ProteinRL: Reinforcement learning with generative protein language models for property-directed sequence design

Matt C. Sternke
Protein Design and Informatics
GSK
matt.c.sternke@gsk.com

Joel D. Karpiak
Protein Design and Informatics
GSK
joel.x.karpiak@gsk.com

Abstract

The overarching goal of protein engineering is the design and optimization of proteins customized for specific purposes. Generative protein language models (PLMs) allow for *de novo* protein sequence generation, however current PLMs lack capabilities for controllable sequence generation of sequences tailored with desired properties. Here we present ProteinRL, a flexible, data-driven reinforcement learning framework for fine-tuning generative PLMs for the *de novo* design of sequences optimized for specific sequence and/or structural properties. We highlight two example cases of realistic protein design goals: a single-objective design for sequences containing unusually high charge content, and a multi-objective design scenario of a hit expansion, diversifying a target sequence with generated sequences having high-confidence structure predictions and high probability predictions of soluble expression. In both cases ProteinRL fine-tuning guides the PLM towards generating sequences optimized for the defined properties, extending to values rarely or never seen in natural sequences or sequences generated without ProteinRL fine-tuning. The demonstrated success and adaptability of the ProteinRL framework allows for the *de novo* design of novel protein sequences optimized for applications across many areas of protein engineering.

1 Introduction

Over the past few years, developments in generative protein language models (PLMs) have led to major advancements in *de novo* protein sequence design. Trained on large databases of natural protein sequences such as UniProt [1], generative PLMs can learn the statistical patterns of amino acids found in natural proteins. Once trained, generative PLMs can be sampled to generate novel *de novo*-designed protein sequences that match those amino acid patterns of natural sequences [2]. While generative PLMs have been developed using various deep learning model architectures [3–6], recent developments in transformer-based PLMs have shown particularly promising success in protein sequence generation capabilities [2, 7–10]. Protein sequences generated from a recent transformer-based generative PLM were found to express, fold into their predicted globular structures, and maintain biological activity [7].

One important remaining challenge with current generative PLMs is the controllable generation of sequences tailored or optimized for specific properties. One approach developed in this direction is training PLMs that couple training sequences to control codes that condition sequences on functional annotations such as Pfam family annotations, Gene Ontology terms, or Enzyme Commission numbers [7, 11]. While such strategies allow for tailoring sequence generations towards specific biological functions, they are limited to properties for which these annotations exist and cannot be adapted to properties that the foundational model was not trained on.

NeurIPS 2023 Generative AI and Biology Workshop.

Outside of the field of protein design, reinforcement learning (RL) has been used as a strategy for task-directed tuning of language models. Famously, OpenAI's ChatGPT was trained to increase human-like quality of text generation by fine-tuning a pre-trained GPT-3.5 model using human feedback as a reward. In the field of chemistry, researchers developed the REINVENT framework to fine-tune a pre-trained recurrent neural network (RNN) chemical language model to generate molecules with desired properties [12]. While RL approaches have been used in other areas protein design [13–15], no current approaches design full protein sequences for specific properties by fine-tuning a generative PLM.

Towards the goal of custom-tailored *de novo* sequence design we developed ProteinRL, a policy-based reinforcement learning approach to fine-tune generative PLMs for the generation of protein sequences that possess specified desired properties. To highlight these capabilities, we show that an agent PLM learns to generate sequences that contain high degrees of either positive or negative net charge. Sequences generated as a result of ProteinRL fine-tuning show net charges at very extreme ends of the distribution of net charges observed among natural sequence and are predicted to maintain the natural protein fold. A multi-objective scoring function allows for simultaneous optimization of multiple protein properties. We demonstrate multi-property optimization capabilities of ProteinRL with an example use case of a hit expansion, where we simultaneously optimize generated sequences for high sequence identity to a specific target sequence, high confidence of structural models, and high predicted probability of soluble expression. Again, ProteinRL guides the PLM to generate sequences with values for all three properties that fewer than 0.2% of natural sequences possess. Our approach allows for the *de novo* design of novel protein sequences optimized for specific properties, with applications across many areas of protein engineering.

2 Methods

2.1 ProteinRL property-directed fine-tuning

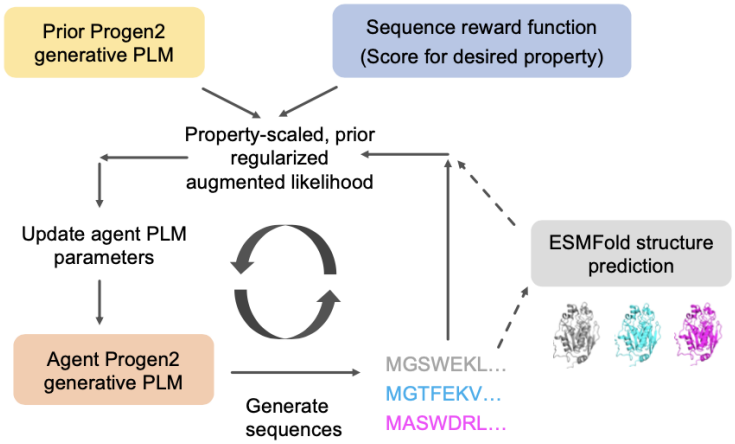

Figure 1: ProteinRL workflow. During training, the agent PLM generates sequences. Structures for generated sequences can be predicted using ESMFold. User-defined sequence and/or structural properties of interest are calculated for generated sequences and used to train agent PLM towards generating sequences with these properties. An untrained, prior PLM regularizes the agent PLM to mitigate catastrophic forgetting.

