
Action Identifiability in Protein-Design Agents

Anonymous Author(s)

Affiliation

Address

email

Abstract

1 Molecular design agents choose interventions by scoring predicted outcomes,
2 which requires the scorer to distinguish the available actions. We test this
3 requirement using 78 COSMIC mutational processes as actions and distance
4 in ESM-2 representation space as the score. Median pairwise distance falls
5 from 0.815 in COSMIC mutational-spectrum space to 0.0068 between ESM-
6 2 outcome regions, a $120\times$ compression, and a composition-matched control
7 reproduces the destinations. A learned policy identifies the true action in
8 only 7.55% of held-out cases, while split-conformal prediction reaches 92.2%
9 coverage only by returning 66.8 of 78 actions on average. At 100 scorer calls,
10 adaptive action selection reaches mean target distance 0.177, compared
11 with 0.165 for a fixed policy and 0.173 for a random-action policy, and an
12 independent Pfam profile hidden Markov model finds no PF17041 hits at
13 the curated threshold. These results separate biological admissibility from
14 scorer-level action identifiability and show why molecular agents should
15 test whether their scorer distinguishes candidate actions before reporting a
16 specific mechanism.

17 1 Introduction

18 Closed-loop molecular agents propose an intervention, score the resulting molecule, and
19 use the score to choose the next action. This procedure can search large spaces efficiently,
20 but only if the scorer resolves the actions available to the policy. If several interventions
21 produce indistinguishable scored outcomes, the agent cannot infer which intervention caused
22 a result. We call this requirement action-space identifiability: outcomes from competing
23 actions must remain distinguishable at the resolution required by the decision. Identifiability
24 differs from biological admissibility. An action may represent a valid biological process
25 even when the scorer discards the information that distinguishes it from other processes.
26 We use 78 COSMIC single-base-substitution signatures as a controlled action space [2, 3].
27 Each signature defines a distribution over 96 trinucleotide substitutions associated with a
28 mutational process. We apply these signatures to coding sequences, translate the mutants,
29 embed the resulting proteins with ESM-2 [4], and score them by cosine distance to a Pfam-
30 domain centroid. This PRISM-based testbed [1] makes both the biological action and the
31 learned score explicit. We ask whether the score preserves action identity, whether closed-
32 loop selection improves on matched controls, and whether an independent model confirms
33 the inferred molecular property.

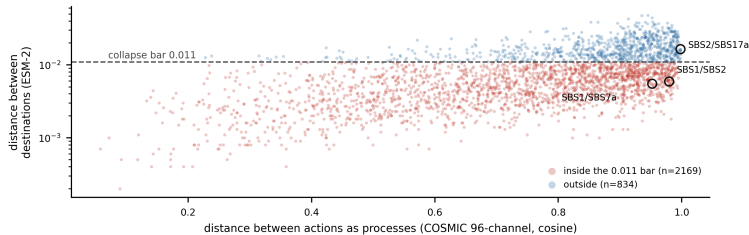


Figure 1: ESM-2 compresses the COSMIC action space. Of 3,003 action pairs, 1,780 remain more than 0.5 apart in COSMIC space but fall within 0.011 between ESM-2 outcome regions.

2 Agent and action space

An action is a COSMIC signature s , represented by $\theta_s \in \Delta^{95}$ over 96 trinucleotide substitution types. Executing s on coding sequence x samples compatible substitutions and produces mutant x' . We translate x' and embed the protein with ESM-2. The primary objective is cosine distance to a target Pfam-domain centroid, where lower distance is better. This score measures movement in representation space. It does not directly measure protein function or acquisition of the target domain.

The optimization loop uses a Metropolis rule on target distance. We reject proteins shorter than 95% of wild-type length and match realized edit counts across proposal arms. An unselected drift control scores each proposal but accepts it regardless of score, testing whether scorer feedback improves on the proposal distribution. A separate supervised policy predicts which COSMIC action should produce a requested Pfam-domain transition.

3 The scorer compresses distinct actions

Across all 3,003 pairs of COSMIC signatures, median distance between native 96-channel spectra is 0.815. Median distance between their ESM-2 outcome regions is 0.0068, making the scored action space approximately 120 $\times$ tighter. In total, 1,780 pairs (59.3%) are more than 0.5 apart in COSMIC space but fall within 0.011 after scoring.

A composition-matched control preserves the realized substitution composition while scrambling trinucleotide context and mutation placement. It reproduces each action’s ESM-2 destination within cosine distance 0.02. Actions with similar broad substitution patterns collapse together, and the same pattern appears in ESM-2 8M and 650M. Thus, the agent has 78 nominal actions but a much coarser effective action space under its scorer.

4 Controls remove the apparent optimization gain

On the PF17041 target, the structured action reduces ESM-2 distance by 0.0736 ± 0.0030 across five seeds, compared with 0.0218 ± 0.0051 for edit-count-matched random mutation. The additional improvement is 0.0519, with paired-bootstrap 95% CI [0.0477, 0.0573].

