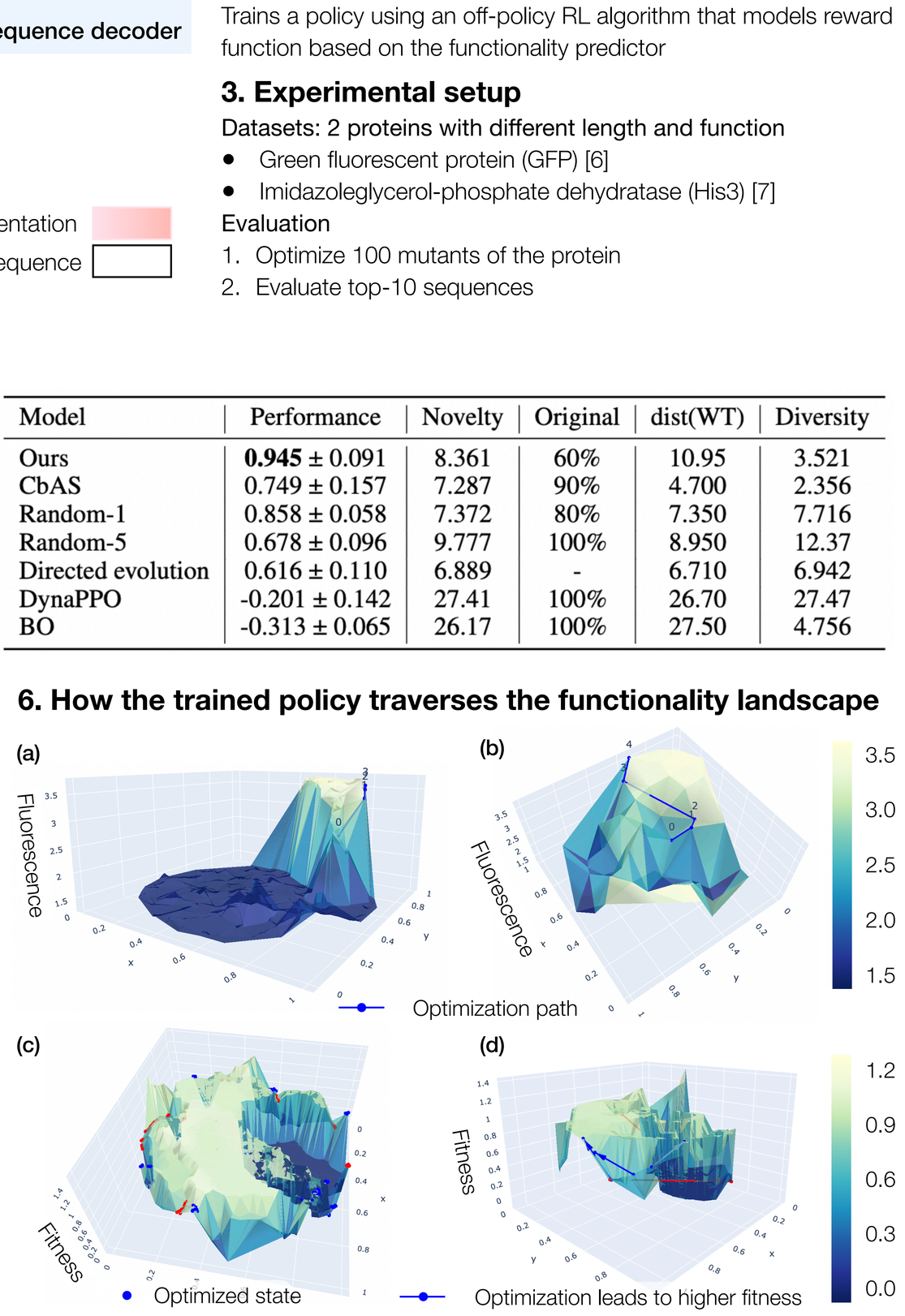
State and action modeling

Model	Performance	Novelty	Diversity	Chromophore
Ours	**3.491** ± 0.352	8.451 ± 2.05	6.311	100%
Directed evolution	3.287 ± 0.237	7.704 ± 2.66	4.858	100%
Swersky et al. (2020) on latent space	2.601 ± 0.912	8.077 ± 2.58	6.600	100%
Random perturbation	1.511 ± 0.797	14.71 ± 5.90	14.616	100%

Representation analysis

State	Action	GFP	His3
Latent vector	Perturbation on latent vector	3.491 ± 0.352	0.945 ± 0.091
Directed evolution		3.287 ± 0.237	0.616 ± 0.110
Sequence	Generate sequence	0.006 ± 0.004	-0.148 ± 0.043
Latent vector	Generate sequence	0.005 ± 0.003	-0.139 ± 0.144
Sequence	Amino acid addition	0.004 ± 0.003	-0.201 ± 0.142

6. How the trained policy traverses the functionality landscape



[1] Angermueller, Christof, et al. "Model-based reinforcement learning for biological sequence design." ICLR (2019)
[2] Belanger, David, et al. "Biological Sequences Design using Batched Bayesian Optimization." (2019)
[3] Stanton, Samuel, et al. "Accelerating Bayesian Optimization for Biological Sequence Design with Denoising Autoencoders." (2022).
[4] Brookes, David, et al. "Conditioning by adaptive sampling for robust design." ICML (2019)
[5] Jain, Moksh, et al. "Biological sequence design with gflownets." ICML (2022)
[6] Sarkisyan, Karen S., et al. "Local fitness landscape of the green fluorescent protein." Nature (2016)
[7] Pokusaeva, Victoria O., et al. "An experimental assay of the interactions of amino acids from orthologous sequences shaping a complex fitness landscape." PLoS genetics (2019)