The workflow for the RL approach is outlined in Figure 1). We adapted the REINVENT [12] RL approach for fine-tuning a pre-trained generative PLM using a reward score calculated for a desired property. At the outset of fine-tuning, two pre-trained PLMs are instantiated: an agent PLM that is fine-tuned by RL and a prior PLM that is not trained, but used to regularize the agent PLM to preserve knowledge of the initial protein sequences and mitigate catastrophic forgetting. We use the pre-trained ProGen2 generative PLM [8] as the foundational PLM for ProteinRL fine-tuning, however we note that any pre-trained generative PLM will work within the ProteinRL framework. During training, the agent PLM performs the episodic task of generating sequences by iterative next residue prediction until an end-of-sequence token is reached. A sequence reward function scores the

generated sequences for the user-specified desired property or properties of interest. If any property for optimization is determined from protein structure, structural models of the generated sequences are determined using the protein structure prediction model ESMFold [16]. Again we note that any protein structure prediction model will work within the ProteinRL framework. A property-scaled, prior-regularized loss function is determined from the generated sequences as:

$$L(\theta) = [\log P_{Agent}(seq) - \log P_{Augmented}(seq)]^2. \tag{1}$$

In Equation 1, $\log P_{Agent}(seq)$ is the sequence log likelihood as determined by the agent PLM, given by:

$$\log P_{Agent}(seq) = \sum_{i=1}^{l} \log P(x_i|x_{<i}) \tag{2}$$

where x_i denotes the residue x at position i over the entire protein sequence length l and $x_{<i}$ indicates the entire sequence preceding position i. The $\log P_{Augmented}(seq)$ term is an augmented log likelihood of a sequence given by:

$$\log P_{Augmented}(seq) = \log P_{Prior}(seq) + \sigma * \phi(R(seq)) \tag{3}$$

composed of two terms. The first term is the log likelihood of the sequence given by the prior model as in Equation 2. The second term is the score determined from the specified sequence reward function $R(seq)$, transformed to the interval [0,1] by transformation function ϕ (where property scores that are favorable are greater), and scaled by a user-defined scaling factor σ that modulates the contribution of the property reward score to the overall loss. Such a loss function guides the agent PLM towards sequence generations that score well by the property reward score function while maintaining knowledge of sequences that are favorable according to the prior PLM.

At each step of the ProteinRL fine-tuning loop, sequences were generated by the agent PLM, scored by the desired sequence property reward function, and the mean loss (by Equation 2) among sequences backpropagated through the agent PLM. Over repeated iterations for fine-tuning, the agent PLM is guided towards sequence generations that score well by the property reward score function(s) but remain favorable by the prior PLM, resulting in quality sequences that are optimized for the specific property or properties of interest.

Further details of ProteinRL finetuning for the two example protein design tasks highlighted below are given in the Supplementary Methods.

3 Results

3.1 Optimization for high net charge content

To demonstrate the ability of ProteinRL to optimize sequences for specific defined properties, we fine-tuned generative PLMs to generate sequences with highly positive or highly negative charge content (Figure 2) The design of proteins with high charge content represents a realistic protein engineering goal, as supercharging proteins has been identified as a strategy to impart high stability [17]. We used β-lactamase as the example protein family for this task.

Throughout the process of fine-tuning, the PLMs learn to generate sequences with extreme positive or negative net charge. The net charges of sequences generated by the PLMs fine-tuned by ProteinRL reach mean values of +18 and -39 for the PLMs optimizing positive and negative net charge content respectively (Figure 2). These values of sequence net charges are at the extreme ends of the distribution of net charges observed among natural β-lactamase sequences (mean net charge of -2, Figure 2), and are more extreme than those observed in 1,000 sequences generated from sampling the prior PLM model that is not fine-tuned towards increasing or decreasing net charge (mean net charge of -3; Figure 2). Importantly, fine-tuning towards extreme charge content does not come at the cost of generating sequences that maintain important sequence and structural features of β-lactamase sequences. Sequences generated by the fine-tuned PLMs have low perplexities as