Composition matching removes this advantage. Reapplying the same amino-acid substitution multiset at random positions produces target improvement 0.0764, compared with 0.0736 for the structured action. Their paired difference is -0.0027 , with 95% CI $[-0.0063, 0.0012]$. We find no evidence that trinucleotide context or mutation placement adds steering power beyond the substitution composition preserved by ESM-2.

Scorer-based selection also adds no measurable benefit. Structured unselected drift improves target distance by 0.0737, compared with 0.0736 under Metropolis selection. Random drift improves it by 0.0236, compared with 0.0218 under selection. An independent PF17041 profile HMM likewise finds no sequence above Pfam’s curated gathering threshold. The evidence supports movement under the ESM-2 objective, not acquisition of PF17041-like sequence properties (Figure 2).

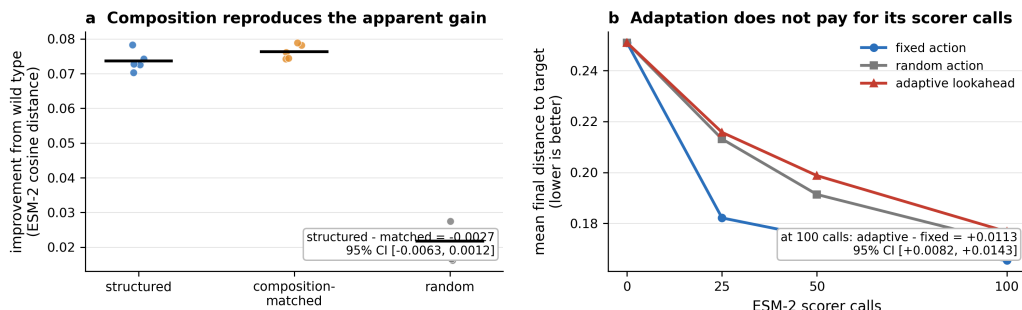


Figure 2: Matched controls test whether the apparent optimization gains require the nominal action or adaptive selection. (a) Each point is one paired seed. Horizontal bars show means. The structured arm does not outperform composition-matched mutation, while both outperform edit-count-matched random mutation. (b) Under equal ESM-2 scorer-call budgets, five-action adaptive lookahead remains farther from the target than fixed or random selection. The budget experiment contains 60 trajectories per budget across four targets, five proteins, and three seeds.

5 The policy cannot identify a specific action

A supervised policy trained on 35,403 realized Pfam-domain transitions reaches 7.55% top-1 accuracy on a protein-grouped split, compared with 6.62% for a train-only per-(origin,target) frequency baseline. The difference is +0.93 percentage points, with protein-bootstrap 95% CI [+0.71, +1.27]. Top-5 accuracy is 19.65%, compared with 16.93% for the baseline. The policy therefore contains information about action choice but predicts the wrong action in 92.45% of held-out cases.

The policy does not generalize to unseen actions. Holding out complete COSMIC processes reduces top-1 accuracy to 0.8%, below the 1.3% uniform baseline. Holding out SBS17a yields 0% recovery. The model interpolates among action classes observed during training.

Set-valued prediction raises coverage by returning most of the catalogue. A direct ESM-2 neighborhood reaches 82.6% coverage with a mean set size of 58.2. Adaptive prediction sets reach 92.2% coverage but return 66.8 actions on average, with median size 78. Abstention produces smaller outputs, but retaining 8.6% of test cases yields only 33.0% conditional accuracy (Figure 3).

Adaptive selection also fails under equal compute. At 100 ESM-2 calls, five-action one-step lookahead reaches mean target distance 0.177, compared with 0.165 for fixed SBS17a and 0.173 for random action selection. Adaptive selection is worse than the fixed policy by +0.0113, with 95% CI [+0.0082, +0.0143], and worse than random by +0.0042, with 95% CI [+0.0021, +0.0057]. Lookahead spends scorer calls without producing enough benefit to offset the lost transitions.

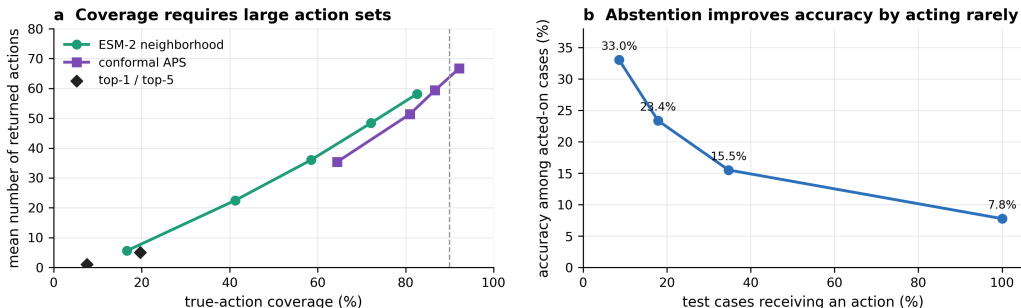


Figure 3: Uncertainty-aware interfaces do not recover a specific useful action. (a) Direct ESM-2 neighborhoods and adaptive prediction sets (APS) increase true-action coverage only by returning much of the 78-action catalogue. The dashed line marks 90% coverage. (b) Validation-frozen abstention thresholds improve conditional accuracy by acting on fewer cases. The point at 100% coverage is the unfiltered policy.

6 Discussion

Biological grounding does not guarantee scorer-level identifiability. COSMIC signatures describe distinct mutational processes, but ESM-2 maps many outcomes to nearly the same region. This loss of resolution prevents reliable action identification and removes the benefit of adaptive selection in this testbed. The results support a three-step evaluation. To start, measure whether the scorer distinguishes the actions available to the policy. Next, compare optimization with controls matched for proposal distribution, edit count, sequence length, and scorer-call budget. Then, validate the molecular claim with a measurement independent of the optimization objective. If the scorer cannot resolve competing actions, the agent should return a useful uncertainty set or abstain. Our conclusions apply to the action-scorer pair studied here. Other molecular actions or scorers may preserve more mechanistic information. Methods that steer protein representations are also dual-use. We restrict the action space to characterized processes and focus on auditing failure modes instead of claiming successful molecular design.

References

- [1] PRISM: Mutational signatures as directional drivers in protein function space. Anonymous, under review.
- [2] J. G. Tate et al. COSMIC: the catalogue of somatic mutations in cancer. *Nucleic Acids Research*, 47(D1):D941–D947, 2019.
- [3] L. B. Alexandrov et al. The repertoire of mutational signatures in human cancer. *Nature*, 578:94–101, 2020.
- [4] Z. Lin et al. Evolutionary-scale prediction of atomic-level protein structure with a language model. *Science*, 379(6637):1123–1130, 2023.
- [5] Y. Geifman and R. El-Yaniv. Selective classification for deep neural networks. *NeurIPS*, 2017.
- [6] Y. Romano, M. Sesia, and E. Candès. Classification with valid and adaptive coverage. *NeurIPS*, 2020.
- [7] A. Angelopoulos and S. Bates. Conformal prediction: A gentle introduction. *Foundations and Trends in Machine Learning*, 16(4), 2023.

A Methods and additional results

Action-space analysis. `cross_signature_steering.py` applies COSMIC actions to proteins and measures the mean ESM-2 representation reached by mutants from each action. We compute distances between all 78 signatures in their native 96-channel parameterization and between their ESM-2 outcome regions. COSMIC and ESM-2 distances have Spearman

125 correlation $\rho = 0.545$. For example, SBS1 and SBS7a are 0.952 apart in COSMIC space
126 but only 0.0055 apart after scoring. We repeat representative analyses with ESM-2 8M and
127 650M. The composition-matched control preserves realized substitution composition while
128 scrambling trinucleotide context and placement.

129 Optimization loop. `prism_proto_design.py` and `proto_ablation.py` implement signa-
130 ture proposals with a Metropolis rule at $T = 0.02$. Matched-random comparisons equalize
131 realized substitution counts. All PF17041 trajectories begin from the same wild type at co-
132 sine distance 0.1757. The corrected comparison uses five seeds. Paired intervals use 100,000
133 bootstrap resamples with seed 20260812. The optimizer accepts approximately 88% of pro-
134 posals at $T = 0.02$.

135 The unselected-drift control uses the same proposals and scoring procedure but accepts
136 every proposal. It differs from the corresponding Metropolis run only in whether the score
137 affects acceptance.

138 External PF17041 validation. `validate_designs_hmmer.py` evaluates the shared wild
139 type and the best sequence from each arm and seed against the PF17041.11 SasG E profile
140 using HMMER 3. No sequence passes Pfam’s curated gathering threshold. Under a per-
141 missive reporting threshold, none of the five structured-action sequences produces a match.
142 Other arms produce only weak, nonsignificant matches, with best $E = 0.34$.

143 Policy. `inverse_design_classifier.py` trains a random forest on ESM-2 displacement
144 and domain-identity features. We use protein-grouped splits and compare against global
145 and per-(origin,target) frequency baselines. Leave-actions-out experiments remove complete
146 COSMIC actions from training.

147 Set-valued prediction and abstention. `equivalence_class_policy.py` combines predicted
148 probabilities with the archived ESM-2 650M outcome-distance matrix. Taking the transi-
149 tive closure at distance 0.011 places all 78 actions in one connected component. Coverage
150 intervals resample the 19 held-out proteins 10,000 times.

151 `selective_action_interface.py` reserves 14 training proteins for calibration, retrains on
152 the remaining 42, and evaluates frozen confidence thresholds, least-ambiguous classification
153 sets, and adaptive prediction sets on 19 disjoint test proteins [5–7]. A threshold chosen to
154 retain 20% of calibration cases accepts 17.9% of test cases at 23.4% conditional accuracy.

155 Adaptive action selection. `adaptive_policy_multibudget.py` compares five-action
156 greedy lookahead with fixed-SBS17a and random-action policies at budgets of 25, 50, and
157 100 scorer calls. Four Pfam targets, five proteins, and three seeds produce 60 trajectories
158 per budget. Proposals shorter than 95% of wild-type length are rejected but retain their
159 allocated call. Confidence intervals resample proteins while retaining their targets and seeds.