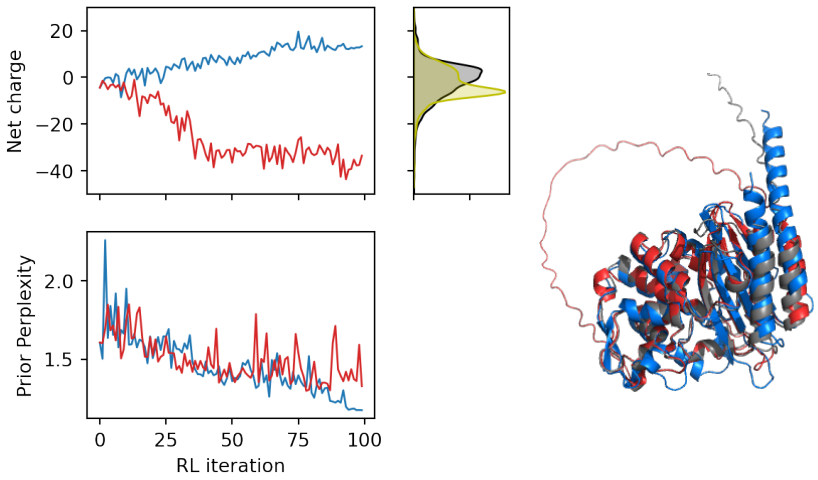

Figure 2: Fine-tuning sequence generation for high net charge content. (Left) Top plot shows mean net charge of sequences generated at each step of ProteinRL fine-tuning to increase (blue) or decrease (red) sequence net charge. Right panel shows distributions of net charges for natural β-lactamase sequences (black) and sequences generated from the prior model (yellow). Y-axes scales are shared. Bottom plot shows mean sequence perplexities determined from the prior model for sequences generated by the agent throughout the iterations of ProteinRL fine-tuning. (Right) Overlaid ESMFold structural models of the natural β-lactamase TEM-1 (gray) and representative sequences generated from the PLM fine-tuned for increasing (blue) or decreasing (red) net charge.

determined by the prior PLM, indicating that they remain high-likelihood β-lactamase sequences (Figure 2). Structural models of representative sequences generated from the fine-tuned PLMs with extreme positive and negative net charges overlay with low RMSD to the natural TEM-1 β-lactamase sequence, despite possessing extreme net charges (+14 for the positively charged sequence, -54 for the negatively charged sequence, -7 for TEM-1) and sharing low sequence identity to TEM-1 (24% identity for the positively charged sequence, 34% identity for the negatively charged sequence) (Figure 2, Figure S2). Active site residues involved in catalysis and substrate recognition are all maintained in generated sequences, and the imparted charged predominantly localize on the protein surface despite no structural information being explicitly included in the design strategy (Figure S3).

3.2 Multi-objective optimization of sequence and structural properties

Next, we tested the capabilities of ProteinRL for simultaneous optimization of multiple properties. We used the scenario of a hit expansion, with the aim of diversifying a target sequence with generated sequences having high-confidence structure predictions and high probability predictions of soluble expression in *E. coli* to enable experimental characterization. For this scenario, we used lysozyme as a test-case, with hen egg white (HEW) lysozyme as the target sequence of interest for expansion. To accomplish this, we defined property reward functions for each of these goals to be combined in a multi-objective reward function (see Supplementary Materials): percent identity of generated sequences to HEW lysozyme, mean pLDDT of residues in ESMFold structural models of generated structures, and predicted probability of soluble expression of generated sequences determined by a BERT-based model that we previous developed [18].

Throughout the multi-objective fine-tuning, the PLM learns to generate sequences that are improved in all three properties (Figure 3 left). Before ProteinRL fine-tuning, sequences generated by the PLM shared 23% identity to the target HEW lysozyme sequence, had mean residue pLDDT of structural models of 60, and had predicted probabilities of expression 84% on average. Throughout fine-tuning, generated sequences reached 85% identity to the HEW lysozyme target sequence, mean residue pLDDT of structural models of 89, and predicted probabilities of expression of 95% on average. Only 12 of 6,507 (<0.2%) of natural lysozyme sequences and no sequences among 1,000 generated from sampling the prior PLM model had values for all three properties greater than these values achieved

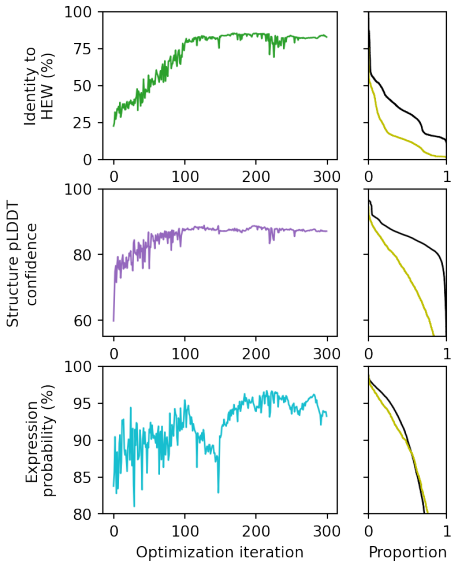

Figure 3: Multi-objective fine-tuning for a hit expansion. (Left) Mean values of sequence identity to the target HEW lysozyme sequence (top), mean pLDDT of ESMFold structural model confidence (middle), and predicted probability of expression determined from an in-house developed model (bottom) for generated sequences throughout iterations of ProteinRL fine-tuning. (Right) Complementary cumulative distributions for property scores for natural lysozyme sequences (black) and sequences generated from the prior model (yellow). The proportion represents the proportion of natural sequences with the indicated score value or greater. Note some distributions truncated with values for the lowest around one-third of natural sequences not shown; full distributions are shown in Figure S4. Y-axes shared with scale on plots to left.

by ProteinRL fine-tuning, indicating that fine-tuning guides the PLM towards a sparsely populated region of sequence space for which all three properties are simultaneously optimized (Figure 3 right). Structural models of sequence generated from the fine-tuned model are predicted to maintain the lysozyme structure (Figure S5), though we note that this is an expected outcome given the goal of increasing sequence identity to the natural lysozyme sequence.

4 Conclusions

ProteinRL is a flexible, data-driven approach for the *de novo* design of protein sequences optimized for specific properties. For two different protein design tasks, including both single- and multi-objective designs, applied to different protein families, ProteinRL showed high success at generating sequences optimized for the desired scenario. In both cases, few or no sequences that possess similar levels of the desired properties were present in natural sequences or sequences generated from PLMs without ProteinRL fine-tuning. Though the cases we highlight here represent realistic goals in protein design, ProteinRL offers great flexibility in the engineering tasks to which it can be applied; any property that can be calculated from proteins sequence or structure can be used in a reward function for property-directed sequence design. We believe ProteinRL is a promising method in the design of fit-for-purpose *de novo* designed proteins for therapeutic, industrial, and biotechnological applications.

5 Acknowledgements

We acknowledge the entire Protein Design and Informatics group at GSK for helpful feedback and discussions in the development of ProteinRL and drafting of this manuscript.

References

[1] The UniProt Consortium. UniProt: the Universal Protein Knowledgebase in 2023. *Nucleic Acids Research*, 51(D1):D523–D531, 11 2022. ISSN 0305-1048. doi: 10.1093/nar/gkac1052. URL https://doi.org/10.1093/nar/gkac1052.

[2] Noelia Ferruz, Steffen Schmidt, and Birte Höcker. Protgpt2 is a deep unsupervised language model for protein design. *Nat Commun*, 13(1):4348, Jul 2022. ISSN 2041-1723 (Electronic); 2041-1723 (Linking). doi: 10.1038/s41467-022-32007-7.

[3] Ethan C Alley, Grigory Khimulya, Surojit Biswas, Mohammed AlQuraishi, and George M Church. Unified rational protein engineering with sequence-based deep representation learning. *Nature methods*, 16(12):1315–1322, 2019.

[4] Donatas Repecka, Vykintas Jauniskis, Laurynas Karpus, Elzbieta Rembeza, Irmantas Rokaitis, Jan Zrimec, Simona Poviloniene, Audrius Laurynenas, Sandra Viknander, Wissam Abuajwa, et al. Expanding functional protein sequence spaces using generative adversarial networks. *Nature Machine Intelligence*, 3(4):324–333, 2021.

[5] Suyue Lyu, Shahin Sowlati-Hashjin, and Michael Garton. Proteinvae: Variational autoencoder for translational protein design. *bioRxiv*, 2023. doi: 10.1101/2023.03.04.531110. URL https://www.biorxiv.org/content/early/2023/03/05/2023.03.04.531110.

[6] Sidney Lyayuga Lisanza, Jake Merle Gershon, Sam Tipps, Lucas Arnoldt, Samuel Hendel, Jeremiah Nelson Sims, Xinting Li, and David Baker. Joint generation of protein sequence and structure with rosettafold sequence space diffusion. *bioRxiv*, 2023. doi: 10.1101/2023.05.08.539766. URL https://www.biorxiv.org/content/early/2023/05/10/2023.05.08.539766.

[7] Ali Madani, Ben Krause, Eric R Greene, Subu Subramanian, Benjamin P Mohr, James M Holton, Jose Luis Jr Olmos, Caiming Xiong, Zachary Z Sun, Richard Socher, James S Fraser, and Nikhil Naik. Large language models generate functional protein sequences across diverse families. *Nat Biotechnol*, 41(8):1099–1106, Aug 2023. ISSN 1546-1696 (Electronic); 1087-0156 (Print); 1087-0156 (Linking). doi: 10.1038/s41587-022-01618-2.

[8] Erik Nijkamp, Jeffrey Ruffolo, Eli N. Weinstein, Nikhil Naik, and Ali Madani. Progen2: Exploring the boundaries of protein language models, 2022.

[9] Daniel Hesslow, Niccoló Zanichelli, Pascal Notin, Iacopo Poli, and Debora Marks. Rita: a study on scaling up generative protein sequence models. *arXiv preprint arXiv:2205.05789*, 2022.

[10] Richard W. Shuai, Jeffrey A. Ruffolo, and Jeffrey J. Gray. Generative language modeling for antibody design. *bioRxiv*, 2022. doi: 10.1101/2021.12.13.472419. URL https://www.biorxiv.org/content/early/2022/12/20/2021.12.13.472419.

[11] Geraldene Munsamy, Sebastian Lindner, Philipp Lorenz, and Noelia Ferruz. Zymctrl: a conditional language model for the controllable generation of artificial enzymes. *Machine Learning for Strucutral Biology Workshop, NeurIPS*, 2022.

[12] Marcus Olivecrona, Thomas Blaschke, Ola Engkvist, and Hongming Chen. Molecular de-novo design through deep reinforcement learning. *Journal of Cheminformatics*, 9(1):48, 2017. doi: 10.1186/s13321-017-0235-x. URL https://doi.org/10.1186/s13321-017-0235-x.

[13] Christof Angermueller, David Dohan, David Belanger, Ramya Deshpande, Kevin Murphy, and Lucy Colwell. Model-based reinforcement learning for biological sequence design. In *International conference on learning representations*, 2019.

[14] Xiaopeng Xu, Tiantian Xu, Juexiao Zhou, Xingyu Liao, Ruochi Zhang, Yu Wang, Lu Zhang, and Xin Gao. Ab-gen: Antibody library design with generative pre-trained transformer and deep reinforcement learning. *Genomics, Proteomics & Bioinformatics*, 2023.

[15] Isaac D. Lutz, Shunzhi Wang, Christoffer Norn, Alexis Courbet, Andrew J. Borst, Yan Ting Zhao, Annie Dosey, Longxing Cao, Jinwei Xu, Elizabeth M. Leaf, Catherine Treichel, Patrisia Litvicov, Zhe Li, Alexander D. Goodson, Paula Rivera-Sánchez, Ana-Maria Bratovianu, Minkyung Baek, Neil P. King, Hannele Ruohola-Baker, and David Baker. Top-down design of protein architectures with reinforcement learning. *Science*, 380(6642):266–273, 2023. doi: 10.1126/science.adf6591. URL https://www.science.org/doi/abs/10.1126/science.adf6591.

[16] Zeming Lin, Halil Akin, Roshan Rao, Brian Hie, Zhongkai Zhu, Wenting Lu, Nikita Smetanin, Robert Verkuil, Ori Kabeli, Yaniv Shmueli, Allan dos Santos Costa, Maryam Fazel-Zarandi, Tom Sercu, Salvatore Candido, and Alexander Rives. Evolutionary-scale prediction of atomic-level protein structure with a language model. *Science*, 379(6637):1123–1130, 2023. doi: 10.1126/science.ade2574. URL https://www.science.org/doi/abs/10.1126/science.ade2574.

[17] Michael S. Lawrence, Kevin J. Phillips, and David R. Liu. Supercharging proteins can impart unusual resilience. *Journal of the American Chemical Society*, 129(33):10110–10112, 2007. doi: 10.1021/ja071641y. URL https://doi.org/10.1021/ja071641y. PMID: 17665911.

[18] Agile Language Transformers for Recombinant Protein Expression Optimization. Jeliazko r. jeliazkov and maxim v. shapovalov and diego del alamo and matt c. sternke and joel d. karpiak. *Machine Learning for Strucutral Biology Workshop, NeurIPS*, 2022.

[19] Baris E. Suzek, Hongzhan Huang, Peter McGarvey, Raja Mazumder, and Cathy H. Wu. UniRef: comprehensive and non-redundant UniProt reference clusters. *Bioinformatics*, 23(10):1282–1288, 03 2007. ISSN 1367-4803. doi: 10.1093/bioinformatics/btm098. URL https://doi.org/10.1093/bioinformatics/btm098.

[20] Martin Steinegger and Johannes Söding. Mmseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. *Nature Biotechnology*, 35(11):1026–1028, 2017. doi: 10.1038/nbt.3988. URL https://doi.org/10.1038/nbt.3988.

[21] Ari Holtzman, Jan Buys, Li Du, Maxwell Forbes, and Yejin Choi. The curious case of neural text degeneration, 2020.

[22] Peter J. A. Cock, Tiago Antao, Jeffrey T. Chang, Brad A. Chapman, Cymon J. Cox, Andrew Dalke, Iddo Friedberg, Thomas Hamelryck, Frank Kauff, Bartek Wilczynski, and Michiel J. L. de Hoon. Biopython: freely available Python tools for computational molecular biology and bioinformatics. *Bioinformatics*, 25(11):1422–1423, 03 2009. ISSN 1367-4803. doi: 10.1093/bioinformatics/btp163. URL https://doi.org/10.1093/bioinformatics/btp163.

[23] Patrick Kunzmann, Tom David Müller, Maximilian Greil, Jan Hendrik Krumbach, Jacob Marcel Anter, Daniel Bauer, Faisal Islam, and Kay Hamacher. Biotite: new tools for a versatile python bioinformatics library. *BMC Bioinformatics*, 24(1):236, 2023. doi: 10.1186/s12859-023-05345-6. URL https://doi.org/10.1186/s12859-023-05345-6.

[24] Schrödinger, LLC. The PyMOL molecular graphics system, version 1.2. November 2015.

[25] Xiaojun Wang, George Minasov, Jesús Blázquez, Emilia Caselli, Fabio Prati, and Brian K. Shoichet. Recognition and resistance in tem beta-lactamase. *Biochemistry*, 42(28):8434–8444, 07 2003. doi: 10.1021/bi034242y. URL https://doi.org/10.1021/bi034242y.

[26] Burkhard Rost and Chris Sander. Conservation and prediction of solvent accessibility in protein families. *Proteins: Structure, Function, and Bioinformatics*, 20(3):216–226, 1994. doi: https://doi.org/10.1002/prot.340200303. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/prot.340200303.

6 Appendix

6.1 Supplementary Methods

6.1.1 Foundational PLM

We used the pre-trained ProGen2-base (764M parameter) model as a foundational generative PLM [8] in ProteinRL. To allow for protein family-specific sequence generation, we first fine-tuned the ProGen2 model on a set of homolog sequences for a particular family. The two protein families used for the two outlined ProteinRL training objective examples were β-lactamase and lysozyme. Homolog sequences were identified by searching the Uniref100 database [19] with the mmseqs2 package [20], using the TEM-1 β-lactamase sequence and hen egg white (HEW) lysozyme sequences as query sequences.

The model was fine-tuned on the sequence homolog set for 2 epochs using a cross entropy loss function, the AdamW optimizer ($\beta_1 = 0.9, \beta_2 = 0.999$), a maximum learning rate of 10^{-4} with a cosine decay, weight decay of 0.1, max gradient norm clipping at 1.0, and an effective batch size of 128 sequences. Fine-tuning was performed using 2 Nvidia A6000 GPUs (48 GB VRAM each). This fine-tuned family-specific Progen2 model was used as the initial generative PLM to be further fine-tuned as an agent PLM for property-directed optimization.

To generate sequences from the generative PLM, we used autoregressive next residue prediction sampling from a specified initial sequence context (given below for specific cases) until an end-of-sequence token was reached. Sequences were generated using nucleus sampling with a sampling temperature of 0.8 using the most probable tokens up to a cumulative probability of 95% [21].

6.1.2 ProteinRL fine-tuning for highly charged β-lactamase sequences

We used ProteinRL to fine-tune a generative PLM to generate β-lactamase sequences with high net charge content. At each step of ProteinRL fine-tuning, 16 sequences were generated from the agent PLM. The initial sequence context of "1MSI" (the first three residues of TEM-1 β-lactamase) was used to seed all sequence generations at every step. We calculated sequence net charge using the `charge_at_pH` (using a pH of 7.4) method within the biopython python package [22]. For fine-tuning for positive and negative net charge, sequence net charges were transformed by the functions:

$$\phi(x)_{positive} = (1 + 10^{10m\frac{x-0.5(x_{high}+x_{low})}{x_{high}-x_{low}}})^{-1} \tag{4}$$

$$\phi(x)_{negative} = (1 + 10^{10m\frac{x-0.5(x_{high}+x_{low})}{x_{high}+x_{low}}})^{-1} \tag{5}$$

where x is the sequence net charge, x_{high} and x_{low} are high and low net charge values observed among natural β-lactamase sequences (black distribution in top right panel of main text Figure 2, and m is the sigmoid slope set to 0.5. Equations 4 and 5 are plotted in Figure S1. A value of σ=120 was used in the augmented log likelihood equation (Equation 3). ProteinRL fine-tuning was run for 100 iterations using an AdamW optimizer ($\beta_1 = 0.9, \beta_2 = 0.999$) and a constant learning rate of 10^{-6}. Fine-tuning was performed using 2 Nvidia A6000 GPUs (48 GB VRAM each) and took around 3 wall hours for each run.

6.1.3 ProteinRL fine-tuning for a lysozyme hit expansion

We used ProteinRL in the scenario of a hit expansion to generete lysozyme sequence optimizing for three properties: high sequence identity to hen egg white (HEW) lysozyme, high confidence of ESMFold structural models, and high predicted probability of expression in *E. coli*. At each step of ProteinRL fine-tuning, 16 sequences were generated using "1MRS" (the first three residues of HEW lysozyme) as the initial sequence context. Sequence identity to HEW lysozyme was calculated using the `align_optimal` method within the biotite python package [23]. The confidence of ESMFold structural models was determined as the mean pLDDT score among C-α residues in the structural model. The predicted probability of expression was determined from a model we

previously developed [18]. Since the goal was to increase values for all properties, similar sigmoidal transformation functions to Equation 4 were used for all three properties.

To allow for multi-objective optimization of all three properties we defined a multi-objective reward function given by the weighted geometric mean of sequence property scores among multiple specified properties as:

$$R_{multi}(seq) = (\prod_{j=1}^{n} \phi_j(R_j(seq))^{w_i})^{\frac{1}{\sum_{j=1}^{n} w_j}} \tag{6}$$

where each specified property receives it's own sequence property reward function $R_j(seq)$, data transformation ϕ_j, and weight w_j that allows for customizable prioritization of the specified properties. Weights for all three properties were set equally at 1.

A value of σ=120 was used in the augmented log likelihood equation (Equation 3). ProteinRL fine-tuning was run for 300 iterations using an AdamW optimizer ($\beta_1 = 0.9, \beta_2 = 0.999$) and a constant learning rate of 10^{-6}. Fine-tuning was performed using 3 Nvidia A6000 GPUs (48 GB VRAM each) and took around 6 wall hours.

6.2 Supplementary Figures

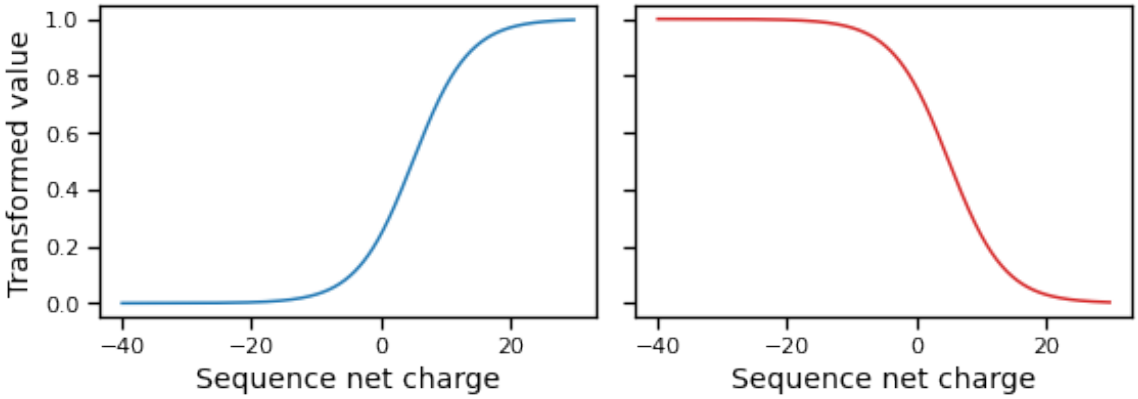

Figure S1: Transformation functions for optimizing β-lactamase sequences for high and low net charges. Functions plotted for Equation 4 (left) and Equation 5 (right).

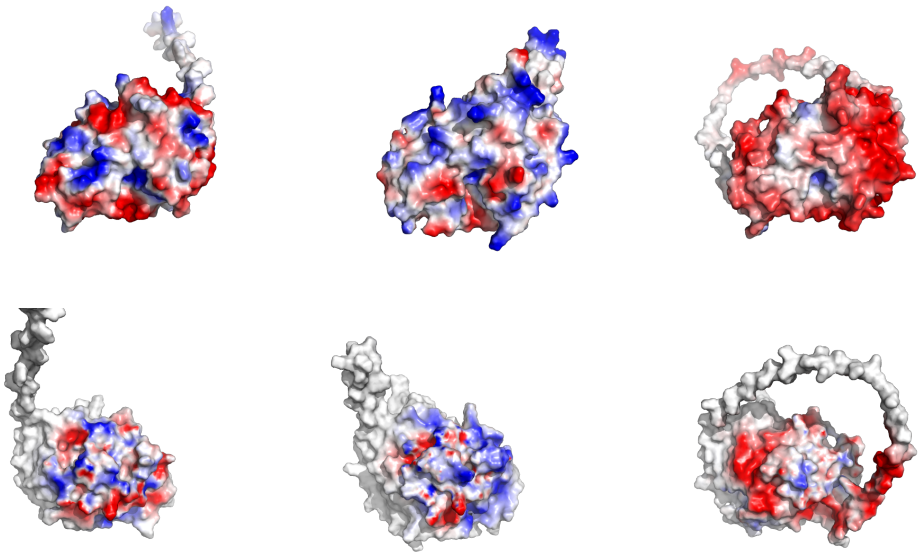

Figure S2: Surface electrostatic potential maps of β-lactamases. Vacuum electrostatic surface potentials generated using PyMOL [24] for of ESMFold structural models for the natural TEM-1 β-lactamase (left) and representative sequences generated from the PLM fine-tuned for increasing net (center) or decreasing net charge (right). Top and bottom rows show two different view of each structure, rotated by 180 degrees.

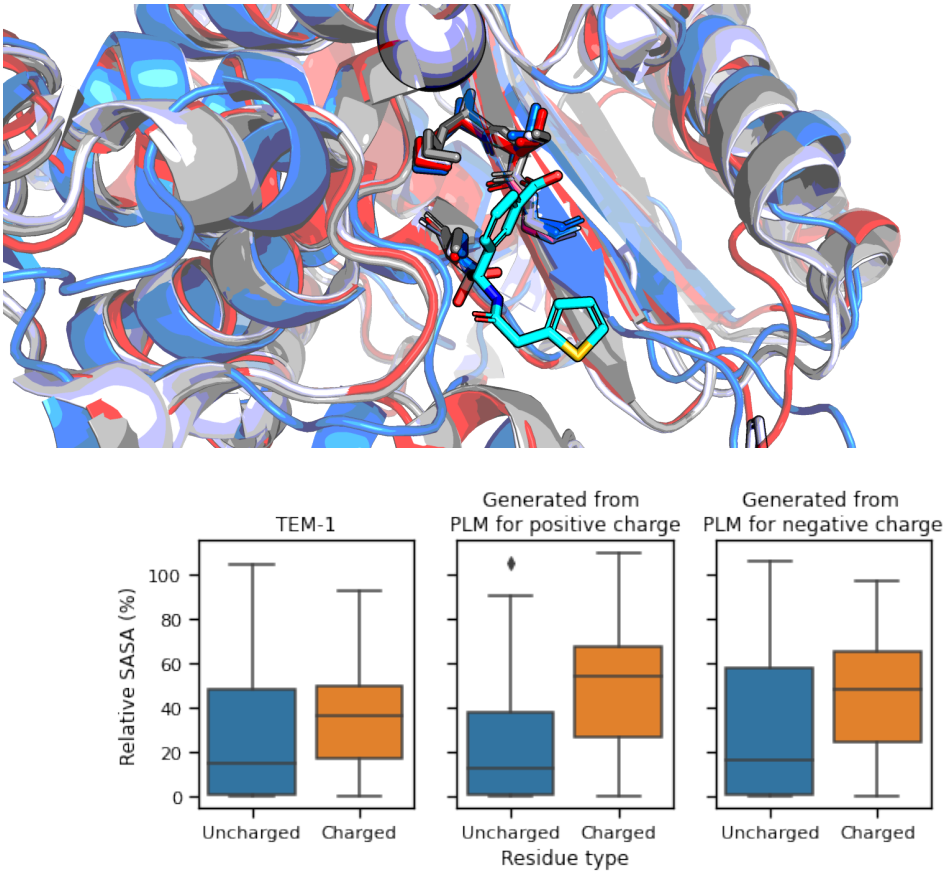

Figure S3: Conservation of β-lactamase sequence features in generated designs. (Top) Snapshot of active sites for overlaid ESMFold structural models of the natural β-lactamase TEM-1 (gray) and representative sequences generated from the PLM fine-tuned for increasing net (blue) or decreasing net charge (red), and a crystal structure of the TEM-1 β-lactamase co-crystallized with a boronic acid inhibitor (protein shown in periwinkle, boronic acid inhibitor shown in cyan, PDB: 1NXY [25]). Residues involved in catalysis and substrate recognition shown as sticks. (Bottom) Box plots distributions of residue relative solvent accessible surface accessibility (SASA) grouped by charged/uncharged residue type for TEM-1β-lactamase (left) and representative sequences generated from the PLM fine-tuned for increasing net (middle) or decreasing net charge (right). Relative SASA is determined as the residue SASA calculated in the structure normalized to the residue exposed maximum SASA described in [26].

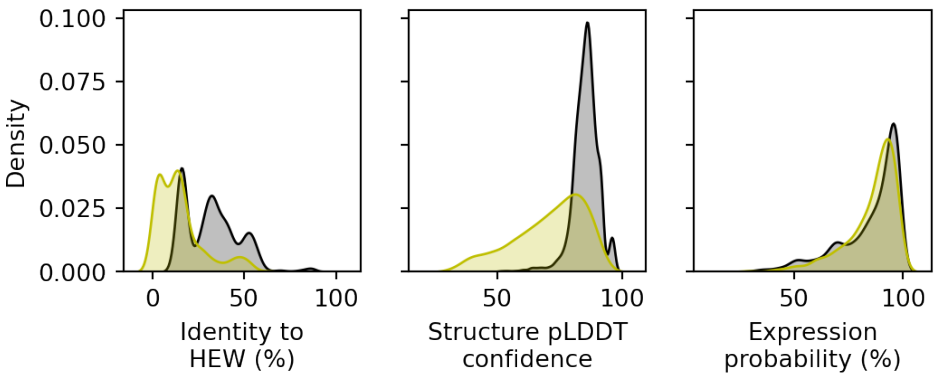

Figure S4: Lysozyme property distributions. Distributions for sequence identities to the target HEW lysozyme sequence (left), mean pLDDT of ESMFold structural model confidence (middle), and predicted probability of expression (right) for natural lysozyme sequences (black) and sequences generated from the prior PLM (yellow).

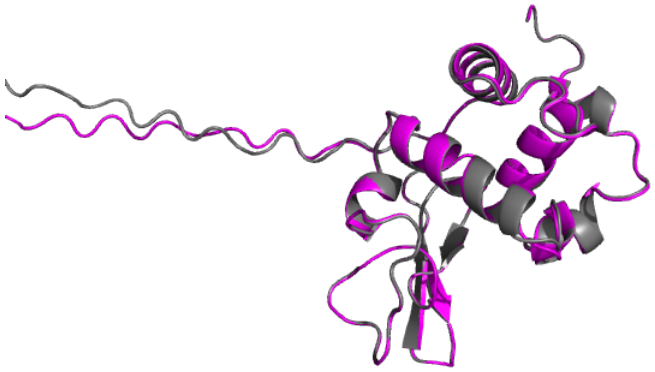

Figure S5: ESMFold structural models of HEW lysozyme (gray) and a representative sequence generated from the PLM fine-tuned with ProteinRL (pink).